NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Regorafenib for previously treated metastatic colorectal cancer [ID4002]

Committee Papers

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Regorafenib for previously treated metastatic colorectal cancer [ID4002]

Contents:

The following documents are made available to consultees and commentators:

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

1. **Company submission** from **Bayer Plc**
2. **Company response to NICE's request for clarification**
3. **Patient group, professional group and NHS organisation submission** from:
   a. Bowel Cancer UK
   b. NCRI-ACP-RCP-RCR
4. **Evidence Review Group report prepared by Kleijnen Systematic Reviews**
5. **Evidence Review Group – factual accuracy check**
6. **Technical engagement response** from **Bayer Plc**
7. **Technical engagement response & expert statement from experts:**
   - Jane Ashford – patient, nominated by Bowel Cancer UK
8. **Technical engagement response from consultees and commentators:**
   - Servier
9. **Evidence Review Group critique of company response to technical engagement prepared by Kleijnen Systematic Reviews**

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

fNATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Regorafenib for treating metastatic colorectal cancer (ID4002)

Document B

Company evidence submission

May 2022

File name	Version	Contains confidential information	Date
Regorafenib_ID4002_DocumentB_25May22_Final_1.0 (fully redacted)	1.0	Yes	25 May 22

Instructions for companies

This is the template for submission of evidence to the National Institute for Health and Care Excellence (NICE) as part of the single technology appraisal (STA) process. Please note that the information requirements for submissions are summarised in this template; full details of the requirements for pharmaceuticals and devices are in the user guide.

This submission must not be longer than 150 pages, excluding appendices and the pages covered by this template. If it is too long it will not be accepted.

Companies making evidence submissions to NICE should also refer to the NICE health technology evaluation guidance development manual.

In this template any information that should be provided in an appendix is listed in a box.

Highlighting in the template (excluding the contents list)

Square brackets and grey highlighting are used in this template to indicate text that should be replaced with your own text or deleted. These are set up as form fields, so to replace the prompt text in [grey highlighting] with your own text, click anywhere within the highlighted text and type. Your text will overwrite the highlighted section.

To delete grey highlighted text, click anywhere within the text and press DELETE.

Grey highlighted text in the footer does not work as an automatic form field, but serves the same purpose – as prompt text to show where you need to fill in relevant details. Replace the text highlighted in [grey] in the header and footer with appropriate text. (To change the header and footer, double click over the header or footer text. Double click back in the main body text when you have finished.)

Contents

List of Tables

List of Figures

Abbreviations

Abbreviation	Definition
5-FU	5-fluorouracil
5-FU/FA	5-fluorouracil/folinic acid
AE	Adverse events
AIC	Akaike information criterion
ALT	Alanine aminotransferase
ASCO	American Society of Clinical Oncology
AST	Aspartate aminotransferase
ATU	Temporary Authorization for Use
BIC	Bayesian information criterion
BMI	Body mass index
BNF	British National Formulary
BSA	Body surface area
BSC	Best supportive care
CAPOX	Capecitabine and oxaliplatin
CC	Chronic constipation
CHMP	Committee for Medicinal Products for Human Use
CI	Confidence interval
CR	Complete response
CrI	Credible Interval
CRC	Colorectal cancer
CSR	Clinical study report
CT	Computed tomography
CTC	Common Toxicity Criteria
CTCAE	Common Terminology Criteria for Adverse Events
DCR	Disease control rate
dMMR	Mismatch repair deficiency
DOR	Duration of response
DSU	Decision Support Unit
ECOG	Eastern Cooperative Oncology Group
eGFR	Estimated glomerular filtration rate
EGFR	Epidermal growth factor receptor
EMA	European Medicines Agency
eMIT	Electronic market information tool
EOT	End of treatment
EPAR	European public assessment report
ERG	Evidence review group
ESMO	European Society for Medical Oncology
ESS	Effective sample size

Abbreviation	Definition
FAS	Full analysis set
FE	Fixed effects
FGFR	Fibroblast growth factor receptor
FOIB	Fluoropyrimidine, oxaliplatin, irinotecan, bevacizumab
FOIBE	Fluoropyrimidine, oxaliplatin, irinotecan, bevacizumab, anti-EGFR antibody
FOLFIRI	Folinic acid, fluorouracil and irinotecan
FOLFOX	Folinic acid, fluorouracil and oxaliplatin
FOLFOXIRI	Folinic acid, fluorouracil, oxaliplatin and irinotecan
FP	Fluoropyrimidine
G-CSF	Granulocyte-colony stimulating factor
GIST	Gastrointestinal stromal tumours
GP	General practitioner
HCC	Hepatocellular carcinoma
HR	Hazard ratio
HRU	Healthcare resource use
HRG	Healthcare Resource Group
HRQL	Health-related quality of life
HUI-3	Health Utilities Index 3
ICER	Incremental cost-effectiveness ratio
INR	International normalized ratio
IQR	Interquartile range
ITC	Indirect treatment comparison
ITT	Intention-to-treat
IVRS	Interactive Voice Response System
KIT	Stem cell factor receptor
KM	Kaplan–Meier
LY	Life year
LYG	Life years gained
mCRC	Metastatic colorectal cancer
MedDRA	Medical Dictionary for Regulatory Activities
MMR	Mismatch repair
MSI	Microsatellite instability
MSS	Microsatellite stable
NCCN	National Comprehensive Cancer Network
NCI	National Cancer Institute
NEL	Non-elective long stay
NES	Non-elective short stay
NHS	National Health Service

Abbreviation	Definition
NICE	National Institute for Health and Care Excellence
NMA	Network meta-analysis
NMB	Net monetary benefit
NR	Not reported
NYHA	New York Heart Association
OD	Once daily
OOR	Objective response rate
OP	Outpatient
OS	Overall survival
OWSA	One-way sensitivity analysis
PartSA	Partitioned survival analysis
PAS	Patient access scheme
PBO	Placebo
PD	Progressive disease
PDGFR	Platelet-derived growth factor receptor
PFS	Progression-free survival
PK	Pharmacokinetic
PO	Per oral (by mouth)
PR	Partial response
PRO	Patient reported outcome
PS	Performance score
PSA	Probabilistic sensitivity analysis
PSS	Personal Social Services
PSSRU	Personal Social Services Research Unit
PTT	Partial thromboplastin time
QALY	Quality-adjusted life year
QoL	Quality of life
RCT	Randomized controlled trial
RDI	Relative dose intensity
RECIST	Response Evaluation Criteria in Solid Tumours
REG	Regorafenib
RET	Rearranged during transfection
RWE	Real-world evidence
SAE	Serious adverse event
SAS	Safety analysis set
seTE	Standard error of treatment effect
SD	Stable disease
SLR	Systematic literature review
SmPC	Summary of product characteristics
STA	Single technology appraisal

Abbreviation	Definition
TA	Technology appraisal
TE	Treatment effect
TEAE	Treatment-emergent adverse event
TKI	Tyrosine kinase inhibitor
TNM	Tumour node metastases
ToT	Time on treatment
TSD	Technical support document
TTF	Time to treatment failure
UK	United Kingdom
ULN	Upper limit of normal
VAS	Visual analogue scale
VEGF	Vascular endothelial growth factor
VEGFR	Vascular endothelial growth factor receptor
WTP	Willingness to pay
XELOX	Oxaliplatin and capecitabine

Executive summary

Colorectal cancer (CRC), also known as bowel cancer, is the fourth most common cancer in the UK. It is a heterogeneous disease characterized by cancerous growths in the large intestine (colon) and rectum[1-3]. Stage IV metastatic CRC (mCRC) is an advanced form of CRC that has metastasized beyond the large intestine and nearby lymph nodes, typically spreading first to the liver.[4, 5] Patients with Stage IV mCRC have a poor prognosis, with 1-year survival rates of approximately 44%, and 5-year survival rates of less than 10%.[6, 7] Survival outcomes in the ≥ 3L setting are particularly poor, ranging between 6–12 months.[8, 9]

In England and Wales, first and second-line treatments for mCRC are chemotherapy and biological therapy. After failure of these therapies patients move to 'subsequent or alternative therapy' (NG151) i.e. NICE TA405 recommends trifluridine/tipiracil, if fluoropyrimidine-, oxaliplatin-, or irinotecan-based chemotherapies, anti-VEGF agents and anti-EFGR agents have failed or when these agents are not suitable.

Bayer is making this submission on the request of physicians who want an alternative to the chemotherapy trifluridine/tipiracil. Physicians consider trifluridine/tipiracil and regorafenib to be of comparable efficacy and comparable tolerability, but with the key difference of having different adverse effect profiles. A medicines adverse effect profile is a significant consideration when making individual treatment decisions.

This is a restricted submission where a recommendation for regorafenib is sought in a ≥3rd line setting i.e. a similar position to trifluridine/tipiracil. As such trifluridine/tipiracil is the key comparator in this submission and is the focus of the economic evidence.

An indirect treatment comparison shows regorafenib and trifluridine/tipiracil to be of comparable efficacy in the key outcomes of overall survival (HR 0.99 CI 0.84, 1.17) and progression free survival (0.93 CI 0.85, 1.03) - the base case point estimates favouring regorafenib to a small but non-significant degree for both outcomes. The findings of the ITC were robust to scenario analyses, with individual analyses numerically, but not statistically, favouring regorafenib or trifluridine/tipiracil.

Regorafenib is available to the NHS with a confidential PAS. [redacted]

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

B.1.1. Decision problem

Since regorafenib received approval in mCRC in 2013, treatment options have increased and clinical practice has changed. Patients are now treated with trifluridine/tipiracil where previously they would have received best supportive care.

Due to these changes in treatment practices, Bayer is not seeking a recommendation for the whole licensed population. This is a restricted submission where a recommendation for regorafenib is sought in a ≥3rd line setting i.e. a similar position to trifluridine/tipiracil. As such trifluridine/tipiracil is the key comparator in this submission and is the focus of the economic evidence.

The proposed position in the treatment pathway is narrower than the marketing authorisation. We are restricting the submission to these patients because physicians have requested an alternative to trifluridine/tipiracil at the third or later line setting. Physicians have indicated that regorafenib, with its comparable efficacy to trifluridine/tipiracil, and different adverse event profile, would provide a valuable additional option for patients.

Table 1: The decision problem

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
Population	Adults with metastatic colorectal cancer who have been previously treated with, or are not considered candidates for, available therapies	Adults with metastatic colorectal cancer who have failed on first-line chemotherapy/first-line biologic and who are being considered for ≥ 3rd-line treatment. Specifically, we are seeking a recommendation for patients for whom treatment with trifluridine/tipiracil is being considered.	Physicians have requested an alternative to trifluridine/tipiracil at the third or later line setting. Physicians have indicated that regorafenib, with its comparable efficacy to trifluridine/tipiracil, but different adverse event profile, would provide an alternative treatment option for these patients and is the patient group for whom regorafenib would be considered.
Intervention	Regorafenib	As per final scope	Not applicable
Comparator(s)	• Single-agent irinotecan (after FOLFOX) • FOLFIRI (after either FOLFOX or CAPOX) • FOLFOX (after either FOLFIRI or CAPOX) • Raltitrexed (if 5-FU/FA are not suitable) • Trifluridine/tipiracil • Best supportive care	• Trifluridine/tipiracil (main comparator – full set of economic analyses) • Best supportive care (minor comparator – reduced set of economic analyses)	**<u>Irinotecan, FOLFIRI, FOLFOX, raltitrexed</u>** These comparators are outside of regorafenib's marketing authorisation (in italics) *Regorafenib is indicated as monotherapy for the treatment of adult patients with metastatic colorectal cancer (mCRC) who have been previously treated with, or are not considered candidates for, available therapies. These include fluoropyrimidine-based chemotherapy, an anti-VEGF therapy and an anti-EGFR therapy.* The listed treatments were available before regorafenib was licensed in mCRC and fall under the definition of "available therapies" in the license wording. Consequently,

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
			regorafenib is not an alternative to these agents and they are not comparators. The registration trials for regorafenib investigated its use in patients who had received these 'available therapies'. Physicians would not consider regorafenib as an alternative to these treatments and would only consider regorafenib after these treatments have failed.
Outcomes	• Overall survival • Progression-free survival • Response rates • Adverse effects of treatment • Health-related quality of life	As per final scope	Not applicable
Key: 5-FU/FA, 5-fluorouracil/folinic acid; CAPOX, capecitabine and oxaliplatin; EGFR, epidermal growth factor receptor; FOLFIRI, folinic acid, fluorouracil and irinotecan; FOLFOX, folinic acid, fluorouracil and oxaliplatin; VEGF, vascular endothelial growth factor.			

B.1.2. Description of the technology being appraised

The summary of product characteristics (SmPC) is in Appendix C.

Regorafenib is an oral tumour deactivation agent that potently blocks multiple protein kinases, including kinases involved in tumour angiogenesis (VEGFR1, -2, -3, TIE2), oncogenesis (KIT, RET, RAF-1, BRAF, $BRAF^{V600E}$), metastasis (VEGFR3, PDGFR, FGFR) and tumour immunity (CSF1R). In particular, regorafenib inhibits mutated KIT, an oncogenic driver, and thereby blocks tumour cell proliferation. In preclinical studies regorafenib has demonstrated potent antitumour activity in a broad spectrum of tumour models including colorectal, gastrointestinal stromal and hepatocellular tumour models which is likely mediated by its anti-angiogenic and anti-proliferative effects. In addition, regorafenib reduced the levels of tumour associated macrophages and has shown anti-metastatic effects in vivo. Major human metabolites (M-2 and M-5) exhibited similar efficacies, compared to regorafenib in *in vitro* and *in vivo* models.

Regorafenib was approved by the European Medicines Agency (EMA) in August 2013 as monotherapy for the treatment of adult patients with metastatic colorectal cancer (mCRC) who have been previously treated with, or are not considered candidates for, available therapies. 'Available therapies' are defined as those available prior to 2013 and include fluoropyrimidine-based chemotherapy, anti-VEGF therapy and anti-epidermal growth factor receptor (anti-EGFR) therapy (Table 2).[10]

Table 2: Technology being evaluated

UK approved name and brand name	UK approved name: Regorafenib Brand name: Stivarga®
Mechanism of action	Regorafenib is an orally available, small molecule, multitargeted TKI of multiple protein kinases, including kinases involved in tumour angiogenesis (VEGFR), oncogenesis (KIT, RET), and the tumour microenvironment (PDGFR, FGFR).[10]
Marketing authorisation/CE mark status	CHMP positive opinion: June 2013 EMA marketing authorization: August 2013
Indications and any restriction(s) as described in the summary of	Regorafenib is indicated as monotherapy for the treatment of adult patients with:

product characteristics (SmPC)	• mCRC who have been previously treated with, or are not considered candidates for, available therapies. These include fluoropyrimidine-based chemotherapy, an anti-VEGF therapy and an anti-EGFR therapy • Unresectable or metastatic GIST who progressed on or are intolerant to prior treatment with imatinib and sunitinib • HCC who have been previously treated with sorafenib
Method of administration and dosage	• Regorafenib should be prescribed by physicians experienced in the administration of anti-cancer therapy • Recommended dose of regorafenib is 160 mg (four tablets of 40 mg) taken once daily for 3 weeks followed by 1 week off therapy. Treatment should continue as long as benefit is observed or until unacceptable toxicity occurs • Dose interruptions and/or dose reductions may be required based on individual safety and tolerability. Dose modifications are to be applied in 40 mg (one tablet) steps. The lowest recommended daily dose is 80 mg. The maximum daily dose is 160 mg (table 1 and table 2 in the SmPC describes different dose adjustments according to AE severity)
Additional tests or investigations	Not applicable
List price and average cost of a course of treatment	List price: £3,744 (84 x 40mg tablets).
Patient access scheme (if applicable)	[redacted]
Key: CHMP, Committee for Medicinal Products for Human Use; EGFR, epidermal growth factor receptor; EMA, European Medicines Agency; FGFR, fibroblast growth factor receptor; GIST, gastrointestinal stromal tumours; HCC, hepatocellular carcinoma; KIT, stem cell factor receptor; mCRC, metastatic colorectal cancer; PDGFR, platelet-derived growth factor receptor; RET, rearranged during transfection; TKI, tyrosine kinase inhibitor; VEGF, vascular endothelial growth factor receptor.	

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

B.1.3.1. Disease overview and epidemiology

Colorectal cancer (CRC; also known as bowel cancer) is a heterogeneous disease characterized by cancerous growths in the large intestine and rectum.[1] The majority of CRCs arise from precursor lesions that have transformed to adenocarcinoma; > 90% are adenocarcinomas originating from epithelial cells.[1] Molecular carcinogenesis pathways have been implicated in the development of CRC including

chromosomal instability (observed in 65–70% of sporadic colorectal tumours) and microsatellite instability (MSI; observed in nearly 15% of sporadic colorectal tumours).[11] In addition, CRC tumours can present with identifiable gene mutations that can act as prognostic factors of prognosis (e.g. *BRAF* and *RAS* V600E mutations).[12]

CRC is the fourth most common cancer in the UK; there were 34,825 new cases in England in 2017, which accounted for 11% of all new cancer cases.[2, 3 20] The current incidence of CRC in England is 68 incidence cases per 100,000.[3]

CRC patient symptomology includes bleeding from the rectum, changes in bowel habit, pain in the abdomen or rectum, involuntary weight loss and anaemia.[13, 14] The incidence of CRC is strongly related to age with approximately 44% of new cases in people aged 75 and over.[2] Other risk factors for developing CRC include male gender, poor diet and obesity.[15-17] However, the overall clinical presentation of CRC depends on the size and location of the tumour, as well as the absence or presence of any metastases. Tumour staging is most often done with the tumour node metastasis (TNM) classification system[18] (Table 3), although other classifications including the Dukes staging[19] system can also be used. The TNM classification system was used to stage patients in the pivotal CORRECT and CONCUR trials (see Section B.2).

Table 3: Tumour staging of CRC

Stage	TNM staging and sites involved
0	Carcinoma in situ (Tis, N0, M0)
I	No nodal involvement, no distant metastases
	Primary tumour invading submucosa/muscle
II	No nodal involvement, no distant metastases
	Primary tumour penetrates into the outermost layers of the colon/rectum
	Primary tumour invades the wall of the colon/rectum but has not grown into other nearby tissues or organs
III	Primary tumour invading submucosa plus metastasis in 4–6 regional lymph nodes
	Primary tumour invading submucosa/muscle plus metastasis in of ≥ 7 regional lymph nodes
	Primary tumour adherent to other organs or structures plus involvement of regional lymph nodes
IV	Any tumour type with distant metastases confined to one organ or site (e.g. liver, lung, ovary, non-regional node)
	Any tumour type with distant metastases in ≥ 1 organ/site

Key: CRC, colorectal cancer; M0, no distant metastases TNM, tumour node metastasis; T0, no evidence of tumour, Tis, tumour in situ (abnormal cells present but may spread to neighbouring tissue, sometimes referred to as pre-invasive cancer); N0, no regional lymph node involvement.
Source: American Cancer Society 2020.[20]

Stage IV mCRC is an advanced form of CRC that has metastasized beyond the large intestine and nearby lymph nodes, typically spreading first to the liver (Table 3).[4, 5] Over the course of the disease, according to the ESMO guidelines, approximately 55% of all patients with Stage II and III CRC will develop metastases and progress to Stage IV mCRC.[21] At initial diagnosis, 23–26% of patients present with Stage IV disease, in which synchronous CRC liver metastases are present in 15–25% of cases, and metastases are confined to the liver in 70–80% of these cases.[22, 23] In England, the estimated annual number of mCRC cases is approximately 22,016 patients; of these, 20–25% will reach the third-line or beyond (≥ 3L) setting (the population of interest to this submission).

B.1.3.2. Survival outcomes for mCRC

Patients with Stage IV mCRC have very poor prognosis, with 1-year survival rates of approximately 44%, and 5-year survival rates of less than 10%.[6, 7] Furthermore, mortality rates for Stage IV mCRC are markedly higher than for those patients diagnosed at earlier stages of the disease.[7] For patients with Stage IV mCRC who do not receive active treatment (i.e. who are given best supportive care [BSC] without chemotherapy), life expectancy is only approximately 6 months.[24-26]

Even when treatment is given, patient prognosis decreases as the treatment line increases. For patients diagnosed at an early stage of the disease, surgery and adjuvant chemotherapy are the main treatments and can be curative in 50% of patients.[27] Upon disease progression or recurrence, FOLFOX (oxaliplatin, leucovorin, and 5-fluorouracil [5-FU]) or FOLFIRI (infusional folic acid/5-FU, and irinotecan) are accepted first- and second-line chemotherapy regimens, and these have demonstrated improved overall survival (OS) outcomes to an average of 20 months in first line, and 9–10 months in second line.[27] Survival outcomes in the ≥ 3L setting (the setting of interest to this submission) are typically in the region of 6–12 months.[8, 9]

The lack of treatment options for ≥ 3L mCRC is well-recognized, with clinicians in agreement that there is a high unmet clinical need for additional treatment options in ≥ 3L mCRC where different tolerability profiles impacts on the choice of medicine.

B.1.3.3. Burden of disease

Along with a significantly poor prognosis compared with earlier stages of the disease (see Section B.1.3.2), patients with Stage IV mCRC experience several burdensome symptoms including change in bowel habits, nausea, vomiting, constipation and fatigue. For those patients with liver metastases, pain on the right side of the abdomen, nausea, ascites and jaundice are also common.[28, 29] The high morbidity and mortality rates in mCRC are associated with a substantial patient burden with patients having a substantially impaired health-related quality of life (HRQL).[30] Psychological distress is common in mCRC with psychological symptoms of depression (ranging between 1.6% and 57.0%) and anxiety (1.0–47.2%) being highly prevalent.[31]

Another burden faced by patients is due to adverse events. Patients differ in their susceptibility and tolerance of different AEs, and it is often tolerability issues with prior therapies that heavily influences the next treatment choice.

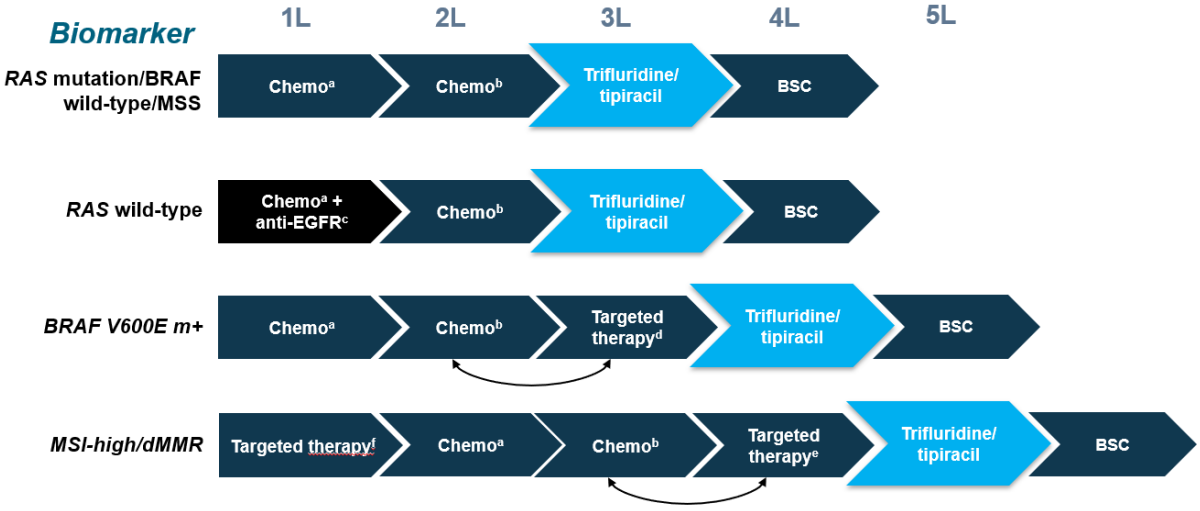
Metastatic CRC also places a substantial impact on caregivers. Patients with mCRC experience multiple symptoms that affect their daily living; therefore, family caregivers have to cope with demanding role changes that may affect their own mental and physical health.[32] Furthermore, increased caregiving responsibilities to the neglect of caregiver vital activities (such as self-care) have been associated with impaired QoL in caregivers.[32]

B.1.3.4. Clinical pathway of care

The management of mCRC is focused on improving QoL and survival through the management of local and metastatic disease.[33]

Figure 1 illustrates the treatment pathway for mCRC in England, based on guidance issued by the National Institute for Health and Care Excellence (NICE). Clinicians have confirmed that they follow the treatment pathway as outlined by NICE. In the NHS in England, treatment decisions for mCRC are based on genetic testing (biomarker driven) and treatment in later-lines is informed by prior therapy. First and second-line treatment of mCRC is dominated by chemotherapy combination regimens, which are typically FOLFOX or XELOX (oxaliplatin and capecitabine), and less commonly FOLFOXIRI (oxaliplatin, leucovorin, 5-FU, and irinotecan) which accounts for only 10% of all front-line treatments for mCRC). For patients with specific mutations, clinicians have the additional choice of including biologics in the first-line setting (Figure 1).

Figure 1: Treatment pathway in mCRC



Key: BSC, best supportive care; chemo, chemotherapy; dMMR, mismatch repair deficient; EGFR, epidermal growth factor receptor; mCRC, metastatic colorectal cancer; MSI-high, microsatellite instability-high; MSS, microsatellite stable.
Notes: [a] typically FOLFOX or XELOX, less common FOLFOXIRI; [b] irinotecan (after FOLFOX) or FOLFIRI (after FOLFOX or XELOX); [c] cetuximab or panitumumab; [d] encorafenib + cetuximab in second- or third-line depending on prior oxaliplatin and irinotecan exposure; [e] nivolumab + ipilimumab after fluoropyrimidine-based combination chemotherapy; [f] pembrolizumab; Figure adapted from the NICE pathway for colorectal cancer[18] and mCRC.[34]

Table 4 presents an overview of the NICE-recommended treatments for later-line mCRC. Trifluridine/tipiracil is recommended for adults who have had previous treatment with available therapies, or for whom these available therapies are not suitable. Trifluridine/tipiracil is the main comparator for regorafenib in previously treated mCRC.

Table 4: NICE recommended treatments for previously treated mCRC (Subsequent or alternative therapy [NG151])

Treatment	Details of recommendation
Biomarker guided	
Nivolumab and ipilimumab	• Nivolumab plus ipilimumab is recommended, within its marketing authorization, as an option for treating metastatic colorectal cancer with high MSI or MMR deficiency after fluoropyrimidine-based combination chemotherapy It is recommended only if the company provides nivolumab and ipilimumab according to the commercial arrangements
Encorafenib plus cetuximab	• Encorafenib plus cetuximab is recommended, within its marketing authorization, as an option for treating BRAF V600E mutation-positive mCRC in adults who have had previous systemic treatment It is recommended only if the company provides it according to the commercial arrangements
Biomarker independent	
Trifluridine/tipiracil	Trifluridine/tipiracil is recommended, within its marketing authorization, as an option for treating mCRC that is: • In adults who have had previous treatment with available therapies including fluoropyrimidine-, oxaliplatin- or irinotecan-based chemotherapies, anti-VEGF agents and anti-EGFR agents, or when these therapies are not suitable, and • Only when the company provides trifluridine–tipiracil with the discount agreed in the patient access scheme
Key: EGFR, epidermal growth factor receptor; mCRC, metastatic colorectal cancer; MSI, microsatellite instability; MMR, mismatch repair; VEGF, vascular endothelial growth factor. **Source:** NICE 2021.[34]	

B.1.3.5. Other clinical guidelines

The European Society for Medical Oncology (ESMO)[21, 27], the American Society of Clinical Oncology (ASCO)[35] and the National Comprehensive Cancer Network (NCCN)[33, 36] have guidelines for the treatment of mCRC. These guidelines are consistent with NICE's treatment pathway with the exception that regorafenib is recommended in later-line mCRC. The ESMO and ASCO guidelines position trifluridine/tipiracil and regorafenib alongside each other as options in the ≥3L setting.

In the ESMO 2016 consensus guidelines, regorafenib is recommended as a third-line treatment for patients previously treated with fluoropyrimidines, oxaliplatin, irinotecan, bevacizumab and EGFR antibodies (in *RAS* wild-type patients).[27] The ASCO guidelines for late-stage CRC published in 2020 also recommend regorafenib

as a third and fourth-line treatment.[35] The NCCN 2021 guidelines for colon cancer recommend regorafenib as a third and fourth-line treatment for patients with mCRC refractory to chemotherapy given before or after trifluridine/tipiracil.[33] Similarly, the NCCN 2021 guidelines for rectal cancer recommend regorafenib as a third and fourth-line treatment under the continuum of care for metastatic disease.[36]

B.1.3.6. Limitations with current treatments and remaining unmet needs in ≥ 3L mCRC

In England and Wales, patients who fail treatment with standard first and second-line therapies (chemotherapy and biological therapy) have limited options i.e.

1) biomarker targeted treatment: MSI or MMR deficiency - nivolumab and ipilimumab; BRAF V600E mutation positive - encorafenib plus cetuximab
2) the chemotherapy trifluridine/tipiracil.

Biomarker targeted therapies

The combination therapies of encorafenib plus cetuximab, and nivolumab plus ipilimumab are biological treatments that are restricted to patients whose tumours contain the relevant molecular alteration (*BRAF* V600E mutation and MSI/mismatch repair [MMR] deficiency, respectively). Thus, these targeted therapies are not available to all patients, as most patients with mCRC do not have these genetic alterations (it is estimated that only 5–21% of patients with mCRC have the *BRAF* mutations[12] and only approximately 4–5% have MSI/MMR deficiency[37]).

Trifluridine/tipiracil

Trifluridine/tipiracil is clinically effective and provides a meaningful increase in overall and progression free survival. However, being a chemotherapy, it is associated with multiple toxicities which are typical of these treatments. A proportion of patients will be reluctant/unable to start another chemotherapy based on their prior experience with this class of therapies and this thereby limits their options and ability to benefit from additional active treatment.

Unmet need

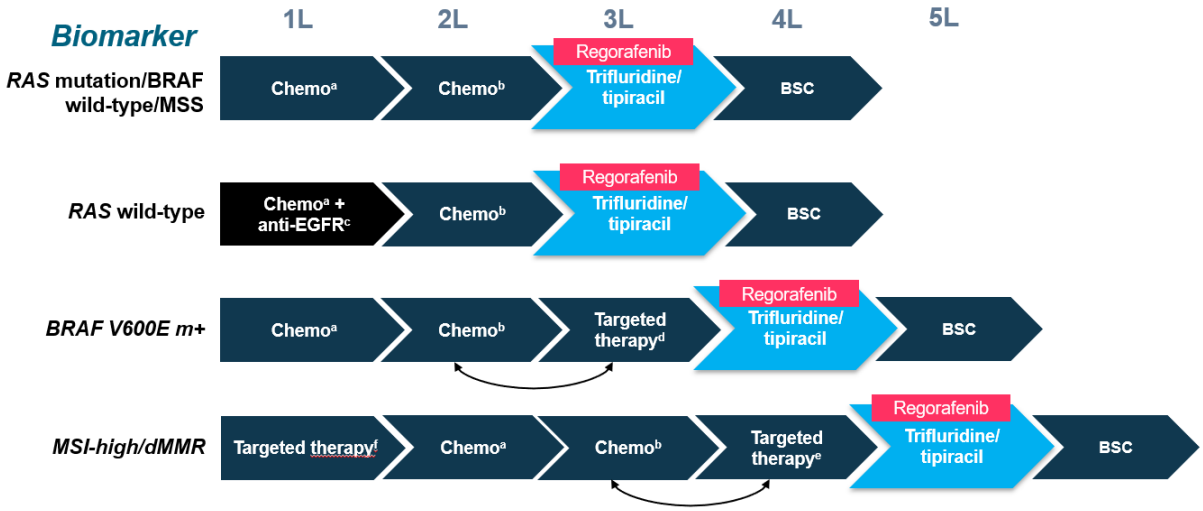
There is a lack of treatment options for patients in the ≥ 3rd line setting. For patients for whom biomarker targeted therapies are not options there is only the chemotherapy trifluridine/tipiracil. There is a need for a chemotherapy-free alternative therapy with a different adverse event profile which would increase the options available.

B.1.3.7. Proposed positioning of regorafenib

Bayer is making this submission on the request of physicians who want an alternative to the chemotherapy trifluridine/tipiracil in the ≥ 3rd line treatment setting. As such, this is a restricted submission and we are seeking a recommendation for regorafenib in a similar position to trifluridine/tipiracil.

Figure 2 presents the proposed positioning of regorafenib in the treatment pathway.

Figure 2: The proposed placement of regorafenib in the clinical pathway in England



Key: BSC, best supportive care; chemo, chemotherapy; dMMR, mismatch repair deficiency; EGFR, epidermal growth factor receptor; L, line; MSI, microsatellite instability; MSS, microsatellite stable.
Notes: [a] typically FOLFOX or XELOX, less common FOLFOXIRI; [b] irinotecan (after FOLFOX) or FOLFIRI (after FOLFOX or XELOX); [c] cetuximab or panitumumab; [d] encorafenib + cetuximab; [e] nivolumab + ipilimumab; [f] pembrolizumab. Figure adapted from the NICE pathway for colorectal cancer[18] and metastatic colorectal cancer.[34]

B.1.4. Equality considerations

No equality issues relating to the use of regorafenib have been identified.

B.2. Clinical effectiveness

B.2.1. Identification and selection of relevant studies

See Appendix D for full details of the process and methods used to identify and select the clinical evidence relevant to regorafenib.

B.2.2. List of relevant clinical effectiveness evidence

Table 5 summarizes the clinical-effectiveness evidence supporting regorafenib for the treatment of adult patients with mCRC receiving ≥ 3L therapy. The key data comes from the two pivotal randomized controlled trials (RCTs): CORRECT and CONCUR. Both trials were used to inform the cost-effectiveness model (see Section B.3).

Table 5: Clinical effectiveness evidence

Study	CORRECT: NCT01103323	CONCUR: NCT01584830
Study design	Phase III, randomized, double-blind, placebo-controlled, multi-centre study	Phase III, randomized, double-blind, placebo-controlled, multi-centre study
Population	Patients (aged ≥ 18 years) with Stage IV mCRC who had progressed on approved, standard treatments	Asian patients (aged ≥ 18 years) with Stage IV mCRC who had failed ≥ 2 lines of prior treatment
Intervention(s)	Regorafenib plus BSC	Regorafenib plus BSC
Comparator(s)	Placebo plus BSC	Placebo plus BSC
Indicate if study supports application for marketing authorisation	Yes	No
Indicate if study used in the economic model	Yes	Yes
Rationale if study not used in the model	Used in the economic model	Used in the economic model
Reported outcomes specified in the decision problem	• **Overall survival** • **Progression-free survival** • Response rate • **Adverse effects of treatment** • **Health-related quality of life**	• **Overall survival** • **Progression-free survival** • Response rate • **Adverse effects of treatment** • **Health-related quality of life**
All other reported outcomes	Not applicable	Not applicable
Sources of evidence	**Published data source:** • Grothey et al., 2013 (primary manuscript)[25] • Su et al., 2021[38] (secondary manuscript) • Battaglin 2020[39], Hofheinz 2019[40], Pericay 2014[41], Pasqualetti 2020[42], Tabernero 2015[43], Yoshino 2015[44] (linked studies)[b]	**Published data source:** • Li et al., 2015 (primary manuscript)[26] • Hofheinz 2019[40] and Xu 2020[47] (linked studies)[b] **Unpublished data sources:** • CSR[48]

Study	CORRECT: NCT01103323	CONCUR: NCT01584830
	Unpublished data sources: • CSR[45] • CSR safety addendum[46]	• CSR safety addendum[49]

Key: BSC, best supportive care; CSR, clinical study report; mCRC, metastatic colorectal cancer; RCT, randomized controlled trial.
Notes: [a] BSC 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 anti-neoplastic chemo/hormonal/immuno-therapy; [b], Note that these linked studies were identified from the clinical systematic literature review as reporting linked data on CORRECT and CONCUR, but they are not citied further within the submission.
Sources: CORRECT Clinical Study Report[45]; CONCUR Clinical Study Report[48]; Grothey et al., 2013[25]; Li et al., 2015.[26]

The key efficacy data presented in this document is taken from the primary publications for both trials, supplemented with data from the unpublished CSRs. Ten observational/RWE (non-randomised) studies were identified in the clinical SLR (Appendix D). These studies were not used to inform the economic model, but the results are presented in Section B.2.10 as supporting evidence to the pivotal trial data for regorafenib. Further details of these studies are presented in Table 17.

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

B.2.3.1. Trial methodology

Table 6 provides details of CORRECT and CONCUR trial methodology, and a brief narrative summary of each trial is presented below.

CORRECT

CORRECT was a Phase III, Global, randomized, double-blind, placebo-controlled trial, designed to evaluate the efficacy and safety of regorafenib and BSC (hereafter termed **regorafenib**) versus placebo plus BSC (hereafter termed **placebo**) in adult (≥ 18 years) patients with Stage IV mCRC (Figure 3). Patients had progressed after approved standard treatments available at the time (including fluoropyrimidine, oxaliplatin, irinotecan, bevacizumab and cetuximab/panitumumab). The study was conducted in 114 centres in 16 countries in North America, Europe, Asia and Australia.

Patient eligibility (Table 6) included histological or cytological documentation of adenocarcinoma of the colon or rectum, Eastern Cooperative Oncology Group performance status (ECOG PS) of 0 or 1, standard therapy discontinuation due to unacceptable toxicity, and disease progression during or within 3 months after the last administration of the last standard therapy. For patients who had received adjuvant oxaliplatin, trial eligibility required progression during or within 6 months of therapy completion, or retreatment with oxaliplatin-based therapy if progression was later than 6 months. Patients with previous regorafenib treatment or uncontrolled medical disorders were ineligible for the trial.

Randomized (2:1) patients received oral regorafenib 160 mg or matching placebo once daily (OD) on Days 1–21 of each 28-day cycle until the determination of Response Evaluation Criteria in Solid Tumours (RECIST)-defined disease progression (PD), clinical progression, the development of severe adverse events (AEs), withdrawal from the study, death, or a decision by the treating physician that discontinuation would be in the patient's best interest (Figure 3). Upon discontinuation of either regorafenib or placebo treatment, all patients were to enter

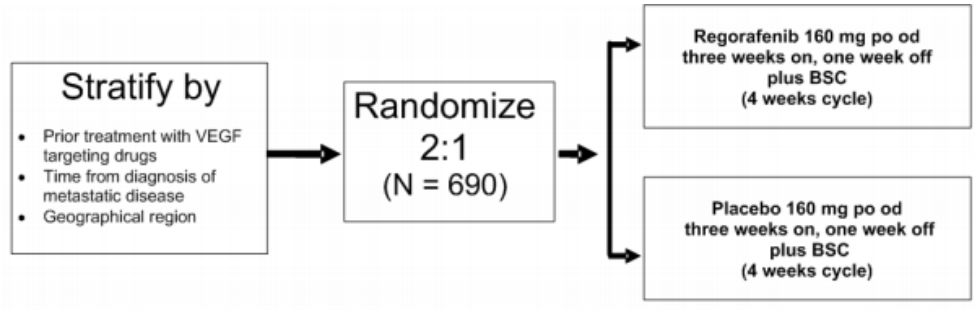
a 30-day safety follow-up period. Regardless of the reason for discontinuation, all patients were followed for survival until death was documented, except for those who specifically withdrew consent to follow-up. Patients who withdrew consent from study drug treatment were allowed to enter the survival follow-up period after signing a separate informed consent for survival follow-up.

All patients received BSC (excluding other investigational anti-tumour agents or antineoplastic chemotherapy, hormonal therapy, or immunotherapy). No crossover between treatment groups was allowed (Table 6). Predefined dose modifications were permitted to manage clinically significant treatment-related toxic effects (see Table 1 and Table 2 of the regorafenib SmPc). Patient randomization was stratified by:

- Previous treatment with VEGF-targeting drugs (yes or no)
- Time from diagnosis of metastatic disease (≥ 18 months versus < 18 months)
- Geographical region (North America, Western Europe, Israel, and Australia; Asia; and South America, Turkey and Eastern Europe)

The primary endpoint in CORRECT was OS, defined as the time from randomization to death from any cause. Secondary endpoints included progression-free survival (PFS), tumour response (objective tumour response rate [ORR]), disease control rate (DCR) and safety. Tertiary endpoints included duration of response (DOR), duration of stable disease and QoL. See Table 6 for details and definitions of all study endpoints.

Figure 3: CORRECT study schematic



Key: BSC, best supportive care; od, once daily; po, per os (by mouth); VEGF, vascular endothelial growth factor.
Source: Figure 7-1 of CORRECT Clinical Study Report (Final).[45]

CONCUR

CONCUR was a Phase III, randomized, double-blind, placebo-controlled, parallel-group trial designed to evaluate the efficacy and safety of regorafenib versus placebo in Asian patients with Stage IV mCRC. Patients had progressed after ≥ 2 lines of previous standard therapy. The trial was designed with a near identical protocol to the CORRECT trial but was exclusively focused on adult (≥ 18 years) Asian patients. The study was conducted at 25 hospitals in mainland China, Hong Kong, South Korea, Taiwan and Vietnam.

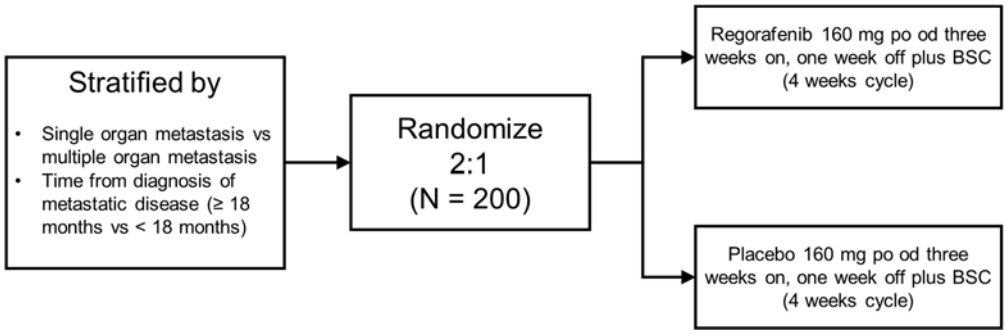
Patient eligibility (Table 6) was as per the CORRECT trial but specified Asian ethnicity and at least two previous treatment lines, including a fluoropyrimidine plus oxaliplatin or irinotecan. Previous treatment with bevacizumab, cetuximab or panitumumab was allowed but was not mandatory. Randomized (2:1) patients received oral regorafenib 160 mg or matching placebo OD on Days 1–21 of each 28-day cycle until the determination of RECIST-defined PD, clinical progression, the development of severe AEs, withdrawal from the study, death, or a decision by the treating physician that discontinuation would be in the patient's best interest (Figure 4). However, patients with PD could continue treatment at the investigator's discretion. Upon discontinuation of either regorafenib or placebo treatment, all patients were to enter a 30-day safety follow-up period. Regardless of the reason for discontinuation, all patients were followed for survival until death was documented, except for those who specifically withdrew consent to follow-up. Patients who withdrew consent from study drug treatment were allowed to enter the survival follow-up period after signing a separate informed consent for survival follow-up.

Similar to CORRECT, all patients received BSC (excluding other investigational anti-tumour agents or antineoplastic chemotherapy, hormonal therapy, or immunotherapy; Table 6) and predefined dose modifications were permitted to manage clinically significant treatment-related toxic effects (see Table 1 and Table 2 of regorafenib SmPc). However, in CONCUR, patient randomization was stratified by:

- Metastatic site (single versus multiple organs); and
- Time from diagnosis of metastatic disease (< 18 months versus ≥ 18 months)

The primary endpoint in CONCUR was OS (defined as the time from randomization to death from any cause) and secondary endpoints included PFS, tumour response (ORR), DCR and safety. Tertiary endpoints included DOR, duration of stable disease and QoL. See Table 6 for details and definitions of all study endpoints.

Figure 4: CONCUR study design



Key: BSC, best supportive care; Od, once daily; po, per os (by mouth).
Source: Figure 7-1 of CONCUR Clinical Study Report (Final).[48]

Table 6: Summary of trial methodology for CORRECT and CONCUR

Trial number (acronym)	CORRECT: NCT01103323	CONCUR: NCT01584830
Location	Global: 105 centres across 15 countries	Asia: Mainland China, Hong Kong, Taiwan, Vietnam, and South Korea
Trial design	Randomized, double-blind, placebo-controlled multi-centre Phase III study	Randomized, double-blind, placebo-controlled multi-centre Phase III study
Eligibility criteria for participants	**Key inclusion criteria:** • Adults (≥ 18 years; both sexes) with mCRC (Stage IV) and measurable or non-measurable disease according to RECIST criteria v1.1 and a life expectancy of at least 3 months • Histological or cytological documentation of adenocarcinoma of the colon or rectum. All other histological types were excluded • Progression during or within 3 months following the last administration of approved standard therapies, which was to include fluoropyrimidine, oxaliplatin, irinotecan, bevacizumab and cetuximab or panitumumab (if *KRAS* WT) • Patients treated with oxaliplatin in an adjuvant setting were to have progressed during or within 6 months of completion of adjuvant therapy • Patients who had progressed more than 6 months after completion of oxaliplatin containing adjuvant treatment were to be retreated with oxaliplatin-based therapy to be eligible • Patients who had withdrawn from standard treatment due to unacceptable toxicity warranting discontinuation of treatment and precluding	**Key inclusion criteria[b]:** • Asian adults (≥ 18 years; both sexes) with mCRC (Stage IV) and measurable or non-measurable disease according to RECIST criteria v1.1 and a life expectancy of at least 3 months • Histological or cytological documentation of adenocarcinoma of the colon or rectum. All other histological types were excluded • At least two lines of prior treatment have failed • Progression during or within 3 months following the last administration of approved standard therapies, which must have included fluoropyrimidine, oxaliplatin, irinotecan • Patients treated with oxaliplatin in an adjuvant setting must have progressed during or within 6 months of completion of adjuvant therapy • Patients who progressed more than 6 months after completion of oxaliplatin containing adjuvant treatment must have been retreated with oxaliplatin-based therapy to be eligible • Patients who had withdrawn from standard treatment due to unacceptable toxicity warranting discontinuation of treatment and precluding retreatment with the same agent prior to progression

Trial number (acronym)	CORRECT: NCT01103323	CONCUR: NCT01584830
	retreatment with the same agent prior to progression of disease were also eligible to enter the study • Patients with an unknown KRAS status at screening were to have received prior anti-EGFR treatment • ECOG PS ≤ 1 • Adequate bone marrow, liver and renal function within 7 days of starting to study treatment **Key exclusion criteria** • Previous treatment with regorafenib • Uncontrolled hypertension, congestive heart failure ≥ NYHA Class 2, unstable angina, arterial or venous thrombotic or embolic events such as cerebrovascular accident, deep-vein thrombosis or pulmonary embolism within the 6 months before start of study medication • Ongoing infection > Grade 2 CTCAE Version 3.0	of disease were also allowed into the study. Patients may have received prior treatment with bevacizumab, and/or cetuximab/ panitumumab (if KRAS WT) • ECOG PS ≤ 1 • Adequate bone marrow, liver and renal function within 7 days of starting to study treatment **Key exclusion criteria** • Previous treatment with regorafenib • Uncontrolled hypertension, congestive heart failure ≥ NYHA Class 2, unstable angina (angina symptoms at rest), new-onset angina (begun within the last 3 months) or myocardial infarction less than 6 months before start of study drug • Ongoing infection > Grade 2 CTCAE Version 3.0
Settings and locations where the data were collected	114 centres across 16 countries (number of centres in brackets): Japan (19), US (17), Germany (15), Italy (9), France (9), Spain (8), Belgium (6), Australia (5), Israel (5), Canada (5), the Czech Republic (2), the Netherlands (2), China (1), Hungary (1), and Switzerland (1). No patients were randomized in South America or Turkey.	Asia (number of patients in brackets): China (129), Hong Kong (23), South Korea (20), Taiwan (20), and Vietnam (12)
Study periods and trial drugs	**Study periods** • Screening: patient eligibility and enrolment • Study drug treatment period: patients underwent evaluations for safety and drug accountability every cycle	**Study periods** • Screening: patient eligibility and enrolment • Study drug treatment period: patients underwent evaluations for safety and drug accountability every cycle

Trial number (acronym)	CORRECT: NCT01103323	CONCUR: NCT01584830
	• 30-day safety follow-up period: All patients entered the follow-up period upon discontinuation of either regorafenib or placebo treatment until death was documented **Trial drugs** • Intervention (n = 505): regorafenib 160 mg od po; 3 weeks on therapy followed by 1 week off therapy to comprise a cycle of 4 weeks plus BSC • Comparator (n = 255): placebo 160 mg po od; 3 weeks on, 1 week off, plus BSC	• 30-day safety follow-up period: All patients entered the follow-up period upon discontinuation of either regorafenib or placebo treatment until death was documented **Trial drugs** • Intervention (n =136): regorafenib 160 mg od orally (po). Three weeks on therapy followed by 1 week off therapy to comprise a cycle of 4 weeks plus BSC • Comparator (n = 68): placebo 160 mg po od 3 weeks on, 1 week off, plus BSC
Concomitant medication	**Permitted concomitant medication** • Standard therapies for concurrent medical conditions. Prophylactic anti-emetics were permitted according to standard practice • Treatment with non-conventional therapies and vitamin/mineral supplements was acceptable provided that they did not interfere with the study endpoints, in the opinion of the Investigator. St John's wort was not permitted • Bisphosphonates • Patients who were therapeutically treated with an agent such as warfarin or heparin were allowed to participate provided that their medication dose and INR/PTT were stable **Non-permissible concomitant medications and procedures** • Systemic anti-cancer therapy including cytotoxic therapy, signal transduction inhibitors, immunotherapy and hormonal therapy • TKIs	**Permitted concomitant medication** • Standard therapies for concurrent medical conditions. Prophylactic anti-emetics were permitted according to standard practice • Treatment with non-conventional therapies and vitamin/mineral supplements was acceptable provided that they did not interfere with the study endpoints, in the opinion of the Investigator. St John's wort was not permitted • Bisphosphonates • Patients who were therapeutically treated with an agent such as warfarin or heparin were allowed to participate provided that their medication dose and INR/PTT were stable **Non-permissible concomitant medications and procedures** • Systemic anti-cancer therapy including cytotoxic therapy, signal transduction inhibitors, immunotherapy and hormonal therapy • TKIs

Trial number (acronym)	CORRECT: NCT01103323	CONCUR: NCT01584830
	• Bone marrow transplant or stem cell rescue • Concomitant palliative radiation therapy was allowed if the target lesion(s) were not included within the radiation field and no more than 10% of the bone marrow was irradiated • Use of biological response modifiers, such as G-CSF within 3 weeks of study entry. G-CSF and other haematopoietic growth factors were permitted during the study in the management of acute toxicity such as febrile neutropenia when clinically indicated or at the discretion of the Investigator. However, they could not be substituted for a required dose reduction. Patients taking chronic erythropoietin were permitted • Patients taking narrow therapeutic index medications were to be monitored proactively • All traditional medicines with an anti-cancer indication, including traditional Chinese medicine	• Bone marrow transplant or stem cell rescue • Concomitant palliative radiation therapy was allowed if the target lesion(s) were not included within the radiation field and no more than 10% of the bone marrow was irradiated • Use of biological response modifiers, such as G-CSF, within 3 weeks of study entry. G-CSF and other haematopoietic growth factors were permitted during the study in the management of acute toxicity such as febrile neutropenia when clinically indicated or at the discretion of the Investigator. However, they could not be substituted for a required dose reduction. Patients taking chronic erythropoietin were permitted • Patients taking narrow therapeutic index medications were to be monitored proactively • All traditional medicines with an anti-cancer indication, including traditional Chinese medicine
Primary outcomes	• OS: defined as the time (days) from randomization to death due to any cause • Patients alive at the time of analysis were censored at the last date known to be alive • If a patient was lost to follow-up and there was no contact after randomization, this patient was censored at Day 1	• OS: defined as the time (days) from randomization to death due to any cause • Patients alive at the time of analysis were censored at the last date known to be alive • If a patient was lost to follow-up and there was no contact after randomization, this patient was censored at Day 1
Other outcomes used in the economic model/specified in the scope	**Secondary efficacy endpoints** • PFS: defined as time (days) from date of randomization to date of first observed disease progression (radiological or clinical) or death due to any cause, if death occurred before progression was	**Secondary efficacy endpoints** • PFS: defined as time (days) from date of randomization to date of first observed disease progression (radiological or clinical) or death due to any cause, if death occurred before progression was

Trial number (acronym)	CORRECT: NCT01103323	CONCUR: NCT01584830
	documented, based on investigator assessment using RECIST v1.1 • ORR: percentage of patients with CR or PR as best overall response based on investigator assessment. A best overall response was defined for all patients, using the RECIST criteria, v1.1 • DCR: Percentage of patients whose best response was CR, PR or SD based on investigator assessment • Safety: type, frequency, and severity of AEs[b] **Tertiary efficacy endpoints** • DOR: defined as time (days) from the first documented objective response of PR or CR, whichever was noted earlier, to disease progression or death (if death occurred before progression) • Duration of stable disease: time (days) from randomization to date of disease progression or death (if death occurred before progression). This variable was only calculated for patients who failed to achieve a best response of PR or CR • PRO: HRQL and health utility values were measured using the EORTC QLQ-C30 (global health status/QoL) and EQ-5D, respectively. For the EORTC QLQ-C30 (range 0–100), higher scores on the functioning scales and the global health status/QoL scale represent a higher level of functioning and better HRQL. A change of ≥ 10 points on the EORTC QLQ-C30 scale is considered as clinically meaningful. For the EQ-5D, higher scores represent better health status. A change of 0.07 to 0.12 points on the EQ-5D index and a change	documented, on investigator assessment using RECIST v1.1 • ORR: percentage of patients with CR or PR as best overall response based on investigator assessment. A best overall response was defined for all patients, using the RECIST criteria, v1.1 • DCR: Percentage of patients whose best response was CR, PR or SD based on investigator assessment • Safety: type, frequency, and severity of AEs[b] **Tertiary efficacy endpoints** • DOR: defined as time (days) from the first documented objective response of PR or CR, whichever was noted earlier, to disease progression or death (if death occurred before progression) • Duration of stable disease: time (days) from randomization to date of disease progression or death (if death occurred before progression). This variable was only calculated for patients who failed to achieve a best response of PR or CR • PRO: HRQL and health utility values were measured using the EORTC QLQ-C30 and EQ-5D, respectively. For the EORTC QLQ-C30 (range 0–100), higher scores on the functioning scales and the global health status/QoL scale represent a higher level of functioning and better HRQL. A change of ≥ 10 points on the EORTC QLQ-C30 scale is considered as clinically meaningful. For the EQ-5D, higher scores represent better health status. A change of 0.07 to 0.12 points on the EQ-5D index and a change of 7 to 12 points on the VAS are considered as clinically meaningful

Trial number (acronym)	CORRECT: NCT01103323	CONCUR: NCT01584830
	of 7 to 12 points on the VAS are considered as clinically meaningful	
Pre-planned subgroups	**Subgroup analyses of OS and PFS** • Demographic information such as race, sex and age group (< 65, ≥ 65 years) • Region: Region 1 (North America, Western Europe, Israel and Australia), Region 2 (Asia) and Region 3 (South America, Turkey, and Eastern Europe) • Time from diagnosis of metastatic disease (≥ 18 months and < 18 months) • Prior systemic anti-cancer therapies: – Prior anti-VEGF therapy (yes/no) – Prior systemic anti-cancer therapies: – Prior anti-cancer drugs, categorized in the following groups: FOIB; FOIBE – Number of prior treatment lines (≤ 3, > 3) – Number of prior treatment lines for metastatic disease (≤ 3, > 3) • Historical *KRAS* mutation status • Further important baseline cancer characteristics of primary site of tumour (e.g. ECOG PS: 0 and 1) **Subgroup analyses of safety** • Age (years): < 65, ≥ 65 • BMI (kg/m^2): < 25, 25 ≤ BMI, < 30, 30 ≤ BMI • Sex • Race • Hepatic function at baseline: maximum of baseline AST and baseline ALT value ≤ 1.5 x ULN, 1.5 x ULN	**Subgroup analyses of OS and PFS** • Demographic information such as race, sex and age group (< 65, ≥ 65 years) • Time from diagnosis of metastatic disease (≥ 18 months and < 18 months) • Single organ metastasis or multiple organ metastasis • Prior systemic anti-cancer therapies (targeted therapies – yes/no) • Prior systemic anti-cancer therapies in the following four groups: – Patients without any preceding targeted therapy (no anti-VEGF, no anti-EGFR therapy) – Patients with prior anti-VEGF treatment but without prior anti-EGFR treatment – Patients with prior anti-EGFR treatment but without prior anti-VEGF treatment – Patients with prior anti-VEGF treatment AND with prior anti-EGFR treatment • Number of prior treatment lines (≤ 3, > 3) • Number of prior treatment lines for metastatic disease (≤ 3, > 3) • *KRAS* mutation status • ECOG PS (0 and 1) • *BRAF* mutation status • Region (China [mainland China, Hong Kong, and Taiwan] and Asia, other than China)

Trial number (acronym)	CORRECT: NCT01103323	CONCUR: NCT01584830
	< maximum of baseline AST and baseline ALT value ≤ 3 x ULN, 3 x ULN < maximum of baseline AST and baseline ALT value • Kidney function at baseline: normal/mildly impaired renal function (estimated glomerular filtration rate, eGFR ≥ 60 mL/min/1.73 m²) • Moderately impaired renal function (eGFR) • ECOG PS at baseline	**Subgroup analyses of safety** • Age (years): < 65, ≥ 65 • BMI (kg/m²): < 25, 25 ≤ BMI, < 30, 30 ≤ BMI • Sex • Race • Hepatic function at baseline: maximum of baseline AST and baseline ALT value ≤ 1.5 x ULN, 1.5 x ULN < maximum of baseline AST and baseline ALT value ≤ 3 x ULN, 3 x ULN < maximum of baseline AST and baseline ALT value • Kidney function at baseline: normal/mildly impaired renal function (estimated glomerular filtration rate, eGFR ≥ 60 mL/min/1.73 m²) • Moderately impaired renal function (eGFR • ECOG PS at baseline

Key: AE, adverse event; ALT, alanine aminotransferase; AST, aspartate aminotransferase; BMI, body mass index; BSC, best supportive care; CTCAE, Common Terminology Criteria for AEs; CR, complete response; DCR, disease control rate; DOR, duration of response; ECOG PS, Eastern Cooperative Oncology Group performance status; EGFR, epidermal growth factor receptor FOIB, fluoropyrimidine, oxaliplatin, irinotecan, bevacizumab; FOIBE, fluoropyrimidine, oxaliplatin, irinotecan, bevacizumab, anti-EGFR antibody; G-CSF, granulocyte-colony stimulating factor; HRQL, heath-related quality of life; INR, international normalized ratio; mCRC, metastatic colorectal cancer; NYHA, New York Heart Association; od, once a day; ORR, overall response rate; OS, overall survival; PFS, progression free survival; po, per os (oral); PR, partial response; PRO, patient reported outcomes; PTT, partial thromboplastin time; RECIST, Response Evaluation Criteria in Solid Tumours; SD, stable disease; TKI, tyrosine kinase inhibitor; ULN, upper limit of normal; VAS, visual analogue scale; VEGF, vascular endothelial growth factor; WT, wild type;
Notes: [a] Due to the similarity of treatment guidelines in Europe, the US and Asia (especially China), the design of this study is similar to that of CORRECT: patients with mCRC were included if their disease progressed during or within 3 months following the last administration of approved standard therapies. However, since the analysis of the real-life-situation in Asia had revealed that not all patients had access to treatment with targeted therapies, the study protocol allowed the inclusion also of those patients who had not been pre-treated with bevacizumab and/or Erbitux® (cetuximab)/Vectibix® (panitumumab) even if these drugs were approved but not available at study entry. The other inclusion and exclusion criteria roughly mirrored those of CORRECT. BSC 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 except other investigational anti-tumour agents or anti-neoplastic chemotherapies / hormonal therapies / immunotherapies. [b] AEs included acute renal failure or severe proteinuria, interstitial lung disease, acute

Trial number (acronym)	CORRECT: NCT01103323	CONCUR: NCT01584830
cardiac failure, clinically significant bleeding and severe skin infections (Stevens-Johnson Syndrome, erythema multiforme and toxic epidermal necrolysis) and acute liver failure. **Source:** CORRECT Clinical Study Report[45]; CONCUR Clinical Study Report[48]; Grothey et al., 2013[25]; Li et al., 2015.[26]		

B.2.3.2. Baseline characteristics

CORRECT

Table 7 provides a summary of the baseline characteristics, including demographic and clinical characteristics, for patients in CORRECT. Overall, baseline characteristics were balanced between the two arms; the majority of patients were men (regorafenib: 62%; placebo: 60%) and white (regorafenib: 78%; placebo: 79%). The median age was 61 years for both groups. The majority of patients had received four or more previous systemic anti-cancer therapies (regorafenib: 49%; placebo: 47%), and all patients had received prior bevacizumab (an anti-VEGF monoclonal antibody). There was a higher proportion of patients in the placebo group who had progression on bevacizumab, irinotecan and oxaliplatin (84%, 90% and 63%, respectively) than in the regorafenib group (80%, 80% and 55%, respectively). The rates of ECOG PS 1 (48% versus 43%) and a positive *BRAF* mutation (4% versus 2%) were similar between arms (Table 7).

Table 7: CORRECT – baseline characteristics (ITT)

Characteristic	Regorafenib (n = 505)	Placebo (n = 255)
Median age, years (IQR)	61 (54.0–67.0)	61 (54.0–68.0)
Sex, n (%)		
Men	311 (62)	153 (60)
Women	194 (38)	102 (40)
Race, n (%)		
White	392 (78)	201 (79)
Black	6 (1)	8 (3)
Asian	76 (15)	35 (14)
Other/ not specified	31 (6)	11 (4)
Region, n (%)		
North America, western Europe, Israel, Australia	420 (83)	212 (83)
Asia	69 (14)	35 (14)
Eastern Europe	16 (3)	8 (3)
ECOG performance status, n (%)		
0	265 (52)	146 (57)
1	240 (48)	109 (43)

Characteristic	Regorafenib (n = 505)	Placebo (n = 255)
Primary site of disease, n (%)[a]		
Colon	323 (64)	172 (68)
Rectum	151 (30)	69 (27)
Colon and rectum	30 (6)	14 (5)
***KRAS* mutation, n (%)[b]**		
No	205 (41)	94 (37)
Yes	273 (54)	157 (62)
Unknown	27 (5)	4 (2)
***BRAF* mutation, n (%)[c]**		
No	322 (96)	163 (98)
Yes	14 (4)	3 (2)
Histology, n (%)		
Adenocarcinoma	493 (98)	245 (96)
Adenocarcinoma in situ	2 (< 1)	3 (1)
Adenosquamous carcinoma	1 (< 1)	1 (< 1)
Carcinoma, not otherwise specified	4 (1)	1 (< 1)
Mucinous carcinoma	5 (1)	4 (2)
Undifferentiated carcinoma	0 (0)	1 (< 1)
Number of previous systemic anti-cancer therapies (on or after diagnosis of metastatic disease), n (%)		
1–2	135 (27)	63 (25)
3	125 (25)	72 (28)
≥ 4	245 (49)	120 (47)
Previous anti-VEGF treatment, n (%)		
Bevacizumab	505 (100)	255 (100)
Patients stopping previous treatment because of progression, n (%)		
Fluoropyrimidine	421 (83)	221 (87)
Bevacizumab	403 (80)	214 (84)
Irinotecan	405 (80)	229 (90)
Oxaliplatin	278 (55)	160 (63)
Panitumumab or cetuximab, or both	219 (43)	107 (42)
Time from diagnosis of metastases		
Median, months (IQR)	31.0 (20.6–43.3)	29.9 (20.2–46.4)
< 18 months, n (%)	91 (18)	49 (19)
≥ 18 months, n (%)	414 (82)	206 (81)

Key: ECOG, Eastern Cooperative Oncology Group; IQR, interquartile range; ITT, intention to treat; VEGF, vascular endothelial growth factor.
Notes: [a] Information missing in one patient in the regorafenib group; [b] *KRAS* mutation status was based on historical patient record; [c] *BRAF* mutation status was determined with plasma DNA samples collected from 336 patients in the regorafenib group and 166 in the placebo group; [d] Five patients in the placebo group (2%) and 16 patients in the regorafenib group (3%) had received only one previous line of treatment for metastatic disease.
Source: Table 1 of Grothey et al., 2013.[25]

CONCUR

Table 8 provides a summary of the baseline characteristics, including demographic and clinical characteristics, for patients in CONCUR (all patients were Asian). Overall, baseline characteristics were balanced between arms with two exceptions: sex and age category. There were more males than females (58% versus 42%) overall in the study; however, there was a higher percentage of males in the regorafenib group than in the placebo group (63% versus 49%). The mean age was 56.5 years; however, there were more patients aged < 65 years old in the placebo group than in the regorafenib group (85% versus 70%). The majority of patients had received four or more systemic anti-cancer therapies (regorafenib: 54%; placebo: 51%), and overall, 40% of patients had not previously received any biological treatment (anti-VEGF or otherwise) before randomization. However, more patients in the placebo arm (25%) than those in the regorafenib arm (18%) received an anti-EGFR but not anti-VEGF as a prior targeted therapy. The rates of ECOG PS 1 (74% versus 78%) and a positive *BRAF* mutation (0% versus 1%) were similar between arms (Table 8).

Table 8: CONCUR – baseline characteristics (ITT)

Characteristic, n (%)	**Regorafenib (n = 136)**	**Placebo (n = 68)**
Age		
Median age, years (IQR)	57.5 (50.0–66.0)	55.5 (48.5–62.0)
< 65, n (%)	95 (70)	58 (85)
≥ 65, n (%)	41 (30)	10 (15)
Sex, n (%)		
Men	85 (63)	33 (49)
Women	51 (38)	35 (51)
Region, n (%)		
China (mainland China, Taiwan, and Hong Kong)	112 (82)	60 (88)
Asia other than China	24 (18)	8 (12)
BMI, kg/m^2 (IQR)	23.1 (20.8–25.5)	22.8 (20.0–25.0)
ECOG PS, n (%)		
0	35 (26)	15 (22)
1	101 (74)	53 (78)
Main site of disease, n (%)		
Colon	79 (58)	48 (71)
Rectum	53 (39)	19 (28)

Characteristic, n (%)	Regorafenib (n = 136)	Placebo (n = 68)
Colon and rectum	4 (3)	1 (1)
***KRAS* mutation, n (%)**		
No	50 (37)	29 (43)
Yes	46 (34)	18 (26)
Unknown	40 (29)	21 (31)
***BRAF* mutation, n (%)**		
No	28 (21)	14 (21)
Yes	0 (0)	1 (1)
Unknown	108 (79)	53 (78)
Histology, n (%)		
Adenocarcinoma	130 (96%)	66 (97)
Mucinous carcinoma	6 (4)	2 (3)
Time from diagnosis of metastatic disease		
Median, months (IQR)	20.3 (13.8–28.8)	19.9 (13.3–27.7)
< 18 months, n (%)	53 (39)	32 (47)
≥ 18 months, n (%)	83 (61)	36 (53)
Number of metastatic sites, n (%)		
Single	28 (21)	15 (22)
Multiple	108 (79)	53 (78)
Previous targeted biological treatment, n (%)		
None	56 (41)	26 (38)
Any (anti-VEGF[a] or anti-EGFR[b], or both)	80 (59)	42 (62)
Anti-VEGF but not anti-EGFR	32 (24)	13 (19)
Anti-EGFR but not anti-VEGF	24 (18)	17 (25)
Anti-VEGF and anti-EGFR	24 (18)	12 (18)
Previous systemic anti-cancer treatment lines, n (%)		
Any intention		
2	31 (23)	14 (21)
3	32 (24)	19 (28)
≥ 4	73 (54)	35 (51)
On or after diagnosis of metastatic disease[c]		
1–2	48 (35)	24 (35)
3	32 (24)	17 (25)
≥ 4	52 (38)	27 (40)

Key: BMI, body mass index; ECOG PS, Eastern Cooperative Oncology Group performance status; EGFR, epidermal growth factor receptor; IQR, interquartile range; ITT, intention-to-treat; VEGF, vascular endothelial growth factor.
Notes: [a] bevacizumab; [b] cetuximab or panitumumab; [c] four patients (3%) in the regorafenib group had not previously received any treatment for metastatic disease.
Source: Table 1 of Li et al., 2015.[26]

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

B.2.4.1. Statistical analyses

Table 9 provides a summary of the statistical analysis and definitions of analysis sets in CORRECT and CONCUR. A brief narrative summary of the statistical analyses of each trial is presented below.

CORRECT

CORRECT was designed to have 90% power to detect a 33.3% increase in median OS, assuming a 4.5-month median OS for the placebo group (i.e. a hazard ratio [HR] of 0.75 for regorafenib over placebo). A total of 582 death events were required (one-sided Type 1 error rate of 0.025) and a total of 690 patients were planned for randomization (2:1). Two formal interim analyses were planned during the study when approximately 30% and 70% of the planned total number of required death events had occurred. Statistical evaluation was performed using the software package SAS® (Version 9.1 or higher). OS and PFS were compared between treatment groups with a stratified log-rank test; HRs (with 95% confidence intervals [CIs]) were calculated with the Cox model, adjusting for stratification factors; and Kaplan–Meier survival estimates were calculated for each treatment group. ORR and DCR were compared between treatment groups with the Cochran–Mantel–Haenszel test, adjusting for stratification factors. Safety data were listed and summarized descriptively for the safety population. AEs were coded according to the Medical Dictionary for Regulatory Activities (MedDRA) v14.1 and graded according to the National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE) v3.0.

The formal efficacy analyses were based on the intention-to-treat (ITT) population (data cut-off 21 July 2011). At the time of the analysis no patients had crossed over from placebo to regorafenib therapy.

Patients were followed on a monthly basis for safety until all patients were off treatment. Safety analyses were based on the safety analysis set (SAS) – data cut-off 27 March 2014.

CONCUR

CONCUR was designed to have 80% power to detect a 33.3% increase in median OS (i.e. an HR of 0.75). A total of 154 deaths were estimated to be required, with randomization planned for approximately 200 patients. Statistical evaluation was performed using the software package SAS. OS and PFS were compared between treatment groups with a stratified log-rank test; HRs (with 95% CIs) were calculated with the Cox model, adjusting for stratification factors; and Kaplan–Meier survival estimates were calculated for each treatment group. ORR and DCR were compared between treatment groups with the Cochran–Mantel-Haenszel test, adjusting for stratification factors. Safety data were listed and summarized descriptively for the safety population. AEs were coded according to MedDRA and graded according to NCI CTCAE v4.0.

The formal efficacy analyses were based on the full analysis set (FAS) – data cut-off 29 November 2013, which was defined as all randomized patients (Table 9).

Patients were followed on a monthly basis for safety until all patients were off treatment. All safety analyses were based on the safety analysis set (SAS) – data cut-off 14 January 2016.

B.2.4.1.1. Patient withdrawals (CORRECT and CONCUR)

Patients were allowed to withdraw from treatment at any time at their own request or withdraw at the discretion of the investigator or sponsor due to safety or behavioural reasons. Patients withdrawn from the study treatment were not replaced.

Table 9: Summary of statistical analyses for CORRECT and CONCUR

Trial	CORRECT	CONCUR
Hypothesis objective	To compare OS between the regorafenib group and placebo group, the following hypothesis was tested: • H_0: HR (regorafenib/placebo) ≥ 1 versus • H_1: HR (regorafenib/placebo) < 1	To compare OS between the regorafenib group and placebo group, the following hypothesis was tested: • H_0: HR (regorafenib/placebo) ≥ 1 versus • H_1: HR (regorafenib/placebo) < 1
Statistical analysis	**Main analyses** OS and PFS were compared between treatment groups with a stratified log-rank test; HRs (with 95% CIs) were calculated with the Cox model, adjusting for stratification factors; and KM survival estimates were calculated for each treatment group. ORR and DCR were compared between treatment groups with the Cochran–Mantel–Haenszel test, adjusting for stratification factors. **Subgroups** Forest plots, descriptive statistics and HR estimates with 95% CIs for OS and PFS were presented for predefined subgroups (provided there was a sufficient number of events in total within the subgroup across the treatment arms). Summaries of AEs were presented according to CTCAE v3.0 and MedDRA. **Sensitivity analyses** OS and PFS were tested with an unstratified log-rank test. Two sensitivity analyses of OS and PFS were performed: one on unstratified data and one using stratification information from the IVRS. In the sensitivity analyses of PFS, all available tumour assessment data were taken into account also from the follow-up period.	**Main analyses** OS and PFS were compared between treatment groups with a stratified log-rank test; HRs (with 95% CIs) were calculated with the Cox model, adjusting for stratification factors; and KM survival estimates were calculated for each treatment group. ORR and DCR were compared between treatment groups with the Cochran–Mantel–Haenszel test, adjusting for stratification factors. **Subgroups** Forest plots, descriptive statistics and HR estimates with 95% CIs for OS and PFS were presented for predefined subgroups. Summaries of AEs were presented according to CTCAE v4.0 and MedDRA. **Sensitivity analyses** OS and PFS were tested with an unstratified log-rank test. Three pre-specified sensitivity analyses of OS were performed: an unstratified analysis of OS, an analysis using stratification information from the IVRS, and an analysis stratified by previous targeted anti-cancer therapy (targeted therapy defined as anti-VEGF or anti-EGFR therapy or both). Two additional analyses of PFS were performed using a definition of PFS that included all assessments from follow-up and one that considered a new treatment initiation date in follow-up as the event date.

Trial	CORRECT	CONCUR
Analysis sets	**ITT analysis set:** all randomized patients including those who withdrew regardless of the reason for withdrawal. This was the primary population for all efficacy analyses.	**FAS:** all randomized patients. This set was the primary population for the efficacy analyses.
	SAS: all patients who received at least one dose of study medication.	**SAS:** all patients who received at least one dose of study medication.
	PK analysis set: all patients with available PK data collected after at least 14 days of uninterrupted stable dosing of regorafenib.	**PK analysis set:** all patients with available PK data collected after at least 14 days of uninterrupted stable dosing of regorafenib.
	Biomarker analysis set: all patients with available biomarker data and signed consent form for the analyses.	**Biomarker analysis set:** all patients with available biomarker data and signed consent form for the analyses.
Sample size, power calculation	**Sample size and power calculation** CORRECT was designed to have 90% power to detect a 33.3% increase in median OS, assuming a 4.5-month median OS for the placebo group (i.e. an HR of 0.75 for regorafenib over placebo). Assuming a one-sided overall α of 0.025, a power of 90%, a randomization ratio of 2:1 between regorafenib and placebo, and two formal interim analyses of OS during the study, with an O'Brien–Fleming-type error spending function, the study required 582 deaths for the final analysis. A total of 690 patients were planned for randomization. **Interim analyses** Two formal interim analyses were planned during the study when approximately 30% (first interim) and 70% (second interim) of the planned total number of required death events had occurred. The first formal interim analysis was for futility only. The second interim analysis was for efficacy and futility. A Lan–Demets alpha spending function determined the monitoring boundary for efficacy, so the overall false positive rate (α) was ≤ 0.025 (one-sided). The alpha spending function was the O'Brien–Fleming type boundary specified. Boundaries were specified to stop the	**Sample size and power calculation** The sample size was based on the primary endpoint of OS. A total of 200 patients and 154 death events were required, assuming a target increase in median OS of 33.3% (i.e. an HR of 0.75, regorafenib over placebo), one-sided overall α of 0.2, a power of 80% and a randomization ratio of 2:1 between regorafenib and placebo. It was projected that 154 events would occur after approximately 19 months, assuming a monthly patient enrolment rate of 33 patients/month and 200 patients were randomized after an initial 6 months ramp-up period, a dropout rate of 3%, exponentially distributed event times for OS, and 4.5 and 6 month median OS time for the placebo and the regorafenib groups, respectively.

Trial	CORRECT	CONCUR
	study for efficacy or futility on the basis of the actual number of events included in the analysis. At the second interim analysis, the study was to be stopped for futility if the HR (regorafenib over placebo) was 0.9006 or greater, and for efficacy if the one-sided p value was less than or equal to 0.009279, roughly corresponding to an HR (regorafenib over placebo) of less than or equal to 0.7864.	
Data management, patient withdrawals	**Censoring methods** For the primary endpoint of OS, patients alive at the time of analysis were censored at the last date they were known to be alive. If a patient was lost to follow-up and there was no contact after randomization, this patient was censored at Day 1. Patients with evidence of being alive as of the database cut-off date were censored using the cut-off date of 21 July 2011. Standard censoring methods were applied to PFS and response (ORR, DOR, duration of stable disease) analyses for those patients without (or missing) evaluable assessments. **Missing data** Patients withdrawn from study treatment were not replaced, and missing data were not estimated or carried forward in any statistical analysis (unless otherwise stated). No imputation was performed for missing assessments.	**Censoring methods** For the primary endpoint of OS, patients alive at the time of analysis were censored at the last date they were known to be alive. Patients with evidence of being alive as of the database cut-off date were censored using the cut-off date of 29 November 2013. Standard censoring methods were applied to PFS and response (ORR, DOR, duration of stable disease) analyses for those patients without (or missing) evaluable assessments. **Missing data** Patients withdrawn from study treatment were not replaced, and missing data were not estimated or carried forward in any statistical analysis (unless otherwise stated). No imputation was performed for missing assessments.

Key: AE, adverse event; CI, confidence interval; CTCAE, Common Terminology Criteria for AEs; DCR, disease control rate; DOR, duration of response; FAS, full analysis set; HR, hazard ratio; ITT, intention to treat; IVRS, Interactive Voice Response System; KM, Kaplan–Meier; MedDRA, Medical Dictionary for Regulatory Activities; ORR, overall response rate; OS, overall survival; PFS, progression-free survival; PK, pharmacokinetic; SAS, safety analysis set.
Sources: Section 7.6 of CORRECT Clinical Study Report (Final)[45]; Section 7.6 of CONCUR Clinical Study Report (Final)[48]; Grothey et al., 2013[25]; Li et al., 2015.[26]

B.2.4.2. Patient disposition data

Appendix D provides the CONSORT diagrams for CORRECT and CONCUR. A brief overview of the patient disposition for each trial is summarized below.

CORRECT

A total of 1,052 patients completed screening, of whom 760 were randomized (regorafenib, n = 505; placebo, n = 255 [ITT population]) between 19 May 2010 and 22 March 2011. Most randomized patients were enrolled in Italy (n = 143), France (n = 116), Japan (n = 100), Belgium (n = 87), Spain (n = 83), US (n = 83) and Germany (n = 64). Overall, 753 patients initiated treatment (regorafenib: n = 500; placebo: n = 253 [safety population]). Of the seven patients who did not receive treatment, four patients (n = 2 for each group) had AEs that were associated with clinical disease progression and one patient (regorafenib group) had an AE that was not associated with clinical disease progression. Of the remaining two patients who did not receive treatment, one patient had an ECOG PS of > 1 after randomization, and the other patient withdrew consent.

In total, 88.7% and 95.7% of patients in each arm (regorafenib versus placebo), respectively, discontinued treatment. 540 patients (71.1%) entered post-treatment survival follow-up (regorafenib: n = 353; placebo: n = 187). In addition, a total of 21 patients were permitted to continue treatment after PD was proven as they derived clinical benefit from treatment. At the efficacy data cut-off date (21 July 2011), 61 patients (regorafenib: n = 52; placebo: n = 9) were still receiving treatment.

CONCUR

A total of 234 patients were screened, of whom 204 were randomized (regorafenib: n = 136; placebo: n = 68 [FAS population]) between 29 April 2012 and 19 January 2013. The primary reasons for screening failure were not meeting study eligibility criteria (15.2%) and patient withdrawal (0.8%). All 204 randomized patients were included in the FAS population and also in the safety population.

At the efficacy data cut-off of 29 November 2013, the majority of patients (regorafenib: 95.6%; placebo: 100%) had discontinued treatment, while six patients

in the regorafenib group continued treatment. By the safety data cut-off date (14 January 2016), all 204 patients in the study had terminated treatment. Safety follow-up data were available for 192 patients.

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

Quality assessment of CORRECT and CONCUR was conducted using the NICE checklist, the full details of which are provided in Appendix D. Both trials were approved by the institutional review board or independent ethics committee for each study centre and were both conducted according to Good Clinical Practice. Overall, both trials were considered to be methodologically robust, high-quality studies with a comprehensive approach to patient allocation, control of confounding factors, and an overall low risk of bias.[25, 26]

Overall, the results of both studies are expected to be generalizable to patients in England and Wales. There were a two main areas where the trials diverged from the baseline characteristic of patients in England and Wales i.e. in respect of ethnicity and prior treatment.

Ethnicity

Regarding applicability of the trials to English clinical practice, CONCUR enrolled exclusively Asian patients, while CORRECT enrolled European and Asian patients from across 16 countries including Western Europe. Thus, it was possible to observe whether Asian and non-Asian patients responded to regorafenib in a similar manner; indeed, subgroup analyses of CORRECT and CONCUR confirmed that race was not a treatment effect modifier for regorafenib[49]. Clinicians have confirmed that ethnicity is not expected to be a treatment effect modifier (see Section B.2.7). It is therefore reasonable to generalize the results of both studies to Western populations.

Prior treatments: Anti-VEGF

In CORRECT, all patients had previously received anti-VEGF targeted therapy (i.e. bevacizumab), whilst in CONCUR, patients were enrolled who had not received previous treatment with bevacizumab.

Bevacizumab (an anti-VEGF) is not recommended by NICE (either as monotherapy or as part of a combination regimen) for patients with mCRC.[34] Thus, of the two trials, CONCUR more closely aligns with clinical practice in England and Wales as it includes a significant proportion of patients who have never received anti-VEGF therapy. Prior treatment with anti-VEGF is relevant as regorafenib has anti-VEGF activity (see Section B.1.2) - the implication of this prior therapy is that regorafenib could be expected to be less effective in patients who have already been treated with, and failed on, an anti-VEGF.

Prior treatments: number received for mCRC

CORRECT recruited a more heavily pre-treated population (Table 7) than CONCUR (Table 8). It is recognised that prognosis, and ability to benefit from treatment diminishes with each additional line of therapy. On this basis a difference in efficacy between CORRECT and CONCUR might be anticipated.

B.2.6. Clinical effectiveness results of the relevant trials

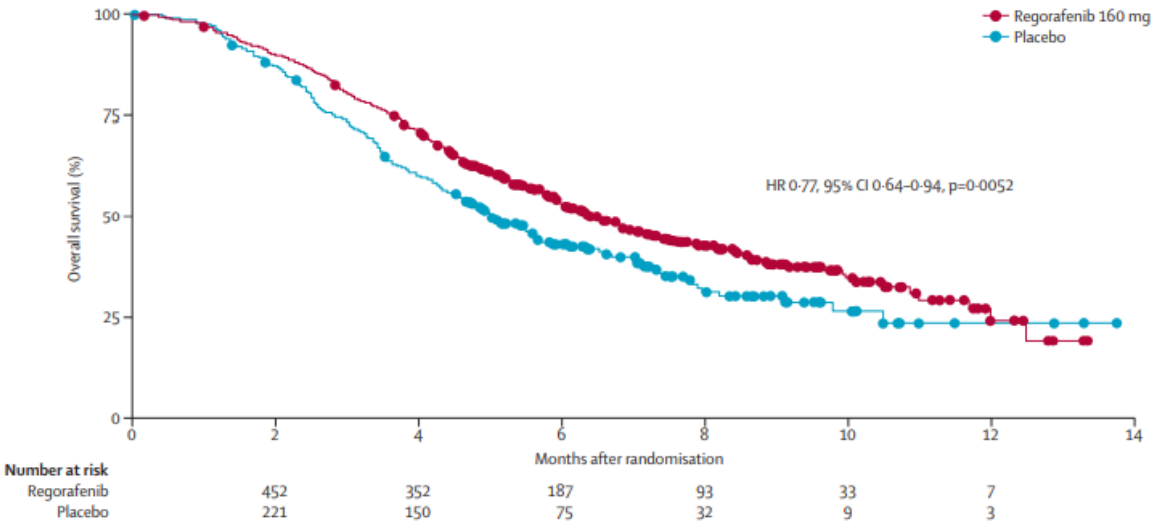
B.2.6.1. CORRECT

B.2.6.1.1. Primary efficacy outcome: overall survival

The median duration of treatment was 2.8 versus 1.8 months for regorafenib versus placebo. A total of 432 death events occurred in the ITT population (n = 760), with the majority occurring in the placebo group (regorafenib: 54.5%; placebo: 61.6%; Appendix M; Table 79).

In CORRECT, the primary endpoint was met. Regorafenib significantly prolonged OS compared with placebo (HR 0.77; 95% CI: 0.64, 0.94; p = 0.0052; Appendix M; Table 79 and Figure 5). [redacted] of patients had experienced an OS event (Appendix M; Table 79), and median OS was 6.4 versus 5.0 months, respectively. The OS rate was higher in the regorafenib group than in the placebo group at 3 (80.3% versus 72.7%), 6 (52.5% versus 43.5%), 9 (24.3% versus 24.0%) and 12 (24.3% versus 24.0%) months post-randomization, respectively. [redacted].

Figure 5: CORRECT – Kaplan–Meier curve of overall survival (ITT)



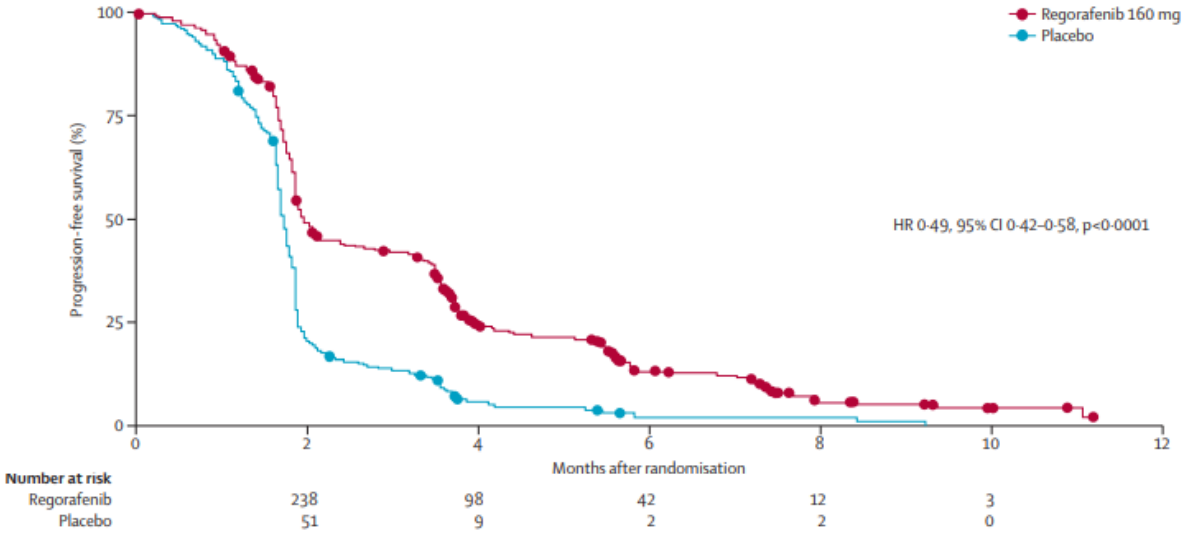
Key: CI, confidence interval; HR, hazard ratio; ITT, intention-to-treat.
Source: Figure 2. Grothey et al., 2013.[25]

B.2.6.1.2. Secondary efficacy outcomes

B.2.6.1.2.1. Progression-free survival

PFS was significantly longer in the regorafenib group compared with placebo (HR 0.49; 95% CI: 0.42, 0.58; p < 0.0001; Figure 6; and Appendix M; Table 81). [redacted] of patients had experienced a PFS event (Appendix M; Table 81), and the median PFS was 1.9 versus 1.7 months, respectively. PFS was consistently higher in the regorafenib group than in the placebo group at 3, 6 and 9 months post-randomization: [redacted] respectively. [redacted].

Figure 6: CORRECT – Kaplan–Meier curve of progression-free survival (ITT)



Key: CI, confidence interval; HR, hazard ratio; ITT, intention-to-treat.
Note: Kaplan–Meier curves based on investigator assessment
Source: Figure 3. Grothey et al., 2013.[25]

B.2.6.1.2.2. Objective response rate and disease control rates

In CORRECT, no patients had a complete response (CR); five patients assigned regorafenib and one patient assigned placebo had a partial response (PR), giving ORRs of 1.0% and 0.4%, respectively (p = 0.19; Table 10). Compared with placebo, regorafenib was associated with significant disease control (41% [n = 207] versus 15% [n = 38]; p < 0.0001). In addition, more patients in the regorafenib group compared with placebo had a shrinkage in tumour size ([redacted]% versus [redacted]%; Appendix M; Table 83).

Table 10: CORRECT – tumour response (ITT)

Response	Regorafenib (n = 505)	Placebo (n = 255)
Best response		
CR, n (%) [95% CI]	[redacted]	[redacted]
PR, n (%) [95% CI]	[redacted]	[redacted]
SD, n (%) [95% CI]	[redacted]	[redacted]
PD, n (%) [95% CI]	[redacted]	[redacted]
Non CR/Non PD, n (%)[a] [95% CI]	[redacted]	[redacted]
ORR and DCR, n (%) [95% CI]		

	Regorafenib (n = 505)	Placebo (n = 255)	P-value
ORR[b]	[redacted]	[redacted]	[redacted]
DCR[c]	[redacted]	[redacted]	[redacted]

Key: CI, confidence interval; CR, complete response; DCR, disease control rate; ITT, intention-to-treat; PD, progressive disease; PR, partial response; SD, stable disease.
Notes: [a] non CR/non PD included in DCR and followed the same criteria as stable disease; [b] percentage of patients with CR or PR; [c] percentage of patients with CR or PR or SD according to RECIST v1.1. Patients with SD as response performed earlier than 6 weeks after randomization were not taken into account. Non-CR/non-PD were included in disease control rate and followed same criteria as stable disease.
Source: Tables 9-6 and 9-7 of CORRECT Clinical Study Report (Final).[45]

B.2.6.1.3. Tertiary outcomes

B.2.6.1.3.1. Duration of response and duration of stable disease

For the [redacted] in the regorafenib group and the [redacted] in the placebo group who had a PR, the median duration of response could not be estimated in the regorafenib group due to the small number of patients in the group (range without censored values: [redacted] days) and was [redacted] days for the one patient in the placebo group (Appendix M; Table 84). Median duration of stable disease was 2.0 months (interquartile range [IQR]: 1.7–4.0) in the regorafenib group and 1.7 months (IQR: 1.4–1.9) in the placebo group (Appendix M; Table 84).

B.2.6.1.3.2. Patient-reported outcomes

Appendix M: Tables 85 and 86 provides a summary of the patient-reported outcomes (PROs); please refer to Table 6 for the PRO instruments ranking rules.

In brief, the mean EORTC QLQ-C30 score at baseline (Day 1, Cycle 1) was 62.6 and 64.7 in the regorafenib and placebo groups, respectively. The mean scores at the end of treatment (EOT) visit were 48.9 and 51.9, respectively. The mean EQ-5D™ index scores were similar between the two groups at baseline (regorafenib: 0.73; placebo: 0.74) and EOT (0.59 for both groups). The mean EQ-5D visual analogue scale (VAS) scores were also similar between arms at baseline (regorafenib, 65.4; placebo, 65.8) and at EOT (regorafenib, 55.5; placebo, 57.3).

The five dimensions of the EQ-5D-3L were further evaluated.[38] In the analysis conducted by Su et al., the authors calculated the proportion of ITT patients reporting 'no problems' in each of the five dimensions (mobility, self-care, usual activities, pain/discomfort and anxiety/depression; dimensions can be graded as 'no problem', 'some problems' or 'severe problems'). At baseline, 95% of the ITT population completed the EQ-5D-3L questionnaire; patients reported the highest percentage of 'no problems' in self-care (87%) and lowest in pain/discomfort (35%) (Table 11). EQ-5D completion rates declined with each cycle of treatment, as did the percentages of patients reporting 'no problems' in each dimension; notably, the declines were more rapid for placebo than for regorafenib. Both completion rate and percentages of 'no problems' were somewhat higher for placebo than for regorafenib at baseline. This continued only until Cycle 2, and reversed completely from Cycle 3 onward – showing an increasingly diminished proportion of those on placebo able to complete the questionnaire and reporting 'no problem' in each dimension. The differences between arms became larger with each cycle beginning Cycle 3 (Table 11). Overall, the results indicate that regorafenib appears to enable patients to maintain their mobility, self-care and usual activity, and to remain pain-free and anxiety-free, as well as their ability to complete the EQ-5D questionnaire, at much higher rates than placebo.[38]

Table 11: CORRECT – patients completing the EQ-5D-3L questionnaire (ITT)

Cycle	EQ-5D Data (% completed)		Mobility No Problem		Self-Care No Problem		Usual Activity No Problem		Pain No Problem		Anxiety No Problem	
	PBO	REG	PBO	REG	PBO	REG	PBO	REG	PBO	REG	PBO	REG
1	95%	95%	70%	65%	88%	86%	60%	56%	37%	34%	62%	57%
2	78%	74%	53%	43%	70%	62%	46%	34%	27%	21%	51%	42%
3	23%	39%	17%	23%	22%	35%	16%	19%	9%	13%	14%	22%
4	13%	35%	10%	19%	12%	33%	8%	17%	6%	11%	9%	22%
5	6%	19%	4%	11%	5%	17%	3%	8%	3%	7%	4%	12%
6	4%	17%	3%	10%	4%	15%	3%	8%	2%	5%	3%	11%
7	1%	9%	1%	6%	1%	8%	1%	5%	1%	3%	1%	6%
8	1%	7%	1%	4%	1%	7%	1%	3%	0%	3%	0%	5%
Source: Su et al., 2021.[38]												

B.2.6.1.4. Overview of efficacy

The Phase III CORRECT trial demonstrated a significant prolongation of survival outcomes with the addition of regorafenib to BSC in patients with relapsed/refractory mCRC who had had the option of receiving all approved standard therapies that were available at the time of study initiation. The primary endpoint of OS was in favour of regorafenib compared with BSC alone with an HR of 0.77 (95% CI: 0.64, 0.94), which translates into a 23% reduction in risk of death with regorafenib during the course of the study. Similarly, analyses of PFS ($p < 0.0001$) and disease control ($p < 0.0001$) were consistently in favour of regorafenib over BSC alone, with the data indicating that the main effect of regorafenib on mCRC appears to be disease stabilization. Sensitivity analyses (both stratified and unstratified) of OS and PFS were consistent with the main findings and predefined subgroup analyses of OS and PFS were aligned with the main findings, demonstrating consistent benefit of regorafenib over BSC.

Overall, the observed benefit in OS and PFS is clinically significant and meaningful in a patient population with very poor prognosis and limited treatment options for whom prior therapies have failed.

B.2.6.2. CONCUR

B.2.6.2.1. Primary efficacy outcome: overall survival

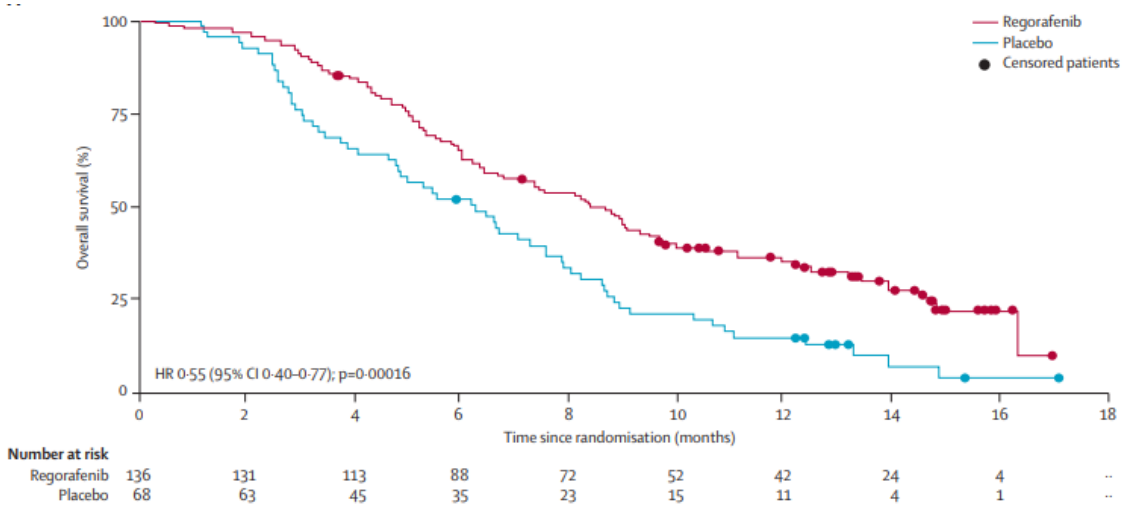
The median duration of treatment was 2.4 versus 1.6 months for regorafenib versus placebo. A total of 155 death events occurred in the FAS population (n = 204), with the majority occurring in the placebo group (regorafenib: 70%; placebo: 88%). Median follow-up for the OS analysis was 7.4 months (IQR: 4.3–12.2).

In CONCUR, the primary endpoint was met. Regorafenib significantly prolonged OS compared with placebo (HR 0.55; 95% CI: 0.40, 0.77; p = 0.00016; Figure 7; and Appendix M; Table 89). [redacted] of patients had experienced an OS event (Appendix M; Table 89), and median OS was 8.8 versus 6.3 months, respectively. The OS rate was higher in the regorafenib group than in the placebo group at 3 ([redacted]% versus [redacted]%), 6 ([redacted]% versus [redacted]%), 9 ([redacted]% versus [redacted]%) and 12 ([redacted]% versus [redacted]%) months post-randomization,

respectively. [redacted]

[redacted]

Figure 7: CONCUR – Kaplan–Meier curve of overall survival (FAS)



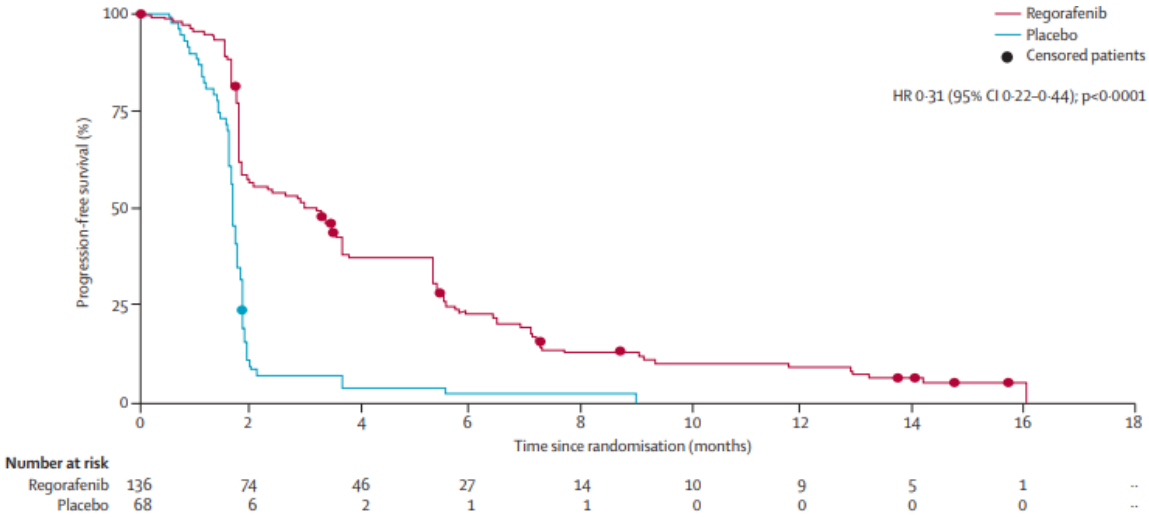
Key: CI confidence interval; FAS, full analysis set; HR, hazard ratio.
Source: Figure 2. Li et al., 2015.[26]

B.2.6.2.2. Secondary efficacy outcomes

B.2.6.2.2.1. Progression-free survival

PFS was significantly longer in the regorafenib group compared with placebo (HR 0.31; 95% CI: 0.22, 0.44; p < 0.0001; Figure 8; and Appendix M; Table 91). [redacted] of patients had experienced a PFS event (Appendix M; Table 91), and the median PFS was 3.2 versus 1.7 months, respectively. [redacted]

Figure 8: CONCUR – Kaplan–Meier curve of progression-free survival (FAS)



Key: CI confidence interval; FAS, full analysis set; HR, hazard ratio.
Note: Kaplan–Meier curves based on investigator assessment
Source: Figure 3. Li et al., 2015.[26]

B.2.6.2.2.2. Objective response rate and disease control rates

In CONCUR, no patients had a CR. Six patients assigned regorafenib had a PR, while no patients receiving placebo had a PR. More patients receiving regorafenib than placebo had stable disease as best overall response (45.6% versus 7.4%), and the ORR was 0% versus 4.4%, respectively (Table 12). Compared with placebo, regorafenib was associated with significant disease control (51.5% [n = 70] versus 7.4% [n = 5]; p < 0.0001). In addition, more patients in the regorafenib group compared with the placebo group had a shrinkage in tumour size ([redacted]% versus [redacted]%; Appendix M; Table 93).

Table 12: CONCUR – tumour response (FAS)

Response[a]	Regorafenib (n = 136)	Placebo (n = 68)	
Best response			
CR, n (%) [95% CI]	[redacted]	[redacted]	
PR, n (%) [95% CI]	[redacted]	[redacted]	
SD[b], n (%) [95% CI]	[redacted]	[redacted]	
PD, n (%) [95% CI]	[redacted]	[redacted]	
Non CR/Non PD, n (%)[a] [95% CI]	[redacted]	[redacted]	
ORR and DCR, n (%) [95% CI]			
	Regorafenib (n = 136)	**Placebo (n = 68)**	**P-value**
ORR[c]	[redacted]	[redacted]	[redacted]
DCR[d]	[redacted]	[redacted]	[redacted]

Key: CI, confidence interval; CR, complete response; DCR, disease control rate; FAS, full analysis set; N, number of patients; ORR, overall response rate; PD, progressive disease; PR, partial response; SD, stable disease.
Notes: [a] 95% CI by exact binomial calculation. Non-CR/non-PD included in disease control rate and followed same criteria as stable disease. PD and non-CR/non-PD assessed after at least 6 weeks. Denominator for rates (%) and 95% CIs based on FAS (all randomized patients); [b] number of weeks without progression; [c] percentage of patients with CR or PR; [d] percentage of patients with CR or PR or SD according to RECIST v1.1. Patients with stable disease as response performed earlier than 6 weeks after randomization were not taken into account. Non-CR/non-PD were included in disease control rate and followed same criteria as stable disease.
Source: Tables 9-6 and 9-7. CONCUR Clinical Study Report (Final).[48]

B.2.6.2.3. Tertiary outcomes

B.2.6.2.3.1. Duration of response and duration of stable disease

The median DOR in the six patients in the regorafenib group who had a PR was 4.8 months. The median duration of stable disease was 3.0 months in the regorafenib group and 1.7 months in the placebo group (Appendix M; Table 94).

B.2.6.2.3.2. Patient-reported outcomes

Appendix M; Tables 95 and 96 provide a summary of the PROs; please refer to Table 6 for the PRO instruments ranking rules.

In brief, the mean EORTC QLQ-C30 score at baseline (Day 1, Cycle 1) was 66.7 and 58.0 in the regorafenib and placebo groups, respectively. The mean scores at the EOT visit were 51.1 and 52.2, respectively. The mean EQ-5D index scores were similar in the two groups at baseline (regorafenib: 0.84; placebo: 0.75)) and EOT (0.57 for both groups). The mean EQ-5D VAS scores were also similar between arms at baseline (regorafenib: 73.4; placebo: 71.4) and EOT (regorafenib: 61.5; placebo: 62.6).

B.2.6.2.4. Overview of efficacy conclusions

The Phase III CONCUR trial demonstrated that, for patients with relapsed/refractory mCRC, the addition of regorafenib to BSC resulted in a clinically relevant and statistically significant prolongation of survival outcomes. Compared with placebo, regorafenib significantly prolonged OS (HR 0.55; 95% CI: 0.40,0.77; p = 0.000159) and PFS (HR 0.31; 95% CI: 0.22,0.44; p < 0.0001) and was associated with statistically significantly improved disease control rates (p < 0.0001). Overall, the subgroup and sensitivity analyses were consistent and supportive of the main survival (OS and PFS) outcomes.

CONCUR and CORRECT were similarly designed studies, however a larger benefit was observed in CONCUR. Clinical experts have advised that efficacy is not expected to differ according to ethnicity so the difference is not explained by CONCUR being conducted in an Asian population. Two possible explanations, aside from random fluctuation (i.e. the confidence intervals are overlapping), are related to prior treatment i.e. the total number of prior treatments received; and the proportion of patients pre-treated with anti-VEGF.

- CORRECT recruited a more heavily pre-treated population (Table 7) than CONCUR (Table 8). It is understood that prognosis, and ability to benefit from treatment diminishes with each additional line of treatment. *Ceteris paribus* you might anticipate a better efficacy from CONCUR on this basis.
- All patients in CORRECT had previously received anti-VEGF therapy (i.e. bevacizumab), whilst in CONCUR, patients were enrolled who had not received previous treatment with bevacizumab. Prior treatment with anti-VEGF is relevant as regorafenib has anti-VEGF activity (see Section B.1.2) - the implication of this prior therapy is that regorafenib could be expected to be

less effective in patients who have already been treated with, and failed on, an anti-VEGF (as was the case in the CORRECT study). As a consequence, the results seen in CONCUR could be more representative of expectations for patients in the NHS.

B.2.7. Subgroup analysis (CORRECT and CONCUR)

In both CORRECT and CONCUR, subgroup analyses of OS and PFS were pre-planned and were prespecified, as listed in the statistical analysis plan (SAP). For both trials, descriptive statistics and HR estimates with 95% CIs for OS and PFS were calculated for each baseline stratification factor and other relevant baseline variables, provided there was a sufficient number of events in total within the subgroup across the treatment arms. Forest plot representations were also provided. Subgroup analyses were performed based on demographic information (e.g. race, sex, age group [< 65 years, ≥ 65 years]), region (Region 1: North America, Western Europe, Israel, and Australia; Region 2; Asia; Region 3, South America, Turkey, and Eastern Europe), time from diagnosis of metastatic disease (≥ 18 months and < 18 months), prior systemic anti-cancer therapies, historical *KRAS* mutation status, and baseline cancer characteristics of primary site of tumour (e.g. ECOG PS: 0 or 1). Subgroup analyses of AEs were also conducted.

Overall, in both trials, the efficacy subgroup analyses demonstrated consistent survival benefits with regorafenib over placebo, with OS and PFS outcomes that were generally comparable with those observed in the overall populations.

A summary of results for subgroups analysed in both trials are provided in Appendix E.

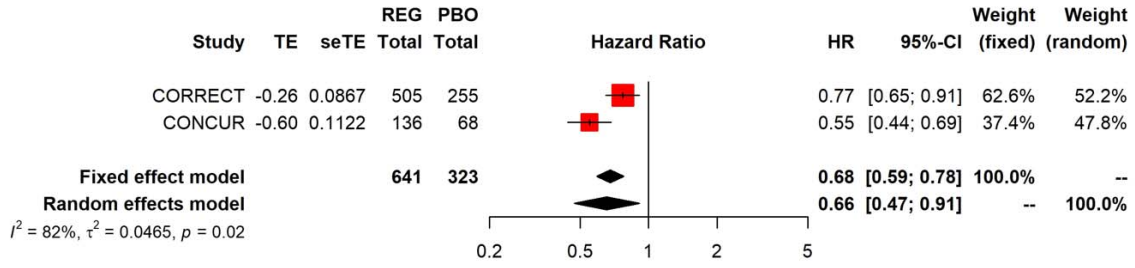
B.2.8. Meta-analysis

A direct meta-analysis using fixed and random effects models for OS and PFS data from CORRECT and CONCUR was performed. The generic inverse variance method for meta-analysis was performed using the '*meta*' package in R.[50] Results are shown in Figure 9 and Figure 10, respectively.

The fixed and random effects models estimated similar OS and PFS HRs for regorafenib versus placebo. For OS, the fixed effects HR was 0.68 (95% CI: 0.58, 0.79) and for PFS it was 0.42 (95% CI: 0.39, 0.45).

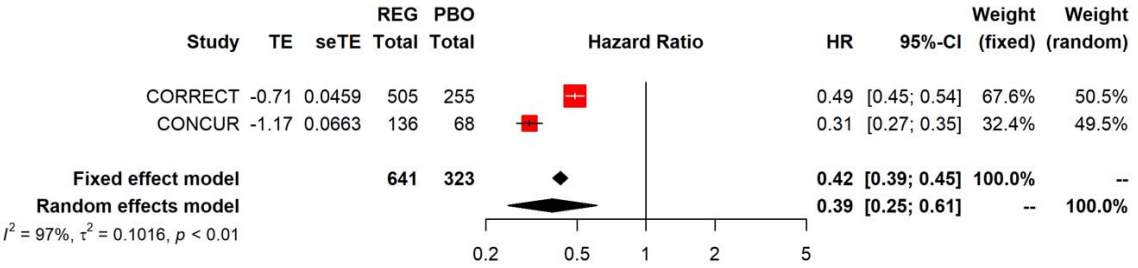
The I2 statistic was estimated at 82% in the analysis of OS and 97% in the analysis of PFS, which suggests a high level of heterogeneity between the studies. Study heterogeneity was explored in detail in the context of indirect treatment comparisons (B.2.9) and studies were considered sufficiently similar for meta-analysis (see Appendix D for further details).

Figure 9: Direct meta-analysis – overall survival – CORRECT and CONCUR



Key: CI, confidence interval; HR, hazard ratio; PBO, placebo; REG, regorafenib; seTE, standard error of treatment effect; TE, treatment effect.

Figure 10: Direct meta-analysis – progression-free survival – CORRECT and CONCUR



Key: CI, confidence interval; HR, hazard ratio; PBO, placebo; REG, regorafenib; seTE, standard error of treatment effect; TE, treatment effect.

B.2.9. Indirect and mixed treatment comparisons

In the absence of direct head-to-head evidence comparing the efficacy of regorafenib directly with that of trifluridine/tipiracil, indirect treatment comparisons (ITCs) were performed in line with NICE Decision Support Unit (DSU) technical support document (TSD) 2 to estimate relative treatment effects for OS and PFS.

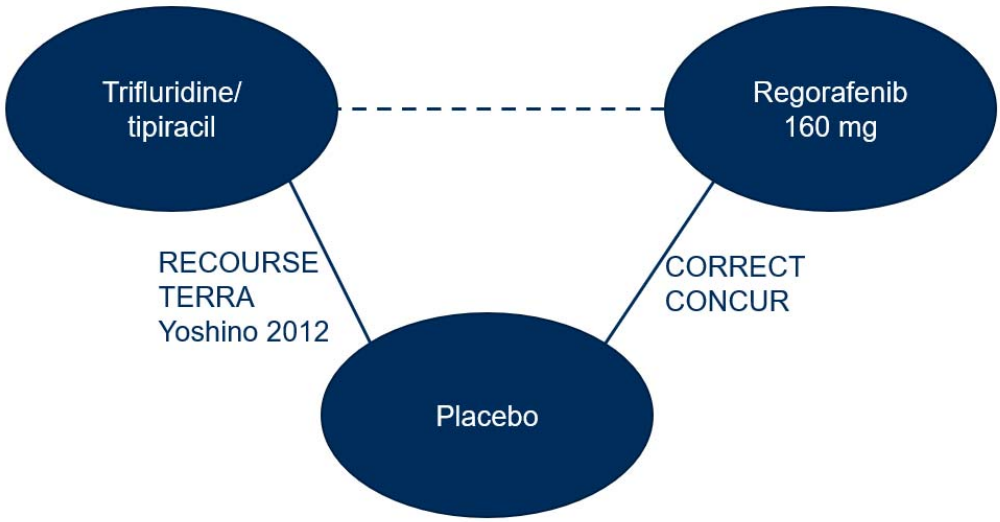
Evidence for the efficacy and safety of regorafenib and trifluridine/tipiracil were identified through a clinical systematic literature review (SLR). Studies included in the ITCs were placebo controlled RCTs. Table 13 summarises the studies included in the ITCs and a summary of the excluded studies can be found in Appendix D (Section B.3.1.8). Data from the primary publications were used for both regorafenib and trifluridine/tipiracil to avoid bias from open-label follow-up. All studies included a best supportive care arm.

Table 13: Summary of the trials used to carry out the indirect treatment comparisons

References of trial	Regorafenib	Trifluridine/ tipiracil	Study Phase	Blinding	Study centre(s)
CORRECT[25]	Yes		III	Double-blind	Multicentre: Europe, North America, Australia, Japan and China
CONCUR[26]	Yes		III	Double-blind	Multicentre: China, Hong Kong, South Korea, Taiwan and Vietnam
RECOURSE[51]		Yes	III	Double-blind	Multicentre: Europe, the US, Australia and Japan
TERRA[52]		Yes	III	Double-blind	Multicentre: China, the Republic of Korea and Thailand
Yoshino 2012[53]		Yes	II	Double-blind	Multicentre: Japan

Assessment of study heterogeneity (see section B.2.9.1) identified some differences in study design, inclusion/exclusion criteria, and baseline characteristics between the five studies outlined in Table 13, however, the studies were considered to sufficiently satisfy the basic assumptions of homogeneity, similarity, and consistency as reported in Song et al. 2009.[54] Fixed effects NMAs were therefore performed using the network presented in Figure 11. A fixed effects NMA, rather than a random effects NMA, was deemed more suitable given the small number of studies included in the network. Full details of the methods are included in Appendix D (B.3.1.8.2)

Figure 11: Network diagram for OS and PFS



Key: OS, overall survival; PFS, progression-free survival.

Table 14 presents the results from the fixed effects NMA of OS comparing regorafenib with trifluridine/tipiracil and placebo in the study ITT populations. Regorafenib was associated with a reduction in the hazard of death compared with placebo (HR: 0.68 [95% CrI: 0.59, 0.78]). This is consistent with the results observed in the clinical trials investigating regorafenib; CORRECT HR: 0.77 (95% CI: 0.66, 0.94) and CONCUR HR: 0.55 (95% CI: 0.39, 0.81). Trifluridine/tipiracil was also associated with a reduction in the hazard of death compared with placebo (HR: 0.68 [95% CrI: 0.62, 0.76]). Finally, results from the NMA suggested that the effectiveness of regorafenib and trifluridine/tipiracil for OS was similar, with a small but non-significant numerical advantage for regorafenib (HR was close to 1 and the CrI included 1).

Table 14: Results of the fixed effects NMA of OS

Comparison	HR (95% CrI)
Regorafenib versus placebo	0.68 (0.59, 0.78)
Trifluridine/tipiracil versus placebo	0.68 (0.62, 0.76)
Regorafenib versus trifluridine/tipiracil	0.99 (0.84, 1.17)
Key: CrI, credible interval; HR, hazard ratio; NMA, network meta-analysis; OS, overall survival.	

Table 15 presents the results from the fixed effects NMA of PFS comparing regorafenib with trifluridine/tipiracil and placebo for the ITT population. Regorafenib was associated with a reduction in the hazard of progression compared with placebo (HR: 0.42 [95% CrI: 0.39, 0.45]). This is consistent with the results observed in the clinical trials; CORRECT HR: 0.49 (95% CI: 0.42, 0.58) and CONCUR HR: 0.31 (95% CI: 0.22, 0.44). Trifluridine/tipiracil was also associated with reduction in the hazard of progression compared with placebo (HR: 0.45 [95% CrI: 0.42, 0.48]). Finally, results from the NMA suggested that the effectiveness of regorafenib and trifluridine/tipiracil for PFS was similar (HR was close to 1 and the CI included 1).

Table 15: Results of the fixed effects NMA of PFS

Comparison	HR (95% CrI)
Regorafenib versus placebo	0.42 (0.39, 0.45)
Trifluridine/tipiracil versus placebo	0.45 (0.42, 0.48)
Regorafenib versus trifluridine/tipiracil	0.93 (0.85, 1.03)
Key: CrI, credible interval; HR, hazard ratio; NMA, network meta-analysis; PFS, progression-free survival.	

To test the sensitivity of the fixed-effects NMAs including the five studies (CORRECT, CONCUR, RECOURSE, TERRA and Yoshino 2012), the analyses as outlined in Table 16 were performed to allow for differences between studies that were identified in the assessment of study heterogeneity.

Table 16. ITC sensitivity analyses

	Regorafenib		Trifluridine/tipiracil			
	CORRECT	CONCUR	RECOURSE	Yoshino 2012	TERRA	Brief Explanation
Base case	x	x	x	x	x	
SA1	x	x	x		x	Removal of Yoshino 2012 as the only phase II trial
SA2	x		x			All patients had been received prior anti-VEGF in these two trials
SA3		x		x	x	These 3 studies included Asian only patients. Treatment most closely aligned with UK clinical practice
SA4		x			x	Phase III trials most closely aligned to UK clinical practice
SA5		x		x		Mainly for completeness i.e. Asian-only study and complements SA4

- Sensitivity analysis 1: Fixed-effects NMA including CORRECT, CONCUR, RECOURSE and TERRA

Yoshino 2012 was removed from the network to explore whether the difference in phase of study impacted the results (CORRECT, CONCUR, RECOURSE and TERRA were all Phase III studies compared with Yoshino 2012 which was Phase II). Yoshino 2021 was also the only study to be conducted in one country only (i.e. in Japan).

- Sensitivity analysis 2: Bucher ITC including CORRECT and RECOURSE

All patients in CORRECT and RECOURSE had received a prior targeted biologic treatment (prior bevacizumab) compared with CONCUR, TERRA and Yoshino 2012 where around 40%, 50% and 20% of patients had received prior targeted biologic treatment, respectively. In the UK bevacizumab is not recommended. Prior targeted

biological treatment was identified as a potential treatment effect modifier by clinical experts. However, further sensitivity analyses adjusting for this variable (using matching-adjusted indirect treatment comparison methods) had very little impact on results (more information on identification of treatment effect modifiers and results from the matching-adjusted indirect treatment comparison is detailed in Appendix D [Section B.3.1.8.1]).

- Sensitivity analysis 3: Fixed-effects NMA including CONCUR, TERRA and Yoshino 2012

CORRECT and RECOURSE included patients from across the world (14% and 34% of patients were Asian, respectively), whereas CONCUR, TERRA and Yoshino 2012 only included Asian patients. Despite there being no evidence that ethnicity is prognostic or a treatment effect modifier this sensitivity analysis was conducted for completeness. In addition, some patients in each of these trials had not received prior treatment with anti-VEGF (anti-VEGF not recommended in the UK) and therefor these 3 trials are more representative of treatment in the UK setting.

- Sensitivity analysis 4: Bucher ITC including CONCUR and TERRA

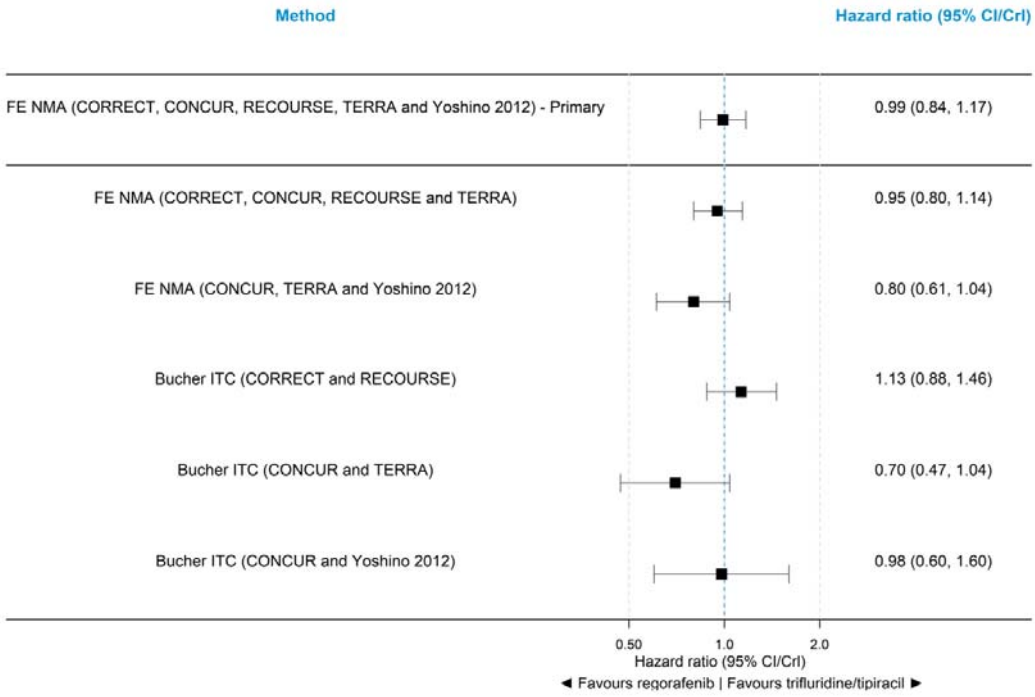
The patient characteristics from TERRA and CONCUR are considered to be most similar to those seen in UK clinical practice in terms of patients being naïve to targeted biologic treatment (52.7% had no prior targeted therapy).

- Sensitivity analysis 5: Bucher ITC including CONCUR and Yoshino 2012

Conducted for completeness in respect of Asian-only trials.

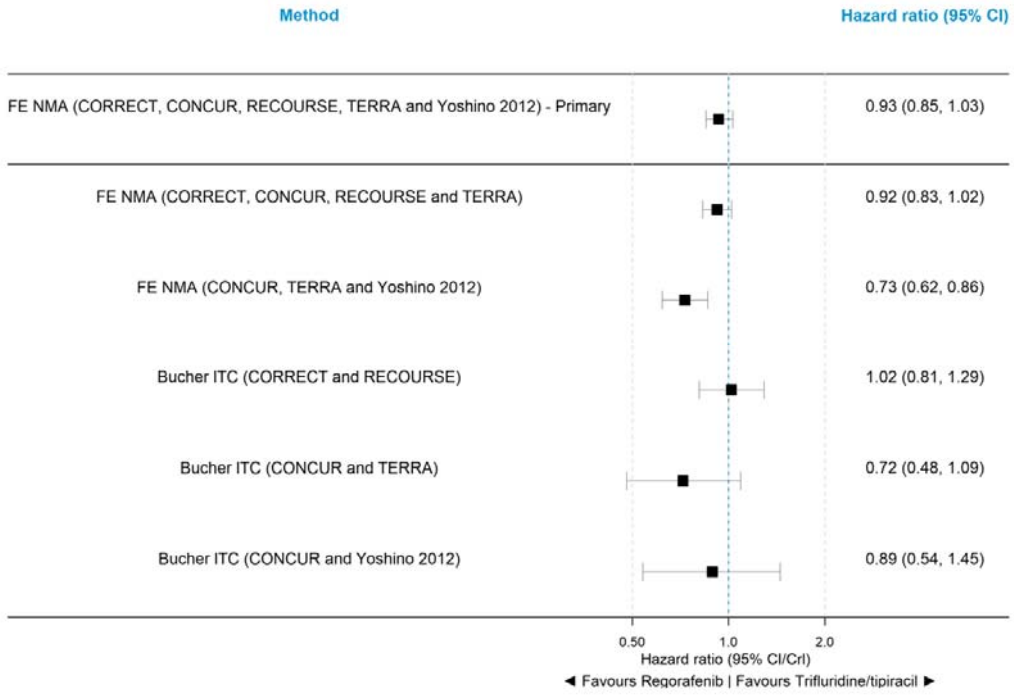
Results for OS and PFS from these sensitivity analyses are summarized in Figure 12 and Figure 13, respectively. Overall, sensitivity analysis results for both OS and PFS were consistent with the primary analysis and indicated comparable efficacy.

Figure 12: Overview of sensitivity analysis results for OS



Key: CI, confidence interval; CrI, credible interval; FE, fixed effects; ITC, indirect treatment comparison; NMA, network meta-analysis; OS, overall survival.

Figure 13: Overview of sensitivity analysis results for PFS



Key: CI, confidence interval; CrI, credible interval; FE, fixed effects; ITC, indirect treatment comparison; NMA, network meta-analysis; PFS, progression-free survival.

B.2.9.1. Uncertainties in the indirect and mixed treatment comparisons

The credible intervals for the comparisons of regorafenib with placebo and trifluridine/tipiracil with placebo from the NMA are slightly smaller than the CIs observed in the clinical trials. This is likely due to the fixed effects NMA underestimating the uncertainty of the point estimates and is a limitation of this NMA. While a random effects NMA could address this slight underestimation, it was not deemed feasible to conduct this NMA given the small number of studies included in the analysis, which would likely lead to implausibly wide credible intervals when using a non-informative prior to estimate the between-trial heterogeneity.

Assessment of study heterogeneity identified some differences across the CORRECT, CONCUR, RECOURSE, TERRA and Yoshino 2012 studies. The main differences related to phase of study (phase II/III), ethnicity of patients (some studies were conducted in Asians only) and prior targeted biologic treatments (see Appendix D) . Only prior targeted biologic treatment was identified as a potential treatment effect modifier and differences in the distributions of patients with prior targeted biologic treatment across studies may have introduced some bias into the analyses. However, across sensitivity analyses (that [1] used studies with similar patients with/without prior targeted biologic treatment and that [2] matched on this variable using matching-adjusted indirect treatment comparison methods – Appendix D) results remained consistent indicating that the primary analysis results were robust despite the differences observed.

B.2.10. Supportive evidence

Ten observational/real-world studies, presenting evidence on the effectiveness of regorafenib in clinical practice, were identified in the clinical SLR (Appendix D). An overview of real-world evidence from these observational studies is presented in Table 17, with a brief overview of the findings in Section B.2.10.1.

Table 17: List of observational studies (real-world evidence) of regorafenib

Study	Design	Objective	Population (and prior treatments)	Prior treatment eligibility criteria	Primary study reference
CORRELATE (NCT02042144)	Observational, prospective study conducted in 13 countries in Europe, Latin America and Asia	To characterize the safety and effectiveness of regorafenib in real-world clinical practice	n = 1,037; mCRC patients previously treated with approved therapies • received ≥4 prior systemic anti-cancer therapies: 39%	The study population comprised patients with mCRC who were previously treated with, or who were not considered candidates for, other approved therapies and for whom a decision was made by the treating physician to treat with regorafenib according to the local health authority approved label.	Ducreux et al., 2019[55]
REBECCA (NCT02310477)	Observational; an ATU cohort study nested within a French compassionate use programme	Designed to evaluate survival, safety and potential prognostic factors for outcomes associated with regorafenib in patients with mCRC refractory to standard therapies	n = 690; patients with histologically proven mCRC enrolled from a compassionate use programme • ≥3 prior lines • of treatment for metastatic disease: 35% • ≥5 prior lines • of treatment for metastatic disease: 15%	Study was designed to designed to evaluate the efficacy and safety of regorafenib in real-life clinical practice for mCRC patients who have been previously treated with or are not considered candidates for standard therapies.	Adenis et al., 2016[56]

Study	Design	Objective	Population (and prior treatments)	Prior treatment eligibility criteria	Primary study reference
			• received prior oxaliplatin and irinotecan: 99% • received prior bevacizumab: 92% • patients with KRAS wild-type tumours had previously received anti-EGFR therapy: 97%		
RECORA (NCT01959269)	Prospective, multi-centre, non-interventional study	To characterize the efficacy and safety of regorafenib under routine daily clinical practice conditions in Germany	n = 481; Patients enrolled across 90 local clinical practice centres in Germany	Not reported	Schulz et al., 2018 (conference proceeding)[57]
CORECT	Czech Republic National Registry	To help describe the use of regorafenib for the treatment of mCRC in clinical practice in the Czech Republic, and to describe the clinical outcomes of patients in terms of safety and survival	n = 555; patients treated with regorafenib in clinical practice Prior therapies *(note: one patient could have had more than one prior therapy)* • bevacizumab: 89.2% • cetuximab: 23.6%	Regorafenib was administered according to the registration label to patients with mCRC previously treated with fluoropyrimidine, oxaliplatin, irinotecan, anti-VEGF therapy, and, in wild-type RAS tumours, also anti-EGFR therapy	Novakova-Jiresova et al., 2020[58]

Study	Design	Objective	Population (and prior treatments)	Prior treatment eligibility criteria	Primary study reference
			• panitumumab: 27.7% • aflibercept: 16.2% • trifluridine / tipiracil: 1.8%		
Tanaka 2018	Retrospective, observational, single-centre study conducted in Japan	To evaluate the efficacy and safety of regorafenib and trifluridine/tipiracil in patients with refractory mCRC	n = 44; 20 patients in regorafenib group Patients had 2–4 prior regimens: • fluoropyrimidine: 100% • oxaliplatin: 100% • irinotecan: 100% • anti-VEGF: 100% • anti-EGFR antibody (Wild KRAS or all-RASa): 45.0%	Patients previously treated with fluoropyrimidine, irinotecan, oxaliplatin, and anti-VEGF antibody (bevacizumab), or anti-epidermal growth factor receptor (EGFR) antibody (cetuximab or panitumumab) for patients who had KRAS exon 2 wild-type tumour	Tanaka et al., 2018[59]
Banzi 2019	Retrospective, observational, multi-centre study conducted in Italy	To evaluate adherence to treatment, effectiveness and safety	n = 144, adult patients with previously treated mCRC. The majority of patients were in third line and further	Not reported	Banzi et al., 2019[60]
Sueda 2016	Retrospective, observational, single-centre	To evaluate the efficacy and safety of regorafenib and	n = 37; 23 patients in the regorafenib group. Prior therapies:	Patients had histologically-confirmed, unresectable, mCRC and received two or	Sueda et al., 2016[61]

Study	Design	Objective	Population (and prior treatments)	Prior treatment eligibility criteria	Primary study reference
	study conducted in Japan	trifluridine/tipiracil in patients with mCRC refractory to standard chemotherapies	• anti-EGFR: 52.2% • anti-VEGF: 100%	more prior regimens of standard chemotherapy	
Huemer 2020	Retrospective, observational, multi-centre study conducted in Austria	To investigate hospitalizations during regorafenib or trifluridine/tipiracil therapy, as well as the impact of hospitalizations on clinical outcome in mCRC beyond second-line therapy	n = 93 mCRC patients; 69 patients in the regorafenib group received treatment at third or fourth line	Prior disease progression on fluorouracil, oxaliplatin, irinotecan, anti-VEGF and/or anti-EGFR (in case of RAS wild-type status) targeted therapy was a prerequisite for the initiation of regorafenib and/or trifluridine/tipiracil	Huemer et al., 2020[62]
Hirano 2015	Retrospective, observational, multi-centre study in Japan	To investigate the efficacy and safety of reduced-dose regorafenib for treatment of mCRC in Japan	n = 32 patients with pathohistologically-proven unresectable, recurrent or mCRC with 2 or more prior treatment regimens. Prior treatments: • bevacizumab: 78% • EGFR antibody: 72%	Not reported	Hirano et al., 2015[63]

Study	Design	Objective	Population (and prior treatments)	Prior treatment eligibility criteria	Primary study reference
Zengin 2018	Retrospective, observational, multi-centre study conducted in Turkey	To investigate the efficacy and safety of initiation of regorafenib at different doses in mCRC	n = 112 patients with mCRC. Median number of prior treatment lines was 3	The study population comprised of patients with mCRC who were previously treated with fluoropyrimidine, irinotecan and oxaliplatin +/- biologic agent	Zengin et al., 2018[64]

B.2.10.1. Real-world evidence

The real-world studies demonstrate efficacy outcomes consistent with those of CORRECT and CONCUR – but, notably, with numerically better median OS and PFS rates across the studies (Table 18). The safety data from these studies are also consistent with that of CORRECT and CONCUR (Section B.2.11).

Table 18: Summary of efficacy outcomes in the real-world studies

Efficacy	REBECCA (n = 654)	CORRELATE (n = 1,037[a])	RECORA (n = 463[b])	CORECT (n = 555)	Tanaka 2018 (n = 20)	Banzi 2019 (n = 144)	Sueda 2016 (n = 23)	Huemer 2020 (n = 69)	Hirano 2015 (n = 32)	Zengin 2018 (n = 112)
OS, median, months	5.6	7.7	5.8	9.3	9.1	5.5	5.8	10.4	7.66[c]	16.56[d]
PFS, median, months	2.7	2.9	3.1	3.5	2.1	-	3.0	-	2.66[c]	2.76[d]
PR, %	-	4	-	2.8	0	-	0	-	0	11.8[e]
SD, %	-	22	-	27.5	75.0	-	30.4	-	31	-
DCR, %	-	26	26.7	30.3	75.0	-	30.4	-	34	-

Key: DCR, disease control rate; ITT, intention to treat; OS, overall survival; PFS, progression-free survival; PR, partial response; SD, stable disease.
Note: Comparison of data across studies is for illustrative purposes only. Studies may differ in terms of population, design and endpoints, and direct comparisons should not be made; [a] response rate data based on 758 patients with at least one tumour assessment; [b] n = 463 is the ITT population; [c] values reported in days and re-calculated as months; [d] values reported in weeks and re-calculated as months; [e] clinical benefit rate was reported as 31.8%.
Sources: Adenis et al., 2016[56]; Ducreux et al., 2019[55]; Schulz et al., 2018[57]; Novakova-Jiresova et al., 2020[58]; Tanaka et al., 2018[59]; Banzi et al., 2019[60]; Sueda et al., 2016[61]; Huemer et al., 2020[62]; Hirano et al., 2015[63]; Zengin et al., 2018.[64]

B.2.11. Adverse reactions

B.2.11.1. CORRECT

B.2.11.1.1. Extent of exposure

Treatment compliance was high in both treatment groups: patients assigned regorafenib received [redacted] of the planned dose during the course of the study (mean daily dose 147.1 mg) compared with 90.1% for the placebo group (mean 159.2 mg).

B.2.11.1.2. Overview of adverse events

Table 19 presents an overview of the safety data up to the data cut-off date of 21 July 2011. Most patients in each group (regorafenib, 99.6%; placebo, 96.8%) experienced at least one treatment-emergent AE (TEAE), the majority of which were mild or moderate events. SAEs were reported at a similar rate in both groups (regorafenib, 43.8%; placebo, 39.5%), and, while low overall, the incidence of treatment-emergent SAEs that were considered drug-related was slightly higher with regorafenib (11.8% versus 3.6%).

Although significantly more patients in the regorafenib arm had dose modifications because of AEs (66.6% versus 22.5%), the difference in the incidence of AEs leading to permanent treatment discontinuation was relatively small (17.6% versus 12.6%), indicating that most AEs could be managed by dose modification and did not result in a permanent discontinuation of study drug.

Table 19: CORRECT – overview of TEAEs (SAS)

Event	Any AE, n (%)		Drug-related AE, n (%)	
	Regorafenib (n = 500)	**Placebo (n = 253)**	**Regorafenib (n = 500)**	**Placebo (n = 253)**
TEAE	498 (99.6)	245 (96.8)	465 (93.0)	154 (60.9)
CTC Grade 1	[redacted]	[redacted]	[redacted]	[redacted]
CTC Grade 2	[redacted]	[redacted]	[redacted]	[redacted]
CTC Grade 3	[redacted]	[redacted]	[redacted]	[redacted]
CTC Grade 4	[redacted]	[redacted]	[redacted]	[redacted]
CTC Grade 5	[redacted]	[redacted]	[redacted]	[redacted]
Treatment emergent SAE	[redacted]	[redacted]	[redacted]	[redacted]

Event	Any AE, n (%)		Drug-related AE, n (%)	
	Regorafenib (n = 500)	Placebo (n = 253)	Regorafenib (n = 500)	Placebo (n = 253)
CTC Grade 1	[redacted]	[redacted]	[redacted]	[redacted]
CTC Grade 2	[redacted]	[redacted]	[redacted]	[redacted]
CTC Grade 3	[redacted]	[redacted]	[redacted]	[redacted]
CTC Grade 4	[redacted]	[redacted]	[redacted]	[redacted]
CTC Grade 5	[redacted]	[redacted]	[redacted]	[redacted]
AE leading to permanent discontinuation	[redacted]	[redacted]	[redacted]	[redacted]
AE leading to dose modification	[redacted]	[redacted]	[redacted]	[redacted]

Key: AE, adverse event; CTC, Common Toxicity Criteria; SAE, serious adverse event; SAS, safety analysis set; TEAE, treatment-emergent adverse event.
Source: Table 10-2 of CORRECT Clinical Study Report (Final).[45]

B.2.11.1.3. Treatment-emergent adverse events

Table 20 presents TEAEs that occurred in ≥ 5% of patients in either treatment group. The most frequent AEs of any grade were fatigue and hand–foot skin reaction (17% each) in the regorafenib group, and fatigue (28%) and anorexia (15%) in the placebo group. Grade 3 or 4 TEAEs occurred at a higher rate in the regorafenib group than in the placebo group (54% versus 14%). The most frequent Grade ≥ 3 regorafenib-related AEs (affecting ≥ 5% of patients) were hand–foot skin reaction (17%), fatigue (< 10%), diarrhoea (< 8%), hypertension (7%), and rash or desquamation (6%) (Table 20).

Table 20: CORRECT – TEAEs occurring in ≥ 5% of patients in either group from start of treatment to 30 days after end of treatment (SAS)

	Regorafenib (n = 500)			Placebo (n = 253)		
	Any grade	Grade 3	Grade 4	Any grade	Grade 3	Grade 4
Any event , n (%)	465 (93)	253 (51)	17 (3)	154 (61)	31 (12)	4 (2)
Clinical adverse event, n (%)						
Fatigue	237 (47)	46 (9)	2 (< 1)	71 (28)	12 (5)	1 (< 1)
Hand–foot skin reaction	233 (47)	83 (17)	0 (0)	19 (8)	1 (< 1)	0 (0)
Diarrhoea	169 (34)	35 (7)	1 (< 1)	21 (8)	2 (1)	0 (0)
Anorexia	152 (30)	16 (3)	0 (0)	39 (15)	7 (3)	0 (0)

	Regorafenib (n = 500)			Placebo (n = 253)		
	Any grade	Grade 3	Grade 4	Any grade	Grade 3	Grade 4
Voice changes	147 (29)	1 (< 1)	0 (0)	14 (6)	0 (0)	0 (0)
Hypertension	139 (28)	36 (7)	0 (0)	15 (6)	2 (1)	0 (0)
Oral mucositis	136 (27)	15 (3)	0 (0)	9 (4)	0 (0)	0 (0)
Rash or desquamation	130 (26)	29 (6)	0 (0)	10 (4)	0 (0)	0 (0)
Nausea	72 (14)	2 (< 1)	0 (0)	28 (11)	0 (0)	0 (0)
Weight loss	69 (14)	0 (0)	0 (0)	6 (2)	0 (0)	0 (0)
Fever	52 (10)	4 (1)	0 (0)	7 (3)	0 (0)	0 (0)
Constipation	42 (8)	0 (0)	0 (0)	12 (5)	0 (0)	0 (0)
Dry skin	39 (8)	0 (0)	0 (0)	7 (3)	0 (0)	0 (0)
Alopecia	36 (7)	0 (0)	0 (0)	1 (< 1)	0 (0)	0 (0)
Taste alteration	35 (7)	0 (0)	0 (0)	5 (2)	0 (0)	0 (0)
Vomiting	38 (8)	3 (1)	0 (0)	13 (5)	0 (0)	0 (0)
Sensory neuropathy	34 (7)	2 (< 1)	0 (0)	9 (4)	0 (0)	0 (0)
Nose bleed	36 (7)	0 (0)	0 (0)	5 (2)	0 (0)	0 (0)
Dyspnoea	28 (6)	1 (< 1)	0 (0)	4 (2)	0 (0)	0 (0)
Muscle pain	28 (6)	2 (< 1)	0 (0)	7 (3)	1 (< 1)	0 (0)
Headache	26 (5)	3 (1)	0 (0)	8 (3)	0 (0)	0 (0)
Pain, abdomen	25 (5)	1 (< 1)	0 (0)	10 (4)	0 (0)	0 (0)
Laboratory abnormalities, n (%)						
Thrombocytopenia	63 (13)	13 (3)	1 (< 1)	5 (2)	1 (< 1)	0 (0)
Hyperbilirubinaemia	45 (9)	10 (2)	0 (0)	4 (2)	2 (1)	0 (0)
Proteinuria	35 (7)	7 (1)	0 (0)	4 (2)	1 (< 1)	0 (0)
Anaemia	33 (7)	12 (2)	2 (< 1)	6 (2)	0 (0)	0 (0)
Hypophosphataemia	25 (5)	19 (4)	0 (0)	1 (< 1)	1 (< 1)	0 (0)

Key: SAS, safety analysis set; TEAE, treatment-related adverse events.
Notes: Data cut-off date: 21 July 2011.
Source: Table 2. Grothey et al., 2013.[25]

B.2.11.1.4. Serious adverse events and deaths

There was a similar incidence of SAEs between groups (regorafenib, 43.8%; placebo, 39.5%; Appendix F; Table 35. The differences regarding Grade 4 and 5 treatment-emergent SAEs between the two treatment groups was small and clinically not relevant. Grade 3 treatment-emergent SAEs were reported at a higher incidence in the regorafenib group than in placebo (18.2% versus 13.8%), although the incidences of Grade 3 and 4 SAEs were similar between groups.

Overall, there were 110 deaths (regorafenib, 13.8% [n = 69]; placebo, 16.2% [n = 41]) reported during the study (i.e. up to within 30 days of last dose). The most common reason for death was PD (regorafenib, 11.6%; placebo, 13.8%); other reasons were reported as due to an AE not associated with clinical disease progression (1.6% versus 1.2%), unknown cause (0.4% versus 0.4%), and other cause (0.8% versus 0.2%). In the regorafenib group, the AEs not associated with disease progression that contributed to death were: pneumonia (n = 2), gastrointestinal bleeding (n = 2), intestinal obstruction (n = 1), pulmonary haemorrhage (n = 1), seizure (n = 1) and sudden death (n = 1). In the placebo group, these AEs were pneumonia (n = 2) and sudden death (n = 1). Occurrence of thromboembolism did not differ between groups (2% for both groups).

B.2.11.1.5. AEs of special interest

Liver dysfunction

The incidence of liver dysfunction was low in both groups (regorafenib, [redacted]%; placebo, [redacted]%). More patients in the regorafenib group than placebo had AEs of liver dysfunction that resulted in fatal outcomes: [redacted] patients versus [redacted] patients. Three events of liver dysfunction that had a fatal outcome in the regorafenib group were assessed by the treating Investigator as related to study drug; for two of these patients, the cause of death as stated by the Investigator was disease progression.

Cardiac ischaemia/infarction and bleeding events

The incidence of cardiac ischaemia/infarction was low in both groups (regorafenib, [redacted]%; placebo, [redacted]%). In most of the cases, patients had existing cardiovascular risk factors. In both groups, there was [redacted] death reported, which was not assessed as drug-related. Regarding bleeding events, the incidence was higher in the regorafenib group than in the placebo group (regorafenib, [redacted]%; placebo, [redacted]%); however, the majority of events were Grade 1 nose bleeds in the regorafenib arm. The incidence of serious bleeding was low in both groups (regorafenib, [redacted]%; placebo, [redacted]%). In total, [redacted] ([redacted]%) patients in the regorafenib group had a bleeding event resulting in death, while no deaths due to bleeding were reported in the placebo group.

Hand–foot skin reaction (palmar–plantar erythrodysesthesia)

The incidence of hand–foot skin reaction was higher in the regorafenib group than in the placebo group (regorafenib, [redacted]%; placebo, [redacted]%). This is expected, as hand–foot skin reaction is a known toxicity for patient treatment in this drug class. The majority of these events were Grade 1 or 2; however, [redacted]% of patients in the regorafenib group had Grade 3 events, compared with [redacted]% in the placebo group. The incidence of drug-related hand–foot skin reaction events was notably higher in the regorafenib group compared with the placebo group ([redacted]% versus [redacted]%). The incidence of SAEs was low ([redacted]% versus [redacted]%, respectively). Permanent discontinuation of treatment due to hand–foot skin reaction was low (regorafenib, [redacted]%; placebo, [redacted]%), with the majority of reactions being managed by dose reductions or interruptions.

Rash

The incidence of rash was higher in the regorafenib group than in the placebo group (regorafenib, [redacted]%; placebo, [redacted]%). This is expected, as rash is a known toxicity for patient treatment in this drug class. The majority of these events were Grade 1 in both treatment groups. The incidence of Grade 3 events was low (regorafenib, [redacted]%; placebo, [redacted]%). Rash treatment-emergent AEs could usually be managed by dose reductions or interruptions, and these TEAEs led to permanent discontinuation of treatment in only [redacted] ([redacted]%) regorafenib-treated patients and [redacted] placebo-treated patients.

Renal failure

The incidence of renal failure was low in both groups (regorafenib, [redacted]%; placebo, [redacted]%). Most events were Grade 3, with [redacted] Grade 4 AE (in the placebo group) and [redacted] Grade 5 AE (in the regorafenib group) reported. TEAEs of renal failure that were assessed as related to treatment were reported at the same incidence in regorafenib-treated patients and placebo-treated patients ([redacted]% in each group).

Proteinuria

The incidence of proteinuria was low in both groups (regorafenib, [redacted]%; placebo, [redacted]%). Most events were Grade 1 or 2, with Grade 3 events reported for [redacted]% and [redacted]%, respectively, in each group. No Grade 4 or Grade 5 events were reported.

B.2.11.1.6. AEs leading to discontinuations and dose modifications

Rates of permanent discontinuations due to TEAEs were low and similar in both groups (regorafenib, [redacted]%; placebo, [redacted]%). The incidence of Grade 3, Grade 4 and Grade 5 AEs leading to permanent discontinuation was similar between the two groups. Most toxicities could be managed by dose modifications.

Overall, more patients receiving regorafenib had AEs that led to dose medications than those on placebo (66.6% versus 22.5%). Of these, dose reductions occurred in 38% versus 3%, respectively, and dose interruption in 61% versus 22%, respectively. The most frequent AEs necessitating dose modification were dermatological, gastrointestinal, constitutional, and metabolic or laboratory events.

B.2.11.1.7. Additional safety data

An addendum to the CORRECT Final CSR was released on 29 August 2014. This addendum presented the cumulative safety results as of the final lock date of the clinical database (27 March 2014), including all data collected through the database cut-off date of 21 July 2011 (as reported in the CORRECT Final CSR), together with the additional safety data collected until all patients were off active treatment (last patient visit on 22 January 2014).

Overall, the results of the addendum were similar (with only slight or no changes) to those reported at the primary efficacy cut-off. In the safety update, there was a similar overall incidence in TEAEs as well as treatment-emergent SAEs for regorafenib versus placebo (Appendix F; Section B.5.1.2). Rates of permanent discontinuations and dose modifications remained low, and no new death events were reported.

B.2.11.1.8. Summary of AEs

Overall, no new safety signals were observed with regorafenib. There was a comparable incidence of TEAEs between groups, and the majority of patients in both groups had ≥ TEAE (regorafenib, 99.6%; placebo, 96.8%). The incidence of SAEs was also comparable between the two treatment groups (regorafenib, 43.8%; placebo, 39.5%). The incidence of TEAEs considered related to study treatment were higher with regorafenib than placebo, but this was as expected for its drug class. The rate of permanent discontinuations due to AEs was low, indicating that

most toxicities could be managed by dose modifications. Overall, there were 110 deaths (regorafenib, n = 69 [13.8%]; placebo, n = 41 [16.2%]) reported during the study. The majority of these cases were due to progression of underlying disease. The additional safety data did not reveal clinically significant changes from the initial data analysis, and no new safety concerns were identified.

In summary, the data suggest that regorafenib was well tolerated in CORRECT. The AEs associated with regorafenib were recognizable and manageable.

B.2.11.2. CONCUR

B.2.11.2.1. Extent of exposure

Treatment compliance was high in both treatment groups: patients assigned regorafenib received [redacted] of the planned dose during the course of the study (mean daily dose: 145.4 mg) compared with [redacted]% for the placebo group (mean daily dose: 160.0 mg).

B.2.11.2.2. Overview of adverse events

Table 21 presents an overview of the safety data up to the data cut-off date of 29 November 2013. Most patients in each group (regorafenib, 100%; placebo, 88.2%) experienced at least one TEAE, the majority of which were mild or moderate events. Of these events, 97.1% and 45.6% in each respective treatment group were considered to be drug-related. SAEs were reported at a similar rate in both groups (regorafenib, 31.6%; placebo, 26.5%), and, while low overall, the incidence of treatment-emergent SAEs that were considered drug-related was higher with regorafenib (8.8% versus 3.4%).

Although more patients in the regorafenib arm had dose modifications because of AEs (71.3% versus 16.2%), the incidence of AEs leading to permanent treatment discontinuation was relatively small (14.0% versus 5.9%), indicating that most AEs could be managed by dose modification and did not result in a permanent discontinuation of study drug.

Table 21: CONCUR – overview of TEAEs (SAS)

Event	Any AE, n (%)		Drug-related AE, n (%)	
	Regorafenib (n = 136)	Placebo (n = 68)	Regorafenib (n = 136)	Placebo (n = 68)
TEAE	136 (100.0)	60 (88.2)	132 (97.1)	31 (45.6)
CTC Grade 1	[redacted]	[redacted]	[redacted]	[redacted]
CTC Grade 2	[redacted]	[redacted]	[redacted]	[redacted]
CTC Grade 3	[redacted]	[redacted]	[redacted]	[redacted]
CTC Grade 4	[redacted]	[redacted]	[redacted]	[redacted]
CTC Grade 5	[redacted]	[redacted]	[redacted]	[redacted]
Treatment-emergent SAE	[redacted]	[redacted]	[redacted]	[redacted]
CTC Grade 1	[redacted]	[redacted]	[redacted]	[redacted]
CTC Grade 2	[redacted]	[redacted]	[redacted]	[redacted]
CTC Grade 3	[redacted]	[redacted]	[redacted]	[redacted]
CTC Grade 4	[redacted]	[redacted]	[redacted]	[redacted]
CTC Grade 5	[redacted]	[redacted]	[redacted]	[redacted]
AE leading to permanent discontinuation	[redacted]	[redacted]	[redacted]	[redacted]
AE leading to dose modification	[redacted]	[redacted]	[redacted]	[redacted]

Key: AE, adverse event; CTC, Common Toxicity Criteria; SAE, serious adverse event; SAS, safety analysis set; TEAE, treatment-emergent adverse event.
Note: For patients experiencing the same AE several times, the AE has been counted only once by the worst severity grade.
Source: Table 10-2. CONCUR Clinical Study Report (Final).[48]

B.2.11.2.3. Treatment-emergent adverse events

Table 22 presents TEAEs that occurred in ≥ 10% of patients in either treatment group. The most frequent AEs of Grade ≥ 3 associated with regorafenib were hand–foot skin reaction (22 [16%]), hypertension (15 [11%]), hyperbilirubinaemia, hypophosphataemia, and alanine aminotransferase concentration increases (nine [7%] each).

Table 22: CONCUR – TEAEs occurring at any grade in ≥ 10% of patients, or at Grade ≥ 3 in any patients in either group, from the start of treatment to 30 days after the end of treatment (SAS)

	Regorafenib (n = 136)				Placebo (n = 68)			
Adverse event, n (%)[a]	**Grade 1–2**	**Grade 3**	**Grade 4**	**Grade 5**	**Grade 1–2**	**Grade 3**	**Grade 4**	**Grade 5**
Any event	58 (43)	67 (49)	5 (4)	2 (1)	21 (31)	9 (13)	1 (1)	0 (0)
Hand–foot skin reaction	78 (57)	22 (16%)	N/A	N/A	3 (4)	0 (0)	N/A	N/A
Hyperbilirubinaemia	41 (30)	6 (4)	3 (2)	N/A	4 (6)	1 (1)	0 (0)	N/A
Alanine aminotransferase concentration increased	23 (17)	9 (7)	0 (0)	N/A	5 (7)	0 (0)	0 (0)	N/A
Aspartate aminotransferase concentration increased	24 (18)	7 (5)	1 (1)	N/A	6 (9)	0 (0)	0 (0)	N/A
Hypertension	16 (12)	15 (11)	0 (0)	0 (0)	1 (1)	2 (3)	0 (0)	0 (0)
Hoarseness	27 (20)	1 (1%)	N/A	N/A	0 (0)	0 (0)	N/A	N/A
Diarrhoea	23 (17)	1 (1)	0 (0)	0 (0)	1 (1)	1 (1)	0 (0)	0 (0)
Fatigue	19 (14)	4 (3)	N/A	N/A	4 (6)	1 (1)	N/A	N/A
Thrombocytopenia	9 (7)	3 (2)	1 (1)	N/	1 (1)	0 (0)	0 (0)	N/A
Hypophosphataemia	4 (3)	9 (7)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Proteinuria	11 (8)	2 (1)	N/A	N/A	0 (0)	1 (1)	N/A	N/A
Maculopapular rash	6 (4)	6 (4)	N/A	N/A	1 (1)	0 (0)	N/A	N/A
Leucopenia	8 (6)	3 (2)	0 (0)	N/A	0 (0)	0 (0)	0 (0)	N/A
Anorexia	9 (7)	1 (1)	0 (0)	0 (0)	3 (4)	0 (0)	0 (0)	0 (0)
Lipase concentration increased	3 (2)	6 (4)	0 (0)	N/A	3 (4)	1 (1)	0 (0)	N/A
Neutropenia	4 (3)	3 (2)	0 (0)	N/A	0 (0)	0 (0)	0 (0)	N/A

	Regorafenib (n = 136)				Placebo (n = 68)			
Adverse event, n (%)[a]	**Grade 1–2**	**Grade 3**	**Grade 4**	**Grade 5**	**Grade 1–2**	**Grade 3**	**Grade 4**	**Grade 5**
Myalgia	6 (4)	1 (1)	N/A	N/A	0 (0)	0 (0)	N/A	N/A
Abdominal pain	5 (4)	1 (1)	N/A	N/A	3 (4)	0 (0)	N/A	N/A
Anaemia	3 (2)	1 (1)	1 (1)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Other investigations[b]	3 (2)	1 (1)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Other skin and subcutaneous tissue disorders	3 (2)	1 (1)	0 (0)	0 (0)	1 (1)	0 (0)	0 (0)	0 (0)
Alkaline phosphatase concentration increased	3 (2)	0 (0)	0 (0)	N/A	0 (0)	1 (1)	0 (0)	N/A
Hypoalbuminaemia	2 (1)	1 (1)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Hypokalaemia	2 (1)	1 (1)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Visceral arterial ischaemia	0 (0)	1 (1)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
γ glutamyl transferase concentration increased	1 (1)	1 (1)	0 (0)	N/A	0 (0)	0 (0)	0 (0)	N/A
Pharyngitis	1 (1)	1 (1)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Atrial fibrillation	1 (1)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (1)	0 (0)
Cardiac arrest	N/A	N/A	0 (0)	1 (1)	N/A	N/A	0 (0)	0 (0)
Oesophageal varices haemorrhage	0 (0)	1 (1)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Death not otherwise specified	N/A	N/A	N/A	1 (1)	N/A	N/A	N/A	0 (0)
Serum amylase concentration increased	1 (1)	0 (0)	0 (0)	N/A	0 (0)	1 (1)	0 (0)	N/A

	Regorafenib (n = 136)				Placebo (n = 68)			
Adverse event, n (%)[a]	**Grade 1–2**	**Grade 3**	**Grade 4**	**Grade 5**	**Grade 1–2**	**Grade 3**	**Grade 4**	**Grade 5**
Wound infection	0 (0)	1 (1)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Flank pain	0 (0)	1 (1)	N/A	N/A	0 (0)	0 (0)	N/A	N/A
Vaginal fistula	0 (0)	1 (1)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Conduction disorder	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (1%)	0 (0)	0 (0)
Heart failure	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (1)	0 (0)
Acute kidney injury	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (1)	0 (0)
Other vascular disorders	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (1)	0 (0)	0 (0)

Key: AE, adverse event; N/A, not applicable; SAS, safety analysis set; TEAE, treatment-related adverse events.
Notes: Data in each column show the number of patients experiencing that grade as their worst severity of the relevant AE. [a] For patients with more than one AE, only the highest grade of the most severe event is shown; [b] Laboratory or diagnostic tests or clinical assessments.
Source: Table 2, Li et al., 2015.[26]

B.2.11.2.4. Serious adverse events and deaths

There was a similar incidence of SAEs between groups (regorafenib, 31.6%; placebo, 26.5%; Appendix F; Table 44). The differences regarding Grade 4 and 5 treatment-emergent SAEs between the two treatment groups was small and clinically not relevant. Grade 3 treatment-emergent SAEs were reported at the same incidence in both groups (11.8%). Most treatment-emergent SAEs were not related to study drug treatment.

Overall, there were 19 deaths reported during the study (regorafenib, n = 12 [8.8%]; placebo, n = 7 [8.8%]). One additional patient in the placebo group died during the follow-up period. Fourteen of these cases were due to progression of underlying disease (regorafenib, n = 8; placebo, n = 6). In the regorafenib group, the deaths of two (1%) patients were deemed to be drug-related within 30 days after the last dose.[26] Brief narratives of the two patients are as follows:

The first patient was a 65-year-old woman who stopped regorafenib treatment during her first cycle as a result of a non-serious Grade 2 increase in bilirubin. One week after stopping treatment, she collapsed at home and had a cardiac arrest

The second patient was a 67-year-old man who received regorafenib for 2 days. On the next day, he had a Grade 4 cardiac arrest, resulting in admission to hospital and death[26]

B.2.11.2.5. AEs of special interest

Acute liver failure

There were no reports of hepatic failure, hepatic necrosis or Grade 2–4 drug-related hepatobiliary disorder AEs in either treatment group. Overall, treatment-emergent hepatobiliary/pancreas SAEs (any grade) were reported with similar incidence in both groups (regorafenib, [redacted]%; placebo, [redacted]%). In the regorafenib group, there was one Grade 4 AE of bile duct stenosis that led to permanent study discontinuation. There were no deaths in either treatment group that resulted from liver dysfunction. No cases of significant transaminase increase or severe drug-induced liver injury have been identified from the ongoing hepatotoxicity monitoring from this study.

Acute cardiac failure and Grade ≥ 3 bleeding events

There was one report of heart failure in the placebo group and one report of cardiac arrest in the regorafenib group. There were no reports of cardiac ischaemia/infarction-related AEs. There was a low incidence of cardiac disorders reported in each group (regorafenib, [redacted]%; placebo, [redacted]%).

Regarding Grade ≥ 3 bleeding events, the overall incidence was higher in the regorafenib group compared with the placebo group ([redacted]% versus [redacted]%); however, the majority of the bleeding events in the regorafenib group were Grade 1 or 2 anaemia. Serious bleeding AEs were only reported in the regorafenib group ([redacted] events). No Grade 5 bleeding events were reported, and there were [redacted] haemorrhage/bleeding events that were the cause of permanent discontinuation of study medication in either treatment group.

Hand–foot skin reaction (palmar–plantar erythrodysesthesia)

The incidence of hand–foot skin reaction was higher in the regorafenib group than placebo (regorafenib, [redacted]%; placebo, [redacted]%). This is expected, as hand–foot skin reaction is a known toxicity for patient treatment in this drug class. The majority of these events were Grade 1 or 2, although Grade 3 events were reported in the regorafenib group ([redacted]% versus [redacted]%, respectively). There were [redacted] SAEs of hand–foot skin reaction in either treatment group. These events led to the permanent discontinuation of treatment in only one (0.7%) regorafenib-treated patient and no placebo-treated patients.

Rash

The incidence of rash was [redacted] ([redacted]; placebo, [redacted]). This is expected, as rash is a known toxicity for patient treatment in this drug class. The majority of events were Grade 1. No TEAEs related to rash led to permanent discontinuation in either treatment group. SAEs of rash were low (regorafenib, [redacted]%; placebo, [redacted]%), and all were reported as Grade 3 events and were considered related to the study drug.

Acute renal failure (any grade) or severe proteinuria (Grade ≥ 3)

The incidence of proteinuria was [redacted].

B.2.11.2.6. AEs leading to dose modifications

Rates of permanent discontinuations due to TEAEs were low (regorafenib, 14.0%; placebo, 5.9%). The incidence of Grade 3 and 4 AEs leading to permanent discontinuation were comparable between groups. There were no Grade 5 AEs leading to permanent discontinuation in either treatment group, and most toxicities could be managed by dose modifications.

Overall, more patients receiving regorafenib had AEs that led to dose modifications than those on placebo (75.0% versus 22.1%). Of these, dose reductions occurred in 39.7% and 0% of patients, and dose interruptions occurred in 62.5% and 16.2% of patients, respectively.

B.2.11.2.7. Additional safety data

An addendum to the CONCUR final CSR was released on 21 March 2017. This addendum presents the cumulative updated safety results, including all data collected as of 29 November 2013 together with the additional safety data collected until all patients were off active treatment (database cut-off date of 14 January 2016). This CSR addendum includes the safety results for all patients, including the six patients (all in the regorafenib group) who were still receiving treatment as of the 29 November 2013 database cut-off date.

At the safety data cut-off date all 204 patients in the study had terminated treatment; the six regorafenib patients who continued treatment past 29 November 2013 discontinued treatment by this point due to disease progression based on radiological evaluation. Since the primary analysis the additional available safety data did not reveal clinically significant changes from the initial data analysis; no new safety concerns were identified from the data updated since the initial analysis (Appendix F; Section B.5.2.2).

B.2.11.2.8. Summary of AEs

Overall, no new safety signals were observed with regorafenib. There was a comparable incidence of TEAEs between groups, and the majority of patients in both groups had one of more TEAEs (regorafenib, 100%; placebo, 88.2%). The incidence of SAEs was also comparable between the two treatment groups (regorafenib, 31.6%; placebo, 26.5%). The incidence of TEAEs considered related to study treatment was higher with regorafenib than with placebo, but this was as expected for its drug class. The rate of permanent discontinuations due to AEs was low, indicating that most toxicities could be managed by dose modifications. Overall, there were 19 deaths (regorafenib, n = 12 [8.8%]; placebo, n = 7 [8.8%]) reported during the study. One additional patient in the placebo group died during the follow-up period. The majority of these deaths were due to progression of underlying disease. The additional safety data from the updated data cut (14 January 2016) did not reveal clinically significant changes from the initial data analysis, and no new safety concerns were identified.

In summary, the data suggest that regorafenib was well tolerated. The AEs associated with regorafenib were recognizable and manageable. The safety profile overall was similar to that observed in earlier Phase I and II studies as well as the Phase III CORRECT study after which the study design for this study was modelled.

B.2.12. Ongoing studies

Both CORRECT and CONCUR are completed trials. No other RCTs of regorafenib in ≥ 3L mCRC are reading out in the next 12 months.

B.2.13. Interpretation of clinical effectiveness and safety evidence

B.2.13.1. Principal findings from the clinical evidence

Data from both the CORRECT and CONCUR trials demonstrated consistent clinically meaningful survival outcomes with regorafenib over placebo in ≥ 3L mCRC. Significantly prolonged OS (CORRECT, HR 0.77 [95% CI: 0.64, 0.94]; CONCUR, HR 0.55 [95% CI: 0.40, 0.77]; primary endpoint met for both trials) and PFS (CORRECT, HR 0.49 [95% CI: 0.42, 0.58]; CONCUR, HR 0.31 [95% CI: 0.22, 0.44]) were demonstrated for both trials. Subgroup analyses of OS and PFS were

consistent with these findings. While the ORR was similar between arms in both trials (CORRECT: p = 0.19; CONCUR, p = 0.045), regorafenib was associated with significant disease control compared with placebo (p < 0.0001 for both trials). Furthermore, regorafenib appears to improve QoL, with patients more likely to stay on treatment when compared with placebo (when assessed by EQ-5D-3L).[38] Indeed, this finding in QoL enhancement is in line with that reported in the literature for regorafenib, where it has been demonstrated that, compared with placebo, regorafenib delays the time until definitive deterioration in HRQL for multiple tumour types (including mCRC using data from CORRECT and CONCUR).[65]

The safety profile of regorafenib in CORRECT and CONCUR was both recognizable and manageable, and no new safety signals were reported. The overall incidence of TEAEs was comparable between the treatment arms in both trials (CORRECT, 99.6% versus 96.8%; CONCUR, 100% versus 88.2%). Dose modifications owing to AEs were higher with regorafenib compared with placebo (CORRECT, 67% versus 23%; CONCUR, 71% versus 16%); however, permanent discontinuations due to TEAEs were relatively small (CORRECT, 17.6% versus 12.6%; CONCUR, 14.0% versus 5.9%). Grade 3 TEAEs events were higher in the regorafenib arm than in the placebo arm (CORRECT, 56.0% versus 26.5%; CONCUR, 52.9% versus 32.4%). In both CORRECT and CONCUR, hand–foot skin reaction and hypertension were common regorafenib-related Grade ≥ 3 TEAEs. A total of 110 (regorafenib, n = 69 [13.8%]; placebo, n = 41 [16.2%]) and 18 (regorafenib, n = 12 [8.8%]; placebo, n = 6 [8.8%]) deaths were reported in CORRECT and CONCUR, respectively. The majority of these cases were due to progression of underlying disease.

Results from network meta-analyses of OS and PFS show similar efficacy between regorafenib and trifluridine/tipiracil. For OS, the hazard ratio (HR) was 0.99 (95% CrI: 0.84, 1.17) and for PFS the HR was 0.93 (95% CrI: 0.85, 1.03). Sensitivity analyses on the NMA confirmed efficacy was similar between regorafenib and trifluridine/tipiracil for OS and PFS. In addition, a MAIC comparison confirmed the results of the standard NMA (Appendix B.3.1.8.2).

B.2.13.2. Strengths and limitations of the evidence base

The CORRECT and CONCUR trials demonstrate the clinical benefit of regorafenib in a group of patients with limited options - patients with mCRC who have had two or

more prior therapies. While the reported survival outcomes were modest over placebo (BSC), they were nevertheless clinically meaningful findings in the ≥ 3L setting in mCRC, a population of patients where survival prognosis is very poor and there is strong clinical consensus of there being a strong clinical need for additional treatment options.

Both CORRECT and CONCUR were relatively large studies (n = 760 and n = 204, respectively) that were methodologically robust, enabling them to read-out high-quality clinical trial data. Furthermore, data from the two trials were consistent with each other, highlighting the efficacy of regorafenib across patients of different ethnicities. In addition, both trials reported HRQL data using validated instruments. This is a notable strength of the evidence base, as QoL assessments for active treatment in ≥ 3L mCRC are not commonly reported in the literature and were not reported for trifluridine/tipiracil in its NICE technology assessment. For regorafenib, evaluation of HRQL demonstrated that it was not associated with impaired HRQL over that of placebo (BSC) but rather that it likely has a positive effect on QoL, as well as on the length of survival.[38]

The evidence base is strengthened by the availability of several observational, real-world studies of regorafenib (Section B.2.10). These trials demonstrate efficacy and safety findings that are consistent with CORRECT and CONCUR.

A limitation of the evidence base is that there are no head-to-head data for regorafenib versus trifluridine/tipiracil, the key comparator in the scope of this appraisal. The CORRECT and CONCUR trials were completed before trifluridine/tipiracil regulatory approval and, therefore, did not include trifluridine/tipiracil as a comparator. To address this limitation, a meta-analysis and ITC were conducted, the findings of which demonstrate that both regorafenib and trifluridine/tipiracil have comparable efficacy (Section B.2.8 and B.2.9).

B.2.13.3. Applicability of clinical evidence to practice

As discussed in Section B.2.5, in general, the trial populations of CORRECT and CONCUR are reflective of patients with ≥ 3L mCRC in England and Wales. Furthermore, as ethnicity is not a recognized prognostic factor in mCRC, there is clinical consensus that of the two trials, the CONCUR cohort (specifically those with

no prior anti-VEGF therapy), despite being conducted solely in Asian patients, is more generalizable to the UK mCRC population. As UK patients will not be as highly pre-treated as the population of CORRECT (given that several targeted therapies, including bevacizumab, are not licensed in the UK but were used in CORRECT), it is expected that their survival outcomes will be slightly better and more closely aligned with those reported in CONCUR.

Regorafenib is an oral therapy that is not associated with any NHS clinical service changes or with any companion diagnostic requirements (Section B.2.3). This means that its approval would introduce a new chemotherapy-free therapy option in ≥ 3L mCRC without additional burden on the patient or NHS (oral medicines are associated with reduced exposure to healthcare facilities).

The primary treatment goal in mCRC is to prolong survival while maintaining QoL - incorporating a chemo-free alternative treatment option into the clinical management pathway will help achieve this goal.[66] In particular, treatments with different adverse event profiles provide valuable options where particular AEs have been troublesome in the past. This means that introducing regorafenib in ≥ 3L mCRC will help ensure patients gain the maximum benefit available to them from active treatment while their performance status is still good.

In conclusion, regorafenib offers a treatment with a different mode of action compared to the chemotherapy tipiracil/trifluridine which is the only alternative in the ≥ 3L mCRC setting. Regorafenib provides a chemotherapy-free treatment option that has the potential to make a substantial impact on patient health-related benefits, as it has a different AE profile compared to chemotherapy and provides meaningful survival outcomes [25, 26]

B.3. Cost effectiveness

B.3.1. Published cost-effectiveness studies

- In appendix G, describe and compare the methods and results of any published cost-effectiveness analyses available for the technology and/or the comparator technologies (relevant to the technology evaluation).
- See section 3.1 of the user guide for full details of the information required in appendix G.

An SLR was conducted to identify and evaluate the available economic evidence in ≥ 3L mCRC. Full details of the SLR methodology are presented in Appendix G. In total 23 cost-effectiveness and 3 budget impact studies, were identified (see Appendix G). All 23 identified cost-effectiveness studies for treatments used in ≥ 3L mCRC, along with a short overview of the approach, data input, and outcomes are shown in Table 23.

Of the identified studies, several studies investigated the cost-effectiveness of trifluridine/tipiracil in previously treated mCRC. These are also the most relevant studies for this submission as they are in line with the decision problem of this appraisal, in terms of comparator and indication. However, none of these studies provide reliable cost-effectiveness results to be used in this appraisal, for reasons detailed below.

Firstly, none of the studies used efficacy data from the 5 trials available for regorafenib and trifluridine/tipiracil in the cost-effectiveness analysis. More specifically, none of the identified studies used data from both CORRECT and CONCUR to inform regorafenib's efficacy, except for Giuliani (2021).[67] However, this study only performed a simple cost calculation, based on median OS and treatment duration, which is too simplistic to accurately assess cost-effectiveness. Similarly, none of the identified studies used the full set of available trifluridine/tipiracil data (i.e. RECOURSE, TERRA, and Yoshino (2012)). Consequently, none of the identified cost-effectiveness studies provide a reliable efficacy comparison using all available regorafenib and trifluridine/tipiracil effectiveness data.

Secondly, none of the identified studies use the appropriate cost data for this decision problem. This stems from the fact that all studies use the regorafenib list price without patient access scheme (PAS) discount for the analysis. Consequently, all identified studies will over-estimate the regorafenib costs, from an NHS perspective. In addition, most of the identified studies do not have a UK perspective, resulting in differences in treatment costs, healthcare resource use (HRU) costs, and AE costs used in these models, further limiting the accuracy of the cost results.

As none of the cost effectiveness studies provide reliable outcomes that are appropriate for the current decision problem, a de-novo model will be required to address the decision problem for this submission.

Table 23: Summary list of published cost-effectiveness studies

Study*	Year	Summary of model	Trials / data informing the model	QALYs (intervention, comparator)	Costs (intervention, comparator)	ICER (per QALY gained)
Giuliani 2021[67] (Italy)[67]	2021	Simple cost calculation using median OS and duration of treatment from the trial publications.	Regorafenib: • CORRECT, CONCUR, ReDOS T/T: • RECOURSE trial	NR	• Regorafenib: €5,818.68 • Trifluridine/ tipiracil: €2,101.50	NR (only presented the difference in costs per month-overall survival (OS) gained)
Guan 2020[68] (China)[68]	2020	Three-state Markov model • **Health states**: progression free survival (PFS) state, disease progression (PD) state and death state • **Time horizon**: Lifetime • **Cycle length**: 1 month	Regorafenib: • CONCUR study Fruquintinib: • FRESCO study	• Fruquintinib: 0.74 • Regorafenib: 0.79	• Fruquintinib cost: ¥151,058 ($22,888) • Regorafenib cost: ¥226,657 ($34,342)	¥1,529,196/QALY ($231,676/QALY)
SMC [Encorafenib] 2021[69] (Scotland)	2021	Cohort semi-Markov model (Partitioned survival cohort simulation model) • **Health states**: Pre-progression (starting health state), Post-progression, Death • **Time horizon**: 10 years • **Cycle length**: 1 month	Encorafenib plus cetuximab: • BEACON-CRC study (May 2020 data cut-off) Trifluridine/tipiracil: • RECOURSE	• Encorafenib plus cetuximab: 0.98 • Trifluridine/tipiracil: 0.26	• Encorafenib plus cetuximab: £67,482 • Trifluridine/Tipiracil: £14,782	Encorafenib plus cetuximab: - Trifluridine/Tipiracil: £72,914

Study*	Year	Summary of model	Trials / data informing the model	QALYs (intervention, comparator)	Costs (intervention, comparator)	ICER (per QALY gained)
Kashiwa 2020[70] (Japan)[70]	2020	Semi-Markov model (Partitioned survival model) • **Health states**: Progression-free survival, Post progression survival PPS, Death • **Time horizon**: 5 years • **Cycle length**: 28 days	Regorafenib: • CORRECT (CONCUR not used in ITC) T/T, BSC: • RECOURSE (Yoshino not used in ITC)	• T/T: 2.749 • Regorafenib: 2.654 • BSC: 2.405	• T/T: $82,049 • Regorafenib: $144,539 • BSC: $36,758	BSC: — Regorafenib: $432,734 T/T: $131,799
NICE TA668[8] (UK)	2020	Partitioned survival model: • **Health states**: mCRC progression-free, progressed, death • **Time horizon**: 10 years • **Cycle length**: 1 month	T/T: • RECOURSE Encorafenib, FOLFIRI: • BEACON	• FOLFIRI: 0.59 • T/T: 0.26 • Encorafenib plus cetuximab: 1.36	• FOLFIRI: £12,204 • T/T: NR • Encorafenib plus cetuximab: redacted	T/T vs FOLFIRI: Dominated Other ICERs are redacted
Zhang 2020[71] (China)	2020	Markov model • **Health states**: NR • **Time horizon**: Lifetime • **Cycle length**: NR	Regorafenib, BSC: • CONCUR	• Regorafenib: 0.62 • Placebo plus BSC: 0.44	• Regorafenib: ¥188,534 • Placebo plus best supportive care (BSC): ¥106,835	(Regorafenib vs. Placebo) ¥444,356/QALY
Almadiyeva 2019[72] (Kazakhstan)	2019	Markov model • **Health states**: NR • **Time horizon**: NR • **Cycle length**: NR	Regorafenib, BSC: • CORRECT	• Compared to BSC, Rego additional 0.025 QALY per patient	• Compared to BSC, Regorafenib additional costs: 5,245,000 KZT ($13,658)	Regorafenib compared to BSC: $552,571

Study*	Year	Summary of model	Trials / data informing the model	QALYs (intervention, comparator)	Costs (intervention, comparator)	ICER (per QALY gained)
Barzi 2019[73] (US)	2019	Markov model • **Health states**: NR • **Time horizon**: NR • **Cycle length**: NR	Regorafenib: • CORRECT • ReDOS BSC: • CORRECT,	• NR	• NR	Regorafenib compared to BSC: $384,687
Chu 2019[74] (US)	2019	State transition cohort or Markov model • **Health states** 1) disease on third-line therapy, 2) Alive with a durable response (remission) on third-line therapy, 3) Alive with disease on fourth-line therapy, 4) Alive on palliative care, 5) Dead • **Time horizon**: Lifetime • **Cycle length**: 1 week	T/T: • RECOURSE, Ipilimumab and nivolumab: • CheckMate 142	• T/T: 0.07 • Nivolumab: 6.76 • Ipilimumab and nivolumab: 9.25	• T/T: $90,700 • Nivolumab: $1,113,400 • Ipilimumab and nivolumab: $1,519,200	Vs T/T : • Nivolumab: $153,000 • Ipilimumab and nivolumab: $162,700
Gourzoulidis[75] 2019 (Greece)	2019	Partitioned survival model • **Health states** pre-progression, progressed disease and death • **Time horizon**: 10 years • **Cycle length**: NR	Regorafenib: • CORRECT T/T, BSC: • RECOURSE, • Yoshino (2012)	• T/T: 0.57 • Best supportive care (BSC): 0.40 • Regorafenib: 0.50	• T/T: €10,087 • Best supportive care (BSC): €1,879 • Regorafenib: €10,850	T/T Vs. BSC: €49,326 T/T Vs. Regorafenib: Dominant
Li 2019[76] (NR)	2019	Markov model • **Health states**: Progression-free survival, Progressive disease, Death	Regorafenib: • CONCUR study Fruquintinib: • FRESCO study	• Fruquintinib: 0.54 • Regorafenib: 0.53	• Fruquintinib: $25,550.15 • Regorafenib: $29,681.52	• Fruquintinib compared to Regorafenib: $-413,137.

Study*	Year	Summary of model	Trials / data informing the model	QALYs (intervention, comparator)	Costs (intervention, comparator)	ICER (per QALY gained)
		• **Time horizon**: NR • **Cycle length**: NR				
Sabater 2019[77] (UK)	2019	Update of a previous T/T model, with updated utility data based on PRECONNECT data	T/T, BSC: • PRECONNECT added to previous model based on RECOURSE, Yoshino, and CORRECT	Incr. T/T results: • 2.1 months vs BSC • 0.8 months vs Rego. (Total QALYs: NR)	NR	£51,792 (T/T vs BSC) T/T dominated regorafenib
Yao 2019[78] (China)	2019	Markov model (mathematical Markov model) • **Health states**: NR • **Time horizon**: NR • **Cycle length**: NR	Regorafenib: • CONCUR study Fruquintinib: • FRESCO study	• Fruquintinib: 0.274 • Regorafenib: 0.246	• Fruquintinib: $33,536 • Regorafenib: $35,607	NR
Bullement 2018[79] (UK)	2018	Partitioned survival model • **Health states**: mCRC progression free, post-progression, death • **Time horizon**: 10 years • **Cycle length**: NR	T/T, BSC: • RECOURSE • Yoshino (2012) Regorafenib: • CORRECT	• T/T: 0.57 • BSC: 0.40 • Rego: 0.51	• T/T: £17,978 • BSC: £9,499 • Rego: £24,112	£51,194 (T/T vs BSC) T/T dominated regorafenib
Cho 2018[80] (US)	2018	Markov model • **Health states** Stable disease state (progression free), Disease progression state, Death • **Time horizon**: 5 years • **Cycle length**: 30 days	Regorafenib: • CORRECT T/T, BSC: • RECOURSE	• Regorafenib: 0.397 • T/T: 0.437 • BSC: 0.339	• Regorafenib: $26,657 • T/T: $43,264 • BSC: $3,879	T/T versus regorafenib: $406,104 Regorafenib versus BSC: $395,223

Study*	Year	Summary of model	Trials / data informing the model	QALYs (intervention, comparator)	Costs (intervention, comparator)	ICER (per QALY gained)
						T/T versus BSC: $399,740
Liu 2018[81] (China)	2018	Markov model • **Health states**: NR • **Time horizon**: NR • **Cycle length**: NR	Regorafenib: • CONCUR	• Regorafenib: 0.68 • Cetuximab plus irinotecan: 0.65	• Regorafenib: ¥221,860 • Cetuximab plus irinotecan: ¥417,616	NR (ICER dominant)
Almeida 2017[82] (Portugal)	2017	Partitioned survival model • **Health states**: NR • **Time horizon**: Lifetime • **Cycle length**: NR	T/T, BSC: • RECOURSE, Yoshino (2012)	QALYs not reported	• T/T: €9,899/patient	ICER only reported for LYs
SMC [Trifluridine/ Tipiracil] 2017[83] (Scotland)	2017	Partitioned survival model • **Health states**: pre-progression (progression free survival), post-progression, death • **Time horizon**: 10 years • **Cycle length**: 1 day	T/T, BSC: • RECOURSE, Yoshino (2012)	• Incremental QALYs (T/T compared to BSC): 0.17	• Incremental costs (T/T compared to BSC): £8,197	T/T as compared to BSC (per QALY): £49,225
Kimura 2016[84] (Japan)	2016	NR	Regorafenib: • CORRECT T/T, BSC: • RECOURSE	NR	• Regorafenib: ¥705,330 • T/T: ¥371,199	Regorafenib: ¥110,208 T/T: ¥52,282
NICE TA405[9] (UK)	2015	Partitioned survival model • **Health states**: mCRC progression-free, post-progression, death	T/T, BSC: • RECOURSE Yoshino (2012)	• T/T: 0.59 • BSC: 0.42	• T/T: £17,859 • BSC: £10,286	£44,032 (T/T vs BSC)

Study*	Year	Summary of model	Trials / data informing the model	QALYs (intervention, comparator)	Costs (intervention, comparator)	ICER (per QALY gained)
		• **Time horizon**: 10 years • **Cycle length**: 1 day				
Deger 2015[85] (Turkey)	2015	Cohort partition model • **Health states**: NR • **Time horizon**: Lifetime • **Cycle length**: NR	Regorafenib: • CORRECT	QALYs not reported	• Regorafenib: €4,234 • Standard treatment: €2,394	Patients who have been treated 3 or more treatment regimens previously: €8,308
Goldstein[86] 2015 (US)	2015	Markov model • **Health states**: 1) third line regorafenib treatment, 2) Best supportive care, 3) Death • **Time horizon**: NR • **Cycle length**: 4 weeks	Regorafenib: • CORRECT	• QALYs gained : 0.04	• 120mg: $32,141 • 147mg: $39,391 • 160mg: $42,838	• 120mg: $732,242 • 147mg: $897,411 • 160mg: $975,954
Seal 2013[87] (US)	2013	Cohort partition model • **Health states**: Alive, dead • **Time horizon**: NR • **Cycle length**: NR	Regorafenib: • CORRECT	• Regorafenib: 0.47 • Best supportive care: 0.37	• Regorafenib: $60,188 • Best supportive care: $28,972	NR

*Only cost-effectiveness studies were included, full overview of SLR results with budget impact studies included (N=3) available in Appendix G.
Key: BSC, best supportive care; cet., cetuximab; enco, encorafenib, ICER, incremental cost-effectiveness ratio; mCRC: metastatic colorectal cancer; NR, not reported; QALYs, quality-adjusted life years; T/T, trifluridine/tipiracil.
Currencies: €, euro; £, pound sterling; $, US dollar; ¥, Yen; KZT, Kazakhstani Tenge

B.3.2. Economic analysis

As discussed in Section B.3.1 none of the identified cost-effectiveness studies appropriately address the current decision problem as 1) none use data from each of the 5 RCTs identified in the SLR, and 2) none uses the regorafenib PAS price. As a result, a de novo cost-effectiveness analysis tailored to the current decision problem was performed.

B.3.2.1. Patient population

In line with the restricted submission and the CORRECT and CONCUR trials[45, 48], the model features adult patients with mCRC who have been previously treated with, or are not considered candidates for, available therapies (i.e. fluoropyrimidine-based chemotherapy, anti-VEGF therapy and anti-EGFR therapy).

B.3.2.2. Model structure

A de novo partitioned survival (area under the curve) model was constructed in Microsoft Excel®. A partitioned survival model was chosen because it allows for a straightforward unbiased estimation of costs, life years (LYs) and quality-adjusted life years (QALYs). Unlike a Markov model, which also uses health states, partitioned survival models do not require the estimation and use of transition probabilities. Instead, the number of patients in each health state is calculated directly from each (extrapolated) OS and PFS curves.

The model consists of the following mutually exclusive health states (see Figure 14):

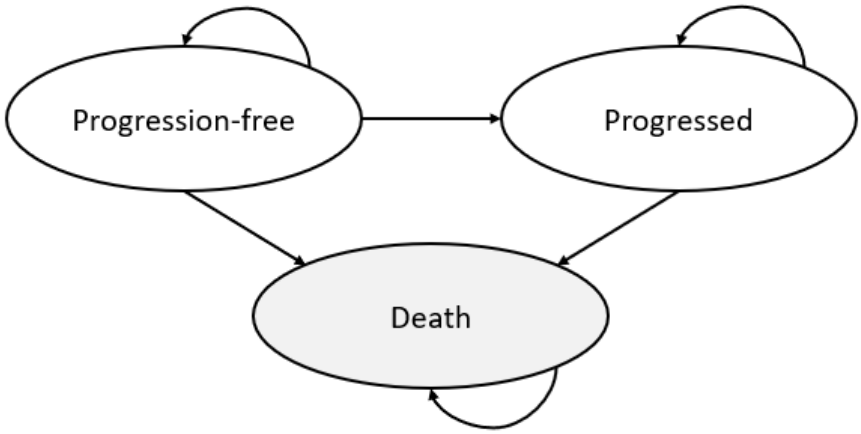
Progression-free: a patient's disease is stable or responding, and not actively progressing. Costs in this health state are associated with treatment, administration and management of disease and AEs. QoL is higher compared with patients with progressed disease

Progressed: a patient's disease is assumed to have progressed. Costs in this health state are only associated with management of disease as the model assumes no active treatment is given after progression. Patients have a lower QoL than in the progression-free state.

Death: an absorbing health state

Model health states were selected in accordance with the clinical pathway of care and are typical of modelling in advanced oncology. This structure has also been used in previous later-line mCRC NICE technology appraisals (TA242[88], TA307[89], TA405[9], TA668[8]). This structure is identical for all patients, regardless of their treatment, as it is based on disease progression rather than treatment. The possible routes patients may flow through the model are presented in Figure 14, below.

Figure 14: Model structure



The likelihood of patients transitioning between the health states is determined via PFS and OS curves that were fitted to the clinical trial data. All patients enter the model in the progression-free health state and remain in this state until disease progression or death. The progression-free health state is designed to capture the benefits from an active treatment whilst the disease is controlled prior to progression, leading to relatively higher QoL. Following progression, patients are unable to transition back from the progressed health state to the progression-free health state. The progressed disease state is designed to capture the relatively poor QoL following disease progression and prior to death. Hence, the model captures the changes in QoL between patients who are progression-free and progressed. Patients can transition to the death health state from any other health state.

The model evaluates the cost-effectiveness of regorafenib in mCRC from a UK NHS and Personal Social Services (PSS) perspective. The model uses a weekly cycle length. This is in contrast to some previous mCRC models that used cycle lengths ranging from 1 day (TA405[9]) to 1 month (TA242[88], TA668[8]). A weekly cycle length

was chosen to provide a good balance between more granularity but without over-complicating the model. The time horizon of the model is 10 years, which corresponds to lifetime as < 1% of patients are alive after 10 years. Costs and QALYs are discounted with an annual rate of 3.5%, in line with the NICE reference case.[90] Half-cycle correction was also applied, to correct for mid-cycle progressions.

Table 24 below summarizes the different elements of the model structure and compares these to past appraisals in later-line mCRC. Not all previous appraisals are equally representative, as they include appraisals in ≥ 2L, ≥ 3L, and mutation-specific later-line mCRC. The only past appraisal in the same ≥ 3L mCRC population as regorafenib is TA405, which assesses trifluridine/tipiracil in ≥ 3L mCRC. TA405 is therefore the most relevant source of comparison and validation for this appraisal.

Table 24: Features of the economic analysis

Factor	Previous appraisals				Current appraisal	
	TA242[88]	TA307[89]	TA405[9]	TA668[8]	Chosen values	Justification
Time horizon	10 years	15 years	10 years	10 years	10 years	A 10-year time horizon corresponds to a lifetime horizon (< 1% alive after 10 years), in line with the NICE reference case
Cycle length	1 month	2 weeks	1 day	1 month	1 week	Cycle length should be short enough to represent the frequency of key clinical events and interventions, but long enough to maintain computational efficiency. One week is sufficiently short to capture key events.
Model approach	PartSA	PartSA	PartSA	PartSA	PartSA	In line with past TAs and allows for unbiased estimation of outcomes
Treatment waning effect?	Not described	Not described	Not described	None	None	Unbiased simplification which applies equally to trifluridine/tipiracil and regorafenib
Source of utilities	CO.17 trial[88]	mCRC utilities study[89]	CORRECT trial[9]	BEACON trial[8]	Pooled CORRECT and CONCUR EQ-5D data	For internal consistency, the model uses pooled data for all inputs
Source of costs and resource use	BNF/NHS reference costs	Clinical study, BNF, PSSRU, NHS reference costs	SLR, BNF, PSSRU, NHS reference costs	SLR, BNF, PSSRU, NHS reference costs, eMIT, clinical input	SLR, BNF, PSSRU, NHS reference costs, clinical input	As per NICE reference case

Key: BNF, British National Formulary; eMIT, electronic market information tool; mCRC; metastatic colorectal cancer; NHS, National Health Service; NICE, National Institute for Health and Care Excellence; PartSA, partitioned survival analysis; PSSRU, Personal Social Services Research Unit; TA, technology appraisal.

B.3.2.3. Intervention technology and comparators

The economic model allows the costs and efficacy of regorafenib to be compared with those of trifluridine/tipiracil and BSC.

A full set of economic analyses is presented against trifluridine/tipiracil as this is the main comparator and in keeping with the requested position i.e. we are seeking a similar position to trifluridine/tipiracil in the ≥3rd line setting. As stated in section B.1.1, Bayer is making this submission on the request of physicians who are seeking another treatment option alongside trifluridine/tipiracil.

A limited set of analyses is presented against BSC, mainly because BSC was the comparator in the pivotal trials for regorafenib, rather than it being considered directly relevant to the requested position. Treatment practices have changed since the completion of the pivotal trials and trifluridine/tipiracil is now used at the third-line or later setting. The model includes the same functionality for comparing against trifluridine/tipiracil or BSC.

Regorafenib is implemented in the model for patients with mCRC at a recommended dose of 160 mg (4 x 40mg tablets) once daily for 3 weeks followed by one week off therapy. This is reflective of the decision problem described in Section B.1.1. Regorafenib efficacy is informed by the two registration Phase III trials: the global CORRECT trial[25] and the CONCUR study.[26] Further details of these trials are discussed in Section B.3.3.

All consulted clinical experts agreed that trifluridine/tipiracil would be the key comparator for regorafenib in mCRC. Trifluridine/tipiracil is an orally administered combination of trifluridine, a thymidine-based nucleic acid analogue, and tipiracil hydrochloride, a thymidine phosphorylase inhibitor.[91] Trifluridine/tipiracil is administered at a dose of 35 mg/m^2 twice daily, 5 days a week, with 2 days of rest, for 2 weeks, followed by a 14-day rest period. This treatment cycle is repeated every 4 weeks.[91] Trifluridine/tipiracil efficacy is informed by the registration Phase III trials RECOURSE[51] and TERRA[52], and the Japanese Phase II trial Yoshino (2012).[53] In the model, trifluridine/tipiracil is implemented as a comparator via an indirect comparison (see Section B.3.3).

As per the trial protocols and licence in mCRC, treatment with regorafenib and trifluridine/tipiracil is continued until the determination of RECIST-defined disease progression, clinical progression, the development of severe AEs, withdrawal from the study, death, or a decision by the treating physician that discontinuation would be in the patient's best interest.[91, 92] To accurately model treatment duration as observed in the CORRECT and CONCUR trials, pooled time on treatment (ToT) data are used to calculate the total active drug costs. Further details on regorafenib and trifluridine/tipiracil drug cost calculation are discussed in Section B.3.5.1.

The model also compares regorafenib against BSC alone (reduced analysis set): the comparator in both the CORRECT and CONCUR studies. BSC in CORRECT and CONCUR included any concomitant medications or treatments: antibiotics, analgesics, radiation therapy for pain control (limited to bone metastases), corticosteroids, transfusions, psychotherapy, growth factors, palliative surgery, and any other symptomatic therapy necessary to provide BSC. Other investigational anti-tumour agents or antineoplastic chemotherapies / hormonal therapies / immunotherapies were not included in BSC.

B.3.3. Clinical parameters and variables

Efficacy data for regorafenib in mCRC are available from two Phase III trials:

1) The CORRECT study (n = 760)
2) The CONCUR study (n = 204)

B.3.3.1. Data pooling

To maximize the data informing the model and to accurately represent the available evidence, data from CORRECT and CONCUR were pooled. However, neither trial is 100% generalisable to the UK setting – the two main differences are described below.

1) Prior anti-VEGF use

All patients in CORRECT received prior therapy with an anti-VEGF agent (i.e. bevacizumab), compared with 41.2% of patients receiving regorafenib and 36.8% of

patients receiving placebo in CONCUR. As anti-VEGF is not recommended by NICE, the CONCUR trial is more reflective of UK clinical practice.

Prior exposure to anti-VEGF therapy may reduce the treatment effect associated with regorafenib over BSC, as regorafenib is a multi-kinase inhibitor with targets that include VEGF.[92] This possibility is supported by the higher point estimate HRs for patients who received prior targeted therapy in CONCUR compared with patients who were naïve to targeted treatment. Similarly, it was supported by nine UK clinical experts who attended a recent advisory board meeting. As a result of this effect, and the absence of anti-VEGF therapy in UK clinical practice, it is possible that the CORRECT study would underestimate the efficacy of regorafenib in the UK population compared with CONCUR.

2) Ethnicity

All patients in CONCUR were Asian which makes CONCUR less representative of the ethnic mix who would receive treatment in the UK. However, none of the consulted clinical experts considered ethnicity to be a treatment effect modifier. This conclusion is supported by the CORRECT study, which reported comparable OS HRs in Asian (0.79) and non-Asian (0.76) groups. CONCUR therefore likely provides efficacy estimates generalizable to the UK population.

Appropriateness of using pooled data

We consider that pooling both trials represents the most robust evidence base for decision-making. Using pooled CORRECT and CONCUR data is also in line with approach in the appraisal of trifluridine/tipiracil (TA405).[9] In TA405 the Committee preferred an economic model that pooled the global RECOURSE study and Yoshino study, as opposed to alternatives based on the single studies themselves despite similar differences in ethnicity and prior treatment as present between CONCUR and CORRECT as described above.

Pooled CORRECT and CONCUR data were therefore used to inform the efficacy in the base case. Where pooled data were not readily available (e.g. relative dose intensity [RDI] data or AE probabilities), a weighted average of data from CORRECT

and CONCUR was used to ensure the model consistently uses pooled data for all inputs.

In summation, the following data were used in the base case:

- PFS was modelled using pooled data from CORRECT and CONCUR
- OS was modelled using pooled data from CORRECT and CONCUR
- ToT (time on treatment) was modelled using pooled data from CORRECT and CONCUR (for drug costing)
- RDI was modelled using a weighted average of data from CORRECT and CONCUR (for drug costing)
- Utility data was modelled using pooled EQ-5D data from CORRECT and CONCUR
- AE probabilities were modelled using a weighted average of data from CORRECT and CONCUR

B.3.3.2. Data extrapolation

Clinical data for time-to-event outcomes were extrapolated to facilitate modelling of survival over a lifetime horizon (i.e. 10 years). Treatment effects were modelled by extrapolating patient-level data from the primary cut-off for each arm separately (in line with NICE Decision Support Unit Technical Support Document 14[93]). The proportional hazards assumption holds for all endpoints, as shown in Appendix O.

For each of the above-mentioned outcomes (PFS, OS, ToT), seven standard parametric models (i.e. exponential, Weibull, log-logistic, log-normal, generalized gamma, gamma and Gompertz) were fitted for each treatment group. To determine the best model fit, the following steps were undertaken:

- **Akaike information criterion (AIC) and Bayesian information criterion (BIC) –** model fits were evaluated using AIC and BIC statistics. Lower AIC and BIC figures are indicative of a better statistical fit of the survival curve
- **Visual inspection** – visual inspection was carried out by plotting the projected survival curves overlaid with the Kaplan–Meier survival functions.
- **Clinical validity** – the clinical plausibility of the extrapolated outcomes was assessed using expert opinion

Inclusion and extrapolation of each clinical outcome is detailed in the following sections.

B.3.3.2.1. Progression-free survival

As discussed in Section B.2.6, PFS in both CONCUR and CORRECT was very mature: 85.1% and 94.5% of regorafenib and BSC patients in CORRECT had experienced a PFS event at the time of primary efficacy cut-off, as had 88.2% and 95.6% of regorafenib and BSC patients, respectively, in CONCUR.[45, 48] Similarly, PFS Kaplan–Meier data was mature, with pooled PFS data available until 2.33% of regorafenib patients and 0.69% of BSC patients are progression-free. The KM data are used directly in the model after which an exponential extrapolation informs the remainder of the model (as the exponential survival estimates are closest to the Kaplan–Meier data at 9.2 and 16.1 months).

The fully parametric survival curves for PFS are shown in appendix N. A scenario using a fully parametric PFS approach is explored in sensitivity analysis.

B.3.3.2.2. Overall survival

Pooled OS data were extrapolated by applying seven standard parametric models (i.e. exponential, Weibull, log-logistic, log-normal, generalized gamma, gamma and Gompertz) to the pooled Kaplan–Meier data. Log-logistic was the model with the best statistical fit for regorafenib, whereas the log-normal curve showed the best statistical fit for BSC (see Table 25, below). Figure 15 to Figure 18 compare the pooled OS Kaplan–Meier data with different extrapolated curves for regorafenib and BSC, respectively. For BSC, the log-normal, log-logistic, and generalized gamma curves each performed well in terms of both visual inspection and statistical fit (see Table 25 and Figure 17). For regorafenib, visual inspection suggests the log-normal distribution fits the Kaplan–Meier data best, particularly in the tail end of the curve, followed by the log-logistic and generalized gamma (see Figure 15).

Table 25: Goodness-of-fit statistics of the pooled overall survival extrapolations

Fitted function	Regorafenib		Statistical rank	BSC		Statistical rank
	AIC	BIC		BIC	AIC	
Log-logistic	2,419.7	2,428.6	1	1,274.6	1,282.1	3
Generalized gamma	2,424.4	2,437.8	2	1,270.6	1,281.9	2
Log-normal	2,428.0	2,437.0	3	1,268.9	1,276.5	1
Gamma	2,428.7	2,437.7	4	1,282.7	1,290.3	4
Weibull	2,437.4	2,446.3	5	1,292.0	1,299.5	5
Gompertz	2,470.0	2,478.9	6	1,316.9	1,324.4	6
Exponential	2,489.2	2,493.7	7	1,328.0	1,331.7	7

Key: AIC, Akaike information criterion; BIC, Bayesian information criterion; BSC, best supportive care.

Figure 15: Short-term parametric fits of regorafenib overall survival extrapolations

Key: KM, Kaplan–Meier; OS, overall survival.

Figure 16: Long-term parametric fits of regorafenib overall survival extrapolations

Key: KM, Kaplan–Meier; OS, overall survival.

Figure 17: Short-term parametric fits of BSC overall survival extrapolations

Key: BSC, best supportive care; KM, Kaplan–Meier; OS, overall survival.

Figure 18: Long-term parametric fits of BSC overall survival extrapolations

Key: BSC, best supportive care; KM, Kaplan–Meier; OS, overall survival.

The validity of the long-term OS estimates was discussed during a clinical advisory board. Considering the small variation in OS estimates, especially for the best fitting OS curves (log-logistic, log-normal, and generalized gamma), the experts felt a case could be made for all curves. The experts considered that consistency with TA405 was important, as trifluridine/tipiracil was the key comparator. They therefore recommended log-logistic to be used in the base-case, in line with the committee and ERG preference in TA405. This approach was confirmed by the consulted health economic expert, who also agreed that log-logistic shows the best statistical fit for regorafenib.

The trifluridine/tipiracil OS results from our model were compared with the trifluridine/tipiracil results reported in TA405.[9] Median OS results were similar across OS distributions and in line with TA405 (see Table 26). For mean OS, the different extrapolations resulted in bigger differences in predicted values, with the log-logistic and log-normal extrapolations predicting mean OS most in line with TA405 ([redacted] versus [redacted] months; Table 26). These data further confirm that the log-logistic and log-normal extrapolations provide the most appropriate OS input for the model. The base-case therefore applies a log-logistic OS extrapolation, with other curves explored in scenarios.

Table 26: Modelled trifluridine/tipiracil outcomes compared to reported outcomes in TA405

	Current model outcomes (months)							TA405 reported outcomes
	Weibull	Log-normal	Log-logistic	Exp.	Gen. gamma	Gompertz	Gamma	
Median OS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	7.4
Mean OS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	11.1

Key: BSC, best supportive care; exp., exponential; gen., generalized; OS, overall survival.
Source: NICE TA405.[9]

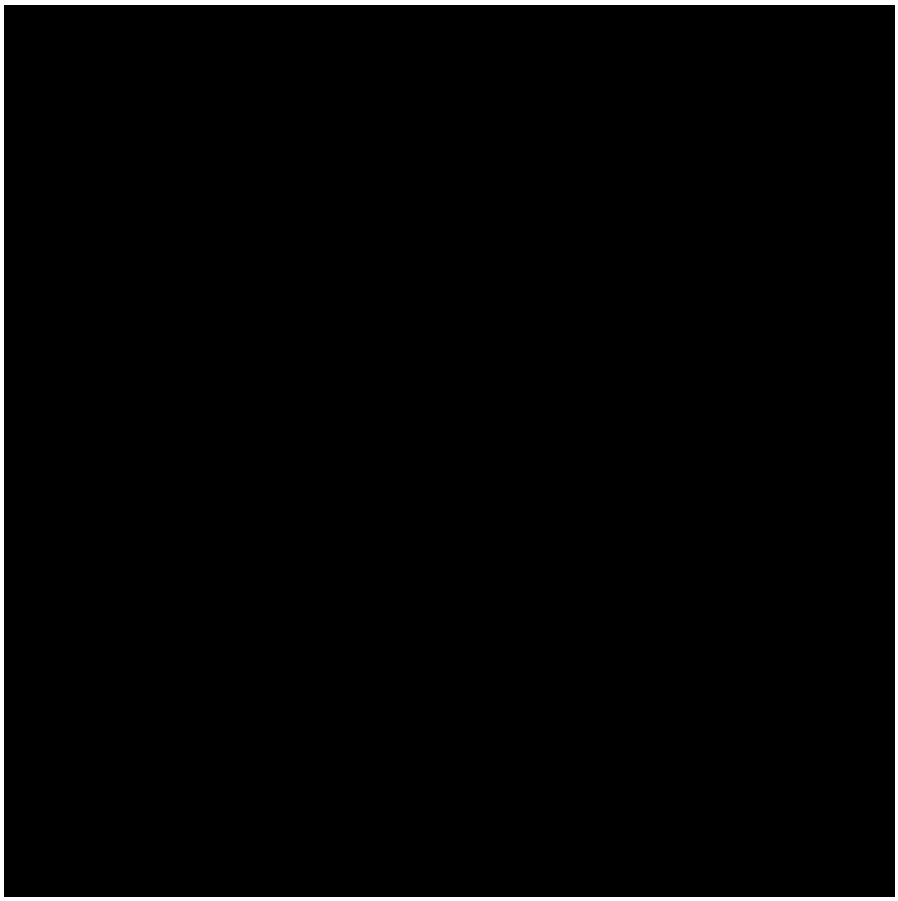
B.3.3.2.3. Time on treatment

As per the licensed indication, patients treated with regorafenib are expected to be treated until disease progression or unacceptable toxicity. Treatment duration can therefore differ slightly from PFS due to early discontinuations caused by AEs and

other reasons for discontinuations before progression (e.g. patients declining therapy). Therefore, a post hoc analysis of ToT (time on treatment) was performed, to be able to model treatment use more accurately. Here, ToT was defined as 'Date of treatment end' – 'Date of treatment start', as recorded in CORRECT and CONCUR.

Similar to PFS and OS, the model uses pooled ToT data from CORRECT and CONCUR to determine the duration of treatment in the base case. Due to the maturity of ToT, Kaplan–Meier data were used directly to inform ToT in the base case, with a parametric approach to ToT explored in the scenario analyses. The model uses a log-logistic extrapolation to inform ToT for the remainder of the model, when Kaplan–Meier data were no longer available, as the log-logistic estimates are closest to the end of the Kaplan–Meier data. By using this approach, the model assumes that ToT on average is similar but slightly shorter than PFS, as would be expected in clinical practice (see Figure 19).

Figure 19: Regorafenib and BSC survival curves used in the model base case



Key: BSC, best supportive care; KM, Kaplan-Meier; OS, overall survival; PFS, progression-free survival; ToT, time on treatment.

B.3.3.3. Comparators not included in CORRECT/CONCUR:

As discussed in Section B.2.9, there is no direct evidence comparing the efficacy of regorafenib directly to that of trifluridine/tipiracil. To be able to compare against trifluridine/tipiracil, an ITC using the primary publications for both treatments was conducted for OS and PFS (see Section B.2.9). The resulting ITC HRs versus regorafenib were applied to the OS and PFS curves used for regorafenib, to model trifluridine/tipiracil's efficacy.

For OS and PFS, the full NMA using all data for regorafenib (CORRECT and CONCUR) and trifluridine/tipiracil (RECOURSE, Yoshino (2012), TERRA) showed both treatments to be comparable, with a small non-significant numerical benefit in favour of regorafenib, with OS and PFS HRs (CrIs) of 0.99 (0.84, 1.17) and 0.93 (0.85, 1.03), respectively. These point estimates for OS and PFS are used to model trifluridine/tipiracil OS, PFS, and ToT in the base case. The resulting trifluridine/tipiracil survival curves used in the model base case are shown in Figure 20.

For ToT, no ITC could be performed, as insufficient ToT data for trifluridine/tipiracil were publicly available. However, considering the similarity between ToT and PFS in clinical practice, since patients are treated until progression, it is reasonable to assume the ToT and PFS HR will be similar as well. This is also illustrated by the PFS and ToT data used in the model, which shows a high degree of similarity (Figure 19). The model therefore assumes the estimated indirect PFS HR of 0.93 also applies to ToT. Similar to OS and PFS, this PFS HR versus regorafenib was applied to the ToT curves used for regorafenib, to model trifluridine/tipiracil's ToT. The resulting trifluridine/tipiracil survival curves are shown in the Figure 20 overleaf.

Figure 20: Trifluridine/tipiracil survival curves used in the model base case

Key: OS, overall survival; PFS, progression-free survival; ToT, time on treatment.

However, as discussed in Section B.2.9, there were some limitations to the ITC, most notably around population differences between the included trials. To explore these uncertainties, the NMA includes several sensitivity analyses, which focussed on a different subsets of the studies. Specifically, this refers to the NMAs of CORRECT vs RECOURSE, CONCUR vs TERRA, CONCUR vs TERRA and Yoshino (2012), and the NMA that excluded Yoshino (2012). The resulting HRs are explored in the model as scenarios. In addition, considering that neither the NMA nor the sensitivity analyses showed any significant difference in OS or PFS, and that the consulted clinical and health economic experts agreed it was plausible to assume regorafenib and trifluridine/tipiracil are equivalent, we also explore a scenario that applies a HR of 1 to OS and PFS.

B.3.4. Measurement and valuation of health effects

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

In the CORRECT and CONCUR trials, patients' QoL and health utility values were measured using the EORTC QLQ-C30 and EQ-5D-3L questionnaires, respectively. As discussed in Section B.2.6, only minor differences in EQ-5D-3L were observed between regorafenib and BSC in both trials, both for patients receiving treatment and at the end of treatment. Although there were some data indicating that patients on regorafenib had better quality of life, most notably from the higher response rates in

CORRECT and data suggesting a superior QoL for patients on regorafenib,[38] no significant difference in QoL was observed in either study.[45, 48] The model therefore assumes that there is no utility benefit attributable to the treatment itself. Utilities are health-state specific with disutilities applied in the presence of adverse events. A average of all EQ-5D-3L index scores from CORRECT and CONCUR was calculated, for all measurements while patients were on treatment, and for patients at the end of treatment. This pooled on-treatment utility value is assumed to represent the mean utility for the pre-progression health state, whereas the end of treatment utility is used for post-progression patients (Table 27).

None of the trifluridine/tipiracil studies included in the model captured EQ-5D. The only reported QoL data for trifluridine/tipiracil was captured in PRECONNECT, an open-label single arm post-registration study, designed to evaluate safety and QoL. This study was used by Sabater to calculate a pre- and post-progression utility for trifluridine/tipiracil of 0.72 and 0.59, respectively.[94] These values match the pre and post-progression values calculated from CORRECT and CONCUR i.e. also 0.72 and 0.59. In the economic model the values from CORRECT and CONCUR are used in the base case for both regorafenib and trifluridine/tipiracil (Table 27).

The use of utility values in the economic model are aligned to the reference case as 1) the EQ5D is preferred by NICE 2) the values are directly elicited from patients 3) the UK tariff was used.

Table 27: Utilities used in the base case of the model

	Utility	SD	N	Source
Pre-progression				
Regorafenib, trifluridine/tipiracil and BSC	0.72	0.26	2600	Pooled average of EQ-5D scores for patients on treatment in the CORRECT and CONCUR trials.
Post-progression				
Regorafenib, trifluridine/tipiracil and BSC	0.59	0.34	570	Pooled average of EQ-5D scores for patients at the end of treatment in the CORRECT and CONCUR trials.

Key: BSC, best supportive care; EQ-5D, Euro quality of life, 5-dimensional; NR, not reported.
Source: CORRECT CSR[45], CONCUR CSR[48]

B.3.4.2. Mapping

No mapping was required, as EQ-5D-3L HRQL data were collected in both CORRECT and CONCUR.

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

In appendix H describe how systematic searches for relevant health-related quality-of-life data were done.

An SLR was performed to identify utility data for patients receiving ≥ 3L treatment for mCRC, and to validate the utility data observed in CORRECT and CONCUR trials. Full details of the search strategy and outcomes are described in Appendix H, and the identified unique studies are described below. Notably, many of the identified economic evaluations applied utility values derived from prior utility sources. Where this is the case, solely the originator study is reported.

In general, the utility values from the SLR are in line with the pooled CORRECT and CONCUR values used in the model (see Table 28). None of the alternative values was considered to be 'better' than using values derived from the CORRECT and CONCUR studies.

Several studies reported slightly higher progression-free utility values (0.70–0.81 in the SLR versus 0.72 in the model). In addition, the identified studies reported similar progression-free utility values between arms, with the active treatment sometimes showing a higher utility value than BSC.

Table 28: Utility systemic literature review results

Source* (year)	Population	Country	Progression-free utility		Progressed utility value		Comments
			Treatment	Utility	Treatment	Utility	
Kashiwa (2021)[95]	Later-line *KRAS* WT mCRC	Japan	Panitumumab + BSC BSC	0.73 0.68	All treatments	0.68	Disutility of skin-toxicity: − 0.033 Disutility of Severe skin-toxicity: − 0.10
TA668 (2021)[8]	*BRAF* mutation-positive mCRC	UK	Enco. + cetuximab FOLFIRI Trifluridine/tipiracil (average):	0.743 0.741 0.742	Enco. + cetuximab FOLFIRI: Trifluridine/tipiracil (average)	0.622 0.631 0.627	ERG rejected the use of average utility values for trifluridine/tipiracil and preferred using CORRECT utility values
Sabater (2019)[77]	Third-line and later mCRC	Multiple (excl. UK)	Trifluridine/tipiracil	0.72	Trifluridine/tipiracil	0.59	Based on PRECONNECT RWE study
Xu (2018)[96]	Later-line *KRAS* WT mCRC	US	Cetuximab Panitumumab	0.74 0.745	Cetuximab Panitumumab	0.65 0.65	Cetuximab utilities were derived with HUI-3
Graham (2016)[97]	Later-line *KRAS* WT mCRC	US	Cetuximab Panitumumab	0.796 0.810	Active treatment BSC	0.749 0.602	
Koukakis (2016)[98]	Third-line and later mCRC		Panitumumab + BSC BSC	0.78 0.73	NR	NR	Only baseline values were reported
TA405 (2016)[9]	Third-line and later mCRC	UK	Trifluridine/tipiracil BSC	0.73 0.74	Trifluridine/tipiracil BSC	0.64 0.64	Progression-free utilities were based on only the baseline utilities, not on all pre-progression values ERG critiqued the use of average CORRECT and TA176 utility values and preferred CORRECT only

Source* (year)	Population	Country	Progression-free utility		Progressed utility value		Comments
			Treatment	Utility	Treatment	Utility	
Chang (2015)[99]	Third-line and later mCRC	Asia	Regorafenib (combined value) BSC (combined value)			0.70 0.74	Study only calculated utility for the full CONCUR trial period
Hoyle (2013)[100]	Third-line and later mCRC	UK	BSC: Cetuximab: Panitumumab:	0.75 0.78 0.78	BSC:	0.69	Cetuximab and panitumumab values were capped at the general population utility
Seal (2013)[87]	Third-line and later mCRC	US	Regorafenib BSC	0.71 0.71	Regorafenib BSC	0.59 0.59	
Siena (2013)[101]	Third-line and later mCRC	NR	Regorafenib (combined value) BSC (combined value)		0.67 0.67		Study only calculated utility for the full CORRECT trial period
Blank (2011)[102]	Third-line and later mCRC	Switzerland	Cetuximab BSC	0.77 0.70	All patients	0.5	Progression-free utility is for responders only
Shiroiwa (2010)[103]	Later-line *KRAS* WT mCRC	Japan	Cetuximab-effective Cetuximab ineffective/ no-cetuximab group	0.7 0.7	NR	NR	

*Only unique utility studies were include here. A full overview of utility SLR results, including studies using the same utility value, is available in Appendix H
Key: BSC, best supportive care, ERG, evidence review group; HUI-3, Health Utilities Index 3; mCRC, metastatic colorectal cancer; NR, not reported; RWE, real-world evidence; WT, wild-type.

B.3.4.4. Adverse reactions

To account for the effects of AEs, the model uses data from the CORRECT and CONCUR CSRs for regorafenib and BSC, and the main publications of RECOURSE, TERRA, and Yoshino for trifluridine/tipiracil to calculate AE rates per cycle.[45, 48, 51-53] The model considers AEs of Grade 3 and higher that occurred in at least 2% of patients in any treatment arm. This cut-off was chosen to ensure that infrequent but costly or severe AEs (e.g. febrile neutropenia, which occurred in 3.8% of trifluridine/tipiracil patients, or anaemia and thrombocytopenia, which occurred in 2.8% of regorafenib patients)[45, 51] are also considered in the model. The average AE rate per cycle per treatment across the different trials was calculated by combining the observed AEs and number of patients per arm, for the different trials. This was divided by the weighted average treatment duration to get an AE rate per treatment cycle and then converted to a probability per week, to be used in the model. These probabilities were combined with the costs per AE (discussed in Section B.3.5.3) to get the average AE cost per patient per week (i.e. per model cycle), which were applied to each model cycle until progression.

The pooled AE rates and resulting probabilities per week used in the model base case are shown in Table 29, below. As discussed in Section B.2.11, AEs in CORRECT and CONCUR were infrequent and mostly well tolerated.[45, 48] Trifluridine/tipiracil showed a different AE profile, with high rates for haematological AEs, as typically observed for chemotherapy. The overall rate of grade 3+ AEs of trifluridine/tipiracil was comparable to regorafenib.

Table 29: Grade 3+ treatment related AEs reported in at least 2% of patients in any treatment arm

Adverse event (Grade 3+)	Regorafenib (CORRECT and CONCUR)		BSC (CORRECT and CONCUR)		Trifluridine /tipiracil (RECOURSE, TERRA and Yoshino)	
	Pooled AE rate	Prob. per week	Pooled AE rate	Prob. per week	Pooled AE rate	Prob. per week
Abdominal pain					1.4%	0.10%
Anaemia	2.2%	0.17%			17.8%	1.29%
Anorexia	2.5%	0.20%	2.2%	0.25%		

Adverse event (Grade 3+)	Regorafenib (CORRECT and CONCUR)		BSC (CORRECT and CONCUR)		Trifluridine /tipiracil (RECOURSE, TERRA and Yoshino)	
	Pooled AE rate	Prob. per week	Pooled AE rate	Prob. per week	Pooled AE rate	Prob. per week
Decreased appetite					2.1%	0.15%
Diarrhoea	5.7%	0.45%	0.6%	0.07%	2.5%	0.18%
Fatigue	8.2%	0.65%	4.4%	0.50%	3.1%	0.22%
Febrile neutropenia					2.7%	0.20%
Hand–foot skin reaction	16.5%	1.30%	0.3%	0.04%		
Hypertension	8.0%	0.63%	1.2%	0.14%		
Leukopenia					21.9%	1.59%
Lymphopenia					5.5%	0.40%
Mucositis	2.4%	0.19%				
Nausea						
Neutropenia					37.8%	2.72%
Rash	5.5%	0.44%				
Thrombocytopenia	2.8%	0.22%	0.3%	0.04%	4.4%	0.32%
Vomiting						
Hyperbilirubinaemia	3.0%	0.24%	0.9%	0.11%	7.0%	0.51%
Hypophosphataemia	4.4%	0.35%	0.3%	0.04%	4.6%	0.33%
Increase in ALT level					1.4%	0.10%
Increase in AST level					3.6%	0.26%
Increase in lipase level	3.5%	0.27%	0.6%	0.07%		

Key: AE, adverse event; ALT, alanine aminotransferase, AST, aspartate aminotransferase; BSC; best supportive care.
Source: CORRECT CSR[45], CONCUR CSR[48], Mayer (2015)[51], Xu (2018)[52], Yoshino (2012)[53]

B.3.4.5. AE utility decrement

To capture the impact of AE on the patient's QoL, utility decrements were also included in the model. Although it is likely that AE disutility is already captured within the utilities observed in the trials, most of the AEs were transient in nature. It is therefore uncertain whether patients would have had any disutility from the experienced AE on the day the EQ5D was administered. AE disutilities were

therefore included in the base case for all treatments, to ensure the different AE profiles are reflected in the utilities used in the model.

The disutilities per AE used in the model are shown in Table 30 below. These utilities were aligned with disutilities used in past TAs. The disutilities were combined with the pooled weekly AE probabilities, as reported in Section B.3.4.4, to calculate the average AE disutility per treatment. In line with TA405, the model assumes an AE duration of 1 week, so the average AE disutility was directly subtracted from the pre-progression utility to generate treatment specific utilities. The resulting AE utilities in the base case were -0.00361, -0.00770, and -0.00124 for regorafenib, trifluridine/tipiracil, and BSC respectively. This is similar to TA405, where the ERG subtracted a disutility of 0.01 from the trifluridine/tipiracil pre-progression health statue utility.[9] The effect of including AE disutilities will also be explored in a scenario where disutilities were excluded.

Table 30: Disutilities per adverse event included in the model

Adverse event (Grade 3+)	Disutility	Source
Abdominal pain	0.103	Assumed equal to diarrhoea (Lloyd et al. (2006))
Anaemia	0.090	Assumed equal to Neutropenia (Nafees et al. (2008))
Anorexia	0.103	Assumed equal to diarrhoea (Lloyd et al. (2006))
Decreased appetite	0.103	Assumed equal to diarrhoea (Lloyd et al. (2006))
Diarrhoea	0.103	Lloyd et al. (2006)
Fatigue	0.115	Lloyd et al. (2006)
Febrile neutropenia	0.115	Assumed equal to fatigue (Lloyd et al. (2006))
Hand–foot skin reaction	0.032	Assumed equal to Skin reactions (Nafees et al. (2008))
Hypertension	0.069	Doyle et al. (2008)
Leukopenia	0.090	Assumed equal to Neutropenia (Nafees et al. (2008))
Lymphopenia	0.090	Assumed equal to Neutropenia (Nafees et al. (2008))
Mucositis	0.032	Assumed equal to Skin reactions (Nafees et al. (2008))
Nausea	0.103	Lloyd et al. (2006)
Neutropenia	0.090	Nafees et al. (2008)

Adverse event (Grade 3+)	Disutility	Source
Rash	0.032	Assumed equal to Skin reactions (Nafees et al. (2008))
Thrombocytopenia	0.090	Assumed equal to Neutropenia (Nafees et al. (2008))
Vomiting	0.103	Lloyd et al. (2006)
Hyperbilirubinaemia	0.090	Assumed equal to Neutropenia (Nafees et al. (2008))
Hypophosphataemia	0.090	Assumed equal to Neutropenia (Nafees et al. (2008))
Increase in ALT level	0.090	Assumed equal to Neutropenia (Nafees et al. (2008))
Increase in AST level	0.090	Assumed equal to Neutropenia (Nafees et al. (2008))
Increase in lipase level	0.090	Assumed equal to Neutropenia (Nafees et al. (2008))
Key: ALT, alanine aminotransferase, AST, aspartate aminotransferase. **Source:** Doyle et al. (2008)[104], Lloyd et al. (2006)[105], Nafees et al. (2008)[106].		

B.3.4.6. Health-related quality-of-life data used in the cost-effectiveness analysis

The utilities used in the model per treatment per health state are described in Section B.3.4.1 and shown in Table 27, with average AE disutility per treatment described in Section B.3.4.1 based on the disutilities in Table 30. These were combined to compute the overall utility value per treatment, per health state, as shown in Section B.3.4.5, below.

Age effect on utilities was not included in the model. Considering that later line mCRC is a severe condition, with the observed median OS in CORRECT and CONCUR ranging from 5.0 to 8.8 months, it was assumed that age would only have a minor impact on model outcomes. This is further illustrated that only 24.3% and 36.6% of regorafenib patients, and 24.0% and 16.6% of BSC patients were observed to survive beyond year 1 in CORRECT and CONCUR, respectively (Appendix M). Any age effect on utility would therefore only apply to a small subset of patients and is unlikely to impact the model results. We therefore decided not to explicitly model age-effect on utilities, as a simplifying assumption.

Table 31: Summary of utility values for cost-effectiveness analysis

State and utility input category	Utility value: mean (standard error)	95% confidence interval	Reference in submission (section and page number)	Justification
Progression free				
Regorafenib, trifluridine/tipiracil, and BSC	0.72 (0.005)	(0.710, 0.730)	Section B.3.4.1, page 123	Utilities from trials used when available
Regorafenib AE disutility (PFS)	-0.00361 (N/A*)	(0.00326, 0.00397)	Section B.3.4.5, page 129	Average disutility per treatment calculated and applied to utility value from trials
Trifluridine/tipiracil AE disutility (PFS)	-0.00770 (N/A*)	(0.00696, 0.00847)		
BSC AE disutility (PFS)	-0.00124 (N/A*)	(0.00112, 0.00137)		
Final regorafenib PFS utility used in model	0.716	N/A	Combination of utility value and disutility, reported in Sections B.3.4.1 and B.3.4.5, pages 123 and 129	
Final trifluridine/tipiracil PFS utility used in model	0.712	N/A		
Final BSC PFS utility used in model	0.719	N/A		
Progressed				
Regorafenib, trifluridine/tipiracil, and BSC	0.59 (0.014)	(0.562, 0.618)	Section B.3.4.1, page 123	Utilities from trials used when available

Key: AE, adverse event; BSC, best supportive care; N/A, not available; PFS, progression-free.
*When no standard error was reported, a assumed standard error of 5% was used instead.
Source: CORRECT Clinical Study Report[45]; CONCUR Clinical Study Report[48], Sabater (2019)[94]

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

There are no NHS reference costs or payment-by-results tariffs specific for costing regorafenib. Details about the cost estimation of treatment with regorafenib, in terms of acquisition and administration, are described below.

There have been several recent NICE technology appraisals in later-line mCRC that are relevant to the decision problem (TA242[88], TA307[89], TA405[9] and TA668[8]). Of these, TA405 is particularly relevant as it evaluated trifluridine/tipiracil, the only other active treatment approved in this population. We have therefore utilized the resource

costs identified within TA405 for this submission. These resource use assumptions were supplemented by literature data where needed and confirmed by clinical experts.

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

B.3.5.1. Intervention and comparators' costs and resource use

B.3.5.1.1. Regorafenib active treatment costs

The model applies a 160 mg dose of regorafenib given each day for three weeks for followed by one week of rest (see Appendix A).[92] This is consistent with the marketing authorization for regorafenib. The list price of a pack of 84 x 40 mg tablets is £3,744.00, which corresponds with one 28-day cycle consisting of 21 days of treatment at a dose of 160mg (4 x 40 mg tablets).[107] However, under the terms of a confidential PAS, [redacted]. This PAS price is used in the model.

In accordance with the anticipated pharmacy administration schedule for regorafenib, the model assumed patients who are yet to cease treatment at the start of each 4-week treatment cycle would be dispensed all the medicine required to last them for the next 4 weeks. This means that if a patient stops treatment halfway through the cycle, the full costs of the ongoing cycle are considered in the model. Clinical experts confirmed that this is in line with the expected treatment use in UK clinical practice.

B.3.5.1.2. Trifluridine/tipiracil active treatment costs

As discussed in Section B.3.2.3, trifluridine/tipiracil dosage is based on body surface area (BSA), with a licensed dose of 35 mg/m^2 in mCRC.[91] To inform the BSA distribution, the model uses Sacco et al., who reported BSA data for adult UK cancer patients.[108] This approach deviates from the manufacturers approach in TA405, where the company directly used BSA data from the RECOURSE trial.[9] In TA405, the ERG concluded that the RECOURSE BSA data likely underestimates a UK patient's BSA due to a higher proportion of Asian participants in the trial, and preferred BSA from a healthy UK population instead. However, the committee felt

that healthy population data is in turn likely to overestimate the BSA of cancer patients. We therefore considered the UK BSA data of cancer patients from Sacco et al. to be a good middle ground between the company's and the ERG's approach in TA405 and aligned with the committees considerations.

Trifluridine/tipiracil is given in 28-day treatment cycles, consisting of 2 weeks of active treatment and two weeks of rest. During the active treatment weeks, trifluridine/tipiracil is given twice daily at 35 mg/m^2 for 5 days, with 2 days of rest per week. Therefore, a full 28-day cycle consists of 20 doses. Trifluridine/tipiracil is available in 15 mg and 20 mg tablets for £500.00 and £666.67 per 20 tablets (NHS list price), respectively.[107] The distribution of trifluridine/tipiracil dosing across BSA categories is given in Table 32. Applying a UK cancer patient's BSA distribution (assuming 55.8% of mCRC patients are male[3]) to this dosing schedule[108] results in an average dose of 1.53 x 15 mg and 2.16 x 20 mg tablets, costing £2,147.40 per cycle; this is used in the model. As per regorafenib, all costs are incurred at the start of a cycle, and wastage is applied when patients discontinue treatment.

Table 32: Trifluridine/tipiracil dosing based on BSA

BSA	Dosage (mg; twice daily)	15 mg units	20 mg units	Cost per cycle (list price)
< 1.07	35	1	1	£1,167
1.07–1.22	40	-	2	£1,333
1.23–1.37	45	3	-	£1,500
1.38–1.52	50	2	1	£1,667
1.53–1.68	55	1	2	£1,833
1.69–1.83	60	-	3	£2,000
1.84–1.98	65	3	1	£2,167
1.99–2.14	70	2	2	£2,333
2.15–2.29	75	1	3	£2,500
≥ 2.30	80	-	4	£2,667

Key: BSA, body surface area; SmPC, Summary of Product Characteristics.
Source: Trifluridine/tipiracil SmPC.[91]

B.3.5.1.3. BSC treatment costs

As discussed in Section B.3.2.3, BSC can consist of a variety of concomitant treatments, procedures and other palliative care. However, in line with the approach in past mCRC appraisals, these costs are assumed to be captured by BSC

healthcare resource use (HRU) costs, discussed below.[8, 9] Treatment costs for BSC are therefore assumed to be £0.

B.3.5.1.4. Missed dose and dose reductions

As discussed in Section B.2, not all patients received the full dose of regorafenib during the clinical trials. In practice, the dose patients received in CORRECT and CONCUR was lower due to dose reductions and cycle delays in response to AEs. The model therefore includes a reduced dose intensity (RDI) to reflect the dose received and to account for how the treatments will be used in clinical practice, and to ensure that modelled dosing is in line with the corresponding efficacy data.

RDI information from the different regorafenib and trifluridine/tipiracil trials is presented in Table 33. The mean dose patients received was similar across CORRECT and CONCUR ([redacted]% versus [redacted]% of the licensed 160 mg). For consistency with the efficacy input, the RDI from CORRECT and CONCUR was pooled, using a simple weighted average.

In the base case, this RDI was applied directly to the received regorafenib cost per cycle. We also explore a scenario in which RDI is applied to the number of pills dispensed, rather than directly to the regorafenib costs. In this scenario regorafenib costs are only saved when the average patient has enough residual pills for a full prescription, in which no costs are incurred for that cycle as there is no need to open a new pack. However, this approach is likely to overestimate regorafenib costs.

Table 33: Dose intensity information used in the model

Study	Relative dose intensity (N)	Dose reduction	Cycle delay
Regorafenib – CORRECT	[redacted] (500)	N/A	N/A
Regorafenib – CONCUR	[redacted] (136)	N/A	N/A
Regorafenib – CORRECT and CONCUR pooled	[redacted]	N/A	N/A
Trifluridine/tipiracil – TA405	N/A	97.4%	2.72 days

Key: NICE, National Institute for Health and Care Excellence.
Source: CORRECT CSR[45]; CONCUR CSR[48]; Mayer et al., 2015[51]; NICE TA405.[9]

For trifluridine/tipiracil, no single RDI measure has been reported. Trifluridine/tipiracil dose reductions and cycle delays were therefore modelled separately instead, with data on the number of dose reductions and cycle delays from TA405 (see Table 33, above).[9] Although the approach differs, using a combination of dose reductions and cycle delays approximates how RDI was assessed for regorafenib in both CORRECT and CONCUR.[45, 48] Importantly, these different modelling approaches are more reflective of clinical practice, as the trifluridine/tipiracil toxicity is managed by delaying subsequent doses rather than reducing the dose. This was also confirmed by the consulted clinical experts who stated that regorafenib and trifluridine/tipiracil RDI cannot be modelled using the same metric, as the dose of regorafenib is generally reduced if toxicities develop, whereas trifluridine/tipiracil is delayed. We again took a conservative approach by assuming that all trifluridine/tipiracil dose reductions were already applied during the first dose and continued for the full course of treatment; in practice, the dose would decline gradually. To account for the uncertainty around the trifluridine/tipiracil RDI estimate, we explored a scenario where we assumed equal RDI between trifluridine/tipiracil and regorafenib.

B.3.5.1.5. Administration costs

All active treatments in the model are oral, and therefore do not incur any administration costs, in line with past TAs.[9] Any costs related to routine visits and dispensing are assumed to be covered by the HRU costs, discussed below.

B.3.5.2. Health-state unit costs and resource use

HRU estimates in the model were informed by the HRU used in the studies identified by the SLR, and confirmed and validated by clinical experts. The SLR identified a total of four studies, which reported HRU rates for later line mCRC: TA405[9], Bullement (2018)[79], Hoyle (2013)[100], and TA668[8]. The different HRU categories and rates reported in these SLR studies are shown in Table 34.

HRU items in the SLR studies were categorized by progression status and whether patients receive active treatment. Although the active treatment differed, the chosen HRU rates were comparable across studies. Most studies included at least 1 oncologist or chemotherapy visit per month, along with a health home visitor for a subset of patients. Both Hoyle 2013 and the ERG in TA405 also assumed 33% of patients would undergo a CT scan per month. Most studies assumed lower HRU

rates for BSC than for the active treatment. The ERG preferred rates from TA405 were therefore chosen for the model base case, as they represent a good middle ground between all identified studies, in terms of combining the rates for active treatment rates and using lower rates for BSC. They are also directly applicable to the positioning of regorafenib and maintain consistency of approach between appraisals.

The chosen HRU rates were also validated during a clinical advisory board. The consulted clinical experts broadly agreed with using the ERG-preferred rates from TA405, with just one expert questioning the post-progression GP surgery visit. However, since this was only mentioned once, we decided to continue with the values as reported in TA405. The experts confirmed that it is reasonable to use the same HRU rates for regorafenib and trifluridine/tipiracil, which are therefore assumed equal.

Table 34: Monthly HRU rates in studies identified by the SLR

HRU category	TA405 ERG preference (Current model's base case)			TA405 company submission. / Bullement (2018)			Hoyle 2013		TA668	
	Lon	BSC	PP	T/T	BSC	PP	Bev	BSC	Enco	PP
Oral chemotherapy day-case	100%	-	-	100%	-	-	-	-	50%	
Medical oncologists OP visit	-	-	-	-	100%	-	200%	-	50%	
GP home consultation	-	-	25%	-	-	25%	-	-		25%
Community nurse specialist visit	-	-	100%	-	-	100%	-	-		100%
Health home visitor	25%	25%	100%	25%	25%	100%	-	-	50%	100%
District nurse visit	-	-	100%	-	-	100%	-	-		100%
GP surgery visit	-	-	100%	-	-	100%	-	-		100%
CT scan	33%	-	-	-	-	-	33%	-		

Key: BSC, best supportive care; CT, computed tomography; Enco., encorafenib; GP, general practitioner, HRU, health resource use, NHS, OP, outpatient, PP, post-progression; Rego., regorafenib, T/T, trifluridine/tipiracil.

Source: NICE TA405[9], Bullement (2018)[79], Hoyle (2013)[100], and TA668[8]

B.3.5.2.1. Progression-free HRU

Patients on regorafenib and trifluridine/tipiracil are assumed to attend an oral chemotherapy outpatient appointment (per treatment cycle). During this appointment they receive treatment for the upcoming cycle, undergo routine tests and see a clinician to review their treatment.[9] Furthermore, 33% of regorafenib and trifluridine/tipiracil patients are assumed to undergo a computed tomography (CT) scan per cycle. BSC patients are not assumed to attend any routine oncologist visits. It was also assumed that 25% of all patients incurred the cost of a health home visitor per treatment cycle, regardless of their treatment, based on expert opinion on palliative care elicited in TA405.[9] The HRU assumptions and costs used in the model base case are summarized in Table 35, below.

B.3.5.2.2. Progressed HRU

Following progression, HRU is expected to change as patients receive more palliative and home-based care. Consequently, patients are no longer assumed to attend day case or outpatient consultations and instead receive care closer to home (i.e. home care, community nurse, general practitioner [GP], etc.). The HRU assumptions and costs used in the model base case are summarized in Table 35, below. All costs were sourced using the latest sources (checked 4 May 2022). Costs were inflated using the 2021 PSSRU price index .[109]

Table 35: HRU assumptions used in the model, based on TA405 ERG report and expert input

Resource item	% of patients using			Unit cost (£)	Cost source
	Prog.-free		Prog.		
	Rego. /Lon.	BSC	All pts.		
Oral chemotherapy outpatient	100%	-	-	£208.24	NHS reference costs 2019/20: Outpatient; SB11Z; Deliver exclusively oral chemotherapy, inflated to 2021
GP home consultation	-	-	25%	£82.80	PSSRU 2021: Calculated based on GP cost per minute (£3.60, without qualifications), assuming out of surgery visit lasting 23 minutes - 10.3b
Community nurse specialist visit	-	-	100%	£55.00	PSSRU 2021: Band 6 Nurse (Community) Cost per hour - 10.1 (contact assumed to last 1 hour)
Health home visitor	25%	25%	100%	£32.00	PSSRU 2021: Home care worker Cost per hour (Face-to-face visit for social services) - 11.5 (contact assumed to last 1 hour)
District nurse visit	-	-	100%	£44.00	PSSRU 2021: Band 5 Nurse (Community) Cost per hour - 10.1 (contact assumed to last 1 hour)
GP surgery visit	-	-	100%	£33.00	PSSRU 2021: GP consultation (Per surgery consultation lasting 9.22 minutes, without qualifications) - 10.3b
CT scan	33%	-	-	£103.52	NHS reference costs 2019-20: RD26Z: Computerised Tomography Scan of Three Areas, with Contrast, Outpatient, inflated to 2021

Key: BSC, best supportive care; CT, computed tomography; GP, general practitioner, HRU, health resource use, NHS, National Health Service; OP, outpatient, Prog., progression; PSSRU, Personal Social Services Research Unit; Rego., regorafenib.
Source: NHS reference costs 2019/20[110], PSSRU 2021[109]

B.3.5.3. Adverse reaction unit costs and resource use

A description of the AEs included in the model, and the corresponding frequencies, are presented in Section B.3.4. AE costs information was mostly obtained from NHS reference costs 2019/2020, and from past NICE TAs when a corresponding NHS reference cost was not available (see Table 36).

Table 36: Costs per adverse event included in the model

Adverse event (Grade 3+)	Cost*	Source
Abdominal pain	£182.99	NHS reference costs 2019/20: Service code: 191, Pain Management, Outpatient Attendance
Anaemia	£1,211.78	NHS reference costs 2019/20; NES and NEL; weighted average of HRG codes: SA04G, H, J, K and L
Anorexia	£182.13	NHS reference costs 2019/20: Service code: 300, General Medicine, Outpatient Attendance
Decreased appetite	£182.13	NHS reference costs 2019/20: Service code: 300, General Medicine, Outpatient Attendance
Diarrhoea	£714.37	NHS reference costs 2019/20; NES; Weighted average of PF26A&B; Paediatric Other Gastrointestinal Disorders with CC Score 1+; Short Stay
Fatigue	£13.86	NICE ERG report abiraterone (TA259), table 24, p. 64, inflated to 2021
Febrile neutropenia	£2,933.55	The NICE DSU report on the cost of febrile neutropenia 2007, inflated to 2021
Hand–foot skin reaction	£182.13	NHS reference costs 2019/20: Service code: 300, General Medicine, Outpatient Attendance
Hypertension	£640.22	NHS reference costs 2019-20: HRG EB04Z, Hypertension
Leukopenia	£180.18	TA405 ERG preferences, inflated to 2021
Lymphopenia	£180.18	Assumed equal to Leukopenia
Mucositis	£182.13	NHS reference costs 2019/20: Service code: 300, General Medicine, Outpatient Attendance
Nausea	£182.13	NHS reference costs 2019-20: Service code: 300, General Medicine, Outpatient Attendance
Neutropenia	£180.18	TA405 ERG preferences, inflated to 2021
Rash	£182.13	NHS reference costs 2019/20: Service code: 300, General Medicine, Outpatient Attendance

Adverse event (Grade 3+)	Cost*	Source
Thrombocytopenia	£1,909.15	NHS reference costs 2019/20; NES and NEL; weighted average of HRG codes: SA12G, H, J, and K
Vomiting	£182.13	NHS reference costs 2019/20: Service code: 300, General Medicine, Outpatient Attendance
Hyperbilirubinaemia	£182.13	NHS reference costs 2019/20: Service code: 300, General Medicine, Outpatient Attendance
Hypophosphataemia	£182.13	NHS reference costs 2019/20: Service code: 300, General Medicine, Outpatient Attendance
Increase in ALT level	£182.13	NHS reference costs 2019/20: Service code: 300, General Medicine, Outpatient Attendance
Increase in AST level	£182.13	NHS reference costs 2019/20: Service code: 300, General Medicine, Outpatient Attendance
Increase in lipase level	£182.13	NHS reference costs 2019/20: Service code: 300, General Medicine, Outpatient Attendance

Key: ALT, alanine aminotransferase, AST, aspartate aminotransferase; CC, chronic constipation DSU, Decision Support Unit; HRG, Healthcare Resource Group; NEL, non-elective long stay, NES, non-elective short stay, NHS, National Health Service; NICE, National Institute for Health and Care Excellence.
*All 2019/20 NHS reference costs were inflated to 2021 using the 2021 PSSRU price index[109]
Source: NHS reference costs 2019/20[110], TA405[9], TA259[111], the NICE DSU report on the cost of febrile neutropenia.[112]

The pooled weekly AE probabilities from Section B.3.4.4 were combined with the cost per AE to determine the average AE costs per treatment per week. These costs were then applied each model cycle (week) that patients were on treatment. This approach was preferred over calculating a one-off AE cost, as using weekly AE costs ensures the modelled AE costs are reflective of the modelled treatment duration, and provides better accuracy in terms of discounting AEs costs. The resulting weekly AE costs were: trifluridine/tipiracil - £39.95; regorafenib - £19.18; and BSC - £3.10 (see Table 37).

Table 37: Aggregate weekly adverse event costs used in the model

Treatment	Weekly AE cost	Source
Regorafenib	£19.18	Pooled CORRECT and CONCUR AE probabilities, regorafenib arm[45, 48]
BSC	£3.10	Pooled CORRECT and CONCUR AE probabilities, BSC arm[45, 48]
Trifluridine/tipiracil	£39.95	Pooled RECOURSE, TERRA, and Yoshino AE probabilities, trifluridine/tipiracil arm[51-53]
Key: AE, adverse event; BSC, best supportive care. **Source**: CORRECT CSR[45], CONCUR CSR[48], Mayer (2015)[51], Xu (2018)[52], Yoshino (2012)[53]		

B.3.5.4. Miscellaneous unit costs and resource use

B.3.5.4.1. End of life costs

End-of-life costs were taken from Round et al.[113]: a modelling study that estimates the cost of end of life caring for people with CRC. The model only considers healthcare and social care costs, as charity and informal care do not fall within NHS or PSS costs; this is also in line with past TAs.[9] Furthermore, costs were inflated to 2021 using the 2021 PSSRU price index[109], resulting in one-time end-of-life costs of £6,832.17, applied for all patients upon death.

B.3.5.4.2. Subsequent treatment

There was some post-progression treatment in the CORRECT and CONCUR trials (CORRECT: Regorafenib 26%, BSC 30%; CONCUR Regorafenib 31%, BSC 43%). However, clinical experts have advised that in England and Wales, patients receiving regorafenib or trifluridine/tipiracil are unlikely to receive further active treatment after progression due to the advanced nature of the disease and limited treatment options available. The estimate of the proportion of patients who might receive anti-cancer therapy after trifluridine/tipiracil or regorafenib was <10%. In the basecase no post-progression treatment is assumed. A scenario using the subsequent treatment costs as reported in TA405 has been explored to test the impact of subsequent treatments on cost-effectiveness estimates.

B.3.6. Severity

Due to the severity of the disease, patients suffering from ≥ 3L mCRC experience a substantial QALY shortfall, compared to the general population. This is illustrated by the QALY shortfall calculations, as shown in Table 40, with the features of this analysis shown in Table 38 and Table 39. The QALY shortfall was calculated using the QALY Shortfall Calculator by Schneider, McNamara and Love-Koh et al.[114] This calculator uses the national life tables for England, 2017-2019 (pooled) to estimate age- and sex-specific survival times[115], and combines this with age- and sex-specific utilities based on Health Survey for England 2017 and 2018 (pooled)[116] and Hernandez Alava et al.'s EQ-5D-5L to 3L mapping algorithm.[117]

The starting age used for the shortfall calculations was based on a weighted average of the mean starting age in CORRECT (60.5, N=760) and CONCUR (56.5, N=204). This resulted in a weighted average mean age of 59.65, which was rounded up to 60. In addition, the calculation assumes 55.8% of mCRC patients are male rounded up to 56%, based on UK cancer registration data for colorectal cancer.[3] Using the calculator by Schneider et al, this results in a healthy population estimate of 12.36 QALYs remaining. The remaining QALYs for patients with mCRC were informed by the base case model results, presented in Section B.3.10 below, as [redacted] for trifluridine/tipiracil and [redacted] for BSC. Consequently, the proportional shortfall for ≥ 3L mCRC patients is [redacted]% and [redacted]% for patients currently treated with trifluridine/tipiracil and BSC, respectively, justifying a 1.7x QALY weight for both comparisons. This QALY weight will be applied indirectly in the base case by using a higher willingness to pay (WTP) threshold of £51,000, rather than adjusting the QALYs themselves. This approach was considered to be the most transparent, as it also directly shows the ICER results without any severity modification applied.

Table 38: Summary features of QALY shortfall analysis

Factor	Value (reference to appropriate table or figure in submission)	Reference to section in submission
Sex distribution	55.8% male	Section B.3.5.1.2
Starting age	60	Section B.3.6

Table 39: Summary of health state benefits and utility values for QALY shortfall analysis

State	Utility value: mean (standard error)	Undiscounted life years
Pre-progression	0.72 (0.005)	[redacted] (trifluridine/tipiracil) [redacted] (BSC)
Progressed	0.59 (0.014)	[redacted] (trifluridine/tipiracil) [redacted] (BSC)

Table 40: Summary of QALY shortfall analysis

Expected total QALYs for the general population	Total QALYs that people living with a condition would be expected to have with current treatment	Proportional QALY shortfall
12.36	[redacted] (trifluridine/tipiracil)	[redacted]%
12.36	[redacted] (BSC)	[redacted]%

B.3.7. Uncertainty

Any uncertainty related to the condition, and data informing the model is discussed in Section B.3.15.2. In addition, structural uncertainty around the data informing the model is explored in the sensitivity analyses described in Sections B.3.11.1 and B.3.11.2 and any uncertainty related to the choice of data input is explored in the scenarios, described in Section B.3.11.3.

B.3.8. Managed access proposal

Not applicable

B.3.9. Summary of base-case analysis inputs and assumptions

B.3.9.1. Summary of base case analysis inputs

A summary of the variables used in the model are shown in Table 41, below.

Table 41: Summary of variables applied in the economic model

Variable	Value	Confidence interval (distribution)	Reference to section in submission
General settings			
Cycle length	1 week	None	Section B.3.2
Time horizon	10 years	None	
Discount rate for effects and costs	3.5%	None	
Efficacy			
Regorafenib OS curve in use	Loglogistic Shape: [redacted] Scale: [redacted]	Varied together using covariance	Section B.3.3
BSC OS curve in use	Loglogistic Shape: [redacted] Scale: [redacted]	Varied together using covariance	
Regorafenib PFS curve in use	Pooled KM data, followed by an exponential extrapolation	None (for KM data, extrapolation varied using covariance)	
BSC PFS curve in use			
Regorafenib ToT curve in use	Pooled KM data, followed by a log-logistic extrapolation		
BSC ToT curve in use			
T/T PFS HR	0.93	0.85 – 1.03 (Log-Normal)	
T/T OS HR	0.99	0.84 – 1.17 (Log-Normal)	
Drug Costing			
Cost per package: regorafenib 40 mg tablet, pack of 84	[redacted]	None	Section B.3.5.1
Cost per pack: Trifluridine/tipiracil 15 mg tablet, pack of 20	£500	None	
Cost per pack: Trifluridine/tipiracil 20 mg tablet, pack of 20	£666.67	None	
Dosing: % males in population	55.80%	50.33% – 61.27% (Normal)	
Dosing: mean BSA (male)	1.93	1.91 – 1.95 (Normal)	
Dosing: mean BSA (female)	1.68	1.66 – 1.70 (Normal)	
Dose intensity: Regorafenib (mean)	[redacted]	[redacted] (Beta)	
Dose reduction: T/T (mean)	97.44%	97.19% – 97.69% (Beta)	

Average cycle delay: T/T (mean)	2.72	2.45 – 2.99 (Normal)	
Healthcare resource use			
Oral chemotherapy day-care – Regorafenib, Pre-progression	100%	90.20% – 109.80% (Log-Normal)	Section B.3.5.2
Health home visitor – Regorafenib, Pre-progression	25%	22.55% – 27.45% (Log-Normal)	
CT scan – Regorafenib, Pre-progression	33%	30.07% – 36.60% (Log-Normal)	
Health home visitor – BSC, Pre-progression	25%	25.50% – 27.45% (Log-Normal)	
Oral chemotherapy day-case – T/T, Pre-progression	100%	90.20% - 109.80% (Log-Normal)	
Health home visitor – T/T, Pre-progression	25%	22.55% – 27.45% (Log-Normal)	
CT scan – T/T, Pre-progression	33%	30.07% - 36.56% (Log-Normal)	
GP home consultation – All Tx, Progressed	25%	22.55% - 27.45% (Log-Normal)	
Community nurse – All Tx, Progressed	100%	90.20% - 109.80% (Log-Normal)	
Health home visitor – All Tx, Progressed	100%	90.20% - 109.80% (Log-Normal)	
District nurse visit – All Tx, Progressed	100%	90.20% - 109.80% (Log-Normal)	
GP surgery visit – All Tx, Progressed	100%	90.20% - 109.80% (Log-Normal)	
End of life costs	£6,832.17	£6,178.95 - £7,517.74 (Gamma)	
Average weekly adverse event costs			
Total aggregate adverse event costs for Regorafenib	£19.18	£17.35 – £21.11 (Gamma)	Aggregate value calculated based on AE rates in Section B.3.4.4 and AE costs in Section B.3.5.3
Total aggregate adverse event costs for BSC	£3.10	£2.80 – £3.41 (Gamma)	
Total aggregate adverse event costs for Trifluridine/tipiracil	£39.95	£36.13 – £43.96 (Gamma)	

Average adverse event disutilities			
Total aggregate adverse event disutility for Regorafenib	0.00361	0.00326 – 0.00397 (Beta)	Aggregate value calculated based on AE rates in Section B.3.4.4 and AE disutilities in Section B.3.4.5
Total aggregate adverse event disutility for BSC	0.00124	0.00112 – 0.00137 (Beta)	
Total aggregate adverse event disutility for Trifluridine/tipiracil	0.00770	0.00696 – 0.00847 (Beta)	
Utilities			
CORRECT and CONCUR utility – Pooled PFS	0.72	0.7088 – 0.7311 (Beta)	Section B.3.4.1
CORRECT and CONCUR utility – Pooled PPS	0.59	0.5620 – 0.6178 (Beta)	
Key: AE, Adverse events; BSA, Body surface area; BSC, Best supportive care; CT, Computerized tomography; HR, Hazard ratio; mg, milligram; OS, Overall survival; PFS, Progression-free survival; PPS, post-progression; T/T, trifluridine/tipiracil.			

B.3.9.2. Assumptions

An overview of the most important model assumptions are shown in Table 42, below.

Table 42: Key model assumptions

Assumption	Justification
It is methodologically sound to pool CORRECT and CONCUR efficacy data	To increase sample size and to make use of all available patient-level data (n = 964), OS, PFS and ToT data from CORRECT and CONCUR were pooled. Despite differences in prior VEGF use and ethnicity, patient characteristics and outcomes of trials were considered sufficiently similar to justify pooling data to maximize the sample size of data informing the model. Pooling was considered to be appropriate by consulted clinical experts. In addition, a similar approach was used in TA405[9], where data of trials with differences in ethnicity and prior VEGF use were also pooled. The impact of using pooled data will be explored in the scenario analyses.
Treatment practice in the UK is expected to be similar as treatment use in CORRECT and CONCUR with respect to dose intensity and ToT	The reported dose intensity was similar between CORRECT and CONCUR, indicating that treatment practice is not impacted by region, thereby increasing the robustness of the dose intensity input. In addition, the small gap between the Kaplan–Meier curves of PFS and ToT are in line with how UK clinical experts expect regorafenib to be used
The frequency of assessing radiographic progression in CORRECT and CONCUR is	The trial protocols dictated that radiographic progression was assessed on an 8-weekly basis. Consulted clinical experts confirmed that this is in line with UK clinical

Assumption	Justification
in line with how progression is assessed in UK practice	practice. Although some centres use a 12-weekly schedule, checks are often moved up, resulting in an 8-weekly schedule in practice. This schedule was also reflected in the HRU assumptions used in the model.
All patients receive BSC post progression, meaning that no active treatment costs are modelled in the progressed health state	Clinical experts confirmed that, considering the progressed nature of regorafenib and trifluridine/tipiracil patients, the majority of patients would not receive further active treatment after progression in UK practice. This is also in line with the limited treatment options available in the UK for late line mCRC patients.
Key: BSC, best supportive care; ITC, indirect treatment comparison; mCRC, metastatic colorectal cancer; OS, overall survival; PFS, progression-free survival; ToT, time on treatment.	

B.3.10. Base case results

B.3.10.1. Base case incremental cost-effectiveness analysis results

The base case analysis shows that regorafenib patients have a mean OS of [redacted] months and a total of [redacted] QALYs. Patients on trifluridine/tipiracil have a shorter modelled OS of [redacted] months and [redacted] QALYs, resulting in an incremental [redacted] QALY benefit in favour of regorafenib. [redacted], with total costs of [redacted] versus [redacted] respectively. Apart from offering a chemotherapy-free alternative to trifluridine/tipiracil, regorafenib also represents [redacted]

Compared to BSC, regorafenib offers a meaningful survival benefit of [redacted] months compared with patients treated with BSC, with a mean OS of [redacted] and [redacted] months for regorafenib and BSC, respectively. Patients treated with regorafenib accrued an additional [redacted] QALYs at an additional [redacted].

The deterministic ICER and net health benefit (NHB) base case results versus trifluridine/tipiracil and BSC are summarized in Table 43 and Table 44 below. Considering the severity of the disease, NHB is also explored at a WTP threshold of £51,000, corresponding to a QALY weight of 1.7. Overall, these results indicate [redacted] In addition, regorafenib is a cost effective alternative to BSC in mCRC, with an ICER of [redacted].

Table 43: Base-case results

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs regorafenib (£)	Incremental LYG regorafenib	Incremental QALYs regorafenib	ICER regorafenib versus Tx (£/QALY)	ICER incremental (£/QALY)
Regorafenib	[redacted]	[redacted]	[redacted]					
T/T	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Key: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years; T/T, trifluridine/tipiracil.

Table 44: Net health benefit

Technologies	Total costs (£)	Total QALYs	Incremental costs regorafenib (£)	Incremental QALYs regorafenib	NHB regorafenib versus Tx at £20,000	NHB regorafenib versus Tx at £30,000	NHB regorafenib versus Tx at £51,000
Regorafenib	[redacted]	[redacted]					
T/T	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Key: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years; T/T, trifluridine/tipiracil.

B.3.11. Exploring uncertainty

B.3.11.1. Probabilistic sensitivity analysis

Probabilistic sensitivity analysis (PSA) was performed to account for multivariate and stochastic uncertainty in the model. The uncertainties in the individual parameters for treatment effect, costs and utilities were characterized using probability distributions and analysed using a Monte Carlo simulation with 5,000 simulations.

An overview of the probabilistic sensitivity analysis results for the cost-effectiveness of regorafenib versus trifluridine/tipiracil and versus BSC are shown in Table 45, below. The results for both comparisons were also plotted on a cost-effectiveness plane, in which each dot resembles one Monte Carlo simulation and the black line represents a WTP threshold of £30,000 per QALY gained (see Figure 21 and Figure 22). The results were also plotted in a cost-effectiveness acceptability curve (CEAC) versus trifluridine/tipiracil, in which each line is assigned to a treatment and is mapped to display that probability of that treatment being the most cost-effective across a range of WTP thresholds (see Figure 23). As trifluridine/tipiracil is the key comparator only regorafenib and trifluridine/tipiracil were included in the CEAC.

The probabilistic sensitivity analysis results are in line with the deterministic outcomes presented in the base case analysis [redacted]. Against BSC the deterministic ICER was [redacted] and the probabilistic ICER was [redacted]. In addition, the CEAC indicated that regorafenib has a probability to be cost-effective compared to trifluridine/tipiracil of [redacted] and [redacted], at WTP thresholds of £30,000 and £51,000 respectively.

Figure 21: Cost effectiveness plane of regorafenib vs trifluridine/tipiracil (NMB calculated using a WTP of £30,000)

Key: NMB, net monetary benefit; QALY, quality-adjusted life year; REG, regorafenib; T/T, trifluridine/tipiracil.

Figure 22: Cost effectiveness plane of regorafenib vs BSC.

Key: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year; REG, regorafenib.
WTP threshold of £30,000 indicated by line

Figure 23: Cost-effectiveness acceptability curve

Key: BSC, best supportive care; QALY, quality-adjusted life year.

Table 45: Probability sensitivity analysis results

Treatment	QALYs (95% CI)	Total costs (95% CI, £)	ICER (95% CI, £)	NMB* (95% CI, £)
Regorafenib	[redacted]	[redacted]	-	-
Trifluridine/ tipiracil	[redacted]	[redacted]	[redacted]	[redacted]
BSC	[redacted]	[redacted]	[redacted]	[redacted]

Key: BSC, best supportive care; CI, confidence interval; ICER, incremental cost-effectiveness ratio; NMB, net monetary benefit; LYG, life years gained; QALYs, quality-adjusted life years; reg., regorafenib; T/T, trifluridine/tipiracil
*NMB calculated assuming a WTP threshold of £30,000

B.3.11.2. Deterministic sensitivity analysis

During the univariate one-way sensitivity analyses (OWSAs), each input parameter was varied to explore the impact of each parameter on model outcomes. Variables for which no CI and/or standard deviation or error was available have been varied

using an arbitrary standard error of 5%. Parameters with no associated uncertainty, such as drug costs, are excluded from the analysis. Interdependent variables that cannot be varied individually, such as efficacy extrapolation parameters, were also excluded. All parameters included in the OWSA along with the used CI are shown in Table 41, in Section B.3.9.1 above.

The parameters that had the biggest impact on model outcomes have been summarized in a table and plotted in a tornado diagram. The net monetary benefit (NMB) and ICER results for regorafenib versus trifluridine/tipiracil are shown in Table 46 and Figure 24 below. A WTP of £30K was used – results using a WTP of £51K have not been presented but are applicable given that mCRC achieves the highest severity QALY weighting. Overall, the spread in outcomes was narrow, confirming the PSA results that indicated limited structural uncertainty in the model results. The most impactful driver of cost-effectiveness vs trifluridine/tipiracil was the OS HR versus regorafenib, followed by PFS HR versus regorafenib. The sensitivity of outcomes to these inputs is to be expected as both treatments are of comparable efficacy, however, these sensitivity analyses explored efficacy inputs which favoured one treatment over the other. In addition, three out of the ten most impactful parameters relate to how the patients' actual dose is modelled (i.e. regorafenib RDI, and trifluridine/tipiracil cycle delay and dose reduction). This indicates the sensitivity of the model to the ITC assumptions and treatment modelling assumptions and the importance of exploring these inputs more extensively in the scenario analyses.

Overall, the OWSA shows that the analysis is robust with most parameters having little impact on model outcomes, and [redacted]

Figure 24: NMB results of regorafenib vs. trifluridine/tipiracil (WTP of £30,000)

Key: AE, adverse event; HR, hazard ratio; NMB, net monetary benefit; OS, overall survival; PFS, progression-free survival, ToT; time on treatment; T/T, trifluridine/tipiracil.

Table 46: NMB and ICER results of regorafenib vs. trifluridine/tipiracil

Parameter (lower input, upper input)	NMB results vs. T/T		ICER results vs. T/T	
Base case				
	Lower NMB*	Upper NMB*	Lower ICER*	Upper ICER*
T/T OS HR (0.84, 1.17)				
T/T PFS HR (0.85, 1.03)				
Dose intensity: Regorafenib (mean) (0.77, 0.81)				
Oral chemotherapy day-case - Regorafenib, Pre-progression (0.9, 1.1)				
Oral chemotherapy day-case - T/T, Pre-progression (0.9, 1.1)				
Adverse event management cost (mean): Trifluridine/tipiracil (36.13, 43.96)				
Adverse event management cost (mean): Regorafenib (17.35, 21.11)				
Average cycle delay: T/T (mean) (2.45, 2.99)				
Dose reduction: T/T (mean) (0.972, 0.977)				
CT scan - Regorafenib, Pre-progression (0.3, 0.37)				

Parameter (lower input, upper input)	NMB results vs. T/T	ICER results vs. T/T
Key: AE, adverse event; HR, hazard ratio; NMB, net monetary benefit; OS, overall survival; PFS, progression-free survival, [redacted]; ToT; time on treatment; T/T, trifluridine/tipiracil. *NMB calculated assuming a WTP threshold of £30.000 [redacted]		

B.3.11.3. Scenario analysis

B.3.11.3.1. Scenario analyses vs trifluridine/tipiracil

To further explore the uncertainty around the modelled results in respect of key inputs and assumptions a series of scenario analyses with alternative modelling assumptions were performed. As trifluridine/tipiracil is the main comparator, a full set of scenarios is explored. These scenarios explore some of the key uncertainties, as discussed throughout this submission (e.g. pooling of CORRECT and CONCUR data, OS curve selection), along with the key outcome drivers as identified by the OWSA e.g. HR versus trifluridine/tipiracil. All performed scenario analyses are briefly summarized in Table 47, below.

Table 47: Scenario analyses performed versus trifluridine/tipiracil

Scenario category		Scenario name	Description and rationale
1	**Regorafenib efficacy input**	Regorafenib OS, PFS, and ToT data from CONCUR data only	Although consulted experts agreed with the pooling of CORRECT and CONCUR data to inform OS, PFS, and ToT for regorafenib and BSC, there is some uncertainty on how representative these data are to UK practice; most notably for CORRECT, due to the high prior anti-VEGF use. These scenarios therefore explore the effect of pooling the data by only using efficacy data from either CORRECT or CONCUR as efficacy input for regorafenib. Trifluridine/tipiracil is modelled using the same HR point estimates from the full NMA as the base case, which is then applied to the updated CORRECT or CONCUR-only curves for regorafenib.
2		Regorafenib OS, PFS, and ToT data from CORRECT only	
3		Weibull OS	The base case uses a log-logistic OS extrapolation, in line with TA405 and expert input. However, as clinical experts stated that all OS extrapolations provided plausible OS predictions, these scenarios explore the other OS options being used as regorafenib OS input.
4		Log-normal OS	
5		Exponential OS	
6		Generalized gamma OS	
7		Gompertz OS	
8		Gamma OS	
9		Parametric PFS and ToT curves used for regorafenib	Due to the maturity of the PFS and ToT data, the base case uses KM data as PFS and ToT input for regorafenib, to which the trifluridine/tipiracil ITC HR is applied. To explore the impact of this assumption, this scenario uses a fully parametric PFS and ToT extrapolation instead. For this scenario, the log-logistic extrapolation was selected for both PFS and ToT, as it showed the best statistical fit to the PFS and ToT data.
10	**Trifluridine/tipiracil ITC input**	NMA without Yoshino	The ITC input used for trifluridine/tipiracil is a key driver of model outcomes. This is also shown in the OWSA, with the OS HR and PFS HR being the most impactful and third most impactful drivers of outcomes identified by the ITC. For the base case, the full ITC, comparing CORRECT and CONCUR for regorafenib to RECOURSE, TERRA, and Yoshino (2012) for trifluridine/tipiracil, was chosen as the most reliable input, as it makes use of all the available efficacy data. However, there is some uncertainty around that ITC, most notably due to differences in trial design (with/without prior VEGF use, combining phase II with phase III studies) and patient characteristics (global vs Asian studies). Considering the model's sensitivity to the selected ITC input, these scenarios explore different ITC inputs, which only utilize subsets of the efficacy data.
11		CORRECT vs RECOURSE ITC	
12		CONCUR vs TERRA and Yoshino ITC	
13		CONCUR vs TERRA ITC	
14		CONCUR vs Yoshino ITC	

Scenario category		Scenario name	Description and rationale
15		Assume equal efficacy between regorafenib and trifluridine/tipiracil (HRs of 1)	The model base case uses point estimate HRs. However, the analysis do not support regorafenib or trifluridine/tipiracil being different in terms of efficacy i.e. CrI included 1. We therefore explore a scenario that applies a HR of 1 to OS and PFS.
16	**Costs**	Equal RDI	The model base-case uses different RDI definitions for regorafenib and trifluridine/tipiracil (i.e. pooled RDI from CORRECT and CONCUR and a combination of dose reduction and cycle delay data from TA405), since no comparable RDI metric is reported for both. Although this approach is in line with how RDI was defined in CORRECT and CONCUR, there is some uncertainty in this approach. This scenario therefore applies the pooled RDI of [redacted]% to both regorafenib and trifluridine/tipiracil.
17		Apply RDI to pack count rather than price per pack	As described in Section B.3.5.1.1, there are two different ways of including RDI in the model, either by applying RDI directly on the modelled price per pack, or by applying RDI on the pack count. RDI applied to the price per pack is applied in the base case. This scenario explores an alternative approach, by applying RDI on the pack count, as described in Section B.3.5.1.1,
18		Include post-progression treatment costs	This scenario includes post-progression costs. Since CORRECT and CONCUR did not report individual post progression treatments used, only broad categories, no costs could be calculated from CORRECT and CONCUR directly. This scenario therefore uses the total post progression treatment costs reported for trifluridine/tipiracil reported in TA405, inflated to 2021 (£1,633.18), and applies this as a one-off costs to regorafenib and trifluridine/tipiracil patients upon progression.
19	**Utilities**	Exclude AE disutilities	The base case includes AE disutilities, as the observed equal utilities between trifluridine/tipiracil and regorafenib may not be fully reflective of the difference AE profiles. However, there is a risk of double counting when applying AE disutilities this way. This scenario therefore explores excluding these disutilities.
20	**Model structure**	5-year time horizon	To account for the long-term uncertainty in the model, two scenarios with a shorter time horizon were explored. These scenarios only capture the short-term benefits.
21		No discounting	In line with NICE methods, and scenario with 0% discounting is explored.

Key: AE, adverse event; HR, hazard ratio; ITC, indirect treatment comparison; KM, Kaplan-Meier; mCRC, metastatic colorectal cancer; OS, overall survival; OWSA, one-way sensitivity analysis; PFS, progression-free survival, RDI, relative dose intensity; ToT; time on treatment; VEGF, vascular endothelial growth factor.

A tornado diagram of the 10 most impactful scenarios is shown in Figure 25 (deterministic), with the full set of deterministic and probabilistic scenario results shown below in Table 48 and Table 49 respectively. Overall, [redacted], indicating that regorafenib is likely to be cost-effective, [redacted]

Figure 25: Tornado diagram of the 10 most influential scenarios versus trifluridine/tipiracil (NMB calculated using a WTP of £30,000)

Key: ITC: indirect treatment comparison; NMA: network meta-analysis; NMB, net monetary benefit; RDI: relative dose intensity; T/T, trifluridine/tipiracil.

In line with the OWSA, the three scenarios with the biggest impact on the NMB all related to the ITC input. The ITC scenarios encompassed both positive and negative effects on the NMB. This indicates that the chosen ITC input for the base case, which combines all available efficacy data, provides a good middle ground of the different ITC options available. In addition, out of all explored ITC scenarios, the only scenario with a substantial negative impact on the NMB was the CORRECT vs RECOURSE ITC. However, this scenario is also likely to be the least representative for UK practice, as both CORRECT and RECOURSE showed the most prior anti-VEGF use out of all included studies. Overall, these results illustrate the robustness of the ITC input and emphasize the conservative nature of including all studies in the

ITC, as the scenarios without CORRECT or RECOURSE had a positive effect on the NMB.

Other impactful scenarios relate to how RDI is applied in the model. Both applying it on the pack count, or assuming the regorafenib RDI to trifluridine/tipiracil decrease the NMB to [redacted] and [redacted]. However, these scenarios are less reflective of UK practice than the base case. Differences in regorafenib and trifluridine/tipiracil RDI are reflective of the differences in treatment practice, as clinical experts confirmed that the dose of regorafenib is generally reduced if toxicities develop, whereas a dose of trifluridine/tipiracil is delayed.

Finally, the choice in OS, PFS, and ToT input data source for regorafenib had a moderate impact on the results, with the scenarios that use only CORRECT and only CONCUR for regorafenib OS, PFS, and ToT input both in the top 10 of most impactful scenarios. Of these, the scenario that only uses CONCUR for regorafenib OS, PFS, and ToT data has the biggest impact on the NMB, raising the benefit to [redacted], whereas using CORRECT only lowers the NMB to [redacted]. As discussed in Section B.3.3.1, CONCUR provides efficacy input that is more reflective of UK practice, as there is less prior anti-VEGF use in CONCUR compared to CORRECT. Based on these results, it could therefore be argued that the modelled base case provides conservative cost-effectiveness results, as the pooled data that is used in the base case results in a slightly lower NMB compared to the CONCUR only scenario.

Table 48: Deterministic scenario results versus trifluridine/tipiracil

		Regorafenib		Trifluridine/tipiracil			
#	**Scenario Name**	**Total costs**	**Total QALYs**	**Total costs**	**Total QALYs**	**ICER**	**NMB**
-	**Base case**	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
1	CONCUR efficacy data only	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
2	CORRECT efficacy data only	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
3	Weibull OS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
4	Log-normal OS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
5	Exponential OS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
6	Generalized gamma OS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
7	Gompertz OS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
8	Gamma OS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
9	Parametric PFS and ToT	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
10	NMA without Yoshino	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
11	CORRECT vs RECOURSE ITC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
12	CONCUR vs TERRA and Yoshino ITC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
13	CONCUR vs TERRA ITC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
14	CONCUR vs Yoshino ITC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
15	Assume equal efficacy	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
16	Equal T/T RDI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
17	Apply RDI on pack count	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
18	Include post-progression treatment costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
19	Exclude AE disutilities	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
20	5 year time horizon	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
21	No discounting	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Key: AE, adverse event; ITC, indirect treatment comparison; NMA: network meta-analysis; OS, overall survival; PFS, progression-free survival, RDI, relative dose intensity; [redacted]; ToT, time on treatment; T/T, trifluridine/tipiracil.

[redacted]

Table 49: Probabilistic scenario results versus trifluridine/tipiracil

		Regorafenib		Trifluridine/tipiracil			
#	**Scenario Name**	**Total costs**	**Total QALYs**	**Total costs**	**Total QALYs**	**ICER**	**NMB**
-	**Base case**	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
1	CONCUR efficacy data only	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
2	CORRECT efficacy data only	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
3	Weibull OS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
4	Log-normal OS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
5	Exponential OS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
6	Generalized gamma OS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
7	Gompertz OS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
8	Gamma OS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
9	Parametric PFS and ToT	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
10	NMA without Yoshino	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
11	CORRECT vs RECOURSE ITC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
12	CONCUR vs TERRA and Yoshino ITC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
13	CONCUR vs TERRA ITC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
14	CONCUR vs Yoshino ITC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
15	Assume equal efficacy	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
16	Equal T/T RDI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
17	Apply RDI on pack count	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
18	Include post-progression treatment costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
19	Exclude AE disutilities	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
20	5 year time horizon	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
21	No discounting	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Key: AE, adverse event; ITC, indirect treatment comparison; NMA: network meta-analysis; OS, overall survival; PFS, progression-free survival, RDI, relative dose intensity; [redacted]; ToT, time on treatment; T/T, trifluridine/tipiracil.
[redacted]

B.3.11.3.2. Scenario analyses vs BSC

For the comparison to BSC, only a selection of key scenarios were explored, as BSC is less relevant to the decision problem in this submission. The main uncertainties versus BSC are the efficacy data input for PFS, OS, and ToT (i.e. using data from only CORRECT, only CONCUR, or pooled data), and which OS extrapolation to use. We therefore explored four scenarios: two scenarios using only CORRECT and only CONCUR for regorafenib and BSC PFS, OS, and ToT inputs, and two scenarios using the OS extrapolations with the biggest net impact on the ICER for both arms (log-normal and generalized gamma), as best and worst case scenarios. The probabilistic and deterministic results of these scenarios are shown in Table 50 and Table 51, and discussed in more detail below.

Table 50: Deterministic scenario results versus BSC

		Regorafenib		BSC		ICER
#	**Scenario Name**	**Total costs**	**Total QALYs**	**Total costs**	**Total QALYs**	
-	**Base case**	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
1	CONCUR efficacy data only	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
2	CORRECT efficacy data only	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
3	Log-normal OS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
4	Generalized gamma OS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Key: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; OS, overall survival; QALY, quality-adjusted life year.						

Table 51: Probabilistic scenario results versus BSC

		Regorafenib		BSC		ICER
#	**Scenario Name**	**Total costs**	**Total QALYs**	**Total costs**	**Total QALYs**	
-	**Base case**	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
1	CONCUR efficacy data only	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
2	CORRECT efficacy data only	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
3	Log-normal OS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
4	Generalized gamma OS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Key: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; OS, overall survival; QALY, quality-adjusted life year.						

The scenario that uses the generalized gamma OS input for both regorafenib and BSC has the biggest impact on the outcomes, raising the deterministic ICER vs BSC to [redacted]. However, these results should be interpreted with caution as the OS curve for BSC exceeds the regorafenib OS curve after 40 months, which is not clinically plausible. Although this can be partially 'corrected' by forcing BSC OS to always be lower than regorafenib (which lowers the ICER in this scenario to [redacted]), this is still concerning, as it indicates that the whole BSC curve overestimates OS. In addition, the generalized gamma curve is likely to underestimate regorafenib survival.

Similar to the comparison with trifluridine/tipiracil, the scenario using CONCUR data only is more favourable for regorafenib, whereas the scenario using CORRECT data only is more favourable to BSC. It could therefore again be argued that the modelled base case provides conservative cost-effectiveness estimates versus BSC, as CONCUR may be more reflective of UK practice.

B.3.12. Subgroup analysis

No subgroup analyses were conducted as part of this submission.

B.3.13. Benefits not captured in the QALY calculation

We do not anticipate there are additional benefits associated with regorafenib above and beyond those captured by the QALY.

B.3.14. Validation

B.3.14.1. Validation of cost-effectiveness analysis

The clinical validity of the model and assumptions was validated by UK clinical experts. The validation consisted of various smaller online and offline one-to-one interactions and an advisory board involving nine clinical experts. A total of nine clinical experts participated in this advisory board.

In addition, we conducted an external health economic validation meeting to confirm whether the modelling approach was appropriate for the decision problem. Both the clinical advisory board and the external health economic validation meeting supported the modelling approach.

B.3.14.2. Internal validity and model functionality

To verify the results of the cost-effectiveness model, internal quality control procedures were undertaken by the model developers to ensure that the mathematical calculations were performed correctly and were consistent with the model specifications.

Health economists not involved in the development of the model reviewed the model for coding errors, inconsistencies and the plausibility of inputs and results. The model has also been subjected to a checklist of known modelling errors, and the assumptions have been questioned. This involved checks on the selection and results of different modelling options, calculation spot checks, cross checks against source data and extreme value scenarios to check if the model behaved logically. The validation identified no major issues with the computational accuracy of the model. A number of small inaccuracies were identified and rectified.

B.3.14.3. Comparison of model and trial outcomes

As part of the validation process, model outcomes for regorafenib and BSC were compared to the pooled clinical trial data.

An overlay of the modelled regorafenib and BSC OS compared to the pooled trial data is shown in Figure 26 and Figure 27 below. In addition, the OS and PFS estimates are compared to the pooled OS and PFS from CORRECT and CONCUR at set timepoints in Table 52. The modelled regorafenib results were closely aligned with the clinical trial data. Although there is some variation in OS estimates, at set time points (Table 52), these are all explained by the use of smoothened extrapolated data rather than KM input, as the used OS data accurately follows the pooled KM data overall (Figure 26).

Figure 26: Regorafenib modelled OS compared to pooled trial results

Key: Rego., regorafenib; OS, overall survival; PFS, progression-free survival.

Figure 27: BSC modelled OS compared to pooled trial results

Key: BSC, best supportive care; OS, overall survival; PFS, progression-free survival.

Table 52: Model outcomes compared to published trial results

Outcome	Source	3 months	6 months	9 months	12 months	15 months
Regorafenib: model vs trial outcomes						
OS (log-logistic extrapolation)	Pooled trial data	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Model	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PFS (pooled KM data)	Pooled trial data	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Model	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BSC: model vs trial outcomes						
OS (log-logistic extrapolation)	Pooled trial data	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Model	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PFS (pooled KM data)	Pooled trial data	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Model	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Key: BSC, best supportive care; KM, Kaplan–Meier; N/A, not available; OS, overall survival; PFS, progression-free survival.

B.3.14.4. External validity and comparison with other cost-effectiveness models

As discussed in the previous sections, we aimed to align the assumptions in the de novo cost-effectiveness analysis with past appraisals in order to ensure external validity and consistency in model outcomes. In particular, TA405 served as an important anchor point as it is the key comparator for this appraisal. In addition, the model used in TA405 and its underlying assumptions were thoroughly reviewed by the ERG and NICE as part of the appraisal process.

B.3.15. Interpretation and conclusions of economic evidence

B.3.15.1. Conclusions

The base case analysis shows that regorafenib is a cost-effective option to treat ≥ 3L mCRC patients. Regorafenib has comparable efficacy to trifluridine/tipiracil [redacted]. Regorafenib is cost-effective versus both trifluridine/tipiracil and BSC.

Compared with trifluridine/tipiracil, regorafenib was associated with [redacted], resulting in a deterministic NMB of [redacted] (using a WTP of £30,000 i.e. not applying the highest QALY weighting of the severity modifier). Regorafenib [redacted] in the majority of the probabilistic scenarios, with a probabilistic NMB of [redacted]. In addition, even at the least favourable ITC input of CORRECT vs RECOURSE, which only includes a subset of all efficacy data, and compares the two studies that are the least representative to the UK in terms of anti-VEGF use, regorafenib [redacted] compared with trifluridine/tipiracil. Furthermore, as discussed in Section B.3.4.4, trifluridine/tipiracil is associated with an AE profile typical of chemotherapy meaning it will not be suitable for all patients. There is therefore a large unmet need for additional treatment options for later line mCRC patients, which regorafenib can address. Regorafenib offers a chemotherapy alternative to trifluridine/tipiracil, providing mCRC patients a valuable extra treatment option [redacted].

For the limited number of patients who would otherwise have received BSC, the base case analysis shows that regorafenib would be a cost-effective alternative to BSC. Regorafenib provides a substantial health benefit of [redacted]incremental QALYs at [redacted] incremental costs, resulting in a deterministic ICER of [redacted]. Considering the severity of mCRC, with the average BSC patient experience is proportional QALY shortfall of over 95% compared to the healthy population, and with that a higher WTP threshold of £51,000, regorafenib would be a cost-effective alternative to BSC for ≥ 3L mCRC patients.

B.3.15.2. Strengths and weaknesses

The base case provides a robust assessment of the cost-effectiveness of regorafenib in ≥ 3L mCRC. Several steps were undertaken to increase the reliability of the analysis:

- An extensive body of clinical data was used with 5 studies used to inform the efficacy of regorafenib and trifluridine/tipiracil. Results were validated across a range of probabilistic and deterministic scenario analyses.

- The clinical validity of the model outcomes was extensively explored through input from clinical experts and comparison with published OS and PFS data.

- The internal validity of the model was checked via a systematic technical review process to ensure the accuracy of the model outcomes

Nevertheless, some uncertainties remain due to limitations in the available data. Most notably all patients in the CORRECT trial received prior treatment with bevacizumab, which is not in line with UK clinical practice. This may have affected the observed treatment effect of regorafenib as both clinical experts and the ITC indicated that prior treatment with an anti-VEGF therapy could be a treatment effect modifier. Similarly, the ITC used to inform the efficacy of trifluridine/tipiracil also included CORRECT and RECOURSE, both of which included prior anti-VEGF treatment. However, given the size of these studies, it would not be sound to exclude them from the analysis, as this would drastically reduce the evidence base supporting this analysis. Furthermore, including CORRECT and RECOURSE only increases the conservative nature of the analysis, as prior anti-VEGF treatment is expected to have a negative effect on the observed efficacy of regorafenib. This is also reflected by the scenarios that only used data from CONCUR or excluded CORRECT and RECOURSE from the ITC, which both resulted in a higher NMB versus trifluridine/tipiracil of [redacted] and [redacted] respectively.

There is some uncertainty around the results versus trifluridine/tipiracil, due to limitations in both the available cost and efficacy data for trifluridine/tipiracil. In terms of costs, no single RDI value was reported for trifluridine/tipiracil, so RDI was approximated by combining data on dose reductions and cycle delays. In terms of

efficacy, the model has to rely on an ITC to compare trifluridine/tipiracil with regorafenib, as no clinical study directly compares regorafenib and trifluridine/tipiracil in mCRC. However, both of these uncertainties were explored in scenarios using different RDI and HR assumptions, and regorafenib [redacted] in all of the scenarios. Therefore, overall, the model and its associated analyses shows regorafenib is a cost-effective option for mCRC patients, despite some limitations in the analyses.

B.3.15.3. Generalizability to UK practice

The base case analysis was designed to provide a cost-effectiveness estimate that is as generalizable to UK practice as practically feasible, utilising all of the available data. Model inputs and assumptions were validated by UK clinical experts. Included clinical studies were carefully analysed in terms of trial design and baseline characteristics to ensure the efficacy estimates from these studies are representative of the expected efficacy in the UK.

Nevertheless, due to limited available data, some compromises were made that reduce the generalizability to the UK. Specifically, this relates to the choice of using pooled CORRECT and CONCUR data to inform the efficacy of regorafenib and BSC patients, and to the inclusion of CORRECT and RECOURSE in the ITC, despite all CORRECT and RECOURSE patients receiving prior anti-VEGF treatment. However, as discussed above, we concluded that the added data and robustness CORRECT and RECOURSE provide outweighs any uncertainty that is introduced by including more data from patients who received prior anti-VEGF treatment. In addition, both the ITC and clinical experts indicate that prior anti-VEGF use is likely to limit the observed treatment effect of regorafenib. This is also reflected by the scenarios that exclude CORRECT and RECOURSE from the analysis, which all resulted in better cost-effectiveness outcomes versus both trifluridine/tipiracil and BSC. Therefore, although inclusion of CORRECT and RECOURSE reduces the generalizability to the UK, the risk and uncertainty associated with including these studies in the analysis was considered limited, as including these studies only increases the conservative nature of the cost effectiveness estimates for regorafenib in UK practice.

B.4. References

1. Fleming M, Ravula S, Tatishchev SF and Wang HL. Colorectal carcinoma: Pathologic aspects. *Journal of gastrointestinal oncology*. 2012; 3(3):153-73.
2. Cancer Research UK. Bowel cancer incidence statistics. Available from https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer/incidence#heading-Zero. 2021. Accessed: 22 June 2021.
3. Office of National Statistics. Cancer registration statistics, England: 2017. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/cancerregistrationstatisticsengland/2017. 2021. Accessed: 09 July 2021.
4. Zarour LR, Anand S, Billingsley KG, et al. Colorectal Cancer Liver Metastasis: Evolving Paradigms and Future Directions. *Cell Mol Gastroenterol Hepatol*. 2017; 3(2):163-73.
5. Cancer Research UK. What is bowel cancer. Available from https://www.cancerresearchuk.org/about-cancer/bowel-cancer/about-bowel-cancer. 2021. Accessed: 22 June 2021.
6. Xynos ID, Kavantzas N, Tsaousi S, et al. Factors Influencing Survival in Stage IV Colorectal Cancer: The Influence of DNA Ploidy. *ISRN Gastroenterology*. 2013; 2013:490578.
7. Cancer Research UK (CRUK). Bowel cancer survival statistics. 2012. Available at: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer/survival. 2021. Accessed: 22 June 2021.
8. NICE. TA668 - Encorafenib plus cetuximab for previously treated BRAF V600E mutation-positive metastatic colorectal cancer. 2021.
9. NICE. TA405 - Trifluridine–tipiracil for previously treated metastatic colorectal cancer. 2016.
10. EMA. Stivarga: EPAR. 2013.
11. Nguyen LH, Goel A and Chung DC. Pathways of Colorectal Carcinogenesis. *Gastroenterology*. 2020; 158(2):291-302.
12. Grothey A, Fakih M and Tabernero J. Management of BRAF-mutant metastatic colorectal cancer: a review of treatment options and evidence-based guidelines. *Ann Oncol*. 2021; 32(8):959-67.
13. Cancer Research UK. Symptoms of bowel cancer. Available from https://www.cancerresearchuk.org/about-cancer/bowel-cancer/symptoms. 2021. Accessed: 29 June 2021.
14. Astin M, Griffin T, Neal RD, et al. The diagnostic value of symptoms for colorectal cancer in primary care: a systematic review. *Br J Gen Pract*. 2011; 61(586):e231-43.
15. Cancer Research UK. Bowel cancer risk. Available at https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer/risk-factors. 2021. Accessed: 30 June 2021.
16. Huxley RR, Ansary-Moghaddam A, Clifton P, et al. The impact of dietary and lifestyle risk factors on risk of colorectal cancer: a quantitative overview of the epidemiological evidence. *Int J Cancer*. 2009; 125(1):171-80.
17. Kolligs FT. Diagnostics and Epidemiology of Colorectal Cancer. *Visc Med*. 2016; 32(3):158-64.

18. National Institute of Health and Care Excellence (NICE). NG151. Colorectal cancer. NICE guideline. 2020.
19. Cancer Research UK. Dukes staging system. 2021. Available at: https://www.cancerresearchuk.org/about-cancer/bowel-cancer/stages-types-and-grades/dukes-staging. Accessed: 17 September 2021.
20. American Cancer Society. Colorectal Cancer Stages. 2020. Available at: https://www.cancer.org/cancer/colon-rectal-cancer/detection-diagnosis-staging/staged.html. Accessed: 10 September 2021.
21. Van Cutsem E, Cervantes A, Nordlinger B, et al. Metastatic colorectal cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2014; 25 Suppl 3:iii1-9.
22. National Cancer Registration and Analysis Service. Cancer survival by stage. Available at: http://www.ncin.org.uk/publications/survival_by_stage. 2021. Accessed: 09 July 2021.
23. Adam R, de Gramont A, Figueras J, et al. Managing synchronous liver metastases from colorectal cancer: a multidisciplinary international consensus. *Cancer Treat Rev*. 2015; 41(9):729-41.
24. Scheithauer W, Rosen H, Kornek GV, et al. Randomised comparison of combination chemotherapy plus supportive care with supportive care alone in patients with metastatic colorectal cancer. *BMJ (Clinical research ed)*. 1993; 306(6880):752-5.
25. Grothey A, Van Cutsem E, Sobrero A, et al. Regorafenib monotherapy for previously treated metastatic colorectal cancer (CORRECT): an international, multicentre, randomised, placebo-controlled, phase 3 trial. *Lancet*. 2013; 381(9863):303-12.
26. Li J, Qin S, Xu R, 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. *Lancet Oncol*. 2015; 16(6):619-29.
27. Van Cutsem E, Cervantes A, Adam R, et al. ESMO consensus guidelines for the management of patients with metastatic colorectal cancer. *Ann Oncol*. 2016; 27(8):1386-422.
28. Cancer Research UK. Symptoms of advanced bowel cancer. Available at https://www.cancerresearchuk.org/about-cancer/bowel-cancer/advanced/symptoms-advanced-cancer. 2021. Accessed: 29 June 2021.
29. Valderrama-Treviño AI, Barrera-Mera B, Ceballos-Villalva JC and Montalvo-Javé EE. Hepatic Metastasis from Colorectal Cancer. *Euroasian journal of hepato-gastroenterology*. 2017; 7(2):166-75.
30. Kohn CG, Nguyen E, Miller JD, et al. The Humanistic Burden of Microsatellite Instability-High (MSI-H) Colorectal Cancer (CRC) and Metastatic CRC (MCRC) Patients Treated with Subsequent Line Therapy: A Systematic Review. *Value Health*. 2017; 20(9):PA454.
31. Peng YN, Huang ML and Kao CH. Prevalence of Depression and Anxiety in Colorectal Cancer Patients: A Literature Review. *Int J Environ Res Public Health*. 2019; 16(3).
32. Mosher CE, Secinti E, Kroenke K, et al. Acceptance and commitment therapy for fatigue interference in advanced gastrointestinal cancer and caregiver burden: protocol of a pilot randomized controlled trial. *Pilot Feasibility Stud*. 2021; 7(1):99.

33. National Comprehensive Cancer Network. Colon Cancer (Version 2.2021). Available at: https://www.nccn.org/professionals/physician_gls/pdf/colon.pdf. 2021. Accessed: 29 June 2021.
34. National Institute of Health and Care Excellence (NICE). NICE: managing metastatic colorectal cancer. 2021.
35. Chiorean EG, Nandakumar G, Fadelu T, et al. Treatment of Patients With Late-Stage Colorectal Cancer: ASCO Resource-Stratified Guideline. *JCO Glob Oncol*. 2020; 6:414-38.
36. National Comprehensive Cancer Network. Rectal Cancer (Version 1.2021). Available at: https://www.nccn.org/professionals/physician_gls/pdf/rectal.pdf. 2021. Accessed: 29 June 2021.
37. Battaglin F, Naseem M, Lenz H-J and Salem ME. Microsatellite instability in colorectal cancer: overview of its clinical significance and novel perspectives. *Clin Adv Hematol Oncol*. 2018; 16(11):735-45.
38. Su Y, Ostojic H and Jaeger S. “No problem” health-related quality of life (HRQoL) in patients treated with regorafenib (Reg) for relapsed metastatic colorectal cancer (mCRC). *Journal of Clinical Oncology*. 2021; 39(15_suppl):e15593-e.
39. Battaglin F, Zhang W, Skubala A, et al. Single nucleotide polymorphism (SNP) analysis identifies potential prognostic and predictive biomarker in patients (pts) with metastatic colorectal cancer (mCRC) treated with regorafenib in the phase III CORRECT trial. *Ann Oncol*. 2020; 31:S439.
40. Hofheinz R, Marian M, Bartsch J and Odom D. Effect of regorafenib in delaying definitive deterioration in health-related quality of life across three tumor types. *Ann Oncol*. 2019; 30:AA37.
41. Pericay C and Gallego-Plazas J. Regorafenib as a new standard of care in advanced colorectal cancer. *Colorectal Cancer*. 2014; 3(5):417-23.
42. Pasqualetti G, Schirripa M, Dochy E, et al. Thyroid hormones ratio is a major prognostic marker in advanced metastatic colorectal cancer: Results from the phase III randomised CORRECT trial. *Eur J Cancer*. 2020; 133:66-73.
43. Tabernero J, Lenz HJ, Siena S, et al. Analysis of circulating DNA and protein biomarkers to predict the clinical activity of regorafenib and assess prognosis in patients with metastatic colorectal cancer: A retrospective exploratory analysis of the CORRECT trial. *Lancet Oncol*. 2015; 16(8):937-48.
44. Yoshino T, Komatsu Y, Yamada Y, et al. Randomized phase III trial of regorafenib in metastatic colorectal cancer: Analysis of the CORRECT Japanese and non-Japanese subpopulations. *Invest New Drugs*. 2015; 33(3):740-50.
45. Bayer Healthcare. CORRECT CSR. No. A53306. 19 March 2012. 2012.
46. Bayer Healthcare. CORRECT CSR Addendum (PH-37845. 29 August 2014. 2014.
47. Xu J, Xu RH, Qin S, et al. Regorafenib in Chinese patients with metastatic colorectal cancer: Subgroup analysis of the phase 3 CONCUR trial. *J Gastroenterol Hepatol*. 2020; 35(8):1307-16.
48. Bayer Healthcare. CONCUR CSR. PH-37592. 20 August 2014. 2014.
49. Bayer Healthcare. CONCUR CSR Addendum. PH-39012. 21 March 2017. 2017.
50. Balduzzi S, Rucker G and Schwarzer G. How to perform a meta-analysis with R: a practical tutorial. *Evid Based Ment Health*. 2019; (22):153-60.
51. Mayer RJ, Van Cutsem E, Falcone A, et al. Randomized trial of TAS-102 for refractory metastatic colorectal cancer. *N Engl J Med*. 2015; 372(20):1909-19.

52. Xu J, Kim TW, Shen L, et al. Results of a Randomized, Double-Blind, Placebo-Controlled, Phase III Trial of Trifluridine/Tipiracil (TAS-102) Monotherapy in Asian Patients With Previously Treated Metastatic Colorectal Cancer: The TERRA Study. *J Clin Oncol*. 2018; 36(4):350-8.
53. Yoshino T, Mizunuma N, Yamazaki K, et al. TAS-102 monotherapy for pretreated metastatic colorectal cancer: a double-blind, randomised, placebo-controlled phase 2 trial. *Lancet Oncol*. 2012; 13(10):993-1001.
54. Song F, Loke YK, Walsh T, et al. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. *BMJ (Clinical research ed)*. 2009; 338:b1147.
55. Ducreux M, Petersen LN, Öhler L, et al. Safety and effectiveness of regorafenib in patients with metastatic colorectal cancer in routine clinical practice in the prospective, observational CORRELATE study. *Eur J Cancer*. 2019; 123:146-54.
56. Adenis A, de la Fouchardiere C, Paule B, et al. Survival, safety, and prognostic factors for outcome with Regorafenib in patients with metastatic colorectal cancer refractory to standard therapies: results from a multicenter study (REBECCA) nested within a compassionate use program. *BMC Cancer*. 2016; 16:412.
57. Schulz R, Janssen J, Strauss UP, et al. Clinical efficacy and safety of regorafenib (REG) in the treatment of metastatic colorectal cancer (mCRC) in daily practice in Germany: Final results of the prospective multicentre non-interventional RECORA study. *J Clin Oncol*. 2018; 36(4_suppl):748-.
58. Novakova-Jiresova A, Kopeckova K, Boublikova L, et al. Regorafenib for Metastatic Colorectal Cancer: An Analysis of a Registry-Based Cohort of 555 Patients. *Cancer Manag Res*. 2020; 12:5365-72.
59. Tanaka A, Sadahiro S, Suzuki T, et al. Retrospective study of regorafenib and trifluridinetipiracil efficacy as a thirdline or later chemotherapy regimen for refractory metastatic colorectal cancer. *Oncology Letters*. 2018; 16(5):6589-97.
60. Banzi M, Romagnani A, Berselli A, et al. Retrospective study on real-world use of regorafenib in colorectal cancer in the emiliaromagna region CoreGI01 preliminary results. *Tumori*. 2019; 105(6):20.
61. Sueda T, Sakai D, Kudo T, et al. Efficacy and safety of regorafenib or TAS102 in patients with metastatic colorectal cancer refractory to standard therapies. *Anticancer Res*. 2016; 36(8):4299-306.
62. Huemer F, Piringer G, Schlintl V, et al. Hospitalizations and clinical outcome in metastatic colorectal cancer during regorafenib or TAS102 therapy. *Cancers*. 2020; 12(10):44207.
63. Hirano G, Makiyama A, Makiyama C, et al. Reduced dose of salvage-line regorafenib monotherapy for metastatic colorectal cancer in Japan. *Anticancer Res*. 2015; 35(1):371-8.
64. Zengin G, Özyurt N, Beypinar I, et al. Efficiency and safety of regorafenib in metastatic colorectal cancer mCRC Real life experience from Turkey. *J Clin Oncol*. 2018; 36(15).
65. Hofheinz R-D, Bruix J, Demetri GD, et al. Effect of Regorafenib in Delaying Definitive Deterioration in Health-Related Quality of Life in Patients with Advanced Cancer of Three Different Tumor Types. *Cancer Manag Res*. 2021; 13:5523-33.
66. Grothey A, Ciardiello F and Marshall JL. How to incorporate a chemo-free interval into the management of metastatic colorectal cancer. *Clin Adv Hematol Oncol*. 2020; 18 Suppl 16(10):1-24.

67. Giuliani J, Fiorica F, Ponturo G, et al. Dose-escalation strategy in refractory metastatic colorectal cancer: A change in terms of cost-effectiveness. *J Oncol Pharm Pract*. 2021; ,.
68. Guan X, Li H, Xiong X, et al. Cost-effectiveness analysis of fruquintinib versus regorafenib as the third-line therapy for metastatic colorectal cancer in China. *J Med Econ*. 2021; 24(1):339-44.
69. Scottish Medicines Consortium (SMC). Encorafenib 50mg and 75mg hard capsules (Braftovi®). 2021. Available at: https://www.scottishmedicines.org.uk/medicines-advice/encorafenib-braftovi-full-smc2312/. Accessed: 5 July 2021.
70. Kashiwa M and Matsushita R. Comparative Cost utility Analysis of Regorafenib and TrifluridineTipiracil in The Treatment of Metastatic Colorectal Cancer in Japan. *Clin Ther*. 2020; 42(7):1376-87.
71. Zhang S MA, Li H, Rascati K. Cost-effectiveness analysis of regorafenib used in Chinese patients with metastatic colorectal cancer. *International Society for Pharmacoeconomics and Outcomes Research (ISPOR) Annual*. Virtual Conference: Orlando, FL, USA, 2020.
72. Almadiyeva A, Abeuov A, Absattarova K and Sultanov M. Cost-effectiveness of regorafenib for metastatic colorectal cancer in the context of the healthcare system of Kazakhstan. *Value Health*. 2019; 22:S480-S1.
73. Barzi A, Cho SK, Hay JW, et al. Costeffectiveness (CE) of regorafenib dose optimization schedule in metastatic colorectal cancer. *J Clin Oncol*. 2019; 37.
74. Chu JN, Choi J, Ostvar S, et al. Cost-effectiveness of immune checkpoint inhibitors for microsatellite instability–high/mismatch repair–deficient metastatic colorectal cancer. *Cancer*. 2019; 125(2):278-89.
75. Gourzoulidis G, Maniadakis N, Petrakis D, et al. Economic evaluation of trifluridine and tipiracil hydrochloride in the treatment of metastatic colorectal cancer in Greece. *J Comp Eff Res*. 2019; 8(3):133-42.
76. Li Q. Fruquintinib or regorafenib as the third-line treatments for metastatic colorectal cancer based on CONCUR and FRESCO trials: A cost-effectiveness analysis. *American Society of Clinical Oncology (ASCO)*. Chicago, IL, USA, 2019.
77. Sabater J, Ralph L, Batteson R, et al. Validation of cost-effectiveness of trifluridine/tipiracil versus best supportive care and regorafenib for previously treated metastatic colorectal cancer in the UK using phase IIIb PRECONNECT early access clinical trial data in the real world setting. *J Clin Oncol*. 2019; 37(4_suppl):639-.
78. Yao L, Han J, She L, et al. Cost-effectiveness for metastatic colorectal cancer. *J Clin Oncol*. 2019; 37.
79. Bullement A, Underhill S, Fougeray R and Hatswell AJ. Cost-effectiveness of Trifluridine/tipiracil for Previously Treated Metastatic Colorectal Cancer in England and Wales. *Clin Colorectal Cancer*. 2018; 17(1):e143-e51.
80. Cho SK, Hay JW and Barzi A. Cost-effectiveness Analysis of Regorafenib and TAS-102 in Refractory Metastatic Colorectal Cancer in the United States. *Clin Colorectal Cancer*. 2018; 17(4):e751-e61.
81. Liu J, Zhu S, Sun W, et al. Cost-effectiveness analysis of regorafenib for third-line metastatic colorectal cancer (mCRC) compared to cetuximab plus irinotecan in China. *Value Health*. 2018; 21:S33.
82. J. Almeida, B. Vandewalle, J. Félix, et al. Cost-effectiveness of trifluridine/tipiracil for the treatment of metastatic colorectal cancer in Portugal. *Value Health*. 2017; 20(9):A437.

83. Scottish Medicines Consortium (SMC). trifluridine/tipiracil (as hydrochloride), 15mg/6.14mg and 20mg/8.19mg film-coated tablets (Lonsurf®). 2017. Available at: https://www.scottishmedicines.org.uk/medicines-advice/trifluridinetipiracil-lonsurf-fullsubmission-122117/. Accessed: 22 March 2021.
84. Kimura M, Usami E, Iwai M, et al. Comparison of cost-effectiveness of regorafenib and trifluridine/tipiracil combination tablet for treating advanced and recurrent colorectal cancer. *Mol Clin Oncol*. 2016; 5(5):635-40.
85. C. Deger, F. Telli, M. Gunaldi, et al. The cost-effectiveness of regorafenib in the treatment of patients with metastatic colorectal cancer (MCRC) who have progressed after standard therapies in turkey. *Value Health*. 2015; 18(7):A458.
86. Goldstein DA, Ahmad BB, Chen Q, et al. Cost-effectiveness analysis of regorafenib for metastatic colorectal cancer. *J Clin Oncol*. 2015; 33(32):3727-32.
87. Seal BS, Özer-Stillman I, Whalen J, et al. Cost-effectiveness of regorafenib for pretreated metastatic colorectal cancer patients in the United States. *J Clin Oncol*. 2013; 31(4_suppl):578-.
88. NICE. TA242 - Cetuximab, bevacizumab and panitumumab for the treatment of metastatic colorectal cancer after first-line chemotherapy: Cetuximab (monotherapy or combination chemotherapy), bevacizumab (in combination with non-oxaliplatin chemotherapy) and panitumumab (monotherapy) for the treatment of metastatic colorectal cancer after first-line chemotherapy. 2012.
89. NICE. TA307 - Aflibercept in combination with irinotecan and fluorouracil-based therapy for treating metastatic colorectal cancer that has progressed following prior oxaliplatin-based chemotherapy. 2014.
90. NICE. Guide to the methods of technology appraisal 2013 [Available from: https://www.nice.org.uk/process/pmg9/chapter/foreword]. . 2013.
91. EMA. Trifluridine tipiracil Summary of Product Characteristics (SmPC). 2016.
92. EMA. Regorafenib Summary of Product Characteristics (SmPC). 2013.
93. NICE. Technical support document 14: survival analysis for economic evaluation alongside clinical trials - extrapolation with patient-level data. 2011.
94. Sabater J, Ralph L, Batteson R, et al. Validation of cost-effectiveness of trifluridine/tipiracil versus best supportive care and regorafenib for previously treated metastatic colorectal cancer in the UK using phase IIIb PRECONNECT early access clinical trial data in the real world setting. *Journal of Clinical Oncology*. 2019; 37(4_suppl):639-.
95. Kashiwa M and Matsushita R. Cost-effectiveness of preemptive skin treatment to prevent skin-toxicity caused by panitumumab in third-line therapy for KRAS wild type metastatic colorectal cancer in Japan. *J Pharm Health Care Sci*. 2021; 7(1):35.
96. Xu Y, Hay JW and Barzi A. Impact of drug substitution on cost of care: an example of economic analysis of cetuximab versus panitumumab. *Cost Eff Resour Alloc*. 2018; 16(1):30.
97. Graham CN, Maglinte GA, Schwartzberg LS, et al. Economic Analysis of Panitumumab Compared With Cetuximab in Patients With Wild-type KRAS Metastatic Colorectal Cancer That Progressed After Standard Chemotherapy. *Clin Ther*. 2016; 38(6):1376-91.
98. Koukakis R, Gatta F, Hechmati G and Siena S. Skin toxicity and quality of life during treatment with panitumumab for RAS wild-type metastatic colorectal carcinoma: results from three randomised clinical trials. *Qual Life Res*. 2016; 25(10):2645-56.

99. Chang J, Ngai C, Odom D, et al. Effects of regorafenib (REG) therapy on health-related quality of life (HRQoL) in patients with metastatic colorectal cancer (mCRC) in the phase III CONCUR trial. *J Clin Oncol*. 2015; 33(15_suppl):3560-.
100. Hoyle M, Peters J, Crathorne L, et al. Cost-Effectiveness of Cetuximab, Cetuximab Plus Irinotecan, and Panitumumab for Third and Further Lines of Treatment for KRAS Wild-Type Patients with Metastatic Colorectal Cancer. *Value Health*. 2013; 16(2):288-96.
101. Siena S, Grothey A, Sobrero A, et al. Effects of regorafenib therapy on health-related quality of life in patients with metastatic colorectal cancer in the phase III CORRECT study. *Eur J Cancer*. 2013; 49:S482.
102. Blank PR, Moch H, Szucs TD and Schwenkglenks M. KRAS and BRAF mutation analysis in metastatic colorectal cancer: a cost-effectiveness analysis from a Swiss perspective. *Clin Cancer Res*. 2011; 17(19):6338-46.
103. Shiroiwa T, Motoo Y and Tsutani K. Cost-effectiveness analysis of KRAS testing and cetuximab as last-line therapy for colorectal cancer. *Mol Diagn Ther*. 2010; 14(6):375-84.
104. Doyle S, Lloyd A and Walker M. Health state utility scores in advanced non-small cell lung cancer. *Lung Cancer*. 2008; 62(3):374-80.
105. Lloyd A, Nafees B, Narewska J, et al. Health state utilities for metastatic breast cancer. *Br J Cancer*. 2006; 95(6):683-90.
106. Nafees B, Stafford M, Gavriel S, et al. Health state utilities for non small cell lung cancer. *Health Qual Life Outcomes*. 2008; 6:84.
107. Royal Pharmaceutical Society of Great Britain. BNF Online [Available from: https://www.nice.org.uk/bnf-uk-only. Accessed 23 June 2021]. . 2021.
108. Sacco JJ, Botten J, Macbeth F, et al. The average body surface area of adult cancer patients in the UK: a multicentre retrospective study. *PloS one*. 2010; 5(1):e8933-e.
109. Unit PSSR. Unit Costs of Health and Social Care 2015 - Available via: https://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-of-health-and-social-care-2021/ [Accessed: 8 April 2022]. 2022.
110. NHS improvement. NHS Reference Costs 2019-2020. Available via: https://www.england.nhs.uk/publication/2019-20-national-cost-collection-data-publication/ [accessed: 10 June 2021]. . 2021.
111. NICE. TA259 - Abiraterone for castration-resistant metastatic prostate cancer previously treated with a docetaxel-containing regimen. 2012.
112. NICE Decision Support Unit. The risk and costs of febrile neutropenia in patients with non small cell lung cancer treated with docetaxel. 2007.
113. Round J, Jones L and Morris S. Estimating the cost of caring for people with cancer at the end of life: A modelling study. *Palliative medicine*. 2015; 29(10):899-907.
114. Schneider P, McNamara S, Love-Koh J, et al. QALY shortfall calculator. Available at: https://r4scharr.shinyapps.io/shortfall/. Accessed: 8 April 2022.
115. Statistics OoN. National Life Tables, England, 1980-1982 to 2017-2019 - Avialable via: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/nationallifetablesenglandreferencetables. 2021.
116. (NatCen) NCfSR. Health Survey for England, 2017 and 2018. UK Data Service (2021) - Available via: http://doi.org/10.5255/UKDA-SN-8488-2. 2021.
117. Hernandez Alava M, Pudney, S., and Wailoo, A. Estimating the relationship between EQ-5D-5L and EQ-5D-3L: results from an English Population Study. *Policy*

Research Unit in Economic Evaluation of Health and Care Interventions Universities of Sheffield and York. 2020; Report 063.

B.5. Appendices

Appendix C: Summary of product characteristics

Appendix D: Identification, selection and synthesis of clinical evidence

Appendix E: Subgroup analysis

Appendix F: Adverse reactions

Appendix G: Published cost-effectiveness studies

Appendix H: Health-related quality-of-life studies

Appendix I: Cost and healthcare resource identification, measurement and valuation

Appendix J: Clinical outcomes and disaggregated results from the model

Appendix K: Price details of treatments included in the submission

Appendix L: Checklist of confidential information

Appendix M: Additional clinical trial data

Appendix N: Progression-free survival and time on treatment extrapolations

Appendix O: Assessment of proportional hazards

Appendix P: References

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Regorafenib for treating metastatic colorectal cancer [ID4002]

Clarification questions

June 2022

File name	Version	Contains confidential information	Date
Regorafenib_ID4002_ClarificationQuestions_7Jul22_(complete)_CIC Redacted	1.0	Yes	07/07/2022

Notes for company

Highlighting in the template

Square brackets and grey highlighting are used in this template to indicate text that should be replaced with your own text or deleted. These are set up as form fields, so to replace the prompt text in [grey highlighting] with your own text, click anywhere within the highlighted text and type. Your text will overwrite the highlighted section.

To delete grey highlighted text, click anywhere within the text and press DELETE.

Section A: Clarification on effectiveness data

Literature searches

A 1. Priority question: The structure for the clinical effectiveness Embase/ Medline search is as follows:

CRC
+
Non-response/Relapsed/refractory/2-4 line/etc
+
Named drugs (regorafenib/ tipiracil/ trifluridine /nivolumab/ ipilimumab/ encorafenib)
+
RCTs/Obs
+
(Limits: No Animals/SRs/letters etc)

The company's submission (CS) reported that the searches were designed from *"a multi-country perspective and therefore included comparators that are not relevant to this appraisal"* (Appendix D, section B.3.1.). It was further stated that results relating to these comparators (nivolumab/ ipilimumab/ encorafenib) were excluded from this appraisal. The searches also included a facet for terms relating to non-response etc refining the number of retrieved studies. However, the ERG is concerned the inclusion of this facet may have been overly restrictive and adversely affected the recall of results, especially

those relating to studies of tipiracil plus trifluridine. A more cautious approach may have been to remove the 3 additional named drugs, which would lower the hits retrieved allowing for the removal of the non-response facet:

CRC
+
Named drugs (regorafenib/ tipiracil/ trifluridine)
+
RCTs/Obs
+
(Limits: No Animals/SRs/letters etc)

Please rerun this search and check that no additional relevant studies have been missed.

The searches were updated for Embase®, MEDLINE® (both via Embase.com) and CENTRAL (via the Cochrane Library) by removing the facet for relapsed/refractory disease and only focussing on regorafenib and tipiracil/trifluridine interventions. The updated searches resulted in 1725 records from Embase and 458 records from CENTRAL and resulted in 193 and 31 records, respectively. After removing the duplicates across these two databases there were 195 records that were unique to this search i.e. were not 'returned' by the submitted search. These 195 records were reviewed against the same inclusion/exclusion criteria as the original search and 174 records were excluded at the title/abstract stage. Full-text review was performed for 21 publications. Overall, the updated search strategy search located the same five studies that were identified in the original search with no additional RCTs relevant to the appraisal being located. The updated search strategies and the list of excluded publications are presented in tables below:

Table A1.1: Embase.com search for clinical effectiveness

S. No.	Query	Search hits	Facet
1.	'colon cancer'/exp OR 'colon carcinoma'/exp OR 'colorectal cancer'/exp OR 'rectum cancer'/exp OR 'rectum adenoma'/exp	351,304	Disease
2.	((cancer* OR carcinoma* OR adenoma* OR adenocarci* OR tumor* OR tumour* OR neoplasm* OR malignan*) NEAR/4 (colorectal OR 'colo-rectal' OR 'colonrectal' OR 'colon rectal' OR 'colon-rectal' OR colon OR rect* OR pararec* OR bowel OR sigmoid)):ab,ti,kw	377,766	
3.	crc:ab,ti,kw OR mcrc:ab,ti,kw OR 'm-crc':ab,ti,kw	74,214	
4.	#1 OR #2 OR #3	466,025	
5.	'randomization'/exp OR 'controlled clinical trial'/exp OR 'controlled clinical trial (topic)'/exp OR 'placebo effect'/exp OR 'placebo'/exp OR 'clinical trial'/exp OR 'clinical trial (topic)'/exp OR 'control group'/exp OR 'randomized controlled trial'/exp OR 'randomized controlled trial (topic)'/exp OR 'controlled clinical trial':ab,ti,kw OR 'controlled clinical trials':ab,ti,kw OR 'randomised controlled trial':ab,ti,kw OR 'randomized controlled trial':ab,ti,kw OR 'randomised controlled trials':ab,ti,kw OR 'randomized controlled trials':ab,ti,kw OR 'randomi?ed controlled trial*' OR rct:ab,ti,kw OR ((random* NEAR/2 (alloca* OR assign* OR distribut* OR group*)):ab,ti,kw) OR (((single OR double OR triple OR treble) NEAR/2 (blind* OR mask*)):ab,ti,kw) OR placebo*:ab,ti,kw OR 'single blind procedure'/exp OR 'crossover procedure'/exp OR 'double blind procedure'/exp OR 'triple blind procedure'/exp OR 'controlled study'/exp	10,423,453	Study design
6.	'clinical article'/exp OR 'clinical trial'/exp OR 'case control study'/exp OR 'longitudinal study'/exp OR 'family study'/exp OR 'retrospective study'/exp OR 'prospective study'/exp OR 'cohort analysis'/exp OR ((cohort NEAR/1 (study OR studies OR trial*)):ab,ti,kw) OR (('case control' NEAR/1 (study OR studies OR trial*)):ab,ti,kw) OR (('follow up' NEAR/1 (study OR studies OR trial*)):ab,ti,kw) OR ((observational NEAR/1 (study OR studies OR trial*)):ab,ti,kw) OR (('cross sectional' NEAR/1 (study OR studies OR trial*)):ab,ti,kw) OR 'comparative study'/exp OR 'follow up'/exp OR retrospectiv*:ab,ti,kw OR 'medical record review'/exp OR 'intervention study'/exp OR 'major clinical study'/exp OR 'open study'/exp OR registr*:ab,ti,kw OR (((hospital OR medical OR electronic) NEAR/2 (record OR chart)):ab,ti,kw) OR 'community trial'/exp OR 'cross-sectional study'/exp OR 'non-random*':ab,ti,kw OR 'non random*':ab,ti,kw OR 'single arm*':ab,ti,kw OR 'observational study'/exp OR 'observational method'/exp OR 'cancer registry'/exp OR 'real world*':ab,ti,kw OR 'real-world*':ab,ti,kw OR 'real life*':ab,ti,kw OR 'real-life*':ab,ti,kw OR claim*:ab,ti,kw OR 'compassionate	11,559,265	

S. No.	Query	Search hits	Facet
	use'/exp OR 'compassionate use':ab,ti,kw OR 'expanded access*':ab,ti,kw		
7.	#5 OR #6	16,586,736	
8.	'case study':it OR 'case report':it OR 'abstract report':it OR editorial:it OR letter:it OR comment:it OR note:it OR 'case report'/exp OR 'case study'/exp OR 'editorial'/exp	5,470,570	
9.	'animal'/exp NOT ('animal'/exp AND 'human'/exp)	5,800,937	
10	(review:it OR 'literature review':it) NOT ('meta-analysis':it OR 'meta-analysis as topic'/mj OR 'systematic review':ti OR 'systematic literature review':ti OR 'meta-analysis':ab,ti OR 'meta analysis':ab,ti,kw)	2,806,077	
11	#8 OR #9 OR #10	13,765,718	
12	#7 NOT #11	12,381,414	
13	'second line chemotherapy' OR 'third line chemotherapy' OR 'fourth line chemotherapy' OR 'second-line' OR 'second line' OR 'third-line' OR 'third line' OR 'fourth-line' OR 'fourth line' OR '2nd line' OR '2nd-line' OR '3rd line' OR '3rd-line' OR '4th line' OR '4th-line' OR 'second or later*' OR 'third or later*' OR 'fourth or later*' OR 'second- or later*' OR 'third- or later*' OR 'fourth- or later*' OR 'second and later*' OR 'third and later*' OR 'fourth and later*' OR 'second- and later*' OR 'third- and later*' OR 'fourth- and later*' OR '2 l' OR '3 l' OR '2l' OR '3l' OR '2-l' OR '3-l' OR '2 line*' OR '2-line*' OR '3 line*' OR '3-line*' OR 'previously treated' OR 'previously-treated' OR 'pre-treated' OR 'pretreated' OR 'failed' OR 'prior treatment' OR 'prior-treatment' OR 'prior treated' OR 'prior-treated' OR 'prior therap*' OR 'prior-therap*' OR 'second-' OR 'third-' OR 'fourth-' OR 'prior' OR 'failure' OR relaps* OR refrac* OR resist* OR recur* OR progress* OR 'cancer recurrence'/exp OR 'relapse'/exp OR 'therapy resistance'/exp OR 'tumor recurrence'/exp OR 'recurrent disease'/exp OR 'patient history of therapy'/exp OR 'cancer resistance'/exp OR 'drug resistance'/exp OR 'treatment failure'/exp OR 'salvage therapy'/exp OR reocur* OR 're occur' OR 're ocur' OR recrudescen* OR ((post* NEAR/4 (chemo* OR line OR therap* OR treat*)):ab,ti,kw) OR (((pre* OR prio* OR prev* OR post* OR heav* OR late* OR receiv* OR subseque*) NEAR/4 (treat* OR therap* OR regim* OR progress* OR fail* OR relaps* OR resis* OR refract* OR line* OR chemo* OR target*)):ab,ti,kw) OR (((lack* OR inadequa*) NEAR/2 respon*):ab,ti,kw) OR nonrespon*:ab,ti,kw OR 'non respon*':ab,ti,kw OR unrespon*:ab,ti,kw OR 'un-respon*':ab,ti,kw OR 'no respon*':ab,ti,kw OR 'not respon*':ab,ti,kw	10,401,568	Relapsed/ refractory
14	'regorafenib'/exp OR 'tipiracil plus trifluridine'/exp OR 'tipiracil'/exp OR 'trifluridine'/exp	7,934	Intervention
15	regorafenib:ab,ti,kw OR 'bay 73 4506':ab,ti,kw OR 'bay 73-4506':ab,ti,kw OR 'bay 734506':ab,ti,kw OR 'bay73 4506':ab,ti,kw OR 'bay73-4506':ab,ti,kw OR bay734506:ab,ti,kw OR stivarga:ab,ti,kw	3,145	

S. No.	Query	Search hits	Facet
16	((tipiracil NEAR/3 trifluridine):ab,ti,kw) OR lonsurf:ab,ti,kw OR 'tas 102':ab,ti,kw OR tas102:ab,ti,kw OR tipiracil:ab,ti,kw OR trifluridine:ab,ti,kw OR thriherpine:ab,ti,kw OR triflumann:ab,ti,kw OR 'trifluor thymidine':ab,ti,kw OR 'trifluoro thymidine':ab,ti,kw OR trifluorodeoxythymidine:ab,ti,kw OR trifluorothymidine:ab,ti,kw OR trifuridine:ab,ti,kw OR triherpin:ab,ti,kw OR triherpine:ab,ti,kw OR viromidin:ab,ti,kw OR virophta:ab,ti,kw OR viroptic:ab,ti,kw	1,519	
17	#14 OR #15 OR #16	8,827	
18	#4 AND #12 AND #13 AND #17	1,532	Disease AND Study design AND Relapsed/ refractory AND Intervention
19	#4 AND #12 AND #17	1,725	Disease AND Study design AND Intervention
20	#19 NOT #18	**193**	**Unique hits without relapsed/ refractory facet**

Table A1.2: The Cochrane Library (CENTRAL) search for clinical effectiveness

S. No.	Query	Search hits	Facet
1.	MeSH descriptor: [Colonic Neoplasms] explode all trees	1,894	Disease
2.	MeSH descriptor: [Colorectal Neoplasms] explode all trees	9,160	
3.	MeSH descriptor: [Rectal Neoplasms] explode all trees	2,005	
4.	((cancer* OR carcinoma* OR adenoma* OR adenocarci* OR tumor* OR tumour* OR neoplasm* OR malignan*) NEAR/4 (colorectal OR colo-rectal OR colonrectal OR colon rectal OR colon-rectal OR colon OR rect* OR pararec* OR bowel OR sigmoid)):ab,ti,kw	24,234	
5.	crc:ab,ti,kw OR mcrc:ab,ti,kw OR "m-crc":ab,ti,kw	4,888	
6.	#1 OR #2 OR #3 OR #4 OR #5	25,241	
7.	(“second line chemotherapy” OR “third line chemotherapy” OR “fourth line chemotherapy” OR second-line OR “second line” OR third-line OR “third line” OR fourth-line OR “fourth line” OR “2nd line”):ab,ti,kw	7,118	Relapsed/ refractory
8.	(2nd-line OR “3rd line” OR 3rd-line OR “4th line” OR 4th-line):ab,ti,kw	907	
9.	(“second or later*” OR “third or later*” OR “fourth or later*” OR “second and later*” OR “third and later*” OR “fourth and later*”):ab,ti,kw	91	
10	(“previously treated” OR previously-treated OR pre-treated OR pretreated OR failed OR “prior treatment” OR prior-treatment OR “prior treated” OR prior-treated)	42,574	
11	(“prior therap*” OR prior-therap* OR prior OR failure OR relaps* OR refrac* OR resist* OR recur* OR progress* OR “cancer recurrence” OR relapse OR “therapy resistance” OR “tumor recurrence” OR “recurrent disease” OR “patient history of therapy” OR “cancer resistance” OR “drug resistance” OR “treatment failure” OR “salvage therapy” OR reocur* OR “re occur” OR “re ocur” OR recrudescen*):ab,ti,kw	438,308	
12	((post* NEAR/4 (chemo* OR line OR therap* OR treat*)):ab,ti,kw) OR (((pre* OR prio* OR prev* OR post* OR heav* OR late* OR receiv* OR subseque*) NEAR/4 (treat* OR therap* OR regim* OR progress* OR fail* OR relaps* OR resis* OR refract* OR line* OR chemo* OR target*)):ab,ti,kw) OR (((lack* OR inadequa*) NEAR/2 respon*):ab,ti,kw) OR nonrespon*:ab,ti,kw OR “non respon*”:ab,ti,kw OR unrespon*:ab,ti,kw OR un-respon*:ab,ti,kw OR “no respon*”:ab,ti,kw OR “not respon*”:ab,ti,kw	385,087	
13	#7 OR #8 OR #9 OR #10 OR #11 OR #12	674,529	
14	MeSH descriptor: [Trifluridine] explode all trees	87	Intervention
15	(regorafenib OR “tipiracil plus trifluridine” OR tipiracil OR encorafenib):ab,ti,kw	734	

S. No.	Query	Search hits	Facet
16	regorafenib:ab,ti,kw OR “bay 73*4506”:ab,ti,kw OR “bay 734506”:ab,ti,kw OR “bay73*4506”:ab,ti,kw OR bay734506:ab,ti,kw OR stivarga:ab,ti,kw	581	
17	((tipiracil NEAR/3 trifluridine):ab,ti,kw) OR lonsurf:ab,ti,kw OR “tas 102”:ab,ti,kw OR tas102:ab,ti,kw OR tipiracil:ab,ti,kw OR trifluridine:ab,ti,kw OR thriherpine:ab,ti,kw OR triflumann:ab,ti,kw OR “trifluor thymidine”:ab,ti,kw OR “trifluoro thymidine”:ab,ti,kw OR trifluorodeoxythymidine:ab,ti,kw OR trifluorothymidine:ab,ti,kw OR trifuridine:ab,ti,kw OR triherpin:ab,ti,kw OR triherpine:ab,ti,kw OR viromidin:ab,ti,kw OR virophta:ab,ti,kw OR viroptic:ab,ti,kw	341	
18	#14 OR #15 OR #16 OR #17	844	
19	#6 AND #13 AND #18, In trials	427	Disease AND Relapsed/ refractory AND Intervention
20	#6 AND #18, In trials	458	Disease AND Intervention
21	**#20 NOT #19**	31	**Unique hits without relapsed/ refractory facet**

Table A1.3: List of excluded studies

S. No.	Authors	Title	Source	Exclusion Reason
1	Nakashima M., Takeuchi M., Kawakami K.	Effectiveness and Safety of Regorafenib vs. Trifluridine/Tipiracil in Unresectable Colorectal Cancer: A Retrospective Cohort Study	Clinical Colorectal Cancer (2020) 19:4 (e208-e225). Date of Publication: 1 Dec 2020	Disease stage
2	Charette N., Vandeputte C., Ameye L., et al	Prognostic value of adipose tissue and muscle mass in advanced colorectal cancer: A post hoc analysis of two non-randomized phase II trials	BMC Cancer (2019) 19:1 Article Number: 134. Date of Publication: 12 Feb 2019	Study design
3	Nakashima M., Ide K., Kawakami K.	Comparison of Standard Initial Dose and Reduced Initial Dose Regorafenib for Colorectal Cancer Patients: A Retrospective Cohort Study	Targeted Oncology (2019) 14:3 (295-306). Date of Publication: 1 Jun 2019	Line of therapy
4	Roberto M., Marchetti P., Arrivi G., et al	The treatment paradigm of right-sided metastatic colon cancer: harboring BRAF mutation makes the difference	International Journal of Colorectal Disease (2020) 35:8 (1513-1527). Date of Publication: 1 Aug 2020	Study design
5	Schröder C., Lawrance M., Li C., Lenain C., et al	Building external control arms from patient-level electronic health record data to replicate the randomized IMblaze370 control arm in metastatic colorectal cancer	JCO Clinical Cancer Informatics (2021) 5 (450-458). Date of Publication: 2021	Study design
6	Hasegawa H., Taniguchi H., Nakamura Y., et al	FMS-like tyrosine kinase 3 (FLT3) amplification in patients with metastatic colorectal cancer	Cancer Science (2021) 112:1 (314-322). Date of Publication: 1 Jan 2021	Study design
7	Tilby M., Escola C., Ellison C., Narramneni L., et al	Trifluridine-tipiracil for the treatment of metastatic colorectal cancer patients: UK multicentre real-world experience	Annals of Oncology (2019) 30 Supplement 4 (iv27). Date of Publication: 1 Jul 2019	Line of therapy
8	Jiang F.-E., Zhang H.-J., Yu C.-Y., Liu A.-N.	Efficacy and safety of regorafenib or fruquintinib plus camrelizumab in patients with microsatellite stable and/or proficient mismatch repair metastatic colorectal cancer: an observational pilot study	Neoplasma (2021) 68:4 (861-866). Date of Publication: 2021	Intervention
9	Rauthan A., Patil P., Somashekhar S.P., Zaveri S.	Real world experience with regorafenib in dose escalation schedule in metastatic colorectal cancer in Indian patients	Annals of Oncology (2018) 29 Supplement 9 (ix37). Date of Publication: 1 Nov 2018	Not relevant to the assessment
10	Hofheinz R.-D., Bruix J., Demetri G.D., et al	Effect of regorafenib in delaying definitive deterioration in health-related quality of life in patients with advanced cancer of three different tumor types	Cancer Management and Research (2021) 13 (5523-5533). Date of Publication: 2021	Study design

11	Sánchez-Camacho A., Herrero Rivera D., Carrasco I., et al	Real World Data (RWD) of patients with chemorefractory metastatic colorectal cancer treated with trifluridine/tipiracil (TAS-102): clinical benefit from a Spanish single institution	Annals of Oncology (2019) 30 Supplement 4 (iv62). Date of Publication: 1 Jul 2019	Identified in original search
12	Yeh K.-H., Yang T.-S., Hsu T.-C., et al	Real-world evidence of the safety and effectiveness of regorafenib in patients with metastatic colorectal cancer (mCRC) from Taiwan: A subgroup analysis from the prospective, observational CORRELATE study	Annals of Oncology (2018) 29 Supplement 9 (ix35-ix36). Date of Publication: 1 Nov 2018	Identified in original search
13	Bazarbashi M.S., Elshenawy M.A., Kandil M.S., et al	Efficacy of regorafenib in metastatic colorectal cancer: A multi-institutional retrospective study	Annals of Oncology (2018) 29 Supplement 9 (ix36). Date of Publication: 1 Nov 2018	Line of therapy
14	Argiles G., Margalef N.M., Valladares-Ayerbes M., et al	Results of REARRANGE trial: A randomized phase 2 study comparing different dosing approaches for regorafenib (REG) during the first cycle of treatment in patients (pts) with metastatic colorectal cancer (mCRC)	Annals of Oncology (2019) 30 Supplement 4 (iv135). Date of Publication: 1 Jul 2019	Identified in original search
15	Ducreux M., O'Connor J., Dochy E., et al	Regorafenib dose escalations in the prospective, observational CORRELATE study in patients with metastatic colorectal cancer	Annals of Oncology (2019) 30 Supplement 4 (iv119). Date of Publication: 1 Jul 2019	Identified in original search
16	Rodriguez-Salas N., Segura M.S., Jimenez-Gordo A., et al	Retrospective analysis of clinical factors associated with a greater benefit with Trifluridine and Tipiracil in metastasic colorectal cancer	Annals of Oncology (2018) 29 Supplement 5 (v75). Date of Publication: 1 Jun 2018	Line of therapy
17	Kotaka M., Ogata M., Ogata T., et al	Trifluridine/tipiracil vs regorafenib as salvage-line treatment in patients with metastatic colorectal cancer: A multicenter retrospective study	Annals of Oncology (2018) 29 Supplement 5 (v65). Date of Publication: 1 Jun 2018	Line of therapy
18	Ducreux M., Petersen L., Öhler L., et al	Safety and effectiveness of regorafenib in patients with metastatic colorectal cancer (mCRC) in routine clinical practice: Final analysis from the prospective, observational CORRELATE study	Annals of Oncology (2018) 29 Supplement 5 (v104). Date of Publication: 1 Jun 2018	Identified in original search
19	Hara H., Fukuoka S., Takahashi N., et al	Regorafenib plus nivolumab in patients with advanced colorectal or gastric cancer: an open-label, dose-finding, and dose-expansion phase 1b trial (REGONIVO, EPOC1603)	Annals of Oncology (2019) 30 Supplement 4 (iv124). Date of Publication: 1 Jul 2019	Disease stage
20	Jakobsen A., Andersen R.F.,	Early ctDNA response to chemotherapy. A potential	European Journal of Cancer (2021) 149 (128-	Line of therapy

	Hansen T.F., et al	surrogate marker for overall survival	133). Date of Publication: 1 May 2021	
21	T Yoshino, H Uetake, N Fujita, et al	TAS-102 Safety in Metastatic Colorectal Cancer: results From the First Postmarketing Surveillance Study	Clin Colorectal Cancer	Line of therapy

A 2.In all search sections (Appendices D, G, H, I). the CS states that the MEDLINE search was conducted via Embase.com. Please can you confirm that by this you mean a search of the Embase database conducted on the understanding that it now contains all records from Medline or was this a separate multi file search of both resources using the same strategy?

The Embase.com platform was used to run a multi-faceted search strategy to identify records from both Embase and MEDLINE databases. MEDLINE In-process records were identified by a separate multi-faceted search conducted on Pubmed.com

A 3. Please provide the date span for all databases searched including Embase/MEDLINE.

For the clinical and utility SLR, all databases were searched from the date of database inception to the date of running the update searches on 22nd February 2022. The economic evaluations and cost & resource use SLR were restricted from 2010 to 22nd February 2022.

A 4.Appendix D section B.3.1.1. reports searches of the following conference proceedings:

a) American Society of Clinical Oncology (ASCO)

b) European Society for Medical Oncology (ESMO)

c) Digestive Disease Week (DDW)

Please can you provide full search strategies and hits per year/conference.

The table overleaf provides the search strategies and hits obtained per year per conference for manual screening of conference proceedings:

Table A4.1: Search tracker for conference proceedings

Conference name and year	Search terms	Number of hits
ASCO 2021	Manually searched	NA
ASCO 2020	colorectal	621
	Bowel cancer	81
	colon cancer	308
	Rectal cancer	578
	CRC	505
ASCO 2019	colorectal	628
	Bowel cancer	86
	colon cancer	307
	Rectal cancer	559
	CRC	531
ESMO 2021	Colorectal	220
	Bowel cancer	2
	colon cancer	2
	Rectal cancer	23
	CRC	156
ESMO 2020	Searched manually	NA
ESMO 2019	Searched manually	NA
DDW 2021	Searched manually	NA
DDW 2020	colorectal	365
	Bowel cancer	8
	colon cancer	81
	Rectal cancer	298
	CRC	225
DDW 2019	colorectal	218
	Bowel cancer	0
	colon cancer	29
	Rectal cancer	159
	CRC	12
ISPOR annual 2021	colorectal	24
	Bowel cancer	2
	colon cancer	5
	Rectal cancer	3
	CRC	8
ISPOR annual 2020	colorectal	34

Conference name and year	Search terms	Number of hits
	Bowel cancer	4
	colon cancer	6
	Rectal cancer	3
	CRC	11
ISPOR annual 2019	colorectal	31
	Bowel cancer	2
	colon cancer	8
	Rectal cancer	5
	CRC	9
ISPOR European 2021	colorectal	13
	Bowel cancer	0
	colon cancer	2
	Rectal cancer	0
	CRC	5
ISPOR European 2020	colorectal	49
	Bowel cancer	8
	colon cancer	20
	Rectal cancer	8
	CRC	18
ISPOR European 2019	colorectal	26
	Bowel cancer	5
	colon cancer	8
	Rectal cancer	5
	CRC	9

Key: ASCO, American Society of Clinical Oncology; CRU, cost and resource use; DDW, Digestive Disease Week; EM, economic modelling; ESMO, European Society for Medical Oncology; ISPOR: International Society for Pharmacoeconomics and Outcomes Research

A 5. Appendix G also reports additional searches for ASCO, ESMO, DDW and ISPOR as well as the HTA agencies NICE, SMC and AWMSG. Whilst keywords were provided for the HTA searches, no strategies or keywords were provided for the conference searches and the PRISMA flow chart contained only a condensed number of hits for all HTAs, conferences etc. This is also the case in the PRISMA flow charts in Appendices H & I. Please provide full search details including hits per resource for each section.

Please refer to the table provided in response Question A4 above for the search strategies and hits obtained per year per conference for manual screening of conference proceedings. Similar representation of search terms and number of hits retrieved for HTA searches is provided in the table below:

Table A5.1: Search tracker for HTA websites

HTA	Search terms	Number of hits
NICE	colorectal	153
	Bowel cancer	41
	colon cancer	12
	Rectal cancer	18
	CRC	7
The Scottish Medicines Consortium (SMC)	colorectal	21
	Bowel cancer	8
	colon cancer	6
	Rectal cancer	0
	CRC	3
All Wales Medicines Strategy Group (AWMSG)	Colorectal	13
	Bowel cancer	179
	colon cancer	173
	Rectal cancer	169
	CRC	0

Key: AWMSG, All Wales Medicines Strategy Group; CRU: cost and resource use; EM: economic modelling, NICE, National Institute for Health and Care Excellence; SMC, Scottish Medicines Consortium

A 6. Whilst the 2022 update search strategies conducted for the clinical effectiveness searches have been provided, these are missing for the economics searches (see Appendices G, H & I). Please provide full strategies, including hits per line, for all missing update searches.

Apologies for missing the updated economic searches previously. Please see below the updated search strategies for the economic reviews.

A 6.1 Economic evaluations

Table A6.1: Embase.com searches

S. No.	Query	De novo searches ran on 05 March 2021	Update searches ran on 22 February 2022
1.	'colon cancer'/exp OR 'colon carcinoma'/exp OR 'colorectal cancer'/exp OR 'rectum cancer'/exp OR 'rectum adenoma'/exp	322, 420	343,507
2.	((cancer* OR carcinoma* OR adenoma* OR adenocarci* OR tumor* OR tumour* OR neoplasm* OR malignan*) NEAR/4 (colorectal OR 'colo-rectal' OR 'colonrectal' OR 'colon rectal' OR 'colon-rectal' OR colon OR rect* OR pararec* OR bowel OR sigmoid)):ab,ti,kw	347,731	370,078
3.	crc:ab,ti,kw OR mcrc:ab,ti,kw OR 'm-crc':ab,ti,kw	64,569	71,622
4.	'second line chemotherapy' OR 'third line chemotherapy' OR 'fourth line chemotherapy' OR 'second-line' OR 'second line' OR 'third-line' OR 'third line' OR 'fourth-line' OR 'fourth line' OR '2nd line' OR '2nd-line' OR '3rd line' OR '3rd-line' OR '4th line' OR '4th-line' OR 'second or later*' OR 'third or later*' OR 'fourth or later*' OR 'second- or later*' OR 'third- or later*' OR 'fourth- or later*' OR 'second and later*' OR 'third and later*' OR 'fourth and later*' OR 'second- and later*' OR 'third- and later*' OR 'fourth- and later*' OR '2 l' OR '3 l' OR '2l' OR '3l' OR '2-l' OR '3-l' OR '2 line*' OR '2-line*' OR '3 line*' OR '3-line*' OR 'previously treated' OR 'previously-treated' OR 'pre-treated' OR 'pretreated' OR 'failed' OR 'prior treatment' OR 'prior-treatment' OR 'prior treated' OR 'prior-treated' OR 'prior therap*' OR 'prior-therap*' OR 'second-' OR 'third-' OR 'fourth-' OR 'prior' OR 'failure' OR relaps* OR refrac* OR resist* OR recur* OR progress* OR 'cancer recurrence'/exp OR 'relapse'/exp OR 'therapy resistance'/exp OR 'tumor recurrence'/exp OR 'recurrent disease'/exp OR 'patient history of therapy'/exp OR 'cancer resistance'/exp OR 'drug resistance'/exp OR	9,857,906	10,179,691

S. No.	Query	De novo searches ran on 05 March 2021	Update searches ran on 22 February 2022
	'treatment failure'/exp OR 'salvage therapy'/exp OR reocur* OR 're occur' OR 're ocur' OR recrudescen* OR ((post* NEAR/4 (chemo* OR line OR therap* OR treat*)):ab,ti,kw) OR (((pre* OR prio* OR prev* OR post* OR heav* OR late* OR receiv* OR subseque*) NEAR/4 (treat* OR therap* OR regim* OR progress* OR fail* OR relaps* OR resis* OR refract* OR line* OR chemo* OR target*)):ab,ti,kw) OR (((lack* OR inadequa*) NEAR/2 respon*):ab,ti,kw) OR nonrespon*:ab,ti,kw OR 'non respon*':ab,ti,kw OR unrespon*:ab,ti,kw OR 'un-respon*':ab,ti,kw OR 'no respon*':ab,ti,kw OR 'not respon*':ab,ti,kw		
5.	'case study':it OR 'case report':it OR 'abstract report':it OR editorial:it OR letter:it OR comment:it OR note:it OR 'case report'/exp OR 'case study'/exp OR 'editorial'/exp	5,144,485	5,389,710
6.	'animal'/exp NOT ('animal'/exp AND 'human'/exp)	5,579,126	5,738,872
7.	(review:it OR 'literature review':it) NOT ('meta-analysis':it OR 'meta-analysis as topic'/mj OR 'systematic review':ti OR 'systematic literature review':ti OR 'meta-analysis':ab,ti OR 'meta analysis':ab,ti,kw)	2,614,371	2,757,565
8.	#5 OR #6 OR #7	13,041,372	13,578,152
9.	#1 OR #2 OR #3	427,252	455,927
10.	'decision theory'/exp OR 'cost effectiveness analysis'/exp OR 'economic evaluation'/exp OR 'cost utility analysis'/exp OR 'cost benefit analysis'/exp OR 'quality adjusted life year'/exp OR 'decision tree'/exp OR 'monte carlo method'/exp OR 'hidden markov model'/exp OR 'sensitivity analysis'/exp OR ((cost NEXT/1 estimate*):ab,ti,kw) OR ((cost NEXT/1 variable*):ab,ti,kw) OR ((unit NEXT/1 cost*):ab,ti,kw) OR economic*:ab,ti,kw OR pharmacoeconomic*:ab,ti,kw OR markov*:ab,ti,kw OR ((decision NEXT/2 tree*):ab,ti,kw) OR ((decision NEXT/2 analy*):ab,ti,kw) OR ((monte NEXT/1 carlo):ab,ti,kw) OR (((incremental OR qaly OR 'quality adjusted life years') NEAR/3 cost):ab,ti,kw) OR ((cost NEAR/3 (effect* OR utility* OR benefit OR conseq* OR minimi* OR increment* OR qaly* OR ly* OR 'quality adjusted life year*' OR 'life year*')):ab,ti,kw) OR icer:ab,ti,kw OR qaly:ab,ti,kw OR 'quality adjusted life year*':ab,ti,kw OR 'life year*':ab,ti,kw OR (((markov* OR simulat* OR	1,456,241	1,577,310

S. No.	Query	De novo searches ran on 05 March 2021	Update searches ran on 22 February 2022
	decisio* OR analy* OR 'area under curve' OR partition* OR survival* OR economic* OR transitio* OR state* OR discrete* OR individual* OR cohort*) NEAR/3 model*):ab,ti,kw) OR 'economic model'/exp OR 'markov chain'/exp OR 'simulation'/exp		
11.	'regorafenib'/exp OR 'tipiracil plus trifluridine'/exp OR 'tipiracil'/exp OR 'trifluridine'/exp OR 'nivolumab'/exp OR 'ipilimumab'/exp OR 'encorafenib'/exp	34,445	41,898
12.	regorafenib:ab,ti,kw OR 'bay 73 4506':ab,ti,kw OR 'bay 73-4506':ab,ti,kw OR 'bay 734506':ab,ti,kw OR 'bay73 4506':ab,ti,kw OR 'bay73-4506':ab,ti,kw OR bay734506:ab,ti,kw OR stivarga:ab,ti,kw	2,651	3,015
13.	((tipiracil NEAR/3 trifluridine):ab,ti,kw) OR lonsurf:ab,ti,kw OR 'tas 102':ab,ti,kw OR tas102:ab,ti,kw OR tipiracil:ab,ti,kw OR trifluridine:ab,ti,kw OR thriherpine:ab,ti,kw OR triflumann:ab,ti,kw OR 'trifluor thymidine':ab,ti,kw OR 'trifluoro thymidine':ab,ti,kw OR trifluorodeoxythymidine:ab,ti,kw OR trifluorothymidine:ab,ti,kw OR trifuridine:ab,ti,kw OR triherpin:ab,ti,kw OR triherpine:ab,ti,kw OR viromidin:ab,ti,kw OR virophta:ab,ti,kw OR viroptic:ab,ti,kw	1,350	1,482
14.	nivolumab:ab,ti OR 'bms 936558':ab,ti OR bms936558:ab,ti OR 'cmab 819':ab,ti OR cmab819:ab,ti OR 'mdx 1106':ab,ti OR mdx1106:ab,ti OR 'ono 4538':ab,ti OR ono4538:ab,ti OR opdivo:ab,ti	12,393	15,221
15.	ipilimumab:ab,ti,kw OR 'bms 734016':ab,ti,kw OR bms734016:ab,ti,kw OR 'mdx 010':ab,ti,kw OR 'mdx 101':ab,ti,kw OR mdx010:ab,ti,kw OR mdx101:ab,ti,kw OR strentarga:ab,ti,kw OR yervoy:ab,ti,kw	7,915	9,165
16.	encorafenib:ab,ti,kw OR 'nvp lgx 818':ab,ti,kw OR 'nvp lgx 818 nxa':ab,ti,kw OR 'nvp lgx818':ab,ti,kw OR 'nvp lgx818 nxa':ab,ti,kw OR braftovi:ab,ti,kw OR 'lgx 818':ab,ti,kw OR lgx818:ab,ti,kw	302	394
17.	#11 OR #12 OR #13 OR #14 OR #15 OR #16	35,657	43,363
18.	#4 AND #9 AND #10 AND #17	179	221
19.	#18 NOT #8	118	148
20.	#18 NOT #8 AND [2010-2020]/py	**114**	-
21.	**#18 NOT #8 AND [01-03-2021]/sd**	**-**	**36**

Table A6.2: Medline In-process: PubMed.com searches

S. No.	Query	De novo searches ran on 19 March 2021	Update searches ran on 22 February 2022
1.	((colorectal neoplasms[MeSH Terms]) OR (colonic neoplasms[MeSH Terms])) OR (rectal neoplasms[MeSH Terms])	207,249	221,340
2.	("rectum adenoma"[Title/Abstract]) OR ("colon carcinoma"[Title/Abstract])	10,772	10,993
3.	(cancer*[Title/Abstract] OR carcinoma*[Title/Abstract] OR adenoma*[Title/Abstract] OR adenocarci*[Title/Abstract] OR tumor*[Title/Abstract] OR tumour*[Title/Abstract] OR neoplasm*[Title/Abstract] OR malignan*[Title/Abstract]) AND (colorectal[Title/Abstract] OR "colo-rectal"[Title/Abstract] OR "colonrectal"[Title/Abstract] OR "colon rectal"[Title/Abstract] OR "colon-rectal"[Title/Abstract] OR colon[Title/Abstract] OR rect*[Title/Abstract] OR pararec*[Title/Abstract] OR bowel[Title/Abstract] OR sigmoid[Title/Abstract])	293,031	312,117
4.	(publisher[sb] NOT pubstatusnihms NOT pubstatuspmcsd NOT pmcbook) OR (pubstatusaheadofprint) OR (inprocess[sb])	372,882	497,769
5.	((Trifluridine[MeSH Terms]) OR (Nivolumab[MeSH Terms])) OR (ipilimumab[MeSH Terms])	4,901	6,073
6.	((((regorafenib[Title/Abstract] OR "bay 73 4506"[Title/Abstract] OR "bay 73-4506"[Title/Abstract] OR "bay 734506"[Title/Abstract] OR "bay73 4506"[Title/Abstract] OR "bay73-4506"[Title/Abstract] OR bay734506[Title/Abstract] OR stivarga[Title/Abstract]) OR (((tipiracil[Title/Abstract] AND trifluridine))[Title/Abstract] OR lonsurf[Title/Abstract] OR "tas 102"[Title/Abstract] OR tas102[Title/Abstract] OR tipiracil[Title/Abstract] OR trifluridine[Title/Abstract] OR thriherpine[Title/Abstract] OR triflumann[Title/Abstract] OR "trifluor thymidine"[Title/Abstract] OR "trifluoro thymidine"[Title/Abstract] OR trifluorodeoxythymidine[Title/Abstract] OR trifluorothymidine[Title/Abstract] OR trifuridine[Title/Abstract] OR triherpin[Title/Abstract] OR triherpine[Title/Abstract] OR viromidin[Title/Abstract] OR virophta[Title/Abstract] OR viroptic[Title/Abstract])) OR (nivolumab:ab,ti[Title/Abstract] OR "bms 936558":ab,ti[Title/Abstract] OR bms936558:ab,ti[Title/Abstract] OR "cmab 819":ab,ti[Title/Abstract] OR cmab819:ab,ti[Title/Abstract] OR "mdx 1106":ab,ti[Title/Abstract] OR	6,681	7,636

S. No.	Query	De novo searches ran on 19 March 2021	Update searches ran on 22 February 2022
	mdx1106:ab,ti[Title/Abstract] OR "ono 4538":ab,ti[Title/Abstract] OR ono4538:ab,ti[Title/Abstract] OR opdivo:ab,ti[Title/Abstract])) OR (ipilimumab[Title/Abstract] OR "bms 734016"[Title/Abstract] OR bms734016[Title/Abstract] OR "mdx 010"[Title/Abstract] OR "mdx 101"[Title/Abstract] OR mdx010[Title/Abstract] OR mdx101[Title/Abstract] OR strentarga[Title/Abstract] OR yervoy[Title/Abstract])) OR (encorafenib[Title/Abstract] OR "nvp lgx 818"[Title/Abstract] OR "nvp lgx 818 nxa"[Title/Abstract] OR "nvp lgx818"[Title/Abstract] OR "nvp lgx818 nxa"[Title/Abstract] OR braftovi[Title/Abstract] OR "lgx 818"[Title/Abstract] OR lgx818[Title/Abstract])		
7.	#5 OR #6	9,535	11,239
8.	#1 OR #2 OR #3	337,393	358,148
9.	#4 AND #7 AND #8	**28**	-
10.	**#4 AND #7 AND #8 Filters: from 2010/1/1 - 2022/2/22**	-	**39**

Table A6.3: EconLit: Ebsco.com searches

S. No.	Query	Search Options	De novo searches ran on 19 March 2021	Update searches ran on 22 February 2022
S1	(“colon cancer” OR “colon carcinoma” OR “colorectal cancer” OR “rectum cancer” OR “rectum adenoma”) OR ((cancer OR carcinoma OR adenoma OR adenocarcinoma OR tumor OR tumour OR neoplasm OR malignant OR malignancy) AND(colorectal OR “colo-rectal” OR “colonrectal” OR “colon rectal” OR“ colon-rectal” OR colon OR rectum OR rectal OR pararectal OR bowel OR sigmoid)) AND (crc OR mcrc OR “m-crc”)	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	**114**	**-**
S1	(“colon cancer” OR “colon carcinoma” OR “colorectal cancer” OR “rectum cancer” OR “rectum adenoma”) OR ((cancer OR carcinoma OR adenoma OR adenocarcinoma OR tumor OR tumour OR neoplasm OR malignant OR malignancy) AND(colorectal OR “colo-rectal” OR “colonrectal” OR “colon rectal” OR“ colon-rectal” OR colon OR rectum OR rectal OR pararectal OR bowel OR sigmoid)) AND (crc OR mcrc OR “m-crc”)	**Expanders** - Also search within the full text of the articles **Search modes** - Find all my search terms **Limiters** - Date Published: 20210101-20221231	**-**	6

Table A6.4: NHSEED and HTAD: CRD York.com searches

S. No.	Query	De novo searches ran on 07 April 2021	Update searches ran on 22 February 2022
1.	MeSH DESCRIPTOR Colorectal Neoplasms EXPLODE ALL TREES IN NHSEED,HTA	673	673
2.	MeSH DESCRIPTOR Rectal Neoplasms EXPLODE ALL TREES IN NHSEED,HTA	70	70
3.	MeSH DESCRIPTOR Colonic Neoplasms EXPLODE ALL TREES IN NHSEED,HTA	92	92
4.	(cancer* OR carcinoma* OR adenoma* OR adenocarci* OR tumor* OR tumour* OR neoplasm* OR malignan*) IN NHSEED, HTA	6,654	6,654
5.	(colorectal OR colo-rectal OR colonrectal OR 'colon rectal' OR colon-rectal OR colon OR rect* OR pararec* OR bowel OR sigmoid) IN NHSEED, HTA	1,429	1,429
6.	#4 AND #5	998	998
7.	(crc OR mcrc OR m-crc):TI IN NHSEED, HTA	9	9
8.	#1 OR #2 OR #3 OR #6 OR #7	1,027	1,027
9.	(second line chemotherapy OR third line chemotherapy OR fourth line chemotherapy OR second-line OR second line OR third-line OR third line OR fourth-line OR fourth line OR 2nd line OR 2nd-line OR 3rd line OR 3rd-line OR 4th line OR 4th-line OR second or later OR third or later OR fourth or later OR second and later OR third and later OR fourth and later OR 2 l OR 3 l OR 2l OR 3l OR 2-l OR 3-l OR 2 line* OR 2-line OR 3 line* OR 3-line* OR previously treated OR previously-treated OR pre-treated OR pretreated OR failed OR prior treatment OR prior-treatment OR prior treated OR prior-treated OR prior therap* OR prior-therap* OR second- OR third- OR fourth- OR prior OR failure OR relaps* OR refrac* OR resist* OR recur* OR progress* OR cancer recurrence OR relapse OR therapy resistance OR tumor recurrence OR recurrent disease OR patient history of therapy OR cancer resistance OR drug resistance OR treatment failure OR salvage therapy OR reocur* OR re occur OR re ocur OR recrudescen*) IN NHSEED, HTA	659	659
10.	#8 AND #9	**34**	-
11.	(#8 AND #9) IN NHSEED, HTA WHERE LPD FROM 01/03/2021 TO 23/02/2022	-	**0**

A 6.2 Cost and resource use

Table A6.5: Embase.com searches

S. No.	Query	De novo searches ran on 05 March 2021	Update searches ran on 22 February 2022
1.	'colon cancer'/exp OR 'colon carcinoma'/exp OR 'colorectal cancer'/exp OR 'rectum cancer'/exp OR 'rectum adenoma'/exp	322,420	343,507
2.	((cancer* OR carcinoma* OR adenoma* OR adenocarci* OR tumor* OR tumour* OR neoplasm* OR malignan*) NEAR/4 (colorectal OR 'colo-rectal' OR 'colonrectal' OR 'colon rectal' OR 'colon-rectal' OR colon OR rect* OR pararec* OR bowel OR sigmoid)):ab,ti,kw	347,731	370,078
3.	crc:ab,ti,kw OR mcrc:ab,ti,kw OR 'm-crc':ab,ti,kw	64,569	71,622
4.	'second line chemotherapy' OR 'third line chemotherapy' OR 'fourth line chemotherapy' OR 'second-line' OR 'second line' OR 'third-line' OR 'third line' OR 'fourth-line' OR 'fourth line' OR '2nd line' OR '2nd-line' OR '3rd line' OR '3rd-line' OR '4th line' OR '4th-line' OR 'second or later*' OR 'third or later*' OR 'fourth or later*' OR 'second- or later*' OR 'third- or later*' OR 'fourth- or later*' OR 'second and later*' OR 'third and later*' OR 'fourth and later*' OR 'second- and later*' OR 'third- and later*' OR 'fourth- and later*' OR '2 l' OR '3 l' OR '2l' OR '3l' OR '2-l' OR '3-l' OR '2 line*' OR '2-line*' OR '3 line*' OR '3-line*' OR 'previously treated' OR 'previously-treated' OR 'pre-treated' OR 'pretreated' OR 'failed' OR 'prior treatment' OR 'prior-treatment' OR 'prior treated' OR 'prior-treated' OR 'prior therap*' OR 'prior-therap*' OR 'second-' OR 'third-' OR 'fourth-' OR 'prior' OR 'failure' OR relaps* OR refrac* OR resist* OR recur* OR progress* OR 'cancer recurrence'/exp OR 'relapse'/exp OR 'therapy resistance'/exp OR 'tumor recurrence'/exp OR 'recurrent disease'/exp OR 'patient history of therapy'/exp OR 'cancer resistance'/exp OR 'drug resistance'/exp OR 'treatment failure'/exp OR 'salvage therapy'/exp OR reocur* OR 're occur' OR 're ocur' OR recrudescen* OR ((post* NEAR/4 (chemo* OR line OR therap* OR treat*)):ab,ti,kw) OR (((pre* OR prio* OR prev* OR post* OR heav* OR late* OR receiv* OR subseque*) NEAR/4 (treat* OR therap* OR regim* OR progress* OR fail* OR relaps* OR resis* OR refract* OR line* OR chemo* OR target*)):ab,ti,kw) OR (((lack* OR inadequa*) NEAR/2 respon*):ab,ti,kw) OR nonrespon*:ab,ti,kw OR 'non respon*':ab,ti,kw OR unrespon*:ab,ti,kw OR 'un-respon*':ab,ti,kw OR 'no respon*':ab,ti,kw OR 'not respon*':ab,ti,kw	9,857,906	10,179,691
5.	'case study':it OR 'case report':it OR 'abstract report':it OR editorial:it OR letter:it OR comment:it OR note:it OR 'case report'/exp OR 'case study'/exp OR 'editorial'/exp	5,144,485	5,389,710
6.	'animal'/exp NOT ('animal'/exp AND 'human'/exp)	5,579,126	5,738,872
7.	(review:it OR 'literature review':it) NOT ('meta-analysis':it OR 'meta-analysis as topic'/mj OR 'systematic review':ti OR	2,614,371	2,757,565

S. No.	Query	De novo searches ran on 05 March 2021	Update searches ran on 22 February 2022
	'systematic literature review':ti OR 'meta-analysis':ab,ti OR 'meta analysis':ab,ti,kw)		
8.	#5 OR #6 OR #7	13,041,372	13,578,152
9.	#1 OR #2 OR #3	427,252	455,927
10.	'pharmacoeconomics'/exp OR 'health insurance'/exp OR 'cost control'/exp OR 'health care cost'/exp OR 'drug cost'/exp OR 'hospital cost'/exp OR 'cost of illness'/exp OR 'health care utilization'/exp OR 'resource management'/exp OR 'resource allocation'/exp OR ((healthcare NEXT/1 cost*):ab,ti,kw) OR ((unit NEXT/1 cost*):ab,ti,kw) OR price*:ab,ti,kw OR pricing:ab,ti,kw OR ((resource* NEXT/2 allocat*):ab,ti,kw) OR ((health*care NEXT/1 (utilisation OR utilization)):ab,ti,kw) OR (('health care' NEXT/1 (utilisation OR utilization)):ab,ti,kw) OR ((resource NEXT/1 (utilisation OR utilization OR use)):ab,ti,kw) OR ((cost* NEAR/3 (treat* OR therap*)):ab,ti,kw) OR (((total OR direct OR indirect OR medical OR drug OR administration OR laborat* OR diagnos* OR productivity OR illness OR transport* OR societ* OR 'out of pocket*') NEAR/2 (cost OR costs)):ab,ti,kw) OR 'hospitalization cost'/exp OR 'length of stay'/exp OR 'economic aspect'/mj OR 'socioeconomics'/mj OR 'financial management'/mj OR 'health care financing'/mj OR 'fee'/exp OR 'budget'/exp OR economic:ab,ti,kw OR economics:ab,ti,kw OR 'absenteeism'/exp OR 'presenteeism'/exp OR 'medical leave'/exp OR 'productivity'/exp OR (((sick* OR illness OR disab*) NEAR/3 leave*):ab,ti,kw) OR ((work* NEAR/3 (absence OR absent OR impair* OR disab*)):ab,ti,kw) OR productivity:ab,ti,kw OR ((burden NEAR/2 (disease* OR illness)):ab,ti,kw) OR 'caregiver burden'/exp OR 'caregiver support'/exp OR carer*:ab,ti,kw OR caregiver*:ab,ti,kw OR 'care giver':ab,ti,kw OR 'care-giver':ab,ti,kw OR 'care givers':ab,ti,kw OR 'care-givers':ab,ti,kw	1,619,270	1,740,581
11.	'united kingdom'/syn OR 'united kingdom'/exp OR 'united kingdom':ab,ti,kw OR 'uk':ab,ti,kw OR 'great britain'/syn OR 'great britain'/exp OR 'great britain':ab,ti,kw OR england:ab,ti,kw OR wales:ab,ti,kw OR scotland:ab,ti,kw OR ireland:ab,ti,kw OR pound*:ab,ti,kw OR gbp*:ab,ti,kw OR scottish*:ab,ti,kw OR irish*:ab,ti,kw OR british*:ab,ti,kw OR britain*:ab,ti,kw OR £ OR albion:ab,ti,kw OR blighty:ab,ti,kw OR sterling:ab,ti,kw OR britann*:ab,ti,kw OR u.k.:ab,ti,kw OR u.k:ab,ti,kw OR limeyland:ab,ti,kw OR 'u. k':ab,ti,kw OR 'u k':ab,ti,kw OR 'u. k.':ab,ti,kw OR bath:ab,ti,kw OR birmingham:ab,ti,kw OR bradford:ab,ti,kw OR brighton:ab,ti,kw OR bristol:ab,ti,kw OR cambridge:ab,ti,kw OR canterbury:ab,ti,kw OR carlisle:ab,ti,kw OR chester:ab,ti,kw OR chichester:ab,ti,kw OR coventry:ab,ti,kw OR derby:ab,ti,kw OR durham:ab,ti,kw OR ely:ab,ti,kw OR exeter:ab,ti,kw OR gloucester:ab,ti,kw OR hereford:ab,ti,kw OR hull:ab,ti,kw OR lancaster:ab,ti,kw OR leeds:ab,ti,kw OR leicester:ab,ti,kw OR lichfield:ab,ti,kw OR lincoln:ab,ti,kw OR liverpool:ab,ti,kw OR london:ab,ti,kw OR manchester:ab,ti,kw OR 'newcastle upon tyne':ab,ti,kw OR norwich:ab,ti,kw OR nottingham:ab,ti,kw OR	9,323,126	9,814,546

S. No.	Query	De novo searches ran on 05 March 2021	Update searches ran on 22 February 2022
	oxford:ab,ti,kw OR peterborough:ab,ti,kw OR plymouth:ab,ti,kw OR portsmouth:ab,ti,kw OR preston:ab,ti,kw OR ripon:ab,ti,kw OR salford:ab,ti,kw OR salisbury:ab,ti,kw OR sheffield:ab,ti,kw OR southampton:ab,ti,kw OR albans:ab,ti,kw OR 'stoke-on-trent':ab,ti,kw OR sunderland:ab,ti,kw OR truro:ab,ti,kw OR wakefield:ab,ti,kw OR wells:ab,ti,kw OR westminster:ab,ti,kw OR winchester:ab,ti,kw OR wolverhampton:ab,ti,kw OR worcester:ab,ti,kw OR york:ab,ti,kw OR bangor:ab,ti,kw OR cardiff:ab,ti,kw OR newport:ab,ti,kw OR 'st davids':ab,ti,kw OR swansea:ab,ti,kw OR aberdeen:ab,ti,kw OR dundee:ab,ti,kw OR edinburgh:ab,ti,kw OR glasgow:ab,ti,kw OR inverness:ab,ti,kw OR stirling:ab,ti,kw OR armagh:ab,ti,kw OR belfast:ab,ti,kw OR londonderry:ab,ti,kw OR lisburn:ab,ti,kw OR newry:ab,ti,kw		
12.	#4 AND #9 AND #10 AND #11	2,839	3,023
13.	#12 NOT #8	2,092	2,234
14.	#12 NOT #8 AND [2010-2020]/py	**1,666**	-
15.	#12 NOT #8 AND [01-03-2021]/sd	-	**242**

Table A6.6: Medline In-process: PubMed.com searches

Sr. No.	Query	De novo searches ran on 19 March 2021	Update searches ran on 22 February 2022
1.	((colorectal neoplasms[MeSH Terms]) OR (colonic neoplasms[MeSH Terms])) OR (rectal neoplasms[MeSH Terms])	207,249	221,340
2.	("rectum adenoma"[Title/Abstract]) OR ("colon carcinoma"[Title/Abstract])	10,772	10,993
3.	(cancer*[Title/Abstract] OR carcinoma*[Title/Abstract] OR adenoma*[Title/Abstract] OR adenocarci*[Title/Abstract] OR tumor*[Title/Abstract] OR tumour*[Title/Abstract] OR neoplasm*[Title/Abstract] OR malignan*[Title/Abstract]) AND (colorectal[Title/Abstract] OR "colo-rectal"[Title/Abstract] OR "colonrectal"[Title/Abstract] OR "colon rectal"[Title/Abstract] OR "colon-rectal"[Title/Abstract] OR colon[Title/Abstract] OR rect*[Title/Abstract] OR pararec*[Title/Abstract] OR bowel[Title/Abstract] OR sigmoid[Title/Abstract])	293,031	312,117
4.	#1 OR #2 OR #3	337,393	358,148
5.	(Economics[Mesh] OR "costs and cost analysis"[Mesh] OR "Cost allocation"[Mesh] OR "Cost-benefit analysis"[Mesh] OR "Cost control"[Mesh] OR "Cost savings"[Mesh] OR "Cost of illness"[Mesh] OR "Cost sharing"[Mesh] OR "deductibles and coinsurance"[Mesh] OR "Medical savings accounts"[Mesh]	1,180,345	1,258,554

Sr. No.	Query	De novo searches ran on 19 March 2021	Update searches ran on 22 February 2022
	OR "Health care costs"[Mesh] OR "Direct service costs"[Mesh] OR " Drug costs"[Mesh] OR "Employer health costs"[Mesh] OR "Hospital costs"[Mesh] OR "Health expenditures"[Mesh] OR "Capital expenditures"[Mesh] OR "Value of life"[Mesh] OR "economics, hospital"[Mesh] OR "economics, medical"[Mesh] OR "Economics, nursing"[Mesh] OR "Economics, pharmaceutical"[Mesh] OR "fees and charges"[Mesh] OR "budgets"[Mesh] OR fiscal [ti] OR funding[ti] OR financial[ti] OR finance[ti] OR economic* OR pharmacoeconomic* OR price* OR pricing[ti]) OR ("utilization"[Subheading] OR "health care use"[tiab] OR "healthcare use"[tiab] OR "health service use"[tiab] OR "health services use"[tiab] OR "health care utilisation"[tiab] OR "healthcare utilisation"[tiab] OR "healthcare utilization"[tiab] OR "health care utilization"[tiab] OR "health resource utilization"[tiab] OR "health resource utilisation"[tiab] OR "health service utilisation"[tiab] OR "health service utilization"[tiab] OR "health services utilisation"[tiab] OR "health services utilization"[tiab] OR "resource use"[tiab] OR "medical leave"[tiab] OR "work disability"[tiab] OR "Absenteeism"[Mesh] OR absenteeism[tiab] OR "disability absence"[tiab] OR "illness day"[tiab] OR "Retirement"[Mesh] OR retirement[tiab] OR "sick day"[tiab] OR "Sick Leave"[Mesh] OR "work absence"[tiab] OR "work day loss"[tiab] OR "work incapacity"[tiab] OR "work loss"[tiab] OR "work time loss"[tiab] OR "Workers' Compensation"[Mesh] OR "workmans' compensation"[tiab] OR "workman's compensation"[tiab] OR "workers' compensation" [tiab] OR "worker's compensation"[tiab] OR "productivity loss"[tiab] OR "work impairment"[tiab] OR "sickness absence"[tiab] OR "lost days"[tiab] OR productivity[tiab])		
6.	(publisher[sb] NOT pubstatusnihms NOT pubstatuspmcsd NOT pmcbook) OR (pubstatusaheadofprint) OR (inprocess[sb])	372,882	497,769
7.	#4 AND #5 AND #6	**93**	**-**
8.	#4 AND #5 AND #6 Filters: from 2010/1/1 - 2022/2/23	**-**	**136**

Table A6.7: EconLit: Ebsco.com searches

S. No.	Query	Search Options	De novo searches ran on 19 March 2021	Update searches ran on 22 February 2022
S1	(“colon cancer” OR “colon carcinoma” OR “colorectal cancer” OR “rectum cancer” OR “rectum adenoma”) OR ((cancer OR carcinoma OR adenoma OR adenocarcinoma OR tumor OR tumour OR neoplasm OR malignant OR malignancy) AND(colorectal OR “colo-rectal” OR “colonrectal” OR “colon rectal” OR“ colon-rectal” OR colon OR rectum OR rectal OR pararectal OR bowel OR sigmoid)) AND (crc OR mcrc OR “m-crc”)	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	**114**	-
S1	(“colon cancer” OR “colon carcinoma” OR “colorectal cancer” OR “rectum cancer” OR “rectum adenoma”) OR ((cancer OR carcinoma OR adenoma OR adenocarcinoma OR tumor OR tumour OR neoplasm OR malignant OR malignancy) AND(colorectal OR “colo-rectal” OR “colonrectal” OR “colon rectal” OR“ colon-rectal” OR colon OR rectum OR rectal OR pararectal OR bowel OR sigmoid)) AND (crc OR mcrc OR “m-crc”)	**Expanders** - Also search within the full text of the articles **Search modes** - Find all my search terms **Limiters** - Date Published: 20210101-20221231	-	**6**

Table A6.8: NHSEED and HTAD: CRD York.com searches

S. No.	Query	De novo searches ran on 07 April 2021	Update searches ran on 22 February 2022
12	MeSH DESCRIPTOR Colorectal Neoplasms EXPLODE ALL TREES IN NHSEED,HTA	673	673
13	MeSH DESCRIPTOR Rectal Neoplasms EXPLODE ALL TREES IN NHSEED,HTA	70	70
14	MeSH DESCRIPTOR Colonic Neoplasms EXPLODE ALL TREES IN NHSEED,HTA	92	92
15	(cancer* OR carcinoma* OR adenoma* OR adenocarci* OR tumor* OR tumour* OR neoplasm* OR malignan*) IN NHSEED, HTA	6,654	6,654
16	(colorectal OR colo-rectal OR colonrectal OR 'colon rectal' OR colon-rectal OR colon OR rect* OR pararec* OR bowel OR sigmoid) IN NHSEED, HTA	1,429	1,429
17	#4 AND #5	998	998
18	(crc OR mcrc OR m-crc):TI IN NHSEED, HTA	9	9

19	#1 OR #2 OR #3 OR #6 OR #7	1,027	1,027
20	(second line chemotherapy OR third line chemotherapy OR fourth line chemotherapy OR second-line OR second line OR third-line OR third line OR fourth-line OR fourth line OR 2nd line OR 2nd-line OR 3rd line OR 3rd-line OR 4th line OR 4th-line OR second or later OR third or later OR fourth or later OR second and later OR third and later OR fourth and later OR 2 l OR 3 l OR 2l OR 3l OR 2-l OR 3-l OR 2 line* OR 2-line OR 3 line* OR 3-line* OR previously treated OR previously-treated OR pre-treated OR pretreated OR failed OR prior treatment OR prior-treatment OR prior treated OR prior-treated OR prior therap* OR prior-therap* OR second- OR third- OR fourth- OR prior OR failure OR relaps* OR refrac* OR resist* OR recur* OR progress* OR cancer recurrence OR relapse OR therapy resistance OR tumor recurrence OR recurrent disease OR patient history of therapy OR cancer resistance OR drug resistance OR treatment failure OR salvage therapy OR reocur* OR re occur OR re ocur OR recrudescen*) IN NHSEED, HTA	659	659
21	#8 AND #9	**34**	-
22	(#8 AND #9) IN NHSEED, HTA WHERE LPD FROM 01/03/2021 TO 23/02/2022	-	**0**

A 6.3 Utilities

Table A6.9: Embase.com searches

Sr. No.	Query	De novo searches ran on 05 March 2021	Update searches ran on 22 February 2022
1.	colon cancer'/exp OR 'colon carcinoma'/exp OR 'colorectal cancer'/exp OR 'rectum cancer'/exp OR 'rectum adenoma'/exp	322,420	343,507
2.	((cancer* OR carcinoma* OR adenoma* OR adenocarci* OR tumor* OR tumour* OR neoplasm* OR malignan*) NEAR/4 (colorectal OR 'colo-rectal' OR 'colonrectal' OR 'colon rectal' OR 'colon-rectal' OR colon OR rect* OR pararec* OR bowel OR sigmoid)):ab,ti,kw	347,731	370,078
3.	crc:ab,ti,kw OR mcrc:ab,ti,kw OR 'm-crc':ab,ti,kw	64,569	71,622
4.	'second line chemotherapy' OR 'third line chemotherapy' OR 'fourth line chemotherapy' OR 'second-line' OR 'second line' OR 'third-line' OR 'third line' OR 'fourth-line' OR 'fourth line' OR '2nd line' OR '2nd-line' OR '3rd line' OR '3rd-line' OR '4th line' OR '4th-line' OR 'second or later*' OR 'third or later*' OR 'fourth or later*' OR 'second- or later*' OR 'third- or later*' OR 'fourth- or later*' OR 'second and later*' OR 'third and later*' OR 'fourth and later*' OR 'second- and later*' OR 'third- and later*' OR 'fourth- and later*' OR '2 l' OR '3 l' OR '2l' OR '3l' OR '2-l' OR '3-l' OR '2 line*' OR '2-line*' OR '3 line*' OR '3-line*' OR 'previously treated' OR 'previously-treated' OR 'pre-	9,857,906	10,179,691

Sr. No.	Query	De novo searches ran on 05 March 2021	Update searches ran on 22 February 2022
	treated' OR 'pretreated' OR 'failed' OR 'prior treatment' OR 'prior-treatment' OR 'prior treated' OR 'prior-treated' OR 'prior therap*' OR 'prior-therap*' OR 'second-' OR 'third-' OR 'fourth-' OR 'prior' OR 'failure' OR relaps* OR refrac* OR resist* OR recur* OR progress* OR 'cancer recurrence'/exp OR 'relapse'/exp OR 'therapy resistance'/exp OR 'tumor recurrence'/exp OR 'recurrent disease'/exp OR 'patient history of therapy'/exp OR 'cancer resistance'/exp OR 'drug resistance'/exp OR 'treatment failure'/exp OR 'salvage therapy'/exp OR reocur* OR 're occur' OR 're ocur' OR recrudescen* OR ((post* NEAR/4 (chemo* OR line OR therap* OR treat*)):ab,ti,kw) OR (((pre* OR prio* OR prev* OR post* OR heav* OR late* OR receiv* OR subseque*) NEAR/4 (treat* OR therap* OR regim* OR progress* OR fail* OR relaps* OR resis* OR refract* OR line* OR chemo* OR target*)):ab,ti,kw) OR (((lack* OR inadequa*) NEAR/2 respon*):ab,ti,kw) OR nonrespon*:ab,ti,kw OR 'non respon*':ab,ti,kw OR unrespon*:ab,ti,kw OR 'un-respon*':ab,ti,kw OR 'no respon*':ab,ti,kw OR 'not respon*':ab,ti,kw		
5.	'case study':it OR 'case report':it OR 'abstract report':it OR editorial:it OR letter:it OR comment:it OR note:it OR 'case report'/exp OR 'case study'/exp OR 'editorial'/exp	4,369,440	5,389,710
6.	'animal'/exp NOT ('animal'/exp AND 'human'/exp)	5,579,126	5,738,872
7.	(review:it OR 'literature review':it) NOT ('meta-analysis':it OR 'meta-analysis as topic'/mj OR 'systematic review':ti OR 'systematic literature review':ti OR 'meta-analysis':ab,ti OR 'meta analysis':ab,ti,kw)	2,614,371	2,757,565
8.	#5 OR #6 OR #7	13,041,372	13,578,152
9.	#1 OR #2 OR #3	427,252	455,927
10.	utility:ab,ti,kw NOT ('clinical utility':ab,ti,kw OR 'diagnostic utility':ab,ti,kw) OR 'utilities':ab,ti,kw OR 'disutility':ab,ti,kw OR 'disutilities':ab,ti,kw OR 'sf 6':ab,ti,kw OR sf6:ab,ti,kw OR 'short form 6':ab,ti,kw OR 'shortform 6':ab,ti,kw OR 'sf six':ab,ti,kw OR sfsix:ab,ti,kw OR 'shortform six':ab,ti,kw OR 'short form six':ab,ti,kw OR euroqol:ab,ti,kw OR 'euro qol':ab,ti,kw OR 'euro-qol':ab,ti,kw OR 'euroqol 5d':ab,ti,kw OR 'euroqol-5d':ab,ti,kw OR 'euroqol 5-d':ab,ti,kw OR eq5d:ab,ti,kw OR 'eq 5d':ab,ti,kw OR 'health utilit* index':ab,ti,kw OR hui:ab,ti,kw OR hui1:ab,ti,kw OR hui2:ab,ti,kw OR 'hui-2':ab,ti,kw OR hui3:ab,ti,kw OR 'hui-3':ab,ti,kw OR 'standard gamble*':ab,ti,kw OR ((standard NEAR/2 gamble*):ab,ti,kw) OR 'time trade off':ab,ti,kw OR 'time tradeoff':ab,ti,kw OR timetradeoff*:ab,ti,kw OR tto:ab,ti,kw OR ((time NEAR/2 trade*):ab,ti,kw) OR 'patient preference'/mj OR 'european quality of life 5 dimension'/exp OR ((euro* NEAR/4 'quality of life*'):ab,ti,kw) OR 'visual analog scale':ab,ti,kw OR hsuv*:ab,ti,kw OR 'health* year* equivalent*':ab,ti,kw OR euroqual:ab,ti,kw OR ((euro* NEAR/4 (5d OR '5 d' OR '5-d' OR qol OR 'ql' OR 'quality of life' OR hrql	381,527	415,083

Sr. No.	Query	De novo searches ran on 05 March 2021	Update searches ran on 22 February 2022
	OR hrqol OR qual OR '5 dimension*' OR '5-dimension*' OR 'five dimension*' OR 'five-dimension*')):ab,ti,kw) OR ((eq* NEAR/4 (5d OR '5 d' OR '5-d' OR '5 dimension*' OR '5-dimension*' OR 'five dimension*' OR 'five-dimension*')):ab,ti,kw) OR ((('short-form*' OR sf* OR 'short form') NEAR/4 (6d OR '6 d' OR '6-d' OR '6 dimension*' OR '6-dimension*' OR 'six dimension*' OR 'six-dimension*')):ab,ti,kw) OR ((quality NEAR/3 (wellbeing OR 'well being' OR 'well-being')):ab,ti,kw) OR qwb:ab,ti,kw OR '15 d':ab,ti,kw OR 15d:ab,ti,kw OR '15-d':ab,ti,kw OR '15 dimension':ab,ti,kw OR 'fifteen dimension*':ab,ti,kw OR 'multi-attribute*':ab,ti,kw OR 'multiattribute*':ab,ti,kw OR 'multi attribute*':ab,ti,kw OR 'aqol-8d':ab,ti,kw OR 'aqol 8d':ab,ti,kw OR ((('quality of life' OR qol* OR eortc OR qlq) NEAR/6 (8d OR '8 d' OR '8-d' OR '8 dimension*' OR '8-dimension*' OR 'eight dimension*' OR 'eight-dimension*')):ab,ti,kw) OR maui*:ab,ti,kw OR mauc*:ab,ti,kw OR 'qlu-c10d':ab,ti,kw OR c10d:ab,ti,kw OR 'eortc-8d':ab,ti,kw OR 'eortc 8d':ab,ti,kw OR 'patient preference':ab,ti,kw OR 'health status indicator'/mj OR vignette*:ab,ti,kw OR crosswalk:ab,ti,kw OR 'cross-walk':ab,ti,kw OR 'cross walk':ab,ti,kw OR valuation*:ab,ti,kw OR ((('health state*' OR 'illness state*') NEAR/2 (utility* OR disutility* OR preferen* OR valu*)):ab,ti,kw) OR 'psychometry'/mj OR ('quality of life'/mj AND ec:ab,ti,kw) OR ('quality of life'/mj AND ((health NEAR/2 status):ab,ti,kw)) OR 'time trade-off method'/exp OR 'standard gamble'/exp		
11.	#4 AND #9 AND #10	3,226	3,423
12.	#11 NOT #8	**2,609**	2,728
13.	#11 NOT #8 AND [01-03-2021]/sd	-	**265**

Table A6.10: Medline In-process: PubMed.com searches

Sr. No.	Query	De novo searches ran on 19 March 2021	Update searches ran on 22 February 2022
11.	((colorectal neoplasms[MeSH Terms]) OR (colonic neoplasms[MeSH Terms])) OR (rectal neoplasms[MeSH Terms])	207,249	221,340
12.	("rectum adenoma"[Title/Abstract]) OR ("colon carcinoma"[Title/Abstract])	10,772	10,993
13.	(cancer*[Title/Abstract] OR carcinoma*[Title/Abstract] OR adenoma*[Title/Abstract] OR adenocarci*[Title/Abstract] OR tumor*[Title/Abstract] OR tumour*[Title/Abstract] OR neoplasm*[Title/Abstract] OR malignan*[Title/Abstract]) AND (colorectal[Title/Abstract] OR "colo-rectal"[Title/Abstract] OR "colonrectal"[Title/Abstract] OR "colon rectal"[Title/Abstract] OR "colon-rectal"[Title/Abstract] OR colon[Title/Abstract] OR rect*[Title/Abstract] OR pararec*[Title/Abstract] OR bowel[Title/Abstract] OR sigmoid[Title/Abstract])	293,031	312,098
14.	#1 OR #2 OR #3	337,393	358,129
15.	"health utility"[tiab] OR "health utilities"[tiab] OR "health state utility"[tiab] OR "health state utilities"[tiab] OR "utility score"[tiab] OR "utility scores"[tiab] OR "utility value"[tiab] OR "utility values"[tiab] OR "utility valuation"[tiab] OR "Standard gamble"[tiab] OR SG[tiab] OR "time trade-off"[tiab] OR "time tradeoff"[tiab] OR TTO[tiab] OR "visual analog scale"[tiab] OR "visual analog scales"[tiab] OR "visual analogue scale"[tiab] OR "visual analogue scales"[tiab] OR "patient preference"[tiab] OR "patient preferences"[tiab] OR preference[tiab] OR preferences[tiab] OR "EQ-5D"[tiab] OR "EQ5D"[tiab] OR EuroQol[tiab] OR "health utilities index"[tiab] OR HUI[tiab] OR SF-6D[tiab] OR "short form 6D"[tiab] OR "quality of well-being scale"[tiab] OR "quality of well-being scales"[tiab] OR "health utility index"[tiab] OR "utility assessment"[tiab] OR QALY[tiab] OR QALYs[tiab] OR "quality adjusted life years"[tiab] OR utility[tiab] OR utilities[tiab] NOT (electric OR electricity)	452,384	489,848
16.	(publisher[sb] NOT pubstatusnihms NOT pubstatuspmcsd NOT pmcbook) OR (pubstatusaheadofprint) OR Iinprocess[sb])	372,882	501,124
17.	#4 AND #5 AND #6	**103**	**175**

Table A6.11: EconLit: Ebsco.com searches

S. No.	Query	Search Options	De novo searches ran on 19 March 2021	Update searches ran on 22 February 2022
S1	(“colon cancer” OR “colon carcinoma” OR “colorectal cancer” OR “rectum cancer” OR “rectum adenoma”) OR ((cancer OR carcinoma OR adenoma OR adenocarcinoma OR tumor OR tumour OR neoplasm OR malignant OR malignancy) AND(colorectal OR “colo-rectal” OR “colonrectal” OR “colon rectal” OR“ colon-rectal” OR colon OR rectum OR rectal OR pararectal OR bowel OR sigmoid)) AND (crc OR mcrc OR “m-crc”)	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	**114**	-
S1	(“colon cancer” OR “colon carcinoma” OR “colorectal cancer” OR “rectum cancer” OR “rectum adenoma”) OR ((cancer OR carcinoma OR adenoma OR adenocarcinoma OR tumor OR tumour OR neoplasm OR malignant OR malignancy) AND(colorectal OR “colo-rectal” OR “colonrectal” OR “colon rectal” OR“ colon-rectal” OR colon OR rectum OR rectal OR pararectal OR bowel OR sigmoid)) AND (crc OR mcrc OR “m-crc”)	**Expanders** - Also search within the full text of the articles **Search modes** - Find all my search terms **Limiters** - Date Published: 20210101-20221231	-	**6**

Table A6.12: NHSEED and HTAD: CRD York.com searches

S. No.	Query	De novo searches ran on 07 April 2021	Update searches ran on 22 February 2022
23.	MeSH DESCRIPTOR Colorectal Neoplasms EXPLODE ALL TREES IN NHSEED,HTA	673	673
24.	MeSH DESCRIPTOR Rectal Neoplasms EXPLODE ALL TREES IN NHSEED,HTA	70	70
25.	MeSH DESCRIPTOR Colonic Neoplasms EXPLODE ALL TREES IN NHSEED,HTA	92	92
26.	(cancer* OR carcinoma* OR adenoma* OR adenocarci* OR tumor* OR tumour* OR neoplasm* OR malignan*) IN NHSEED, HTA	6,654	6,654
27.	(colorectal OR colo-rectal OR colonrectal OR 'colon rectal' OR colon-rectal OR colon OR rect* OR pararec* OR bowel OR sigmoid) IN NHSEED, HTA	1,429	1,429
28.	#4 AND #5	998	998
29.	(crc OR mcrc OR m-crc):TI IN NHSEED, HTA	9	9

30.	#1 OR #2 OR #3 OR #6 OR #7	1,027	1,027
31.	(second line chemotherapy OR third line chemotherapy OR fourth line chemotherapy OR second-line OR second line OR third-line OR third line OR fourth-line OR fourth line OR 2nd line OR 2nd-line OR 3rd line OR 3rd-line OR 4th line OR 4th-line OR second or later OR third or later OR fourth or later OR second and later OR third and later OR fourth and later OR 2 l OR 3 l OR 2l OR 3l OR 2-l OR 3-l OR 2 line* OR 2-line OR 3 line* OR 3-line* OR previously treated OR previously-treated OR pre-treated OR pretreated OR failed OR prior treatment OR prior-treatment OR prior treated OR prior-treated OR prior therap* OR prior-therap* OR second- OR third- OR fourth- OR prior OR failure OR relaps* OR refrac* OR resist* OR recur* OR progress* OR cancer recurrence OR relapse OR therapy resistance OR tumor recurrence OR recurrent disease OR patient history of therapy OR cancer resistance OR drug resistance OR treatment failure OR salvage therapy OR reocur* OR re occur OR re ocur OR recrudescen*) IN NHSEED, HTA	659	659
32.	#8 AND #9	34	-
33.	**(#8 AND #9) IN NHSEED, HTA WHERE LPD FROM 01/03/2021 TO 23/02/2022**	-	**0**

Decision problem

A 7. Priority question: The population in the final scope by NICE is "*Adults with metastatic colorectal cancer who have been previously treated with, or are not considered candidates for, available therapies.*" (p. 2). This might imply that the only eligible comparator is BSC, notwithstanding the fact that the NICE scope also lists various comparators including T/T in addition to BSC. However, 'Available therapies' are defined more precisely in the CS as "*those available prior to 2013 and include fluoropyrimidine-based chemotherapy, anti-VEGF therapy and anti-epidermal growth factor receptor (anti-EGFR) therapy*" (p. 16). The population in this decision problem is distinct in that it is patients for whom one alternative treatment - T/T – remains: " *Specifically, we are seeking a recommendation for patients for whom treatment with trifluridine/tipiracil is being considered."* (Table 1, CS) Also, according to Figure 1, T/T precedes BSC and is also recommended for adults who have had previous treatment with available therapies, or for whom these available therapies are not suitable.

a) Please clarify that 'available therapies' include all comparators in the scope other than T/T.

Yes, "available therapies" includes all comparators in the scope other than trifluridine/tipiracil. Regorafenib is only considered after failure of these agents.

The inclusion criteria in the CORRECT study, in relation to prior treatment was that patients should have progressed following the administration of approved standard therapies which was to include fluoropyrimidine, oxaliplatin, irinotecan, bevacizumab and cetuximab or panitumumab (if KRAS wild type).

The inclusion criteria in the CONCUR study, in relation to prior treatment was that patients should have progressed following the administration of approved standard therapies which was to include fluoropyrimidine, oxaliplatin and irinotecan.

In respect of raltitrexed, this is only indicated where 5-FU (fluoropyrimidine) and folinic acid based regimens are either not tolerated or inappropriate. Raltitrexed is an alternative to these agents and not an alternative to regorafenib.

b) Please clarify whether the population in the decision problem is intended to be an earlier line than BSC, thus ruling out BSC as a comparator. If so, then please explain what is meant by BSC being a "*minor comparator*" (Table 1, CS).

Yes, the population is intended to be earlier than BSC thus ruling out BSC as a comparator. Please also see our response to B4.

The description of BSC as a 'minor' comparator was a poor choice of words as we don't consider it to be a comparator. Rather, we included a limited set of analysis against BSC as it was included in the CORRECT and CONCUR trials and its inclusion was useful from a model validation perspective. In hindsight it would have been clearer if no analyses against BSC had been presented in the submission.

c) Please clarify that the patients in this population would not be eligible for any treatment other than T/T.

The recommendation for trifluridine/tipiracil (TA405) is after available therapies and an option before BSC. Trifluridine/tipiracil is used as a last-line option i.e. after other options have been exhausted.

If regorafenib is considered alongside trifluridine/tipiracil then it follows that the patients do not have other options. Physicians requested we complete a submission as they wanted an alternative option alongside trifluridine/tipiracil.

d) If the only comparator is T/T then please explain how patients would be identified in UK clinical practice to be only eligible for T/T in terms of treatment history and any other clinical characteristics.

As a last-line treatment before BSC, trifluridine/tipiracil is identified as a treatment option if earlier lines have failed or are not appropriate. In addition, the patient would need to be considered 'fit' enough to receive chemotherapy and to be able to tolerate the adverse events typical of chemotherapies.

e) Given that BSC is included in the scope, in the CONCUR and CORRECT studies as comparator and concomitant to regorafenib, in the economic model as concomitant to regorafenib and T/T and as subsequent treatment, please provide a detailed description of BSC as would be observed in UK clinical practice and how it might vary depending on whether concomitant and on line of therapy

According to NICE "Cancer Service Guidance – improving supportive and palliative care for adults with cancer", patients should have access to a range of services to improve their quality of life such as physical, psychological, spiritual or emotional support. People with cancer may also need assistance with symptom management, either regularly or from time to time. Those working with patients should assess their needs for help with a broad range of symptoms, such as pain, fatigue or breathlessness, and set up a plan to manage these.

We are not aware of any sources of data which provide information on exactly what supportive care is provided to patients with cancer in England. However, clinically there is no reason why it would differ meaningfully from the supportive care received in CORRECT and CONCUR i.e. 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.

Best supportive care in the regorafenib trials excluded other investigational anti-tumour agents or anti-neoplastic chemo/hormonal/immuno-therapy – this exclusion is common in context of clinical trials. Patients in England might be candidates for such investigational therapy.

We are not aware of any guidance on BSC specifically related to line of therapy, however, it could be expected that patients require more medicines for pain control and palliative care as the disease progresses (as indicated by each additional line of therapy).

f) Please make a detailed comparison between BSC as observed in the regorafenib, RECOURSE, TERRA, and Yoshino (2012) trials and BSC as would be observed in UK clinical practice.

It is not possible to provide all of the requested information - detailed data on exactly what is provided as best supportive care in the UK, and in what proportion to patients, is not available (see part e above). Furthermore, no detail on best supportive care is available to Bayer for the trifluridine/tipiracil trials. Table A7.1 provides a topline description of BSC from the trials. Table A7.2 provides detail on the concomitant medicines received in the CORRECT trial for both arms – please note that not all of this medication is necessarily BSC but also includes medications necessary to manage the patients co-morbidities. Table A7.3 provides detail on the concomitant medicines received in the CONCUR trial.

Table A7.1. Description of best supportive care in the regorafenib and trifluridine/tipiracil trials

	Description of best supportive care	
CORRECT	BSC 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. Best supportive care excluded other investigational anti-tumour agents or anti-neoplastic chemo/hormonal/immuno-therapy.	
CONCUR	BSC 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.	

	Best supportive care excluded other investigational anti-tumour agents or anti-neoplastic chemo/hormonal/immuno-therapy.	
RECOURSE	Definition not available from primary publication	In ERG report to TA405 (section 3.4) it is stated that the ERG questioned how each trial defined best supportive care. The company clarified that there was no internationally accepted definition of best supportive care, but that all necessary support was provided to patients, except therapies that were not permitted in trial protocols.
Yoshino	Definition not available from primary publication	
TERRA	Definition not available from primary publication	

Table A7.2: CORRECT - Number of subjects who took at least one concomitant medication (ITT analysis set)

ATC CLASSIFICATION SUBCLASS WHO-DD Version 3Q2005	Placebo N=255 (100%)	Regorafenib 160 mg N=505 (100%)	Total N=760 (100%)
Number of subjects (%) with at least one concomitant medication	254 (99.6%)	505 (100.0%)	759 (99.9%)
ALIMENTARY TRACT AND METABOLISM	214 (83.9%)	472 (93.5%)	686 (90.3%)
ALIMENTARY TRACT AND METABOLISM	0	1 (0.2%)	1 (0.1%)
ALPHA GLUCOSIDASE INHIBITORS	1 (0.4%)	3 (0.6%)	4 (0.5%)
ALUMINIUM COMPOUNDS	1 (0.4%)	3 (0.6%)	4 (0.5%)
AMINO ACIDS AND DERIVATIVES	0	3 (0.6%)	3 (0.4%)
AMINOSALICYLIC ACID AND SIMILAR AGENTS	0	2 (0.4%)	2 (0.3%)
ANTACIDS	1 (0.4%)	6 (1.2%)	7 (0.9%)
ANTACIDS WITH ANTIFLATULENTS	1 (0.4%)	2 (0.4%)	3 (0.4%)
ANTACIDS WITH SODIUM BICARBONATE	4 (1.6%)	16 (3.2%)	20 (2.6%)
ANTIALLERGIC AGENTS, EXCL. CORTICOSTEROIDS	1 (0.4%)	1 (0.2%)	2 (0.3%)
ANTIBIOTICS	11 (4.3%)	45 (8.9%)	56 (7.4%)
ANTIDIARRHEAL MICROORGANISMS	10 (3.9%)	22 (4.4%)	32 (4.2%)
ANTIEMETICS AND ANTINAUSEANTS	1 (0.4%)	3 (0.6%)	4 (0.5%)
ANTIINFECT. AND ANTISEPT. FOR LOCAL ORAL TREATMENT	15 (5.9%)	58 (11.5%)	73 (9.6%)
ANTIPROPULSIVES	28 (11.0%)	138 (27.3%)	166 (21.8%)
ANTISPASMODICS IN COMBINATION WITH OTHER DRUGS	4 (1.6%)	1 (0.2%)	5 (0.7%)
ANTISPASMODICS, PSYCHOLEPTICS, ANALGESICS IN COMB	1 (0.4%)	1 (0.2%)	2 (0.3%)
APPETITE STIMULANTS	9 (3.5%)	20 (4.0%)	29 (3.8%)
ASCORBIC ACID (VITAMIN C), INCL. COMBINATIONS	0	1 (0.2%)	1 (0.1%)
ASCORBIC ACID (VITAMIN C), PLAIN	5 (2.0%)	6 (1.2%)	11 (1.4%)
BELLADONNA ALKALOIDS SEMISYNT,QUATER AMMONIUM COMP	12 (4.7%)	23 (4.6%)	35 (4.6%)
BELLADONNA ALKALOIDS, TERTIARY AMINES	2 (0.8%)	1 (0.2%)	3 (0.4%)
BELLADONNA AND DERIVATIVES IN COMB.WITH PSYCHOLEP.	1 (0.4%)	0	1 (0.1%)
BIGUANIDES	13 (5.1%)	29 (5.7%)	42 (5.5%)
BILE ACID PREPARATIONS	3 (1.2%)	15 (3.0%)	18 (2.4%)
BULK PRODUCERS	2 (0.8%)	7 (1.4%)	9 (1.2%)
CALCIUM	4 (1.6%)	23 (4.6%)	27 (3.6%)
CALCIUM COMPOUNDS	2 (0.8%)	8 (1.6%)	10 (1.3%)
CALCIUM, COMBINATIONS WITH OTHER DRUGS	2 (0.8%)	12 (2.4%)	14 (1.8%)
CHARCOAL PREPARATIONS	2 (0.8%)	1 (0.2%)	3 (0.4%)
COMB AND COMPL. OF ALUMIN., CALC. AND MAGNES. COMP	7 (2.7%)	28 (5.5%)	35 (4.6%)
COMBINATIONS OF ORAL BLOOD GLUCOSE LOWERING DRUGS	4 (1.6%)	1 (0.2%)	5 (0.7%)
COMBINATIONS OF VITAMINS	3 (1.2%)	4 (0.8%)	7 (0.9%)
CONTACT LAXATIVES	23 (9.0%)	52 (10.3%)	75 (9.9%)
CORTICOSTEROIDS ACTING LOCALLY	44 (17.3%)	144 (28.5%)	188 (24.7%)
CORTICOSTEROIDS FOR LOCAL ORAL TREATMENT	45 (17.6%)	102 (20.2%)	147 (19.3%)
DIGESTIVES, INCL. ENZYMES	2 (0.8%)	2 (0.4%)	4 (0.5%)
DIPEPTIDYL PEPTIDASE 4 (DPP-4) INHIBITORS	1 (0.4%)	3 (0.6%)	4 (0.5%)
DRUGS FOR FUNCTIONAL GASTROINTESTINAL DISORDERS	0	2 (0.4%)	2 (0.3%)
DRUGS FOR PEPTIC ULCER AND GORD	5 (2.0%)	21 (4.2%)	26 (3.4%)

ENEMAS	7 (2.7%)	16 (3.2%)	23 (3.0%)
ENZYME PREPARATIONS	0	4 (0.8%)	4 (0.5%)
H2-RECEPTOR ANTAGONISTS	21 (8.2%)	39 (7.7%)	60 (7.9%)
INSULINS AND ANALOGUES	3 (1.2%)	2 (0.4%)	5 (0.7%)
INSULINS AND ANALOGUES, FAST-ACTING	19 (7.5%)	23 (4.6%)	42 (5.5%)
INSULINS AND ANALOGUES, LONG-ACTING	3 (1.2%)	8 (1.6%)	11 (1.4%)
INTESTINAL ADSORBENTS	0	2 (0.4%)	2 (0.3%)
LAXATIVES	15 (5.9%)	54 (10.7%)	69 (9.1%)
LIVER THERAPY	2 (0.8%)	8 (1.6%)	10 (1.3%)
MAGNESIUM	7 (2.7%)	19 (3.8%)	26 (3.4%)
MAGNESIUM COMPOUNDS	21 (8.2%)	44 (8.7%)	65 (8.6%)
MINERAL SUPPLEMENTS	1 (0.4%)	2 (0.4%)	3 (0.4%)
MULTIVITAMINS WITH MINERALS	4 (1.6%)	6 (1.2%)	10 (1.3%)
MULTIVITAMINS, OTHER COMBINATIONS	2 (0.8%)	3 (0.6%)	5 (0.7%)
MULTIVITAMINS, PLAIN	5 (2.0%)	21 (4.2%)	26 (3.4%)
ORAL BLOOD GLUCOSE LOWERING DRUGS	1 (0.4%)	2 (0.4%)	3 (0.4%)
ORAL REHYDRATION SALT FORMULATIONS	3 (1.2%)	4 (0.8%)	7 (0.9%)
OSMOTICALLY ACTING LAXATIVES	48 (18.8%)	100 (19.8%)	148 (19.5%)
OTHER AGENTS FOR LOCAL ORAL TREATMENT	23 (9.0%)	40 (7.9%)	63 (8.3%)
OTHER ALIMENTARY TRACT AND METABOLISM PRODUCTS	1 (0.4%)	4 (0.8%)	5 (0.7%)
OTHER ANTIDIARRHEALS	3 (1.2%)	5 (1.0%)	8 (1.1%)
OTHER ANTIEMETICS	25 (9.8%)	29 (5.7%)	54 (7.1%)
OTHER DRUGS FOR FUNCTIONAL BOWEL DISORDERS	11 (4.3%)	27 (5.3%)	38 (5.0%)
OTHER DRUGS FOR PEPTIC ULCER AND GORD	5 (2.0%)	15 (3.0%)	20 (2.6%)
OTHER INTESTINAL ANTIINFECTIVES	0	3 (0.6%)	3 (0.4%)
OTHER MINERAL PRODUCTS	0	7 (1.4%)	7 (0.9%)
OTHER MINERAL SUPPLEMENTS	1 (0.4%)	1 (0.2%)	2 (0.3%)
OTHER ORAL BLOOD GLUCOSE LOWERING DRUGS	2 (0.8%)	3 (0.6%)	5 (0.7%)
OTHER PLAIN VITAMIN PREPARATIONS	15 (5.9%)	46 (9.1%)	61 (8.0%)
PAPAVERINE AND DERIVATIVES	0	2 (0.4%)	2 (0.3%)
PERIPHERAL OPIOID RECEPTOR ANTAGONISTS	0	2 (0.4%)	2 (0.3%)
POTASSIUM	19 (7.5%)	51 (10.1%)	70 (9.2%)
PROPULSIVES	49 (19.2%)	109 (21.6%)	158 (20.8%)
PROSTAGLANDINS	1 (0.4%)	1 (0.2%)	2 (0.3%)
PROTON PUMP INHIBITORS	112 (43.9%)	244 (48.3%)	356 (46.8%)
SELENIUM	2 (0.8%)	2 (0.4%)	4 (0.5%)
SEROTONIN (5HT3) ANTAGONISTS	31 (12.2%)	39 (7.7%)	70 (9.2%)
SODIUM	1 (0.4%)	0	1 (0.1%)
SOFTENERS, EMOLLIENTS	9 (3.5%)	9 (1.8%)	18 (2.4%)
STOMATOLOGICAL PREPARATIONS	5 (2.0%)	37 (7.3%)	42 (5.5%)
SULFONAMIDES, UREA DERIVATIVES	10 (3.9%)	23 (4.6%)	33 (4.3%)
SYNT ANTICHOLIN,ESTERS WITH TERTIARY AMINO GROUP	8 (3.1%)	10 (2.0%)	18 (2.4%)
SYNT ANTICHOLINERGICS,QUATERNARY AMMONIUM COMPOUND	2 (0.8%)	3 (0.6%)	5 (0.7%)
SYNTHETIC ANTICHOLIN. AGENTS IN COMB. W/ANALGESICS	0	2 (0.4%)	2 (0.3%)
SYNTHETIC ANTICHOLIN. AGENTS IN COMB. W/PSYCHOLEPT	0	1 (0.2%)	1 (0.1%)
THIAZOLIDINEDIONES	1 (0.4%)	2 (0.4%)	3 (0.4%)
TONICS	0	1 (0.2%)	1 (0.1%)
VARIOUS ALIMENTARY TRACT AND METABOLISM PRODUCTS	1 (0.4%)	2 (0.4%)	3 (0.4%)
VIT B1 IN COMB WITH VITAMIN B6 AND/OR VITAMIN B12	1 (0.4%)	4 (0.8%)	5 (0.7%)
VITAMIN A AND D IN COMBINATION	0	1 (0.2%)	1 (0.1%)
VITAMIN A, PLAIN	1 (0.4%)	1 (0.2%)	2 (0.3%)
VITAMIN B-COMPLEX WITH MINERALS	2 (0.8%)	1 (0.2%)	3 (0.4%)
VITAMIN B-COMPLEX, PLAIN	3 (1.2%)	7 (1.4%)	10 (1.3%)
VITAMIN B1, PLAIN	2 (0.8%)	3 (0.6%)	5 (0.7%)
VITAMIN D AND ANALOGUES	6 (2.4%)	12 (2.4%)	18 (2.4%)
VITAMINS	0	2 (0.4%)	2 (0.3%)
VITAMINS WITH MINERALS	5 (2.0%)	2 (0.4%)	7 (0.9%)
VITAMINS, OTHER COMBINATIONS	0	2 (0.4%)	2 (0.3%)
ZINC	0	1 (0.2%)	1 (0.1%)
ANTIINFECTIVES FOR SYSTEMIC USE	78 (30.6%)	219 (43.4%)	297 (39.1%)
AMPHENICOLS	1 (0.4%)	0	1 (0.1%)
ANTIBIOTICS	0	1 (0.2%)	1 (0.1%)
BETA-LACTAMASE INHIBITORS	2 (0.8%)	2 (0.4%)	4 (0.5%)
BETA-LACTAMASE RESISTANT PENICILLINS	0	1 (0.2%)	1 (0.1%)
BETA-LACTAMASE SENSITIVE PENICILLINS	1 (0.4%)	0	1 (0.1%)
CARBAPENEMS	6 (2.4%)	10 (2.0%)	16 (2.1%)
COMB OF PENICILLINS, INCL. BETA-LACTAMASE INHIB.	13 (5.1%)	65 (12.9%)	78 (10.3%)
COMB.SULFONAMIDES & TRIMETHOPRIM INCL. DERIVATIVES	5 (2.0%)	2 (0.4%)	7 (0.9%)
COMBINATIONS OF ANTIBACTERIALS	2 (0.8%)	5 (1.0%)	7 (0.9%)
FIRST-GENERATION CEPHALOSPORINS	4 (1.6%)	9 (1.8%)	13 (1.7%)
FLUOROQUINOLONES	37 (14.5%)	119 (23.6%)	156 (20.5%)
FOURTH-GENERATION CEPHALOSPORINS	5 (2.0%)	8 (1.6%)	13 (1.7%)
GLYCOPEPTIDE ANTIBACTERIALS	0	2 (0.4%)	2 (0.3%)
IMMUNOGLOBULINS, NORMAL HUMAN	1 (0.4%)	0	1 (0.1%)
INFLUENZA VACCINES	0	5 (1.0%)	5 (0.7%)
MACROLIDES	11 (4.3%)	14 (2.8%)	25 (3.3%)
NEURAMINIDASE INHIBITORS	0	1 (0.2%)	1 (0.1%)
NITROFURAN DERIVATIVES	1 (0.4%)	3 (0.6%)	4 (0.5%)
NUCLEOSIDE REVERSE TRANSCRIPTASE INHIBITORS	1 (0.4%)	0	1 (0.1%)
NUCLEOSIDES AND NUCLEOTIDES EXCL REV.TRANSCR.INHIB	1 (0.4%)	2 (0.4%)	3 (0.4%)
OTHER AMINOGLYCOSIDES	0	1 (0.2%)	1 (0.1%)
OTHER ANTIBACTERIALS	0	7 (1.4%)	7 (0.9%)
OTHER BETA-LACTAM ANTIBACTERIALS	1 (0.4%)	2 (0.4%)	3 (0.4%)
PENICILLINS WITH EXTENDED SPECTRUM	15 (5.9%)	37 (7.3%)	52 (6.8%)
PNEUMOCOCCAL VACCINES	0	2 (0.4%)	2 (0.3%)
QUINOLONE ANTIBACTERIALS	0	1 (0.2%)	1 (0.1%)

SECOND-GENERATION CEPHALOSPORINS	4 (1.6%)	17 (3.4%)	21 (2.8%)
STREPTOGRAMINS	0	1 (0.2%)	1 (0.1%)
THIRD-GENERATION CEPHALOSPORINS	11 (4.3%)	37 (7.3%)	48 (6.3%)
TRIAZOLE DERIVATIVES	1 (0.4%)	7 (1.4%)	8 (1.1%)
TRIMETHOPRIM AND DERIVATIVES	0	1 (0.2%)	1 (0.1%)
NTINEOPLASTIC AND IMMUNOMODULATING AGENTS	6 (2.4%)	8 (1.6%)	14 (1.8%)
ANTI-ESTROGENS	1 (0.4%)	0	1 (0.1%)
ANTINEOPLASTIC AGENTS	1 (0.4%)	0	1 (0.1%)
COLONY STIMULATING FACTORS	0	3 (0.6%)	3 (0.4%)
CYTOKINES AND IMMUNOMODULATORS	1 (0.4%)	0	1 (0.1%)
OTHER ANTINEOPLASTIC AGENTS	3 (1.2%)	5 (1.0%)	8 (1.1%)
OTHER CYTOKINES AND IMMUNOMODULATORS	1 (0.4%)	0	1 (0.1%)
NTIPARASITIC PRODUCTS,INSECTICIDES AND REPELLENTS	0	2 (0.4%)	2 (0.3%)
OTHER AGENTS AGAINST AMOEBIASIS & OTH PROTOZO. DIS	0	1 (0.2%)	1 (0.1%)
OTHER ANTINEMATODALS	0	1 (0.2%)	1 (0.1%)
LOOD AND BLOOD FORMING ORGANS	114 (44.7%)	258 (51.1%)	372 (48.9%)
AMINO ACIDS	3 (1.2%)	10 (2.0%)	13 (1.7%)
ANTIINFECTIVES	0	1 (0.2%)	1 (0.1%)
ANTITHROMBOTIC AGENTS	2 (0.8%)	3 (0.6%)	5 (0.7%)
BLOOD AND RELATED PRODUCTS	18 (7.1%)	46 (9.1%)	64 (8.4%)
BLOOD SUBSTITUTES AND PERFUSION SOLUTIONS	3 (1.2%)	9 (1.8%)	12 (1.6%)
BLOOD SUBSTITUTES AND PLASMA PROTEIN FRACTIONS	6 (2.4%)	20 (4.0%)	26 (3.4%)
ELECTROLYTE SOLUTIONS	2 (0.8%)	6 (1.2%)	8 (1.1%)
ENZYMES	2 (0.8%)	1 (0.2%)	3 (0.4%)
FOLIC ACID AND DERIVATIVES	0	4 (0.8%)	4 (0.5%)
HEPARIN GROUP	65 (25.5%)	125 (24.8%)	190 (25.0%)
I.V. SOLUTION ADDITIVES	2 (0.8%)	2 (0.4%)	4 (0.5%)
I.V. SOLUTIONS	4 (1.6%)	7 (1.4%)	11 (1.4%)
IRON BIVALENT, ORAL PREPARATIONS	6 (2.4%)	22 (4.4%)	28 (3.7%)
IRON IN OTHER COMBINATIONS	0	1 (0.2%)	1 (0.1%)
IRON TRIVALENT, ORAL PREPARATIONS	0	3 (0.6%)	3 (0.4%)
IRON TRIVALENT, PARENTERAL PREPARATIONS	0	3 (0.6%)	3 (0.4%)
OTHER ANTIANEMIC PREPARATIONS	2 (0.8%)	15 (3.0%)	17 (2.2%)
OTHER ANTITHROMBOTIC AGENTS	0	4 (0.8%)	4 (0.5%)
OTHER IRRIGATING SOLUTIONS	1 (0.4%)	2 (0.4%)	3 (0.4%)
OTHER SYSTEMIC HEMOSTATICS	1 (0.4%)	5 (1.0%)	6 (0.8%)
PLATELET AGGREGATION INHIBITORS EXCL. HEPARIN	12 (4.7%)	16 (3.2%)	28 (3.7%)
PROTEINASE INHIBITORS	1 (0.4%)	3 (0.6%)	4 (0.5%)
SALT SOLUTIONS	3 (1.2%)	3 (0.6%)	6 (0.8%)
SOLUTIONS AFFECTING THE ELECTROLYTE BALANCE	15 (5.9%)	40 (7.9%)	55 (7.2%)

SOLUTIONS FOR PARENTERAL NUTRITION	12 (4.7%)	36 (7.1%)	48 (6.3%)
SOLUTIONS PRODUCING OSMOTIC DIURESIS	9 (3.5%)	43 (8.5%)	52 (6.8%)
VITAMIN B12 (CYANOCOBALAMIN AND ANALOGUES)	1 (0.4%)	9 (1.8%)	10 (1.3%)
VITAMIN K	6 (2.4%)	16 (3.2%)	22 (2.9%)
VITAMIN K ANTAGONISTS	5 (2.0%)	27 (5.3%)	32 (4.2%)
CARDIOVASCULAR SYSTEM	164 (64.3%)	375 (74.3%)	539 (70.9%)
ACE INHIBITORS AND CALCIUM CHANNEL BLOCKERS	3 (1.2%)	3 (0.6%)	6 (0.8%)
ACE INHIBITORS AND DIURETICS	5 (2.0%)	18 (3.6%)	23 (3.0%)
ACE INHIBITORS, COMBINATIONS	0	1 (0.2%)	1 (0.1%)
ACE INHIBITORS, PLAIN	43 (16.9%)	84 (16.6%)	127 (16.7%)
ADRENERGIC AND DOPAMINERGIC AGENTS	2 (0.8%)	6 (1.2%)	8 (1.1%)
AGENTS ACTING ON THE RENIN-ANGIOTENSIN SYSTEM	1 (0.4%)	1 (0.2%)	2 (0.3%)
ALDOSTERONE ANTAGONISTS	15 (5.9%)	31 (6.1%)	46 (6.1%)
ALPHA AND BETA BLOCKING AGENTS	1 (0.4%)	6 (1.2%)	7 (0.9%)
ALPHA-ADRENORECEPTOR ANTAGONISTS	4 (1.6%)	12 (2.4%)	16 (2.1%)
ANGIOTENSIN II ANTAGONISTS AND CALCIUM CHANNEL BLO	0	1 (0.2%)	1 (0.1%)
ANGIOTENSIN II ANTAGONISTS AND DIURETICS	7 (2.7%)	17 (3.4%)	24 (3.2%)
ANGIOTENSIN II ANTAGONISTS, COMBINATIONS	1 (0.4%)	2 (0.4%)	3 (0.4%)
ANGIOTENSIN II ANTAGONISTS, PLAIN	16 (6.3%)	62 (12.3%)	78 (10.3%)
ANTIARRHYTHMICS, CLASS I AND III	3 (1.2%)	6 (1.2%)	9 (1.2%)
ANTIARRHYTHMICS, CLASS IB	7 (2.7%)	27 (5.3%)	34 (4.5%)
ANTIARRHYTHMICS, CLASS IC	0	1 (0.2%)	1 (0.1%)
ANTIARRHYTHMICS, CLASS III	1 (0.4%)	5 (1.0%)	6 (0.8%)
ANTIHEMORRHOIDALS FOR TOPICAL USE	0	2 (0.4%)	2 (0.3%)
ANTIVARICOSE THERAPY	0	1 (0.2%)	1 (0.1%)
BENZOTHIAZEPINE DERIVATIVES	0	10 (2.0%)	10 (1.3%)
BETA BLOCKING AGENTS, NON-SELECTIVE	1 (0.4%)	11 (2.2%)	12 (1.6%)
BETA BLOCKING AGENTS, SELECTIVE	43 (16.9%)	78 (15.4%)	121 (15.9%)
BETA BLOCKING AGENTS, SELECTIVE, AND THIAZIDES	3 (1.2%)	3 (0.6%)	6 (0.8%)
BETA BLOCKING AGENTS,SELECTIVE,AND OTHER DIURETICS	1 (0.4%)	2 (0.4%)	3 (0.4%)
BILE ACID SEQUESTRANTS	0	2 (0.4%)	2 (0.3%)
BIOFLAVONOIDS	4 (1.6%)	3 (0.6%)	7 (0.9%)
CARDIAC THERAPY	0	1 (0.2%)	1 (0.1%)
DIGITALIS GLYCOSIDES	1 (0.4%)	10 (2.0%)	11 (1.4%)
DIHYDROPYRIDINE DERIVATIVES	35 (13.7%)	126 (25.0%)	161 (21.2%)
DIURETICS	0	1 (0.2%)	1 (0.1%)
FIBRATES	3 (1.2%)	5 (1.0%)	8 (1.1%)
HIGH-CEILING DIURETICS AND POTASSIUM-SPARING AGENT	0	2 (0.4%)	2 (0.3%)
HMG COA REDUCTASE INHIBITORS	40 (15.7%)	46 (9.1%)	86 (11.3%)
HYDRAZINOPHTHALAZINE DERIVATIVES	0	1 (0.2%)	1 (0.1%)
IMIDAZOLINE RECEPTOR AGONISTS	2 (0.8%)	8 (1.6%)	10 (1.3%)
LOW-CEILING DIURETICS AND POTASSIUM-SPARING AGENTS	2 (0.8%)	12 (2.4%)	14 (1.8%)
ORGANIC NITRATES	5 (2.0%)	6 (1.2%)	11 (1.4%)

OTHER ANTIHEMORRHOIDALS FOR TOPICAL USE	1 (0.4%)	1 (0.2%)	2 (0.3%)
OTHER CAPILLARY STABILIZING AGENTS	1 (0.4%)	0	1 (0.1%)
OTHER CARDIAC COMBINATION PRODUCTS	1 (0.4%)	2 (0.4%)	3 (0.4%)
OTHER CARDIAC PREPARATIONS	25 (9.8%)	59 (11.7%)	84 (11.1%)
OTHER CHOLESTEROL AND TRIGLYCERIDE REDUCERS	3 (1.2%)	6 (1.2%)	9 (1.2%)
OTHER LOW-CEILING DIURETICS	1 (0.4%)	0	1 (0.1%)
OTHER SCLEROSING AGENTS	0	1 (0.2%)	1 (0.1%)
OTHER VASODILATORS USED IN CARDIAC DISEASES	2 (0.8%)	4 (0.8%)	6 (0.8%)
PHENYLALKYLAMINE DERIVATIVES	3 (1.2%)	7 (1.4%)	10 (1.3%)
PRODUCTS CONTAINING CORTICOSTEROIDS	1 (0.4%)	15 (3.0%)	16 (2.1%)
PRODUCTS CONTAINING LOCAL ANESTHETICS	0	3 (0.6%)	3 (0.4%)
PURINE DERIVATIVES	1 (0.4%)	1 (0.2%)	2 (0.3%)
RENIN-INHIBITORS	0	1 (0.2%)	1 (0.1%)
SULFONAMIDES, PLAIN	34 (13.3%)	71 (14.1%)	105 (13.8%)
THIAZIDES, PLAIN	4 (1.6%)	17 (3.4%)	21 (2.8%)
XANTHINE DERIVATIVES	1 (0.4%)	2 (0.4%)	3 (0.4%)
DERMATOLOGICALS	62 (24.3%)	250 (49.5%)	312 (41.1%)
ANESTHETICS FOR TOPICAL USE	0	2 (0.4%)	2 (0.3%)
ANTIBIOTICS	1 (0.4%)	1 (0.2%)	2 (0.3%)
ANTIBIOTICS FOR TOPICAL USE	0	1 (0.2%)	1 (0.1%)
ANTIFUNGALS FOR TOPICAL USE	0	1 (0.2%)	1 (0.1%)
ANTIHISTAMINES FOR TOPICAL USE	6 (2.4%)	25 (5.0%)	31 (4.1%)
ANTIINFECTIVES FOR TREATMENT OF ACNE	2 (0.8%)	8 (1.6%)	10 (1.3%)
ANTISEPTICS AND DISINFECTANTS	0	2 (0.4%)	2 (0.3%)
ANTIVIRALS	1 (0.4%)	4 (0.8%)	5 (0.7%)
BIGUANIDES AND AMIDINES	0	2 (0.4%)	2 (0.3%)
CARBAMIDE PRODUCTS	0	1 (0.2%)	1 (0.1%)
CORTICOSTEROIDS, COMBINATIONS WITH ANTIBIOTICS	1 (0.4%)	0	1 (0.1%)
CORTICOSTEROIDS, DERMATOLOGICAL PREPARATIONS	0	5 (1.0%)	5 (0.7%)
CORTICOSTEROIDS, MODERAT. POTENT, COMB W/ANTIBIOT.	1 (0.4%)	2 (0.4%)	3 (0.4%)
CORTICOSTEROIDS, MODERATELY POTENT (GROUP II)	0	8 (1.6%)	8 (1.1%)
CORTICOSTEROIDS, MODERATELY POTENT, OTHER COMB.	0	1 (0.2%)	1 (0.1%)
CORTICOSTEROIDS, PLAIN	1 (0.4%)	4 (0.8%)	5 (0.7%)
CORTICOSTEROIDS, POTENT (GROUP III)	7 (2.7%)	46 (9.1%)	53 (7.0%)
CORTICOSTEROIDS, POTENT, COMB WITH ANTIBIOTICS	1 (0.4%)	7 (1.4%)	8 (1.1%)
CORTICOSTEROIDS, POTENT, OTHER COMBINATIONS	0	3 (0.6%)	3 (0.4%)
CORTICOSTEROIDS, VERY POTENT (GROUP IV)	2 (0.8%)	23 (4.6%)	25 (3.3%)
CORTICOSTEROIDS, WEAK (GROUP I)	18 (7.1%)	61 (12.1%)	79 (10.4%)
DERMATOLOGICALS	0	3 (0.6%)	3 (0.4%)
EMOLLIENTS AND PROTECTIVES	0	16 (3.2%)	16 (2.1%)
IMIDAZOLE AND TRIAZOLE DERIVATIVES	7 (2.7%)	29 (5.7%)	36 (4.7%)
IODINE PRODUCTS	3 (1.2%)	2 (0.4%)	5 (0.7%)
OTHER ANTIBIOTICS FOR TOPICAL USE	9 (3.5%)	11 (2.2%)	20 (2.6%)

OTHER ANTIFUNGALS FOR TOPICAL USE	2 (0.8%)	15 (3.0%)	17 (2.2%)
OTHER ANTIPRURITICS	0	4 (0.8%)	4 (0.5%)
OTHER ANTISEPTICS AND DISINFECTANTS	0	4 (0.8%)	4 (0.5%)
OTHER CHEMOTHERAPEUTICS	0	1 (0.2%)	1 (0.1%)
OTHER CICATRIZANTS	2 (0.8%)	8 (1.6%)	10 (1.3%)
OTHER DERMATOLOGICAL PREPARATIONS	0	2 (0.4%)	2 (0.3%)
OTHER DERMATOLOGICALS	3 (1.2%)	7 (1.4%)	10 (1.3%)
OTHER EMOLLIENTS AND PROTECTIVES	6 (2.4%)	43 (8.5%)	49 (6.4%)
PREPARATIONS FOR TREATMENT OF WOUNDS AND ULCERS	0	4 (0.8%)	4 (0.5%)
PROTEOLYTIC ENZYMES	0	1 (0.2%)	1 (0.1%)
QUATERNARY AMMONIUM COMPOUNDS	1 (0.4%)	3 (0.6%)	4 (0.5%)
QUINOLINE DERIVATIVES	0	1 (0.2%)	1 (0.1%)
RETINOIDS FOR TREATMENT OF PSORIASIS	0	1 (0.2%)	1 (0.1%)
SALICYLIC ACID PREPARATIONS	0	6 (1.2%)	6 (0.8%)
SOFT PARAFFIN AND FAT PRODUCTS	4 (1.6%)	8 (1.6%)	12 (1.6%)
SOFT PARAFFIN DRESSINGS	0	1 (0.2%)	1 (0.1%)
SULFONAMIDES	0	4 (0.8%)	4 (0.5%)
ZINC PRODUCTS	0	3 (0.6%)	3 (0.4%)
GENITO URINARY SYSTEM AND SEX HORMONES	18 (7.1%)	51 (10.1%)	69 (9.1%)
3-OXOANDROSTEN (4) DERIVATIVES	1 (0.4%)	0	1 (0.1%)
ALPHA-ADRENORECEPTOR ANTAGONISTS	7 (2.7%)	13 (2.6%)	20 (2.6%)
ANDROGENS AND ESTROGENS	0	1 (0.2%)	1 (0.1%)
ANTIINFLAMMATORY PRODUCTS FOR VAGINAL ADMINISTRAT.	3 (1.2%)	12 (2.4%)	15 (2.0%)
DRUGS USED IN ERECTILE DYSFUNCTION	1 (0.4%)	0	1 (0.1%)
ESTREN DERIVATIVES	1 (0.4%)	1 (0.2%)	2 (0.3%)
ESTROGENS	0	1 (0.2%)	1 (0.1%)
IMIDAZOLE DERIVATIVES	0	1 (0.2%)	1 (0.1%)
NATURAL AND SEMISYNTHETIC ESTROGENS, PLAIN	0	1 (0.2%)	1 (0.1%)
OTHER ANTIINFECTIVES AND ANTISEPTICS	0	1 (0.2%)	1 (0.1%)
OTHER UROLOGICALS	0	9 (1.8%)	9 (1.2%)
PREGNEN (4) DERIVATIVES	0	1 (0.2%)	1 (0.1%)
PROGESTOGENS	1 (0.4%)	2 (0.4%)	3 (0.4%)
PROGESTOGENS AND ESTROGENS, FIXED COMBINATIONS	0	1 (0.2%)	1 (0.1%)
SYMPATHOMIMETICS, LABOUR REPRESSANTS	1 (0.4%)	0	1 (0.1%)
TESTOSTERONE-5-ALPHA REDUCTASE INHIBITORS	2 (0.8%)	5 (1.0%)	7 (0.9%)
URINARY ANTISPASMODICS	2 (0.8%)	7 (1.4%)	9 (1.2%)
URINARY CONCREMENT SOLVENTS	0	2 (0.4%)	2 (0.3%)
MUSCULO-SKELETAL SYSTEM	70 (27.5%)	156 (30.9%)	226 (29.7%)
ACETIC ACID DERIVATIVES AND RELATED SUBSTANCES	26 (10.2%)	40 (7.9%)	66 (8.7%)
ANTIINFL. PREP., NON-STEROIDS FOR TOPICAL USE	1 (0.4%)	1 (0.2%)	2 (0.3%)
ANTIINFLAM/ANTIRHEUM. AGENTS IN COMB.W/CORTICOSTER	0	2 (0.4%)	2 (0.3%)

ANTIINFLAMMATORY AND ANTIRHEUMATIC PRODUCTS	1 (0.4%)	1 (0.2%)	2 (0.3%)
ANTIINFLAMMATORY/ANTIRHEUMATIC PROD.,NON-STEROIDS	1 (0.4%)	2 (0.4%)	3 (0.4%)
BISPHOSPHONATES	12 (4.7%)	19 (3.8%)	31 (4.1%)
COXIBS	1 (0.4%)	2 (0.4%)	3 (0.4%)
FENAMATES	0	1 (0.2%)	1 (0.1%)
OTHER ANTIINFL./ANTIRHEUMATIC AGENTS, NON-STEROIDS	11 (4.3%)	22 (4.4%)	33 (4.3%)
OTHER CENTRALLY ACTING AGENTS	7 (2.7%)	13 (2.6%)	20 (2.6%)
OTHER QUATERNARY AMMONIUM COMPOUNDS	1 (0.4%)	1 (0.2%)	2 (0.3%)
OTHER TOPICAL PRODUCTS FOR JOINT AND MUSCULAR PAIN	0	1 (0.2%)	1 (0.1%)
OXICAMS	4 (1.6%)	4 (0.8%)	8 (1.1%)
PREPARATIONS INHIBITING URIC ACID PRODUCTION	5 (2.0%)	19 (3.8%)	24 (3.2%)
PREPARATIONS W. NO EFFECT ON URIC ACID METABOLISM	1 (0.4%)	1 (0.2%)	2 (0.3%)
PROPIONIC ACID DERIVATIVES	26 (10.2%)	75 (14.9%)	101 (13.3%)
QUININE AND DERIVATIVES	0	2 (0.4%)	2 (0.3%)
TOPICAL PRODUCTS FOR JOINT AND MUSCULAR PAIN	1 (0.4%)	0	1 (0.1%)
NERVOUS SYSTEM	199 (78.0%)	419 (83.0%)	618 (81.3%)
AMIDES	4 (1.6%)	7 (1.4%)	11 (1.4%)
ANALGESICS	1 (0.4%)	2 (0.4%)	3 (0.4%)
ANILIDES	76 (29.8%)	217 (43.0%)	293 (38.6%)
ANTIEPILEPTICS	21 (8.2%)	39 (7.7%)	60 (7.9%)
ANTIVERTIGO PREPARATIONS	1 (0.4%)	1 (0.2%)	2 (0.3%)
BARBITURATES AND DERIVATIVES	0	1 (0.2%)	1 (0.1%)
BARBITURATES, COMBINATIONS	0	1 (0.2%)	1 (0.1%)
BENZODIAZEPINE DERIVATIVES	72 (28.2%)	157 (31.1%)	229 (30.1%)
BENZODIAZEPINE RELATED DRUGS	29 (11.4%)	50 (9.9%)	79 (10.4%)
BENZOMORPHAN DERIVATIVES	1 (0.4%)	1 (0.2%)	2 (0.3%)
BUTYROPHENONE DERIVATIVES	13 (5.1%)	21 (4.2%)	34 (4.5%)
CARBAMATES	0	1 (0.2%)	1 (0.1%)
CARBOXAMIDE DERIVATIVES	0	3 (0.6%)	3 (0.4%)
CENTRALLY ACTING SYMPATHOMIMETICS	0	2 (0.4%)	2 (0.3%)
DIAZEPINES, OXAZEPINES AND THIAZEPINES	0	4 (0.8%)	4 (0.5%)
DIPHENYLMETHANE DERIVATIVES	6 (2.4%)	14 (2.8%)	20 (2.6%)
DIPHENYLPROPYLAMINE DERIVATIVES	3 (1.2%)	7 (1.4%)	10 (1.3%)
DOPA AND DOPA DERIVATIVES	1 (0.4%)	1 (0.2%)	2 (0.3%)
DOPAMINE AGONISTS	0	2 (0.4%)	2 (0.3%)
DRUGS USED IN NICOTINE DEPENDENCE	1 (0.4%)	0	1 (0.1%)
DRUGS USED IN OPIOID DEPENDENCE	1 (0.4%)	2 (0.4%)	3 (0.4%)
FATTY ACID DERIVATIVES	1 (0.4%)	1 (0.2%)	2 (0.3%)
HALOGENATED HYDROCARBONS	2 (0.8%)	1 (0.2%)	3 (0.4%)
HYDANTOIN DERIVATIVES	0	3 (0.6%)	3 (0.4%)
LITHIUM	0	1 (0.2%)	1 (0.1%)
MONOAMINE OXIDASE B INHIBITORS	0	1 (0.2%)	1 (0.1%)
NATURAL OPIUM ALKALOIDS	107 (42.0%)	213 (42.2%)	320 (42.1%)

NERVOUS SYSTEM	2 (0.8%)	0	2 (0.3%)
NON-SELECTIVE MONOAMINE REUPTAKE INHIBITORS	6 (2.4%)	13 (2.6%)	19 (2.5%)
OPIOID ANESTHETICS	52 (20.4%)	76 (15.0%)	128 (16.8%)
OPIOIDS	1 (0.4%)	0	1 (0.1%)
ORIPAVINE DERIVATIVES	4 (1.6%)	6 (1.2%)	10 (1.3%)
OTHER ANALGESICS AND ANTIPYRETICS	8 (3.1%)	29 (5.7%)	37 (4.9%)
OTHER ANTI-DEMENTIA DRUGS	1 (0.4%)	0	1 (0.1%)
OTHER ANTIDEPRESSANTS	7 (2.7%)	25 (5.0%)	32 (4.2%)
OTHER ANTIEPILEPTICS	0	5 (1.0%)	5 (0.7%)
OTHER ANTIPSYCHOTICS	4 (1.6%)	1 (0.2%)	5 (0.7%)
OTHER ANXIOLYTICS	0	1 (0.2%)	1 (0.1%)
OTHER GENERAL ANESTHETICS	3 (1.2%)	7 (1.4%)	10 (1.3%)
OTHER HYPNOTICS AND SEDATIVES	2 (0.8%)	0	2 (0.3%)
OTHER OPIOIDS	37 (14.5%)	79 (15.6%)	116 (15.3%)
OTHER PSYCHOSTIMULANTS AND NOOTROPICS	0	1 (0.2%)	1 (0.1%)
PHENOTHIAZINES WITH ALIPHATIC SIDE-CHAIN	2 (0.8%)	6 (1.2%)	8 (1.1%)
PHENYLPIPERIDINE DERIVATIVES	2 (0.8%)	4 (0.8%)	6 (0.8%)
PSYCHOLEPTICS	0	1 (0.2%)	1 (0.1%)
PYRAZOLONES	17 (6.7%)	43 (8.5%)	60 (7.9%)
SALICYLIC ACID AND DERIVATIVES	0	4 (0.8%)	4 (0.5%)
SELECTIVE SEROTONIN (5HT1) AGONISTS	0	2 (0.4%)	2 (0.3%)
SELECTIVE SEROTONIN REUPTAKE INHIBITORS	19 (7.5%)	37 (7.3%)	56 (7.4%)
TERTIARY AMINES	0	1 (0.2%)	1 (0.1%)
RESPIRATORY SYSTEM	49 (19.2%)	154 (30.5%)	203 (26.7%)
ADRENERGICS AND OTH.DRUGS FOR OBSTRUCT.AIRWAY DIS.	7 (2.7%)	15 (3.0%)	22 (2.9%)
ADRENERGICS, INHALANTS	0	1 (0.2%)	1 (0.1%)
ALPHA- AND BETA-ADRENORECEPTOR AGONISTS	1 (0.4%)	0	1 (0.1%)
ANESTHETICS, LOCAL	4 (1.6%)	5 (1.0%)	9 (1.2%)
ANTIALLERGIC AGENTS, EXCL. CORTICOSTEROIDS	0	1 (0.2%)	1 (0.1%)
ANTICHOLINERGICS	2 (0.8%)	11 (2.2%)	13 (1.7%)
ANTIHISTAMINES FOR SYSTEMIC USE	0	5 (1.0%)	5 (0.7%)
ANTISEPTICS	0	6 (1.2%)	6 (0.8%)
CORTICOSTEROIDS	1 (0.4%)	2 (0.4%)	3 (0.4%)
COUGH AND COLD PREPARATIONS	1 (0.4%)	0	1 (0.1%)
COUGH SUPPRESSANTS AND EXPECTORANTS, COMBINATIONS	1 (0.4%)	0	1 (0.1%)
DECONGESTANTS AND OTHER NASAL PREP.FOR TOPICAL USE	0	1 (0.2%)	1 (0.1%)
DRUGS FOR OBSTRUCTIVE AIRWAY DISEASES	0	1 (0.2%)	1 (0.1%)
EXPECTORANTS	1 (0.4%)	3 (0.6%)	4 (0.5%)
EXPECTORANTS,EXCL COMBINATIONS WITH COUGH SUPPR.	0	1 (0.2%)	1 (0.1%)
LEUKOTRIENE RECEPTOR ANTAGONISTS	2 (0.8%)	1 (0.2%)	3 (0.4%)
MUCOLYTICS	7 (2.7%)	23 (4.6%)	30 (3.9%)
NASAL DECONGESTANTS FOR SYSTEMIC USE	0	1 (0.2%)	1 (0.1%)
OPIUM ALKALOIDS AND DERIVATIVES	17 (6.7%)	35 (6.9%)	52 (6.8%)
OPIUM DERIVATIVES AND EXPECTORANTS	3 (1.2%)	7 (1.4%)	10 (1.3%)
OTHER ANTIHISTAMINES FOR SYSTEMIC USE	6 (2.4%)	24 (4.8%)	30 (3.9%)
OTHER COUGH SUPPRESSANTS	6 (2.4%)	8 (1.6%)	14 (1.8%)
OTHER COUGH SUPPRESSANTS AND EXPECTORANTS	0	1 (0.2%)	1 (0.1%)
OTHER RESPIRATORY SYSTEM PRODUCTS	3 (1.2%)	10 (2.0%)	13 (1.7%)
PHENOTHIAZINE DERIVATIVES	1 (0.4%)	2 (0.4%)	3 (0.4%)
PIPERAZINE DERIVATIVES	4 (1.6%)	30 (5.9%)	34 (4.5%)
RESPIRATORY SYSTEM	0	1 (0.2%)	1 (0.1%)
SELECTIVE BETA-2-ADRENORECEPTOR AGONISTS	7 (2.7%)	19 (3.8%)	26 (3.4%)
SUBSTITUTED ALKYLAMINES	3 (1.2%)	17 (3.4%)	20 (2.6%)
SYMPATHOMIMETICS, PLAIN	2 (0.8%)	1 (0.2%)	3 (0.4%)
XANTHINES	0	1 (0.2%)	1 (0.1%)
SENSORY ORGANS	3 (1.2%)	13 (2.6%)	16 (2.1%)
CARBONIC ANHYDRASE INHIBITORS	0	1 (0.2%)	1 (0.1%)
CORTICOSTEROIDS AND ANTIINFECTIVES IN COMBINATION	0	2 (0.4%)	2 (0.3%)
OTHER ANTIALLERGICS	1 (0.4%)	4 (0.8%)	5 (0.7%)
OTHER OPHTHALMOLOGICAL AND OTOLOGICAL PREPARATIONS	0	1 (0.2%)	1 (0.1%)
OTHER OPHTHALMOLOGICALS	2 (0.8%)	2 (0.4%)	4 (0.5%)
PROSTAGLANDIN ANALOGUES	0	3 (0.6%)	3 (0.4%)
SYST. HORMONAL PREP. , EXCL. SEX HORM. AND INSULIN	22 (8.6%)	58 (11.5%)	80 (10.5%)
ANTIGROWTH HORMONE	7 (2.7%)	12 (2.4%)	19 (2.5%)
GLUCOCORTICOIDS	3 (1.2%)	3 (0.6%)	6 (0.8%)
GLYCOGENOLYTIC HORMONES	0	1 (0.2%)	1 (0.1%)
SULFUR-CONTAINING IMIDAZOLE DERIVATIVES	1 (0.4%)	4 (0.8%)	5 (0.7%)
THYROID HORMONES	12 (4.7%)	37 (7.3%)	49 (6.4%)
VASOPRESSIN AND ANALOGUES	0	1 (0.2%)	1 (0.1%)
UNCLASSIFIABLE	2 (0.8%)	2 (0.4%)	4 (0.5%)
UNCLASSIFIABLE	2 (0.8%)	2 (0.4%)	4 (0.5%)
VARIOUS	239 (93.7%)	484 (95.8%)	723 (95.1%)
ALL OTHER THERAPEUTIC PRODUCTS	0	3 (0.6%)	3 (0.4%)
AMINO ACIDS, INCL. COMBINATIONS WITH POLYPEPTIDES	1 (0.4%)	6 (1.2%)	7 (0.9%)
AMINO ACIDS/CARBOHYDRATES/MINERALS/VITAMINS, COMB	2 (0.8%)	6 (1.2%)	8 (1.1%)
ANTIDOTES	1 (0.4%)	7 (1.4%)	8 (1.1%)
BARIUM SULFATE CONTAINING X-RAY CONTRAST MEDIA	12 (4.7%)	20 (4.0%)	32 (4.2%)
CARBOHYDRATES/PROTEINS/MINERALS/VITAMINS, COMB	1 (0.4%)	0	1 (0.1%)
CONTRAST MEDIA	59 (23.1%)	126 (25.0%)	185 (24.3%)
DRUGS FOR TREATM. OF HYPERKAL. & HYPERPHOSPHAT.	2 (0.8%)	3 (0.6%)	5 (0.7%)
FAT/CARBOHYDRATES/PROTEINS/MINERALS/VITAMINS, COMB	0	3 (0.6%)	3 (0.4%)

OTHER COMBINATIONS OF NUTRIENTS	5 (2.0%)	6 (1.2%)	11 (1.4%)
OTHER DIAGNOSTIC AGENTS	0	1 (0.2%)	1 (0.1%)
OTHER MAGNETIC RESONANCE IMAGING CONTRAST MEDIA	1 (0.4%)	0	1 (0.1%)
OTHER NUTRIENTS	6 (2.4%)	28 (5.5%)	34 (4.5%)
OTHER THERAPEUTIC PRODUCTS	9 (3.5%)	19 (3.8%)	28 (3.7%)
PARAMAGNETIC CONTRAST MEDIA	16 (6.3%)	50 (9.9%)	66 (8.7%)
PROTEIN SUPPLEMENTS	1 (0.4%)	2 (0.4%)	3 (0.4%)
SOLVENTS AND DILUTING AGENTS,INCL IRRIGAT SOLUT	0	1 (0.2%)	1 (0.1%)
SUPERPARAMAGNETIC CONTRAST MEDIA	0	5 (1.0%)	5 (0.7%)
TECHNETIUM (99MTC) COMPOUNDS	0	2 (0.4%)	2 (0.3%)
UNSPECIFIED HERBAL	20 (7.8%)	33 (6.5%)	53 (7.0%)
VARIOUS THYROID DIAGNOSTIC RADIOPHARMACEUTICALS	1 (0.4%)	0	1 (0.1%)
WATERSOL.,NEPHROTROPIC, LOW OSM. X-RAY CONTR.MEDIA	159 (62.4%)	324 (64.2%)	483 (63.6%)
WATERSOL.,NEPHROTROPIC,HIGH OSM. X-RAY CONTR.MEDIA	21 (8.2%)	42 (8.3%)	63 (8.3%)
X-RAY CONTRAST MEDIA, IODINATED	1 (0.4%)	0	1 (0.1%)

Source: CORRECT CSR Table 14.1 / 23

Table A7.3: CONCUR - Number of subjects who took at least one concomitant medication (Full analysis set)

Table 14.1 / 27: Number of subjects who took at least one concomitant medication (Full analysis set)

ATC CLASSIFICATION SUBCLASS WHO-DD Version 3Q2005	Placebo N=68 (100%)	Regorafenib 160 mg N=136 (100%)	Total N=204 (100%)
Number of subjects (%) with at least one concomitant medication	68 (100.0%)	136 (100.0%)	204 (100.0%)
ALIMENTARY TRACT AND METABOLISM	49 (72.1%)	112 (82.4%)	161 (78.9%)
ALIMENTARY TRACT AND METABOLISM	1 (1.5%)	1 (0.7%)	2 (1.0%)
ALPHA GLUCOSIDASE INHIBITORS	0	2 (1.5%)	2 (1.0%)
ALUMINIUM COMPOUNDS	1 (1.5%)	0	1 (0.5%)
AMINO ACIDS AND DERIVATIVES	3 (4.4%)	10 (7.4%)	13 (6.4%)
ANDROSTAN DERIVATIVES	0	1 (0.7%)	1 (0.5%)
ANTACIDS	0	1 (0.7%)	1 (0.5%)
ANTACIDS WITH ANTIFLATULENTS	0	1 (0.7%)	1 (0.5%)
ANTACIDS WITH SODIUM BICARBONATE	0	4 (2.9%)	4 (2.0%)
ANTACIDS, OTHER COMBINATIONS	0	1 (0.7%)	1 (0.5%)
ANTIBIOTICS	0	2 (1.5%)	2 (1.0%)
ANTIDIARRHEAL MICROORGANISMS	1 (1.5%)	1 (0.7%)	2 (1.0%)
ANTIEMETICS AND ANTINAUSEANTS	0	1 (0.7%)	1 (0.5%)
ANTIINFECT. AND ANTISEPT. FOR LOCAL ORAL TREATMENT	2 (2.9%)	12 (8.8%)	14 (6.9%)
ANTIPROPULSIVES	6 (8.8%)	27 (19.9%)	33 (16.2%)
APPETITE STIMULANTS	4 (5.9%)	10 (7.4%)	14 (6.9%)
ASCORBIC ACID (VITAMIN C), PLAIN	5 (7.4%)	5 (3.7%)	10 (4.9%)
BELLADONNA ALKALOIDS SEMISYNT,QUATER AMMONIUM COMP	0	5 (3.7%)	5 (2.5%)
BELLADONNA ALKALOIDS, TERTIARY AMINES	1 (1.5%)	4 (2.9%)	5 (2.5%)
BIGUANIDES	3 (4.4%)	11 (8.1%)	14 (6.9%)
BILE ACID PREPARATIONS	4 (5.9%)	11 (8.1%)	15 (7.4%)
BILE AND LIVER THERAPY	3 (4.4%)	4 (2.9%)	7 (3.4%)
BULK PRODUCERS	1 (1.5%)	1 (0.7%)	2 (1.0%)
CALCIUM	0	2 (1.5%)	2 (1.0%)
CALCIUM, COMBINATIONS WITH OTHER DRUGS	1 (1.5%)	3 (2.2%)	4 (2.0%)
CHARCOAL PREPARATIONS	0	1 (0.7%)	1 (0.5%)
COMB AND COMPL. OF ALUMIN., CALC. AND MAGNES. COMP	2 (2.9%)	11 (8.1%)	13 (6.4%)
COMBINATIONS OF ORAL BLOOD GLUCOSE LOWERING DRUGS	0	1 (0.7%)	1 (0.5%)
COMBINATIONS OF VITAMINS	1 (1.5%)	1 (0.7%)	2 (1.0%)
CONTACT LAXATIVES	6 (8.8%)	18 (13.2%)	24 (11.8%)
CORTICOSTEROIDS ACTING LOCALLY	5 (7.4%)	19 (14.0%)	24 (11.8%)
CORTICOSTEROIDS FOR LOCAL ORAL TREATMENT	12 (17.6%)	19 (14.0%)	31 (15.2%)
DIGESTIVES, INCL. ENZYMES	1 (1.5%)	0	1 (0.5%)
DIPEPTIDYL PEPTIDASE 4 (DPP-4) INHIBITORS	0	2 (1.5%)	2 (1.0%)
DRUGS FOR FUNCTIONAL GASTROINTESTINAL DISORDERS	1 (1.5%)	1 (0.7%)	2 (1.0%)
DRUGS FOR PEPTIC ULCER AND GORD	3 (4.4%)	4 (2.9%)	7 (3.4%)
ELECTROLYTES WITH CARBOHYDRATES	0	1 (0.7%)	1 (0.5%)

ATC CLASSIFICATION SUBCLASS WHO-DD Version 3Q2005	Placebo N=68 (100%)	Regorafenib 160 mg N=136 (100%)	Total N=204 (100%)
ENEMAS	5 (7.4%)	7 (5.1%)	12 (5.9%)
ENZYME PREPARATIONS	2 (2.9%)	4 (2.9%)	6 (2.9%)
H2-RECEPTOR ANTAGONISTS	8 (11.8%)	18 (13.2%)	26 (12.7%)
INSULINS AND ANAL.,INTERM.-ACTING,COMB.W/FAST ACT.	0	2 (1.5%)	2 (1.0%)
INSULINS AND ANALOGUES, FAST-ACTING	5 (7.4%)	5 (3.7%)	10 (4.9%)
INSULINS AND ANALOGUES, LONG-ACTING	1 (1.5%)	2 (1.5%)	3 (1.5%)
INTESTINAL ADSORBENTS	0	1 (0.7%)	1 (0.5%)
LAXATIVES	4 (5.9%)	14 (10.3%)	18 (8.8%)
LIVER THERAPY	13 (19.1%)	25 (18.4%)	38 (18.6%)
MAGNESIUM	0	1 (0.7%)	1 (0.5%)
MAGNESIUM COMPOUNDS	5 (7.4%)	9 (6.6%)	14 (6.9%)
MINERAL SUPPLEMENTS	0	1 (0.7%)	1 (0.5%)
MULTIVITAMINS, COMBINATIONS	0	4 (2.9%)	4 (2.0%)
MULTIVITAMINS, OTHER COMBINATIONS	2 (2.9%)	2 (1.5%)	4 (2.0%)
MULTIVITAMINS, PLAIN	0	3 (2.2%)	3 (1.5%)
ORAL REHYDRATION SALT FORMULATIONS	0	1 (0.7%)	1 (0.5%)
OSMOTICALLY ACTING LAXATIVES	11 (16.2%)	18 (13.2%)	29 (14.2%)
OTHER AGENTS FOR LOCAL ORAL TREATMENT	4 (5.9%)	16 (11.8%)	20 (9.8%)
OTHER ALIMENTARY TRACT AND METABOLISM PRODUCTS	2 (2.9%)	1 (0.7%)	3 (1.5%)
OTHER ANABOLIC AGENTS	1 (1.5%)	1 (0.7%)	2 (1.0%)
OTHER ANTIDIARRHEALS	0	2 (1.5%)	2 (1.0%)
OTHER ANTIEMETICS	0	6 (4.4%)	6 (2.9%)
OTHER DRUGS FOR FUNCTIONAL BOWEL DISORDERS	2 (2.9%)	11 (8.1%)	13 (6.4%)
OTHER DRUGS FOR PEPTIC ULCER AND GORD	1 (1.5%)	7 (5.1%)	8 (3.9%)
OTHER INTESTINAL ADSORBENTS	1 (1.5%)	0	1 (0.5%)
OTHER INTESTINAL ANTIINFECTIVES	0	1 (0.7%)	1 (0.5%)
OTHER MINERAL SUPPLEMENTS	1 (1.5%)	1 (0.7%)	2 (1.0%)
OTHER ORAL BLOOD GLUCOSE LOWERING DRUGS	2 (2.9%)	1 (0.7%)	3 (1.5%)
OTHER PLAIN VITAMIN PREPARATIONS	6 (8.8%)	11 (8.1%)	17 (8.3%)
PAPAVERINE AND DERIVATIVES	0	1 (0.7%)	1 (0.5%)
POTASSIUM	6 (8.8%)	23 (16.9%)	29 (14.2%)
PROPULSIVES	16 (23.5%)	26 (19.1%)	42 (20.6%)
PROTON PUMP INHIBITORS	15 (22.1%)	15 (11.0%)	30 (14.7%)
SELENIUM	1 (1.5%)	0	1 (0.5%)
SEROTONIN (5HT3) ANTAGONISTS	6 (8.8%)	10 (7.4%)	16 (7.8%)
SODIUM	0	1 (0.7%)	1 (0.5%)
SOFTENERS, EMOLLIENTS	1 (1.5%)	1 (0.7%)	2 (1.0%)
STOMATOLOGICAL PREPARATIONS	0	4 (2.9%)	4 (2.0%)
SULFONAMIDES, UREA DERIVATIVES	2 (2.9%)	9 (6.6%)	11 (5.4%)
SYNT ANTICHOLIN,ESTERS WITH TERTIARY AMINO GROUP	1 (1.5%)	2 (1.5%)	3 (1.5%)
SYNT ANTISPASMODICS, AMIDES WITH TERTIARY AMINES	0	1 (0.7%)	1 (0.5%)
THIAZOLIDINEDIONES	0	1 (0.7%)	1 (0.5%)
TONICS	0	5 (3.7%)	5 (2.5%)

ATC CLASSIFICATION SUBCLASS WHO-DD Version 3Q2005	Placebo N=68 (100%)	Regorafenib 160 mg N=136 (100%)	Total N=204 (100%)
VITAMIN B-COMPLEX, OTHER COMBINATIONS	0	1 (0.7%)	1 (0.5%)
VITAMIN B-COMPLEX, PLAIN	1 (1.5%)	5 (3.7%)	6 (2.9%)
VITAMIN B1, PLAIN	2 (2.9%)	1 (0.7%)	3 (1.5%)
VITAMIN D AND ANALOGUES	0	1 (0.7%)	1 (0.5%)
VITAMINS	0	1 (0.7%)	1 (0.5%)
VITAMINS WITH MINERALS	0	1 (0.7%)	1 (0.5%)
VITAMINS, OTHER COMBINATIONS	0	3 (2.2%)	3 (1.5%)
ANTIINFECTIVES FOR SYSTEMIC USE	14 (20.6%)	59 (43.4%)	73 (35.8%)
BETA-LACTAMASE RESISTANT PENICILLINS	0	1 (0.7%)	1 (0.5%)
BETA-LACTAMASE SENSITIVE PENICILLINS	2 (2.9%)	2 (1.5%)	4 (2.0%)
CARBAPENEMS	0	2 (1.5%)	2 (1.0%)
COMB OF PENICILLINS, INCL. BETA-LACTAMASE INHIB.	3 (4.4%)	17 (12.5%)	20 (9.8%)
COMB.SULFONAMIDES & TRIMETHOPRIM INCL. DERIVATIVES	0	1 (0.7%)	1 (0.5%)
COMBINATIONS OF ANTIBACTERIALS	1 (1.5%)	2 (1.5%)	3 (1.5%)
FIRST-GENERATION CEPHALOSPORINS	0	9 (6.6%)	9 (4.4%)
FLUOROQUINOLONES	5 (7.4%)	19 (14.0%)	24 (11.8%)
GLYCOPEPTIDE ANTIBACTERIALS	0	1 (0.7%)	1 (0.5%)
HEPATITIS VACCINES	0	1 (0.7%)	1 (0.5%)
INFLUENZA VACCINES	1 (1.5%)	2 (1.5%)	3 (1.5%)
MACROLIDES	0	6 (4.4%)	6 (2.9%)
MONOBACTAMS	0	1 (0.7%)	1 (0.5%)
NUCLEOSIDE REVERSE TRANSCRIPTASE INHIBITORS	1 (1.5%)	4 (2.9%)	5 (2.5%)
NUCLEOSIDES AND NUCLEOTIDES EXCL REV.TRANSCR.INHIB	0	2 (1.5%)	2 (1.0%)
OTHER AMINOGLYCOSIDES	0	1 (0.7%)	1 (0.5%)
OTHER ANTIMYCOTICS FOR SYSTEMIC USE	0	1 (0.7%)	1 (0.5%)
OTHER BETA-LACTAM ANTIBACTERIALS	1 (1.5%)	0	1 (0.5%)
OTHER QUINOLONES	1 (1.5%)	1 (0.7%)	2 (1.0%)
PENICILLINS WITH EXTENDED SPECTRUM	3 (4.4%)	12 (8.8%)	15 (7.4%)
SECOND-GENERATION CEPHALOSPORINS	4 (5.9%)	11 (8.1%)	15 (7.4%)
THIRD-GENERATION CEPHALOSPORINS	3 (4.4%)	14 (10.3%)	17 (8.3%)
TRIAZOLE DERIVATIVES	0	1 (0.7%)	1 (0.5%)
ANTINEOPLASTIC AND IMMUNOMODULATING AGENTS	14 (20.6%)	19 (14.0%)	33 (16.2%)
ANTHRACYCLINES AND RELATED SUBSTANCES	0	3 (2.2%)	3 (1.5%)
ANTIMETABOLITES	1 (1.5%)	0	1 (0.5%)
ANTINEOPLASTIC AGENTS	2 (2.9%)	1 (0.7%)	3 (1.5%)
COLONY STIMULATING FACTORS	0	5 (3.7%)	5 (2.5%)
CYTOKINES AND IMMUNOMODULATORS	1 (1.5%)	1 (0.7%)	2 (1.0%)
FOLIC ACID ANALOGUES	1 (1.5%)	0	1 (0.5%)
IMMUNOSTIMULANTS	1 (1.5%)	0	1 (0.5%)

ATC CLASSIFICATION SUBCLASS WHO-DD Version 3Q2005	Placebo N=68 (100%)	Regorafenib 160 mg N=136 (100%)	Total N=204 (100%)
INTERLEUKINS	1 (1.5%)	1 (0.7%)	2 (1.0%)
MONOCLONAL ANTIBODIES	0	1 (0.7%)	1 (0.5%)
OTHER ALKYLATING AGENTS	0	1 (0.7%)	1 (0.5%)
OTHER ANTINEOPLASTIC AGENTS	8 (11.8%)	10 (7.4%)	18 (8.8%)
OTHER CYTOKINES AND IMMUNOMODULATORS	5 (7.4%)	1 (0.7%)	6 (2.9%)
PLATINUM COMPOUNDS	1 (1.5%)	1 (0.7%)	2 (1.0%)
PYRIMIDINE ANALOGUES	2 (2.9%)	3 (2.2%)	5 (2.5%)
ANTIPARASITIC PRODUCTS,INSECTICIDES AND REPELLENTS	0	1 (0.7%)	1 (0.5%)
ECTOPARASITICIDES, INCL. SCABICIDES	0	1 (0.7%)	1 (0.5%)
BLOOD AND BLOOD FORMING ORGANS	19 (27.9%)	70 (51.5%)	89 (43.6%)
AMINO ACIDS	4 (5.9%)	12 (8.8%)	16 (7.8%)
BLOOD AND RELATED PRODUCTS	2 (2.9%)	7 (5.1%)	9 (4.4%)
BLOOD SUBSTITUTES AND PLASMA PROTEIN FRACTIONS	5 (7.4%)	9 (6.6%)	14 (6.9%)
ELECTROLYTE SOLUTIONS	0	1 (0.7%)	1 (0.5%)
ENZYMES	0	3 (2.2%)	3 (1.5%)
FOLIC ACID AND DERIVATIVES	1 (1.5%)	1 (0.7%)	2 (1.0%)
HEPARIN GROUP	0	7 (5.1%)	7 (3.4%)
I.V. SOLUTIONS	0	1 (0.7%)	1 (0.5%)
IRON BIVALENT, ORAL PREPARATIONS	0	3 (2.2%)	3 (1.5%)
IRON IN COMBINATION WITH FOLIC ACID	0	1 (0.7%)	1 (0.5%)
IRON TRIVALENT, ORAL PREPARATIONS	1 (1.5%)	1 (0.7%)	2 (1.0%)
OTHER ANTITHROMBOTIC AGENTS	0	1 (0.7%)	1 (0.5%)
OTHER SYSTEMIC HEMOSTATICS	2 (2.9%)	7 (5.1%)	9 (4.4%)
PLATELET AGGREGATION INHIBITORS EXCL. HEPARIN	4 (5.9%)	2 (1.5%)	6 (2.9%)
PROTEINASE INHIBITORS	0	1 (0.7%)	1 (0.5%)
SOLUTIONS AFFECTING THE ELECTROLYTE BALANCE	1 (1.5%)	13 (9.6%)	14 (6.9%)
SOLUTIONS FOR PARENTERAL NUTRITION	9 (13.2%)	20 (14.7%)	29 (14.2%)
SOLUTIONS PRODUCING OSMOTIC DIURESIS	1 (1.5%)	31 (22.8%)	32 (15.7%)
VITAMIN B12 (CYANOCOBALAMIN AND ANALOGUES)	2 (2.9%)	2 (1.5%)	4 (2.0%)
VITAMIN K	2 (2.9%)	1 (0.7%)	3 (1.5%)
VITAMIN K ANTAGONISTS	1 (1.5%)	0	1 (0.5%)
CARDIOVASCULAR SYSTEM	31 (45.6%)	75 (55.1%)	106 (52.0%)
ACE INHIBITORS AND CALCIUM CHANNEL BLOCKERS	0	1 (0.7%)	1 (0.5%)
ACE INHIBITORS, PLAIN	2 (2.9%)	12 (8.8%)	14 (6.9%)
ADRENERGIC AND DOPAMINERGIC AGENTS	2 (2.9%)	2 (1.5%)	4 (2.0%)
ALDOSTERONE ANTAGONISTS	3 (4.4%)	7 (5.1%)	10 (4.9%)
ALPHA-ADRENORECEPTOR ANTAGONISTS	0	5 (3.7%)	5 (2.5%)
ANGIOTENSIN II ANTAGONISTS AND DIURETICS	1 (1.5%)	0	1 (0.5%)
ANGIOTENSIN II ANTAGONISTS, COMBINATIONS	0	1 (0.7%)	1 (0.5%)
ANGIOTENSIN II ANTAGONISTS, PLAIN	3 (4.4%)	9 (6.6%)	12 (5.9%)

ATC CLASSIFICATION SUBCLASS WHO-DD Version 3Q2005	Placebo N=68 (100%)	Regorafenib 160 mg N=136 (100%)	Total N=204 (100%)
ANTIARRHYTHMICS, CLASS I AND III	1 (1.5%)	2 (1.5%)	3 (1.5%)
ANTIARRHYTHMICS, CLASS IB	2 (2.9%)	0	2 (1.0%)
ANTIARRHYTHMICS, CLASS IC	1 (1.5%)	0	1 (0.5%)
ANTIARRHYTHMICS, CLASS III	1 (1.5%)	2 (1.5%)	3 (1.5%)
ANTIHEMORRHOIDALS FOR TOPICAL USE	0	1 (0.7%)	1 (0.5%)
ANTIHYPERTENSIVES	0	2 (1.5%)	2 (1.0%)
BETA BLOCKING AGENTS, NON-SELECTIVE	0	2 (1.5%)	2 (1.0%)
BETA BLOCKING AGENTS, SELECTIVE	5 (7.4%)	15 (11.0%)	20 (9.8%)
BIOFLAVONOIDS	0	1 (0.7%)	1 (0.5%)
DIGITALIS GLYCOSIDES	2 (2.9%)	1 (0.7%)	3 (1.5%)
DIHYDROPYRIDINE DERIVATIVES	10 (14.7%)	36 (26.5%)	46 (22.5%)
FIBRATES	0	1 (0.7%)	1 (0.5%)
HMG COA REDUCTASE INHIBITORS	4 (5.9%)	3 (2.2%)	7 (3.4%)
ORGANIC NITRATES	2 (2.9%)	2 (1.5%)	4 (2.0%)
OTHER ANTIHEMORRHOIDALS FOR TOPICAL USE	0	2 (1.5%)	2 (1.0%)
OTHER CAPILLARY STABILIZING AGENTS	0	1 (0.7%)	1 (0.5%)
OTHER CARDIAC PREPARATIONS	8 (11.8%)	12 (8.8%)	20 (9.8%)
OTHER PERIPHERAL VASODILATORS	1 (1.5%)	1 (0.7%)	2 (1.0%)
OTHER POTASSIUM-SPARING AGENTS	0	1 (0.7%)	1 (0.5%)
PRODUCTS CONTAINING CORTICOSTEROIDS	1 (1.5%)	6 (4.4%)	7 (3.4%)
PRODUCTS CONTAINING LOCAL ANESTHETICS	1 (1.5%)	4 (2.9%)	5 (2.5%)
PROSTAGLANDINS	0	2 (1.5%)	2 (1.0%)
RAUWOLFIA ALKALOIDS	1 (1.5%)	2 (1.5%)	3 (1.5%)
SULFONAMIDES, PLAIN	9 (13.2%)	17 (12.5%)	26 (12.7%)
THIAZIDES, PLAIN	0	3 (2.2%)	3 (1.5%)
VASOPROTECTIVES	2 (2.9%)	1 (0.7%)	3 (1.5%)
XANTHINE DERIVATIVES	1 (1.5%)	3 (2.2%)	4 (2.0%)
DERMATOLOGICALS	13 (19.1%)	45 (33.1%)	58 (28.4%)
ANESTHETICS FOR TOPICAL USE	0	2 (1.5%)	2 (1.0%)
ANTIHISTAMINES FOR TOPICAL USE	2 (2.9%)	5 (3.7%)	7 (3.4%)
ANTIINFECTIVES FOR TREATMENT OF ACNE	0	4 (2.9%)	4 (2.0%)
ANTIPSORIATICS FOR TOPICAL USE	0	1 (0.7%)	1 (0.5%)
ANTIVIRALS	1 (1.5%)	2 (1.5%)	3 (1.5%)
CHEMOTHERAPEUTICS FOR TOPICAL USE	1 (1.5%)	0	1 (0.5%)
CORTICOSTEROIDS, DERMATOLOGICAL PREPARATIONS	1 (1.5%)	3 (2.2%)	4 (2.0%)
CORTICOSTEROIDS, MODERATELY POTENT (GROUP II)	0	1 (0.7%)	1 (0.5%)
CORTICOSTEROIDS, POTENT (GROUP III)	1 (1.5%)	5 (3.7%)	6 (2.9%)
CORTICOSTEROIDS, VERY POTENT (GROUP IV)	2 (2.9%)	6 (4.4%)	8 (3.9%)
CORTICOSTEROIDS, WEAK (GROUP I)	2 (2.9%)	3 (2.2%)	5 (2.5%)
DERMATOLOGICALS	0	1 (0.7%)	1 (0.5%)
EMOLLIENTS AND PROTECTIVES	0	1 (0.7%)	1 (0.5%)
IMIDAZOLE AND TRIAZOLE DERIVATIVES	3 (4.4%)	1 (0.7%)	4 (2.0%)
IODINE PRODUCTS	1 (1.5%)	1 (0.7%)	2 (1.0%)

ATC CLASSIFICATION SUBCLASS WHO-DD Version 3Q2005	Placebo N=68 (100%)	Regorafenib 160 mg N=136 (100%)	Total N=204 (100%)
OTHER ANTIBIOTICS FOR TOPICAL USE	1 (1.5%)	12 (8.8%)	13 (6.4%)
OTHER ANTIFUNGALS FOR TOPICAL USE	0	3 (2.2%)	3 (1.5%)
OTHER ANTIPRURITICS	1 (1.5%)	2 (1.5%)	3 (1.5%)
OTHER ANTISEPTICS AND DISINFECTANTS	0	2 (1.5%)	2 (1.0%)
OTHER CICATRIZANTS	0	1 (0.7%)	1 (0.5%)
OTHER DERMATOLOGICALS	0	1 (0.7%)	1 (0.5%)
OTHER EMOLLIENTS AND PROTECTIVES	0	5 (3.7%)	5 (2.5%)
PHENOL AND DERIVATIVES	0	1 (0.7%)	1 (0.5%)
PREPARATIONS FOR TREATMENT OF WOUNDS AND ULCERS	0	1 (0.7%)	1 (0.5%)
PROTEOLYTIC ENZYMES	0	1 (0.7%)	1 (0.5%)
SALICYLIC ACID PREPARATIONS	0	2 (1.5%)	2 (1.0%)
SOFT PARAFFIN AND FAT PRODUCTS	2 (2.9%)	9 (6.6%)	11 (5.4%)
ZINC PRODUCTS	0	1 (0.7%)	1 (0.5%)
GENITO URINARY SYSTEM AND SEX HORMONES	4 (5.9%)	12 (8.8%)	16 (7.8%)
ALPHA-ADRENORECEPTOR ANTAGONISTS	1 (1.5%)	4 (2.9%)	5 (2.5%)
ANTIINFLAMMATORY PRODUCTS FOR VAGINAL ADMINISTRAT.	1 (1.5%)	5 (3.7%)	6 (2.9%)
ANTISEPTICS AND CORTICOSTEROIDS	1 (1.5%)	0	1 (0.5%)
IMIDAZOLE DERIVATIVES	1 (1.5%)	1 (0.7%)	2 (1.0%)
PROGESTOGENS	0	1 (0.7%)	1 (0.5%)
PROGESTOGENS AND ESTROGENS, FIXED COMBINATIONS	0	1 (0.7%)	1 (0.5%)
URINARY ANTISPASMODICS	0	1 (0.7%)	1 (0.5%)
MUSCULO-SKELETAL SYSTEM	6 (8.8%)	27 (19.9%)	33 (16.2%)
ACETIC ACID DERIVATIVES AND RELATED SUBSTANCES	3 (4.4%)	11 (8.1%)	14 (6.9%)
ANTIINFL. PREP., NON-STEROIDS FOR TOPICAL USE	1 (1.5%)	0	1 (0.5%)
BISPHOSPHONATES	0	7 (5.1%)	7 (3.4%)
CARBAMIC ACID ESTERS	1 (1.5%)	0	1 (0.5%)
FENAMATES	0	4 (2.9%)	4 (2.0%)
OTHER ANTIINFL./ANTIRHEUMATIC AGENTS, NON-STEROIDS	0	1 (0.7%)	1 (0.5%)
OTHER CENTRALLY ACTING AGENTS	2 (2.9%)	4 (2.9%)	6 (2.9%)
OXAZOL, THIAZINE, AND TRIAZINE DERIVATIVES	1 (1.5%)	1 (0.7%)	2 (1.0%)
OXICAMS	2 (2.9%)	2 (1.5%)	4 (2.0%)
PREPARATIONS WITH SALICYLIC ACID DERIVATIVES	0	1 (0.7%)	1 (0.5%)
PROPIONIC ACID DERIVATIVES	1 (1.5%)	1 (0.7%)	2 (1.0%)
NERVOUS SYSTEM	42 (61.8%)	80 (58.8%)	122 (59.8%)
ANALGESICS	2 (2.9%)	3 (2.2%)	5 (2.5%)
ANILIDES	13 (19.1%)	39 (28.7%)	52 (25.5%)
ANTIEPILEPTICS	1 (1.5%)	7 (5.1%)	8 (3.9%)

ATC CLASSIFICATION SUBCLASS WHO-DD Version 3Q2005	Placebo N=68 (100%)	Regorafenib 160 mg N=136 (100%)	Total N=204 (100%)
ANTIVERTIGO PREPARATIONS	1 (1.5%)	0	1 (0.5%)
BENZODIAZEPINE DERIVATIVES	9 (13.2%)	17 (12.5%)	26 (12.7%)
BENZODIAZEPINE RELATED DRUGS	1 (1.5%)	7 (5.1%)	8 (3.9%)
CENTRALLY ACTING SYMPATHOMIMETICS	0	1 (0.7%)	1 (0.5%)
DIAZEPINES, OXAZEPINES AND THIAZEPINES	0	1 (0.7%)	1 (0.5%)
DIPHENYLMETHANE DERIVATIVES	1 (1.5%)	4 (2.9%)	5 (2.5%)
NATURAL OPIUM ALKALOIDS	26 (38.2%)	46 (33.8%)	72 (35.3%)
NON-SELECTIVE MONOAMINE REUPTAKE INHIBITORS	1 (1.5%)	3 (2.2%)	4 (2.0%)
OPIOID ANESTHETICS	5 (7.4%)	12 (8.8%)	17 (8.3%)
ORIPAVINE DERIVATIVES	0	1 (0.7%)	1 (0.5%)
OTHER ANALGESICS AND ANTIPYRETICS	0	3 (2.2%)	3 (1.5%)
OTHER ANTIDEPRESSANTS	0	1 (0.7%)	1 (0.5%)
OTHER OPIOIDS	15 (22.1%)	38 (27.9%)	53 (26.0%)
OTHER PSYCHOSTIMULANTS AND NOOTROPICS	1 (1.5%)	1 (0.7%)	2 (1.0%)
PHENOTHIAZINES WITH ALIPHATIC SIDE-CHAIN	0	1 (0.7%)	1 (0.5%)
PHENYLPIPERIDINE DERIVATIVES	0	4 (2.9%)	4 (2.0%)
PYRAZOLONES	1 (1.5%)	1 (0.7%)	2 (1.0%)
SALICYLIC ACID AND DERIVATIVES	0	4 (2.9%)	4 (2.0%)
SELECTIVE SEROTONIN REUPTAKE INHIBITORS	0	1 (0.7%)	1 (0.5%)
RESPIRATORY SYSTEM	18 (26.5%)	50 (36.8%)	68 (33.3%)
ADRENERGICS AND OTH.DRUGS FOR OBSTRUC.AIRWAY DISEA	0	1 (0.7%)	1 (0.5%)
ADRENERGICS AND OTH.DRUGS FOR OBSTRUCT.AIRWAY DIS.	2 (2.9%)	4 (2.9%)	6 (2.9%)
ANESTHETICS, LOCAL	4 (5.9%)	6 (4.4%)	10 (4.9%)
ANTISEPTICS	1 (1.5%)	1 (0.7%)	2 (1.0%)
COUGH AND COLD PREPARATIONS	0	2 (1.5%)	2 (1.0%)
EXPECTORANTS	2 (2.9%)	8 (5.9%)	10 (4.9%)
EXPECTORANTS,EXCL COMBINATIONS WITH COUGH SUPPR.	1 (1.5%)	0	1 (0.5%)
LEUKOTRIENE RECEPTOR ANTAGONISTS	1 (1.5%)	1 (0.7%)	2 (1.0%)
MUCOLYTICS	0	6 (4.4%)	6 (2.9%)
NASAL DECONGESTANTS FOR SYSTEMIC USE	0	1 (0.7%)	1 (0.5%)
OPIUM ALKALOIDS AND DERIVATIVES	6 (8.8%)	14 (10.3%)	20 (9.8%)
OPIUM DERIVATIVES AND EXPECTORANTS	2 (2.9%)	11 (8.1%)	13 (6.4%)
OTHER ANTIHISTAMINES FOR SYSTEMIC USE	3 (4.4%)	11 (8.1%)	14 (6.9%)
OTHER COUGH SUPPRESSANTS	0	3 (2.2%)	3 (1.5%)
OTHER NASAL PREPARATIONS	1 (1.5%)	6 (4.4%)	7 (3.4%)
PIPERAZINE DERIVATIVES	0	9 (6.6%)	9 (4.4%)
RESPIRATORY STIMULANTS	0	1 (0.7%)	1 (0.5%)
SELECTIVE BETA-2-ADRENORECEPTOR AGONISTS	2 (2.9%)	4 (2.9%)	6 (2.9%)
SUBSTITUTED ALKYLAMINES	5 (7.4%)	18 (13.2%)	23 (11.3%)
THROAT PREPARATIONS	0	1 (0.7%)	1 (0.5%)

ATC CLASSIFICATION SUBCLASS WHO-DD Version 3Q2005	Placebo N=68 (100%)	Regorafenib 160 mg N=136 (100%)	Total N=204 (100%)
XANTHINES	1 (1.5%)	2 (1.5%)	3 (1.5%)
SENSORY ORGANS	0	3 (2.2%)	3 (1.5%)
CORTICOSTEROIDS AND ANTIINFECTIVES IN COMBINATION	0	1 (0.7%)	1 (0.5%)
OPHTHALMOLOGICALS	0	1 (0.7%)	1 (0.5%)
SYMPATHOMIMETICS USED AS DECONGESTANTS	0	1 (0.7%)	1 (0.5%)
SYST. HORMONAL PREP., EXCL. SEX HORM. AND INSULIN	3 (4.4%)	19 (14.0%)	22 (10.8%)
ANTIGROWTH HORMONE	0	5 (3.7%)	5 (2.5%)
GLUCOCORTICOIDS	1 (1.5%)	1 (0.7%)	2 (1.0%)
SYST. HORMONAL PREP., EXCL. SEX HORM. AND INSULIN	0	1 (0.7%)	1 (0.5%)
THYROID HORMONES	1 (1.5%)	12 (8.8%)	13 (6.4%)
VASOPRESSIN AND ANALOGUES	1 (1.5%)	1 (0.7%)	2 (1.0%)
UNCLASSIFIABLE	1 (1.5%)	0	1 (0.5%)
UNCLASSIFIABLE	1 (1.5%)	0	1 (0.5%)
VARIOUS	51 (75.0%)	108 (79.4%)	159 (77.9%)
ALL OTHER THERAPEUTIC PRODUCTS	6 (8.8%)	12 (8.8%)	18 (8.8%)
AMINO ACIDS, INCL. COMBINATIONS WITH POLYPEPTIDES	1 (1.5%)	0	1 (0.5%)
ANTIDOTES	7 (10.3%)	15 (11.0%)	22 (10.8%)
BARIUM SULFATE CONTAINING X-RAY CONTRAST MEDIA	4 (5.9%)	11 (8.1%)	15 (7.4%)
CONTRAST MEDIA	1 (1.5%)	1 (0.7%)	2 (1.0%)
DRUGS FOR TREATM. OF HYPERKAL. & HYPERPHOSPHAT.	0	1 (0.7%)	1 (0.5%)
NON-WATERSOLUBLE X-RAY CONTRAST MEDIA	0	2 (1.5%)	2 (1.0%)
OTHER COMBINATIONS OF NUTRIENTS	1 (1.5%)	4 (2.9%)	5 (2.5%)
OTHER DIAGNOSTIC AGENTS	0	1 (0.7%)	1 (0.5%)
OTHER NUTRIENTS	0	1 (0.7%)	1 (0.5%)
OTHER THERAPEUTIC PRODUCTS	15 (22.1%)	44 (32.4%)	59 (28.9%)
PARAMAGNETIC CONTRAST MEDIA	1 (1.5%)	3 (2.2%)	4 (2.0%)
UNSPECIFIED HERBAL	9 (13.2%)	17 (12.5%)	26 (12.7%)
WATERSOL.,NEPHROTROPIC, LOW OSM. X-RAY CONTR.MEDIA	38 (55.9%)	88 (64.7%)	126 (61.8%)
WATERSOL.,NEPHROTROPIC,HIGH OSM. X-RAY CONTR.MEDIA	5 (7.4%)	5 (3.7%)	10 (4.9%)

Medications that are ongoing at, began after the start of study drug, or medications that were started after end of study drug, are included in this table.
Multiple ATC codes per drug are possible. Therefore, the same drug may be counted in more than one category for the same subject.
Global Biostatistics: /by-sasp/patdb/projects/734506/15808/stat/prod_interim_01/pgms/t-cm-10.sas etget 18JUL2014 19:03
End of table

Source: CONCUR CSR Table 14.1 / 27

A 8. Priority question: The decision problem defines the population as "Adults with metastatic colorectal cancer who have failed on first-line chemotherapy/first-line biologic and who are being considered for ≥ 3rd-line treatment." (Table 1, p. 14). Please provide further details on the second line of therapy and how it relates to the choice of subsequent treatments. Specifically, please indicate how this choice would limit patients to be only eligible for T/T.

NICE's treatment pathway for managing metastatic colorectal cancer is shown in figure A8.1. Several treatments are available as first-line chemotherapy/biological options.

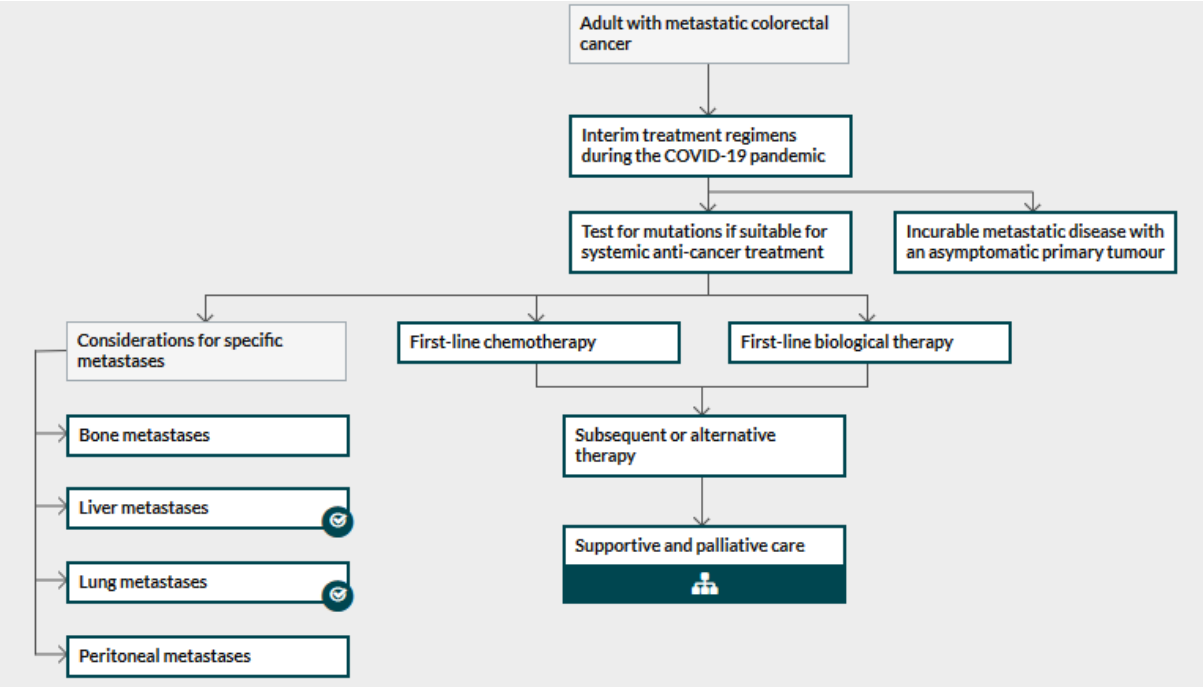
When a patient has failed on, or is not eligible for the first-line chemotherapy/biological therapies, their next step in the treatment pathway is

“Subsequent or alternative” therapy (In this pathway “Subsequent or alternative” is synonymous with ≥3rd-line treatment as patients may have received multiple lines of treatment within what is categorised as first-line chemotherapy/ first-line biological therapy). Patients move to this part of the treatment pathway when prior options are exhausted.

The only recommended “Subsequent or alternative therapy” which is in the final scope is trifluridine/tipiracil. Patients will be considered for trifluridine/tipiracil provided they are ‘fit’ enough and provided they are able to tolerate the side effects of chemotherapy. A patient’s ability to tolerate chemotherapy-related side effects is informed by their experience of chemotherapy in the past.

‘Nivolumab and ipilimumab’ and ‘Encorafenib plus cetuximab’, are also “Subsequent or alternative therapies”. However, these were not included in the final scope as these treatments are only used in the presence of specific mutations. If the mutations are present they are always used ahead of trifluridine/tipiracil (and similarly would be used ahead of regorafenib).

Figure A8.1 – Managing Metastatic colorectal cancer



Source: https://pathways.nice.org.uk/pathways/colorectal-cancer/managing-metastatic-colorectal-cancer (Accessed November 2021)

A 9. Priority question: The company stated that "There was some post-progression treatment in the CORRECT and CONCUR trials (CORRECT: Regorafenib 26%, BSC 30%; CONCUR Regorafenib 31%, BSC 43%)." (p. 142). Also, a scenario analysis of the economic evaluation with subsequent treatment included was conducted. However, Figure 1 in the CS suggests that only BSC would follow T/T or regorafenib and the CS states *"… clinical experts have advised that in England and Wales, patients receiving regorafenib or trifluridine/tipiracil are unlikely to receive further active treatment after progression due to the advanced nature of the disease and limited treatment options available."* (p. 142)

a) Please provide the details of any post-progression treatment provided in the two trials.

CORRECT

Appendix D (Section B.5.1.2 – page 79) provides details on the systemic anti-cancer therapy received during follow-up in the CORRECT trial. This information is repeated below. In addition, this has been supplemented with more granular detail (Table A9.2).

Systemic anti-cancer therapy during follow-up (up to primary data cut-off)

Following discontinuation of study drug patients could receive systemic anti-cancer treatment as determined by the treating physician.

Slightly more patients in the placebo + BSC group received systemic anti-cancer therapy during follow-up compared with patients in the regorafenib + BSC group (29.8% vs 25.9%). The most common (≥5% patients in total) agents were: pyrimidine analogues (20.4% in the placebo + BSC group vs 18.6% in the regorafenib + BSC group), other cytotoxic antibiotics (11.4% vs 7.5%), monoclonal antibodies (8.6% 7.7%), folic acid and derivatives (8.2% vs 5.5%), and platinum compounds (5.5% vs 6.9%).

For patients with a KRAS mutation reported at baseline, more patients in the placebo + BSC group received anti-cancer therapy during follow-up compared with patients in the regorafenib + BSC group (52/157, 33.1% vs 68/273, 24.9%). The most common

(≥5% patients in total) agents were: pyrimidine analogues (22.3% in the placebo + BSC group vs 19.4% in the regorafenib + BSC group), other cytotoxic antibiotics (12.7% vs 8.1%), monoclonal antibodies (7.0% 6.2%), platinum compounds (6.4% vs 5.9%), and folic acid and derivatives (8.3% vs 4.0%).

For patients with no KRAS mutation reported at baseline, fewer patients in the placebo + BSC group received anti-cancer therapy during follow-up compared with patients in the regorafenib group (24/94, 25.5% vs 60/205, 29.3%). The most common (≥5% patients in total) agents were: pyrimidine analogues (18.1% in the placebo + BSC group vs 18.5% in the regorafenib + BSC group), monoclonal antibodies (11.7% 10.7%), folic acid and derivatives (8.5% vs 8.3%), other cytotoxic antibiotics (9.6% vs 7.3%), and platinum compounds (4.3% vs 9.3%).

Table A9.1. Systemic anti-cancer therapy during follow-up (ITT) (Source: Table 41 Appendix D)

Table 8-9 Systemic anti-cancer therapy during follow-up (ITT)

	Placebo + BSC n (%)	Regorafenib + BSC n (%)
Number of subjects (%) with at least one medication	76/255 (29.8)	131/505 (25.9)
Baseline KRAS Mutation		
No	24/94 (25.5)	60/205 (29.3)
Yes	52/157 (33.1)	68/273 (24.9)
Unknown	0/4	3/27 (11.1)

Table A9.2 provides more granular detail on the systemic anti-cancer treat received to the data cut date of 21 July 2011.

Table A9.2: Systemic anti-cancer therapy during follow-up (ITT analysis: data cut-off 21 July 2011)

ATC CLASSIFICATION SUBCLASS WHO-DD Version 3q2005	Placebo N=255 (100%)	Regorafenib 160 mg N=505 (100%)	Total N=760 (100%)
Number of subjects (%) with at least one medication	76 (29.8%)	131 (25.9%)	207 (27.2%)
ANTIINFECTIVES FOR SYSTEMIC USE	1 (0.4%)	1 (0.2%)	2 (0.3%)
VACCINES	1 (0.4%)	1 (0.2%)	2 (0.3%)
ANTINEOPLASTIC AND IMMUNOMODULATING AGENTS	74 (29.0%)	130 (25.7%)	204 (26.8%)
ANTHRACYCLINES AND RELATED SUBSTANCES	0	2 (0.4%)	2 (0.3%)
ANTIMETABOLITES	0	5 (1.0%)	5 (0.7%)
ANTINEOPLASTIC AGENTS	7 (2.7%)	11 (2.2%)	18 (2.4%)
COMBINATIONS OF ANTINEOPLASTIC AGENTS	4 (1.6%)	0	4 (0.5%)
FOLIC ACID ANALOGUES	2 (0.8%)	3 (0.6%)	5 (0.7%)
MONOCLONAL ANTIBODIES	22 (8.6%)	39 (7.7%)	61 (8.0%)
NITROGEN MUSTARD ANALOGUES	0	3 (0.6%)	3 (0.4%)
OTHER ALKYLATING AGENTS	1 (0.4%)	3 (0.6%)	4 (0.5%)
OTHER ANTINEOPLASTIC AGENTS	18 (7.1%)	15 (3.0%)	33 (4.3%)
OTHER CYTOKINES AND IMMUNOMODULATORS	1 (0.4%)	0	1 (0.1%)
OTHER CYTOTOXIC ANTIBIOTICS	29 (11.4%)	38 (7.5%)	67 (8.8%)
OTHER PLANT ALKALOIDS AND NATURAL PRODUCTS	1 (0.4%)	0	1 (0.1%)
PLATINUM COMPOUNDS	14 (5.5%)	35 (6.9%)	49 (6.4%)
PROTEIN KINASE INHIBITORS	0	3 (0.6%)	3 (0.4%)
PYRIMIDINE ANALOGUES	52 (20.4%)	94 (18.6%)	146 (19.2%)
BLOOD AND BLOOD FORMING ORGANS	21 (8.2%)	28 (5.5%)	49 (6.4%)
FOLIC ACID AND DERIVATIVES	21 (8.2%)	28 (5.5%)	49 (6.4%)
DERMATOLOGICALS	1 (0.4%)	0	1 (0.1%)
IMIDAZOLE AND TRIAZOLE DERIVATIVES	1 (0.4%)	0	1 (0.1%)
MUSCULO-SKELETAL SYSTEM	1 (0.4%)	1 (0.2%)	2 (0.3%)
BISPHOSPHONATES	1 (0.4%)	1 (0.2%)	2 (0.3%)
UNCLASSIFIABLE	2 (0.8%)	2 (0.4%)	4 (0.5%)
INVESTIGATIONAL DRUG	0	2 (0.4%)	2 (0.3%)
UNCLASSIFIABLE	2 (0.8%)	0	2 (0.3%)

Source – CSR 28Mar12 table 14.1/21

<u>Systemic anti-cancer therapy during safety follow-up (including period after primary data cut-off)</u>

The use of systemic anti-cancer therapy during follow-up overall was slightly increased: compared to that reported in CSR A53306 dated 19 MAR 2012: 164 (32.5%) patients in the regorafenib + BSC group, 1 (25.0%) patient in the placebo-regorafenib + BSC crossover group, and 80 (31.9%) patients in the placebo + BSC group in the current analysis vs 131 (25.9%) patients in the regorafenib + BSC group and 76 (29.8%) patients in the placebo + BSC group in CSR A53306 dated 19 MAR 2012.

The most common (≥5% patients in total) agents were similar to those reported previously in CSR A53306 dated 19 MAR 2012: pyrimidine analogues (114 patients, 22.6% in the regorafenib + BSC group; 1 patient, 25.0% in the placebo-regorafenib + BSC crossover group; and 54 patients, 21.5% in the placebo + BSC group [94 patients, 18.6% in the regorafenib + BSC group and 52 patients, 20.4% in the

placebo + BSC group]), monoclonal antibodies (57 patients, 11.3%; 1 patient, 25.0%; and 23 patients, 9.2% [39 patients, 7.7% in the regorafenib + BSC group and 22 patients, 8.6% in the placebo + BSC group]), other cytotoxic antibiotics (44 patients, 8.7%; 0 patients, 0%; and 30 patients, 12.0% [38 patients, 7.5% in the regorafenib + BSC group and 29 patients, 11.4% in the placebo + BSC group]), folic acid and derivatives (38 patients, 7.5%; 1 patient, 25.0%; and 22 patients, 8.8% [28 patients, 5.5% in the regorafenib + BSC group and 21 patients, 8.2% in the placebo + BSC group]), platinum compounds (44 patients, 8.7%; 0 patients, 0%; and 16 patients, 6.4% [35 patients, 6.9% in the regorafenib + BSC group and 14 patients, 5.5% in the placebo + BSC group]), and other antineoplastic agents (28 patients, 5.5%; 0 patients, 0%; and 18 patients, 7.2% [15 patients, 3.0% in the regorafenib + BSC group and 18 patients, 7.1% in the placebo + BSC group]).

CONCUR

Appendix D (Section B.5.2.2 – page 93) provides details on the systemic anti-cancer therapy received during follow-up in the CONCUR trial. This information is repeated below. Table A9.4 provides more granular detail on the systemic anti-cancer treat received to the data cut date of 29Nov13.

<u>Systemic anti-cancer therapy during follow-up (up to primary data cut-off)</u>

Systemic anti-cancer therapy during follow-up is summarized by prior targeted. There was an imbalance in systemic anti-cancer therapy use during follow-up between treatment groups in the prior anti-EGFR treatment subgroup but not the prior anti-VEGF subgroup. The difference between the treatment groups was small when subjects had not received targeted therapy during follow-up (placebo 23.1%, regorafenib 28.6%). The difference was more pronounced in the subgroups where subjects did receive some type of targeted anticancer therapy during follow up: prior anti-VEGF but no anti-EGFR – placebo 53.8% versus regorafenib 18.8%; prior anti-EGFR but no anti-VEGF – placebo 52.9% versus regorafenib 33.3%; and prior anti-VEGF treatment and prior anti-EGFR treatment – placebo 54.8% versus 32.5%. Sample sizes were small across all the subgroups.

Table A9.3. Systemic anti-cancer therapy during follow-up – (FAS) (Source: table 48 appendix D

Table 8–11: Systemic anti-cancer therapy during follow-up – (FAS)

Subgroup by prior targeted therapy	Placebo (N=68)	Regorafenib (N=136)
No prior targeted therapy (neither anti-VEGF nor anti-EGFR)	6/26 (23.1%)	16/56 (28.6%)
Prior anti-VEGF treatment but no prior anti-EGFR	7/13 (53.8%)	6/32 (18.8%)
Prior anti-EGFR treatment but no prior anti-VEGF	9/17 (52.9%)	8/24 (33.3%)
Prior anti-VEGF treatment and prior anti-EGFR	7/12 (58.3%)	12/24 (50.0%)
Any prior targeted therapy (anti-VEGF, anti-EGFR, or both)	23/42 (54.8%)	26/80 (32.5%)
Without prior anti-VEGF therapy (with or without anti-EGFR)	15/43 (34.9%)	24/80 (30.0%)
With prior anti-VEGF therapy (with or without anti-EGFR)	14/25 (56.0%)	18/56 (32.1%)
Total across all subgroups during Follow-up	29/68 (42.6%)	42/136 (30.9%)

Abbreviations: EGFR = endothelial growth factor receptor; FAS = full analysis set; FU = follow-up; N = number of subjects; VEGF = vascular endothelial growth factor
a The source tables contain details on all anti-neoplastic and immunomodulating agents used.
Note: Subjects may have had more than one entry.

Table A9.4 provides details on systemic anti-cancer therapy irrespective of prior targeted therapy.

Table A9.4. Systemic anti-cancer therapy during follow-up (full analysis set)

ATC CLASSIFICATION SUBCLASS WHO-DD Version 3q2005	Placebo N=68 (100%)	Regorafenib 160 mg N=136 (100%)	Total N=204 (100%)
Number of subjects (%) with at least one medication	29 (42.6%)	42 (30.9%)	71 (34.8%)
ANTINEOPLASTIC AND IMMUNOMODULATING AGENTS	27 (39.7%)	39 (28.7%)	66 (32.4%)
ANTIMETABOLITES	8 (11.8%)	5 (3.7%)	13 (6.4%)
ANTINEOPLASTIC AGENTS	0	5 (3.7%)	5 (2.5%)
FOLIC ACID ANALOGUES	10 (14.7%)	10 (7.4%)	20 (9.8%)
MONOCLONAL ANTIBODIES	11 (16.2%)	15 (11.0%)	26 (12.7%)
OTHER ANTINEOPLASTIC AGENTS	7 (10.3%)	17 (12.5%)	24 (11.8%)
OTHER CYTOTOXIC ANTIBIOTICS	0	4 (2.9%)	4 (2.0%)
PLATINUM COMPOUNDS	10 (14.7%)	14 (10.3%)	24 (11.8%)
PROTEIN KINASE INHIBITORS	2 (2.9%)	2 (1.5%)	4 (2.0%)
PYRIMIDINE ANALOGUES	14 (20.6%)	23 (16.9%)	37 (18.1%)
BLOOD AND BLOOD FORMING ORGANS	4 (5.9%)	11 (8.1%)	15 (7.4%)
FOLIC ACID AND DERIVATIVES	4 (5.9%)	11 (8.1%)	15 (7.4%)
VARIOUS	4 (5.9%)	5 (3.7%)	9 (4.4%)
OTHER THERAPEUTIC PRODUCTS	4 (5.9%)	5 (3.7%)	9 (4.4%)

Subject may have more than one entry .

Source table: 14.1/24

<u>Systemic anti-cancer therapy during follow-up (including period after primary data cut-off)</u>

Twenty-nine (42.6%) subjects in the placebo group and 45 (33.1%) subjects [42 (30.9%) subjects] in the regorafenib group received systemic anti-cancer therapy during follow-up. The most common systemic anti-cancer therapy reported during

follow-up based on ATC classification in both treatment groups was antineoplastic and immunomodulating agents (39.7% of subjects in the placebo group and 30.9% of subjects [28.7% of subjects] in the regorafenib group).

Table A9.5 provides more granular detail.

Table A9.5. Systemic anti-cancer therapy during follow-up (Full analysis set)

ATC CLASSIFICATION SUBCLASS WHO-DD Version 3q2005	Placebo N=68 (100%)	Regorafenib 160 mg N=136 (100%)	Total N=204 (100%)
Number of subjects (%) with at least one medication	29 (42.6%)	45 (33.1%)	74 (36.3%)
ALIMENTARY TRACT AND METABOLISM	0	1 (0.7%)	1 (0.5%)
VITAMIN B1, PLAIN	0	1 (0.7%)	1 (0.5%)
ANTINEOPLASTIC AND IMMUNOMODULATING AGENTS	27 (39.7%)	42 (30.9%)	69 (33.8%)
ANTIMETABOLITES	9 (13.2%)	6 (4.4%)	15 (7.4%)
ANTINEOPLASTIC AGENTS	0	4 (2.9%)	4 (2.0%)
FOLIC ACID ANALOGUES	11 (16.2%)	11 (8.1%)	22 (10.8%)
MONOCLONAL ANTIBODIES	11 (16.2%)	16 (11.8%)	27 (13.2%)
OTHER ANTINEOPLASTIC AGENTS	8 (11.8%)	19 (14.0%)	27 (13.2%)
OTHER CYTOTOXIC ANTIBIOTICS	0	4 (2.9%)	4 (2.0%)
PLATINUM COMPOUNDS	11 (16.2%)	15 (11.0%)	26 (12.7%)
PROTEIN KINASE INHIBITORS	2 (2.9%)	4 (2.9%)	6 (2.9%)
PYRIMIDINE ANALOGUES	14 (20.6%)	23 (16.9%)	37 (18.1%)
TAXANES	0	1 (0.7%)	1 (0.5%)
BLOOD AND BLOOD FORMING ORGANS	4 (5.9%)	11 (8.1%)	15 (7.4%)
FOLIC ACID AND DERIVATIVES	4 (5.9%)	11 (8.1%)	15 (7.4%)
DERMATOLOGICALS	0	1 (0.7%)	1 (0.5%)
OTHER ANTIPSORIATICS FOR TOPICAL USE	0	1 (0.7%)	1 (0.5%)
VARIOUS	5 (7.4%)	5 (3.7%)	10 (4.9%)
OTHER THERAPEUTIC PRODUCTS	5 (7.4%)	5 (3.7%)	10 (4.9%)

Subject may have more than one entry .

Source: CSR addendum Table 14.1/10

b) Please clarify the extent to which subsequent active therapy reflects UK clinical practice.

There is a small minority of patients who, after failure of the last line of recommended treatment, might be both 'fit' enough and able to tolerate further anti-cancer treatment. As no further recommended line of treatment exists, this would likely be re-challenge with previous failed lines of therapy for which benefit had been observed in the past. Some patients might be considered for enrolment in clinical trials. Experts consulted by Bayer suggested that <10% of patients would be fit enough for subsequent active treatment.

We do not have details regarding active treatment beyond trifluridine/tipiracil in RECOURSE, Yoshino or TERRA. There is data from the named patient programme

for trifluridine/tipiracil prior to its licensing in 2016 (Iverson 2020 https://doi.org/10.1186/s12885-020-6577-1). This publication relates to a selected cohort of patients who were considered fit enough for further treatment – as a consequence of this being a selected group the proportions receiving 4/5th/… line of active therapy cannot be considered representative of the population of all patients failing the last recommended line of therapy.

Data is provided on 4th-line treatment and beyond for the patient group who received trifluridine/tipiracil and for those who didn't. Figure A9.1 shows the proportions of patients receiving different therapies. It is not known the extent to which this data from a clinician-selected group of patients represents current treatment practices.

Figure A9.1. Prior treatment lines in patients with mCRC enrolled in UK trifluridine/tipiracil named patient programme.

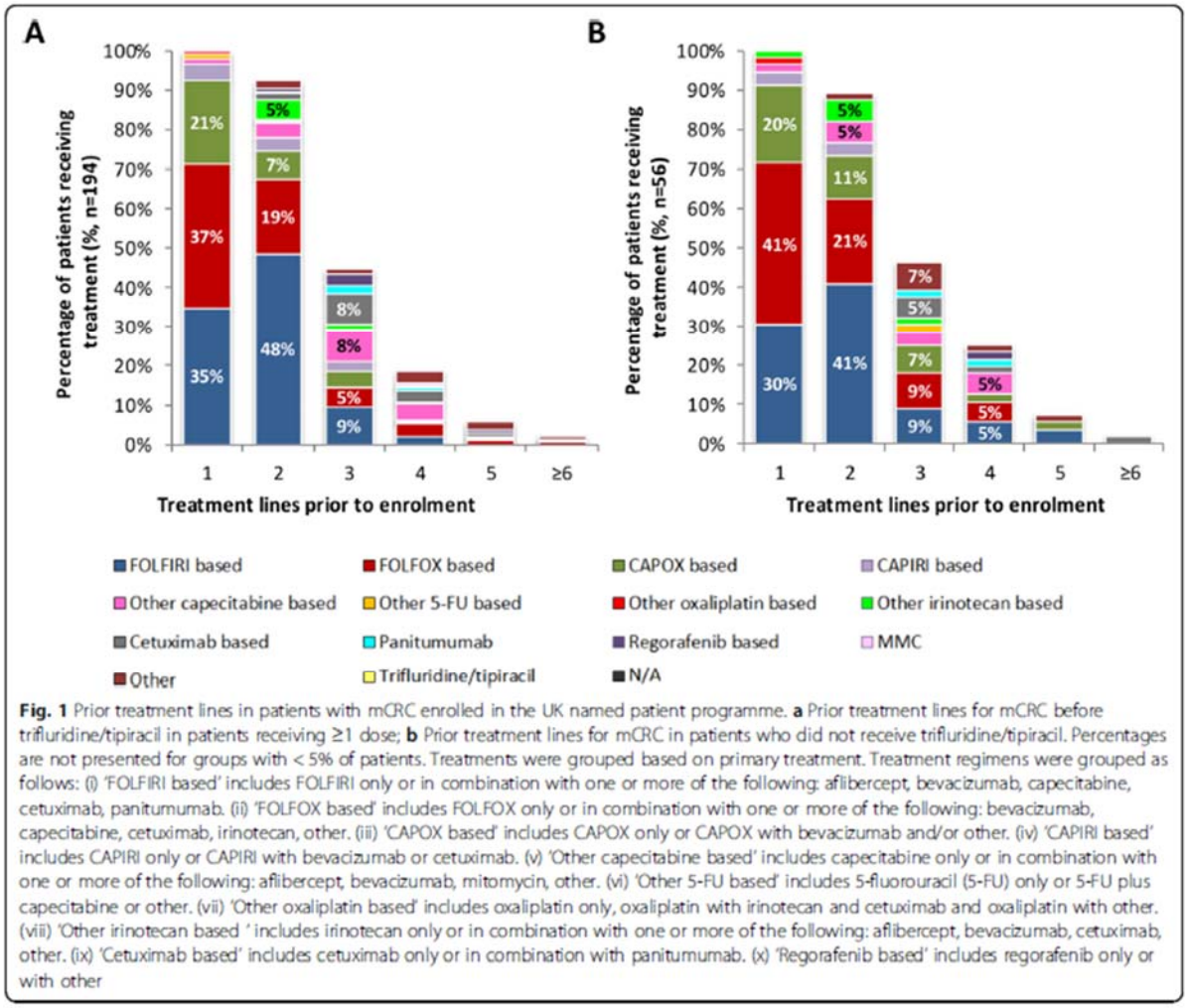


Fig. 1 Prior treatment lines in patients with mCRC enrolled in the UK named patient programme. **a** Prior treatment lines for mCRC before trifluridine/tipiracil in patients receiving ≥1 dose; **b** Prior treatment lines for mCRC in patients who did not receive trifluridine/tipiracil. Percentages are not presented for groups with < 5% of patients. Treatments were grouped based on primary treatment. Treatment regimens were grouped as follows: (i) 'FOLFIRI based' includes FOLFIRI only or in combination with one or more of the following: aflibercept, bevacizumab, capecitabine, cetuximab, panitumumab. (ii) 'FOLFOX based' includes FOLFOX only or in combination with one or more of the following: bevacizumab, capecitabine, cetuximab, irinotecan, other. (iii) 'CAPOX based' includes CAPOX only or CAPOX with bevacizumab and/or other. (iv) 'CAPIRI based' includes CAPIRI only or CAPIRI with bevacizumab or cetuximab. (v) 'Other capecitabine based' includes capecitabine only or in combination with one or more of the following: aflibercept, bevacizumab, mitomycin, other. (vi) 'Other 5-FU based' includes 5-fluorouracil (5-FU) only or 5-FU plus capecitabine or other. (vii) 'Other oxaliplatin based' includes oxaliplatin only, oxaliplatin with irinotecan and cetuximab and oxaliplatin with other. (viii) 'Other irinotecan based ' includes irinotecan only or in combination with one or more of the following: aflibercept, bevacizumab, cetuximab, other. (ix) 'Cetuximab based' includes cetuximab only or in combination with panitumumab. (x) 'Regorafenib based' includes regorafenib only or with other

c) Please conduct an analysis of OS and PFS in CORRECT and CONCUR to estimate the effect of subsequent systemic anti-cancer treatment.

PFS

In CORRECT and CONCUR, patients in both arms were treated with Regorafenib + BSC or Placebo + BSC until progression. After progression, patients could receive further anti-cancer treatment (i.e. post-progression treatment) at the discretion of their physician. As a result of treatment being initiated after progression there is no effect on progression free survival.

OS

We have not conducted any analyses to estimate the effect of post-progression treatment on OS. After investigating this question we consider that any attempts to isolate and quantify the effect of post-progression treatment would be flawed. We do however report below a post-hoc exploratory analysis considering post progression treatment.

Investigations

In the CONCUR CSR there is a post-hoc exploratory analysis considering post-progression treatment – this is presented below. No equivalent analyses are available from CORRECT.

CONCUR – Analysis of overall survival censored at the start of new anti-cancer treatment

To investigate the potential impact of post study anti-cancer on the OS HR for regorafenib, exploratory analyses of OS with censoring at the start of new anti-cancer treatment were performed. The post-hoc analysis of OS confirmed that subjects treated with regorafenib had a prolongation of OS when compared with the placebo group (Table A9.6). The estimated HR for the FAS was 0.413 (95% CI: 0.274 to 0.623).

Table A9.6 Overall survival – censoring at new treatment – (FAS)

Table 9–20: Overall survival - censoring at new treatment - (FAS)

		Placebo (N=68)	Regorafenib (N=136)
N		68	136
Number (%) of subjects with event		37 (54.4%)	71 (52.2%)
Number (%) of subjects censored		31 (45.6%)	65 (47.8%)
Median [95% CI]	Days	147 (105, 203)	269 (225, 297)
Range (including censored values)	Days	(20[a]-523[a])	(6-520[a])
Range (excluding censored values)	Days	(35-406)	(6-500)
Overall Survival rate at	Month 3 [95% CI]	0.678 (0.544,0.811)	0.896 (0.843,0.950)
	Month 6 [95% CI]	0.383 (0.231,0.534)	0.668 (0.581,0.755)
	Month 9 [95% CI]	0.179 (0.046,0.313)	0.479 (0.380,0.578)
	Month 12 [95% CI]	0.072 (0.000,0.165)	0.335 (0.235,0.435)
Hazard ratio[b] (regorafenib/placebo)		0.413	
95% CI for hazard ratio		0.274, 0.623	
One-sided p-value from log rank test		0.000007	

Abbreviations: CI = confidence interval; FAS = full analysis set; N = number of subjects
a censored observation
b Hazard ratio and its 95% CI was based on Cox Regression Model, stratified by single organ metastasis vs. multiple organ metastasis and time from diagnosis of metastatic disease (>=18 months vs <18 months).
Note: Median, percentile and other 95% CIs computed using Kaplan-Meier estimates. A Hazard ratio < 1 indicates superiority of Regorafenib over Placebo

Further exploratory analyses, stratified by prior targeted therapy, were conducted with censoring at the time of post-progression treatment. Table A9.7 summarises the new systemic anti-cancer therapies initiated during follow-up by previous anti-VEGF/anti-EGFR treatment exposure. Table A9.8 shows overall survival after censoring and stratification by prior targeted anti-cancer therapy.

What these analyses indicate is that regorafenib is beneficial irrespective of post-progression treatment. As more patients in the placebo group received post-progression therapy, the results suggest that removal of this potential benefit from the placebo arm resulted in a slightly more favourable hazard ratio for regorafenib. However, we consider that it would not be advisable to interpret the results of this exploratory analysis too literally. The analyses, considered as a whole and particularly in relation to the stratification according to prior treatment, serves to highlight the complexities encountered as you explore the data in ever increasing levels of detail. It is evident that there are multiple factors that are ‘entangled’ with post-progression treatment that contribute to overall survival e.g. prior treatment received, the fitness of the patient, and highly likely additional confounders that are unknown.

We believe that any estimations of post-progression treatment benefit would be flawed. Subjects who are censored (receive post-progression treatment) are different to those who are not censored (do not receive post-progression treatment) and they are not a random subset. This difference results in informative censoring leading to bias in any results.

As post-progression treatment will likely be composed of previous failed therapy, it stands to reason that the efficacy of the re-challenge will inevitably be significantly less than for the initial therapy. However, we do not consider it possible to isolate and quantify the benefit of post-progression treatment.

Table A9.7: Systemic anti-cancer therapy during follow-up by previous anti-VEGF/anti-EGFR treatment – (FAS)

Table 9–19: Systemic anti-cancer therapy during follow-up by previous anti-VEGF/anti-EGFR treatment – (FAS)

Subgroup by prior targeted therapy	Placebo (N=68)	Regorafenib (N=136)
No prior targeted therapy (neither anti-VEGF nor anti-EGFR)	**(N=26)**	**(N=56)**
Number of subjects (%) with at least one medication during FU	6 (23.1%)	16 (28.6%)
Anti-neoplastic and immunomodulating agents[a]	5 (19.2%)	14 (25.0%)
Blood and blood-forming organs	0	2 (3.6%)
Various	1 (3.8%)	4 (7.1%)
Prior anti-VEGF treatment but no prior anti-EGFR	**(N=13)**	**(N=32)**
Number of subjects (%) with at least one medication during FU	7 (53.8%)	6 (18.8%)
Anti-neoplastic and immunomodulating agents[a]	6 (46.2%)	5 (15.6%)
Blood and blood-forming organs	0	2 (6.3%)
Various	1 (7.7%)	1 (3.1%)
Prior anti-EGFR treatment but no prior anti-VEGF	**(N=17)**	**(N=24)**
Number of subjects (%) with at least one medication during FU	9 (52.9%)	8 (33.3%)
Anti-neoplastic and immunomodulating agents[a]	9 (52.9%)	8 (33.3%)
Blood and blood-forming organs	3 (17.6%)	4 (16.7%)
Various	2 (11.8%)	0
Prior anti-VEGF treatment and prior anti-EGFR	**(N=12)**	**(N=24)**
Number of subjects (%) with at least one medication during FU	7 (58.3%)	12 (50.0%)
Anti-neoplastic and immunomodulating agents[a]	7 (58.3%)	12 (50.0%)
Blood and blood-forming organs	1 (8.3%)	3 (12.5%)
Any prior targeted therapy (anti-VEGF, anti-EGFR, or both)	**(N=42)**	**(N=80)**
Number of subjects (%) with at least one medication during FU	23 (54.8%)	26 (32.5%)
Anti-neoplastic and immunomodulating agents[a]	22 (52.4%)	25 (31.3%)
Blood and blood-forming organs	4 (9.5%)	9 (11.3%)
Various	3 (7.1%)	1 (1.3%)
Without prior anti-VEGF therapy (with or without anti-EGFR)	**(N=43)**	**(N=80)**
Number of subjects (%) with at least one medication during FU	15 (34.9%)	24 (30.0%)
Anti-neoplastic and immunomodulating agents[a]	14 (32.6%)	22 (27.5%)
Blood and blood-forming organs	3 (7.0%)	6 (7.5%)
Various	3 (7.0%)	4 (5.0%)
With prior anti-VEGF therapy (with or without anti-EGFR)	**(N=25)**	**(N=56)**
Number of subjects (%) with at least one medication during FU	14 (56.0%)	18 (32.1%)
Anti-neoplastic and immunomodulating agents[a]	13 (52.0%)	17 (30.4%)
Blood and blood-forming organs	1 (4.0%)	5 (8.9%)
Various	1 (4.0%)	1 (1.8%)
Total across all subgroups during Follow-up	**(N=68)**	**(N=136)**
Number of subjects (%) with at least one medication during FU	29 (42.6%)	42 (30.9%)
Anti-neoplastic and immunomodulating agents[a]	27 (39.7%)	39 (28.7%)
Blood and blood-forming organs	4 (5.9%)	11 (8.1%)
Various	4 (5.9%)	5 (3.7%)

Abbreviations: EGFR = endothelial growth factor receptor; FAS = full analysis set; FU = follow-up; N = number of subjects; VEGF = vascular endothelial growth factor

a The source tables contain details on all anti-neoplastic and immunomodulating agents used.

Note: Subjects may have had more than one entry.

Table A9.8: Subgroup analysis of overall survival by prior targeted anti-cancer therapies – censoring at new treatment (FAS)

Table 9–21: Subgroup analyses of overall survival by prior targeted anti-cancer therapies - censoring at new treatment (FAS)

Subgroup	N	# Events	# Censored	Hazard Ratio (Reg/Pla) Estimate	95% CI	Median (Days) Placebo	Regorafenib
No prior targeted therapy (neither anti-VEGF nor anti-EGFR therapy)	82	46	36	0.270	(0.148, 0.493)	146	323
Prior anti-VEGF therapy but no prior anti-EGFR therapy	45	27	18	0.409	(0.159, 1.055)	105	162
Prior anti-EGFR therapy but no prior anti-VEGF therapy	41	19	22	0.594	(0.219, 1.612)	202	256
Prior anti-VEGF therapy and prior anti-EGFR therapy	36	16	20	0.505	(0.158, 1.612)	203	267
Any prior targeted therapy (anti-VEGF or anti-EGFR therapy or both)	122	62	60	0.572	(0.321, 1.021)	148	225
Without prior anti-VEGF therapy (with or without anti-EGFR therapy)	123	65	58	0.350	(0.211, 0.581)	146	295
With prior anti-VEGF therapy (with or without anti-EGFR therapy)	81	43	38	0.498	(0.243, 1.022)	148	225

Abbreviations: CI = confidence interval; EGFR = epidermal growth factor receptor; FAS = full analysis set; N = number of subjects; Pla = placebo; Reg = regorafenib; VEGF = vascular endothelial growth factor

Note: A hazard ratio <1 indicates superiority of regorafenib over placebo. Hazard ratio and CIs are based on unstratified Cox Regression Model.

d) Please adjust OS and PFS in order to better reflect current UK clinical practice in line with NICE TSD 16 (also see Ouwens M, Darilay A, Zhang Y, Mukhopadhyay P, Mann H, Ryan J, et al. Assessing the influence of subsequent immunotherapy on overall survival in patients with unresectable stage III non-small cell lung cancer from the PACIFIC study. Current Therapeutic Research, Clinical and Experimental 2021;95:100640.)

Please see response to part c – we do not believe these analyses are possible.

e) According to the company, regarding clinical practise in England and Wales, clinical experts have advised that "the proportion of patients who might receive anti-cancer therapy after T/T or regorafenib was <10%." (p. 142). Please provide details and available evidence on the provided anti-cancer therapy to these patients in UK practise.

We do not have the data requested. In our response to A9b we have provided some information on possible post-progression treatment but there are significant caveats with the data provided.

A 10. Document B of the CS states that "*First and second-line treatment of mCRC is dominated by chemotherapy combination regimens, which are typically FOLFOX or XELOX (oxaliplatin and capecitabine), and less commonly FOLFOXIRI (oxaliplatin, leucovorin, 5-FU, and irinotecan) which accounts for only 10% of all front-line treatments for mCRC).*" (p. 21). In Figure 1 irinotecan is presented as second- and third-line therapy (for RAS mutation/BRAF wild-type/MSS, RAS wild-type, BRAF V600e m+ and MSI-high/dMMR) but it is not mentioned in Table 4 as an option of NICE recommended treatments for previously treated mCRC. Please provide further details on the use of irinotecan in UK usual care and how it relates to the current submission.

[Company: please enter your answer to this question here]

Table 4 is based on NG151 and relates to 'Subsequent or alternative therapy' i.e. the position for regorafenib in this submission. Irinotecan is not a treatment recommended as 'Subsequent or alternative therapy' but a treatment earlier in the clinical pathway and is not relevant to table 4 of the CS.

Patients entered into CORRECT and CONCUR had progressed after approved standard treatments available at the time including irinotecan.

Irinotecan forms a treatment option before either trifluridine/tipiracil or regorafenib and beyond this is of no further relevance to this submission.

Systematic review

A 11. Priority question: As per section B.3.1.2. of the Appendix D, a set of "UK-specific inclusion and/or exclusion criteria were applied". Please clarify what is meant by that.

To determine the clinical evidence base that is applicable to the appraisal, inclusion and exclusion criteria were applied to the literature search results as per Table 1, Section B.3.1.2, Appendix D and aligned with the decision problem as per Table 1, Section B.1.1, Document B. We used the term "UK-specific" to clarify that the evidence base aligns with the decision problem. For example, the broader search included terms for encorafenib, a treatment not included in the scope and not relevant to the decision problem. Trials for encorafenib were excluded as part of the "UK-specific" criteria.

A 12.Priority question: Please provide further details regarding the following eligibility criteria for the clinical SLR:

a) As per Table 1 of the Appendix D, patients with "*early-stage mCRC*" were excluded. Please define early stage.

The exclusion criteria is more correctly "Early-stage CRC" - defined as those with non-metastatic disease stage I, II and III.

b) Please clarify what is meant by "*author defined best supportive care*".

There is a lack of standard definition for best supportive care (BSC) in mCRC and it varies according to individual symptoms and needs. BSC can include physical, psychological, social, and spiritual support. What constitutes BSC is seldom described in publications, hence it was decided to consider a treatment as BSC based purely on the author defining it as such.

c) It is unclear why RCTs and non-RCTs were restricted to current second line and later treatment, while observational and RWE studies were restricted to current third line and later treatment. Please explain.

The proposed position for regorafenib is as an alternative to trifluridine/tipiracil which is used as ≥3rd-line treatment (please see our response to A8). Restricting to second-line and later treatment is sufficiently broad not to miss any RCTs of relevance to the decision problem. The search conducted in response to A1 confirms that the studies relevant to the decision problem have been located.

RWE provides supporting information only and therefore, for pragmatic reasons, the search was more restrictive than for RCTs, but still in line with the decision problem.

d) Please explain why single arm trials were excluded while RWE were included.

Single-arm studies do not provide evidence of comparative effectiveness and cannot be used in indirect treatment comparisons. RWE tends not to be multi-arm, and therefore in order to provide real-world evidence it was necessary to be less restrictive.

e) Please comment on the implications of the exclusion of non-English studies.

The exclusion criteria was to exclude non-English publications, not non-English studies. We don't anticipate any impact of this exclusion criteria as studies of high quality are most likely to be reported in English language publications. This is a commonly applied exclusion criteria in systematic literature reviews.

A 13. Priority question: As per section B.3.1.2. (p. 7-8), REDOS and REARRANGE were excluded as not being relevant to the decision problem. Please include REDOS and REARRANGE for safety outcomes.

REDOS

ReDOS was a Phase II dose-escalation study of regorafenib in patients with mCRC which evaluated two dosing strategies. The primary endpoint was the proportion of patients in each arm starting Cycle 3 of treatment.

The adverse events from REDOS are presented in table A13.1 below.

Table A13.1: REDOS – Adverse Events

	Dose-escalation group (n=54)				Standard-dose group (n=62)			
	Grade 1–2	Grade 3	Grade 4	Grade 5	Grade 1–2	Grade 3	Grade 4	Grade 5
Fatigue	42 (78%)	7 (13%)	0	0	44 (71%)	11 (18%)	0	0
Hand-foot skin reaction	27 (50%)	8 (15%)	0	0	33 (53%)	10 (16%)	0	0
Hypertension	34 (63%)	4 (7%)	0	0	29 (47%)	9 (15%)	0	0
Nausea	23 (43%)	0	0	0	31 (50%)	0	0	0
Diarrhoea	23 (43%)	1 (2%)	0	0	25 (40%)	2 (3%)	0	0
Anorexia	14 (26%)	1 (2%)	0	0	16 (26%)	2 (3%)	0	0
Rash maculopapular	10 (19%)	0	0	0	16 (26%)	3 (5%)	0	0
Vomiting	13 (24%)	0	0	0	14 (23%)	1 (2%)	0	0
Blood bilirubin increased	7 (13%)	2 (4%)	0	0	13 (21%)	5 (8%)	0	0
Anaemia	12 (22%)	1 (2%)	0	0	12 (19%)	1 (2%)	0	0
Aspartate aminotransferase increased	8 (15%)	1 (2%)	0	0	12 (19%)	4 (6%)	0	0
Alkaline phosphatase increased	6 (11%)	3 (6%)	0	0	10 (16%)	1 (2%)	1 (2%)	0
Abdominal pain	1 (2%)	9 (17%)	0	0	5 (8%)	4 (6%)	0	0
Dyspnoea	5 (9%)	1 (2%)	1 (2%)	0	8 (13%)	4 (6%)	0	0
Alanine aminotransferase increased	8 (15%)	0	0	0	8 (13%)	1 (2%)	0	0
Hoarseness	8 (15%)	0	0	0	8 (13%)	0	0	0
Weight loss	4 (7%)	1 (2%)	0	0	10 (16%)	1 (2%)	0	0
Hyponatremia	0	2 (4%)	1 (2%)	0	7 (11%)	4 (6%)	1 (2%)	0
Platelet count decreased	7 (13%)	0	0	0	8 (13%)	0	0	0
Mucositis oral	4 (7%)	1 (2%)	0	0	8 (13%)	1 (2%)	0	0
Hypoalbuminemia	5 (9%)	1 (2%)	0	0	7 (11%)	0	0	0
Peripheral sensory neuropathy	6 (11%)	0	0	0	6 (10%)	0	0	0
Lymphocyte count decreased	1 (2%)	4 (7%)	0	0	6 (10%)	0	0	0
Hypocalcaemia	6 (11%)	0	0	0	3 (5%)	1 (2%)	0	0
Hypokalaemia	3 (6%)	1 (2%)	0	0	5 (8%)	0	1 (2%)	0
Generalised muscle weakness	5 (9%)	1 (2%)	0	0	2 (3%)	1 (2%)	0	0
Myalgia	0	1 (2%)	0	0	6 (10%)	2 (3%)	0	0
Pain	5 (9%)	0	0	0	3 (5%)	1 (2%)	0	0
Dehydration	1 (2%)	0	0	0	2 (3%)	5 (8%)	0	0
Investigations, other (specified)	3 (6%)	0	0	0	4 (6%)	1 (2%)	0	0
Back pain	1 (2%)	1 (2%)	0	0	5 (8%)	0	0	0
Dry skin	1 (2%)	1 (2%)	0	0	3 (5%)	0	0	0
Neoplasms: benign, malignant, unspecified, other (specified)	0	0	0	2 (4%)	0	0	0	2 (3%)
Colonic obstruction	0	3 (6%)	0	0	0	0	0	0
Hyperglycaemia	1 (2%)	1 (2%)	0	0	0	1 (2%)	0	0
Hyperkalaemia	1 (2%)	0	0	0	1 (2%)	1 (2%)	0	0
Sinus tachycardia	0	1 (2%)	0	0	1 (2%)	1 (2%)	0	0
Ascites	1 (2%)	1 (2%)	0	0	0	0	0	0
Chest wall pain	0	1 (2%)	0	0	1 (2%)	0	0	0
Death not otherwise specified	0	0	0	1 (2%)	0	0	0	1 (2%)
Encephalopathy	0	0	0	0	0	2 (3%)	0	0
Respiratory failure	0	0	1 (2%)	0	0	0	0	1 (2%)
Sepsis	0	0	1 (2%)	0	0	0	1 (2%)	0
Thromboembolic event	1 (2%)	1 (2%)	0	0	0	0	0	0
Abdominal infection	0	1 (2%)	0	0	0	0	0	0
Adult respiratory distress syndrome	0	0	1 (2%)	0	0	0	0	0
Atelectasis	0	0	0	0	0	1 (2%)	0	0
Colitis	0	1 (2%)	0	0	0	0	0	0
Confusion	0	0	0	0	0	1 (2%)	0	0
Hepatic failure	0	0	0	0	0	0	1 (2%)	0

(Table 2 continues on next page)

	Dose-escalation group (n=54)				Standard-dose group (n=62)			
	Grade 1–2	Grade 3	Grade 4	Grade 5	Grade 1–2	Grade 3	Grade 4	Grade 5
(Continued from previous page)								
Hepatobiliary disorders, other (specified)	0	1 (2%)	0	0	0	0	0	0
Hypoxia	0	0	0	0	0	1 (2%)	0	0
Infections and infestations, other (specified)	0	0	0	0	0	1 (2%)	0	0
Increased international normalised ratio	0	0	0	0	0	1 (2%)	0	0
Lower gastrointestinal haemorrhage	0	0	0	0	0	1 (2%)	0	0
Lung infection	0	0	1 (2%)	0	0	0	0	0
Myocardial infarction	0	0	0	0	0	0	0	1 (2%)
Rectal fistula	0	1 (2%)	0	0	0	0	0	0
Rectal obstruction	0	1 (2%)	0	0	0	0	0	0
Syncope	0	1 (2%)	0	0	0	0	0	0
Urinary retention	0	1 (2%)	0	0	0	0	0	0

Events are listed in order of overall frequency. All grade 3, 4, and 5 events are shown as well as grade 1 and 2 occurrence of these events. For other grade 1–2 events, only events occurring in at least 10% of patients are included.

Table 2: Adverse events

Source: Bekaii-Saab TS et al. Lancet Oncol 2019;20:1070-82 (http://dx.doi.org/10.1016/)

REARRANGE

REARRANGE investigated the impact of initial flexible dosing on regorafenib tolerability.

Refractory mCRC pts were randomized 1:1:1 to standard dose 160 mg/day 3 weeks (w) on 1 w off (SD), reduced dose 120 mg/day 3 w on 1 w off (RD) or intermittent dose 160 mg 1 w on 1 w off (ID). Pts in RD or ID escalated to SD after cycle 1 if no limiting toxicity occurred. Primary endpoint was % of pts with G3/4 treatment related adverse events (AE) on each arm.

From Jul 2016 to Sept 2017, 299 pts were randomized. Safety population set was: 100 SD, 98 RD, 99 ID. Median number of prior lines and age were 4 and 64 years. % of pts with G3/G4 AE were: 60 SD, 56 RD, 55 ID.

Source: Argiles G et al. Results of REARRANGE trial: A randomized phase 2 study comparing different dosing approaches for regorafenib (REG) during the first cycle of treatment in patients (pts) with metastatic colorectal cancer (mCRC). Ann Oncol. 2019; 30:aa135.

A 14. Data extraction. As per section B.3.1.2. "*all extracted data were verified against the original source paper* …" (p. 7 and 11). Please clarify whether data were extracted by two independent researchers.

Data extraction was performed by two independent reviewers where initially data from each study was extracted by one researcher and the extracted data was quality checked against the original source paper by an independent second researcher.

A 15. As per section B.3.1.5. "…*included RCTs were critically appraised using the National Health and Care Excellence (NICE) methodology checklist in line with the Cochrane risk of bias tool*".

a) Please clarify how these two instruments are in line with each other.

The quality assessment of included studies was performed according to the risk of bias questions provided in the NICE user guide for company evidence submission template (https://www.nice.org.uk/process/pmg24/chapter/clinical-effectiveness#critical-appraisal-of-the-relevant-clinical-effectiveness-evidence). These questions are similar to the risk of bias assessment checklist for RCTs suggested by the Cochrane Collaboration who also provide guidance to answering these questions (https://training.cochrane.org/handbook/current/chapter-08).

b) Please clarify whether one on more reviewers were involved in the quality assessment process.

The quality assessment of each included study was performed by two independent reviewers where one reviewer initially performed the quality assessment and this was verified by an independent second reviewer.

A 16. As per Table 6 of the Appendix D, patient-level data were utilised from the CORRECT and CONCUR studies for ITC. However, the remaining three studies (RECOURSE, Yoshino 2012 and TERRA) used the Kaplan-Meier curves for both OS and PFS. Please explain how these data were compatible with the ITC.

The ITCs presented in Section B.2.9 used published aggregate data (hazard ratios and associated confidence intervals) identified via the clinical systematic literature review. The patient-level data from CORRECT and CONCUR and the Kaplan-Meier plots from RECOURSE, Yoshino 2021 and TERRA were not used in these ITCs. Patient-level data from CORRECT and CONCUR were used in the anchored MAICs

presented in Appendix D and also in response to these questions. The aim of Table 6 was to summarize data availability only.

A 17. In Section B.2.8 the company states that "*meta-analysis was performed using the 'meta' package in R*". Please provide the names of any additional packages that were used in the analysis, along with the code and the datasets. Preferably provide the R script and the datasets separately.

The description of methods provided in Section B.2.8, detail of packages and functions, alongside the datasets should enable the meta-analysis to be conducted.

Detail provided in Section B.2.8

The generic inverse variance method for meta-analysis was performed using the 'meta' package in R. Fixed and random effects models were fitted to the data.

Packages

Additional packages used were: *'mautils'* (a public repository owned by RichardBirnie)

Functions

[redacted]

Data


OS data - CORRECT and CONCUR.xlsx


PFS data - CORRECT and CONCUR.xlsx

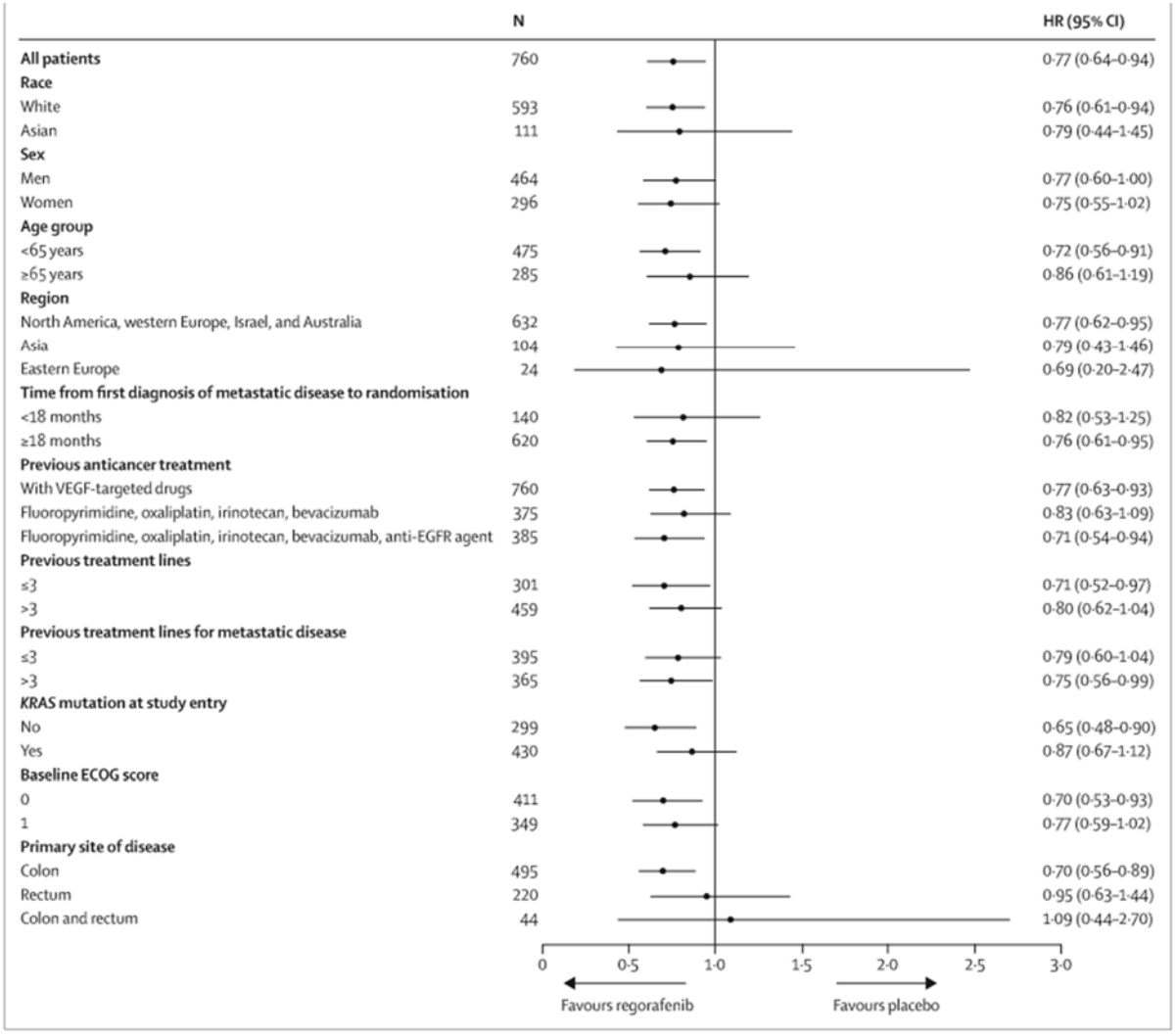
Clinical effectiveness evidence

A 18. Priority question: No UK patients were included in CORRECT or CONCUR. Please discuss with objective evidence how the study data has relevance to UK clinical practice and provide the supporting evidence.

CORRECT

CORRECT was a multinational study conducted in locations across 16 countries and included patients with different ethnicities. The benefits in respect of improved overall survival was similar regardless of region or ethnicity – see figure A18.1.

Figure A18.1. CORRECT - Overall survival; subgroup analysis



Source: Grothey A et al. Lancet. 2013; 381(9863):303-12 (figure 2)

As no interaction is observed for region or race the results of CORRECT are considered to be generalisable to UK patients.

CONCUR

CONCUR was conducted in an exclusively Asian population and therefore has an ethnic profile which is different to the UK. However, subgroup analysis of CORRECT (see above) showed no difference in efficacy as a result of ethnicity. Therefore it is reasonable to generalise the result of this trial to the UK setting.

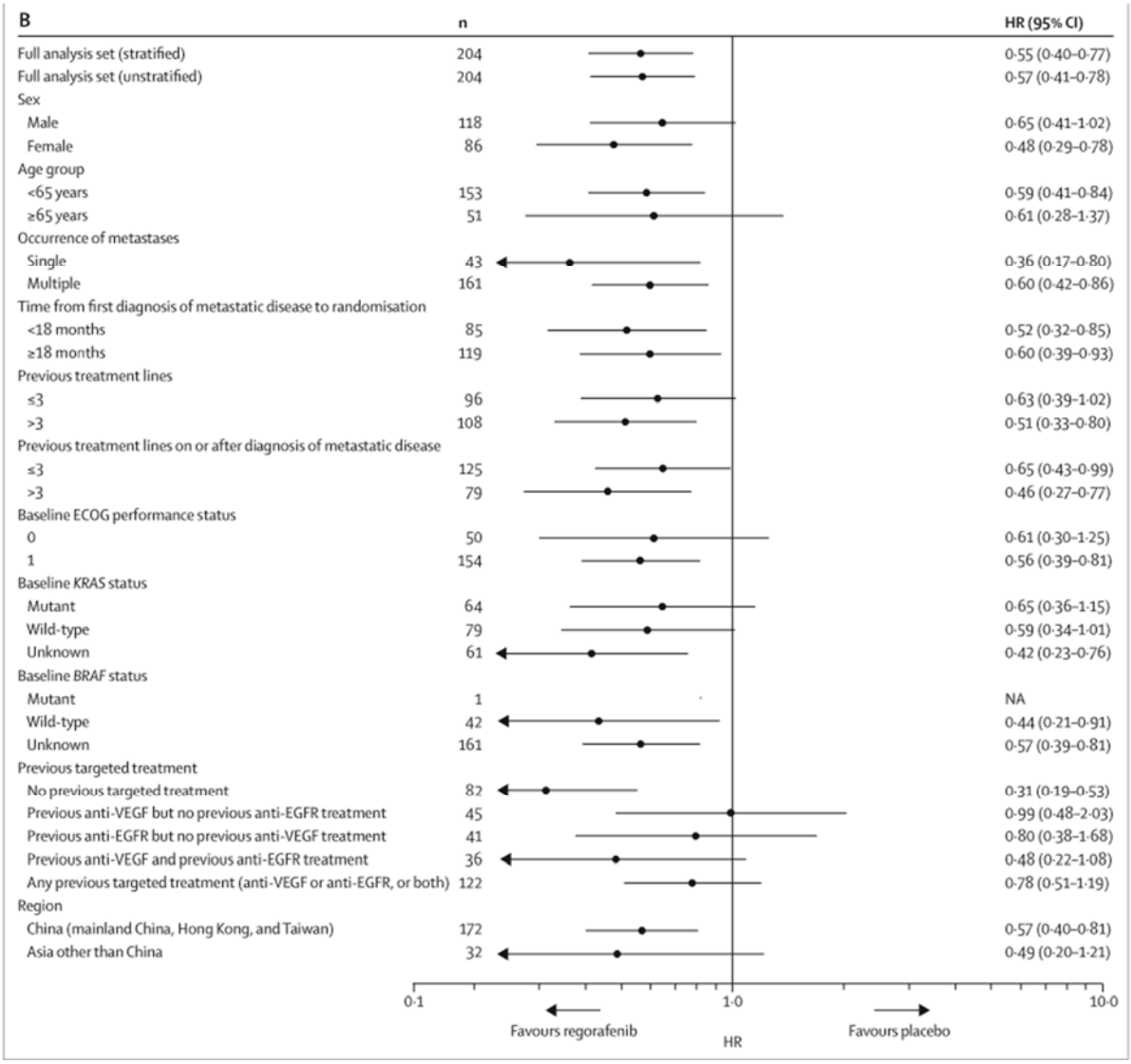
In TA405 the committee considered that RECOURSE and Yoshino – trials that are similar to CORRECT and CONCUR - were both relevant to the decision problem.

A 19. Priority question: CONCUR only included Asian patients while CORRECT had a mix of Asian (14%) and non-Asian patients. The company states in Document B of the CS that "*Regarding applicability of the trials to English clinical practice, CONCUR enrolled exclusively Asian patients, while CORRECT enrolled European and Asian patients from across 16 countries including Western Europe. Thus, it was possible to observe whether Asian and non-Asian patients responded to regorafenib in a similar manner; indeed, subgroup analyses of CORRECT and CONCUR confirmed that race was not a treatment effect modifier for regorafenib.*" (p. 52). Please confirm that subgroup analysis regarding race was indeed only available for CORRECT but not for CONCUR?

Please refer to figure A18.1 from question A18 in respect of CORRECT.

In respect of CONCUR, the study was conducted in 25 hospitals in mainland China, Hong Kong, South Korea, Taiwan, and Vietnam i.e. the study only included Asian patients. Figure A19.1 shows efficacy by region i.e. China (mainland China, Hong Kong, and Taiwan) versus 'Asia other than China' – similar efficacy was observed.

Figure A19.1: CONCUR. Overall survival – subgroup analysis



Source: Li et al. Lancet Oncol 2015; 16:619-29

A 20. Priority question: Please provide further details of the relevant experience of any clinical experts whose opinion was sought in preparing the CS and methods of elicitation of clinical expert opinion including any advisory boards.

An advisory board was conducted with 10 consultant medical & clinical oncologists who currently treat patients with mCRC. The oncologists were from hospitals across the UK (London, Southampton, Cardiff, Manchester, Sheffield, Scotland). The advisory board was chaired by an oncologist and had the following objectives

- Understanding UK clinical practice (treatments received, assessment of progression, resource use)
- Gathering feedback on regorafenib's trial data and generalisability to the UK
- Gathering clinical views of the relative efficacy of regorafenib and trifluridine/tipiracil
- Understanding the place of regorafenib in the treatment pathway
- Assessing the appropriateness of the economic model.

The meeting was structured around the presentation of trial data and group discussion on the above topics.

A 21. Priority question: The company suggests that one possible explanation for the different results between CORRECT and CONCUR, one of which is the total number of prior treatments received. In CORRECT, OS HR was lower in the ≤ 3 subgroup 0.709 (95% CI 0.521, 0.967) vs the >3 subgroup 0.804 (95% CI 0.624, 1.038) and the same trend was reported for PFS HRs. On the other hand, in CONCUR, OS HR was higher in the ≤ 3 subgroup 0.629 (95% CI 0.388, 1.019) vs the >3 subgroup 0.514 (95% CI 0.330, 0.800). Regarding >3 number of prior treatments on or after diagnosis of metastatic disease the difference between CORRECT and CONCUR was higher (48% vs 38.8%), while OS and PFS HRs were higher in the ≤ 3 subgroups in both studies. The only HRs that did not overlap was for PFS comparing ≤ 3 vs >3 number of prior treatments on or after diagnosis of

metastatic disease in CONCUR 0.369 (95% CI 0.244, 0.557) vs 0.195 (95% CI 0.108, 0.350).

a) Please explain the differences in the efficacy results between the two studies.

It is clinically accepted that as a patient progresses through each line of therapy the potential to benefit from the next treatment is, on average, diminished. This is a result of drug resistance which accumulates with each successive treatment, and failure of that treatment.

Although the direction of diminishing ability to benefit from successive lines of treatment is clinically accepted, the effect observed in CORRECT and CONCUR is counter to expectations (as pointed out by the EAG). The same counterintuitive result was also observed in the RECOURSE trial for trifluridine/tipiracil (Mayer 2015 – figure B). However, the differences are non-significant and confidence intervals relatively wide.

The counterintuitive results cannot be explained. However, there is a risk in overinterpreting point estimates of subgroup results from trials that are powered at the overall population level. The point estimates from these subgroup analyses should be viewed in the context of their confidence intervals.

The results observed in CONCUR should be taken as confirmation of the treatment benefit for regorafenib which was observed in CORRECT. The benefit was greater in the CONCUR trial but there is no clear explanation for this. Similarly, the trials for trifluridine/tipiracil show a difference in OS benefit (RECOURSE 0.68, Yoshino 0.56, TERRA 0.79) without a clear explanation for the difference. In this context, the best indication of benefit is the average result across the respective trials.

b) Please discuss the implications for generalisability to the decision problem.

We don't believe that the subgroup results referred to in this question have any implications for generalisability. On face-value, CONCUR may be more generalisable to the UK in terms of prior treatments (compared to CORRECT), however, as mentioned above there is no clear explanation for the difference in

results observed between the two trials. We consider that the results of both trials are generalisable to England and that the best estimate of effect probably lies somewhere between both studies and should be estimated via meta-analysis.

We note that during the trifluridine/tipiracil appraisal (TA405) similar questions were raised in respect of the trials for trifluridine/tipiracil which were conducted in different countries and differed in prior treatment. The committee considered that the results of both trials (only RECOURSE and Yoshino 2012 were available at the time) were generalisable to NHS patients in England.

A 22. Priority question: The company stated that "Prior treatment with anti-VEGF is relevant as regorafenib has anti-VEGF activity (see Section B.1.2) - the implication of this prior therapy is that regorafenib could be expected to be less effective in patients who have already been treated with, and failed on, an anti-VEGF" (p. 53). The subgroup analysis results for both OS and PFS did not present the subgroup of interest i.e. patients previously treated with an anti-VEGF treatment. Instead, the company has provided results on five subgroups involving combinations with an anti-EGFR, which produce what appear to be some counterintuitive results. No targeted treatment gives the lowest HR (most effective) for both OS and PFS, which makes sense. However, for OS, although previous anti-VEGF and no previous anti-EGFR gives the highest HR (least effective), no previous anti-VEGF and previous anti-EGFR gives the second highest HR. For PFS, the highest HR is produced by no previous anti-VEGF and previous anti-EGFR and there is no overlap in the 95% CIs with no previous treatment.

a) Please explain the apparent inconsistency of these results

In respect of prior anti-VEGF treatment versus no prior treatment - the OS hazard ratio for patients who had not received anti-VEGF was lower (HR: 0.470; 95% CI: 0.309, 0.714) compared to patients who received anti-VEGF (HR: 0.726; 95% CI: 0.430, 1.224). These results are supportive of a greater potential to benefit in patients who have not received prior anti-VEGF treatment. Although these analyses

are supportive, they are post-hoc in nature and caution should be exercised when interpreting any post-hoc results.

Apparent inconsistency of results referred to in the question

The subgroup results show an overall consistency of benefit across a wide range of patient characteristics. However, we believe that the results in the subgroups should not be overinterpreted as the trials were powered at the ITT level - not the subgroup level. Numbers in some subgroups are relatively small. It is unrealistic to expect that point estimates across 'related' subgroups will always be clinically logical - there will always be the play of chance. Furthermore, treatment groups were randomised at the overall population level ensuring important characteristics were well balanced. Some imbalances may arise between treatment groups in subpopulations which could influence outcomes.

b) Please provide further subgroup analysis comparing the subgroup of patients that have received an anti-VEGF vs patients who have not received an anti-VEGF, regardless of anti-EGFR experience.

The post-hoc subgroup analysis for overall survival comparing patients who had not received prior anti-VEGF therapy versus patients who received prior anti-VEGF therapy, regardless of anti-EGFR therapy, is presented in the submission in Table 32, Section B.4.2.3 in Appendix E. The OS hazard ratio for patients who had not received anti-VEGF was lower (HR: 0.470; 95% CI: 0.309, 0.714) compared to patients who received prior anti-VEGF (HR: 0.726; 95% CI: 0.430, 1.224), although in both cases the benefit was in favour of regorafenib.

c) Please discuss the implications for generalisability of each of the trials to UK clinical practice.

Regorafenib has been shown to be effective in the mCRC population overall and in subgroup and exploratory analyses. As a consequence of powering and low numbers in some subgroups overinterpretation should be avoided. The results from CONCUR (and CORRECT) are generalisable to the UK and individual subgroup results supports a consistent effect across a wide range of subgroups.

A 23. Priority question: The anti-VEGF medication bevacizumab is not recommended by NICE for patients with mCRC. The entire population in CORRECT and 39.7% of the patients in CONCUR were previously treated with an anti-VEGF. The company states that the trial CONCUR "…*more closely aligns with clinical practice in England and Wales as it includes a significant proportion of patients who have never received anti-VEGF therapy*". (p. 53).

a) Please confirm that all the CORRECT patients received bevacizumab: if not then provide the percentage who received another anti-VEGF instead.

We confirm that 100% of patients in both the regorafenib and placebo groups received prior bevacizumab in CORRECT (see Table 7, Section B.2.3.2 of Document B).

b) Please discuss how previous treatment with bevacizumab might affect generalisability to UK clinical practice.

Please also see our response to A22 as there is significant overlap.

Prior exposure to anti-VEGF therapy (i.e. bevacizumab) may reduce the treatment effect associated with regorafenib, as regorafenib is a multi-kinase inhibitor with targets that include VEGF. Consequently, prior treatment and failure on anti-VEGF treatment could, *ceteris paribus,* have a downward effect on the efficacy of regorafenib. The implication could be better efficacy in UK clinical practice (compared to the results from CORRECT) as patients won't have received prior treatment with bevacizumab.

We believe the results of CORRECT and CONCUR to be generalisable to the UK. The 'best' estimate of efficacy would be from meta-analysis of CORRECT and CONCUR.

A 24. In the CORRECT and CONCUR trials the median duration of treatment was 2.8 and 2.4 months in the regorafenib group and 1.8 and 1.6 months in the placebo group. It is possible that this difference could contribute to considerable performance bias in favour of regorafenib. Based on the KM plots, it is highly unlikely that this difference can be explained by differential death rates in the early part of the studies. Please explain the source of this discrepancy and the likely effect on outcomes.

In the management of mCRC (and typically all cancers), patients are treated until progression is observed. Progression signifies the cancer has become resistant to that treatment. If the patient is earlier in the treatment pathway then progression would result in the stopping of that treatment and the start of the next line of therapy.

Due to the anti-cancer activity of regorafenib it took longer for progression to take place which explains the difference in median treatment duration between arms – there is no bias that is introduced.

A 25. In the progression-free survival KM curve for CORRECT (p55) there appear to be periodic increases in the negative gradient of the curve, corresponding to the periods immediately before the 2, 4, 6, 8 and 10 month follow up points, and these are particularly marked in the regorafenib group. Is this an artefact of the timing of outcome evaluation? Please explain this phenomenon because the shape of the curves has an impact on interpretation.

In CORRECT and CONCUR patients were assessed every 8 weeks to determine if the cancer had progressed. This 8-weekly assessment matches clinical practice in the UK. The 'stepped' nature of the KM curve reflects the timepoints of clinical assessment. However, the 'step' is not perfectly vertical as there was a one-week window either side of 8-week timepoint where assessments could take place.

A 26. Although the scope in the CS is ≥3rd line setting the majority of patients in both trials were at 5th line (49% and 47%, 54% and 51% in CORRECT and CONCUR, respectively). The company recognized that "*prognosis, and ability to benefit from treatment diminishes with each additional line of therapy*" (p. 53). Please discuss how number of prior treatments might have affected clinical effectiveness outcomes.

As discussed in our response to A21 it is clinically accepted that as a patient receives, and fails on, successive treatments their ability to benefit from the next treatment diminishes (all else being equal). This is partly as a result of the build-up of treatment resistances.

The CORRECT and CONCUR trials enrolled patients with differing numbers of prior treatments. Regorafenib was shown to be effective irrespective of the number of prior treatments received.

A 27. Please specify whether patients with Eastern Cooperative Oncology Group (ECOG) status 2 and above would be offered regorafenib.

Patients who will be considered fit enough for active treatment will typically have an ECOG status of 0 /1. Patients with ECOG status 2 or above are generally considered too unwell for active therapy – however this would be an individualised decision.

Indirect treatment comparison (ITC)

A 28. Priority question. According to the company the unmet need that regorafenib would cover is that of "*a chemotherapy-free alternative therapy with a different adverse event profile which would increase the options available*." (p.25). Also, in Section B.2.11.1.3 the company states that in CORRECT "Grade 3 or 4 TEAEs occurred at a higher rate in the regorafenib group than in the placebo group (54% versus 14%)" in ≥ 5% of patients (p 81), while in Section B.2.11.2.3 regarding CONCUR (in Table 22) it is reported that Grade 3 or 4 TEAEs also occurred at a higher rate in the regorafenib group than in the placebo group (53% versus 14%) in ≥ 10% of patients.

a) Please discuss how this safety profile compares to that for trifluridine /tipiracil.

Grade ≥ 3 TEAEs occurred at a higher rate in the trifluridine/tipiracil group than in the placebo group in RECOURSE (trifluridine/tipiracil: 49.0%; placebo: 10.2%) and TERRA (trifluridine/tipiracil: 45.8%; placebo: 10.4%). The incidence of Grade ≥ 3 TEAEs and the difference between treatments groups are similar to those observed in CORRECT and CONCUR. In Yoshino 2012, absolute values were not given for incidence of any Grade ≥ 3 TEAEs. However, it was noted that Grade 3 or worse adverse events were uncommon in the placebo group when compared to the trifluridine/tipiracil group.

Trifluridine/tipiracil has a different AE profile compared to regorafenib; trifluridine/tipiracil is associated with higher haematological AEs such as Grade ≥ 3 neutropenia (33–50%), leukopenia (21–28%) and anaemia (17–18%), whereas regorafenib is associated with higher Grade ≥ 3 hand–foot skin reactions (16–17%) and hypertension (7–11%). The different safety profile of regorafenib increases the options available i.e. patients who are unlikely to tolerate the adverse event profile of trifluridine/tipiracil, as indicated by experience with prior chemotherapy, may be better on regorafenib. Conversely, patients more susceptible to regorafenib's AE profile might be better on trifluridine/tipiracil.

Table A28.1: Grade ≥ 3 treatment-related adverse events in > 2% in regorafenib and trifluridine/tipiracil studies

AE Grade ≥ 3	Regorafenib		Trifluridine/tipiracil		
	CORRECT	CONCUR	RECOURSE	TERRA	Yoshino 2012
Abdominal pain			2.4%		
Anaemia	2.8%		18.0%	17.7%	16.8%
Anorexia	3.2%				4.4%
Asthenia			3.4%		
Decreased appetite			3.6%		
Diarrhoea	7.2%		3.0%		6.2%
Fatigue	9.6%	2.9%	3.9%		6.2%
Febrile neutropenia		2.2%	3.8%		4.4%
Hand–foot skin reaction	16.6%	16.2%			
Hypertension	7.2%	11.0%			
Leukopenia		2.2%	21.2%	20.7%	28.3%
Lymphopenia				14.4%	9.7%
Mucositis	3.0%				
Nausea			1.9%		4.4%
Neutropenia		2.2%	37.5%	33.2%	50.4%
Rash	5.8%	4.4%			
Thrombocytopenia	2.8%	2.9%	5.1%	3.0%	4.4%
Vomiting			2.1%		3.5%
Hyperbilirubinaemia	2.0%	6.6%	8.4%	7.0%	
Hypophosphataemia	3.8%	6.6%	7.9%		
Increase in ALT level		6.6%	1.9%	1.1%	
Increase in AST level		5.9%	4.3%	3.7%	
Increase in lipase level	3.2%	4.4%			

Key: AE, adverse event; ALT, alanine aminotransferase; AST, aspartate aminotransferase.
Source: Grothey et al., 2013{Grothey, 2013 #4}; Li et al., 2015{Li, 2015 #2}; Mayer et al., 2015{Mayer, 2015 #9}; Xu et al., 2018{Xu, 2018 #61}; Yoshino et al., 2012.{Yoshino, 2012 #35}

b) In section B.3.4.4, it states that the AE rates were pooled for trifluridine /tipiracil. Please describe the method of pooling and whether randomisation was preserved.

AE rates were pooled using a weighted proportion. This simple approach was preferred over more complex methods such as NMA or logistic regression as AE rates were very low and a continuity correction was needed. Adding in a continuity correction introduces events where none were observed which would not be reflective of the data. A simple approach was also preferred given the AE rates do not have a high impact on the cost-effectiveness modelling, with the modelled total

AE costs only amounting to [redacted] for regorafenib and trifluridine/tipiracil respectively.

c) Please carry out and present a NMA analysis for any Grade 3+ treatment emergent adverse events.

To estimate the relative safety of regorafenib compared with trifluridine/tipiracil and placebo in Grade 3 or 4 AEs, a fixed effects NMA was fitted to the data presented in A28b. Yoshino 2012 could not be included in this analysis as data were not reported.

Table A28.2: Data included in fixed effects NMA in treatment related Grade 3 or 4 AEs

Study Name	Treatment	Safety N	Number of treatment related Grade 3 or 4 AEs: n(%)
CORRECT	Regorafenib 160 mg	500	270 (54.0)
	Placebo	253	35 (13.8)
CONCUR	Regorafenib 160 mg	136	74 (54.4)
	Placebo	68	10 (14.7)
RECOURSE	Trifluridine/tipiracil	533	261 (49.0)
	Placebo	265	27 (10.2)
TERRA	Trifluridine/tipiracil	271	124 (45.8)
	Placebo	135	14 (10.4)
Key: AEs, adverse events; CI, confidence interval; N, total number of patients; n, number of patients with an event; NMA, network meta-analysis.			

Table A28.3 presents the results of the fixed effects NMA in treatment related Grade 3 or 4 AEs. The NMA suggested similar odds of experiencing a Grade 3 or 4 adverse event for regorafenib and trifluridine/tipiracil (OR: 0.90 [95% credible interval: 0.55, 1.47]).

Table A28.3: Results of the fixed effect NMA in treatment related Grade 3 or 4 adverse events

Comparison	OR (95% CrI) - NMA
Regorafenib vs placebo	7.32 (5.19, 10.44)
Trifluridine/tipiracil vs placebo	8.11 (5.74, 11.65)
Regorafenib vs trifluridine/tipiracil	0.90 (0.55, 1.47)
Key: CrI, credible interval; NMA, network meta-analysis; OR, odds ratio; PFS, progression-free survival.	

d) Please carry out and present a NMA analysis for all treatment emergent AEs included in the economic model.

Treatment emergent adverse events are not generally modelled given they include mild and moderate events which are not expected to impact cost and quality of life. Grade 3/4 adverse events, however, will have an impact on cost and quality of life.

To estimate the relative safety of regorafenib compared with trifluridine/tipiracil and placebo in all treatment emergent AEs, a fixed effect NMA was fitted to the data presented in A28.4. Yoshino 2012 could not be included as data were not reported.

Table A28.4: Data included in fixed effects NMA in all treatment emergent AEs

Study Name	Treatment	Safety N	Number of all treatment emergent AEs: n (%)
CORRECT	Regorafenib 160 mg	500	465 (93.0)
	Placebo	253	154 (60.9)
CONCUR	Regorafenib 160 mg	136	132 (97.1)
	Placebo	68	31 (45.6)
RECOURSE	Trifluridine/tipiracil	533	458 (85.9)
	Placebo	265	146 (55.1)
TERRA	Trifluridine/tipiracil	271	244 (90.0)
	Placebo	135	70 (51.9)
Key: AEs, adverse events; CI, confidence interval; N, total number of patients; n, number of patients with an event; NMA, network meta-analysis.			

Table A28.5 presents the results of the fixed effects NMA for all treatment emergent AEs. The NMA suggested higher odds of treatment emergent AEs for patients

treated with regorafenib compared with trifluridine/tipiracil (OR: 1.94 [95% CrI: 1.20, 3.17]).

Table A28.5: Results of the fixed effect NMA in all treatment emergent AEs

Comparison	OR (95% CrI) - NMA
Regorafenib vs placebo	11.42 (7.78 to 17.10)
Trifluridine/tipiracil vs placebo	5.90 (4.43 to 7.89)
Regorafenib vs trifluridine/tipiracil	1.94 (1.20 to 3.17)
Key: CrI, credible interval; NMA, network meta-analysis; OR, odds ratio; PFS, progression-free survival.	

e) Please carry out and present a NMA for discontinuation due to AEs.

To estimate the relative safety of regorafenib compared with trifluridine/tipiracil and placebo in discontinuation due to AEs, a fixed effect NMA was fitted to the data presented in Table A28.6.

Table A28.6: Data included in fixed effects NMA of discontinuations due to AEs

Study Name	Treatment	Safety N	Discontinuation due to AEs: n (%)
CORRECT	Regorafenib 160 mg	500	85 (17.0)
	Placebo	253	30 (11.9)
CONCUR	Regorafenib 160 mg	136	19 (14.0)
	Placebo	68	4 (5.9)
RECOURSE	Trifluridine/tipiracil	533	19 (3.6)
	Placebo	265	4 (1.5)
TERRA	Trifluridine/tipiracil	271	24 (8.9)
	Placebo	135	11 (8.1)
Yoshino 2012	Trifluridine/tipiracil	113	1 (0.9)
	Placebo	57	4 (7.0)
Key: AEs, adverse events; CI, confidence interval; N, total number of patients; n, number of patients with an event; NMA, network meta-analysis.			

Table A28.7 presents the results of the fixed effects NMA in discontinuation due to AEs. The NMA suggested similar odds of discontinuation due to an adverse event for regorafenib and trifluridine/tipiracil (OR: 1.10 [95% credible interval: 0.53, 2.24]).

Table A28.7: Results of the fixed effect NMA in discontinuation due to AEs

Comparison	OR (95% CrI) - NMA
Regorafenib vs placebo	1.66 (1.11 to 2.56)
Trifluridine/tipiracil vs placebo	1.51 (0.86 to 2.78)
Regorafenib vs trifluridine/tipiracil	1.10 (0.53 to 2.24)
Key: CrI, credible interval; NMA, network meta-analysis; OR, odds ratio; PFS, progression-free survival.	

f) Please comment on and provide evidence on how the safety results of regorafenib offer an alternative adverse event profile compared to T/T.

Overall, the results above indicate similar proportions of patients experiencing adverse events which have implications on quality of life and costs (Grade 3+). The proportion of patients experiencing any TEAE was higher for regorafenib compared to trifluridine/tipiracil, however this analysis includes grade 1 and 2 events which are classed as mild/moderate and are not expected to impact quality of life or costs.

Discontinuations due to AEs was comparable between the two medicines (see table A28.7).

The total proportions of patients experiencing events overall was comparable, however the types of events differed between the medicines (see table A28.1). This difference in adverse event profile is of clinical importance when selecting treatment for individual patients but has little impact on relative cost-effectiveness.

A 29. Priority question: There are differences in the populations between CORRECT and CONCUR, regarding race (Asians vs non-Asians) and patients previously treated with an anti-VEGF agent.

a) Please provide separate ITC efficacy analyses including only one of the studies at a time.

Sensitivity analyses accounting for differences in race (Asians vs non-Asians) and prior anti-VEGF treatment were provided in the CS: see Section B.2.9 (Table 16 summarizes the sensitivity analyses performed). CORRECT only was compared with RECOURSE given all patients had received prior anti-VEGF in these two trials and both trials were multicentre international (both studies included patients in North America, Europe, Australia and Japan). It would not be appropriate to compare CORRECT with TERRA or Yoshino as a sensitivity to explore similar race and prior anti-VEGF treatment given the latter two studies included only Asian patients and 50% and 20% of patients had received prior targeted biologic treatment, respectively. Likewise, sensitivity analyses were performed comparing CONCUR only with TERRA, Yoshino and both TERRA and Yoshino given all three studies included Asian patients and a proportion had received prior targeted biologic treatment.

b) In addition, for the ITC including CORRECT, please also include anchored matching-adjusted indirect comparison (MAIC), adjusting for race and region.

An anchored matching-adjusted indirect treatment comparison requires all treatment effect modifiers (balanced or imbalanced) to be included in the matching. No purely prognostic variables should be included to avoid inflating the standard error (over matching). Given age and gender were identified as potential treatment effect modifiers, these have also been included in an anchored MAIC comparing evidence from CORRECT and RECOURSE. Race and region were not identified as potential treatment effect modifiers and therefore the inclusion of these variables may lead to over matching.

CORRECT versus RECOURSE

The population of CORRECT was matched to the population of RECOURSE based on the proportion of male patients, the proportion of patients aged<65 years, the proportion of patients whose race was white and proportion of patients not located in Asia. Table A29.1 presents the patient characteristics of CORRECT before and after matching, and the ESS obtained after matching. The ESS after matching (561) was 73.8% the original sample size of 760 patients.

Table A29.1: Matching CORRECT (ITT) to RECOURSE (ITT)

	Prior to matching		After matching
	CORRECT	RECOURSE	CORRECT
N/ESS	N = 760	N = 800	ESS = 561
% male patients	61.05%	61.38%	61.38%
% patients < 65	62.5%	56.0%	56.0%
% White	78.0%	57.6%	57.6%
% Area not Asia	86.3%	66.8%	66.8%
Key: ESS, effective sample size; ITT, intention-to-treat.			

Results of the anchored MAIC for OS are presented in Table A29.2, along with an unadjusted Bucher comparison performed to assess the impact of matching on results. Results show weighting had minimal impact on the unadjusted HRs. Results of the anchored MAIC suggested regorafenib had a slightly higher hazard of death compared with trifluridine/tipiracil with an HR of 1.15 (0.86, 1.54).

Table A29.2: Results of OS – CORRECT (ITT) versus RECOURSE (ITT)

Comparison	HR (95% CI)	Source
Regorafenib vs placebo	0.77 (0.64, 0.94)	CORRECT
Trifluridine/tipiracil vs placebo	0.68 (0.58, 0.81)	RECOURSE
Regorafenib vs trifluridine/tipiracil	1.13 (0.88, 1.46)	Bucher ITC
Regorafenib vs placebo (adjusted)	0.78 (0.62, 1.01)	CORRECT (weighted)
Regorafenib vs trifluridine/tipiracil (adjusted)	1.15 (0.86, 1.54)	Anchored MAIC
Key: CI, confidence interval; HR, hazard ratio; ITC, indirect treatment comparison; ITT, intention-to-treat; OS, overall survival; vs, versus.		

Results of the anchored MAIC for PFS are presented in Table A29.3, along with an unadjusted Bucher comparison to assess the impact of matching on results. Given the same set of weights was used to adjust PFS as OS, matching resulted in a minimal impact on PFS. Results of the anchored MAIC suggested patients treated by regorafenib had a very similar hazard of progression compared with trifluridine/tipiracil, with an HR of 1.05 (0.81, 1.35).

Table A29.3: Results: PFS – CORRECT (ITT) versus RECOURSE (ITT)

Comparison	HR (95% CI)	Source
Regorafenib vs placebo	0.49 (0.42, 0.58)	CORRECT
Trifluridine/tipiracil vs placebo	0.48 (0.41, 0.57)	RECOURSE
Regorafenib vs trifluridine/tipiracil	1.02 (0.81, 1.29)	Bucher ITC
Regorafenib vs placebo (adjusted)	0.50 (0.41, 0.60)	CORRECT (weighted)
Regorafenib vs trifluridine/tipiracil (adjusted)	1.05 (0.81, 1.35)	Anchored MAIC
Key: CI, confidence interval; HR, hazard ratio; ITC, indirect treatment comparison; MAIC, matching-adjusted indirect comparison; PFS, progression-free survival; vs, versus.		

c) In addition, for the ITC including CONCUR, please also include anchored matching-adjusted indirect comparison (MAIC) adjusting for anti-VEGF therapy.

An anchored matching-adjusted indirect treatment comparison requires all treatment effect modifiers (balanced or imbalanced) to be included in the matching. No purely prognostic variables should be included to avoid inflating the standard error (over matching). Given age and gender, as well as prior targeted biological treatment, were identified as potential treatment effect modifiers, these have all been included as matching variables in anchored MAICs comparing: (1) CONCUR versus TERRA and (2) CONCUR versus Yoshino 2012.

CONCUR versus TERRA

The population of CONCUR was matched to the population of TERRA based on the proportion of patients without prior targeted therapy, the proportion of male patients and the proportion of patients aged <65 years. Table A29.4 presents the patient characteristics of CONCUR before and after matching, and the ESS obtained after matching. The ESS after matching (189) was very similar to the original sample size of 204 patients, which is not surprising given the similarity of the characteristics between the populations prior to matching.

Table A29.4: Matching CONCUR (ITT) to TERRA (ITT)

	Prior to matching		After matching
	CONCUR	TERRA	CONCUR
N/ESS	N = 204	N = 406	ESS = 189
% male patients	57.8%	62.6%	62.6%
% patients < 65	75.0%	76.1%	76.1%
% no prior targeted biological treatment	40.2%	52.7%	52.7%
Key: ESS, effective sample size; ITT, intention-to-treat.			

Results of the anchored MAIC for OS are presented in Table A29.5, along with an unadjusted Bucher comparison performed to assess the impact of matching on

results. Results show weighting had minimal impact on the unadjusted HRs, which is consistent with the ESS. Results of the anchored MAIC suggested regorafenib decreased the hazard of death compared with trifluridine/tipiracil with an HR of 0.64 (0.42, 0.97).

Table A29.5: Results of OS – CONCUR (ITT) versus TERRA (ITT)

Comparison	HR (95% CI)	Source
Regorafenib vs placebo	0.55 (0.40, 0.77)	CONCUR
Trifluridine/tipiracil vs placebo	0.79 (0.62, 0.99)	TERRA
Regorafenib vs trifluridine/tipiracil	0.70 (0.47, 1.04)	Bucher ITC
Regorafenib vs placebo (adjusted)	0.50 (0.36, 0.71)	CONCUR (weighted)
Regorafenib vs trifluridine/tipiracil (adjusted)	0.64 (0.42, 0.97)	Anchored MAIC
Key: CI, confidence interval; HR, hazard ratio; ITC, indirect treatment comparison; ITT, intention-to-treat; OS, overall survival; vs, versus.		

Results of the anchored MAIC for PFS are presented in Table A29.6, along with an unadjusted Bucher comparison to assess the impact of matching on results. Given the same set of weights was used to adjust PFS as OS, matching resulted in a minimal impact on PFS. Results of the anchored MAIC suggested regorafenib resulted in a decrease in the hazard of progression compared with trifluridine/tipiracil, with an HR of 0.65 (0.41–1.01).

Table A29.6: Results: PFS – CONCUR (ITT) versus TERRA (ITT)

Comparison	HR (95% CI)	Source
Regorafenib vs placebo	0.31 (0.22, 0.44)	CONCUR
Trifluridine/tipiracil vs placebo	0.43 (0.34, 0.54)	TERRA
Regorafenib vs trifluridine/tipiracil	0.72 (0.48, 1.09)	Bucher ITC
Regorafenib vs placebo (adjusted)	0.28 (0.18, 0.39)	CONCUR (weighted)
Regorafenib vs trifluridine/tipiracil (adjusted)	0.65 (0.41, 1.01)	Anchored MAIC
Key: CI, confidence interval; HR, hazard ratio; ITC, indirect treatment comparison; MAIC, matching-adjusted indirect comparison; PFS, progression-free survival; vs, versus.		

CONCUR versus Yoshino 2012

The population of CONCUR was matched to the population of Yoshino 2012 based on the proportion of patients without prior targeted therapy, the proportion of male patients and the proportion of patients aged <65 years. Table A29.7 presents the patient characteristics of CONCUR before and after matching and the ESS obtained after matching.

The proportion of patients with prior targeted therapy in Yoshino 2012 was based on the number of patients receiving prior bevacizumab.

Table A29.7: Matching CONCUR (ITT) to Yoshino 2012 (ITT)

	Prior to matching		After matching
	CONCUR	Yoshino 2012	CONCUR
N/ESS	N = 204	N = 169	ESS = 143
% male patients	57.8%	54.4%	54.4%
% patients < 65	75.0%	55.6%	55.6%
% no prior targeted biological treatment	40.2%	20.7%*	20.7%
Key: ESS, effective sample size; VEGF, Vascular endothelial growth factor. **Notes:** * Based on prior anti-VEGF; information on prior anti-EGFR is not known.			

Results of the anchored MAIC for OS are presented in Table A29.8, along with an unadjusted Bucher comparison performed to assess the impact of matching on results. While weighting has a larger impact in this analysis compared with the previous matching performed, the adjusted HR of regorafenib compared with placebo was fairly similar to the HR observed in CONCUR (the weighting resulted in an increase in the HR by 0.08). This increase in the HR comparing regorafenib with placebo is likely due to the reduction in the proportion of patients in CONCUR without prior targeted therapy. The HR comparing regorafenib with trifluridine/tipiracil was 1.13 with a relatively large confidence interval containing 1 (95% CI: 0.64, 1.99).

Table A29.8: Results: OS – CONCUR (ITT) versus Yoshino 2012 (ITT)

Comparison	HR (95% CI)	Source
Regorafenib vs placebo	0.55 (0.40, 0.77)	CONCUR
Trifluridine/tipiracil vs placebo	0.56 (0.39, 0.81)	Yoshino 2012
Regorafenib vs trifluridine/tipiracil	0.98 (0.60, 1.60)	Bucher ITC
Regorafenib vs placebo (adjusted)	0.63 (0.41, 0.96)	CONCUR (weighted)
Regorafenib vs trifluridine/tipiracil (adjusted)	1.13 (0.64, 1.99)	Anchored MAIC
Key: CI, confidence interval; HR, hazard ratio; ITC, indirect treatment comparison; MAIC, matching-adjusted indirect comparison; OS, overall survival; vs, versus.		

Results of the anchored MAIC for PFS are presented in Table A29.9, along with an unadjusted Bucher comparison to assess the impact of matching on results. For regorafenib versus placebo, the adjusted HR for PFS (0.38 [0.20–0.57]) increased by 0.07 compared with the HR observed in CONCUR, which is consistent with the OS results. Again, this is likely due to reducing the proportion of patients without prior targeted therapy. Results of the anchored MAIC showed regorafenib with a similar hazard of progression compared with trifluridine/tipiracil, with a HR of 1.09 and relatively large confidence interval containing 1 (95% CI: 0.58–2.04).

Table A29.9: Results: PFS – CONCUR (ITT) versus Yoshino 2012 (ITT)

Comparison	HR (95% CI)	Source
Regorafenib vs placebo	0.31 (0.22, 0.44)	CONCUR
Trifluridine/tipiracil vs placebo	0.35 (0.25, 0.50)	Yoshino 2012
Regorafenib vs trifluridine/tipiracil	0.89 (0.54, 1.45)	Bucher ITC
Regorafenib vs placebo (adjusted)	0.38 (0.20, 0.57)	CONCUR (weighted)
Regorafenib vs trifluridine/tipiracil (adjusted)	1.09 (0.58, 2.04)	Anchored MAIC
Key: CI, confidence interval; HR, hazard ratio; ITC, indirect treatment comparison; MAIC, matching-adjusted indirect comparison; PFS, progression-free survival; vs, versus.		

A 30. Priority question: The CS uses fixed-effects models for the ITC efficacy analyses due to the small number of included studies. Please provide additional analyses using random-effect models.

The base case analyses (data from the primary analysis) presented in Section B.2.9 (Table 14 and Table 15) have been rerun using a random effects model. Results for OS and PFS are presented in Table A30.1 and Table A30.2, respectively. Point estimates are very similar across fixed effect and random effects models and confidence intervals are wider from the random effects models.

Table A30.1: Results of the random effects versus fixed effect NMA of OS

Comparison	Random effects model: HR (95% CrI)	Fixed effect model: HR (95% CrI)
Regorafenib versus placebo	0.66 (0.34, 1.26)	0.68 (0.59, 0.78)
Trifluridine/tipiracil versus placebo	0.68 (0.39, 1.15)	0.68 (0.62, 0.76)
Regorafenib versus trifluridine/tipiracil	0.98 (0.41, 2.26)	0.99 (0.84, 1.17)
Key: CrI, credible interval; HR, hazard ratio; NMA, network meta-analysis; OS, overall survival.		

Table A30.2: Results of the random effects NMA of PFS

Comparison	Random effects model: HR (95% CrI)	Fixed effect model: HR (95% CrI)
Regorafenib versus placebo	0.39 (0.19, 0.80)	0.42 (0.39, 0.45)
Trifluridine/tipiracil versus placebo	0.44 (0.24, 0.79)	0.45 (0.42, 0.48)
Regorafenib versus trifluridine/tipiracil	0.89 (0.35, 2.26)	0.93 (0.85, 1.03)
Key: CrI, credible interval; HR, hazard ratio; NMA, network meta-analysis; PFS, progression-free survival.		

A 31. Priority question: Please provide additional sensitivity analysis for the pooled and individual studies (CORRECT and CONCUR) ITCs using the method of anchored matching-adjusted indirect comparison (MAIC), adjusting for: ECOG PS, previous treatment lines, KRAS status and time from diagnosis.

The pooled population of CORRECT and CONCUR was matched to the population of pooled RECOURSE and TERRA based on the proportion of ECOG PS 0 patients, the proportion of patients having KRAS mutation, the proportion of patients whose time from diagnosis of first metastases >= 18 months and proportion of patients whose previous treatment line <4. Yoshino 2012 did not report time from diagnosis of metastases and previous treatment lines information, so it was not included in the analysis. As stated in question A29, given this analysis did not include age and gender in the matching, the assumptions required for anchored MAIC are unlikely to hold.

Table A31.1 presents the patient characteristics for both populations before and after matching, and the ESS obtained after matching. The ESS after matching (851) was 88.3% the original sample size of 964 patients, which is not surprising given the similarity of all variables except proportion of previous treatment lines < 4 between the populations prior to matching.

Table A31.1: Matching CORRECT and CONCUR (ITT) to RECOURSE and TERRA (ITT)

	Prior to matching		After matching
	CORRECT + CONCUR	RECOURSE + TERRA	CORRECT + CONCUR
N/ESS	N = 964	N = 1206	ESS = 851
% ECOG PS 0	47.8%	44.9%	44.9%
% KRAS status – yes	51.2%	46.1%	46.1%
% time from first diagnosis metastases >=18 months	76.7%	70.8%	70.8%
% previous treatment lines < 4	51.2%	42.3%	42.3%
Key: ESS, effective sample size; ITT, intention-to-treat.			

Results of the anchored MAIC for OS are presented in Table A31.2, along with an unadjusted Bucher comparison performed to assess the impact of matching on results. Results showed weighting had minimal impact on the HR. Results of the anchored MAIC suggested regorafenib had a very similar hazard of death compared with trifluridine/tipiracil with an HR of 0.95 (0.77, 1.18).

Table A31.2: Results of OS – CORRECT (ITT) + CONCUR (ITT) versus RECOURSE (ITT) + TERRA (ITT)

Comparison	HR (95% CI)	Source
Regorafenib vs placebo	0.71 (0.60, 0.84)	CORRECT + CONCUR
Trifluridine/tipiracil vs placebo	0.72 (0.63, 0.83)	RECOURSE + TERRA
Regorafenib vs trifluridine/tipiracil	0.99 (0.79, 1.22)	Bucher ITC
Regorafenib vs placebo (adjusted)	0.68 (058, 0.81)	CORRECT + CONCUR (weighted)
Regorafenib vs trifluridine/tipiracil (adjusted)	0.95 (0.77, 1.18)	Anchored MAIC
Key: CI, confidence interval; HR, hazard ratio; ITC, indirect treatment comparison; ITT, intention-to-treat; OS, overall survival; vs, versus.		

Results of the anchored MAIC for PFS are presented in Table A31.3, along with an unadjusted Bucher comparison to assess the impact of matching on results. Given the same set of weights was used to adjust PFS as OS, matching resulted in a minimal impact on PFS. Results of the anchored MAIC suggested patients treated by regorafenib had a very similar hazard of progression compared with trifluridine/tipiracil, with an HR of 1.00 (0.81, 1.22).

Table A31.3: Results: PFS – CORRECT + CONCUR (ITT) versus RECOURSE + TERRA (ITT)

Comparison	HR (95% CI)	Source
Regorafenib vs placebo	0.45 (0.39, 0.52)	CORRECT + CONCUR
Trifluridine/tipiracil vs placebo	0.44 (0.39, 0.51)	RECOURSE + TERRA
Regorafenib vs trifluridine/tipiracil	1.02 (0.84, 1.25)	Bucher ITC
Regorafenib vs placebo (adjusted)	0.44 (0.38, 0.52)	CORRECT + CONCUR (weighted)
Regorafenib vs trifluridine/tipiracil (adjusted)	1.00 (0.81, 1.22)	Anchored MAIC
Key: CI, confidence interval; HR, hazard ratio; ITC, indirect treatment comparison; MAIC, matching-adjusted indirect comparison; PFS, progression-free survival; vs, versus.		

A 32. Priority question: The CS presents a set of sensitivity analysis of the fixed-effects NMA in order to investigate the differences between the studies. In the PFS sensitivity analysis presented in the forest plot on Figure 13, the fixed-effects NMA scenario including only the CONCUR, TERRA and Yoshino 2012 studies (sensitivity analysis 3), was the only sensitivity analysis that significantly improved PFS results of regorafenib over T/T. According to the company the three trials in this analysis "… ***are more representative of treatment in the UK setting*** **" (p. 70). However, these trials all included only Asian patients and some of them had received prior anti-VEGF treatment.**

a) Please comment further on these results and specifically on how they relate to the efficacy of regorafenib regarding race and previous anti-VEGF treatment.

General comment on the results

The five randomised trials for regorafenib (CORRECT, CONCUR) and trifluridine/tipiracil (RECOURSE, Yoshino 2012, TERRA) all show benefit of active treatment, but with differences in the absolute benefit observed between trials – see Table A32.2 and table A32.2. As a consequence of the different results between

trials it can be anticipated that including/excluding different trials as a part of sensitivity analyses would provide different estimates of relative efficacy between regorafenib and trifluridine/tipiracil. However, the difference in efficacy between trials for the same treatment cannot be explained. The most reliable estimate of efficacy is that provided by meta-analysis of all trials.

The sensitivity analysis referred to in the question was the only one returning a significant result, which happened to favour regorafenib. As described in the CS (p70), "despite there being no evidence that ethnicity is prognostic or a treatment effect modifier this sensitivity analysis was conducted for completeness". Although we believe the prior treatments received in these three trials make them more representative of clinical practice in England (specifically a significant proportion of patients had not received anti-VEGF), we consider the benefit favouring regorafenib in the sensitivity analysis referred to in the question to be a chance-effect. The clinicians we have consulted consider the two treatments to be comparable.

Table A32.1: Hazard ratios for OS reported across studies (Table 7 from CS)

Study Name	Treatment	N	Hazard ratio (95% CI)
CORRECT	Regorafenib 160 mg	505	0.77 (0.64, 0.94)
	Placebo	255	-
CONCUR	Regorafenib 160 mg	136	0.55 (0.40, 0.77)
	Placebo	68	-
RECOURSE	Trifluridine/tipiracil	534	0.68 (0.58, 0.81)
	Placebo	266	-
TERRA	Trifluridine/tipiracil	271	0.79 (0.62, 0.99)
	Placebo	135	-
Yoshino 2012	Trifluridine/tipiracil	135	0.56 (0.39, 0.81)
	Placebo	157	-
Key: CI, confidence interval; NMA, network meta-analysis; OS, overall survival.			

Table A32.2: Hazard ratios for PFS reported across studies (Table 8 from CS)

Study Name	Treatment	N	Hazard ratio (95% CI)
CORRECT	Regorafenib 160 mg	505	0.49 (0.42, 0.58)
	Placebo	255	-
CONCUR	Regorafenib 160 mg	136	0.31 (0.22, 0.44)
	Placebo	68	-
RECOURSE	Trifluridine/tipiracil	534	0.48 (0.41, 0.57)
	Placebo	266	-
TERRA	Trifluridine/tipiracil	271	0.43 (0.34, 0.54)
	Placebo	135	-
Yoshino 2012	Trifluridine/tipiracil	135	0.41 (0.28, 0.59)
	Placebo	157	-
Key: CI, confidence interval; NMA, network meta-analysis; PFS, progression-free survival.			

Race and efficacy of regorafenib

This question overlaps with that of A18 and A19 which should also be referred to. We don't consider that race/ethnicity has any relevance in terms of the generalisability of results to clinical practice in England.

Prior anti-VEGF therapy and efficacy of regorafenib

This question overlaps with A22 which should also be referred to.

From a mechanism of action perspective is might be anticipated that patients who have not received anti-VEGF treatment have a greater potential to benefit from regorafenib as a result of regorafenib's anti-VEGF activity. However, we do not believe that this potential would explain the benefit of regorafenib observed in this sensitivity analysis. We believe the result to be a chance effect and that regorafenib and trifluridine/tipiracil are comparable in respect of PFS and OS with a HR on or around the null.

b) Please conduct a subgroup analysis of the ITC, which includes only patients with no prior anti-VEGF treatment.

Limited data are reported for patients with no prior anti-VEGF treatment (a summary has been presented in Table A32.3) and where data is reported sample sizes are low. The only evidence available for trifluridine/tipiracil was for patients with no prior bevacizumab treatment from Yoshino 2012 and sample sizes were considered too

small to use this study in an ITC (only 25 patients in the trifluridine/tipiracil arm and 10 patients in the placebo arm).

Table A32.3: Summary of data available for patients with no prior anti-VEGF treatment

	Number of patients with no prior anti-VEGF treatment			No prior anti-VEGF subgroup: HR (95% CI)	
	REG	**T/T**	**PBO**	**OS**	**PFS**
CORRECT	0	NA	0	NA	NA
CONCUR	80	NA	43	0.47 (0.31, 0.71)	0.28 (0.18, 0.42)
RECOURSE	NA	0	1	NA	NA
TERRA	NA	194	91	NR	NR
Yoshino 2012	NA	25	10	0·37 (0·16, 0·86)*	NR
Key: CI, confidence interval; HR, hazard ratio; NA, not applicable; NR, not reported; OS, overall survival; PBO, placebo; PFS, progression-free survival; REG, regorafenib; T/T, trifluridine/tipiracil **Note:** * Data based on patients with no prior bevacizumab treatment					

A 33. Comparability of populations for ITC:

a) As per Table 9 (section B.3.1.8.) the median age of patients in the placebo arm of RECOURSE was 63 years whereas it was only 55.5 years in the placebo arm of CONCUR study (85% vs 56%).

b) As per Table 11 (section B.3.1.8.), 55% of patients in the treatment arm of the TERRA study did not receive previous targeted biological treatment compared with 0% in e.g., CORRECT or RECOURSE.

c) As per Table 14 (section B.3.1.8.), 78% of patients had the ECOG PS 1, versus only 33% in the Yoshino 2012.

d) As per Table 17 (section B.3.1.8.), 24% of patients had received 3 prior treatment lines in the CONCUR study compared with 85% in Yoshino 2012.

e) Please indicate how these populations were comparable in the ITC.

As described in B.3.1.8, it is well acknowledged that there are differences between studies in terms of age (percentages of patients < 65 years), prior targeted biological treatment and ECOG PS. Data reported in Table 17 for prior treatment lines should

be interpreted with high levels of caution given that for CONCUR and CORRECT the data are for previous treatment lines on or after diagnosis of metastases and the data from RECOURSE and TERRA were for the total number of prior regimens. Further, the 3 prior lines percentages for Yoshino also include patients with 4+ prior lines (i.e. 24% is not comparable with 85%).

Differences between studies are important in an anchored ITC when characteristics are identified as treatment effect modifiers. Only prior targeted treatment, age and gender were identified as potential treatment effect modifiers. Anchored MAICs adjusting for these characteristics had very little impact on the results. It was therefore concluded that the differences observed between populations were not important enough to discredit an ITC using NMA methods.

Further, the sensitivity analyses presented in B.2.9 investigated some of the differences highlighted: average age was relatively similar between CORRECT and RECOURSE; prior targeted treatment was the same between CORRECT and RECURSE and also similar between CONCUR and TERRA; and ECOG PS was more similar between CONCUR and TERRA. Overall results from sensitivity analyses were consistent with the primary analysis which further supports the use of an NMA.

A 34. In Appendix D, in the summary of treatment effect modifiers, the company stated "*Prior targeted treatment, age and gender were identified as potential treatment effect modifiers through a combination of published subgroup analyses, exploring the CORRECT and CONCUR patient-level data and clinical opinion. However, the evidence for treatment effect modification for all three characteristics was relatively weak."* (p. 24). On page 25 (section B.3.1.8.), mentions that "*no characteristics were identified as potential effect modifiers for OS"*. However, age and ECOG PS were identified as potential effect modifiers for PFS. Please explain this discrepancy. Please provide further discussion on this statement.

Prior targeted treatment, age and gender were identified as the final list of potential treatment effect modifiers as these characteristics were identified as having stronger evidence for treatment effect modification through the investigations into individual trials (interaction p-values <0.1) and were also validated by clinicians.

Page 25 – page 26 of the CS appendices summarizes the investigations of each trial individually and for the investigations into the CORRECT data – it is stated that "for PFS, age and ECOG PS were noted as potential effect modifiers, however, only age had an interaction p-value < 0.1".

A 35. On page 25 (section B.3.1.8.), the company states that "*clinical experts highlighted that overall treatment benefit appear to be reasonably consistent across all subgroups*". Please supply relevant data to support this assertion.

Appendix E, pages 53 to 67, provides the forest plots for OS and PFS for a wide range of subgroups. In the advisory board the clinical experts, having seen these data, agreed that overall treatment benefit appeared to be reasonably consistent.

It must be considered however, that the CORRECT and CONCUR trials were powered and randomised at the ITT level. In this context the results in individual subgroups should not be overinterpreted.

A 36. The company has provided Table 22 in Appendix D, presenting the results of quality assessment of studies included in the ITC. Please provide a commentary for the outcomes of the assessment.

The majority of studies reported low risk of bias for all quality assessment criteria. All studies were judged as having low risk of bias in terms of appropriateness of randomization or allocation concealment, similarity of baseline characteristics across treatment group within each study, study blinding, any unexpected imbalances in withdrawals between groups and statistical analysis methodologies. In terms of outcome selection and reporting bias the risk of bias was low in 80% of the studies as a study by Yoshino and colleagues measured more outcomes than reported in the clinical registry protocol (Yoshino et al 2012). Also, please see below an updated summary of quality assessment of studies included in the ITC (the question "Did the analysis include an intention-to-treat analysis?" marked as Yes for Yoshino 2012 in this updated table):

Table A36.1: Quality assessment of studies included in the ITC using NICE checklist

Study details	Randomization appropriate?	Allocation concealment adequate?	Groups similar at the outset of the study in terms of prognostic factors?	Blinding to treatment allocation?	Unexpected imbalances in drop-outs between groups?	Authors measured more outcomes than they reported?	Did the analysis include an intention-to-treat analysis?
Grothey 2013	Yes	Yes	Yes	Yes	No	No	Yes
Li 2015	Yes	Yes	Yes	Yes	No	No	Yes
Mayer 2015	Yes	Yes	Yes	Yes	No	No	Yes
Xu 2017	Yes	Yes	Yes	Yes	No	No	Yes
Yoshino 2012	Yes	Yes	Yes	Yes	No	Yes	Yes
Key: NICE, National Institute for Health and Care Excellence							

A 37. Priority question. Please provide the code used for executing the ITC analysis in R along with the information on any additional packages used (besides the reported 'gemtc' package) and the datasets.

We have provided the JAGS model script ("fixed_effects_consistency_model.txt") used to perform the continuous endpoint (PFS and OS) meta-analyses. Input datasets ("PFS NMA data.xlsx" and "OS NMA data.xlsx") are also provided. We are unable to share the R code used to execute these NMAs, however the analyses can be replicated with the provided input datasets, JAGS model script, and the following parameterisation details.

Fixed effects models, assuming normal likelihood with an identity link function, were fit to the difference in mean change from baseline data for the PFS and OS endpoints. Vague priors were used [N(0, 5625), om.scale argument of mtc.model() function set to 5]. Three Markov chains were simulated. 50000 burnin and 50000 sampling iterations were run with thinning interval of 1. Autocorrelation, BGR, posterior density plots, and trace plots, were assessed.

The three main functions used for model fitting were: *mtc.network(), mtc.model()* and *mc.run()*

In addition to 'gemtc' the following R packages were also utilised: 'dplyr', 'mautils', 'readxl', 'Hmisc', 'XLConnect', 'officer', 'flextable' and 'magrittr'.

A 38. As per subgroup analyses (Figure 2, section B.3.1.8. of Appendix D) there was a higher ratio of patients with primary cancer site of both colon and rectum experiencing benefit in the placebo group compared with regorafenib. Please comment on these results.

[Company: please enter your answer to this question here]

The 'colon and rectum' subgroup had a small number of patients (n = 44) and events (n = 22). The trial was not powered or designed to assess efficacy according to the primary site of the disease and the results of individual subgroups should not be overinterpreted.

Adverse events

A 39. As per the section B.5.1.2. of the Appendix F– 27 March 2014 safety the final lock date is used. Please provide more recent data.

The most recent data has been provided in the CS.

A 40. Priority question. In the principal findings in Section B.2.13.1 the company states that "Dose modifications owing to AEs were higher with regorafenib compared with placebo (CORRECT, 67% versus 23%; CONCUR, 71% versus 16%); however, permanent discontinuations due to TEAEs were also higher for regorafenib (CORRECT, 17.6% versus 12.6%; CONCUR, 14.0% versus 5.9%)." (p. 95). Given the high proportion of patients in both trials experiencing AEs leading to dose modification or discontinuation:

a) Please comment on the effect on efficacy in the context of dose modification or discontinuation due to AEs.

As indicated in section B.2.4.1 of the CS, the formal efficacy analyses were based on the ITT populations. As a result of this the efficacy analyses already accounts for dose modifications and discontinuations. The efficacy of regorafenib is not reliant on receiving the full licensed dose.

b) Please discuss the generalisability of dose modification or discontinuation to UK clinical practice and the effect on efficacy that any differences might cause.

The tolerability of regorafenib is not expected to be different in the UK population, and consequently dose modifications (reductions or interruptions) can be expected to be broadly in line with those of the CORRECT and CONCUR studies.

A 41. In Section B.2.11.1.5 and in Section B.2.11.2.5 of Document B, AEs of special interest are reported. Please clarify how these AE were defined.

The AEs of special interest were defined based on data from Phase I/II studies with regorafenib and from the known pharmacological properties of other small molecule tyrosine kinase inhibitors in this drug class. In CORRECT and CONCUR, the investigator immediately notified the sponsor when any of the following occurred: acute renal failure (any grade) or severe proteinuria (Grade 3); interstitial lung disease; acute cardiac failure; clinically significant bleeding (Grade ≥ 3); potentially severe skin reactions (Stevens-Johnson syndrome, toxic epidermal necrolysis, erythema multiforme); and acute liver failure.

Section B: Clarification on cost-effectiveness data

Population

B 1. Priority question. As indicated in the company submission anti-VGEF is not indicated for patients which fall into the scope of this submission in the UK. However, a considerable part of the treatment population in both the CORRECT and CONCUR trial received anti-VGEF treatment. Please conduct a subgroup analysis for those with no prior anti-VEGF treatment and incorporate the results of any subgroup analysis of the ITC, as requested in A32.

No subgroup analysis for the anti-VEGF naïve population is feasible versus trifluridine/tipiracil – this was indicated in our response to question A32b. The reason no ITC was possible is because the only evidence available for trifluridine/tipiracil was for patients with no prior bevacizumab treatment from Yoshino 2012 - and sample sizes were considered too small to use this study in an ITC (only 25 patients in the trifluridine/tipiracil arm and 10 patients in the placebo arm).

A post-hoc subgroup analysis (from the CONCUR trial) by anti-VEGF treatment was presented in question A22. The OS hazard ratio for patients who had not received anti-VEGF was lower (HR: 0.470; 95% CI: 0.309, 0.714) compared to patients who received prior anti-VEGF (HR: 0.726; 95% CI: 0.430, 1.224). In the UK anti-VEGF is not recommended and therefore patients will not have received this as prior treatment.

Model structure

B 2. The NICE Decision Support Unit (DSU) technical support document (TSD) 19 recommended the use of state transition models (STMs) alongside partitioned survival models (PSMs) to verify the plausibility of PSM extrapolations and to explore key clinical uncertainties in the extrapolation period.

a) Please justify the use of a partitioned survival approach given the issues highlighted in NICE DSU TSD 19, particularly regarding the extrapolation of PFS and OS while assuming structural independence between these endpoints.

Whilst our partitioned survival approach does not explicitly assume structural dependence between PFS and OS, it does implicitly capture the dependence between PFS and OS as observed in CORRECT and CONCUR (i.e. we don't assume a specific form of dependence because we model observed data which is itself a function of a latent dependency). Given the relative maturity of the CORRECT and CONCUR data, we believe it is preferable to utilise this data directly using a partitioned survival approach rather than a state transition model reliant on an assumed structural dependence.

When considering this issue, it should be noted that use of a partitioned survival approach is consistent with models accepted in prior comparable NICE mCRC appraisals (e.g., TA668: Encorafenib plus cetuximab for previously treated BRAF V600E mutation-positive metastatic colorectal cancer; TA405: Trifluridine–tipiracil for previously treated metastatic colorectal cancer).

Given these factors, we consider our modelling approach to be appropriate, and preferable to the use of a state transition approach.

b) If deemed necessary, please use state transition modelling to assist in verifying the plausibility of the PSM extrapolations and to address uncertainties in the extrapolation period (NICE DSU TSD 19, recommendation 11).

We do not consider this to be necessary for the reasons outlined in response to question B2a above.

B 3. In the CS, the progression-free health state was defined as "a patient's disease is stable or responding, and not actively progressing".

a) Please elaborate on the exact definition of PFS as applied in the economic model.

In the trials Progressive disease is defined by Response Evaluation Criteria in Solid Tumors (RECIST) criteria , v1.1 (see below), or clinical progression (clinical progression based on the judgement of the investigator).

The definition of PFS as applied in the economic model is identical to that applied in CONCUR and CORRECT.

RECIST Criteria v 1.1

(Eur J Cancer. 2009 Jan;45(2):228-47. doi: 10.1016/j.ejca.2008.10.026)

Progressive Disease (PD): At least a 20% increase in the sum of diameters of target lesions taking as reference the smallest sum on study (this includes the baseline sum if that is the smallest on study). In addition to the relative increase of 20%, the sum must also demonstrate an absolute increase of at least 5mm. Unequivocal progression of existing non-target lesions or the appearance of one or more new lesions will also constitute progressive disease. (Note: the appearance of one or more new lesions is also considered progression). Ascites or pleural effusion will be recorded as disease progression only if proven malignant.

Lymph nodes identified as target lesions should always have the actual short axis measurement recorded (measured in the same anatomical plane as the baseline examination), even if the nodes regress to below 10 mm on study. This means that when lymph nodes are included as target lesions, the 'sum' of lesions may not be

zero even if complete response criteria are met, since a normal lymph node is defined as having a short axis of <10mm.

To achieve unequivocal progression in patients with measurable disease on the basis of the nontarget disease, there must be an overall level of substantial worsening in non-target disease such that, even in presence of SD or PR in target disease, the overall tumour burden has increased sufficiently to merit discontinuation of therapy. A modest 'increase' in the size of one or more nontarget lesions is usually not sufficient to qualify as unequivocal.

In the absence of measurable disease, the same general concepts apply here as noted above.

b) How does this definition compare to PFS in UK clinical practice and PFS in the clinical trials that were used to inform the economic model?

In UK clinical practice progression is also defined by the RECIST criteria i.e. aligning with the definition used in CORRECT and CONCUR.

Intervention and comparators

B 4. Priority question. The company stated that best supportive care (BSC) was included as a comparator in the economic model, *"…mainly because BSC was the comparator in the pivotal trials for regorafenib, rather than it being considered directly relevant to the requested position."* (p. 110) Please elucidate whether BSC was included in the CS and economic model for validation reasons, or whether the ERG should consider BSC as a relevant comparator for regorafenib.

BSC is not a comparator - we are seeking a recommendation for regorafenib as a treatment option alongside trifluridine/tipiracil (i.e. the position for regorafenib in clinical guidelines). This submission was made in response to physician's requests for an alternative to trifluridine/tipiracil. Therefore BSC isn't a comparator.

Treatment effectiveness

B 5.Priority question PFS and ToT are modelled using piece-wise models rather than fully parametric models.

a) Assumption 3 in Table 42 indicates that progression was only assessed every 8 weeks in the CORRECT and CONCUR trials. If this is correct, using KM data to model PFS and ToT may be questionable, especially considering a weekly cycle length and the short median survival. Please justify using KM curves under these circumstances.

We did not use piecewise models in our submission. In our base-case, the full KM data for PFS and ToT were applied directly in the model as this data was nearly complete, and we felt that it was preferable to utilise the observed trial data directly, rather replace it with fully modelled curves. The model only uses parametric models when KM data was no longer available. However, in contrast to piecewise models, these parametric models were fitted using the entire duration of PFS and ToT data.

It should be noted that in clinical practice, evaluation of progression is typically based on an 8-week scan frequency: the same utilised in CONCUR and CORRECT. As a result, real-world evaluated PFS, and subsequently ToT, is likely to mirror the shape of the curves observed in the trials (i.e. evaluation of progression in clinical practice, and so progression to a future progressed disease resource-use health state, will not be smooth because real-world PFS assessments are not continuous). Given this, our modelling of the Kaplan-Meier data will be a better reflection of real life than use of smooth fully parametric curves.

When considering this issue, it should be noted that a sensitivity analysis utilising fully parametric functions for PFS and ToT was provided in the submission (CS document B, Table 48 page 160). This scenario had an immaterial impact on the model results for the comparison of regorafenib to trifluridine/tipiracil: decreasing the NMB from [redacted]

b) Please clarify how many events occurred before and after the cut-point of the KM curves and the parametric models.

As described in the submission (Section B.3.3.2), we used the observed PFS and ToT KM data when this was available, and then extrapolated from this time point onwards using parametric functions (the parametric function was fitted using the entire duration of the data). As the end of the KM data was utilised as the transition to the parametric tails, no events occurred after this point for either ToT or PFS.

c) Please clarify how many patients were at risk at the cut-point.

When the model switches to a parametric function, 0 patients were at risk for PFS for regorafenib and BSC and 1 patient was at risk for ToT in the regorafenib arm.

d) As stated in NICE DSU TSD 21 on flexible methods for survival analysis: "Where a piecewise model is fitted to a single dataset, splitting the data into sections according to time means that sample sizes are reduced in later segments of the curve. This is a particular issue in later sections of the curve, where patient numbers at risk may be very small and the number of observed events may be low, leading to large standard errors and uncertainty when fitting survival models. A key point is that it is the model fitted to the latest section of the curve that is used for extrapolation". Please justify the plausibility of the (extrapolation) approach used for the estimated piecewise models, given the number of patients at risk and observed events (both per treatment) to estimate the tail.

Piecewise models were not used. Parametric fits were based on the full duration of trial data. For further details, see submission (Section B.3.3.2), or our response to question B.5b.

e) As stated in NICE DSU TSD 21 on flexible methods for survival analysis: "the cut-points for the various intervals may be arbitrary and may importantly influence the results of an analysis". Please justify the selected cut-point given the responses above and provide an updated economic model as well as scenario analyses assuming different cut-points (with the parametric survival models estimated from the specific cut-point).

Piecewise models were not used. Parametric fits were based on the full duration of trial data with no cut-point applied. For further details, see submission (Section B.3.3.2), or our response to question B.5b. In addition, a scenario based on fully parametric PFS and ToT input only had a small effect on the results versus trifluridine/tipiracil, decreasing the NMB from [redacted]. So, any scenario that introduces an arbitrary cut-off point to switch from KM data to a parametric function is likely to have a similarly small effect on the outcomes.

To further explore the impact of piecewise modelling on the model outcomes, we have run two additional scenarios in which we only used the first 3 and 6 months of the available KM data, to approximate the impact on piecewise modelling at different cut-off points. This resulted in comparable NMBs of [redacted] for the 3- and 6-month scenario. However, it should be noted that these analyses are not part of the original model design. The analyses were performed manually by using the KM data as input and deleting the KM data after 3 months and 6 months and forcing the model to switch to the log-logistic parametric function once the KM data stops. This is therefore not an available analysis in the submitted model, as this is not part of the intended model design, and the results are only an illustrative approximation of the potential results when piecewise modelling would be used.

f) As stated in NICE DSU TSD 21 on flexible methods for survival analysis: "piecewise models may appear clinically unjustifiable and implausible, if sudden changes in hazards are modelled". Please justify that the piecewise models are clinically justifiable and plausible in this respect.

Piecewise models were not used. Parametric fits were based on the full duration of trial data with no cut-point applied. For further details, see submission (Section B.3.3.2), or question B.5b.

g) Please justify the use of the piecewise models given the responses to the preceding (sub-) questions.

Piecewise models were not used. Parametric fits were based on the full duration of trial data with no cut-point applied. For further details, see submission (Section B.3.3.2), or question B.5b.

h) In line with OS, please implement fully parametric survival models for PFS and ToT individually and present the results.

A sensitivity analysis utilising fully parametric functions for PFS and ToT was provided in the submission (Section B.3.10.3). For this scenario, the log-logistic extrapolation was selected for both PFS and ToT, as it showed the best statistical fit to the PFS and ToT data. This scenario had virtually no impact on the model results for the comparison of regorafenib to trifluridine/tipiracil: decreasing the NMB from [redacted]

To further explore the impact of using fully parametric PFS and ToT functions, we have explored additional scenarios with the other available parametric PFS and ToT options available. When using parametric functions for both PFS and ToT, we applied the same extrapolation for both. None of these scenarios had a significant impact on the model outcomes, with the resulting NMBs ranging from £[redacted] (Table B5.1). In addition, we explored a number of scenarios where PFS and ToT were varied individually, so a parametric function was only used for ToT (Table B5.2) or PFS (Table B5.3), and KM data was used for the other. These resulted in a slightly wider NMB range, from [redacted]. Overall, these scenarios illustrate the small impact using parametric PFS and ToT functions has on the model outcomes.

Table B5.1 : Exploratory model outcomes using fully parametric PFS and ToT functions

Fitted PFS and ToT function	Regorafenib		T/T		NMB
	Total cost	Total QALYs	Total cost	Total QALYs	
KM (base case)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Weibull	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gen. gamma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gamma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Key: NMB, net monetary benefit; PFS, progression-free survival; QALY, quality-adjusted life year; T/T, trifluridine/tipiracil; ToT, time on treatment.

Table B5.2: Exploratory model outcomes using fully parametric functions for ToT only

Fitted ToT function	Regorafenib		T/T		NMB
	Total cost	Total QALYs	Total cost	Total QALYs	
KM (base case)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Weibull	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gen. gamma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gamma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Key: NMB, net monetary benefit; PFS, progression-free survival; QALY, quality-adjusted life year; T/T, trifluridine/tipiracil; ToT, time on treatment.

Table B5.3: Exploratory model outcomes using fully parametric functions for PFS only

Fitted PFS function	Regorafenib		T/T		NMB
	Total cost	Total QALYs	Total cost	Total QALYs	
KM (base case)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Weibull	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gen. gamma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gamma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Key: NMB, net monetary benefit; PFS, progression-free survival; QALY, quality-adjusted life year; T/T, trifluridine/tipiracil; ToT, time on treatment.

i) Please justify your choice of fully parametric curves for PFS and ToT based on the NICE DSU TSD14 criteria assessment.

We believe the direct KM data is the most suitable PFS and ToT input for the modelled base case, as it reflects the trial data and clinical practice.

For the parametric PFS and ToT scenarios described in Section B.3.10.3 of the company submission and in question B.5h, a log-logistic curve was used for both PFS and ToT, as these showed the best statistical fit to the data. Here, statistical fit was considered the best justification, as PFS and ToT were both very mature, leaving only a small tail which has little impact on cost-effectiveness results. Using the AIC BIC data in Table B5.4 and Table B5.5 below, log-logistic was selected as the best fitting curve for both PFS and ToT, in this scenario. However, as shown in B.5h, the model was not sensitive to the choice of PFS or ToT input, with the different parametric PFS and ToT functions resulting NMBs ranging from £[redacted]

Table B5.4: Goodness-of-fit statistics of the pooled progression-free survival extrapolations

Fitted function	Regorafenib		Statistical rank	BSC		Statistical rank
	AIC	BIC		BIC	AIC	
Weibull	2452.934	2461.86	5	871.5151	879.0704	5
Log-normal	2348.506	2357.432	2	803.2672	810.8225	2
Log-logistic	2345.376	2354.302	1	751.179	758.7343	1
Exponential	2529.705	2534.168	7	1022.65	1026.428	7
Generalized gamma	2348.828	2362.217	3	802.6254	813.9584	3
Gompertz	2521.584	2530.51	6	980.991	988.5463	6
Gamma	2412.393	2421.319	4	816.293	823.8483	4
Key: AIC, Akaike information criterion; BIC, Bayesian information criterion; BSC, best supportive care.						

Table B5.5: Goodness-of-fit statistics of the pooled time on treatment extrapolations

Fitted function	Regorafenib		Statistical rank
	AIC	BIC	
Weibull	2540.677	2549.588	5
Log-normal	2483.435	2492.346	2
Log-logistic	2477.802	2486.713	1
Exponential	2554.352	2558.807	6
Generalized gamma	2483.577	2496.943	3
Gompertz	2556.112	2565.023	7
Gamma	2525.736	2534.646	4
Key: AIC, Akaike information criterion; BIC, Bayesian information criterion; BSC, best supportive care.			

B 6. Priority question. It is unclear whether the estimation of parametric survival models is consistent with reported guidance from NICE DSU TSD 14 and 21 on (flexible methods for) survival analyses. Please justify whether the survival analyses in the economic model are currently based on joint or individual modelling and why (please enable both options in the updated economic model). Please provide for the individual and joint survival modelling of OS, PFS and ToT for regorafenib and the comparators, for all full parametric survival models and piece-wise models:

Individual parametric survival models were fitted to the originally submitted pooled regorafenib and placebo data from CORRECT and CONCUR. This was considered appropriate because when patient-level data are available it is unnecessary to rely upon the proportional hazards assumption and apply a proportional hazards (i.e. joint) modelling approach (NICE DSU TSD 14). Both options have been enabled in the economic model.

a) Tables with the numbers of patients at risk, per 3 months.

Numbers at risk from the pooled CORRECT and CONCUR population versus RECOURSE, TERRA and Yoshino have been summarized for OS and PFS in Table B6.1 and Table B6.2, respectively. Kaplan-Meier data for ToT were not available for RECOURSE, TERRA and Yoshino.

Table B6.1: Numbers at risk (OS)

Month	CORRECT and CONCUR (pooled): numbers at risk		RECOURSE: numbers at risk		TERRA: numbers at risk		Yoshino: numbers at risk	
	REG	PBO	T/T	PBO	T/T	PBO	T/T	PBO
0	641	323	534	266	271	135	112	57
3	526	235	459	198	237	114	104	46
6	275	110	294	107	176	78	77	31
9	121	37	137	47	108	48	55	18
12	48	14	64	24	53	19	23	4
15	9	2	23	9	22	7	6	1
18	0	0	7	3	8	1	0	0
21	0	0	0	0	4	0	0	0
24	0	0	0	0	3	0	0	0
27	0	0	0	0	2	0	0	0
Key: PBO, placebo; REG, regorafenib; T/T, trifluridine/tipiracil.								

Table B6.2: Numbers at risk (PFS)

Month	CORRECT and CONCUR: numbers at risk		RECOURSE: numbers at risk		TERRA: numbers at risk		Yoshino: numbers at risk	
	REG	PBO	T/T	PBO	T/T	PBO	T/T	PBO
0	641	323	534	266	271	135	112	57
3	265	37	NR	NR	108	14	31	4
6	69	3	66	2	31	0	17	1
9	22	1	NR	NR	19	0	4	0
12	9	0	5	1	7	0	1	0
15	2	0	NR	NR	0	0	0	0
Key: NR, not reported; PBO, placebo; REG, regorafenib; T/T, trifluridine/tipiracil.								

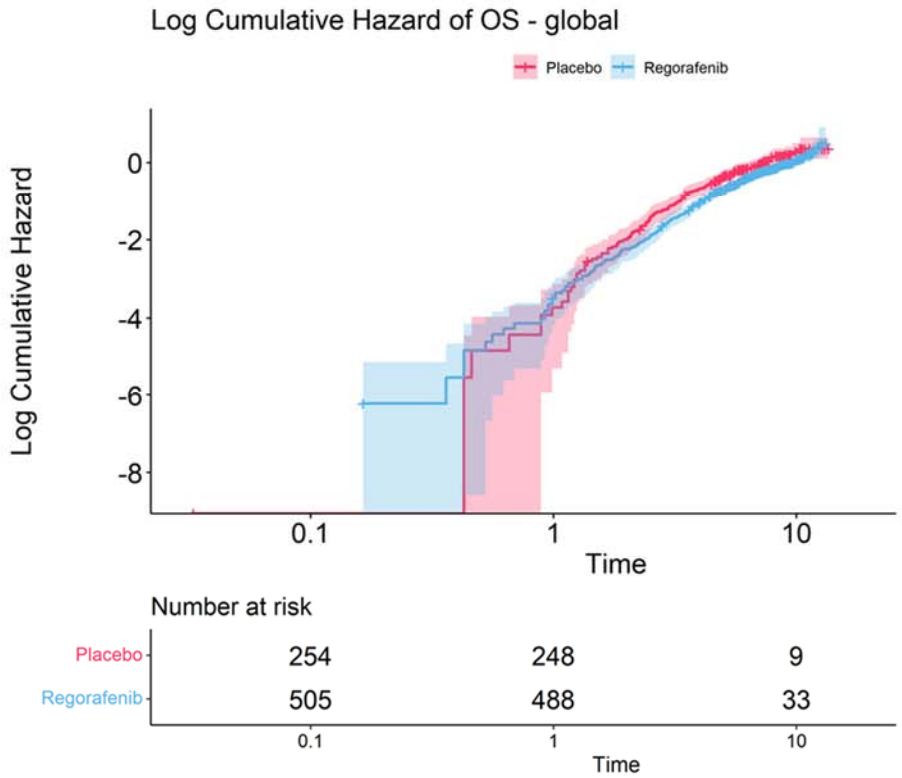
b) To examine the proportional hazard assumption:

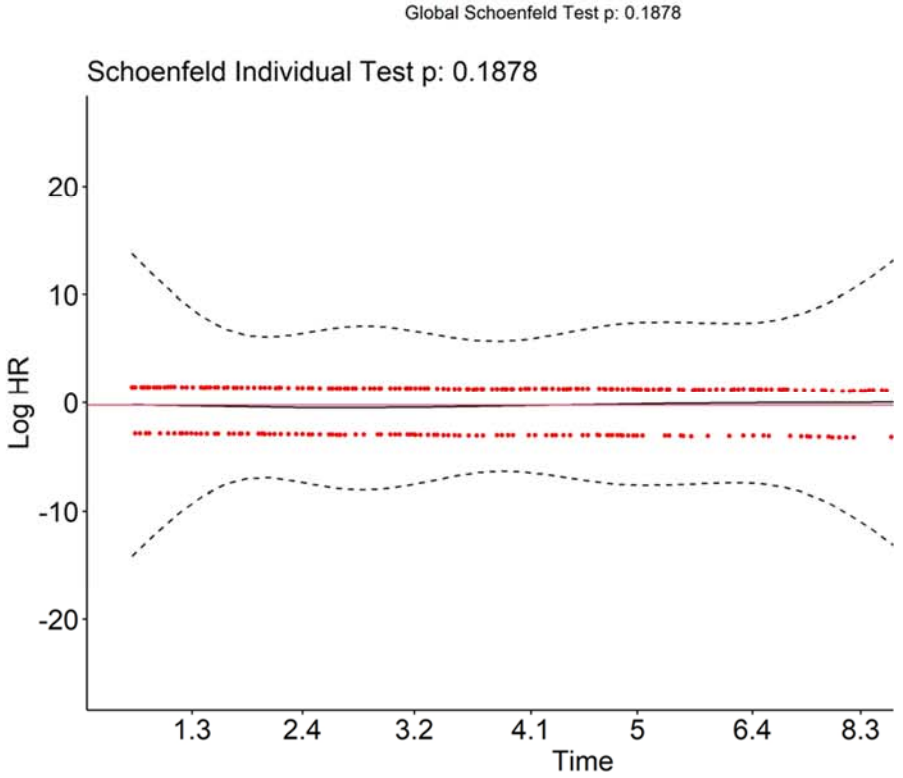
i. Plot the scaled Schoenfeld residuals versus time (all survival curves)

ii. Plot the log cumulative hazard versus log time

Schoenfeld residual plots and log-cumulative hazard plots for OS and PFS from CORRECT and CONCUR were presented in Appendix O of the CS – these have been repeated below. The proportional hazards assumption was shown to hold for all endpoints.

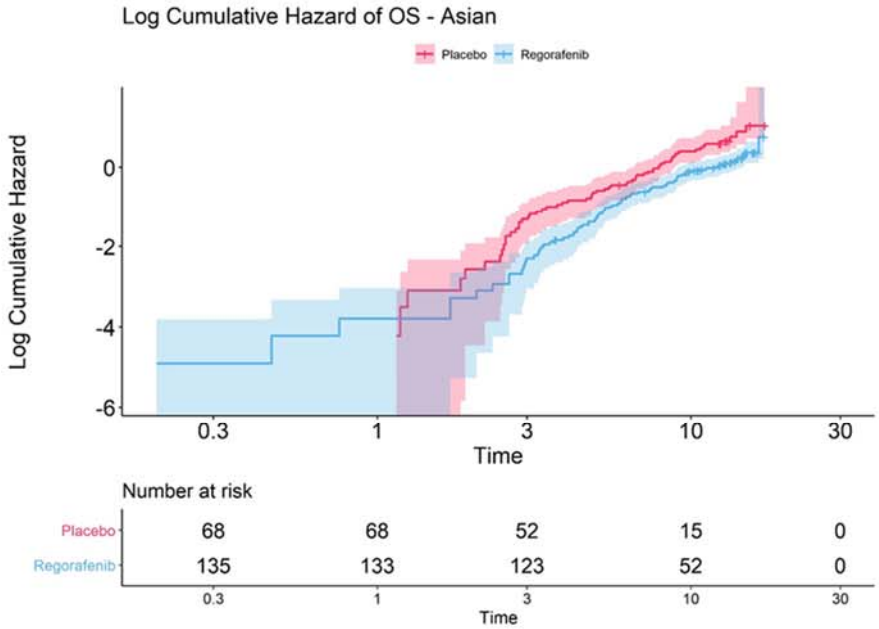
CORRECT OS – log-cumulative hazards plot and Schoenfeld individual test

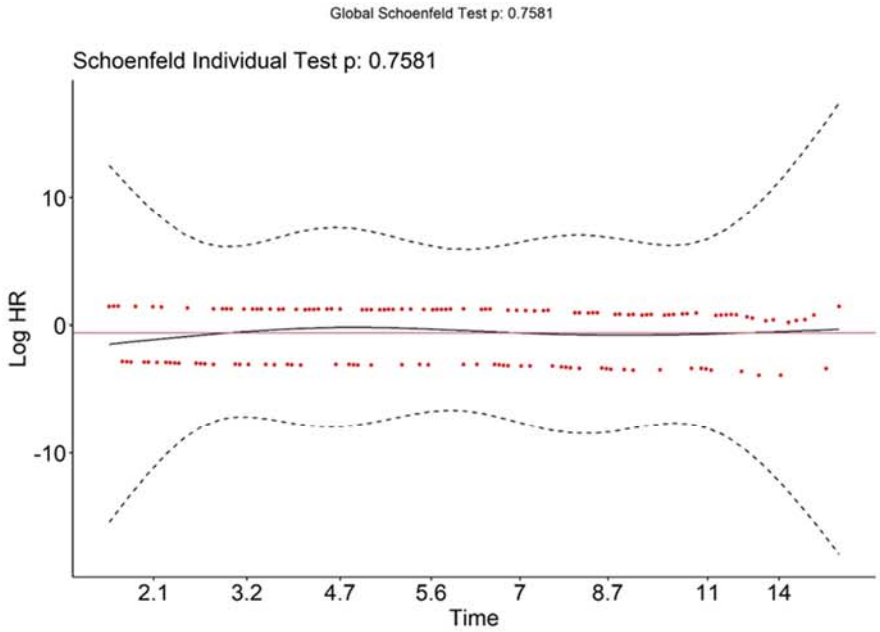




Key: OS, overall survival.

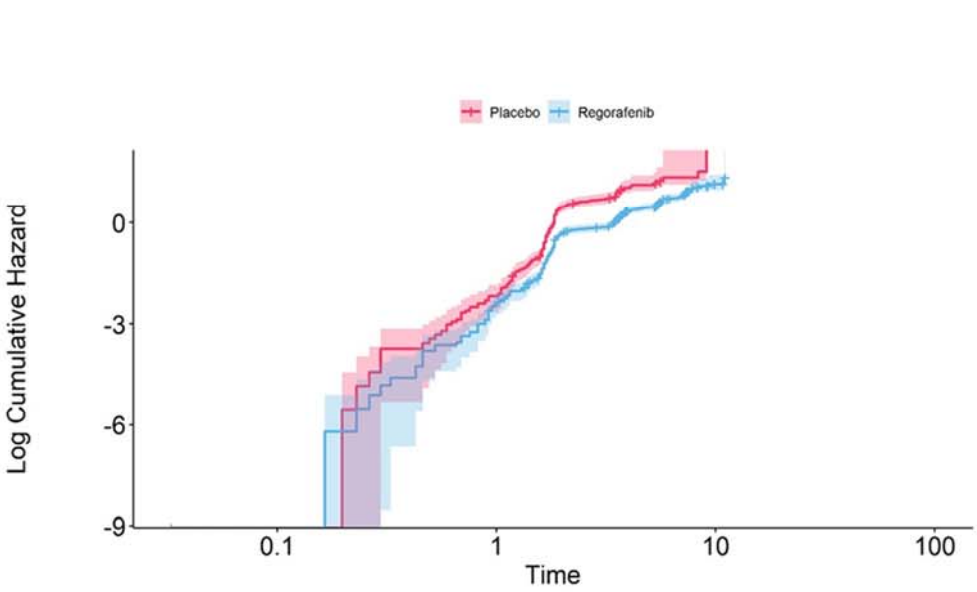
CONCUR OS – Log-cumulative hazards plot and Schoenfeld individual test

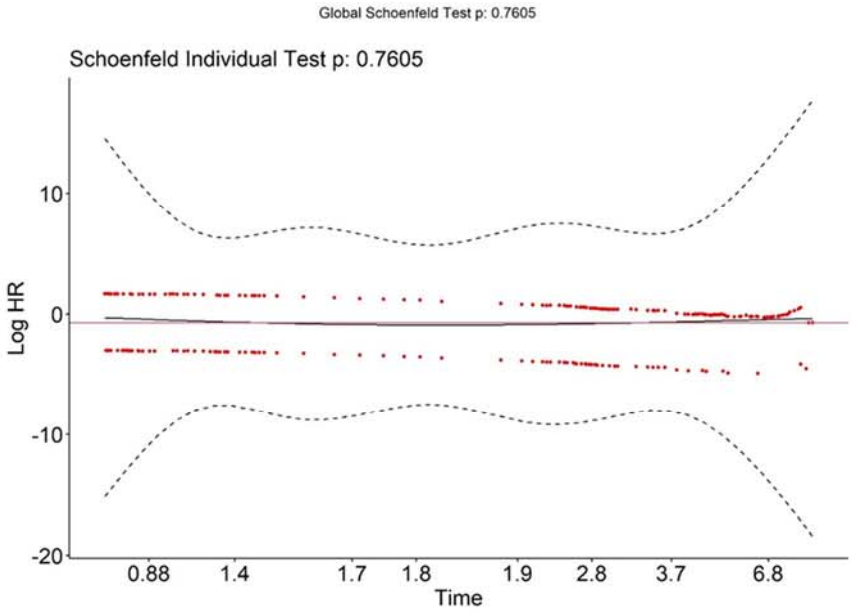




Key: OS, overall survival.

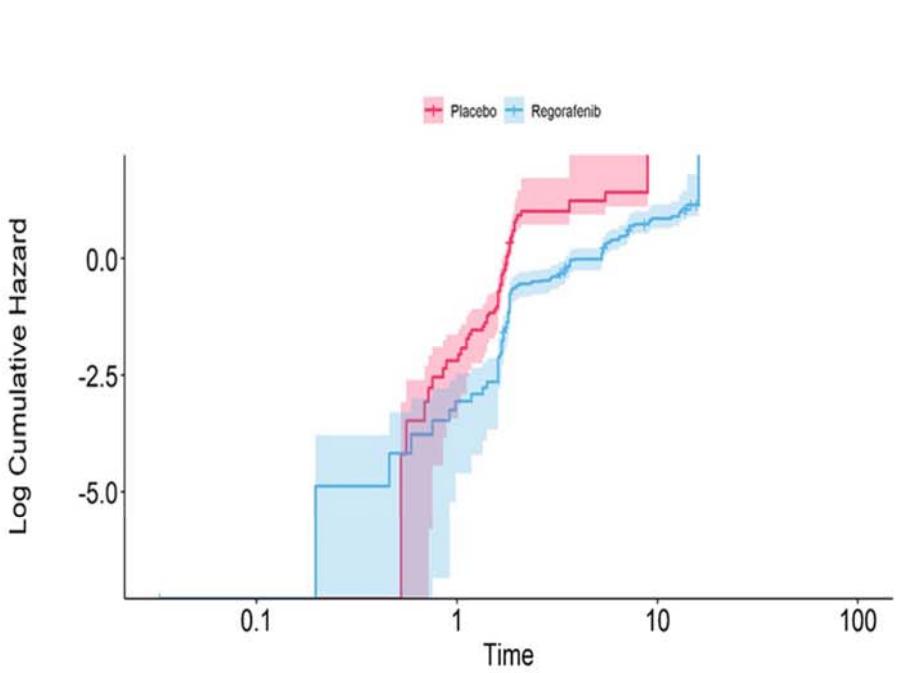
CORRECT PFS – Log-cumulative hazards plot and Schoenfeld individual test

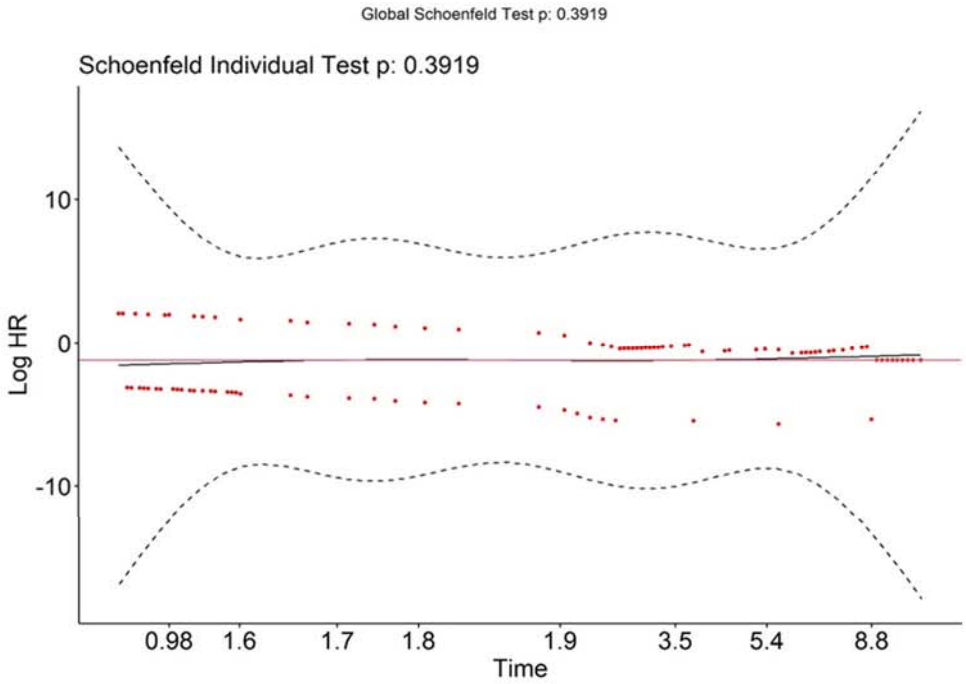




Key: PFS, progression-free survival.

CONCUR PFS – Log-cumulative hazards plot and Schoenfeld individual test





Key: PFS, progression-free survival.

Schoenfeld residual plots and log-cumulative hazard plots for ToT from CORRECT and CONCUR are presented in Figure B6.1 and Figure B6.2.

Figure B6.1: Schoenfeld residual plot and log-cumulative hazard plot for ToT from CORRECT

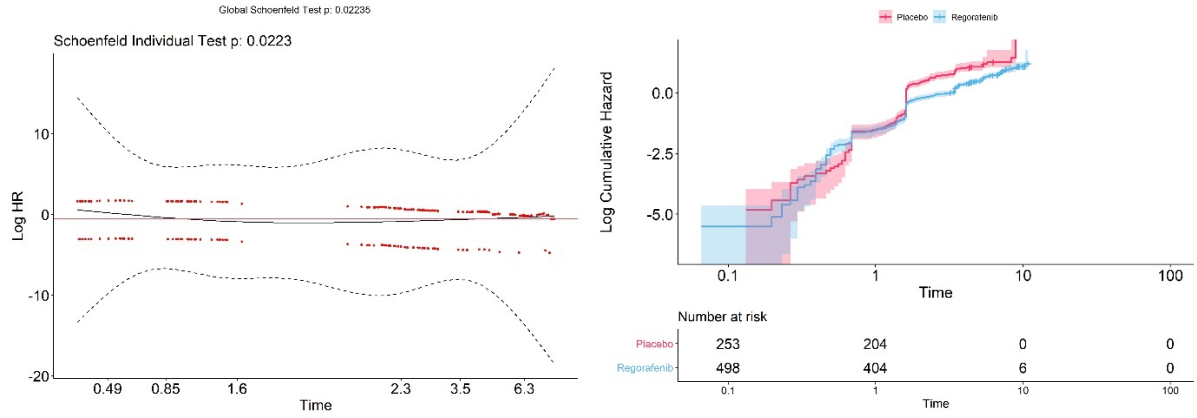
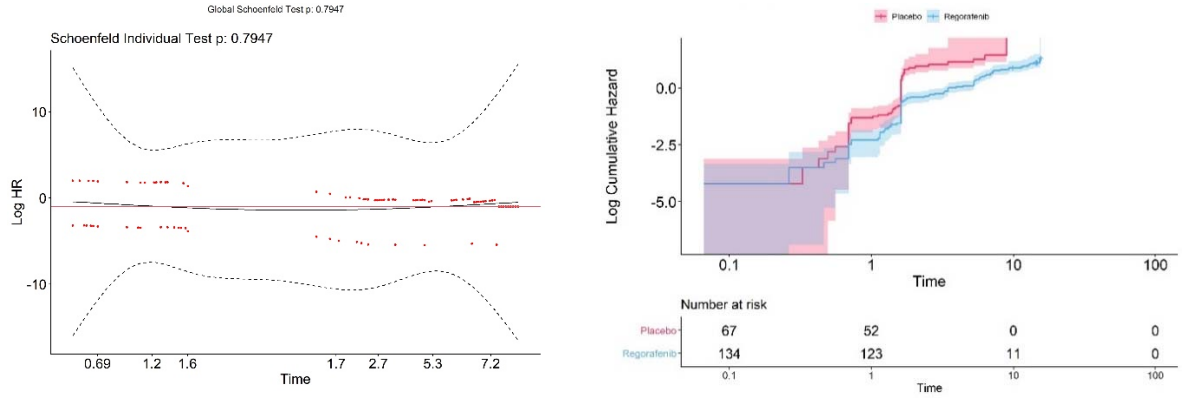


Figure B6.2: Schoenfeld residual plot and log-cumulative hazard plot for ToT from CONCUR



Schoenfeld residual plots and log-cumulative hazard plots for OS and PFS from digitized data from RECOURSE, TERRA and Yoshino 2012 are presented in Figure B6.3 to Figure B6.8. Overall, the plots support that the proportional hazards assumption holds (the log cumulative hazard plots are parallel for the majority of the observed time period and if crossing of curves occurs, this is at the very start of the time period). In TA405, the ERG also noted that the log-cumulative hazard plots for overall survival and for progression-free survival indicated that the proportional hazards assumption would hold.

Figure B6.3: Schoenfeld residual plot and log-cumulative hazard plot for OS from RECOURSE

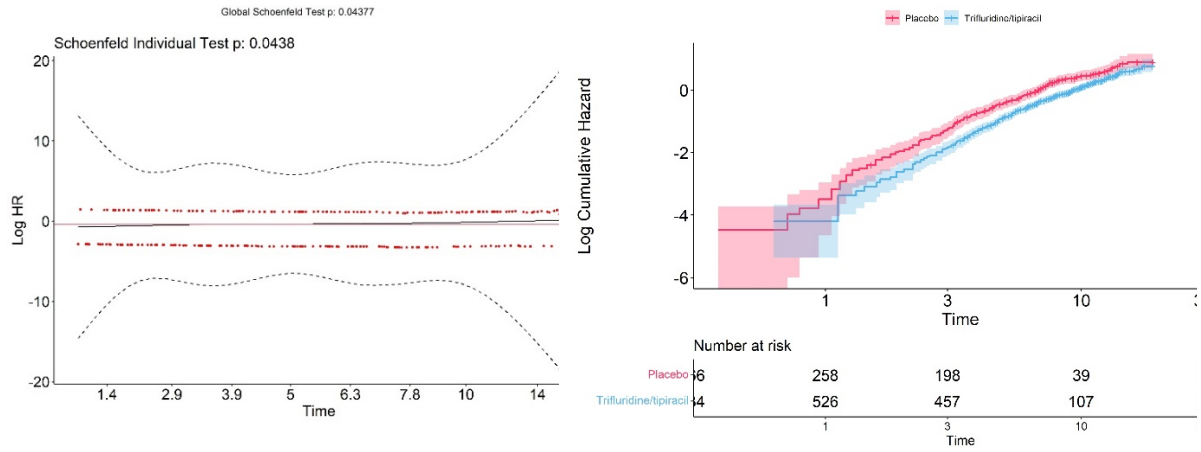


Figure B6.4: Schoenfeld residual plot and log-cumulative hazard plot for PFS from RECOURSE

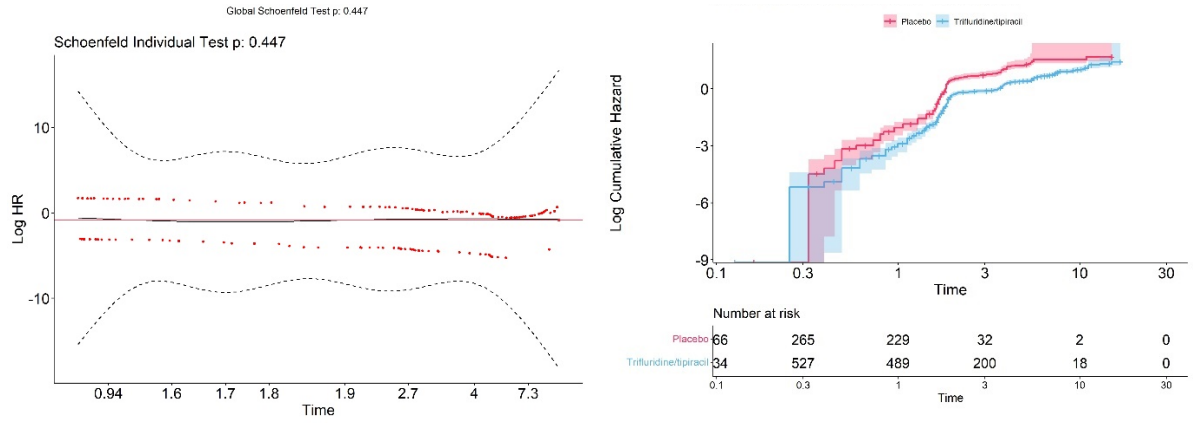


Figure B6.5: Schoenfeld residual plot and log-cumulative hazard plot for OS from TERRA

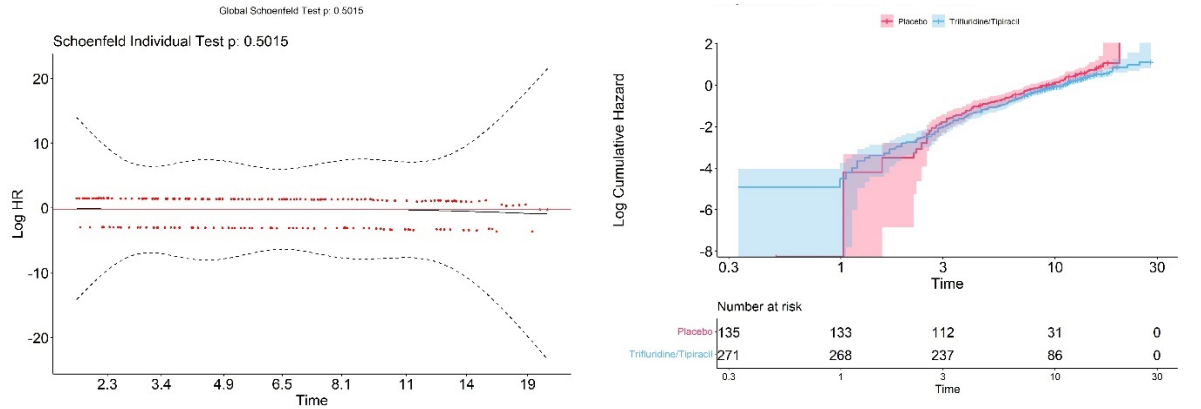


Figure B6.6: Schoenfeld residual plot and log-cumulative hazard plot for PFS from TERRA

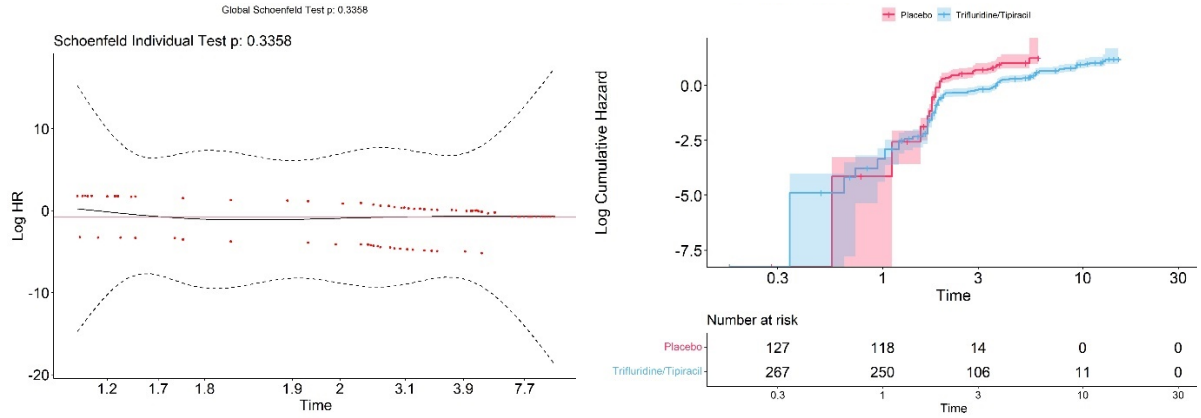


Figure B6.7: Schoenfeld residual plot and log-cumulative hazard plot for OS from Yoshino 2012

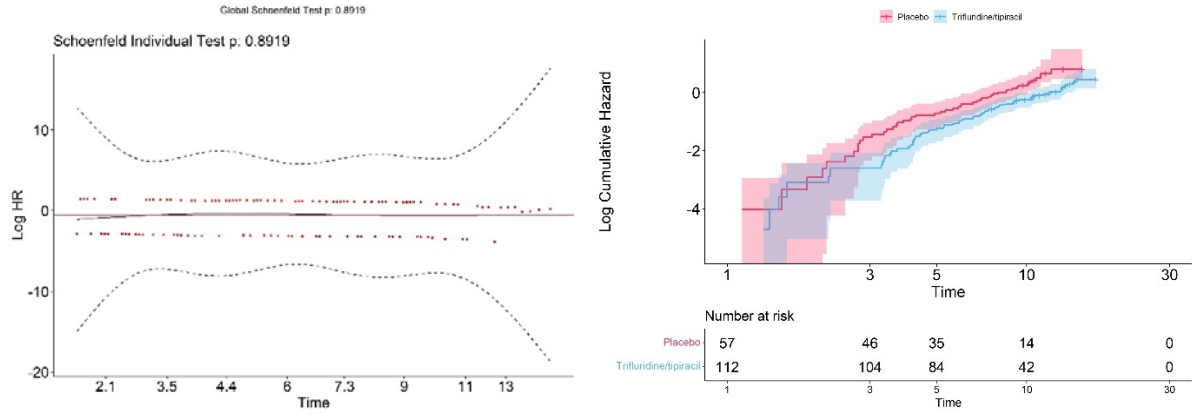
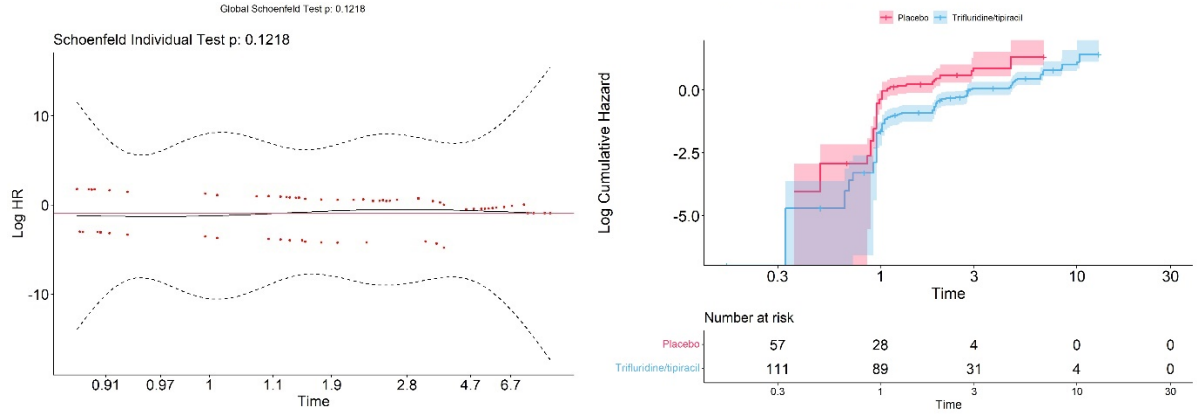


Figure B6.8: Schoenfeld residual plot and log-cumulative hazard plot for PFS from Yoshino 2012



c) To examine the change in hazard function over time:

i. Plot the smoothed hazards over time

Plots of the log smoothed hazard over time have been provided in response to question B6(d).

d) To examine diagnostics of parametric survival models (using the observed data):

i. Plot the cumulative hazard versus time

ii. Plot the log smoothed hazard versus time

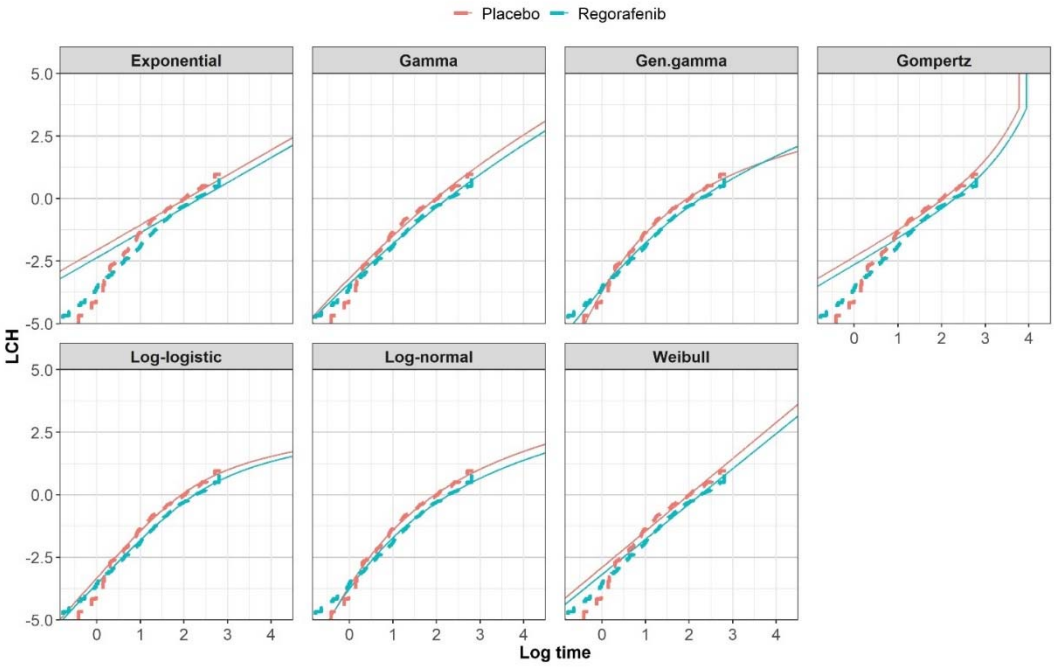
iii. Plot the standard normal quartiles versus log time

iv. Plot the log survival odds versus log time

Overall survival

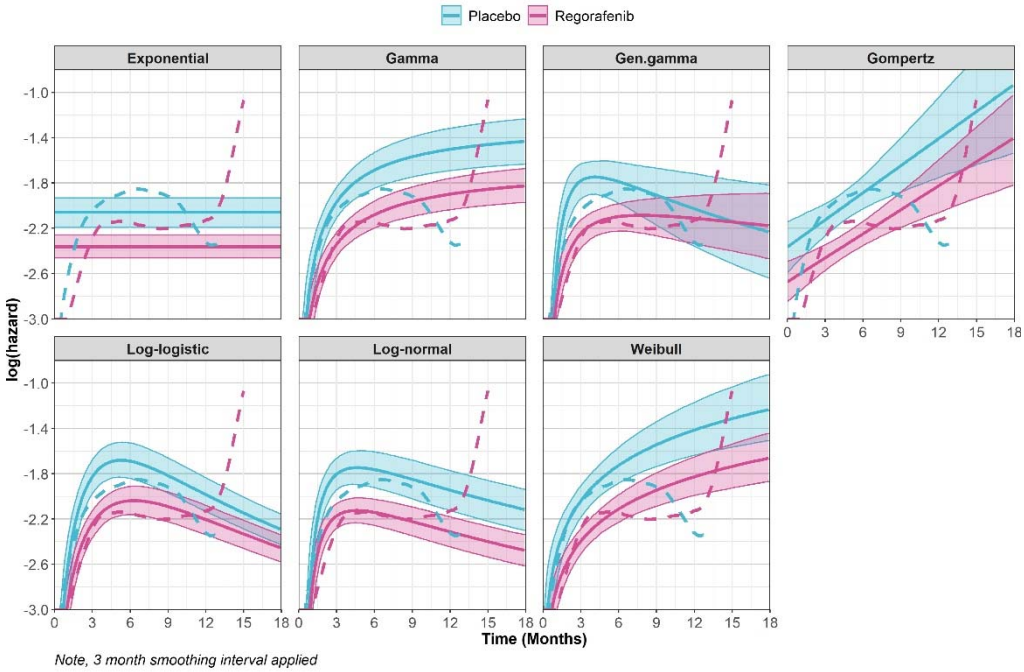
The four plots for the pooled CORRECT and CONCUR OS data are presented in Figure B6.9 to Figure B6.12. Overall, the plots suggest that either the log-logistic, log-normal or generalized gamma models fit the data reasonably well. There is a greater level of uncertainty associated with the plots of the log smoothed hazard versus time as the smoothed hazard relies on the level of smoothing applied. In these plots, there is also more uncertainty towards the end of the data where patient numbers are lower.

Figure B6.9: Cumulative hazard versus time – pooled CORRECT and CONCUR OS data



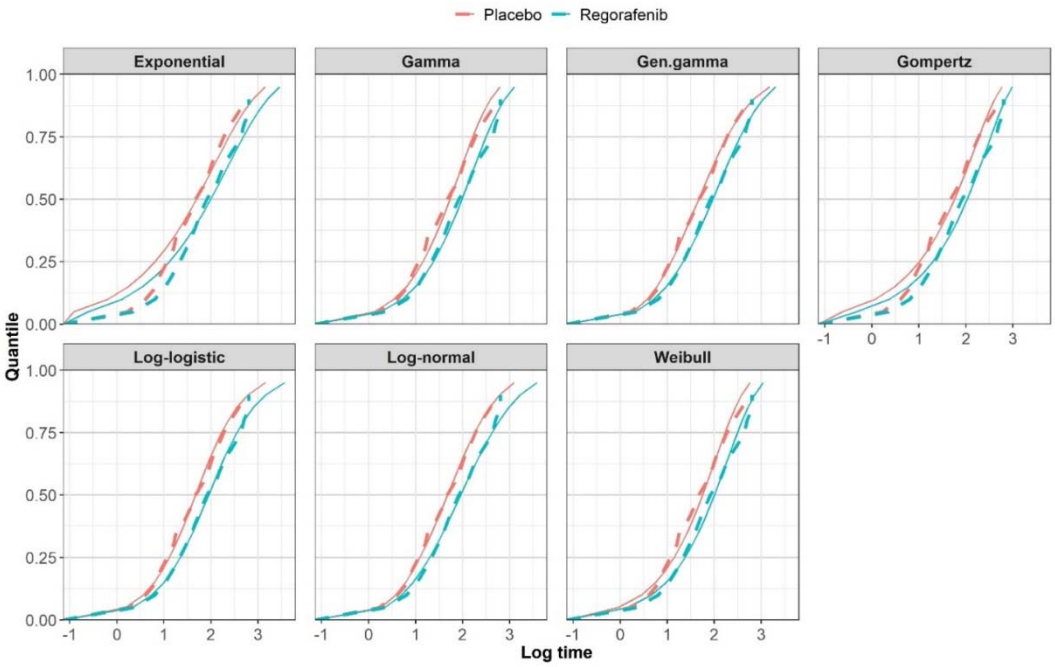
Key: Gen, generalized; LCH, log-cumulative hazard; OS, overall survival.

Figure B6.10: Log smoothed hazard versus time – pooled CORRECT and CONCUR OS data



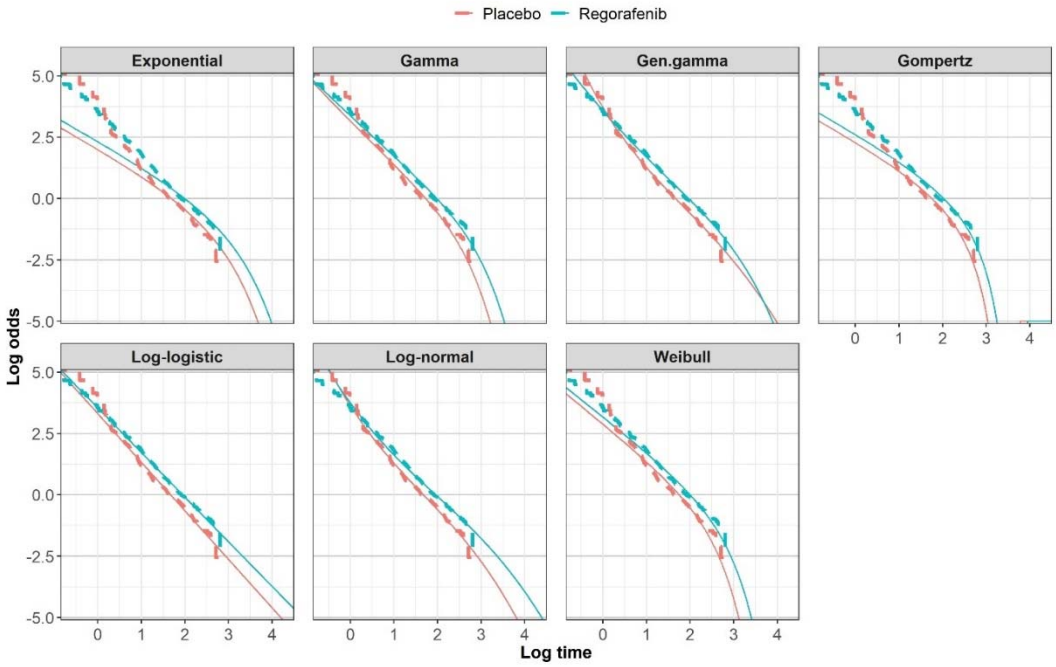
Key: Gen, generalized; OS, overall survival.

Figure B6.11: Standard normal quartiles versus log time – pooled CORRECT and CONCUR OS data



Key: Gen, generalized; OS, overall survival.

Figure B6.12: Log survival odds versus log time – pooled CORRECT and CONCUR OS data

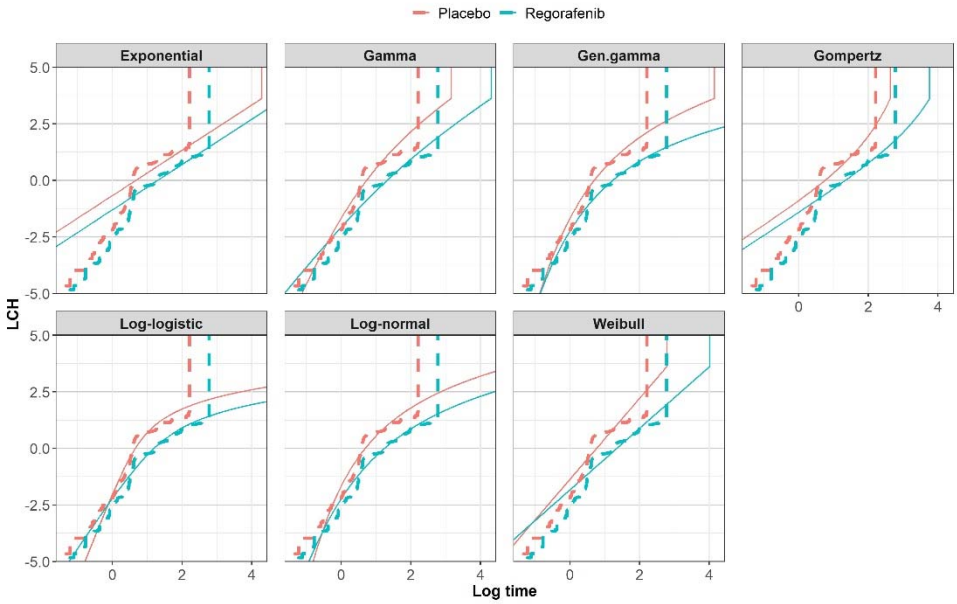


Key: Gen, generalized; OS, overall survival.

PFS

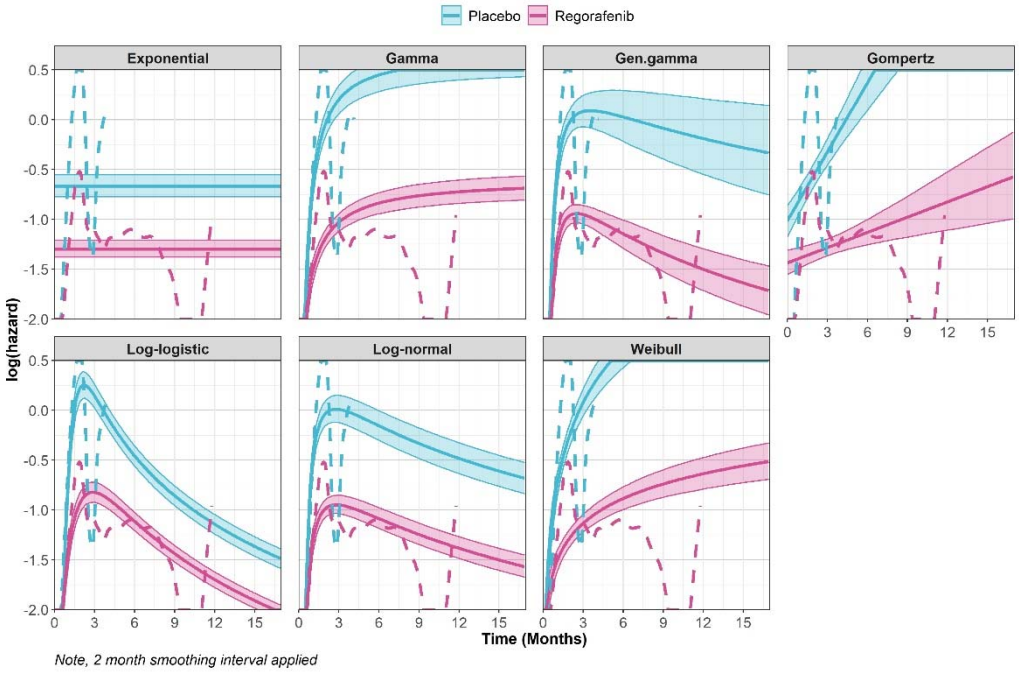
The four plots for the pooled CORRECT and CONCUR PFS data are presented in Figure B6.13 to Figure B6.16. As for OS, plots suggest that the best fitting models are likely to be the log-logistic, log-normal or generalized gamma models. However, the models do not fit the PFS data as well as for the OS data, supporting the use of the Kaplan-Meier data in the economic model.

Figure B6.13: Cumulative hazard versus time – pooled CORRECT and CONCUR PFS data



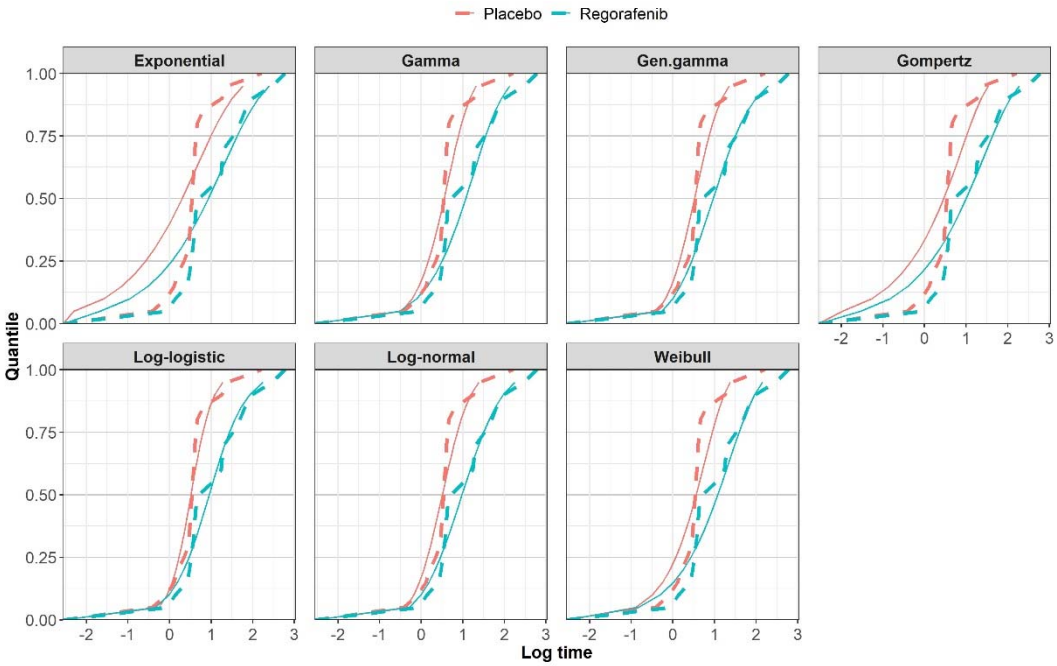
Key: Gen, generalized; LCH, log-cumulative hazard; PFS, progression-free survival.

Figure B6.14: Log smoothed hazard versus time – pooled CORRECT and CONCUR PFS data



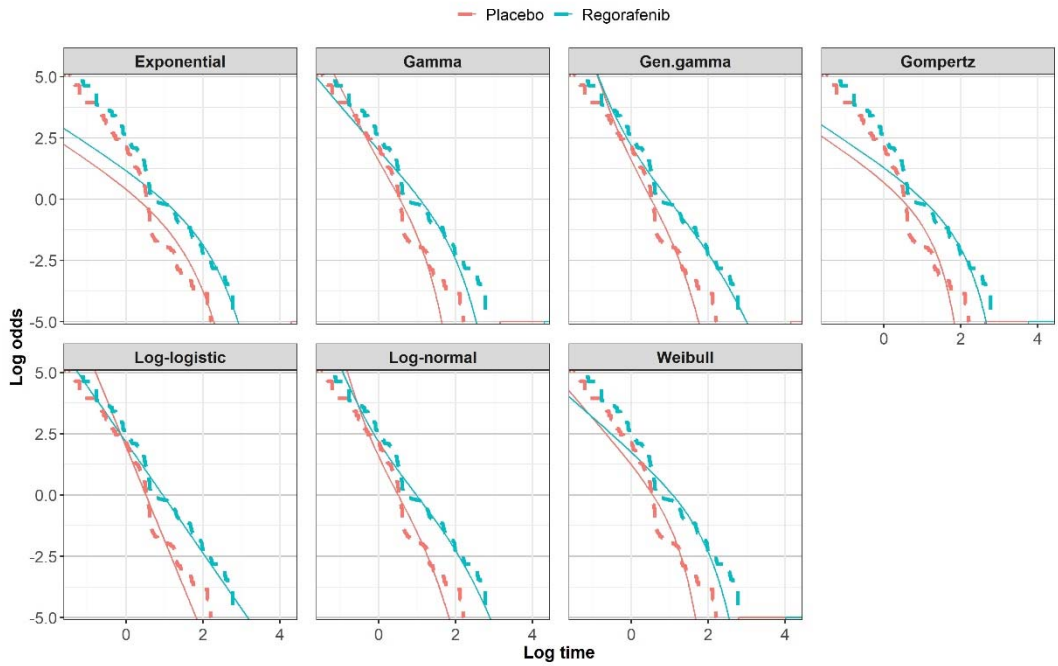
Key: Gen, generalized; PFS, progression-free survival.

Figure B6.15: Standard normal quartiles versus log time – pooled CORRECT and CONCUR PFS data



Key: Gen, generalized; PFS, progression-free survival.

Figure B6.16: Log survival odds versus log time – pooled CORRECT and CONCUR PFS data

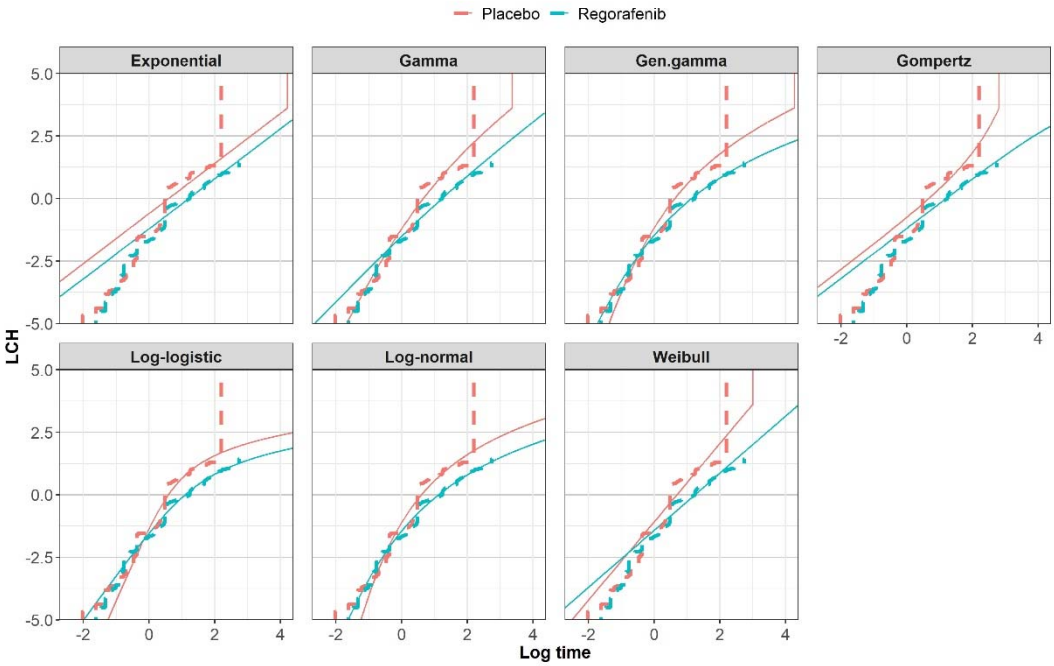


Key: Gen, generalized; PFS, progression-free survival.

ToT

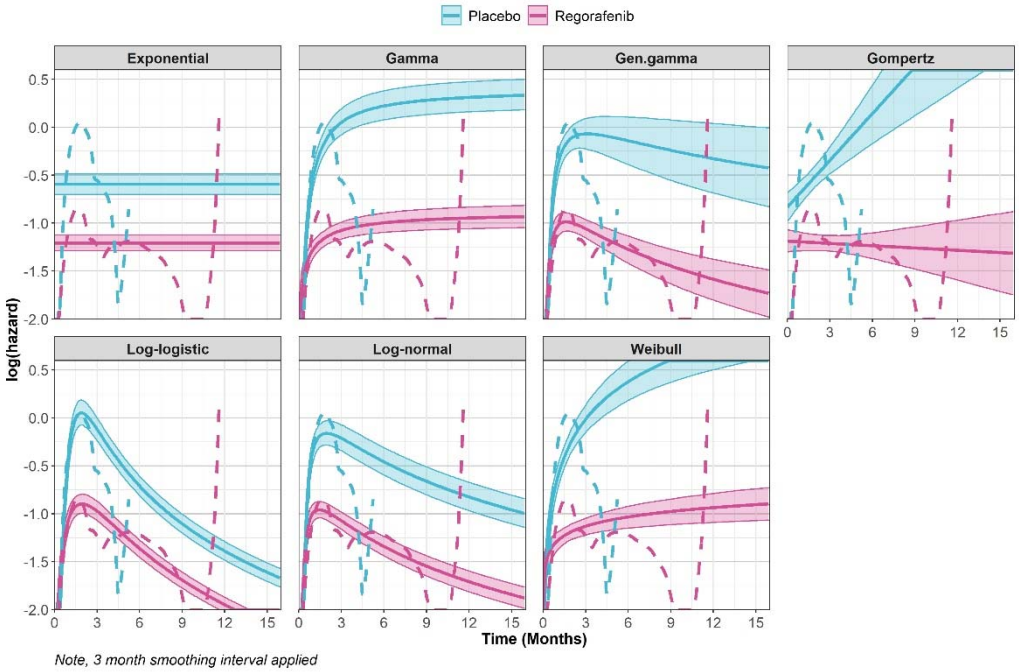
The plots for the pooled CORRECT and CONCUR ToT data are presented in Figure B6.17 to Figure B6.20. Similar conclusions can be made for ToT as were made for PFS (plots suggest that the best fitting models are likely to be the log-logistic, log-normal or generalized gamma models).

Figure B6.17: Cumulative hazard versus time – pooled CORRECT and CONCUR ToT data



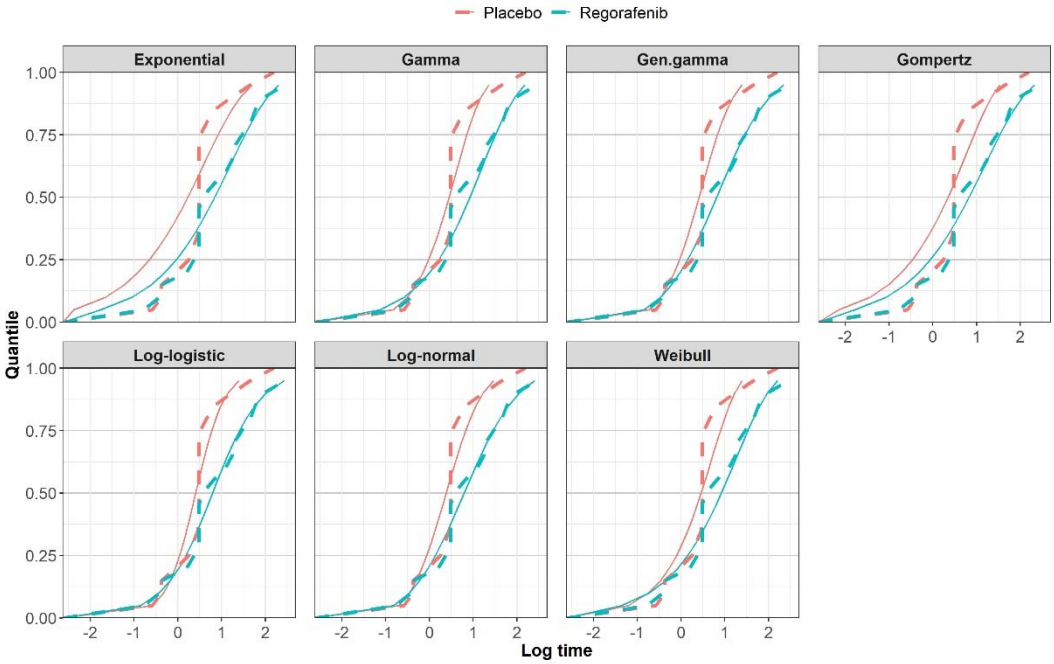
Key: Gen, generalized; LCH, log-cumulative hazard; ToT, time on treatment.

Figure B6.18: Log smoothed hazard versus time – pooled CORRECT and CONCUR ToT data



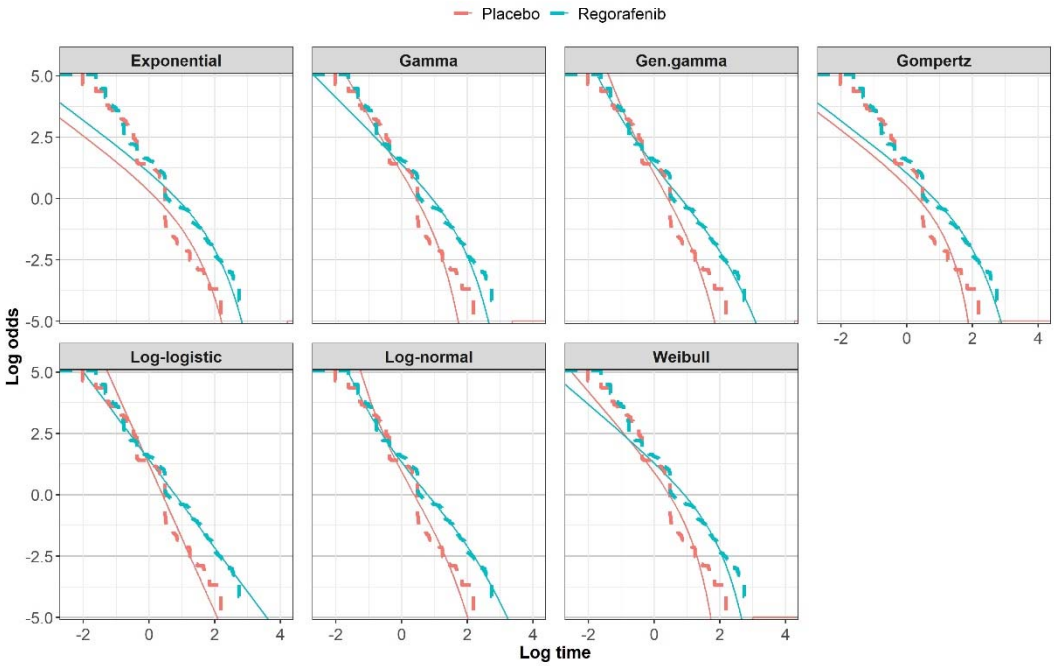
Key: Gen, generalized; ToT, time on treatment.

Figure B6.19: Standard normal quartiles versus log time – pooled CORRECT and CONCUR ToT data



Key: Gen, generalized; ToT, time on treatment.

Figure B6.20: Log survival odds versus log time – pooled CORRECT and CONCUR ToT data



Key: Gen, generalized; ToT, time on treatment.

e) To examine the validity of the extrapolation beyond the data, please provide supporting evidence that the extrapolations are consistent with relevant external data and/or expert opinion. In case of expert opinion, please provide a full description of the methods and results of the expert consultation conducted.

In the submission we used the published primary efficacy data from the trials for regorafenib and trifluridine/tipiracil. This data is the least biased data as it was from the double-blinded period and was not confounded by differences in open-label post-progression treatment or crossover.

In appendix F a post-hoc analysis of survival was presented for both CORRECT (B.5.1.2 – page 79) and CONCUR (B.5.2.2 – page 93). These analyses were based on long-term safety follow-up data. In the same section the post-progression treatment received following discontinuation of study treatment is also presented.

In respect of OS

For OS, we have used the long-term safety follow-up KM data as an extra validation of the parametric curves from the CS. These data are very mature – for pooled OS data there were 94.7% and 94.4% of patients with events for regorafenib and BSC respectively). The safety follow-up KM data is largely consistent with the log-logistic OS modelling (submitted basecase). However, the generalized gamma function may provide a visually closer fit to this more mature data (see Figure B6.21 and Figure B6.22). Based on these data, the log-logistic curve slightly overestimates survival compared to the generalised gamma function. Implementing the generalised function in the model has a minor impact on cost-effectiveness, the NMB vs trifluridine/tipiracil decreases slightly from [redacted] (log-logistic, base case) to [redacted] (generalized gamma OS).

It should be noted that the long-term safety follow-up data included crossover and post-progression anticancer treatment.

Figure B6.21: Overlay of the submitted OS extrapolations and long-term safety follow-up KM data for regorafenib

Figure B6.22: Overlay of the submitted OS extrapolations and long-term safety follow-up KM data for BSC

For PFS and ToT, statistical fit was considered the best justification, as PFS and ToT were both very mature, leaving only a small tail for which there was not data to guide the extrapolation. Therefore, we did not explore any external data to validate these output.

f) Please justify the selection of the approaches to estimate and extrapolate OS, PFS, and TTD, taking into account the responses to the preceding questions as well as the “Survival Model Selection Process Algorithm” provided in NICE DSU TSD 14.

As discussed in question B.5a, in our base-case, KM data for PFS and ToT were applied directly in the model. As this data was nearly complete, we felt that it was preferable to utilise the observed trial data directly, rather replace it with fully modelled curves. In addition, using the KM data directly preserves the stepwise nature of PFS and ToT, which matches how progression is assessed in clinical practice and subsequent treatment decisions. We consider direct implementation of KM data for both of these variables is the best approach.

For OS, the updated safety follow-up data from CORRECT and CONCUR were used to validate our OS input in the model. The generalized gamma OS curve showed a close fit to these data. However, this may be a conservative estimate, as the long-term OS data were captured after unblinding, so there was some cross-over from BSC to regorafenib during the long-term follow and differences in other post-progression anti-cancer treatment. Based on these data, and the additional data provided above, we expect the NMB to be between the log-logistic and generalised gamma estimates i.e. between [redacted] (basecase log-logistic) and [redacted] (Generalised gamma).

g) As suggested in NICE DSU TSD 14, please provide “substantial justification” if different types of parametric models are used for different treatment arms.

Different parametric models were not used for different arms.

B 7. Priority question. Table 24 indicates that no treatment waning was applied in the model.

a) Please justify not implementing treatment waning in this model.

The pooled CORRECT and CONCUR survival data used in the submission provide a robust and mature evidence base, with 54.5% and 61.6% patients having an death event, and KM data available until 9.59% and 7.19% of patients were at risk for BSC and regorafenib respectively. Although there is some inherent uncertainty regarding the remaining 7.19%-9.59% of patients for whom no KM data are available, we did not consider it appropriate to model treatment waning for this small subset of patients. This is also supported by past colorectal cancer submissions (e.g. TA668, TA405), in which no waning was applied either, while using comparably or less mature OS data.

b) To assess the need for treatment waning, please provide a plot of hazard ratio (HR) versus time for both the KM data (using smoothed hazards) and all semi-parametric or parametric functions.

The longer-term safety follow-up survival data from CONCUR and CORRECT was used to validate the original long term OS extrapolations used in our submission in question B.6e. This longer-term data is very mature and captures the treatment effect over the long-term thus removing the need for hypothetical waning analyses.

c) Please conduct an analysis including a hypothetical effect of treatment waning in the model, which considers clinical plausibility and the results of the HR plots.

The longer-term safety follow-up survival data from CONCUR and CORRECT was used to validate the original long term OS extrapolations used in our submission in question B.6e. This longer-term data is very mature and captures the treatment effect over the long-term thus removing the need for hypothetical waning analyses. The longer-term KM data is consistent with our log-logistic basecase in the CS, although the generalized gamma OS function visually provides an improved fit to the data. Versus trifluridine/tipiracil, the basecase NMB was [redacted] (log-logistic) and [redacted] using the generalized gamma function.

It should be noted that the longer-term data includes open-label post-progression anti-cancer treatment and some crossover from placebo to regorafenib.

B 8. Priority Question: As per NICE DSU TSD 14, exponential, Weibull, Gompertz, log-logistic, log normal and Generalised Gamma parametric models should all be considered when performing survival analysis modelling. Please assess the suitability of said distributions by providing the following information.

a) Please rank the parametric survival models for OS, PFS, ToT according to their visual fit for regorafenib, T/T, BSC using joint parametric models for regorafenib and T/T.

No models were fitted to data for trifluridine/tipiracil. This comparator was implemented via an indirect comparison, and application of the estimated HRs for PFS, ToT, and OS to the regorafenib baseline curves. As a result, these questions will only focus on the curve fits for regorafenib and BSC.

The curve overlays of the joint parametric models for the survival data used in the CS, are provided below for regorafenib and BSC. As visual fit is a subjective matter, it is not feasible to rank all parametric models based on visual fit. However, overall,

the visual fit of these models is in line with the statistical fit (AIC/BIC) data provided in B8c, with log-logistic, log-normal and generalized gamma showing the best visual fit for all endpoints.

Figure B8.1: Joint parametric models fitted to Regorafenib OS (short-term and long-term)

Figure B8.2: Joint parametric models fitted to BSC OS (short-term and long-term)

Figure B8.3: Joint parametric models fitted to Regorafenib PFS

Figure B8.4 Joint parametric models fitted to BSC PFS

Figure B8.5: Joint parametric models fitted to Regorafenib ToT

Figure B8.6: Joint parametric models fitted to BSC ToT

b) Please rank the parametric survival models for OS, PFS, ToT according to their visual fit for regorafenib, T/T,BSC using individual parametric models.

No models were fitted to data for trifluridine/tipiracil. This comparator was implemented via an indirect comparison, and application of the estimated HRs for

PFS and OS to the regorafenib baseline curves. As a result, these questions will only focus on the curve fits for regorafenib and BSC.

The curve overlays of the individual parametric models for the original OS, PFS, and ToT data on which the submission was based, are provided below for regorafenib and BSC. Clinical input was sought on these extrapolations, and clinical experts confirmed that all these curves resulted in plausible survival predictions, based on their experience with metastatic colorectal cancer. Because they considered all curves clinically plausible, they recommended to align our approach with TA405, in which a log-logistic was used for OS and PFS, and generalized gamma for ToT.

As visual fit is a subjective matter, it is not feasible to rank all parametric models based on visual fit. However, overall, the visual fit of these models is in line with the statistical fit (AIC/BIC) data provided in B8c, with log-logistic, log-normal and generalized gamma showing the best visual fit for all endpoints.

In addition, the visual fit was generally better for the individual parametric models, than for the joint models presented above. Especially for regorafenib OS, the best fitting joint models show a poor visual fit to the tail end of the KM curve, compared to the individual models. In addition, individual models are also more appropriate when you have access to patient level data, as they rely on fewer assumptions than joint models (in line with NICE DSU TSD 14).

Figure B8.7: Individual parametric models fitted to Regorafenib OS (short-term and long-term)

Figure B8.8: Individual parametric models fitted to BSC OS (short-term and long-term)

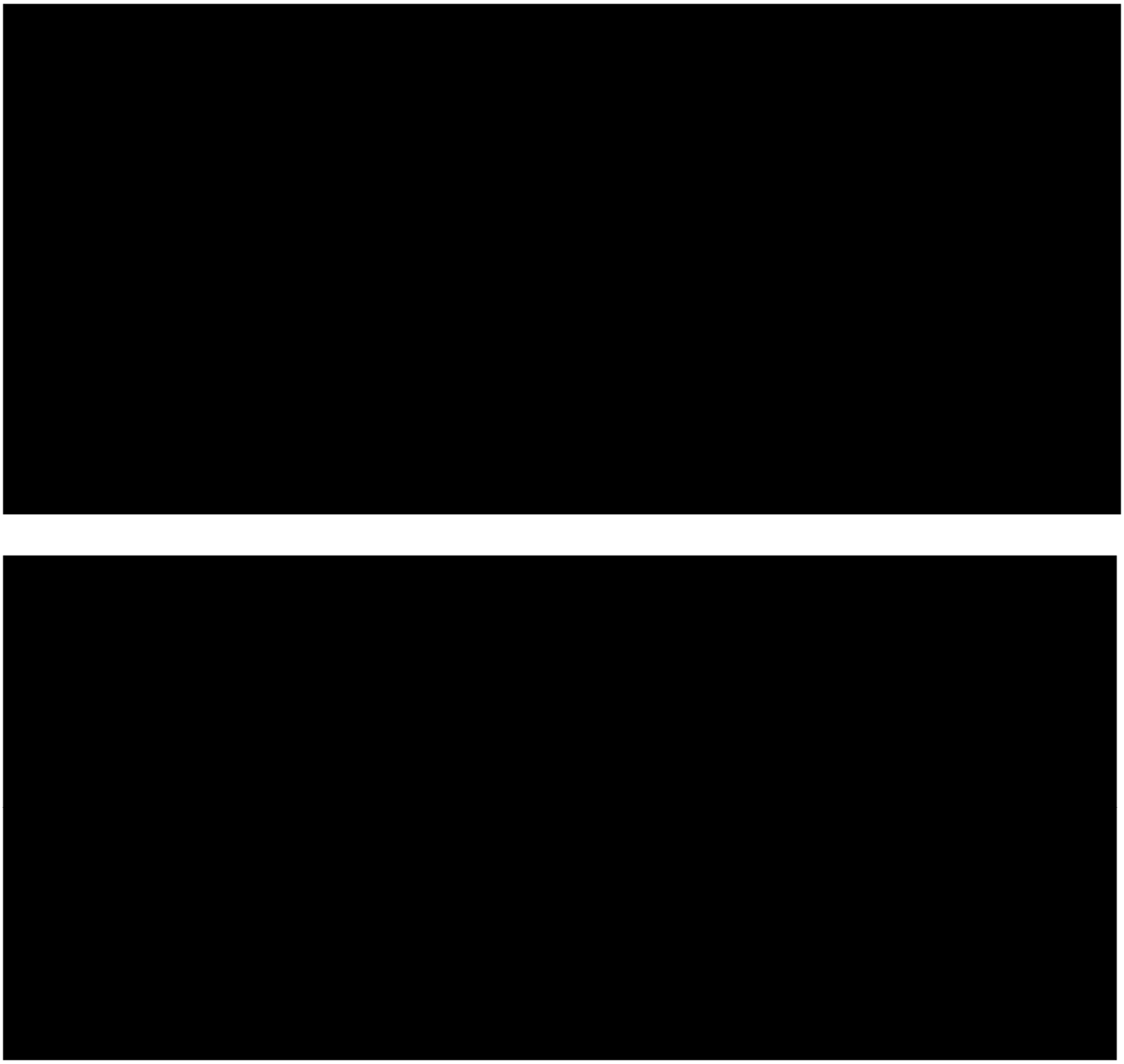


Figure B8.9: Individual parametric models fitted to Regorafenib PFS

Figure B8.10: Individual parametric models fitted to BSC PFS

Figure B8.11: Individual parametric models fitted to Regorafenib ToT

Figure B8.12: Individual parametric models fitted to BSC ToT

c) For the model with joint parametric models enabled, please fill in the following table with the Akaike's Information Criterion (AIC) and the Bayesian Information Criterion (BIC) for each of the distributions.

Table B8.1: Goodness of fit statistics (AIC/BIC) for the joint OS extrapolations

Fitted function (OS)	Regorafenib & BSC (joint model)		Statistical rank
	AIC	BIC	
Weibull	3727.51	3742.13	5
Log-norm.	3699.07	3713.68	2
Log-logistic.	3693.61	3708.22	1
Exponential.	3817.18	3826.92	7
Generalised. Gamma.	3696.73	3716.22	3
Gompertz	3784.93	3799.54	6
Gamma	3710.26	3724.87	4

Key: AIC, Akaike information criterion; BIC, Bayesian information criterion T/T, trifluridine/tipiracil OS, Overall survival

Table B8.2: Goodness of fit statistics (AIC/BIC) for the joint PFS extrapolations

Fitted function (PFS)	Regorafeni & BSC (joint model)		Statistical rank
	AIC	BIC	
Weibull	3351.19	3365.80	5
Log-norm.	3198.28	3212.90	2
Log-logistic.	3162.20	3176.82	1
Exponential.	3552.35	3562.10	7
Generalised. Gamma.	3200.21	3219.69	3
Gompertz	3524.54	3539.15	6
Gamma	3272.39	3287.00	4

Key: AIC, Akaike information criterion; BIC, Bayesian information criterion T/T, trifluridine/tipiracil PFS, progression-free survival

Table B8.3: Goodness of fit statistics (AIC/BIC) for the joint ToT extrapolations

Fitted function (ToT)	Regorafeni & BSC (joint model)		Statistical rank
	AIC	BIC	
Weibull	3477.72	3492.31	5
Log-norm.	3403.43	3418.03	3
Log-logistic.	3366.58	3381.17	1
Exponential.	3550.44	3560.17	6
Generalised. Gamma.	3397.11	3416.57	2
Gompertz	3550.78	3565.38	7
Gamma	3441.38	3455.98	4

Key: AIC, Akaike information criterion; BIC, Bayesian information criterion T/T, trifluridine/tipiracil ToT, Time on treatment

d) For the model with individual curves enabled, please fill in the following table with the Akaike's Information Criterion (AIC) and the Bayesian Information Criterion (BIC) for each of the distributions.

Table B8.4: Goodness of fit statistics (AIC/BIC) for the individual OS extrapolations

Fitted function (OS)	Regorafenib (Individual models)		Statistical rank	BSC (Individual models)		Statistical rank
	AIC	BIC		AIC	BIC	
Log-logistic	2,419.7	2,428.6	1	1,274.6	1,282.1	3
Generalized gamma	2,424.4	2,437.8	2	1,270.6	1,281.9	2
Log-normal	2,428.0	2,437.0	3	1,268.9	1,276.5	1
Gamma	2,428.7	2,437.7	4	1,282.7	1,290.3	4
Weibull	2,437.4	2,446.3	5	1,292.0	1,299.5	5
Gompertz	2,470.0	2,478.9	6	1,316.9	1,324.4	6
Exponential	2,489.2	2,493.7	7	1,328.0	1,331.7	7

Key: AIC, Akaike information criterion; BIC, Bayesian information criterion; BSC, best supportive care; T/T, trifluridine/tipiracil OS, Overall survival

Table B8.5: Goodness of fit statistics (AIC/BIC) for the individual PFS extrapolations

Fitted function (PFS)	Regorafenib (Individual models)		Statistical rank	BSC (Individual models)		Statistical rank
	AIC	BIC		AIC	BIC	
Weibull	2452.934	2461.86	5	871.5151	879.0704	5
Log-normal	2348.506	2357.432	2	803.2672	810.8225	2
Log-logistic	2345.376	2354.302	1	751.179	758.7343	1
Exponential	2529.705	2534.168	7	1022.65	1026.428	7
Generalized gamma	2348.828	2362.217	3	802.6254	813.9584	3
Gompertz	2521.584	2530.51	6	980.991	988.5463	6
Gamma	2412.393	2421.319	4	816.293	823.8483	4

Key: AIC, Akaike information criterion; BIC, Bayesian information criterion; BSC, best supportive care; T/T, trifluridine/tipiracil OS, Overall survival

Table B8.6: Goodness of fit statistics (AIC/BIC) for the individual ToT extrapolations

Fitted function (ToT)	Regorafenib (Individual models)		Statistical rank
	AIC	BIC	
Weibull	2540.677	2549.588	5
Log-normal	2483.435	2492.346	2
Log-logistic	2477.802	2486.713	1
Exponential	2554.352	2558.807	6
Generalized gamma	2483.577	2496.943	3
Gompertz	2556.112	2565.023	7
Gamma	2525.736	2534.646	4

Key: AIC, Akaike information criterion; BIC, Bayesian information criterion; ToT, time on treatment

e) For joint parametric models, to assess the clinical plausibility of the selected distributions, please fill in the following table:

Regorafenib	Weibull	Log-normal	Log-logistic	Exponential	Generalised gamma	Gompertz	Trial data	Expert opinion
OS median	7.36	6.90	6.90	7.36	7.13	7.59	6.90	N/A (Joint models were not discussed at the advisory board)
OS 6 months	59.99%	56.52%	57.14%	56.89%	57.39%	59.94%	55.42%	
OS 1 year	25.36%	28.28%	26.62%	32.37%	27.14%	27.06%	30.46%	
OS 5 years	0.00%	1.10%	1.74%	0.36%	0.36%	0.00%	0.00%	
PFS median	2.99	2.76	2.53	2.53	2.53	2.76	2.07	Experts preferred using the direct KM data over parametric models
PFS 6 months	15.57%	12.09%	9.94%	19.53%	12.11%	18.84%	15.15%	
PFS 1 year	0.54%	1.51%	1.78%	3.82%	1.56%	1.30%	4.67%	
PFS 5 years	0.00%	0.00%	0.03%	0.00%	0.00%	0.00%	0.00%	
ToT median	2.53	2.07	2.07	2.30	2.30	2.30	1.61	Experts preferred using the direct KM data over parametric models
ToT 6months	14.60%	12.72%	11.34%	16.75%	12.48%	16.60%	14.85%	
ToT 1 year	0.99%	2.69%	2.97%	2.81%	1.99%	2.22%	4.68%	
ToT 5 years	0.00%	0.01%	0.11%	0.00%	0.00%	0.00%	0.11%	

BSC	Weibull	Log-normal	Log-logistic	Exponential	Generalised gamma	Gompertz	Trial data	Expert opinion
OS median	5.75	5.52	5.29	5.52	5.52	5.75	5.29	N/A (Joint models were not discussed at the advisory board)
OS 6 months	47.92%	46.15%	44.81%	46.50%	46.04%	48.44%	45.58%	
OS 1 year	13.88%	20.17%	18.10%	21.62%	18.02%	15.71%	18.82%	
OS 5 years	0.00%	0.54%	1.07%	0.05%	0.12%	0.00%	0.00%	
PFS median	1.61	1.61	1.61	1.38	1.61	1.38	1.84	Experts preferred using the direct KM data over parametric models
PFS 6 months	0.75%	3.31%	3.49%	4.66%	3.42%	2.94%	2.08%	
PFS 1 year	0.00%	0.23%	0.59%	0.22%	0.25%	0.01%	0.22%	
PFS 5 years	0.00%	0.00%	0.01%	0.00%	0.00%	0.00%	0.00%	
ToT median	1.38	1.38	1.61	1.15	1.38	1.38	1.61	Experts preferred using the direct KM data over parametric models
ToT 6months	1.38%	5.53%	5.42%	3.67%	4.05%	3.30%	3.07%	
ToT 1 year	0.00%	0.86%	1.35%	0.13%	0.36%	0.07%	1.35%	
ToT 5 years	0.00%	0.00%	0.05%	0.00%	0.00%	0.00%	0.05%	

f) For individual parametric models, to assess the clinical plausibility of the selected distributions, please fill in the following table:

Regorafenib	Weibull	Log-normal	Log-logistic	Exponential	Gen. gamma	Gompertz	Trial data	Expert opinion
OS median	7.36	7.13	7.13	7.36	7.13	7.59	6.90	Experts considered all curves plausible and preferred log-logistic to align with TA405
OS 6 months	59.86%	56.91%	57.22%	56.89%	57.82%	59.79%	55.42%	
OS 1 year	25.58%	29.75%	27.45%	32.37%	27.64%	27.27%	30.46%	
OS 5 years	0.00%	1.50%	1.98%	0.36%	0.28%	0.00%	0.00%	
PFS median	2.99	2.76	2.53	2.53	2.53	2.76	2.07	Experts preferred using the direct KM data over parametric models, so no further input was sought
PFS 6 months	16.28%	14.82%	13.11%	19.53%	15.11%	18.98%	15.15%	
PFS 1 year	0.95%	2.61%	3.04%	3.82%	3.09%	1.99%	4.67%	
PFS 5 years	0.00%	0.00%	0.08%	0.00%	0.01%	0.00%	0.00%	
ToT median	2.53	2.30	2.30	2.30	2.30	2.30	1.84	Experts preferred using the direct KM data over parametric models, so no further input was sought
ToT 6months	15.45%	15.50%	14.40%	16.75%	15.13%	16.83%	14.85%	
ToT 1 year	1.63%	4.25%	4.63%	2.81%	3.64%	3.08%	4.68%	
ToT 5 years	0.00%	0.04%	0.27%	0.00%	0.01%	0.00%	0.27%	

BSC	Weibull	Log-normal	Log-logistic	Exponential	Generalised gamma	Gompertz	Trial data	Expert opinion
OS median	5.75	5.29	5.29	5.52	5.29	5.75	5.29	Experts considered all curves plausible and preferred log-logistic to align with TA405
OS 6 months	48.06%	44.79%	44.17%	46.50%	44.56%	48.64%	45.58%	
OS 1 year	13.50%	17.46%	16.71%	21.62%	18.27%	15.24%	18.82%	
OS 5 years	0.00%	0.25%	0.82%	0.05%	0.48%	0.00%	0.00%	
PFS median	1.84	1.61	1.61	1.38	1.61	1.61	1.84	Experts preferred using the direct KM data over parametric models, so no further input was sought
PFS 6 months	0.20%	0.81%	0.75%	4.66%	0.53%	0.93%	2.08%	
PFS 1 year	0.00%	0.01%	0.05%	0.22%	0.00%	0.00%	0.22%	
PFS 5 years	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	
ToT median	1.61	1.38	1.61	1.15	1.61	1.38	1.61	Experts preferred using the direct KM data over parametric models, so no further input was sought
ToT 6months	0.38%	1.63%	1.57%	3.67%	0.87%	1.23%	3.07%	
ToT 1 year	0.00%	0.07%	0.20%	0.13%	0.01%	0.00%	0.20%	
ToT 5 years	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	

g) Please justify the choice of the selected distributions based on the criteria described above.

As discussed in B.6, in our base-case KM data for PFS and ToT were applied directly in the model, as this data was nearly complete and we felt that it was preferable to utilise the observed trial data directly rather replace it with fully modelled curves. In addition, using the KM data directly preserves the stepwise nature of PFS and ToT, which matches how progression is assessed in clinical practice. We believe this was the correct decision and consider direct application of KM data to be the most reliable and clinically plausible PFS and ToT input.

For OS, the longer-term safety follow-up data from CORRECT and CONCUR were used in question B.6 to validate our OS input in the model. In general, this provides a more complete picture to guide curve selection, than a comparison to trial data at set timepoints, as requested in B.8e and B.8f. Based on these OS data, we expect regorafenib's OS to fall in between the log-logistic and generalized gamma curve, resulting in and NMB range vs trifluridine/tipiracil of [redacted] to [redacted].

B 9. Priority question. For T/T, the company assumed the HR of ToT to be equal to the HR of PFS. For this assumption to hold, patients should discontinue upon disease progression and AE profiles should be similar between the intervention and comparator. However, there are a number of reasons why treatment might be discontinued, such as an AE, which means that it might not coincide with a progression or death and the company makes the point that "*Trifluridine/tipiracil showed a different AE profile*" (p. 128)

a) Please provide additional evidence that the reasons for discontinuation including any stopping rules are similar for regorafenib and T/T.

Clinically there is no difference between the medicines in respect of when treatment is advised to be stopped and they can be considered to be identical in this respect. Treatment for both medicines should continue whilst benefit is observed or until the patient experiences unacceptable toxicity (see below). This supports the use of the PFS HR to model ToT:

- **Regorafenib:** Treatment should continue as long as benefit is observed or until unacceptable toxicity occurs (SmPC)

- **Trifluridine/tipiracil:** Treatment with trifluridine/tipiracil is continued until disease progression or unacceptable toxicity (criteria in trifluridine/tipiracil clinical trials).

b) Given that the AE profiles are different, please justify the assumption that the HR is similar in ToT and PFS.

To explore the impact on the AE profiles of regorafenib and trifluridine/tipiracil on time on treatment, NMAs on Grade 3+ treatment emergent adverse events and discontinuations due to AEs were performed in question A28. The NMAs suggested similar odds of experiencing a Grade 3 or 4 adverse event (OR: 0.90 [95% credible interval: 0.55, 1.47]) or discontinuation due to an adverse event (OR: 1.10 [95% credible interval: 0.53, 2.24]) for regorafenib and trifluridine/tipiracil. These results, along with the similarity in how long patients should take either treatment (part B9a) , supports using the PFS HR to model ToT.

c) Please provide an analysis where ToT incorporates the results of the NMA for discontinuation due to AEs as requested in A28.

The results of the NMA's from question A28 indicate that Grade 3 / 4 AEs (i.e. AEs that have an impact on costs and quality of life) are comparable, as is the odds of discontinuation between the two treatments. With comparability demonstrated for both outcomes no effect on the relative cost-effectiveness would be anticipated.

The results of the NMA requested in A28 are Odds Ratios, as is common for a NMA for a safety outcome. To our knowledge, there is no robust method of using an OR to adjust survival data (HRs are required). Consequently, the results of A28 cannot be used directly to model ToT in the model.

Quality of Life

B 10. Utility values for the post-progression health state were derived from the pooled values for end of treatment utility from the CORRECT and CONCUR trials.

a) Please provide further justification as to the plausibility of using end of treatment utility values as a proxy for post-progression utility, contrary to data from the informing studies.

In the basecase the pre-progression utility value used was 0.72 and 0.59 for post-progression utility. The value of 0.59 for post-progression utility is within the range of published values reported in the CS (document B table 28) which supports its appropriateness.

In TA405 the ERGs preferred utility value for the post-progression health state was 0.59.

At the end-of-treatment visit the majority of patients will be stopping treatment due to progression and consequently the utility value from this visit captures the quality of life of a progressed patient.

b) Please elaborate on the plausibility of the relatively low post-progression utility value (used in the CS base-case) compared to the post-progression utility values identified in the literature.

The post-progression value used in the basecase is within the range of other post-progression values from the literature (see document B table 28) and there is no reason to suspect its plausibility. The use of EQ5D values from patients in CORRECT and CONCUR, and weighted according to the UK tariff, fits the reference case.

In TA405, the ERG preferred the use of utility values from CORRECT (CONCUR was not available at the time) – the values used for post-progression was 0.59 which matches what has been used in this appraisal.

c) Please provide an updated model and scenario analyses informing the post-progression utility using utilities derived from other relevant TAs and provide a justification of how these compare to the end of treatment utility currently used.

We have not provided these analyses as the utility values that we have used are appropriate and align closely to what was used in TA405. However, we consider that the results presented in response to question B12 are relevant as they show that versus trifluridine/tipiracil, the model is not sensitive to an analysis where the difference in pre and post progression utility is significantly widened.

B 11. Slight differences exist in the utility values captured in the CORRECT and CONCUR studies. Utility values used in the economic model were derived from pooling the EQ-5D-3L index scores reported in CORRECT and CONCUR. Please provide an updated model and scenario analyses whereby utility values for each study were used separately.

The post-progression value used in the basecase are within the range of other post-progression values from the literature (see document B table 28). The use of EQ5D values from patients in CORRECT and CONCUR, and weighted according to the UK tariff, fits the reference case.

We have not provided analyses using CORRECT and CONCUR individually as the base case effectiveness pools the efficacy from both studies and therefore it is most appropriate to pool utility values also. However, similar to B10, we consider that the results presented in response to question B12 are relevant as they show that versus trifluridine/tipiracil, the model is not sensitive to an analysis where the difference in pre/post progression utility is significantly widened.

B 12. Table 28 of the CS highlights utility values identified in a systematic literature review for patients receiving ≥ 3rd line treatment for mCRC. These values were not used in the economic model. Please provide an updated model and scenario analyses using utilities for identified studies with the smallest and largest differences in progression-free and progressed disease utility values.

The largest pre-progression utility value from table 28 is 0.810 (Graham 2016 for panitumumab). The lowest post-progression value from table 28 is 0.5 (Bland 2011).

All other values and inputs are the same as the submitted basecase. The NMB using these largest (pre-progression) and smallest values (post-progression) is [redacted] trifluridine/tipiracil ([redacted] higher than in our basecase) which shows that pre and post-progression values are not a driver of cost-effectiveness. This is to be expected given that trifluridine/tipiracil and regorafenib have comparable efficacy.

B 13. The CS highlights uncertainty as to whether any disutility from the experienced AEs would be present at the time of EQ-5D-3L data collections.

a) Please indicate at what points in time EQ-5D-3L were collected?

In both CORRECT and CONCUR, EQ-5D-3L was administered at baseline (Day 1 of Cycle 1), Day 1 of Cycles 2-4, then Day 1 of every other cycle (Cycles 6, 8, etc). Finally, it was administered at the end of treatment visit.

b) Please provide explanation as to how missing data was handled.

No imputation methods were used to account for any missing utility data.

B 14. Adverse event disutilities were sourced from past TAs and combined with the pooled weekly AE probabilities.

a) Please provide justification for the choice of sources and assumptions made in Table 30 of the CS. Please justify whether these disutilities are representative of the AEs for this disease.

The AE disutilities were not subject to a systematic literature review, rather they were selected from past appraisals if it was considered at face value that they were reasonable. We have no reason to suspect they are not representative of AEs for mCRC.

A scenario was presented in the submission where AE disutility was removed entirely (scenario 19 – document B table 47). In this scenario there was practically no impact of removing AE disutility on the NMB versus trifluridine/tipiracil: changing NMB marginally from [redacted] to [redacted] – indicating it is not a driver of the model.

We believe we have been conservative by including AE disutility separately in the economic model – this is because the impact of AEs on a patient's quality of life will be, to some extent, already 'captured' in the patients EQ5D responses i.e. if the

patient is experiencing an AE at the time of the EQ5D response then its impact on QOL will be included. In this regard, the explicit addition of AE disutilities may have an element of double-counting.

b) Please provide an updated model and scenario analyses using AE disutilities from more recent TAs that are relevant to the target population.

We have not been able to review TAs as has been requested. However, the NMA for grade 3+ AEs (see question A28 – table A8.3) showed trifluridine/tipiracil and regorafenib to be comparable in terms of adverse events which have an impact on costs and QOL (grade 3+). As both treatments are comparable then changes to AE disutilities would not be expected to affect cost-effectiveness estimates.

c) Please justify why this approach to incorporating AE disutilities was taken and why disutilities were not applied as a one-off disutility.

We acknowledge that it may be argued that AEs typically occur at the initiation of treatment, and that, as a result, it may be more appropriate to model a one-off disutility rather than a continuous one. However, as demonstrated in our response to B14a, above, excluding AE disutilities entirely has a marginal impact on the model results. Given this, moving from continuous to a one-off disutility would similarly have a marginal impact.

Adverse Events

B 15. Priority question. Please update all cost effectiveness analyses with the results of any NMAs as requested in A28.

NMAs were run in A28 for TEAEs, Grade 3 / 4 adverse events and for discontinuations. However, none of these were suitable for including in the cost effectiveness model, as discussed below.

<u>Discontinuations</u>

The results of the NMA on discontinuation due to AEs were Odds Ratios, as is common for a NMA for a safety outcome. To our knowledge, there is no robust method of using an OR to adjust survival data (HRs are required). Consequently, the results of A28 cannot be used directly to model discontinuation (ToT) in the model.

However, the results of the NMA indicate that regorafenib and trifluridine/tipiracil are comparable in respect of discontinuations.

TEAEs

A cost-effectiveness analysis using the NMA results for TEAEs has not been conducted. This is because TEAEs includes grade 1 (mild) and grade 2 (moderate) events which are not expected to have an impact on costs or quality of life.

Grade 3 /4 adverse events

In order to run this analysis an assumption would be needed that the grade 3 / 4 adverse events that are observed for regorafenib, are common to trifluridine/tipiracil (and vice versa). It would also need to be assumed that the only difference was in the proportion of patients experiencing the events. Neither of these assumptions are supported by table A28.1. We therefore consider that as the necessary assumptions are not supported that any results of such an analysis would be unreliable. In addition, the odds ratio generated in this analysis is not suitable for adjusting survival data.

If an analysis could be run, the direction of effect on the cost-effectiveness result can reasonably be predicted to be in favour of regorafenib as a consequence of the OR favouring regorafenib. However, in respect of the NMA the credible interval crosses 1 indicating no significant difference between the treatments in respect of the occurrence of grade 3 / 4 events i.e. 0.90 (CI 0.55, 1.47). The results of the NMA support that regorafenib and trifluridine/tipiracil are tolerated to a comparable extent – however, table 28.1 indicates that the profile of adverse events is different. We therefore consider that modelling the adverse events based on table A28.1 (as in the CS basecase) was appropriate.

B 16. The model only included Grade 3 or above events that occurred in at least 2% of the population. Please provide a scenario analysis with the inclusion of lower grade AEs.

In our economic model we included AEs that were Grade 3 or higher. These are the adverse events that are sufficiently severe to have an impact on costs and quality of life.

Grade 1 (mild) and grade 2 (moderate) adverse events are not expected to have an impact on costs or quality of life and are generally not modelled in oncology.

We have not been able to conduct the analysis requested in the time available. However, we do not expect any meaningful impact on cost-effectiveness. Sensitivity analyses on AEs, presented in the CS, show that adverse events are not a driver of cost-effectiveness.

B 17. Table 29 of the CS contains missing data. Please provide detail on the reason(s) for this missing data.

The empty cells in Table 29 represent AEs that were not observed for that respective treatment. They were included as 0% in the model

Costs and resource use

B 18. Priority question. As per the CS, two different approaches were used in the company's base-case to represent relative dose intensity (RDI) for regorafenib and T/T. This was partially justified due to insufficient data in the T/T trials and a possible difference in clinical practice: while the regorafenib dose is expected to be reduced when toxicities develop, T/T is expected to be delayed, according to the CS. However, in the CORRECT and CONCUR trial, the regorafenib dose was lower due to both dose reductions and cycle delays in response to AEs.

a) Please provide further clinical justification supported by external data and expert opinion for not implementing the same approach on regorafenib and T/T. In based of expert opinion, please provide a full description of the methods and results of the expert consultation conducted.

In respect of both regorafenib and trifluridine/tipiracil, their respective SmPCs recommend dose reductions and/or delays in order to manage specific adverse events. Therefore the approach to managing adverse events is similar. The main difference is that although both treatments use reductions and delays, regorafenib

tends more towards reductions and trifluridine/tipiracil towards delays. The different implementation of RDI was therefore driven by data availability, as detailed below.

For regorafenib, a single all-encompassing RDI value was available from the CSRs where dose reductions and delays where combined in the single metric of 'RDI'. This is presented in table B18.1 below. This single RDI input was considered the most reliable and straight forward input to model dose intensity, as it reflects the actual dose received for the whole CORRECT and CONCUR populations.

In respect of trifluridine/tipiracil, a single RDI metric was not available for any of the clinical trials. We therefore used the reported average cycle delay and calculated the average proportion of the full dose received, based on data from TA405, to model trifluridine/tipiracil RDI (see table B18.1). We consider that the implementation in respect of trifluridine/tipiracil best reflected the data available.

Table B18.1: Dose intensity information used in the model (table 33 from CS, Document B)

Study	Relative dose intensity (N)	Dose reduction	Cycle delay
Regorafenib – CORRECT	[redacted] (500)		
Regorafenib – CONCUR	[redacted] (136)		
Regorafenib – CORRECT and CONCUR pooled	[redacted]		
Trifluridine/tipiracil – TA405		97.4%	2.72 days

Key: NICE, National Institute for Health and Care Excellence.
Source: CORRECT CSR[45]; CONCUR CSR[48]; Mayer et al., 2015[51]; NICE TA405.[9]

b) Given that regorafenib is supposed to have a better AE profile than T/T, one might expect that T/T has more missed and delay doses due to AEs, and overall lower RDI. Please provide further clinical justification supported by external data or clinical opinion for why the CS reported the opposite effect.

As discussed in clarification question A28, regorafenib has a different (not better) AE profile compared to trifluridine/tipiracil and similar incidences of Grade ≥ 3 AEs. During the advisory board, clinicians agreed that although regorafenib and trifluridine/tipiracil have different safety profiles, overall they are both well-tolerated and similar incidences of Grade ≥ 3 AEs are expected in UK clinical practice. This was also confirmed by the NMA for experiencing a Grade 3 or 4 adverse event performed in A28, which suggested similar odds for regorafenib and trifluridine/tipiracil (OR: 0.90 [95% credible interval: 0.55, 1.47]).

The best indication of tolerability is the incidence of grade 3 / 4 adverse events, which we have already shown to be comparable. A comparison of RDI is not a suitable metric by which to compare tolerability – rather RDI is a manifestation of dose reductions/interruptions that are recommended for different adverse events. For example, if drug A and drug B have an identical incidence of a grade 3 adverse event then they can be considered identical in this aspect. The fact that drug A is recommended to have a higher dose reduction than drug B in relation to that event does not mean that drug A is less tolerable.

c) The company did a scenario analysis applying RDI to the number of pills dispensed, rather than directly to the regorafenib costs, however it is argued that this approach is likely to overestimate regorafenib costs. Given that "*the dose of regorafenib is generally reduced if toxicities develop*", please provide clinical justification on why applying RDI to the number of pills would overestimate the costs.

On average, patients finish a treatment cycle with tablets remaining (due to reductions/interruptions). In the scenario we conducted the patient fills a prescription for a full pack for the next cycle anyway. This continues until enough remaining tablets have been accumulated from prior cycles (i.e. a full pack) such that the next cycle can be completed without the need for a new prescription. In this scenario it is only at the point of accumulating a full pack that a saving is realised.

We consider the scenario to overestimate costs, as in practice, tablets accumulated in one cycle reduce the tablets required and prescribed in the next cycle i.e. a saving is realised on a prescription by prescription basis. Our scenario only allows a saving to be realised at the point a full pack has been accumulated from prior cycles, and as not all patients will stay on treatment for enough cycles to accumulate a full pack only a fraction of costs are saved – therefore this is can be considered a conservative scenario.

d) The company argued that the application of T/T dose reductions during the first dose and continued for the full course of treatment was a conservative approach, given that the "*dose would decline gradually*". This seems to contradict the assumption that T/T would be delayed due to AEs. Please provide further clinical justification supported by external data or expert opinion on this assumption. Furthermore, please elaborate on how this approach had a significant impact on the NMB when explored in a scenario analysis.

As discussed in B18a, adverse events for trifluridine/tipiracil are managed through a combination of dose reductions and dose interruptions/cycle delays. Although a similar approach is followed for regorafenib, there are relatively more dose reductions for regorafenib and relatively more cycle delays for trifluridine/tipiracil. We therefore modelled trifluridine/tipiracil RDI indirectly by applying a combination of dose reductions and cycle delays. Cycle delays had a larger impact on trifluridine/tipiracil costs relative to dose reductions.

However, as dose reductions alone only reduces the trifluridine/tipiracil costs to 97.4%, any over(under)-estimation in dose reduction is not likely to have a significant impact on the NMB. This is also confirmed by the OWSA, in which varying the dose reduction for trifluridine/tipiracil up or down only resulted in a small spread in the [redacted] i.e. there was not a significant impact on the NMB.

B 19. T/T dosage is based on body surface area (BSA). According to the CS, the model used data from Sacco et al., on BSA data for adult UK cancer. However, this study presents the BSA for different types of cancer (e.g., adjuvant and palliative breast cancer, adjuvant and palliative colon cancer, head and neck cancer) and a combined BSA for all groups. Please clarify which data was used in the model (i.e., combined results or only colon cancer results). In case of having implemented the combined results, please provide a scenario analysis and updated economic model using the BSA based on colon cancer patients only.

The model applies BSA data derived solely from colon cancer patients receiving palliative chemotherapy (see Table 2 of Sacco et al.).

B 20. Healthcare resource use (HRU) costs were informed by the four studies identified in the SLR. However, HRU rates used in the model base-case were only derived from TA405 as it "represents a good middle ground".

a) Please provide further clinical justification supported by external data or expert opinion for not implementing the rates from Bullement (2018), Hoyle (2013), TA668, nor a combination of all of them. In case of expert opinion, please provide a full description of the methods and results of the expert consultation conducted.

The ERG preferred HRU from TA405 were simply presented to clinical experts. The values from Bullement, Hoyle and TA668 were not discussed. The experts confirmed that the ERGs preferred HRU from TA405 were appropriate and applicable to both trifluridine/tipiracil and regorafenib. On the basis of being preferred by the ERG in TA405 and being validated by clinical experts they were used in the basecase.

b) Please provide more detailed minutes of the validation of the chosen HRU during the clinical advisory board.

A slide with the ERGs preferred HRUs from TA405 was presented followed by a group discussion. As stated in the CS (document B page 137) the "experts broadly agreed with using the ERG-preferred rates from TA405, with just one expert questioning the post-progression GP surgery visit. However, since this was only mentioned once, we decided to continue with the values as reported in TA405."

c) Please provide further clinical justification for assuming the same HRU rates for regorafenib and T/T based on external data or expert opinion. In case of expert opinion, please provide a full description of the methods and results of the expert consultation conducted.

As mentioned above, the ERGs preferred HRUs from TA405 was presented at the advisory board and this was followed by a group discussion regarding their appropriateness in the mCRC setting.

Using the values that were preferred in TA405 is appropriate as these are estimates related to trifluridine/tipiracil which is the comparator in this submission. Clinical experts did not expect HRU to differ between regorafenib and trifluridine/tipiracil.

d) CS Table 34 included the rates of medical oncologists OP visits. However, this category is missing in CS Table 35. Please update Table 35 to include medical oncologists OP visits and provide scenario analyses including its HRU rate as described in Bullement (2018), Hoyle (2013), and TA668.

Medical oncologists visits is missing from table 35 as they were not applicable i.e. medical oncologist visits were not part of the ERG preferences in TA405. The requested analyses have been run using a medical oncologist visit cost of £193.33 (NHS reference cost 2019-20 (inflated): service code 370, medical oncology, outpatient attendance).

<u>Medical oncologist OP visit as per Bullement</u>

Bullement et al. assume medical oncologist visits are solely required for patients receiving BSC alone, and not trifluridine/tipiracil or regorafenib. As a result, including

their preferred medical oncologist visit rate does not impact the NMB of regorafenib versus trifluridine/tipiracil).

Medical oncologist OP visit as per Hoyle 2013

The Hoyle et al. evaluation assumed patients who received active therapy would experience two medical oncologist visits per month. Applying this rate to regorafenib and trifluridine/tipiracil results in an estimated NMB (versus trifluridine/tipiracil) of [redacted], compared to a base case of [redacted]. This indicates the model is not sensitive to the inclusion/exclusion of this cost.

Medical oncologist OP visit as per TA668

In TA558, it was assumed patients receiving encorafenib in combination with cetuximab would receive one medical oncologist visit every 2 months. Applying this rate to regorafenib and trifluridine/tipiracil results in an estimated NMB (versus trifluridine/tipiracil) of [redacted], compared to a base case of [redacted]. This indicates the model is not sensitive to the inclusion/exclusion of this cost.

B 21. **Priority question. As per the CS, BSC in the CORRECT and CONCUR trials included several concomitant medications and treatments (e.g., antibiotics, radiation therapy for pain control, transfusions, palliative surgery). However, for the CS base-case treatment costs for BSC were assumed to be £0, as they would have been captured by the BSC HRU costs. Nonetheless, according to CS Table 35, the only BSC HRU assumption in the model was that 25% of patients would receive a monthly health home visitor.**

a) Please provide information on all the concomitant medications, treatments, and procedures that BSC entailed for the CORRECT, CONCUR, RECOURSE, TERRA, and Yoshino (2012) trials. In addition, please justify if the use of said BSC reflects UK clinical practice for the population of interest.

For RECOURSE, TERRA, and Yoshino (2012) data on exact concomitant medication use were not publicly available. Concomitant medicines received in the CORRECT and CONCUR trials are provided in table A7.2 and Table A7.3 – please note, as described in our response to A7, not all of these medicines fall under the umbrella of BSC but also include medicines to manage comorbidities. We don't have details on what exactly is prescribed as BSC in the UK but have no clinical reason to suspect it would differ meaningfully from was received in CORRECT and CONCUR.

As can be seen from A7.2 and Table A7.3 all concomitant medicines are captured – it is not possible to discriminate between treatments for comorbidities or palliative care nor is it possible to do a full costing of these medicines. BSC includes medicines for pain relief, proton pump inhibitors, benzodiazepines etc and these are inexpensive. In the absence of being able to do a costing exercise for BSC medicines in the UK we have performed a pragmatic scenario analyses with an assumed BSC-medicines cost of £50 per 28-day treatment cycle. Please note that as medicines for BSC are typically inexpensive we consider this to be on the 'high' side. Explicitly including these costs, in addition to the HRU already in the model, has a negligible effect on cost-effectiveness i.e. NMB decreases by [redacted] (from a

basecase of [redacted] Overall these results confirm that modelling BSC costs only have a minimal impact on the model outcomes.

b) Please elaborate on how the BSC costs for both regorafenib and T/T can be captured using the assumptions and rates defined in CS Table 35 (i.e., 25% of patients receiving a monthly health home visitor).

As described in CS Section B.3.5.2 and Table 35, the HRU rates for trifluridine/tipiracil and regorafenib were assumed equal, and informed by the ERG preferred HRU rates in TA405. In addition to the 25% of patients receiving a monthly health home visitor, the HRU also consist of a monthly oral chemotherapy outpatient visit incurred by 100% of patients and CT scan incurred by 33% of patients. Considering the low cost of the concomitant medication and its negligible impact on cost-effectiveness (question B.21a) we consider that the approach followed in the CS is appropriate and proportionate and aligns with what was done in TA405.

c) Please provide an updated economic model including costs for BSC based on all medications, treatment and procedures defined in each trial for BSC and their weekly rate.

We have not been able to do a full costing exercise as requested but hope that our response above assures the EAG that changes to BSC are not impactful in terms of cost-effectiveness.

B 22. According to the CS, progression-free patients attend an oral chemotherapy outpatient appointment, in which they undergo routine tests and see a clinician to review their treatment, apart from receiving the treatment for the upcoming cycle. Please provide further detail on the routine test included in this appointment and include its respective costs.

'Routine' tests include blood tests to check renal and liver function as well as markers such as C-Reactive Protein. Patients may also have full blood counts. These tests are conducted according to the needs of the individual. These tests are inexpensive and have not been costed individually. In order to answer this question

we would need to seek clinical advice regarding the exact tests performed and the proportion of patients who receive each test as it not be correct to assume all tests are carried out for all patients. We have not been able to gather the information from clinicians in order to address this question.

However, as these tests are inexpensive and are anticipated to be broadly the same for patients on trifluridine/tipiracil and regorafenib, and these two treatments are of comparable efficacy, no impact on cost-effectiveness is expected.

B 23. Table 35 of the CS summarises the HRU assumptions and costs used in the company's base-case for the progression-free and progressed health states.

a) Please elaborate and provide clinical justification based on external data and/or expert opinion on the following base-case assumptions: i) BSC patients do not attend any routine oncologist visits; ii) 25% of all patients incurred the costs of a health home visitor per treatment cycle, regardless of their treatment; iii) after progression, patients are no longer assumed to attend day case or outpatient consultations (i.e., only care closer to home).

We are not able to provide justification over and above the HRUs being preferred by the ERG in TA405, and experts having advised us that in current clinical practice they are applicable to both regorafenib and trifluridine/tipiracil.

b) . Per resource item, please provide details on the source of the percentages of patients using the resource item in CS Table 35 (i.e. literature or expert input).

The percentages are from the ERGs preferences in TA405

c) As per Table 34 and 35 from CS, please define the abbreviations of "Lon", "Bev" and "All pts".

Lon = Lonsurf i.e. the brand name for trifluridine/tipiracil

All pts = all patients

The inclusion of "Bev" (i.e. bevacizumab) is an error in these tables. These resource use estimates are the Hoyle et al. values for cetuximab and panitumumab, not bevacizumab. We apologise for this error.

B 24. Priority question. Post-progression treatment was given to a substantial number of patients in the CORRECT (Regorafenib 26%, BSC 30%) and CONCUR (Regorafenib 31%, BSC 43%) trials. However, the modelled company's base-case does not include any subsequent treatment based on clinical expert opinion.

a) Please elaborate on whether patients in the T/T trials (i.e. RECOURSE, TERRA, and Yoshino (2012)) received post-progression subsequent treatments. If so, please provide details of the type of subsequent treatments and the proportions of patients that received it.

<u>RECOURSE</u>

Information was provided in the committee papers in respect of TA405. In TA405 post-progression treatment costs were estimated to be £1,528 and applied as a lump sum on progression. It is stated in the submission that 42% of patients in RECOURSE went on to receive non-study anti-tumour treatments – table B24.1 provides further details. In TA405, further information on the treatments received and costing analyses were provided in Appendix 4 and 11 of the CS, however these are not available to Bayer.

Table B24.1: Non-study anti-tumour treatments received after the end of the treatment period in RECOURSE (ITT population).

Table 30: Non-study anti-tumour treatments received after the end of the treatment period in RECOURSE (ITT population)

Treatment	Number (%) of patients		
	Trifluridine/tipiracil (N = 534)	Placebo (N = 266)	Total (N = 800)
Surgery	6 (1.1)[a]	5 (1.9)	11 (1.4)
Surgery or systemic anti-cancer therapy	224 (41.9)	118 (44.4)	342 (42.8)
Radiotherapy	0	0	0
Any systemic therapy	222 (41.6)	113 (42.5)	335 (41.9)
Number of regimens			
1	170 (31.8)	88 (33.1)	258 (32.3)
2	41 (7.7)	22 (8.3)	63 (7.9)
≥3	11 (2.1)	3 (1.1)	14 (1.8)
Any regorafenib containing regimen	84 (15.7)	41 (15.4)	125 (15.6)
No regorafenib containing regimens	138 (25.8)	72 (27.1)	210 (26.3)

Key: ITT, intention-to-treat.
Notes: [a] Includes four patients who had surgery plus other systemic anti-cancer therapy, and two patients who had surgery only.
Source: RECOURSE CSR.[12]

Yoshino

Based on the publication by Yoshino post-progression treatment was available to patients i.e.

"subsequent treatments that could be potential confounders of an overall survival endpoint, such as cytotoxic and molecular targeting agents, were given to similar or greater proportions of patients in the placebo group than in the TAS-102 group"

Bayer does not have further details.

TERRA

According to the publication by Xu 2017 "in the trifluridine/tipiracil and placebo arms, after discontinuation of trifluridine/tipiracil or placebo, a small proportion of patients

continued treatment with biologic therapy with or without chemotherapy". Details were provided – see table B24.2. A greater proportion of patients in the placebo arm (42.5%) versus the trifluridine/tipiracil arm (37.6%) received at least one anticancer therapy after discontinuation criteria were met. However, the subsequent anti-cancer treatment duration could not be evaluated.

Table B24.2: post-study treatment

Table 3. Post-Study Treatment

Parameter	No. of Patients (%)	
	Trifluridine/Tipiracil (n = 271)	Placebo (n = 135)
Biologic therapy*	38 (14)	22 (16)
Anti-VEGF but not anti-EGFR	27 (10)	16 (12)
Anti-EGFR but not anti-VEGF	10 (4)	5 (4)
Both anti-VEGF and anti-EGFR	1 (< 1)	1 (1)
Chemotherapy	39 (14)	27 (20)
Investigational drug	39 (14)	22 (16)

Abbreviations: EGFR, epidermal growth factor receptor; VEGF, vascular endothelial growth factor.
*Anti-VEGF therapy includes axitinib, bevacizumab, regorafenib, and sorafenib; anti-EGFR therapy includes only cetuximab; biologic therapy might include a combination of chemotherapeutic agents; chemotherapy includes only chemotherapeutic agents.

b) Please provide further clinical justification on whether subsequent treatments and their costs should be incorporated in the economic model.

Our submission included zero costs for post-progression treatment (base case) and also provided a scenario including post-progression costs taken from TA405 (inflated to 2021 i.e. £1,633,18). These were applied equally to regorafenib and trifluridine/tipiracil. The basecase [redacted] for the scenario analysis – demonstrating that the inclusion (non-inclusion) of post-progression treatment costs is unimportant in respect of cost-effectiveness. Based on these results, varying the costs between £0 and £1,633.18 to account for different proportions of patients receiving post-progression treatment in the UK clinical setting would have no meaningful impact.

c) Given the substantial percentage of patients receiving subsequent treatment post-progression in the CORRECT and CONCUR trials, please provide a scenario analysis and updated economic model including the costs of subsequent treatments in both comparators and intervention as described in the trials.

There is a lack of data on which to base an analysis that could be considered 'superior' to what has been provided in the CS. In terms of post-progression treatment there is no reliable way to determine if it would differ between trifluridine/tipiracil and regorafenib in clinical practice. In respect of the trials the information available is not sufficient to provide the analysis requested as the proportion and types of post-progression treatment differed by treatment arm and across trials. We also do not have data on the durations for which treatments were received.

Our base case NMB and scenario analysis (presented in the CS and in part b above) demonstrate that post-progression treatment costs are not a driver of cost-effectiveness. In the absence of being able to provide the exact analyses requested, we have conducted two additional scenario's where an arbitrary difference in costs of 10% between the regorafenib and trifluridine/tipiracil is implemented. However, in the absence of data to suggest post-progression treatment would differ between regorafenib and trifluridine/tipiracil we consider this to be academic rather than particularly informative in the context of this appraisal.

<u>Scenario 1 – post-progression cost: regorafenib £1,633,18; trifluridine/tipiracil £1,796.50 (£1,633,18 x 110%)</u>

In this scenario the NMB increases from [redacted]

<u>Scenario 2 – post-progression cost: regorafenib £1,796.50; trifluridine/tipiracil £1,633.18</u>

In this scenario the NMB decrease from [redacted]

The above scenarios indicate that inclusion of differential post-progression treatment costs has a small impact on NMB. However, to be holistic, the scenarios would also need to account for any impact of post-progression treatment on OS as well AEs. For the reasons outlined in A9c and B24a this is not possible.

d) Please incorporate the results of any adjustments to OS and PFS for the removal of subsequent treatment as requested in A9.

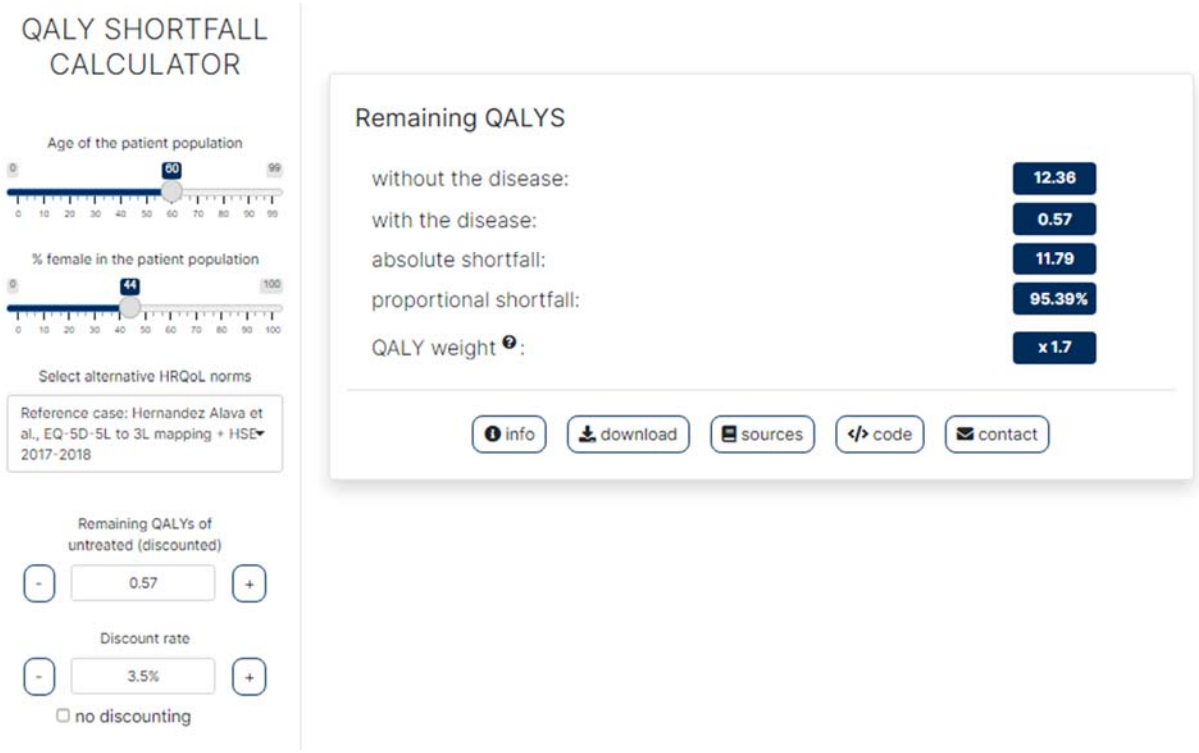
Please see our answer to A9c. After considering the data available we do not believe it is possible to isolate and quantify the effect of post-progression treatment - the trials were not designed for this purpose. We expect that post-progression treatment will have some potential benefit but that this is not quantifiable.

Severity

B 25. The ERG reproduced the shortfall analysis reported in CS section B.3.6. The reported proportional QALY shortfall (CS Table 40) and the 1.7x QALY weight were successfully reproduced.

a) The healthy population estimate of QALYs remaining calculated by the ERG was 12.39 while 12.36 was reported in the CS. Please confirm that the company's estimated is indeed 12.36 and provide instructions to reproduce this estimate.

We have replicated our shortfall estimation and produced the same result: an absolute shortfall of 12.36 QALYs. The parameters utilised to generate this estimate are provided below. We note that using a 50:50 male/female population split would produce an absolute shortfall estimate of 12.39, and wonder if this may be the difference between the ERG and our estimates.



b) The estimated absolute QALY shortfall was not provided in CS section B.3.6. Please provide the estimated absolute QALY shortfall.

We estimated an absolute shortfall of 11.79 QALYs.

Results

B 26. Considering the CS base-case results.

a) Please provide a comparison of the observed OS as well as PFS (e.g. using restricted mean survival time; RMST) and the undiscounted life years (LYs) as well as undiscounted progression free LYs (estimated in the model) by filling out the table below using different periods/truncation points (with justification) to calculate the RMST.

We apologise but we have not been able to provide these results due to time constraints as we focused on the analyses required for priority questions. However, we hope that the analyses using more mature data presented in response to B7 is informative.

	Observed	Modelled	
	Restricted mean survival time (RMST)	Estimated (lifetime time horizon)	Proportion beyond observed data
OS - RMST period / truncation point: XX months			
Regorafenib			
Comparator			
Increment			
OS - RMST period / truncation point: XX months			
Regorafenib			
Comparator			
Increment			
OS - RMST period / truncation point: XX months			
Regorafenib			
Comparator			
Increment			
PFS - RMST period / truncation point: XX months			
Regorafenib			
Comparator			
Increment			
PFS - RMST period / truncation point: XX months			
Regorafenib			
Comparator			
Increment			
PFS - RMST period / truncation point: XX months			
Regorafenib			
Comparator			
Increment			

b) Please elaborate on the plausibility of the differences between observed and modelled outcomes (proportion accumulated beyond observed data) for:

i. Regorafenib

ii. the comparator

iii. the increment

We apologize but we have not been able to provide these results.

c) Regarding the model estimated differences between the intervention and the comparator (in terms of PFS, LYs and quality-adjusted life years (QALYs)); please provide an explanation of the mechanism by which the model generated these differences as well as a justification for why they are plausible based upon available evidence (NICE DSU TSD 19 recommendation 13).

We apologise but we have not been able to provide these results.

Scenario- and Sensitivity-analysis

B 27. A number of scenario analyses are conducted for the T/T comparison which are not conducted for the BSC comparison. Please conduct all scenario analyses which are conducted for the T/T comparison also for the BSC comparison.

BSC is not a comparator - we are seeking a recommendation for regorafenib as a treatment option alongside trifluridine/tipiracil (i.e. the position for regorafenib in clinical guidelines). This submission was made in response to physician's requests for an alternative to trifluridine/tipiracil. Therefore BSC isn't a comparator.

As BSC is not a comparator we have not conducted any further analyses against BSC.

B 28. Please conduct all scenario analyses with the fully parametric curves selected for PFS and ToT.

In our response to B5h we have provided cost-effectiveness results using different curves for PFS and ToT. As the tables show the choice of curve has very little impact on cost-effectiveness.

We have not conducted scenarios for each of these curves due to time constraints. However, in respect of PFS and ToT we consider that implementation of the KM data directly provides the most robust estimates and that the requested analyses are not warranted given the lack of impact curve selection has.

Validation and transparency

B 29. The results of the validity assessments and detailed validation exercises are not described (i.e. specific black-box tests). In CS section B.3.14 it is described that the model has "been subjected to a checklist of known modelling errors, and the assumptions have been questioned".

a) Please provide a detailed description of the validity assessment performed (including the checklist mentioned in the CS) as well as the results.

The model was validated using the Lumanity checklist: a quality control procedure developed using publicly available checklists from Drummond and Philips. The Lumanity checklist includes all checks listed in the published TechVER checklist.

The QC begins with basic validity checks, for example:

•Setting all costs to zero (the total cost should be equal to zero)
•Increasing drug costs (the total costs should increase)
•Checking that results match (i.e. total costs in one table are reflected in total costs in another table)
•Repeating the above tests for utilities and/or risks of clinical outcomes
•Setting all costs and outcomes to be the same for all treatment arms and checking that each treatment arm generates the same results
•Changing the time horizon of the model (if variable) and checking that outputs/results change accordingly

Subsequently the 'Costs' section of the checklist involves checks to identify any problems with references used for costs and any issues with how costs are applied. Example checks include:

•Are all costs taken from the latest available publication?
•Are all costs presented using the same price year (i.e. are inflation indices correctly applied to costs for older data sources) and is this year clearly stated in the model?

The ‘Utilities’ section of the checklist ensures that utilities modelled reflect the population concerned and are correctly applied in the model. Example checks include:

•Are the utilities in the model higher than that of patients in the general population of the same age?
•Are utilities for adverse events included in the model? If not, is a justification provided as to why it is appropriate not to include them?

Clinical input quality checks assess the validity and accuracy of the clinical outcome data used in the model to calculate outcome differences between treatment arms. Example checks include:

•Are the efficacy inputs applied correctly based on the description provided in the model?
•Is mortality applied correctly based on the description provided in the model?
•If the model is not supposed to apply a lifetime horizon, is justification provided for the time horizon specified?

Checks concerning the model settings are then conducted to ensure all switches and settings are working correctly. Example checks include:

•Switches/settings are placed appropriately, and are easy to find and understand
•Switches/settings relate to the correct set of inputs, and changes in these settings are reflected in the model results

Sensitivity analysis checks are then carried out to ensure these are working correctly and that results are meaningful. Example checks include:

•Checking that all parameters used in sensitivity analysis have a plausible associated distribution and maximum/minimum values (e.g. probabilities are between 0 and 1)
•Running simulations to ensure that the output has no erroneous results
•If automated two-way sensitivity analysis or scenario analyses are included, running

these using the code and manually to check the results are correct and plausible

This section checks the background Visual Basic for Application (VBA) coding used in the models, including macros and user forms. Example checks include:

•Checking macros function as described
•Checking that superfluous or unused code (e.g. from the ‘record’ function) or user forms have been removed
•Checking that all code is appropriately commented
•Thoroughly testing any user forms in the model

The QC was led by an experienced health economist who was not been involved in the development of the original model. In addition to the checks outlined above, the health economist carrying out the QC also reviewed the model sheet by sheet to identify any additional errors or inconsistencies that may not have already been identified. On completion of the QC, the model was updated to correct any identified problems.

b) Please provide complete the TECH-VER checklist (Büyükkaramikli et al. 2019, https://pubmed.ncbi.nlm.nih.gov/31705406/) and provide the results.

The Lumanity QC checklist includes all checks in the TECH-VER checklist. Any issues identified via this QC were resolved prior to submission.

B 30. Please provide cross validations, i.e. comparisons with other relevant NICE TAs focussed on similar, potentially relevant, diseases (e.g. related NICE recommendations and NICE Pathways listed in the final scope, including those mentioned in CS Table 24) and elaborate on the identified differences regarding:

a) Input parameters related to:
    i. Clinical effectiveness
    ii. Resource use and costs

Comparisons with trifluridine/tipiracil (TA405) have been made throughout the CS. Apologies but we have not been able to provide a comparison versus non-comparator treatments in the time available.

b) Estimated (disaggregated) outcomes per comparator/ intervention
    i. Life years
    ii. QALYs
    iii. Costs

Apologies but we have not been able to provide this information in the time available

B 31. Please report on the face validity assessment (mentioned in CS B.3.14.1) of the model structure, model assumptions, model inputs, intermediate outcomes as well as final outcomes in more detail (including what aspects were assessed and what were the considerations as well as conclusions). Further validation of modelled effectiveness would be desirable.

The model has a structure which is common in oncology and is entirely appropriate for comparing trifluridine/tipiracil and regorafenib. The validity of the model structure was confirmed by clinical experts and matches the model structure used in TA405.

B 32. Page 99 of the CS: “most of the identified studies do not have a UK perspective, resulting in differences in treatment costs, healthcare resource use (HRU) costs, and AE costs used in these models, further limiting the accuracy of the cost results.”. Please indicate what the implications of this uncertainty is and how this can potentially be examined.

One of the major limitations of the published cost-effectiveness studies is that none of them included all 5 trials for trifluridine/tipiracil and regorafenib. The best assessment of comparative effectiveness is achieved by utilising each of the 5 trials as we have done in our submission.

In respect of the uncertainties listed in the question we believe there is limited usefulness in informing the value of a treatment in the UK, by examining health care resource used in different healthcare systems.

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Regorafenib for previously treated metastatic colorectal cancer [ID4002]

Patient Organisation Submission

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on conditions and their treatment that is not typically available from other sources.

To help you give your views, please use this questionnaire with our guide for patient submissions.

You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type. [Please note that declarations of interests relevant to this topic are compulsory].

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 10 pages.

NICE National Institute for Health and Care Excellence

About you

1.Your name	[redacted]
2. Name of organisation	Bowel Cancer UK
3. Job title or position	[redacted]
4a. Brief description of the organisation (including who funds it). How many members does it have?	We are the UK's leading bowel cancer charity. We are determined to save lives and improve the quality of life of everyone affected by bowel cancer by championing early diagnosis and access to best treatment and care. We support and fund targeted research, provide expert information and support to patients and their families, educate the public and professionals about the disease and campaign for early diagnosis and access to best treatment and care. The majority of our income is generated from individual, corporate and trust fundraisers. A small proportion (£125,791) was given by pharmaceutical companies between 1 April 2021 – 31 March 2022 in support of healthcare professional and patient awareness days, health information publications and our colonoscopy confidence campaign.
4b. Has the organisation received any funding from the company bringing the treatment to NICE for evaluation or any of the comparator treatment companies in the last 12 months? [Relevant companies are listed in the appraisal stakeholder list.]	No
4c. Do you have any direct or indirect links	No

with, or funding from, the tobacco industry?	
5. How did you gather information about the experiences of patients and carers to include in your submission?	The information we provide as part of this appraisal was gathered from a survey of people with advanced bowel cancer who have been previously treated and those caring for advanced bowel cancer patients. The survey was shared on our community forum, advanced bowel cancer Facebook group, website and social media channels. We received responses from 97 people affected by advanced bowel cancer. Seventy-one responses were from patients who have been diagnosed and have received treatment. Eighteen responses were from carers of those with advanced bowel cancer. The majority of responses are from patients and carers who have broader experience of a range of treatments for their advanced bowel cancer, with four respondents having experience of Regorafenib specifically. We have also included pre-existing information we have gathered from our patient community.

Living with the condition

6. What is it like to live with the condition? What do carers experience when caring for someone with the condition?	A diagnosis of bowel cancer is life changing and can affect almost every aspect of daily life, not only for the individual diagnosed but also for their family and loved ones. This is even more acute for those diagnosed at the later stages of the disease, when we know it is harder to treat and the chance of survival is low. Patients experience numerous difficulties and challenges across the pathway, from initial diagnosis, to treatment, and care. In particular, these relate to the impact and reality of an advanced bowel cancer diagnosis, the difficulty and complexity in navigating treatment and care pathways and the impact treatment can have on quality of life. The impact on the patient's daily life and mental health was a major theme in our responses. The diagnosis was described by many as life changing. While some people talked about trying to stay positive, many described living with advanced bowel cancer as *'a daily struggle'* with a *'constant pressure of anxiety'* and *'fear'*. The stress created by the uncertainty of the condition was described by one patient *'Anticipatory grief is the biggest black cloud to live with. It's like you're on a countdown clock but you can't see it'* and by another *'There isn't a morning when I don't wake up and aren't aware of it. The uncertainty is awful'*. One carer described it as *'Living in limbo, in fear and hoping and praying his treatment works.*' The physical impact of the disease and its treatment was highlighted by patients and carers in our survey. Many talked about being limited by pain and extreme fatigue that impacted their ability to work and care for family members. Side effects of treatments described include fatigue, diarrhoea, constipation, piles, sore skin around bottom, skin tears, bleeding and having a stoma. The majority of patients found the side effects of treatments physically hard to cope with, but some reported few side-effects and continued to work throughout treatment. We know each treatment for bowel cancer has a range of potential side effects including changes in bowel habit and control (including diarrhoea, constipation, wind, urgency, frequency, LARS etc), changes in sexual function and fertility, urinary function, nerve damage, fatigue and emotional wellbeing. We know from our wider patient community that regaining bowel control can be one of the biggest challenges patients face after treatment for bowel cancer. Each chemotherapy drug or combination has its own side effects, which can often be controlled by medicines. Not everyone will have the same side effects, but the most common side effects include diarrhoea, increased risk of infection, a sore mouth and feeling being sick. Chemotherapy can also cause temporary or permanent infertility, depending on the drugs and doses used.

	One patient told us *'I carried a pair of gloves with me at all times when I was on oxaliplatin and wore extra thick slipper socks to keep the tingling/burning sensation at bay during cold weather.'* We have also heard previously from a number of advanced bowel cancer patients who experience painful side effects while going through treatment with cetuximab and panitumumab as first line treatment. Prolonged use of these drugs can cause a number of skin toxicities and side effects including extremely painful red skin rashes and fissures, dry and peeling skin across hands, feet and face, cystic painful acne-like spots, severe paronychia, loss of eye lashes and eye soreness, nausea, diarrhoea and reduced appetite. Patients have also emphasised the psychological impact continued treatment has had. Many patients have described how their side effects have left them feeling debilitated, isolated and self-conscious. Unfortunately, we hear that too often patients do not get access to the right support to alleviate these side effects. The psychological impact of treatment options available to advanced bowel cancer patients was raised with one respondent describing how it made them feel *'More options give hope and mentally the more options available the less panicked and heartbroken you feel. It is a constant tight rope'.* The impact of an advanced bowel cancer diagnosis on a patient's family and loved ones was described by patients and carers. One respondent told us *'It's devastating for everyone involved with that person. It changes your life forever.'* Patients described feelings of guilt and worry about their ability to care for children or elderly family after their diagnosis.

7. What do patients or carers think of current treatments and care available on the NHS?	Survival rates for advanced bowel cancer patients are poor, with less than one in ten people surviving more than five years. These patients deserve access to the best quality treatment and care. For some patients the drugs available can be lifesaving, while for others they can prolong life, resulting in more time to spend with loved ones. Therefore, it is essential patients gain timely access to the treatments that their clinicians feel could benefit them. The majority of these patients may be offered chemotherapy combinations as first and second-line treatments for the disease. For patients with specific cancer cell mutations, including RAS, BRAF and mismatch repair (MMR) genes, patients may have an additional option of more targeted treatments. However, patients whose treatment is unsuccessful with these therapies have very limited options at third-line and beyond. Patients who have been previously treated told us that they feel that treatment options available on the NHS are 'limited' or 'inadequate' for those with advanced bowel cancer. This can be psychologically detrimental, with patients unable to access further treatments that could prolong their life and give them the best possible outcome. This also has financial implications for patients and their families, with many having to resort to fundraising or borrowing money in order to fund treatments privately. For patients and their families, this inequity of access causes unnecessary stress, worry and anxiety when they are already struggling to come to terms with being diagnosed advanced bowel cancer. Limiting access in this way means that patients may miss out on treatments that could extend their life. One patient told us *'Make a wider range of treatment available. It's extremely stressful knowing the treatment currently available will sooner, or later stop working and know that unless you can pay you will die'.*

8. Is there an unmet need for patients with this condition?	Bowel Cancer UK argue there is an unmet need for this specific patient population as third line treatment options are limited. The main treatment used in the third-line and beyond for advanced bowel cancer is the chemotherapy tipiracil/trifluridine. Regorafenib offers a new treatment with a different side effect profile giving patients more treatments options. In our survey, respondents agreed that a wider range of drugs should be made available on the NHS and more open and transparent conversations need to be had with clinicians regarding all the options available, including drugs available privately. Patients and carers told us: *'It may be that the only alternative is a drug not approved by NICE and then possibly financing it myself. Thus could more drugs be made available on the NHS for my treatment?'* *'Allow advanced bowel cancer patients to have personalised care, fund drugs that are proven to be effective and remove unnecessary rules - this will enable oncologists to deliver the best care.'* *'Also opening up the availability of more lines of treatment rather than just the usual three lines. Perhaps offering a wider more informative selection of treatment options initially for patients to decide with their oncologist which would be the better choice.'*

Advantages of the technology

9. What do patients or carers think are the advantages of the technology?	Giving patients an additional treatment option that may prolong their life and allow more time with family is welcomed. The offer of a treatment with a different side effect profile for those who may react badly to tipiracil/trifluridine give patients more options. One advantage of Regorafenib is that it is a tablet that is taken orally and so is convenient to take at home, avoiding regular, multiple trips to a hospital meaning patients don't have to deal with the time and stress of frequent travel. Of the 97 respondents to the survey 4 people had direct experience of being treated with Regorafenib. Their responses in direct quotes are below. *'I have only had one cycle so far, but it has been pretty positive - milder side effects and some evidence it is stabilising my cancer. I have had more energy to take part in normal activities. No disadvantages as yet.'* *'The first 3 months of this treatment allowed me to have a normal life. Tablets at home taken before bed. The worst side effect was an upset tummy but this was easy to deal with. Great drug, no long days in hospital and a stable scan after 3 months. I am now having another 3 months.'* *'Provided another line of defence. Instead of only 3 lines of treatment, some comfort knowing there is another effective drug to try.'* *'I participated in a trial for Regorafenib (that is still ongoing) - part of Leap 17 at UCL. Unfortunately it has not been successful for me and I have recently been removed from the trial. Obviously it didn't work to shrink mets I had, and I have today been diagnosed with a collapsed lung but could not say if this was as a result of the trial.'*

Disadvantages of the technology

10. What do patients or carers think are the disadvantages of the technology?	For advantages and disadvantages please see quotes above in answer to question number 9.

Patient population

11. Are there any groups of patients who might benefit more or less from the technology than others? If so, please describe them and explain why.	No comment.

Equality

12. Are there any potential equality issues that should be taken into account when considering this condition and the technology?	Patients and carers who responded to our survey felt that equal access to these drugs on the NHS was of the upmost importance.

Other issues

13. Are there any other issues that you would like the committee to consider?	A personalised treatment approach provides a significant opportunity to improve outcomes for patients with advanced bowel cancer. By understanding the genetic make-up of a patient's tumour at an early stage, healthcare professionals can better target treatments that are more likely to work. In doing so, we can maximise outcomes for patients and also ensure they do not unnecessarily have to undergo the often gruelling side-effects of treatment. Access to personalised medicines for people with bowel cancer require patients to be tested for a range of genetic biomarkers at diagnosis, so that they are given the most appropriate treatment. Embedding routine genetic testing for bowel cancer patients into the patient pathway is vital to ensure everyone has access to the latest and most effective treatment options. Research into genetic biomarkers and new methods to aid personalised treatments herald a brighter future for bowel cancer treatments. It is vital that new, personalised, approaches to treatment are developed and approved by NICE.

Key messages

24. In up to 5 bullet points, please summarise the key messages of your submission.	• A diagnosis of bowel cancer can be life changing for those diagnosed, as well as their friends and family, and is even more acute for those at later stages of the disease when it is harder to treat and there is a low chance of survival. • We hear from patients and carers that they feel that treatment options approved for use on the NHS for advanced bowel cancer are limited, with many unable to access treatments that they believe could prolong their life. • Bowel Cancer UK argue there is an unmet need for this specific patient population as third line treatment options are very limited. • All patients should have access to personalised, tailored treatment that is right for them in order to improve poor outcomes for people with advanced bowel cancer.

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

Please select YES if you would like to receive information about other NICE topics - YES

For more information about how we process your personal data please see our privacy notice.

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Regorafenib for previously treated metastatic colorectal cancer [ID4002]

Professional organisation submission

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on the technology in the context of current clinical practice that is not typically available from the published literature.

To help you give your views, please use this questionnaire. You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 13 pages.

About you

1. Your name	[redacted]
2. Name of organisation	NCRI-ACP-RCP-RCR
3. Job title or position	[redacted]
4. Are you (please select Yes or No):	An employee or representative of a healthcare professional organisation that represents clinicians? Yes
5a. Brief description of the organisation (including who funds it).	NCRI-ACP-RCP-RCR
5b. Has the organisation received any funding from the manufacturer(s) of the technology and/or comparator products in the last 12 months? [Relevant manufacturers are listed in the appraisal matrix.] **If so, please state the name of manufacturer, amount, and purpose of funding.**	No
5c. Do you have any direct or indirect links with, or funding from, the tobacco industry?	No

The aim of treatment for this condition

6. What is the main aim of treatment? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability.)	The main aims of treatment are to: 1. To prolong overall and progression-free survival by delaying progression of metastatic colorectal cancer (mCRC) in the last (i.e. third/fourth) line setting 2. To prolong the disease-control rate (i.e. proportion of those treated who have responding and stable disease) 3. To improve symptom control 4. To improve quality of life
7. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by x cm, or a reduction in disease activity by a certain amount.)	In this advanced and heavily-pretreated population, the priorities are to improve overall survival and quality of life. Radiological response rates are very low in this setting with regorafenib and also with licensed and NICE approved comparator drugs (i.e. trifluridine/tipracil). The key clinically significant outcome measure is improved disease control rate.
8. In your view, is there an unmet need for patients and healthcare professionals in this condition?	Yes, there is an unmet need as large numbers of UK patients with mCRC each year have disease progression on all earlier lines of therapy, with a maintained performance status and who wish to have further systemic anti-cancer therapy. These patients are usually offered referral for early phase trials, but access to these is very variable and often with very significant geographical constraints across the UK. Also, all UK early phase cancer trials units have waiting lists of these people with refractory mCRC in whom the number of possible treatment slots is very limited.

What is the expected place of the technology in current practice?

9. How is the condition currently treated in the NHS?	Following disease progression or intolerance of earlier-line therapies, the only two options are supportive and palliative care and referral for consideration of early phase trials.

9a. Are any clinical guidelines used in the treatment of the condition, and if so, which?	Yes, in the UK we use the NICE colorectal cancer guidelines (NG151, last updated on 15/12/21). Many oncologists also use the ACP guidelines 2017, ESMO 2014 guidelines (Van Cutsem et al, Annals of Oncology) and the US NCCN guidelines (last updated January 2022) but not all the treatment options in either of these latter two documents are available to UK patients with mCRC. Most cancer centres and units in the UK also have local and/or regional guidelines for these patients.
9b. Is the pathway of care well defined? Does it vary or are there differences of opinion between professionals across the NHS? (Please state if your experience is from outside England.)	Yes, it is well-defined. Earlier lines of treatment are now very complex with use of single agent and combination therapeutics, de-escalation maintenance strategies, treatment breaks and localised metastasis directed therapies such as surgery, ablation and stereotactic radiotherapy. Once these have been exhausted, then the options in the third-line are trifluridine/tipracil, referral for early phase trials or supportive/palliative care. There are no significant differences of professional opinion across the UK on this pathway. Most oncologists nowadays restrict use of trifluridine/tipracil to those with good performance status and with clear evidence of response to earlier lines of therapy.
9c. What impact would the technology have on the current pathway of care?	This would offer an alternative to trifluridine/tipracil in the third-line setting to the patients who are deemed unlikely on the criteria above to have benefit from this agent. It would also offer a new therapy following trifluridine/tipracil in the fourth-line setting for a subgroup of fit and motivated patients.
10. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	It could be used as above as an alternative third-line treatment in selected patients.
10a. How does healthcare resource use differ between the technology and current care?	Like trifluridine/tipracil, it is an oral agent delivered in 4 weekly cycles so has a similar footfall in hospitals for 4 weekly clinical assessment (which nowadays is often performed remotely)
10b. In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.)	In specialist oncology clinics by appropriately trained doctors and non-medical prescribers (i.e. nurses and pharmacists)

10c. What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	There is no significant need for investment as the data on this agent is mature with no relevant ongoing studies in mCRC as monotherapy, and the drug is used routinely in other cancers such as gastrointestinal stromal tumours and hepatocellular cancer. Hence, there is no need for additional facilities, equipment and training.
11. Do you expect the technology to provide clinically meaningful benefits compared with current care?	Yes. A minority of patients will derive clear clinical benefit with acceptable levels of drug-induced toxicities.
11a. Do you expect the technology to increase length of life more than current care?	Yes, since this is in large measure supportive and palliative care. In the pivotal trial for registration, regorafenib provided a 23% reduction in the risk of death in this heavily pre-treated group of patients.
11b. Do you expect the technology to increase health-related quality of life more than current care?	The answer to this is nuanced. The QoL data from the pivotal trial and subsequent published real world experience shows that there is no detriment overall in patients' QoL compared to placebo. However, there are a group of patients who derive no benefit and have increased symptoms and decreased QoL (and who stop the drug quickly). There is a second group who live longer with no deterioration in their quality of life. There is a third group who live longer with improved symptoms and quality of life. The specific HQQoL outcomes for these subgroups are missed when the overall HRQoL data is provided for the whole patient group.
12. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	Currently, we have no predictive markers for those in whom the technology would be more or less effective i.e. those who will benefit in terms of both prolonged survival and improved/maintained QoL.

The use of the technology

13. Will the technology be easier or more difficult to use for patients or healthcare professionals than current care? Are there any practical implications for its use (for example, any concomitant treatments needed, additional clinical requirements, factors affecting patient acceptability or ease of use or additional tests or monitoring needed.)	These patients have no current NHS standard care other than supportive and palliative care. They will require in person or remote 4 weekly assessments prior to prescription of regorafenib, and 4 weekly blood tests as with triflurifdine/tipracil. Patients will have disease response assessment through use of 8-12 weekly CT scans, again as with triflurifdine/tipracil.
14. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	Treatment will start based on standard clinical assessment of patients with mCRC in the late-line setting. Treatment will stop on clinical deterioration or radiological progression, and on unacceptable toxicities despite dose reductions and delays. There have been interesting and useful real world data (REDOS study, Bekaii-Saab et al, Lancet Oncology 2019) on the use of an alternate dosing strategy of 80mg daily for one week, then 120 mg daily for one week then 160 mg daily for one week then a break for a week followed by full dosing for subsequent cycles i.e. 160 mg orally daily for days 1 to 21 repeated every 428 days (with dose reduction and delay as needed as per standard dosing). This alternate dosing strategy provided comparable activity and a lower incidence of adverse events with a recommendation that this could be implemented in clinical practice. Our clinical colleagues tell us that this alternate dosing strategy is commonly used in the UK in other licensed and NICE approved indications for regorafenib.

15. Do you consider that the use of the technology will result in any substantial health-related benefits that are unlikely to be included in the quality-adjusted life year (QALY) calculation?	No.
16. Do you consider the technology to be innovative in its potential to make a significant and substantial impact on health-related benefits and how might it improve the way that current need is met?	This drug approval would allow UK patients with mCRC to access a NICE-approved anti-angiogenic for the first time (with this class of agents being commonly used for these patients across the world).
16a. Is the technology a 'step-change' in the management of the condition?	Yes in that there are no alternative agents in third-line for those in whom trifluridine-tipracil is contra-indicated or thought to have minimal chance of benefit. There are no alternative agents in the fourth-line setting.
16b. Does the use of the technology address any particular unmet need of the patient population?	This is end of life cancer care.
17. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	Toxicities are well-known and management of these is through dose reduction and delay and use of appropriate supportive oral and topical medications.

Sources of evidence

18. Do the clinical trials on the technology reflect current UK clinical practice?	This would offer an alternative to trifluridine/tipracil in third-line to the patients who are deemed unlikely on the criteria above to have benefit from this agent. It would also offer a new therapy following trifluridine/tipracil in the fourth-line setting for a subgroup of fit and motivated patients.
18a. If not, how could the results be extrapolated to the UK setting?	N/A
18b. What, in your view, are the most important outcomes, and were they measured in the trials?	The most important outcomes were overall survival, progression-free survival, disease-control rate, symptom control and quality of life. These were all measured in the pivotal CORRECT trial.
18c. If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes?	These are not used.
18d. Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	No. The new data has been on the reduced toxicity of the alternate REDOS dosing strategy.
19. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	There are papers on real world experience of regorafenib monotherapy in late line mCRC with relatively small numbers that may be missed but no significant evidence that we are aware of.
20. Are you aware of any new evidence for the comparator treatments since the publication of NICE technology	No

appraisal guidance 405 and 307?	
21. How do data on real-world experience compare with the trial data?	These are very similar. The results in Asian-only populations have slightly better outcomes than those with a mixed population and with predominantly/exclusively Caucasian populations (see e.g. CONCUR trial, Li et al, Lancet Oncology 2015).

Equality

22a. Are there any potential equality issues that should be taken into account when considering this treatment?	Yes - there are no alternative agents in the third-line mCRC setting for those in whom trifluridine-tipracil is contra-indicated or thought to have minimal chance of benefit.
22b. Consider whether these issues are different from issues with current care and why.	No different

Topic-specific questions

23 How common is retreatment with first line treatments, such as single-agent irinotecan, FOLFIRI, FOLFOX or raltitrexed?	Retreatment with such agents is often done in late line where there has been a good response or prolonged stable disease from the drug(s) previously used, and there has then been a treatment break with use of alternative agent(s) on disease progression.

Key messages

24. In up to 5 bullet points, please summarise the key messages of your submission.	• Regorafenib can provide increased durations of disease control with prolonged PFS and OS in fit and motivated patients with mCRC after third-line treatment where there is no licensed and NICE approved alternative treatments • Regorafenib can provide increased durations of disease control with prolonged PFS and OS in fit and motivated patients with mCRC as third-line treatment where the licensed and NICE approved alternative treatment of trifluridine-tipracil is contra-indicated or felt to be futile. • Altered dosing strategies and careful dosing adjustment during treatment can significantly reduce toxicities of regorafenib treatment, hence improving patient QoL • Many UK patients with mCRC each year with refractory disease after exposure to all the licensed and NICE approved treatments are still fit and well and seek further therapy. They have limited access to early phase trials and so are offered only supportive and palliative care.

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

Please select YES if you would like to receive information about other NICE topics - YES

For more information about how we process your personal data please see our privacy notice.

in collaboration with:

Maastricht University

Regorafenib for treating metastatic colorectal cancer (ID4002)

Produced by Kleijnen Systematic Reviews (KSR) Ltd, in collaboration with Maastricht University Medical Centre (UMC)

Authors
Nigel Armstrong, Health Economics Manager, KSR Ltd, United Kingdom (UK)
Willem Witlox, Health Economist, Maastricht UMC, the Netherlands
Mark Perry, Reviews Manager, KSR Ltd, UK
Bram Ramaekers, Health Economist, Maastricht UMC, the Netherlands
Evangelos Danopoulos, Systematic Reviewer, KSR Ltd, UK
Bradley Sugden, Health Economist, Maastricht UMC, the Netherlands
Pawel Posadzki, Reviews Manager, KSR Ltd, UK
Andrea Fernandez Coves, Health Economist, Maastricht UMC, the Netherlands
Teebah Abu-Zahra, Health Economist, Maastricht UMC, the Netherlands
Thomas Otten, Health Economist, Maastricht UMC, the Netherlands
Caro Noake, Information Specialist, KSR Ltd, UK
Manuela Joore, Health Economist, Maastricht UMC, the Netherlands
Jos Kleijnen, Founder and Owner, KSR Ltd, UK

Correspondence to
Nigel Armstrong, Kleijnen Systematic Reviews Ltd
Unit 6, Escrick Business Park
Riccall Road, Escrick
York, YO19 6FD
United Kingdom

Date completed 27/07/2022

Source of funding: This report was commissioned by the National Institute for Health Research (NIHR) Evidence Synthesis Programme as project number STA 13/54/72.

Declared competing interests of the authors None.

Acknowledgements

Commercial in Confidence (CiC) data are highlighted in blue throughout the report.

Academic in Confidence (AiC) data are highlighted in yellow throughout the report.

Confidential comparator prices are highlighted in green throughout the report.

Any de-personalised data are highlighted in pink throughout the report.

Rider on responsibility for report

The views expressed in this report are those of the authors and not necessarily those of the NIHR Evidence Synthesis Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:

Armstrong N, Witlox W, Perry M, Ramaekers B, Danopoulos E, Sugden B, Posadzki P, Fernandez Coves A, Abu-Zahra T, Otten T, Noake C, Joore MA, Kleijnen J. Regorafenib for treating metastatic colorectal cancer (ID4002): a Single Technology Assessment. York: Kleijnen Systematic Reviews Ltd, 2022.

Contributions of authors

Nigel Armstrong acted as project lead, systematic reviewer and health economist on this assessment, critiqued the clinical effectiveness methods and evidence and company's economic evaluation and contributed to the writing of the report. Willem Witlox acted as health economic project lead, critiqued the company's economic evaluation and contributed to the writing of the report. Bram Ramaekers, Andrea Fernandez Coves, Bradley Sugden, Teebah Abu-Zahra and Thomas Otten acted as health economists on this assessment, critiqued the company's economic evaluation and contributed to the writing of the report. Mark Perry, Evan Danopoulos and Pawel Posadzki acted as systematic reviewers, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Caro Noake critiqued the search methods in the submission and contributed to the writing of the report. Manuela Joore acted as health economist on this assessment, critiqued the company's economic evaluation, contributed to the writing of the report and provided general guidance. Jos Kleijnen critiqued the company's definition of the decision problem and their description of the underlying health problem and current service provision, contributed to the writing of the report and supervised the project.

Abbreviations

AIC	Akaike information criterion
AE	Adverse event
AEOSI	Adverse events of special interest
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
BIC	Bayesian information criterion
BMI	Body mass index
BNF	British National Formulary
BRAF	Mutation in the B-Raf proto-oncogene
BSC	Best supportive care
CADTH	Canadian Agency for Drugs and Technologies in Health
CE	Cost effectiveness
CEA	Cost effectiveness analysis
CENTRAL	Cochrane Central Register of Controlled Trials
CI	Confidence interval
CR	Complete response
CrI	Credible interval
CRD	Centre for Reviews and Dissemination
CS	Company submission
CSR	Clinical study report
CT	Computerised tomography
CTC	Common Toxicity Criteria
CTCAE	Common Terminology Criteria for Adverse Events
DCR	Disease control rate
DOR	Duration of response
DSA	Deterministic sensitivity analyses
DSU	Decision Support Unit
eMIT	Electronic market information tool
ECOG	Eastern Cooperative Oncology Group
ECOG PS	Eastern Cooperative Oncology Group Performance Status
eGFR	Estimated glomerular filtration rate
EGFR	Epidermal growth factor receptor
EORTC	European Organisation for Research and Treatment of Cancer
EORTC QLQ C30	European Organization for Research and Treatment of Cancer Quality of Life Questionnaire (30 items)
EOT	End of treatment
EQ-5D	European Quality of Life-5 Dimensions
ERG	Evidence Review Group
ESS	Effective sample size
FAS	Full analysis set
FE	Fixing errors
FOIB	Fluoropyrimidine, oxaliplatin, irinotecan, bevacizumab
FV	Fixing violations
G-CSF	Granulocyte-colony stimulating factor
HR	Hazard ratio
HRQoL	Health-related quality of life
H_0	Null hypothesis
H_1	Alternative hypothesis
HSUV	Health state utility values
HTA	Health Technology Assessment
IC	Indirect comparison
ICER	Incremental cost-effectiveness ratio

ICTRP	International Clinical Trials Registry Platform
iNMB	Incremental net monetary benefit
INR	International normalised ratio
ITC	Indirect treatment comparison
ITT	Intention-to-treat
IVRS	Interactive Voice Response System
KM	Kaplan-Meier
KSR	Kleijnen Systematic Reviews Ltd
KRAS	Kirsten rat sarcoma viral oncogene homologue
MAIC	Matching-adjusted indirect comparison
mCRC	Metastatic colorectal cancer
MD	Mean difference
MedDRA	Medical Dictionary for Regulatory Activities
MeSH	Medical subject headings
MJ	Matters of judgement
MSS	Microsatellite status
N	Number of patients
NA	Not applicable
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
NMA	Network meta-analysis
NR	Not reported
NYHA	New York Heart Association
Od	Once a day
ORR	Overall response rate
OS	Overall survival
PAS	Patient Access Scheme
PBO	Placebo
PD	Progressed disease
PF	Progression-free
PFS	Progression-free survival
PLD	Patient level data
PK	Pharmacokinetic
Po	Per os (oral)
PPS	Post-progression survival
PR	Partial response
PRESS	Peer Review of Electronic Search Strategies
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-analyses
PRO	Patient reported outcome
PSA	Probabilistic sensitivity analysis
PSM	Partitioned survival model
PSS	Personal Social Services
PSSRU	Personal Social Services Research Unit
PTT	Partial thromboplastin time
QALY	Quality-adjusted life year
QLQ-C30	Quality of Life Questionnaire
QoL	Quality of life
RCT	Randomised controlled trial
RDI	Relative dose intensity
RECIST	Response Evaluation Criteria in Solid Tumours
RMST	Restricted mean survival time
RR	Relative risk; risk ratio
RWE	Real-world evidence
SA	Sensitivity analysis

SAE	Serious adverse events
SAS	Safety analysis set
SD	Stable disease
SD	Standard deviation
SeTE	Standard error of treatment effect
SLR	Systematic literature review
SmPC	Summary of product characteristics
STA	Single Technology Appraisal
STM	State transition model
T/T	Trifluridine/tipiracil
TA	Technology Assessment
TAS-102	Trifluridine/tipiracil
TEAE	Treatment emergent adverse events
TKI	Tyrosine kinase inhibitor
ToT	Time on treatment
tpCR	Total pathological complete response
TRAEs	Treatment related adverse events
TSD	Technical Support Document
ULN	Upper limit of normal
UK	United Kingdom
UMC	University Medical Centre
US	United States
USA	United States of America
UTD	Unable to determine
VAS	Visual analogue scale
VEGF	Vascular endothelial growth factor
WHO	World Health Organization
WT	Wild type
WTP	Willingness-to-pay

Table of Contents

Table of Tables

Table of Figures

1. EXECUTIVE SUMMARY

This summary provides a brief overview of the key issues identified by the Evidence Review Group (ERG) as being potentially important for decision making. If possible, it also includes the ERG's preferred assumptions and the resulting incremental cost effectiveness ratios (ICERs).

Section 1.1 provides an overview of the key issues. Section 1.2 presents the key model outcomes. Section 1.3 discusses the decision problem, Section 1.4 issues relate to the clinical effectiveness, and Section 1.5 issues relate to the cost effectiveness. A summary is presented in Section 1.6.

Background information on the condition, technology and evidence and information on key as well as non-key issues are in the main ERG report, see Sections 2 (decision problem), 3 (clinical effectiveness) and 4, 5 and 6 (cost effectiveness) for more details.

All issues identified represent the ERG's view, not the opinion of the National Institute for Health and Care Excellence (NICE).

1.1 Overview of the ERG's key issues

Table 1.1: Summary of key issues

ID4002	Summary of issue	Report Sections
1	Potential ambiguity of the population in the decision problem, which has implications for the comparators.	2.1, 2.3, 3.3, 3.4
2	Potential difference between subsequent treatment use in the randomised controlled trials (RCTs) and National Health Service (NHS) clinical practice with unknown effect of subsequent treatment.	2.1, 3.2. 3.4
3	Lack of external validity of and comparability between the regorafenib and trifluridine/tipiracil (T/T) RCTs.	3.2, 3.3, 3.4
4	Difference in the treatment effect of regorafenib versus T/T depending on evidence source.	3.2, 3.3, 3.4, 4.2.6
5	The company implemented Kaplan-Meier (KM) curves instead of parametric survival models for the survival analyses of progression-free survival (PFS) and time on treatment (ToT).	4.2.6
6	Low grade adverse events (AEs) may also be relevant, but the company did not consider these in their economic model.	4.2.7
7	The company assumed different relative dose intensity (RDI) for regorafenib and T/T. However, a large observational study by Nakashima 2020[1], directly comparing regorafenib to T/T, reported comparable dose reductions (54% and 48% respectively).	4.2.9
8	In response to the clarification letter, the company did not comply with three requests: 1) to provide a table presenting information about the observed and modelled overall survival (OS) and PFS using restricted mean survival time (RMST) 2) to provide all scenario analyses for the best supportive care (BSC) comparison that were conducted for T/T comparison, and 3) to provide all scenario analyses with the fully parametric models applied for PFS and ToT.	5.2

ID4002	Summary of issue	Report Sections
9	The results of the probabilistic sensitivity analysis are slightly different when running the same analysis multiple times (without changing model settings), likely due to the lack of a fixed random seed in the model probabilistic sensitivity analysis (PSA).	5.3

AEs = adverse events; BSC = best supportive care; KM = Kaplan-Meier; NHS = National Health Service; OS = overall survival; PFS = progression-free survival; PSA = probabilistic sensitivity analysis; RCTs = randomised controlled trials; RDI = relative dose intensity; RMST = restricted mean survival time; T/T = trifluridine/tipiracil; ToT = time on treatment

1.2 Overview of key model outcomes

NICE technology appraisals compare how much a new technology improves length (overall survival) and quality of life (QoL) in a quality-adjusted life year (QALY). An ICER is the ratio of the extra cost for every QALY gained.

Overall, the model is set to affect QALYs:

- Compared to trifluridine/tipiracil (T/T), regorafenib increased the progression-free life-years by [redacted]. This resulted in a QALY increase of [redacted].
- Compared to best supportive care (BSC), regorafenib increased progression-free and progressed disease life-years by [redacted] years pre-progression and [redacted] years post-progression). This resulted in an overall QALY increase of [redacted].

Overall, the model is set to affect costs:

- Compared to T/T, [redacted] ([redacted]). This cost comprised [redacted] of the incremental costs.
- Compared to BSC, regorafenib mainly increased the treatment costs by [redacted] and increased pre-progression care by [redacted], together comprising [redacted] of the incremental costs.

The company performed and presented the results of probabilistic sensitivity analyses (PSA), deterministic sensitivity analyses (DSA) as well as scenario analyses.

The parameters that had the greatest effect on the ICER based on the company's DSA compared to T/T were:

- The hazard ratio (HR) used for overall survival (OS)
- The HR used for progression-free survival (PFS)
- The relative dose intensity (RDI) of regorafenib

The parameters that had the greatest effect on the ICER based on the company's DSA compared to BSC were:

- The RDI of regorafenib
- The proportion of patients in the regorafenib arm which received daily chemotherapy
- The cost of adverse event (AE) management in the regorafenib arm

The company further conducted several scenario analyses, of which most were only presented for the comparison with T/T. The following were most influential:

- Using only the CORRECT and RECOURSE trials to inform the indirect treatment comparison (ITC) [redacted] the incremental net monetary benefit (iNMB) to [redacted].
- Using only the CONCUR and Yoshino trials to inform the ITC [redacted] the iNMB to [redacted].

1.3 The decision problem: summary of the ERG's key issues

Table 1.2: Key issue 1: Potential ambiguity of the population in the decision problem, which has implications for the comparators.

Report Sections	**2.1, 2.3, 3.3, 3.4**
Description of issue and why the ERG has identified it as important	There does appear to be a match between the scope and at least one of the indications for regorafenib, although the decision problem population is more precisely defined: failed on first line treatment and at least third line is narrower than *"previously treated with, or are not considered candidates for, available therapies"* in that such patients might not have failed first line but could have been intolerant to it and might theoretically include second line. However, it is defined in a second way i.e., in terms of eligibility for trifluridine/tipiracil (T/T), the National Institute for Health and Care Excellence Technology Assessment (NICE TA) 405 recommendation for which also specifies that patients need to have been previously treated with, or are not considered candidates for, available therapies. The company initially stated that best supportive care (BSC) was a 'minor comparator', but in response to clarification stated that this was a mistake. Therefore, although the NICE scope included more comparators than T/T including BSC, the company have argued that T/T is the only comparator because the population is defined according to eligibility for this treatment and that BSC would be at the next line.
What alternative approach has the ERG suggested?	The Evidence Review Group (ERG) considers that if the population is defined essentially according to the comparator T/T then the only patients who should be considered for regorafenib are those who might be considered for T/T and no other treatment. The results of the comparison with BSC from the cost-effectiveness analysis (CEA) are reproduced in the ERG report for completeness.
What is the expected effect on the cost effectiveness estimates?	In the company and ERG base-case, regorafenib would not be cost effective versus BSC with a willingness-to-pay (WTP) of £30,000 per quality adjusted life year (QALY), but it would be if WTP was £51,000 i.e., with the 1.7 x QALY weight applied.
What additional evidence or analyses might help to resolve this key issue?	No additional evidence or analyses are needed unless the NICE recommendation is to include patients who might not get T/T e.g., those not fit enough to receive it and who might only be eligible for BSC.
BSC = best supportive care; CEA = cost effectiveness analysis; ERG = Evidence Review Group; NICE = National Institute for Health and Care Excellence; QALY = quality adjusted life year; TA = Technology Assessment; T/T = trifluridine/tipiracil; WTP = willingness-to-pay	

Table 1.3: Key issue 2: Potential difference between subsequent treatment use in the RCTs and NHS clinical practice with unknown effect of subsequent treatment

Report Sections	**2.1, 2.3, 3.3, 3.4**
Description of issue and why the ERG has identified it as important	The company suggested that best supportive care (BSC) would follow either trifluridine/tipiracil (T/T) or regorafenib, but it also stated that <10% would be fit enough to receive subsequent active treatment in clinical practice and many patients in the

Report Sections	2.1, 2.3, 3.3, 3.4
	CORRECT and CONCUR trials received systematic anticancer therapy.
What alternative approach has the ERG suggested?	The Evidence Review Group (ERG) requested an estimate of the effect of subsequent therapy in CORRECT and CONCUR and adjustment to better reflect National Health Service (NHS) clinical practice, but the company refused to do this and stated that it would favour BSC given that more patients in the BSC only arm received it.
What is the expected effect on the cost effectiveness estimates?	Uncertain.
What additional evidence or analyses might help to resolve this key issue?	The ERG would still recommend an analysis to estimate the effect of subsequent anticancer treatment and an adjustment to the CONCUR and CORRECT trial outcomes even if this is a challenge to apply to the T/T trials in the network meta-analysis (NMA). Estimates of the use of subsequent anticancer treatment in the T/T trials could assist the estimation of at least the direction of effect on the comparison with T/T.
BSC = best supportive care; ERG = Evidence Review Group; NHS = National Health Service; NMA = network meta-analysis; T/T = trifluridine/tipiracil	

1.4 The clinical effectiveness evidence: summary of the ERG's key issues

Table 1.4: Key issue 3: Lack of external validity of and comparability between the regorafenib and trifluridine/tipiracil RCTs

Report Sections	3.2, 3.3, 3.4
Description of issue and why the ERG has identified it as important	It is uncertain as to the size and direction of subgroup differences in terms of treatment experience and race in the CORRECT and CONCUR trials. There is also a disparity between the trials and National Health Service (NHS) clinical practice. This also has implications for the comparability with and external validity of the trifluridine/tipiracil (T/T) trials included in the network meta-analysis (NMA).
What alternative approach has the ERG suggested?	The ERG supports the sensitivity analyses in the NMA employing various combinations of trials.
What is the expected effect on the cost effectiveness estimates?	Uncertain.
What additional evidence or analyses might help to resolve this key issue?	Further sensitivity analyses for the NMA with different combinations of trials could be conducted, although considering the lack of difference between the combinations already examined this is unlikely to be particularly informative.
ERG = Evidence Review Group; NHS = National Health Service; NMA = network meta-analysis; T/T = trifluridine/tipiracil	

Table 1.5: Key issue 4: Difference in the treatment effect of regorafenib versus trifluridine/tipiracil depending on evidence source

Report Sections	3.2, 3.3, 3.4, 4.2.6
Description of issue and why the ERG has identified it as important	The network meta-analysis (NMA) combines ostensibly the highest quality evidence in the form of randomised controlled trials (RCTs) for the comparison of regorafenib with trifluridine/tipiracil (T/T) and seems to indicate that, depending on which trials are included, there is little difference in either overall survival (OS) or progression-free survival (PFS). However, there is doubt as to the quality of these RCTs and their comparability. Also, although three of the comparative observational studies provide general support for this conclusion, the largest and arguable best quality comparative observational study by Nakashima 2020 showed an advantage to T/T in OS (hazard ratio (HR) 1.515 versus indirect treatment comparison (ITC) base-case of 0.99), comparison of PFS being impossible.
What alternative approach has the ERG suggested?	The ERG has presented a critique and comparison of all comparative evidence in order that the committee can make as informed a judgment as possible of the treatment effect of regorafenib versus T/T.
What is the expected effect on the cost effectiveness estimates?	A scenario using the HR estimated from Nakashima 2020 shows that regorafenib would not be cost effective: it results in negative incremental net monetary benefit (iNMB) (willingness-to-pay (WTP) thresholds of 30,000 and 51,000 per quality adjusted life year (QALY) gained) for regorafenib versus T/T.
What additional evidence or analyses might help to resolve this key issue?	Although not routinely done, the RCTs and observational comparative studies could be combined in a NMA, using methods as described in Technical Support Document (TSD) TSD 20.[2]

ERG = Evidence Review Group; HR = hazard ratio; iNMB = incremental net monetary benefit; ITC = indirect treatment comparison; NMA = network meta-analysis; OS = overall survival; PFS = progression-free survival; QALY = quality adjusted life year; RCTs = randomised controlled trials; T/T = trifluridine/tipiracil; TSD = Technical Support Document; WTP = willingness-to-pay

1.5 The cost effectiveness evidence: summary of the ERG's key issues

A full summary of the cost effectiveness evidence review conclusions can be found in Section 6.4 of this report. The company's cost effectiveness (CE) results are presented in Section 5, the ERG's summary and detailed critique in Section 4, and the ERG's amendments to the company's model and results are presented in Section 6. The main ERG results are reproduced using confidential Patient Access Schemes (PAS) in a confidential appendix. The key issues in the CE evidence are discussed in Tables 1.6 to 1.10.

Table 1.6: Key issue 5: Treatment effectiveness and extrapolation

Report Section	4.2.6
Description of issue and why the ERG has identified it as important	The company implemented Kaplan-Meier (KM) curves instead of parametric survival models for the survival analyses of progression-free survival (PFS) and time on treatment (ToT) for regorafenib and best supportive care (BSC), leading to potential overfitting of modelled survival outcomes.

Report Section	4.2.6
What alternative approach has the ERG suggested?	Implement parametric models based on National Institute for Health and Clinical Excellence Decision Support Unit Technical Support Document 14 (NICE DSU TSD 14) for survival analyses of PFS and ToT.
What is the expected effect on the cost effectiveness estimates?	Implementing the fully parametric models chosen by the ERG decreases the incremental net monetary benefit (iNMB) of regorafenib compared to trifluridine/tipiracil (T/T) and BSC.
What additional evidence or analyses might help to resolve this key issue?	Not applicable.
BSC = best supportive care; ERG = Evidence Review Group; iNMB = incremental net monetary benefit; KM = Kaplan-Meier; NICE DSU TSD 14 = National Institute for Health and Clinical Excellence, Decision Support Unit, Technical Support Document 14; PFS = progression-free survival; ToT = time on treatment; T/T = trifluridine/tipiracil	

Table 1.7: Key issue 6: Adverse events

Report Section	4.2.7
Description of issue and why the ERG has identified it as important	The company positions regorafenib as being a chemotherapy-free alternative therapy that has a different adverse event (AE) profile compared to trifluridine/tipiracil (T/T). Hence, next to grade 3+ AEs, low grade AEs may also be relevant, but the company did not consider these in their economic model.
What alternative approach has the ERG suggested?	A scenario analysis and updated economic model including grade 1 and 2 AEs.
What is the expected effect on the cost effectiveness estimates?	Unclear.
What additional evidence or analyses might help to resolve this key issue?	A scenario analysis and updated economic model including grade 1 and 2 AEs.
AEs = adverse events; ERG = Evidence Review Group; T/T = trifluridine/tipiracil	

Table 1.8: Key issue 7: Resource use and costs

Report Section	4.2.9
Description of issue and why the ERG has identified it as important	The company assumed different relative dose intensity (RDIs) for regorafenib and trifluridine/tipiracil (T/T). However, a large observational study by Nakashima 2020[1], directly comparing regorafenib to T/T, reported comparable dose reductions (54% and 48% respectively).

Report Section	4.2.9
What alternative approach has the ERG suggested?	Assuming equal RDIs for regorafenib and T/T (the pooled RDI of [redacted] from CONCUR and CORRECT).
What is the expected effect on the cost effectiveness estimates?	Assuming equal RDIs for regorafenib and T/T decreases the incremental net monetary benefit (iNMB).
What additional evidence or analyses might help to resolve this key issue?	Not applicable.
ERG = Evidence Review Group; iNMB = incremental net monetary benefit; RDI = relative dose intensity; T/T = trifluridine/tipiracil	

Table 1.9: Key issue 8: Company's sensitivity analyses

Report Section	5.2
Description of issue and why the ERG has identified it as important	In response to the clarification letter, the company did not comply with three requests: 1) to provide a table presenting information about the observed and modelled overall survival (OS) and progression-free survival (PFS) using restricted mean survival time (RMST), 2) to provide all scenario analyses for the best supportive care (BSC) comparison that were conducted for trifluridine/tipiracil (T/T) comparison, and 3) to provide all scenario analyses with the fully parametric models applied for PFS and time on treatment (ToT).
What alternative approach has the ERG suggested?	The Evidence Review Group (ERG) requested the company to provide the table with observed and modelled OS and PFS using RMST, and all scenario analyses as stated above.
What is the expected effect on the cost effectiveness estimates?	Unclear.
What additional evidence or analyses might help to resolve this key issue?	A table with observed and modelled OS and PFS using RMST, and all scenario analyses as stated above.
BSC = best supportive care; ERG = Evidence Review Group; OS = overall survival; PFS = progression-free survival; RMST = restricted mean survival time; ToT = time on treatment; T/T = trifluridine/tipiracil	

Table 1.10: Key issue 9: Lack of a fixed random seed in model PSA

Report Section	5.3
Description of issue and why the ERG has identified it as important	The results of the probabilistic sensitivity analysis (PSA) are slightly different when running the same analysis multiple times (without

Report Section	5.3
	changing model settings), likely due to the lack of a fixed random seed in the model PSA.
What alternative approach has the ERG suggested?	Implement fixed random seed to model PSA.
What is the expected effect on the cost effectiveness estimates?	The implementation of a fixed random seed will make the results of the model PSA reproducible.
What additional evidence or analyses might help to resolve this key issue?	Implement a fixed random seed to the model PSA.
ERG = Evidence Review Group; PSA = probabilistic sensitivity analysis	

1.6 Summary of the ERG's view

The company submission (CS) base-case probabilistic incremental net monetary benefits (iNMBs) of regorafenib versus T/T were [redacted] (willingness-to-pay (WTP) £30,000 per QALY gained) and [redacted] (WTP £51,000 per QALY gained), respectively. For regorafenib versus BSC, these were [redacted] (WTP £30,000 per QALY) and [redacted] (WTP £51,000 per QALY gained), respectively. The estimated ERG base-case iNMBs for regorafenib versus T/T (probabilistic), based on the ERG preferred assumptions highlighted in Section 6.1, were [redacted] (WTP £30,000 per QALY) and [redacted] (WTP £51,000 per QALY). For regorafenib versus BSC, these were [redacted] (WTP £30,000 per QALY) and [redacted] (WTP £51,000 per QALY). The most influential adjustment was the assumption of equal relative dose intensities (RDIs) for regorafenib and T/T. The scenario analysis using an alternative HR for OS in the T/T arm had a large impact on the iNMB of regorafenib versus T/T.

In conclusion, there remains uncertainty about the effectiveness and CE of regorafenib, which can be at least partly resolved by the company by conducting further analyses (e.g., incorporation of low-grade AEs into the economic model). Moreover, the contradicting estimations of the relative effectiveness of regorafenib versus T/T in terms of OS (based on the ITC conducted by the company and the direct comparison from the literature) adds substantial uncertainty to the effectiveness and CE of regorafenib. Therefore, the ERG believes that the CS nor the ERG report contains an unbiased CE estimation of regorafenib compared with relevant comparators.

Table 1.11: Summary of ERG's preferred assumptions and ICER

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER[1] (£/QALY)	iNMB[2] (£30,000)	iNMB[3] (£51,000)
CS base-case							
Regorafenib	[redacted]	[redacted]					
T/T	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Matter of judgement (1-Log-normal for OS BSC instead of log-logistic)							
Regorafenib	[redacted]	[redacted]					
T/T	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER[1] (£/QALY)	iNMB[2] (£30,000)	iNMB[3] (£51,000)
Matter of judgement (2-Implementation of parametric survival curves for PFS)							
Regorafenib	[redacted]	[redacted]					
T/T	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Matter of judgement (3-Implementation of parametric survival curves for ToT)							
Regorafenib	[redacted]	[redacted]					
T/T	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Matter of judgement (4-Equal RDIs for regorafenib and T/T)							
Regorafenib	[redacted]	[redacted]					
T/T	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Deterministic ERG base-case							
Regorafenib	[redacted]	[redacted]					
T/T	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Probabilistic ERG base-case							
Regorafenib	[redacted]	[redacted]					
T/T	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

BSC = best supportive care; CS = company submission; ERG = Evidence Review Group; ICER = incremental cost-effectiveness ratio; iNMB = incremental net monetary benefit; OS = overall survival; PFS = progression-free survival; RDI = relative dose intensity; ToT = time on treatment; T/T = trifluridine/tipiracil; QALY = quality adjusted life year; WTP = willingness-to-pay

[1]ICER versus regorafenib

[2]iNMB for WTP of £30,000 per QALY

[3]iNMB for WTP of £51,000 per QALY

2. CRITIQUE OF COMPANY'S DEFINITION OF DECISION PROBLEM

Table 2.1: Statement of the decision problem (as presented by the company)

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG comment
Population	Adults with metastatic colorectal cancer (mCRC) who have been previously treated with or are not considered candidates for available therapies.	Adults with mCRC who have failed on first-line chemotherapy/first-line biologic and who are being considered for ≥ third-line treatment. Specifically, we are seeking a recommendation for patients for whom treatment with trifluridine/tipiracil (T/T) is being considered.	Physicians have requested an alternative to T/T at the third or later line setting. Physicians have indicated that regorafenib, with its comparable efficacy to T/T, but different adverse event profile, would provide an alternative treatment option for these patients and is the patient group for whom regorafenib would be considered.	The population is more precise/narrower than the scope except that best supportive care (BSC) is included as a 'minor comparator', which would contradict the definition according to eligibility for T/T.
Intervention	Regorafenib	As per final scope	Not applicable	The intervention is in line with the National Institute for Health and Care Excellence (NICE) scope
Comparator(s)	Single-agent irinotecan (after FOLFOX) FOLFIRI (after either FOLFOX or CAPOX) FOLFOX (after either FOLFIRI or CAPOX) Raltitrexed (if 5-FU/FA are not suitable) T/T BSC	T/T (main comparator – full set of economic analyses) BSC (minor comparator – reduced set of economic analyses)	Irinotecan, FOLFIRI, FOLFOX, raltitrexed These comparators are outside of regorafenib's marketing authorisation (in italics) *Regorafenib is indicated as monotherapy for the treatment of adult patients with mCRC who have been previously treated with, or are not considered candidates for, available therapies. These include fluoropyrimidine-based chemotherapy, an anti-vascular endothelial growth*	The company seems to have provided adequate justification for excluding the chemotherapies, except T/T. However, as mentioned for the population, it is unclear how BSC can also be a comparator.

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG comment
			factor (VEGF) therapy and an anti-epidermal growth factor receptor (EGFR) therapy. The listed treatments were available before regorafenib was licensed in mCRC and fall under the definition of "available therapies" in the license wording. Consequently, regorafenib is not an alternative to these agents and they are not comparators. The registration trials for regorafenib investigated its use in patients who had received these 'available therapies'. Physicians would not consider regorafenib as an alternative to these treatments and would only consider regorafenib after these treatments have failed.	
Outcomes	Overall survival (OS) Progression-free survival (PFS) Response rates Adverse effects of treatment Health-related quality of life (HRQoL)	As per final scope	Not applicable	Only OS and PFS are including in the comparison with T/T.
Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost	Not reported in Table 1.	Not applicable	In line with the reference case.

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG comment
	per quality-adjusted life year (QALY). The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. Costs will be considered from a National Health Service (NHS) and Personal Social Services (PSS) perspective. The availability of any commercial arrangements for the intervention, comparator and subsequent treatment technologies will be taken into account.			
Other considerations	If the evidence allows subgroups will be considered based on previous treatment received for mCRC.	Not reported in Table 1.	Subgroup analyses were presented for the regorafenib trials, CORRECT and CONCUR, but not for the comparison with T/T in terms of clinical effectiveness or cost effectiveness.	The Evidence Review Group (ERG) requested a subgroup analysis for the comparison with T/T in terms of clinical effectiveness or cost effectiveness by prior anti-VEGF treatment given that it is not recommended in the United Kingdom (UK) but experienced by a large proportion of patients in CORRECT and CONCUR.

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG comment
Based on Table 1 of the CS[3] BSC = best supportive care; CS = company submission; ERG = Evidence Review Group; HRQoL = health-related quality of life; mCRC = metastatic colorectal cancer; NHS = National Health Service; NICE = National Institute of Health and Care Excellence; OS = overall survival; PFS = progression-free survival; PSS = Personal Social Services; QALY = quality-adjusted life year; T/T = trifluridine/tipiracil; UK = United Kingdom				

2.1 Population

The population defined in the scope is: adults with metastatic colorectal cancer (mCRC) who have been previously treated with, or are not considered candidates for, available therapies.[4] The population in the company submission (CS) is limited to *"Adults with metastatic colorectal cancer who have failed on first-line chemotherapy/first-line biologic and who are being considered for ≥ 3rd-line treatment. Specifically, we are seeking a recommendation for patients for whom treatment with trifluridine/tipiracil is being considered"*.[3]

In 2013, a European Marketing Authorisation was granted for regorafenib "*...for the treatment of adult patients with:*

- *mCRC who have been previously treated with, or are not considered candidates for, available therapies. These include fluoropyrimidine-based chemotherapy, an anti-VEGF therapy and an anti-EGFR therapy*
- *Unresectable or metastatic GIST who progressed on or are intolerant to prior treatment with imatinib and sunitinib*
- *HCC who have been previously treated with sorafenib"* (page 16)[3]

ERG comment:

There does appear to be a match between the National Institute for Health and Care Excellence (NICE) scope and at least one of the indications for regorafenib. However, the decision problem population is more precisely defined. Failed on first line treatment and at least third line is narrower than "previously treated with, or are not considered candidates for, available therapies" in that such patients might not have failed first line but could have been intolerant to it and might theoretically include second line. However, it is defined in a second way i.e., in terms of eligibility for trifluridine/tipiracil (T/T), which begs the question of how this population is defined. The NICE Technology Assessment TA405 recommends T/T for mCRC *"...in adults who have had previous treatment with available therapies including fluoropyrimidine-, oxaliplatin- or irinotecan-based chemotherapies, anti-vascular endothelial growth factor (VEGF) agents and anti-epidermal growth factor receptor (EGFR) agents, or when these therapies are not suitable,..."* (page 4)[5] This does seem to match the population in the scope at least superficially, but 'available therapies' as specified in the scope implies any therapy, which could include T/T, which logically implies that the only remaining option would be best supportive care (BSC). Effectively, this would mean that the population would effectively be a later line than the one specified by the company where patients are still eligible for at least one more active treatment i.e., T/T. This is consistent with Figure 1 in the CS, which the company attributes to the NICE pathways for colorectal cancer and mCRC i.e., BSC follows T/T, which occurs at either third or fourth line depending on biomarker status (and resulting treatment history). However, the NICE scope does list more than BSC as a comparator i.e., some active treatments including T/T and so the Evidence Review Group (ERG) therefore considers that, if T/T can be a comparator then it is reasonable for the company to define the population as those eligible for T/T and that this would be prior to BSC. Unfortunately, the company seem to undermine their definition of the population in terms of T/T eligibility by stating in Table 1 that BSC is a "minor comparator" for which they carry out come economic analyses.

The ERG requested clarification on 'available therapies' other than T/T to which the company responded that "*available therapies" includes all comparators in the scope other than trifluridine/tipiracil. Regorafenib is only considered after failure of these agents*."[6] The ERG also requested clarification as to what is meant by BSC being a minor comparator to which they responded that: *"the population is intended to be earlier than BSC thus ruling out BSC as a comparator."* They

went on to argue that T/T is recommended as a "*last-line option",* which implies that if regorafenib is compared to it, then it is also similarly placed.

The ERG also asked how patients would be identified in United Kingdom (UK) clinical practice to be only eligible for T/T in terms of treatment history and any other clinical characteristics to which the company replied that earlier treatments would have to failed or been inappropriate and patients would have to be *"'fit' enough to receive chemotherapy and to be able to tolerate the adverse events typical of chemotherapies."* The company stated that this tolerance could be established according to prior experience with chemotherapies.[6] The implication of this is that patients who might not possess this tolerance would have BSC, but these patients are not eligible for either T/T or regorafenib.

In conclusion, the ERG considers that if the population is defined essentially according to the comparator T/T then only patients who might be considered for T/T and no other treatment should be considered for regorafenib.

2.1.1 Subsequent treatment

The ERG also noted that there appeared to be a discrepancy between what would be expected in UK clinical practice at the next line of therapy i.e., following T/T or regorafenib and what occurred in CORRECT and CONCUR. The company stated that *"There was some post-progression treatment in the CORRECT and CONCUR trials (CORRECT: Regorafenib 26%, BSC 30%; CONCUR Regorafenib 31%, BSC 43%)."* (page 142). Also, a scenario analysis of the economic evaluation with subsequent treatment included was conducted. However, Figure 1 in the CS suggests that only BSC would follow T/T or regorafenib and the CS states *"... clinical experts have advised that in England and Wales, patients receiving regorafenib or trifluridine/tipiracil are unlikely to receive further active treatment after progression due to the advanced nature of the disease and limited treatment options available."* (page 142) In response to clarification request the company reproduced data from Appendix D, which showed systemic anticancer therapy use to be: 29.8% versus 25.9% in CORRECT 42.6% versus 30.9% in CONCUR for placebo versus regorafenib respectively.[6, 7] In contrast the company stated that in UK clinical practice: *"Experts consulted by Bayer suggested that <10% of patients would be fit enough for subsequent active treatment."* The ERG requested data to support the opinion of <10% use to which the company responded: *"We do not have the data requested."* The company also stated that no comparable data were available for the T/T trials: they did provide some data from the *"...named patient programme for trifluridine/tipiracil prior to its licensing in 2016 (Iverson 2020 https://doi.org/10.1186/s12885-020-6577-1)."* However, the ERG notes that these seemed to be prior to T/T as opposed to as subsequent treatment.

Given the apparently large discrepancy between subsequent therapy use in the regorafenib trials and UK clinical practice, the ERG requested an estimate of the effect of subsequent treatment. The company responded by stating that there would be no effect on progression-free survival (PFS) given that subsequent treatment was administered post-progression. For overall survival (OS), the company stated: *"After investigating this question we consider that any attempts to isolate and quantify the effect of post-progression treatment would be flawed."*[6] They further explained that this was due to informative censoring. They then presented an analysis where those receiving subsequent treatment were censored. The ERG need to point out that, paradoxically, of all methods that might be employed to estimate the effect of or adjust for subsequent treatment, the one used by the company has been identified as prone to bias due to informative censoring, as described in NICE Technical Support Document (TSD) 16.[8] Indeed, despite the ERG referring to TSD 16, which describes methods for reducing this bias, the company stated: *"we do not believe these analyses are possible."* Of course, despite the apparent discrepancy between the trial and UK clinical practice, the company make the point that it might favour

BSC: *"As more patients in the placebo group received post-progression therapy, the results suggest that removal of this potential benefit from the placebo arm resulted in a slightly more favourable hazard ratio for regorafenib."*[6] The ERG agrees that this is probably the case with regards to the treatment effect of regorafenib versus BSC, although the effect on the treatment effect of regorafenib versus T/T remains unknown. In conclusion, the ERG considers that this uncertainty in effect of subsequent treatment remains a key issue, although without the data on subsequent therapy use in the T/T trials, the ERG cannot think of any way of reducing this uncertainty.

2.2 Intervention

The intervention regorafenib is in line with the scope. The CS states:

- "Regorafenib should be prescribed by physicians experienced in the administration of anti-cancer therapy.
- Recommended dose of regorafenib is 160 mg (four tablets of 40 mg) taken once daily for 3 weeks followed by 1 week off therapy. Treatment should continue as long as benefit is observed or until unacceptable toxicity occurs.
- Dose interruptions and/or dose reductions may be required based on individual safety and tolerability. Dose modifications are to be applied in 40 mg (one tablet) steps. The lowest recommended daily dose is 80 mg. The maximum daily dose is 160 mg (Table 1 and Table 2 in the summary of product characteristics (SmPC) describes different dose adjustments according to adverse event (AE) severity)" (Table 2, CS)[3]

ERG comment:

It is important to note that the intervention is regorafenib plus BSC, just as the comparator also includes BSC in some form, as distinct from BSC only. The presence of BSC as comparator in the randomised controlled trials (RCTs) and concomitant therapy prompted the ERG to request a comparison between BSC in the trials and BSC in UK clinical practice to which the company responded: *"It is not possible to provide all of the requested information - detailed data on exactly what is provided as best supportive care in the UK, and in what proportion to patients, is not available (see part e above). Furthermore, no detail on best supportive care is available to Bayer for the trifluridine/tipiracil trials."*[6] The company did provide a description of BSC for the CORRECT and CONCUR trials (see Table 2.2).

Table 2.2: Description of best supportive care in the regorafenib trials (CORRECT, CONCUR)

Description of best supportive care
Best supportive care (BSC) 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. Best supportive care excluded other investigational anti-tumour agents or anti-neoplastic chemo/hormonal/immuno-therapy.
BSC = best supportive care

The ERG considers that there might be some uncertainty as to the precise nature of BSC, and so there is uncertainty as to whether there is sufficient similarity between the regorafenib and T/T trials such that, even if there is a differential effect on prognosis, there is no substantial impact on the treatment effect between regorafenib and T/T (See Sections 3.3 and 3.4).

2.3 Comparators

The description of the comparators in the NICE scope is as follows:[4]

- Single-agent irinotecan (after FOLFOX)
- FOLFIRI (after either FOLFOX or CAPOX)
- FOLFOX (after either FOLFIRI or CAPOX)
- Raltitrexed (if 5-FU/FA are not suitable)
- T/T
- BSC

The company have chosen T/T as "main comparator" and BSC as "minor comparator".[3] They state that the other comparators are precluded by the marketing authorisation.

ERG comment: Trifluridine/tipiracil and BSC as the only comparators might be seen as insufficient, although the ERG agrees with the company that FOLFOX and FOLFIRI are excluded by the marketing authorisation. The other types of chemotherapy, irinotecan and raltitrexed would probably also be excluded as they would be considered 'available therapies'. In relation to the remaining comparators, as described in Section 2.1, there is a potential contradiction between the comparator implied by the scope population i.e., only BSC and this list in the scope. There is also a potential contradiction between the comparator implied by the decision problem i.e., T/T and the inclusion of BSC as "minor comparator". The ERG considers that, notwithstanding this latter contradiction, given that T/T is within scope and if the population is defined as only those who are eligible by the company, then T/T as the only comparator seems reasonable. However, as stated in Section 2.1, clarification has been requested regarding BSC as comparator, as well as its nature following T/T or regorafenib and how this compares to BSC as comparator and concomitant therapy in the trials. The company responded as reported in Section 2.1 that BSC is ruled out as a comparator and that because T/T is a comparator and it is a "*last-line option"* then there can be no other comparators.[6]

In conclusion, it seems reasonable to the ERG that T/T is the only comparator given that the company have essentially defined the population to be only those patients who would be eligible for T/T.

2.4 Outcomes

The NICE final scope lists the following outcome measures:

- OS
- PFS
- Response rates
- AEs
- Health-related quality of life (HRQoL)

These were all assessed in the company trials of regorafenib plus BSC versus BSC i.e., CORRECT and CONCUR (see Section 3.2). However, only OS and PFS were included in the comparison with T/T i.e., no data were presented for response rates, AEs or HRQoL (see Section 3.3).

ERG comment: Given that T/T is the comparator in the company decision problem, all outcomes in the scope would be expected to have been presented for this comparison. In Section B.3.4.4 of the CS, it states that the AE rates were pooled for T/T. Therefore, in the clarification letter, the ERG has requested that the company describe the method of pooling and whether randomisation was preserved. The ERG has also requested that a comparison with T/T be presented for AEs, including an indirect

treatment comparison (ITC) of any grade 3+ treatment emergent adverse events (TEAEs), all TEAEs used in the economic model and discontinuation due to AEs. The company responded by including three ITCs of AE outcomes (see Section 3.4).[6]

2.5 Other relevant factors

The NICE scope stated: *"If the evidence allows subgroups will be considered based on previous treatment received for metastatic colorectal cancer."*[4] However, although some subgroup analyses were conducted in CONCUR and CORRECT and CONCUR, none were conducted for the ITC versus T/T or for the cost-effectiveness analyses (CEAs). Indeed, anti-VEGF treatments are not indicated for patients which fall into the scope of this submission in the UK, but considerable part of the treatment population in both the CORRECT and CONCUR trials received anti-VEGF treatment. Therefore, the ERG requested a subgroup analysis of the ITC and the CEA based on prior anti-VEGF, given the potential role of prior anti-VEGF as a treatment effect modifier. The company responded that patient numbers were insufficient for a subgroup analysis according to anti-VEGF treatment.[6] The ERG consider that this is a reasonable conclusion, although the role of anti-VEGF treatment together with other patient characteristics in affecting the comparability between the regorafenib and T/T RCTs as well as the regorafenib and T/T cohorts in the comparative observational studies is discussed in more detail in Sections 3.2.3 and 3.4.4.

3. CLINICAL EFFECTIVENESS

3.1 Critique of the methods of review(s)

3.1.1 Searches

The following paragraphs contain summaries and critiques of all searches related to clinical effectiveness presented in the CS. The Canadian Agency for Drugs and Technologies in Health (CADTH) evidence-based checklist for the Peer Review of Electronic Search Strategies (PRESS), was used to inform this critique.[9, 10] The ERG has presented only the major limitations of each search strategy in the report

Appendix D of the CS detailed the systematic literature review (SLR) undertaken to identify all relevant clinical information from evidence related to regorafenib or T/T for the treatment of mCRC in the third- or later-line setting. The searches were conducted in March 2021 and updated February 2022. A summary of the sources searched is provided in Table 3.1.

Table 3.1: Data sources for the clinical effectiveness systematic review (as reported in the CS)

Resource	Host/Source	Date ranges	Dates searched
Electronic databases			
MEDLINE	Embase.com	From inception	5.3.21 Updated 22.2.22
Embase	Embase.com	From inception	5.3.21 Updated 22.2.22
MEDLINE In-Process	Pubmed	From inception	19.3.21 Updated 22.2.22
CDSR	Cochrane library (Wiley)	From inception	7.4.21 Updated 22.2.22
CENTRAL	Cochrane library (Wiley)	From inception	7.4.21 Updated 22.2.22
Conferences			
ASCO		2019-2022	2022/02
ESMO		2019-2022	2022/02
DDW		2019-2022	2022/02
Additional searches			
Handsearching	Bibliographies of key systematic review and meta-analysis articles were screened to fully evaluate the available clinical evidence		

ASCO = American Society of Clinical Oncology; CDSR = Cochrane Database of Systematic Reviews; CENTRAL = Cochrane Central Register of Controlled Trials; CS = company submission; ESMO = European Society for Medical Oncology; DDW = Digestive Disease Week

ERG comment:

- The CS and response to clarification provided sufficient details for the ERG to appraise the literature searches.

- A good range of databases and relevant conference proceedings were searched. Strategies contained a good mix of free text and subject headings.
- The CS stated that the MEDLINE search was conducted via Embase.com. At clarification the ERG asked the company to confirm whether this was a search of the Embase database conducted on the understanding that it now contains all records from MEDLINE or a multifile search where both resources were searched simultaneously using the same strategy. The company responded that *"The Embase.com platform was used to run a multi-faceted search strategy to identify records from both Embase and MEDLINE databases"*.[6] Unfortunately, the ERG was unclear as to what was meant by the term "multi-faceted", as that terminology usually relates to concepts being combined within a search strategy, not the resources being searched, which is what is meant by "multi-file" searching. For clarity the ERG considers it preferable to conduct a separate companion MEDLINE search. Other concerns have also been raised over the combined search options. The Cochrane Handbook refers to potential limitations regarding searches of MEDLINE content on Embase: "*In addition, a recent study found that records from MEDLINE were not always retrieved when searched through Embase due to MeSH not being available in Embase (Bramer et al 2017a). Although it is, therefore, technically possible to search across all MEDLINE records in Embase (note, not all PubMed records), it is recommended that both databases be searched separately*."[11]. With regard to the multifile approach, although simultaneous searching of Embase.com should automatically identify and search for equivalent MEDLINE medical subject heading (MeSH) terms, it is not clear if this is the case for all potentially useful MeSH terms. Given the potential limitations of this approach it may be safer to ensure inclusion of both Emtree and MeSH terms in the search strategy. Given the lack of relevant MeSH terms such as Colorectal Neoplasms/ or Rectal Neoplasms/ in the reported strategy the ERG is unclear what impact this may have had on the overall recall of results.
- The CS reported that MEDLINE In-Process was searched using PubMed. However, it appears that the search limit used in PubMed identifies recently added records, not in-process records: (publisher[sb] NOT pubstatusnihms NOT pubstatuspmcsd NOT pmcbook) OR (pubstatusaheadofprint). The correct subset to use is 'inprocess[sb]'.[12] Therefore, whilst the company's PubMed search would potentially find relevant records not retrieved by the Embase search, it would not retrieve in-process records. This omission was corrected in the update searches for the costs effectiveness Section but was not reported here.
- The company did not report any searches of clinical trials registries, such as ClinicalTrials.com or the World Health Organization (WHO) International Clinical Trials Registry Platform (ICTRP) to identify completed and ongoing clinical trials. Whilst Cochrane Central Register of Controlled Trials (CENTRAL) now contains records from Clinical Trials.gov, research suggests that this approach to searching for ongoing trials is insufficient, and that searching CENTRAL should not be a substitute for searching ClinicalTrials.gov and/or ICTRP to identify unpublished trials.[13] However, given the other searches reported this is unlikely to have affected the overall recall of results.
- The Embase/MEDLINE search for the clinical effectiveness search was structured as follows:

CRC
+
Non-response/Relapsed/refractory/2-4 line/etc
+
Named drugs (regorafenib/tipiracil/trifluridine/nivolumab/ipilimumab/encorafenib)
+

RCTs/Obs
+
(Limits: No animals/SRs/letters etc)

The CS reported that the searches were designed from "a multi-country perspective and therefore included comparators that are not relevant to this appraisal" (see Appendix D, Section B.3.1).[7] It was further stated that results relating to the comparators (nivolumab/ipilimumab/encorafenib) were excluded from this appraisal. The searches also included a facet for terms relating to non-response etc refining the number of retrieved studies. However, the ERG was concerned the inclusion of this facet may have been overly restrictive and adversely affected the recall of results, especially those relating to studies of tipiracil plus trifluridine. In the clarification letter the ERG suggested that a more cautious approach may have been to remove the three additional named drugs, which would have lowered the hits retrieved allowing for the removal of the non-response facet resulting in the following structure:

CRC
+
Named drugs (regorafenib/tipiracil/trifluridine)
+
RCTs/Obs
+
(Limits: No animals/SRs/letters etc)

The ERG asked the company to rerun this search to ensure that no additional relevant studies had been missed. The company reran both the Embase/MEDLINE and CENTRAL searches with the suggested format and reported the following: "*After removing the duplicates across these two databases there were 195 records that were unique to this search i.e., were not 'returned' by the submitted search. These 195 records were reviewed against the same inclusion/exclusion criteria as the original search and 174 records were excluded at the title/abstract stage. Full-text review was performed for 21 publications. Overall, the updated search strategy search located the same five studies that were identified in the original search with no additional RCTs relevant to the appraisal being located.*"[6] However, rescreening of the 21 papers by the ERG identified one additional relevant study, which was an observational comparative study by Nakashima 2020,[1] (see below for further details).

- The ERG requested full search strategies for the conference proceedings reported in the CS, these were provided at clarification. On closer inspection the ERG noticed that the original date span of 2016-2022/02 appeared to be a typographical error as the provided searches were from 2019-2021. Whilst no strategies for 2022 were provided, an update was reported in the Preferred Reporting Items for Systematic Revies and Meta-analyses (PRISMA) flow chart which retrieved no additional proceedings. The numbers provided in the strategies also did not match the search flow, but the ERG presumes that the PRISMA numbers reported the number of relevant papers after screening, not the total identified as reported in the strategy.

3.1.2 Inclusion criteria

The eligibility criteria used in the search strategy for RCTs and non-RCTs as well as observational evidence is presented in Table 3.2. Two reviewers independently of each other screened the titles and abstracts and full text papers. To reach consensus, discrepancies in screening results were checked by involving a third independent reviewer.

Table 3.2: Eligibility criteria used in search strategy for RCT and non-RCT evidence

	Inclusion criteria
Population	Patients with relapsed/refractory metastatic colorectal cancer (mCRC) previously treated with standard therapies e.g., fluorouracil, capecitabine, oxaliplatin, irinotecan or cetuximab monotherapy, or combination therapy
Interventions	• Regorafenib • Trifluridine/tipiracil (T/T)
Comparators	• Placebo • Best supportive care (BSC) (author-defined) • Any other pharmacological/non-pharmacological intervention
Outcomes	**Efficacy:** response rate (complete, partial, overall), survival (overall survival (OS), progression-free survival (PFS), event-free survival (EFS)); duration of response **Safety:** incidence of adverse events (AEs); treatment discontinuation **Utilities:** health-related quality of life (HRQoL)
Study design	• Randomised controlled trials (RCTs) • Non-RCTs • Retrospective and prospective cohort studies • Real-world evidence studies
Language restrictions	English language only
Exclusion criteria	
Population	• Healthy volunteers • Paediatric population • Treatment-naïve patients with mCRC • Early-stage mCRC • Disease other than mCRC
Interventions	• Non-pharmacological interventions (e.g., herbal remedies) • Interventions not included in the list • Surgery
Comparators	None
Outcomes	• Studies assessing only pharmacodynamics • Studies assessing outcomes not relevant to the review
Study design	• Single-arm trials • Reviews, letters, comments and editorials • Case studies or case reports of <10 patients
Language restrictions	Non-English

Source: Table 1 of the Appendix D[3]

AE = adverse event; BSC = best supportive care; CS = company submission; EFS = event-free survival; HRQoL = health-related quality of life; mCRC = metastatic colorectal cancer; OS = overall survival; PFS = progression-free survival; RCT = randomised controlled trial; T/T = trifluridine/tipiracil

ERG comments:

- As per Section B.3.1.2 of the Appendix D, a set of UK-specific inclusion and/or exclusion criteria were applied. The company was asked to clarify this, and responded by stating that, "*To determine the clinical evidence base that is applicable to the appraisal, inclusion and exclusion criteria were applied to the literature search results as per Table 1, Section B.3.1.2, Appendix D and aligned with the decision problem as per Table 1, Section B.1.1, Document B. We used the term "UK-specific" to clarify that the evidence base aligns with the decision problem. For example, the broader search included terms for encorafenib, a treatment not included in the scope and not relevant to the decision problem. Trials for encorafenib were excluded as part of the "UK-specific" criteria*." [6]
- Population: as specified in the NICE scope, the population was: "*Adults with metastatic colorectal cancer who have been previously treated with, or are not considered candidates for, available therapies"*. However, patients with "early-stage mCRC" were excluded: it was initially unclear what was meant by 'early stage'. The company response was that "*the exclusion criteria is more correctly "Early-stage CRC" - defined as those with non-metastatic disease stage I, II and III…"*.[6] The ERG agreed that the company's amendment of the exclusion criterion definition removed the apparent contradiction.
- Choice of comparators: the list seems to have been broad enough to include any of the comparators in the NICE scope, although T/T seems to be the main one in the decision problem, notwithstanding the doubt about BSC (see Sections 2.1 and 2.3). The ERG was uncertain about the precise meaning of "author defined best supportive care". The company response was "*There is a lack of standard definition for best supportive care (BSC) in mCRC and it varies according to individual symptoms and needs. BSC can include physical, psychological, social, and spiritual support. What constitutes BSC is seldom described in publications, hence it was decided to consider a treatment as BSC based purely on the author defining it as such*." [6] The ERG judged this response as reasonable, although it would stress that this means that 'author defined best supportive care' is a very broad concept, and such care might vary greatly in utility, and therefore efficacy with implications for the comparison with T/T via the network meta-analysis (NMA) (see Sections 3.2, 3.3 and 3.4).
- Study designs: RCTs and non-RCTs were restricted to current second line and later treatment, while observational and real-world evidence (RWE) studies were restricted to current third line and later treatment for mCRC. Responding to a request for further clarification, the company stated that, "*The proposed position for regorafenib is as an alternative to trifluridine/tipiracil which is used as ≥3rd-line treatment (please see our response to A8). Restricting to second line and later treatment is sufficiently broad not to miss any RCTs of relevance to the decision problem. The search conducted in response to A1 confirms that the studies relevant to the decision problem have been located. RWE provides supporting information only and therefore, for pragmatic reasons, the search was more restrictive than for RCTs, but still in line with the decision problem.*"[6] The ERG is not fully convinced by this explanation, particularly since a search where 'second line and later treatment' is used as a term might actually be *more* likely to miss studies looking at third line and later treatment. Also, as mentioned below, the RWE studies do seem to have substantial value in the context of an NMA with potential comparability issues (see Sections 3.2, 3.3 and 3.4).
- The ERG does not fully understand this criterion: in Table 17 of the CS, CORRELATE was listed as an observational study of regorafenib (first line).[3] Also, it is unclear why single-arm trials were excluded while observational and RWE studies were admissible. The company was asked to clarify this, and their response was that "*Single-arm studies do not provide evidence of comparative effectiveness and cannot be used in indirect treatment comparisons. RWE tends not to be multi-*

arm, and therefore in order to provide real-world evidence it was necessary to be less restrictive."[6] The ERG regarded the argument of comparative effectiveness as a good rationale.

- Language restrictions: the ERG considers the application of non-English language restrictions as a source of potential bias i.e.; some potentially relevant studies might have been missed. The company response was, "*the exclusion criteria was to exclude non-English publications, not non-English studies. We don't anticipate any impact of this exclusion criteria as studies of high quality are most likely to be reported in English language publications. This is a commonly applied exclusion criteria in systematic literature reviews*"[6] The ERG does not believe that the company has justified its position sufficiently: high quality science and English-language reporting have no natural association, and justification by appeal to common practice is an established fallacy. Therefore, doubt about the wisdom of the company's approach remains.

3.1.3 Critique of data extraction

As per Section B.3.1.2 of the Appendix D,[3] "all extracted data were verified against the original source paper by a second researcher". It is unclear whether two reviewers independently of each other extracted the data which is considered to be the *gold standard* in the SLR process. In the clarification letter, the company stated that, "*data extraction was performed by two independent reviewers where initially data from each study was extracted by one researcher and the extracted data was quality checked against the original source paper by an independent second researcher.*"[6] Therefore, the gold-standard process was not followed, but the ERG is satisfied that a reasonable standard of practice was achieved.

3.1.4 Quality assessment

As per Section B.3.1.5 of the Appendix D,[3] *"Included RCTs were critically appraised using the National Institute for Health and Care Excellence (NICE) methodology checklist in line with the Cochrane risk of bias tool"*. The ERG is concerned with the assumed compatibility of the NICE checklist and the Cochrane risk of bias tool. Also, the appropriate reference for the Cochrane risk of bias tool is missing.

3.1.5 Evidence synthesis

In the abstract/title screening phase of the CS[3] SLR, 1,679 records were excluded and 626 were retained for full text screening, including an additional four articles found from the grey literature. From these, five RCTs were included in the SLR, alongside 10 observational and RWE studies. Two RCTs evaluating regorafenib (REDOS2 and REARRANGE3) were excluded. The company explained that this was because they compared different dosing approaches and so are not relevant to the decision problem. A further 578 of the 626 trials were also excluded. Reasons included inappropriateness of comparator (n=17), disease (n=27), disease stage (n=15), intervention (n=136), line of therapy (n=73), outcome (n=17), population (n=9), or study design (n=259), or because the article was a review/editorial (n=25). Because of the large number of excluded studies, only included studies are featured in Table 3.3.

Table 3.3: Trials included in CS SLR

Trial name	Drugs evaluated	Primary study
RCTs		
CORRECT (NCT01103323)	Regorafenib versus placebo	Grothey 2013[14]
CONCUR (NCT01584830)	Regorafenib versus placebo	Li 2015[15]

Trial name	Drugs evaluated	Primary study
RECOURSE (NCT01607957)	Trifluridine/tipiracil versus placebo	Mayer 2015[16]
TERRA (NCT01955837)	Trifluridine/tipiracil versus placebo	Xu 2018[17]
N/A	Trifluridine/tipiracil versus placebo	Yoshino 2012[18]
Observational studies		
REBECCA (NCT02310477)	Regorafenib	Adenis 2016[19]
CORRELATE (NCT02042144)	Regorafenib	Ducreux 2019[20]
RECORA (NCT01959269	Regorafenib	Schulz 2018[21]
CORRECT	Regorafenib	Novakova-Jiresova 2020[22]
NA	Regorafenib versus trifluridine tipiracil	Tanaka 2018[23]
NA	Regorafenib	Banzi 2019[24]
NA	Regorafenib versus TAS-102	Sueda 2016[25]
NA	Regorafenib versus TAS-102	Huemer 2020[26]
NA	Regorafenib	Hirano 2015[27]
NA	Regorafenib	Zengin 2018[28]

Source: Tables 2 and 3 in CS appendices[7]
CS = company submission; NA = not applicable; RCT = randomised controlled trial; TAS-102 = trifluridine/tipiracil

ERG comment: REDOS2 and REARRANGE3 may not have been relevant to the decision problem but could have been useful for providing safety data. In the clarification letter, the company was asked to include these studies for the safety outcomes. The company have provided safety data, which has been added as Section 3.2.6.7.[6]

3.2 Critique of trials of the technology of interest, their analysis and interpretation (and any standard meta-analyses of these)

In the CS[3], the company considered the CORRECT and CONCUR trials to be the only randomised studies identified by the clinical SLR to explore the effectiveness and safety of regorafenib for the treatment of adult patients with mCRC receiving ≥ third line therapy.

In addition, 10 observational reports on the efficacy of regorafenib were included in the SLR. These are summarised in Table 3.3 above. Because these were only used in the CS[3] as supportive evidence, and because they constitute a lower level of evidence than the available randomised evidence, results from most of these are not included in this report, other than the three studies that directly compared regorafenib to T/T[23, 25, 26]. Details of these three studies are provided in Section 3.2.1.2.

A further relevant observational trial[1] was found after an additional search requested during clarification (see Section 3.1.1). This trial had been missed by the company's previous searches. This has been added to the evidence in the report in Section 3.2.1.2.

3.2.1 Details of the included trials

3.2.1.1 Randomised trials

The two trials evaluating regorafenib (plus BSC) versus placebo (plus BSC) were CORRECT and CONCUR. Both were phase III randomised double-blind placebo-controlled multi-centre studies, comprising patients (aged ≥18 years) with stage IV mCRC. Both used the same outcomes of OS, PFS, response rate, AEs and HRQoL. They differed in that the CORRECT study contained patients from Asian and non-Asian countries who had progressed on approved, standard treatments, whereas CONCUR contained patients exclusively from Asian countries who had failed ≥two lines of prior treatment. Best supportive care excluded other investigational anti-tumour agents or antineoplastic chemotherapy, hormonal therapy, or immunotherapy. A summary of the study methodology from CORRECT and CONCUR is presented in Table 3.4.

Table 3.4: Study methodology for CORRECT and CONCUR

Trial number (acronym)	CORRECT: NCT01103323	CONCUR: NCT01584830
Location	Global: 105 centres across 15 countries	Asia: Mainland China, Hong Kong, Taiwan, Vietnam, and South Korea
Trial design	Randomised, double-blind, placebo-controlled multi-centre Phase III study	Randomised, double-blind, placebo-controlled multi-centre Phase III study
Eligibility criteria for participants	**Key inclusion criteria:** Adults (≥18 years; both sexes) with mCRC (Stage IV) and measurable or non-measurable disease according to RECIST criteria version 1.1 and a life expectancy of at least 3 months. Histological or cytological documentation of adenocarcinoma of the colon or rectum. All other histological types were excluded. Progression during or within 3 months following the last administration of approved standard therapies, which was to include fluoropyrimidine, oxaliplatin, irinotecan, bevacizumab and cetuximab or panitumumab (if *KRAS* WT). Patients treated with oxaliplatin in an adjuvant setting were to have progressed during or within 6 months of completion of adjuvant therapy. Patients who had progressed more than 6 months after completion of oxaliplatin containing adjuvant treatment were to be retreated with oxaliplatin-based therapy to be eligible. Patients who had withdrawn from standard treatment due to unacceptable toxicity warranting discontinuation of treatment and precluding retreatment with the same agent prior to progression of disease were also eligible to enter the study. Patients with an unknown KRAS status at screening were to have received prior anti-EGFR treatment.	**Key inclusion criteria[b]:** Asian adults (≥18 years; both sexes) with mCRC (Stage IV) and measurable or non-measurable disease according to RECIST criteria v1.1 and a life expectancy of at least 3 months. Histological or cytological documentation of adenocarcinoma of the colon or rectum. All other histological types were excluded. At least two lines of prior treatment have failed Progression during or within 3 months following the last administration of approved standard therapies, which must have included fluoropyrimidine, oxaliplatin, irinotecan. Patients treated with oxaliplatin in an adjuvant setting must have progressed during or within 6 months of completion of adjuvant therapy. Patients who progressed more than 6 months after completion of oxaliplatin containing adjuvant treatment must have been retreated with oxaliplatin-based therapy to be eligible. Patients who had withdrawn from standard treatment due to unacceptable toxicity warranting discontinuation of treatment and precluding retreatment with the same agent prior to progression of disease were also allowed into the study. Patients may have received prior treatment with bevacizumab, and/or cetuximab/ panitumumab (if KRAS WT).

Trial number (acronym)	CORRECT: NCT01103323	CONCUR: NCT01584830
	ECOG PS ≤ 1. Adequate bone marrow, liver and renal function within 7 days of starting to study treatment. **Key exclusion criteria** Previous treatment with regorafenib. Uncontrolled hypertension, congestive heart failure ≥ NYHA Class 2, unstable angina, arterial or venous thrombotic or embolic events such as cerebrovascular accident, deep-vein thrombosis or pulmonary embolism within the 6 months before start of study medication. Ongoing infection > Grade 2 CTCAE Version 3.0	ECOG PS ≤ 1 Adequate bone marrow, liver and renal function within 7 days of starting to study treatment. **Key exclusion criteria** Previous treatment with regorafenib. Uncontrolled hypertension, congestive heart failure ≥ NYHA Class 2, unstable angina (angina symptoms at rest), new-onset angina (begun within the last 3 months) or myocardial infarction less than 6 months before start of study drug. Ongoing infection > Grade 2 CTCAE Version 3.0.
Settings and locations where the data were collected	114 centres across 16 countries (number of centres in brackets): Japan (19), US (17), Germany (15), Italy (9), France (9), Spain (8), Belgium (6), Australia (5), Israel (5), Canada (5), the Czech Republic (2), the Netherlands (2), China (1), Hungary (1), and Switzerland (1). No patients were randomised in South America or Turkey.	Asia (number of patients in brackets): China (129), Hong Kong (23), South Korea (20), Taiwan (20), and Vietnam (12)
Study periods and trial drugs	**Study periods** Screening: patient eligibility and enrolment Study drug treatment period: patients underwent evaluations for safety and drug accountability every cycle 30-day safety follow-up period: All patients entered the follow-up period upon discontinuation of either regorafenib or placebo treatment until death was documented **Trial drugs** Intervention (n=505): regorafenib 160 mg od po; 3 weeks on therapy followed by 1 week off therapy to comprise a cycle of 4 weeks plus BSC.	**Study periods** Screening: patient eligibility and enrolment Study drug treatment period: patients underwent evaluations for safety and drug accountability every cycle 30-day safety follow-up period: All patients entered the follow-up period upon discontinuation of either regorafenib or placebo treatment until death was documented **Trial drugs** Intervention (n=136): regorafenib 160 mg od orally (po). Three weeks on therapy followed by 1 week off therapy to comprise a cycle of 4 weeks plus BSC.

Trial number (acronym)	CORRECT: NCT01103323	CONCUR: NCT01584830
	Comparator (n = 255): placebo 160 mg po od; 3 weeks on, 1 week off, plus BSC.	Comparator (n=68): placebo 160 mg po od 3 weeks on, 1 week off, plus BSC.
Concomitant medication	**Permitted concomitant medication** Standard therapies for concurrent medical conditions. Prophylactic anti-emetics were permitted according to standard practice. Treatment with non-conventional therapies and vitamin/mineral supplements was acceptable provided that they did not interfere with the study endpoints, in the opinion of the Investigator. St John's wort was not permitted. Bisphosphonates. Patients who were therapeutically treated with an agent such as warfarin or heparin were allowed to participate provided that their medication dose and INR/ PTT were stable. **Non-permissible concomitant medications and procedures** Systemic anti-cancer therapy including cytotoxic therapy, signal transduction inhibitors, immunotherapy and hormonal therapy. TKIs Bone marrow transplant or stem cell rescue. Concomitant palliative radiation therapy was allowed if the target lesion(s) were not included within the radiation field and no more than 10% of the bone marrow was irradiated Use of biological response modifiers, such as G-CSF within 3 weeks of study entry. G-CSF and other haematopoietic growth factors were permitted during the study in the management of acute toxicity such as febrile neutropenia when clinically indicated or at the discretion of the Investigator. However, they could not be substituted for a	**Permitted concomitant medication** Standard therapies for concurrent medical conditions. Prophylactic anti-emetics were permitted according to standard practice. Treatment with non-conventional therapies and vitamin/mineral supplements was acceptable provided that they did not interfere with the study endpoints, in the opinion of the Investigator. St John's wort was not permitted. Bisphosphonates. Patients who were therapeutically treated with an agent such as warfarin or heparin were allowed to participate provided that their medication dose and INR/PTT were stable. **Non-permissible concomitant medications and procedures** Systemic anti-cancer therapy including cytotoxic therapy, signal transduction inhibitors, immunotherapy and hormonal therapy. TKIs. Bone marrow transplant or stem cell rescue. Concomitant palliative radiation therapy was allowed if the target lesion(s) were not included within the radiation field and no more than 10% of the bone marrow was irradiated Use of biological response modifiers, such as G-CSF, within 3 weeks of study entry. G-CSF and other haematopoietic growth factors were permitted during the study in the management of acute toxicity such as febrile neutropenia when clinically indicated or at the discretion of the Investigator. However, they could not be substituted for a

Trial number (acronym)	CORRECT: NCT01103323	CONCUR: NCT01584830
	required dose reduction. Patients taking chronic erythropoietin were permitted. Patients taking narrow therapeutic index medications were to be monitored proactively. All traditional medicines with an anti-cancer indication, including traditional Chinese medicine.	required dose reduction. Patients taking chronic erythropoietin were permitted. Patients taking narrow therapeutic index medications were to be monitored proactively. All traditional medicines with an anti-cancer indication, including traditional Chinese medicine.
Primary outcomes	**OS**: defined as the time (days) from randomisation to death due to any cause. Patients alive at the time of analysis were censored at the last date known to be alive. If a patient was lost to follow-up and there was no contact after randomisation, this patient was censored at Day 1.	**Overall survival (OS)**: defined as the time (days) from randomisation to death due to any cause. Patients alive at the time of analysis were censored at the last date known to be alive. If a patient was lost to follow-up and there was no contact after randomisation, this patient was censored at Day 1.
Other outcomes used in the economic model/specified in the scope	**Secondary efficacy endpoints** PFS: defined as time (days) from date of randomisation to date of first observed disease progression (radiological or clinical) or death due to any cause, if death occurred before progression was documented, based on investigator assessment using RECIST version 1.1. ORR: percentage of patients with complete response (CR) or partial response (PR) as best overall response based on investigator assessment. A best overall response was defined for all patients, using the RECIST criteria, version 1.1. DCR: percentage of patients whose best response was CR, PR or SD based on investigator assessment. Safety: type, frequency, and severity of adverse events (AEs)[b] **Tertiary efficacy endpoints** DOR: defined as time (days) from the first documented objective response of PR or CR, whichever was noted earlier,	**Secondary efficacy endpoints** PFS: defined as time (days) from date of randomization to date of first observed disease progression (radiological or clinical) or death due to any cause, if death occurred before progression was documented, on investigator assessment using RECIST version 1.1. ORR: percentage of patients with CR or PR as best overall response based on investigator assessment. A best overall response was defined for all patients, using the RECIST criteria, version 1.1. DCR: percentage of patients whose best response was CR, PR or SD based on investigator assessment Safety: type, frequency, and severity of AEs[b] **Tertiary efficacy endpoints** DOR: defined as time (days) from the first documented objective response of PR or CR, whichever was noted earlier,

Trial number (acronym)	CORRECT: NCT01103323	CONCUR: NCT01584830
	to disease progression or death (if death occurred before progression). Duration of stable disease: time (days) from randomisation to date of disease progression or death (if death occurred before progression). This variable was only calculated for patients who failed to achieve a best response of PR or CR. PRO: HRQoL and health utility values were measured using the EORTC QLQ-C30 (global health status/QoL) and EQ-5D, respectively. For the EORTC QLQ-C30 (range 0–100), higher scores on the functioning scales and the global health status/QoL scale represent a higher level of functioning and better HRQoL. A change of ≥ 10 points on the EORTC QLQ-C30 scale is considered as clinically meaningful. For the EQ-5D, higher scores represent better health status. A change of 0.07 to 0.12 points on the EQ-5D index and a change of 7 to 12 points on the VAS are considered as clinically meaningful.	to disease progression or death (if death occurred before progression). Duration of stable disease: time (days) from randomisation to date of disease progression or death (if death occurred before progression). This variable was only calculated for patients who failed to achieve a best response of PR or CR. PRO: HRQoL and health utility values were measured using the EORTC QLQ-C30 and EQ-5D, respectively. For the EORTC QLQ-C30 (range 0–100), higher scores on the functioning scales and the global health status/QoL scale represent a higher level of functioning and better HRQoL. A change of ≥ 10 points on the EORTC QLQ-C30 scale is considered as clinically meaningful. For the EQ-5D, higher scores represent better health status. A change of 0.07 to 0.12 points on the EQ-5D index and a change of 7 to 12 points on the VAS are considered as clinically meaningful.
Pre-planned subgroups	**Subgroup analyses of OS and PFS** Demographic information such as race, sex and age group (<65, ≥65 years). Region: Region 1 (North America, Western Europe, Israel and Australia), Region 2 (Asia) and Region 3 (South America, Turkey, and Eastern Europe). Time from diagnosis of metastatic disease (≥ 18 months and < 18 months). Prior systemic anti-cancer therapies: Prior anti-VEGF therapy (yes/no) Prior systemic anti-cancer therapies: Prior anti-cancer drugs, categorised in the following groups: FOIB; FOIBE Number of prior treatment lines (≤3, >3)	**Subgroup analyses of OS and PFS** Demographic information such as race, sex and age group (<65, ≥65 years). Time from diagnosis of metastatic disease (≥ 18 months and < 18 months) Single organ metastasis or multiple organ metastasis Prior systemic anti-cancer therapies (targeted therapies – yes/no) Prior systemic anti-cancer therapies in the following four groups: Patients without any preceding targeted therapy (no anti-VEGF, no anti-EGFR therapy) Patients with prior anti-VEGF treatment but without prior anti-EGFR treatment

Trial number (acronym)	CORRECT: NCT01103323	CONCUR: NCT01584830
	Number of prior treatment lines for metastatic disease (≤3, >3) Historical *KRAS* mutation status Further important baseline cancer characteristics of primary site of tumour (e.g., ECOG PS: 0 and 1) **Subgroup analyses of safety** Age (years): < 65, ≥ 65 BMI (kg/m^2): < 25, 25 ≤ BMI, < 30, 30 ≤ BMI Sex Race Hepatic function at baseline: maximum of baseline AST and baseline ALT value ≤ 1.5 x ULN, 1.5 x ULN < maximum of baseline AST and baseline ALT value ≤ 3 x ULN, 3 x ULN < maximum of baseline AST and baseline ALT value Kidney function at baseline: normal/mildly impaired renal function (estimated glomerular filtration rate, eGFR ≥ 60 mL/min/1.73 m²) Moderately impaired renal function (eGFR) ECOG PS at baseline	Patients with prior anti-EGFR treatment but without prior anti-VEGF treatment Patients with prior anti-VEGF treatment AND with prior anti-EGFR treatment Number of prior treatment lines (≤ 3, > 3) Number of prior treatment lines for metastatic disease (≤ 3, > 3) *KRAS* mutation status ECOG PS (0 and 1) *BRAF* mutation status Region (China [mainland China, Hong Kong, and Taiwan] and Asia, other than China) **Subgroup analyses of safety** Age (years): < 65, ≥ 65 BMI (kg/m^2): < 25, 25 ≤ BMI, < 30, 30 ≤ BMI Sex Race Hepatic function at baseline: maximum of baseline AST and baseline ALT value ≤ 1.5 x ULN, 1.5 x ULN < maximum of baseline AST and baseline ALT value ≤ 3 x ULN, 3 x ULN < maximum of baseline AST and baseline ALT value Kidney function at baseline: normal/mildly impaired renal function (estimated glomerular filtration rate, eGFR ≥ 60 mL/min/1.73 m²) Moderately impaired renal function (eGFR) ECOG PS at baseline

Source: Table 6, CS[3]

AE = adverse event; ALT = alanine aminotransferase; AST = aspartate aminotransferase; BMI= body mass index; BRAF = mutation in the B-Raf proto-oncogene; BSC = best supportive care; CTCAE = Common Terminology Criteria for AEs; CR = complete response; CS = company submission; DCR = disease control rate; DOR = duration of response; ECOG PS = Eastern Cooperative Oncology Group Performance Status; eGFR = estimated glomerular filtration rate; EGFR= epidermal growth factor receptor; EORTC QLQ C30 = European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire; EQ-5D = European Quality of Life-5 Dimensions; FOIB

Trial number (acronym)	CORRECT: NCT01103323	CONCUR: NCT01584830
= fluoropyrimidine, oxaliplatin, irinotecan, bevacizumab; FOIBE = fluoropyrimidine, oxaliplatin, irinotecan, bevacizumab, anti-EGFR antibody; G-CSF = granulocyte-colony stimulating factor; HRQoL = heath-related quality of life; INR = international normalised ratio; KRAS = Kirsten rat sarcoma viral oncogene homologue; mCRC = metastatic colorectal cancer; NYHA = New York Heart Association; od = once a day; ORR = overall response rate; OS = overall survival; PFS = progression-free survival; po = per os (oral); PR = partial response; PRO = patient reported outcomes; PTT = partial thromboplastin time; QoL = quality of life; RECIST = Response Evaluation Criteria in Solid Tumours; SD = stable disease; TKI = tyrosine kinase inhibitor; ULN = upper limit of normal; US = United States; VAS = visual analogue scale; VEGF = vascular endothelial growth factor; WT = wild type **Notes:** [a] Due to the similarity of treatment guidelines in Europe, the US and Asia (especially China), the design of this study is similar to that of CORRECT: patients with mCRC were included if their disease progressed during or within 3 months following the last administration of approved standard therapies. However, since the analysis of the real-life-situation in Asia had revealed that not all patients had access to treatment with targeted therapies, the study protocol allowed the inclusion also of those patients who had not been pre-treated with bevacizumab and/or Erbitux® (cetuximab)/Vectibix® (panitumumab) even if these drugs were approved but not available at study entry. The other inclusion and exclusion criteria roughly mirrored those of CORRECT. The BSC 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 except other investigational anti-tumour agents or anti-neoplastic chemotherapies/hormonal therapies/immunotherapies. [b] The AEs included acute renal failure or severe proteinuria, interstitial lung disease, acute cardiac failure, clinically significant bleeding and severe skin infections (Stevens-Johnson Syndrome, erythema multiforme and toxic epidermal necrolysis) and acute liver failure.		

ERG comment:

No UK patients were included in CORRECT or CONCUR. In the clarification letter, the company were asked to discuss with objective evidence how the study data has relevance to UK clinical practice and provide the supporting evidence. In the clarification response the company refers to the subgroup data from the OS analysis in CORRECT to appropriately show that race was not an important factor influencing OS. They stated that "*as no interaction is observed for region or race the results of CORRECT are considered to be generalisable to UK patients*".[6] However, the company failed to refer to the subgroup data for PFS, which shows a point estimate difference between the subgroups indicating that Asian participants may have better PFS when on regorafenib. Although there is uncertainty in this result, it is unlikely that the subgroups were powered to detect differences, and so it is appropriate for the ERG to be vigilant for possible type II errors and to make the committee aware of them (see ERG comment in Section 3.2.5 for a fuller discussion of this issue).

3.2.1.2. Comparative observational studies

Nakashima 2020,[1] Tanaka 2018,[23] Huemer 2020[26] and Sueda 2016[25] were retrospective observational studies directly comparing regorafenib to T/T. Nakashima 2020[1] was a nationwide claims database study conducted in Japan, Tanaka 2018[23] and Sueda 2016[25] were both single-centre studies conducted in Japan, whereas Huemer 2020[26] was a multi-centre study conducted in Austria. Table 3.5 provides further detail of these studies, which includes information on baseline equivalence, prior drug regimens and statistical methodology.

Table 3.5: Details of the observational studies of regorafenib versus trifluridine/tipiracil

Study	Design and objective	Population (and prior treatments)	Treatments	Outcomes	Statistical analysis
Nakashima 2020[1]	Retrospective, comparative, observational, nationwide claims database study conducted in Japan; to evaluate the efficacy and safety of regorafenib and trifluridine/tipiracil (T/T) in patients with unresectable colorectal cancer.	Used the Japanese medical claims database maintained by Medical Data Vision Company. The database consists of data from 269 hospitals in different regions covering about 17% of Japanese acute care hospitals. Inclusion: aged 20 or over; patients who received regorafenib and/or T/T and included patients using regorafenib and/or T/T who were or were not diagnosed with metastasis. Exclusion: (1) regorafenib- or T/T - treated patients with diseases other than colorectal cancer, (2) patients who received other chemotherapy simultaneously with regorafenib or T/T. The two groups were similar with respect to most baseline variables. Gender: regorafenib 62% male, T/T 60% male; age: regorafenib 66 years, T/T 68 years; primary site of disease: regorafenib 59% colon, T/T 57% colon; number of metastatic sites ≥3: regorafenib 30%, T/T 28%; number of previous anti-cancer agents: regorafenib 4, T/T 4; previous systemic anticancer agents used by both arms	Regorafenib versus T/T. Doses not provided (presumably as these varied across patients, although no central tendency or variance measure is provided). Two sets of data were analysed: 1) regorafenib monotherapy versus T/T monotherapy, and 2) regorafenib first (whether followed by best supportive care (BSC) or by T/T) versus T/T first (whether followed by BSC or by regorafenib). Only the data for the first 'monotherapy'	Overall survival (OS) defined as the time from regorafenib or T/T initiation until death from any cause. Adverse events (AEs).	OS was compared using the Cox proportional hazard model and log-rank test for univariate analysis. OS was also presented using the Kaplan-Meier (KM) curve. Although not reported in the methodology section, it appears from the results section that the hazard ratio (HR) has been adjusted for confounding but the methods and variables for which adjustment has been made are not reported. A propensity score analysis was undertaken but this was only performed for the regorafenib first and T/T first analysis, not the relevant regorafenib monotherapy and T/T monotherapy analysis. Therefore, the relevant monotherapy analysis has *not* been subject to propensity score

Study	Design and objective	Population (and prior treatments)	Treatments	Outcomes	Statistical analysis
		were fluorouracil (regorafenib: 61%/ T/T 61%), capecitabine (31%/39%), tegafur/gimeracil/oteracil (28%/33%), tegafur/uracil (7%/7%), oxaliplatin (66%/72%), irinotecan (76%/76%), bevacizumab (72%/74%), cetuximab (15%/11%), Panitumumab (27%/26%), afibercept (1%/1%), ranibizumab (7%/7%).	analysis are used for this report. This is because outcomes in the regorafenib first or T/T first groups would be affected by *both* drugs and therefore prohibit a meaningful comparison of the efficacy of the drugs *per se.*		matching but does appear to have been statistically adjusted for confounding.
Tanaka 2018[23]	Retrospective, comparative, observational, single-centre study conducted in Japan; to evaluate the efficacy and safety of regorafenib and T/T in patients with refractory metastatic colorectal cancer (mCRC).	Patients with unresectable mCRC treated at Tokai University Hospital Japan between 2013-2015. N=44; 20 patients in regorafenib group; sex (male): regorafenib 65%, T/T 62.5%; age (years): regorafenib 68, T/T 64 Inclusion: histological confirmation of adenocarcinoma of the colon or rectum, alongside existence of unresectable metastatic disease; previous treatments with fluoropyrimidine, irinotecan, oxaliplatin, and anti-vascular endothelial growth factor (VEGF) antibody (bevacizumab), or anti-epidermal growth factor receptor (EGFR) antibody (cetuximab or panitumumab) for patients who had	Regorafenib (160 mg) once daily on days 1-21, with 7 days of rest. TAS-102 (35 mg/m^2) twice daily 5 days a week, with 2 days of rest, for 2 weeks, followed by a 14-day rest period. Both treatment regimens repeated every 4 weeks. Treatments continued until progression of	All patients had computed tomography (CT) every 8 weeks to assess tumour responses to therapy according to the Response Evaluation Criteria in Solid Tumours (RECIST) version 1.1. Progression-free survival (PFS) was defined as the interval from the first administration of the primary treatment to the first radiologic or clinical observation of	Data compared using a log-rank test with 95% confidence intervals (95% CIs). The results of OS were plotted against the total delivered dose for each drug and fitted to a simple linear regression model to calculate the regression coefficient. A Cox proportional hazards regression model was used to test each candidate variable predictor associated with OS using stepwise model selection according to

Study	Design and objective	Population (and prior treatments)	Treatments	Outcomes	Statistical analysis
		Kirsten rat sarcoma viral oncogene homologue (KRAS) exon 2 wild type (WT) tumour; Eastern Cooperative Oncology Group Performance Status (ECOG PS) of 0 to 2; adequate bone-marrow, liver, and renal function at onset of treatment. Exclusion: previous treatment with regorafenib or TAS-102, or uncontrolled medical disorders Patients had 2–4 prior regimens: fluoropyrimidine: 100% (both arms) oxaliplatin: 100% (both arms) irinotecan: 100% (both arms) anti-VEGF: 100% (regorafenib); 95.8% (T/T) anti-EGFR antibody (wild KRAS or all-RASa): 45.0% (regorafenib); 45.8% (T/T). Similar across arms for some baseline variables, but differed for ECOG 0 status: regorafenib 30%, T/T 58.3%; primary site of disease (right colon): regorafenib 20%, T/T 41.7%; KRAS exon 2 status (wild): regorafenib 45%, T/T 58.3%; number of metastatic sites ≥3: regorafenib 10%, T/T 29.2%.	disease, death, toxicity, withdrawal of consent, or decision by the treating physician. Patients whose starting dose was reduced at the discretion of the treating physician were retained. Patients with dose reductions could re-increase the dose up to the recommended starting dose if toxicity resolved. All patients had the BSC available. Other antitumor agents, hormonal therapy, or immunotherapy were disallowed. Reasons for chosen drug allocation not given.	disease progression or death from any cause, whichever came first. OS was defined as the time between the administration date of the primary treatment and the date of death from any cause, and the median PFS and OS were estimated using the KM method. AEs were classified and graded according to the National Cancer Institute Common Terminology Criteria for Adverse Events, Version 4.03	Akaike's information criterion. HRs were calculated by taking the exponentials of the ß coefficients of Cox models. Model discrimination was done by calculating the Harrell's C (for concordance) index, which is the area under the receiver operator curve. HRs of covariates were rounded to the nearest integer to construct score weights. The range of possible total score weights was divided into three groups to stratify patients into poor-, intermediate- and long-survival tertiles. No statistical adjustment was carried out for evaluation of the main outcomes in relation to treatment.

Study	Design and objective	Population (and prior treatments)	Treatments	Outcomes	Statistical analysis
			A cross-over of treatments was also used (secondary treatment) but the results of this are not provided in this report. Only primary treatments (from initial stage) are presented.		
Sueda 2016[25]	Retrospective, observational, single-centre study conducted in Japan; to evaluate the efficacy and safety of regorafenib and T/T in patients with mCRC refractory to standard chemotherapies.	Patients with histologically-confirmed, unresectable mCRC; N=37; 23 patients in the regorafenib group. Sex (male): regorafenib 52.2%, T/T 71.4%; age (years): regorafenib 59, T/T 66. Inclusion: patients with histologically-confirmed, unresectable, mCRC; received ≥2 prior regimens of standard chemotherapy; ECOG PS of 0 to 2; adequate organ function. Prior therapies: anti-EGFR: regorafenib 52.2%, T/T 64.3% anti-VEGF: regorafenib 100%, T/T 71.4% Differed across arms for all baseline variables.	Regorafenib given at 160 mg once daily on days 1-21 of every 28-day cycle. T/T given at 35 mg/m^2 orally, twice daily, on days 1-5 and 8-12 of every 28-day cycle. Treatment continued until patients had confirmed disease progression, toxicity, withdrew consent, or stopped treatment at the	AEs, treatment response PFS and OS were collected from medical records. The tumour response rate was the proportion of patients with complete response (CR) or partial response (PR), and the disease control rate (DCR) was the proportion of patients with a best response of CR or PR or stable disease (SD). Tumour response and progression were radiologically assessed by	OS was the duration of time from the start of regorafenib or T/T to death from any cause. Survival curves were generated using the KM method. Differences in survival were evaluated with the log-rank test.

Study	Design and objective	Population (and prior treatments)	Treatments	Outcomes	Statistical analysis
		Age: regorafenib 59 years, T/T 66 years; sex (male) regorafenib 52.2%, T/T 71.4%; ECOG 0 status: regorafenib 43.5%, T/T 7.2%; primary site of disease (colon): regorafenib 56.5%, T/T 71.4%; KRAS exon 2 status (wild): regorafenib 52.2%, T/T 64.3%; number of people with metastatic sites in lung: regorafenib 52%, T/T 64%; number of people with metastatic sites in lymph nodes: regorafenib 13%, T/T 50%; number of people with metastatic sites in peritoneum: regorafenib 9%, T/T 14%; number of people with metastatic sites in bone: regorafenib 13%, T/T 29%.	investigator's discretion.	investigators with the RECIST criteria (version 1.1). Patients underwent safety assessments during every cycle, according to the National Cancer Institute Common Terminology Criteria for Adverse Events version 4.0.	
Huemer 2020[26]	Retrospective, observational, multi-centre study conducted in Austria; to investigate hospitalizations during regorafenib or T/T therapy, as well as the impact of hospitalisations on clinical outcome in mCRC beyond second-line therapy.	mCRC patients starting systemic third line therapy with regorafenib or T/T at the tertiary cancer centres in Salzburg and Wels-Grieskirchen (Austria) between 2013-2019. N=93; 69 patients in the regorafenib group. Prior disease progression on fluorouracil, oxaliplatin, irinotecan, anti-VEGF and/or anti-EGFR (in case of RAS wild type status) targeted therapy was a prerequisite for the initiation of regorafenib and/or T/T.	Regorafenib was given at an oral daily dose of 160 mg for the first three weeks of each four-week cycle, or at a starting dose of 80 mg per day with weekly dose increase to a target dose of 160 mg. T/T was orally given twice daily	The only outcome relevant to the National Institute for Health and Care Excellence (NICE) remit was OS.	Multivariate analysis was based on a Fine–Gray proportional sub-distribution hazards regression model. For multivariate analysis covariate selection, a backward stepwise regression for competing risks regression was performed. OS was calculated from the start of third-line treatment until the date of death or date of last known

Study	Design and objective	Population (and prior treatments)	Treatments	Outcomes	Statistical analysis
		Differed across arms for some but not all baseline variables. Age: regorafenib 65 years, T/T 68 years; sex (male) regorafenib 58%, T/T 58%; ECOG 0 status: regorafenib 40%, T/T 21%; Ascites: regorafenib 6%, T/T 17%; RAS status (wild type): regorafenib 49%, T/T 50%; mutation in the B-Raf proto-oncogene (BRAF) status (wild type): regorafenib 100%, T/T 95%; microsatellite status (MSS): regorafenib 98%, T/T 93%; metastatic pattern (liver): regorafenib 81%, T/T 67%; metastatic pattern (lung): regorafenib 71%, T/T 63%; metastatic pattern (peritoneum): regorafenib 14%, T/T 25%.	at a dose of 35 mg/m^2 five days a week, with 2 days of rest, for 2 weeks, followed by a 14-day rest period, and repeated every four weeks.		follow-up. Patients alive at the last contact were censored. Survival curves were estimated by the KM method. For survival analysis, median follow-up was calculated from initiation of third-line treatment with either regorafenib or TAS-102 using the KM estimator with reversed status indicators (death and censored). For survival analysis according to treatment groups, adjusted survival curves using Cox proportional hazards models were created. The likelihood-ratio test was used to compare survival distributions between patient groups. In order to avoid immortal time bias, cross-over was taken into account as time-dependent covariate beginning with the start of fourth-line treatment.

Study	Design and objective	Population (and prior treatments)	Treatments	Outcomes	Statistical analysis
Based on the original sources: Nakashima 2020,[1] Tanaka 2018,[23] Sueda 2016,[25] Huemer 2020,[26] and for the latter three studies, Table 17, CS [3] AEs = adverse events; BRAF = Mutation in the B-Raf proto-oncogene; CI = confidence interval; CR =complete response; CS = company submission; CT = Computerised Tomography; DCR = disease control rate; ECOG PS = Eastern Cooperative Oncology Group Performance Status; EGFR = epidermal growth factor receptor; HR = hazard ratio; KM = Kaplan-Meier; KRAS = Kirsten rat sarcoma viral oncogene homologue; mCRC = metastatic colorectal cancer; NICE = National Institute for Health and Care Excellence; OS = overall survival; PFS = progression-free survival; PR = partial response; RAS = ras proteins are proto-oncogenes that are frequently mutated in human cancers; RECIST = Response Evaluation Criteria in Solid Tumours; SD = stable disease; T/T = trifluridine/tipiracil**;** TAS-102 = trifluridine/tipiracil; VEGF = vascular endothelial growth factor; WT = wild type					

ERG comment:

Nakashima 2020[1] did not exclude those without metastases, but it appears that most participants had metastases. In the regorafenib and T/T groups there were 64% and 62% respectively with liver metastases alone, 46% and 48% with lung metastases alone, 23% and 23% with lymph nodes metastases alone, 29% and 30% with peritoneum metastases alone, 16% and 15% with bone metastases alone, 6% and 5% with brain metastases alone and 10% and 11% with other metastases alone. Even allowing for considerable overlap between different metastases categories, it is highly likely that the number without any metastases was small (although the amount of overlap will have been limited because only 30% and 28% of participants had three or more metastatic sites). It is also the case that both regorafenib and T/T are licensed only for mCRC. Therefore, this study has been retained in the report. In three of the four observational studies there were baseline variables that, as expected, differed across arms. For example, in Tanaka 2018[23], the arms differed for the proportion of people with a baseline ECOG score of zero, with 30% in the regorafenib arm and 58% in the T/T arm. This might confer an advantage to the T/T arm. In contrast, the arms differed for the proportion of people with three or more metastatic sites, with 10% in the regorafenib arm but 29.2% in the T/T arm, so this might confer an advantage to the regorafenib arm. In Sueda 2016 [25] there were differences in most baseline variables, and most of these could be argued to favour the regorafenib group, with the regorafenib arm having lower rates of metastases, lower age and a higher number of people with an ECOG score of zero. In Huemer 2020[26] there were some differences between arms, particularly in ECOG status and Ascites, with these differences again being more likely to favour the regorafenib group. In neither Tanaka 2018[23] nor Sueda 2016[25] were outcomes relevant to the NICE scope adjusted for any confounding, and so any bias relating to such baseline inequivalence may affect estimates. In Huemer 2020[26], the only reported outcome relevant to the NICE scope – PFS – appears to have been adjusted for confounding from 'immortal time bias', but it does not appear to have been specifically adjusted for the differing baseline variables. In Nakashima 2020[1] there appeared to be far better comparability between arms, despite no propensity matching in the analysis pertinent to this report. Nevertheless, because the allocation to groups was non-random it is likely that there exist differences in non-measured covariates that may have impaired internal validity. Therefore, it is likely that all of the four studies had impaired internal validity. No UK patients were included in the four non-randomised studies, also reducing the external validity of these data, although this also applies to the CORRECT and CONCUR studies.

3.2.2 Statistical analyses of the included randomised trials

The statistical analyses used for the main analyses in the CS[3], alongside the sample size calculations and methods for handling missing data are presented in Table 3.6.

Table 3.6: Summary of statistical analyses for the primary analysis in the CORRECT and CONCUR trials

Trial	CORRECT	CONCUR
Hypothesis objective	To compare overall survival (OS) between the regorafenib group and placebo group, the following hypothesis was tested: H_0: hazard ratio (HR) (regorafenib/placebo) ≥1 versus H_1: HR (regorafenib/placebo) <1	To compare OS between the regorafenib group and placebo group, the following hypothesis was tested: H_0: HR (regorafenib/placebo) ≥1 versus H_1: HR (regorafenib/placebo) <1

Trial	CORRECT	CONCUR
Statistical analysis	**Main analyses** OS and progression-free survival (PFS) were compared between treatment groups with a stratified log-rank test; HRs (with 95% confidence intervals (CIs)) were calculated with the Cox model, adjusting for stratification factors; and Kaplan-Meier (KM) survival estimates were calculated for each treatment group. Overall response rate (ORR) and disease control rate (DCR) were compared between treatment groups with the Cochran–Mantel–Haenszel test, adjusting for stratification factors. **Subgroups** Forest plots, descriptive statistics and HR estimates with 95% CIs for OS and PFS were presented for predefined subgroups (provided there was a sufficient number of events in total within the subgroup across the treatment arms). Summaries of adverse events (AEs) were presented according to Common Terminology Criteria for Adverse Events (CTCAE) version 3.0 and Medical Dictionary for Regulatory Activities (MedDRA). **Sensitivity analyses** OS and PFS were tested with an unstratified log-rank test. Two sensitivity analyses of OS and PFS were performed: one on unstratified data and one using stratification information from the Interactive Voice Response System (IVRS). In the sensitivity analyses of PFS, all available tumour assessment data were taken into account also from the follow-up period.	**Main analyses** OS and PFS were compared between treatment groups with a stratified log-rank test; HRs (with 95% CIs) were calculated with the Cox model, adjusting for stratification factors; and KM survival estimates were calculated for each treatment group. ORR and DCR were compared between treatment groups with the Cochran–Mantel–Haenszel test, adjusting for stratification factors. **Subgroups** Forest plots, descriptive statistics and HR estimates with 95% CIs for OS and PFS were presented for predefined subgroups. Summaries of AEs were presented according to CTCAE version 4.0 and MedDRA. **Sensitivity analyses** OS and PFS were tested with an unstratified log-rank test. Three pre-specified sensitivity analyses of OS were performed: an unstratified analysis of OS, an analysis using stratification information from the IVRS, and an analysis stratified by previous targeted anti-cancer therapy (targeted therapy defined as anti-vascular endothelial growth factor (VEGF) or anti-epidermal growth factor receptor (EGFR) therapy or both). Two additional analyses of PFS were performed using a definition of PFS that included all assessments from follow-up and one that considered a new treatment initiation date in follow-up as the event date.
The Analysis sets	**Intention-to-treat (ITT) analysis set:** all randomised patients including those who withdrew regardless of the reason for withdrawal. This was the primary population for all efficacy analyses.	**Full analysis set (FAS):** all randomised patients. This set was the primary population for the efficacy analyses.
	Safety analysis set (SAS): all patients who received at least one dose of study medication.	**SAS:** all patients who received at least one dose of study medication.
	Pharmacokinetic (PK) analysis set: all patients with available PK data collected after at least 14 days of	**PK analysis set:** all patients with available PK data collected after at least 14 days of uninterrupted stable dosing of regorafenib.

Trial	CORRECT	CONCUR
	uninterrupted stable dosing of regorafenib.	
	Biomarker analysis set: all patients with available biomarker data and signed consent form for the analyses.	**Biomarker analysis set:** all patients with available biomarker data and signed consent form for the analyses.
Sample size, power calculation	**Sample size and power calculation** CORRECT was designed to have 90% power to detect a 33.3% increase in median OS, assuming a 4.5-month median OS for the placebo group (i.e., an HR of 0.75 for regorafenib over placebo). Assuming a one-sided overall α of 0.025, a power of 90%, a randomization ratio of 2:1 between regorafenib and placebo, and two formal interim analyses of OS during the study, with an O'Brien–Fleming-type error spending function, the study required 582 deaths for the final analysis. A total of 690 patients were planned for randomisation. **Interim analyses** Two formal interim analyses were planned during the study when approximately 30% (first interim) and 70% (second interim) of the planned total number of required death events had occurred. The first formal interim analysis was for futility only. The second interim analysis was for efficacy and futility. A Lan–Demets alpha spending function determined the monitoring boundary for efficacy, so the overall false positive rate (α) was ≤ 0.025 (one-sided). The alpha spending function was the O'Brien–Fleming type boundary specified. Boundaries were specified to stop the study for efficacy or futility on the basis of the actual number of events included in the analysis. At the second interim analysis, the study was to be stopped for futility if the HR (regorafenib over placebo) was 0.9006 or greater, and for efficacy if the one-sided p value was less than or equal to 0.009279, roughly corresponding to an HR (regorafenib over placebo) of less than or equal to 0.7864.	**Sample size and power calculation** The sample size was based on the primary endpoint of OS. A total of 200 patients and 154 death events were required, assuming a target increase in median OS of 33.3% (i.e., an HR of 0.75, regorafenib over placebo), one-sided overall α of 0.2, a power of 80% and a randomization ratio of 2:1 between regorafenib and placebo. It was projected that 154 events would occur after approximately 19 months, assuming a monthly patient enrolment rate of 33 patients/month and 200 patients were randomized after an initial 6 months ramp-up period, a dropout rate of 3%, exponentially distributed event times for OS, and 4.5 and 6 month median OS time for the placebo and the regorafenib groups, respectively.
Data management,	**Censoring methods**	**Censoring methods**

Trial	CORRECT	CONCUR
patient withdrawals	For the primary endpoint of OS**,** patients alive at the time of analysis were censored at the last date they were known to be alive. If a patient was lost to follow-up and there was no contact after randomisation, this patient was censored at day 1. Patients with evidence of being alive as of the database cut-off date were censored using the cut-off date of 21 July 2011. Standard censoring methods were applied to PFS and response (ORR, DOR, duration of stable disease) analyses for those patients without (or missing) evaluable assessments. **Missing data** Patients withdrawn from study treatment were not replaced, and missing data were not estimated or carried forward in any statistical analysis (unless otherwise stated). No imputation was performed for missing assessments.	For the primary endpoint of OS**,** patients alive at the time of analysis were censored at the last date they were known to be alive. Patients with evidence of being alive as of the database cut-off date were censored using the cut-off date of 29 November 2013. Standard censoring methods were applied to PFS and response (ORR, duration of response (DOR), duration of stable disease) analyses for those patients without (or missing) evaluable assessments. **Missing data** Patients withdrawn from study treatment were not replaced, and missing data were not estimated or carried forward in any statistical analysis (unless otherwise stated). No imputation was performed for missing assessments.

Source: Table 9 of CS[3]
AE = adverse event; CS = company submission; H_0 = null hypothesis; H_1 = alternative hypothesis; CI = confidence interval; CTCAE = Common Terminology Criteria for AEs; DCR = disease control rate; DOR = duration of response; EGFR = Epidermal growth factor receptor; FAS = full analysis set; HR = hazard ratio; ITT = intention to treat; IVRS = Interactive Voice Response System; KM = Kaplan–Meier; MedDRA = Medical Dictionary for Regulatory Activities; ORR, = overall response rate; OS = overall survival; PFS = progression-free survival; PK = pharmacokinetic; SAS = safety analysis set; VEGF = vascular endothelial growth factor

ERG comment:

An intention-to-treat (ITT) analysis was not reported for CONCUR in the CS, but perusal of the original source confirms that an ITT analysis was indeed used. This clarifies that the risk of attrition bias in the CONCUR study was reduced.

3.2.3 Baseline characteristics

3.2.3.1 Randomised trials

A summary of the baseline characteristics of patients in the CORRECT and CONCUR trials are presented in Tables 3.7 and 3.8 respectively.

Table 3.7: Baseline characteristics of patients in the CORRECT trial

Characteristic	Regorafenib (n=505)	Placebo (n=255)
Median age, years (IQR)	61 (54.0–67.0)	61 (54.0–68.0)
Sex, n (%)		
Men	311 (62)	153 (60)
Women	194 (38)	102 (40)

Characteristic	Regorafenib (n=505)	Placebo (n=255)
Race, n (%)		
White	392 (78)	201 (79)
Black	6 (1)	8 (3)
Asian	76 (15)	35 (14)
Other/ not specified	31 (6)	11 (4)
Region, n (%)		
North America, western Europe, Israel, Australia	420 (83)	212 (83)
Asia	69 (14)	35 (14)
Eastern Europe	16 (3)	8 (3)
ECOG PS , n (%)		
0	265 (52)	146 (57)
1	240 (48)	109 (43)
Primary site of disease, n (%)[a]		
Colon	323 (64)	172 (68)
Rectum	151 (30)	69 (27)
Colon and rectum	30 (6)	14 (5)
***KRAS* mutation, n (%)[b]**		
No	205 (41)	94 (37)
Yes	273 (54)	157 (62)
Unknown	27 (5)	4 (2)
***BRAF* mutation, n (%)[c]**		
No	322 (96)	163 (98)
Yes	14 (4)	3 (2)
Histology, n (%)		
Adenocarcinoma	493 (98)	245 (96)
Adenocarcinoma in situ	2 (< 1)	3 (1)
Adenosquamous carcinoma	1 (< 1)	1 (< 1)
Carcinoma, not otherwise specified	4 (1)	1 (< 1)
Mucinous carcinoma	5 (1)	4 (2)
Undifferentiated carcinoma	0 (0)	1 (< 1)
Number of previous systemic anti-cancer therapies (on or after diagnosis of metastatic disease), n (%)		
1–2	135 (27)	63 (25)
3	125 (25)	72 (28)
≥ 4	245 (49)	120 (47)
Previous anti-VEGF treatment, n (%)		
Bevacizumab	505 (100)	255 (100)
Patients stopping previous treatment because of progression, n (%)		
Fluoropyrimidine	421 (83)	221 (87)

Characteristic	Regorafenib (n=505)	Placebo (n=255)
Bevacizumab	403 (80)	214 (84)
Irinotecan	405 (80)	229 (90)
Oxaliplatin	278 (55)	160 (63)
Panitumumab or cetuximab, or both	219 (43)	107 (42)
Time from diagnosis of metastases		
Median, months (IQR)	31.0 (20.6–43.3)	29.9 (20.2–46.4)
< 18 months, n (%)	91 (18)	49 (19)
≥ 18 months, n (%)	414 (82)	206 (81)

Source: Table 7 in CS[3]
BRAF = Mutation in the B-Raf proto-oncogene; CS= company submission; ECOG PS = Eastern Cooperative Oncology Group Performance Status; IQR = interquartile range; ITT = intention to treat; KRAS = Kirsten rat sarcoma viral oncogene homologue; VEGF = vascular endothelial growth factor

Notes:
[a] Information missing in one patient in the regorafenib group
[b] *KRAS* mutation status was based on historical patient record
[c] *BRAF* mutation status was determined with plasma DNA samples collected from 336 patients in the regorafenib group and 166 in the placebo group
[d] Five patients in the placebo group (2%) and 16 patients in the regorafenib group (3%) had received only one previous line of treatment for metastatic disease

Table 3.8: Baseline characteristics of patients in the CONCUR trial

Characteristic, n (%)	Regorafenib (n=136)	Placebo (n=68)
Age		
Median age, years (IQR)	57.5 (50.0–66.0)	55.5 (48.5–62.0)
< 65, n (%)	95 (70)	58 (85)
≥ 65, n (%)	41 (30)	10 (15)
Sex, n (%)		
Men	85 (63)	33 (49)
Women	51 (38)	35 (51)
Region, n (%)		
China (mainland China, Taiwan, and Hong Kong)	112 (82)	60 (88)
Asia other than China	24 (18)	8 (12)
BMI, kg/m^2 (IQR)	**23.1 (20.8–25.5)**	**22.8 (20.0–25.0)**
ECOG PS, n (%)		
0	35 (26)	15 (22)
1	101 (74)	53 (78)
Main site of disease, n (%)		
Colon	79 (58)	48 (71)
Rectum	53 (39)	19 (28)
Colon and rectum	4 (3)	1 (1)
***KRAS* mutation, n (%)**		
No	50 (37)	29 (43)

Characteristic, n (%)	Regorafenib (n=136)	Placebo (n=68)
Yes	46 (34)	18 (26)
Unknown	40 (29)	21 (31)
***BRAF* mutation, n (%)**		
No	28 (21)	14 (21)
Yes	0 (0)	1 (1)
Unknown	108 (79)	53 (78)
Histology, n (%)		
Adenocarcinoma	130 (96%)	66 (97)
Mucinous carcinoma	6 (4)	2 (3)
Time from diagnosis of metastatic disease		
Median, months (IQR)	20.3 (13.8–28.8)	19.9 (13.3–27.7)
< 18 months, n (%)	53 (39)	32 (47)
≥ 18 months, n (%)	83 (61)	36 (53)
Number of metastatic sites, n (%)		
Single	28 (21)	15 (22)
Multiple	108 (79)	53 (78)
Previous targeted biological treatment, n (%)		
None	56 (41)	26 (38)
Any (anti-VEGF[a] or anti-EGFR[b], or both)	80 (59)	42 (62)
Anti-VEGF but not anti-EGFR	32 (24)	13 (19)
Anti-EGFR but not anti-VEGF	24 (18)	17 (25)
Anti-VEGF and anti-EGFR	24 (18)	12 (18)
Previous systemic anti-cancer treatment lines, n (%)		
Any intention		
2	31 (23)	14 (21)
3	32 (24)	19 (28)
≥ 4	73 (54)	35 (51)
On or after diagnosis of metastatic disease[c]		
1–2	48 (35)	24 (35)
3	32 (24)	17 (25)
≥ 4	52 (38)	27 (40)

Source: Table 8 in CS[3]
BMI = body mass index; BRAF = Mutation in the B-Raf proto-oncogene; CS = company submission; ECOG PS = Eastern Cooperative Oncology Group Performance Status; EGFR = epidermal growth factor receptor; IQR = interquartile range; ITT = intention-to-treat; KRAS = Kirsten rat sarcoma viral oncogene homologue; VEGF = vascular endothelial growth factor

Notes:
[a] bevacizumab
[b] cetuximab or panitumumab
[c] four patients (3%) in the regorafenib group had not previously received any treatment for metastatic disease

ERG comment:

In terms of internal validity, although in general there was reasonable comparability between treatment groups within the two regorafenib versus placebo trials, there were some isolated differences in certain characteristics that could create bias. In the CORRECT trial, there were group differences in ECOG status, with a greater proportion of the regorafenib group having ECOG status 1. In the CONCUR study there were differences in age and gender between treatment groups, with a larger proportion of the regorafenib group being ≥65 years and male. However, the company's subgroup analyses show that these differences would have created a bias that favoured the placebo group, rather than the regorafenib group, implying that the observed measure of effect may have been deflated rather than inflated by these differences. In summary, the ERG is not concerned by these baseline differences in terms of internal validity as they have probably weakened the magnitude of observed results in favour of the regorafenib; in other words, had these baseline differences not occurred, an even greater benefit for the regorafenib might have been observed.

In terms of external validity (comparability between trials and with NHS clinical practice), the CORRECT and CONCUR trails obtained very different results, with better results in favour of regorafenib for the CONCUR trial in both OS and PFS (see Section 3.2.5). There are three main differences between cohorts that could explain this:

1) Prior anti-VEGF treatment in the form of bevacizumab
2) Number of prior treatments
3) Race

The results of relevant subgroup analyses are reported in Section 3.2.5.6.

1) The anti-VEGF medication bevacizumab is not recommended by NICE for patients with mCRC, but it appears that the entire population in CORRECT and 39.7% of the patients in CONCUR were previously treated with an anti-VEGF. The company had stated that *"Prior treatment with anti-VEGF is relevant as regorafenib has anti-VEGF activity - the implication of this prior therapy is that regorafenib could be expected to be less effective in patients who have already been treated with, and failed on, an anti-VEGF"* page 53, CS[3]. In the clarification response the company confirmed this. On being asked on how previous treatment with bevacizumab might affect generalisability to UK clinical practice, they responded by stating that: *"The implication could be better efficacy in UK clinical practice (compared to the results from CORRECT) as patients won't have received prior treatment with bevacizumab."* [6] However, the ERG notes that apparently strong, albeit counterintuitive, effects were observed for previous targeted treatment in OS, which has implications for the applicability of regorafenib. The subgroup analysis results for both OS and PFS did not present patients previously treated with an anti-VEGF treatment, and instead provided results on five subgroups involving combinations of anti-VEGF with an anti-EGFR, which produce what appear to be some counterintuitive results. No targeted treatment gave the lowest HR (most effective) for both OS and PFS, which makes sense. However, for OS, although previous anti-VEGF and no previous anti-EGFR gives the highest HR (least effective), no previous anti-VEGF and previous anti-EGFR gives the second highest HR. For PFS, the highest HR is produced by no previous anti-VEGF and previous anti-EGFR and there is no overlap in the 95% CIs with no previous treatment. The company was asked to explain the apparent inconsistency of these results, and responded as follows, helpfully providing a new subgroup analysis that categorised patients into those that had received anti-VEGF versus those who had not, regardless of anti-EGFR experience: "*In respect of prior anti-VEGF treatment versus no prior treatment - the OS hazard ratio for patients who had not received anti-VEGF was lower (HR: 0.470; 95% CI: 0.309, 0.714)*

compared to patients who received anti-VEGF (HR: 0.726; 95% CI: 0.430, 1.224). These results are supportive of a greater potential to benefit in patients who have not received prior anti-VEGF treatment." [6]

2) Patients in CONCUR generally receive fewer lines of therapy from diagnosis of metastatic disease: 38% to 40% versus 47% to 49% received at least four (>3) prior lines in CONCUR versus CORRECT, depending on arm. Subgroup analyses show that within each cohort, there is also an association observed between number of prior treatments and OS and PFS: the point estimate for the HR is lower for >3 than for ≤3 prior lines for metastatic disease. This is therefore counter to what would be expected if this were an explanation for the greater effectiveness of regorafenib in the CONCUR trial, although this does not provide consistent evidence of a convincing effect modifier given the overlap in the 95% CIs. On this matter in the company stated in response to the clarification letter: "*Although the direction of diminishing ability to benefit from successive lines of treatment is clinically accepted, the effect observed in CORRECT and CONCUR is counter to expectations (as pointed out by the EAG). The same counterintuitive result was also observed in the RECOURSE trial for trifluridine/tipiracil (Mayer 2015 – figure B). However, the differences are non-significant and confidence intervals relatively wide. The counterintuitive results cannot be explained. However, there is a risk in overinterpreting point estimates of subgroup results from trials that are powered at the overall population level. The point estimates from these subgroup analyses should be viewed in the context of their confidence intervals. The results observed in CONCUR should be taken as confirmation of the treatment benefit for regorafenib which was observed in CORRECT. The benefit was greater in the CONCUR trial but there is no clear explanation for this. Similarly, the trials for trifluridine/tipiracil show a difference in OS benefit (RECOURSE 0.68, Yoshino 0.56, TERRA 0.79) without a clear explanation for the difference. In this context, the best indication of benefit is the average result across the respective trials*." The company goes on to conclude that "*We don't believe that the subgroup results referred to in this question have any implications for generalisability. On face-value, CONCUR may be more generalisable to the UK in terms of prior treatments (compared to CORRECT), however, as mentioned above there is no clear explanation for the difference in results observed between the two trials. We consider that the results of both trials are generalisable to England and that the best estimate of effect probably lies somewhere between both studies and should be estimated via meta-analysis*". The ERG notes that there is a contradiction in the company's argument that the trials are underpowered to detect subgroup differences and yet the differences should be viewed in the context of their CIs. If the trials are underpowered then the CIs will be wider than they would have otherwise been and so the CIs are no longer a reliable means to infer population difference. In an underpowered trial then the onus is on the reviewer to be vigilant for possible type II errors which will involve careful interpretation of point estimates (alongside, of course, consideration of the confidence intervals and sample size).
3) CONCUR comprised of participants wholly from Asian countries, whereas CORRECT participants were drawn from a worldwide base. Given that CONCUR had better results for both OS and PFS, this initially suggests that race may be an important covariate. The CORRECT subgroup analyses for PFS support this to some extent, where Asian participants appeared to do slightly better (although there is much uncertainty). This may mean that using the overall pooled results from CORRECT and CONCUR, which have a greater proportion of Asian participants overall than the UK population, may lead to an over-optimistic impression of the benefits of regorafenib in the UK population in terms of PFS. This is despite the assertion of the clinical experts* cited by the company, who state that race does not affect the performance of the drug. For the outcome of OS, it is clear that race cannot explain the better OS results for the CONCUR patients overall, as the subgroup analyses showed that Asian patients actually did slightly worse in terms of OS (although

again there was much uncertainty). In the clarification response the company refers to the subgroup data from the OS analysis in CORRECT to appropriately show that race was not an important factor influencing PFS. They stated that "*as no interaction is observed for region or race the results of CORRECT are considered to be generalisable to UK patients*". [6] However, the company failed to refer to the subgroup data for PFS, which shows a point estimate difference between the subgroups indicating that Asian participants may have better PFS when on regorafenib. Although there is uncertainty in this result, it is unlikely that the subgroups were powered to detect differences, and so it is appropriate for the ERG to be vigilant for possible type II errors and to make the committee aware of them.

Generally, when challenged on the implications for generalisability the company stated that: *"Although these analyses are supportive, they are post-hoc in nature and caution should be exercised when interpreting any post-hoc results. The subgroup results show an overall consistency of benefit across a wide range of patient characteristics. However, we believe that the results in the subgroups should not be overinterpreted as the trials were powered at the ITT level - not the subgroup level. Numbers in some subgroups are relatively small. It is unrealistic to expect that point estimates across 'related' subgroups will always be clinically logical - there will always be the play of chance. Furthermore, treatment groups were randomised at the overall population level ensuring important characteristics were well balanced. Some imbalances may arise between treatment groups in subpopulations which could influence outcomes… Regorafenib has been shown to be effective in the mCRC population overall and in subgroup and exploratory analyses. As a consequence of powering and low numbers in some subgroups overinterpretation should be avoided. The results from CONCUR (and CORRECT) are generalisable to the UK and individual subgroup results supports a consistent effect across a wide range of subgroups".*[6] As explained previously, the ERG holds the view that low powering of subgroup analyses does not remove the need to look for possible subgroup differences. Because of the lack of statistical power there is a need to interpret point estimates more broadly, especially where the magnitude of between-subgroup difference is large, with the onus on being vigilant for possible type II errors. Because it is uncertain as to the size and direction of subgroup differences, it is likely that unknown covariates may therefore be causing the difference in outcomes between CONCUR and CORRECT. If these unknown factors differ between the studies and the UK population then this again reduces the applicability of the study findings. The difference in treatment effect between the two trials and the disparity between treatment experience and race in the trials and NHS clinical practice, this is a key issue. This is also the justification for the sensitivity analyses in the NMA with various combinations of both regorafenib and T/T trials (see Sections 3.3 and 3.4).

*The reference to the clinical experts at this point prompted the ERG to ask the company to provide further details of any clinical experts, and how opinion was elicited. The company response was as follows: "*an advisory board was conducted with 10 consultant medical & clinical oncologists who currently treat patients with mCRC. The oncologists were from hospitals across the UK (London, Southampton, Cardiff, Manchester, Sheffield, Scotland). The advisory board was chaired by an oncologist and had the following objectives:*

- *Understanding UK clinical practice (treatments received, assessment of progression, resource use)*
- *Gathering feedback on regarafenib's [sic] trial data and generalisability to the UK*
- *Gathering clinical views of the relative efficacy of regorafenib and trifluridine/tipiracil*
- *Understanding the place of regorafenib in the treatment pathway*
- *Assessing the appropriateness of the economic model.*

The meeting was structured around the presentation of trial data and group discussion on the above topics."[6] The ERG would have preferred this information to have been accompanied by a date, venue and meeting notes.

3.2.3.2 Comparative observational studies

Tables 3.9 to 3.12 show the baseline characteristics from the comparative observational studies.

Table 3.9: Baseline characteristics of patients in Tanaka 2018

Characteristic	Regorafenib (n=20)	T/T (n=24)
Age, median (range)	68 [57-78]	64 [44-86]
Sex Male	13 (65.0)	15 (62.5)
Female	7 (35.0)	9 (37.5)
ECOG PS 0	6 (30.0)	14 (58.3)
1	12 (60.0)	6 (25.0)
2	2 (10.0)	4 (16.7)
Primary site of disease Right colon	4 (20.0)	10 (41.7)
Left colon	7 (35.0)	3 (12.5)
Rectum	9 (37.5)	11 (45.8)
KRAS exon 2 status Wild	9 (45.0)	14 (58.3)
Mutation	11 (55.0)	10 (41.7)
Number of prior regimens 2	12 (60.0)	12 (50.0)
3	8 (40.0)	11 (45.8)
≥4	0 (0)	1 (4.2)
Number of metastatic sites 1	6 (30.0)	6 (25.0)
2	12 (60.0)	11 (45.8)
≥3	2 (10.0)	7 (29.2)
Metastatic site Liver	16 (80.0)	19 (79.2)
Lung	10 (50.0)	13 (54.2)
Peritoneum	6 (30.0)	4 (16.7)
Lymph node	2 (10.0)	8 (33.3)
Others	2 (10.0)	8 (33.3)
Time from initiation of first-line chemotherapy ≤18 months	5 (25.0)	6 (25.0)
>18 months	15 (75.0)	18 (75.0)
History of systemic anticancer agents Fluoropyrimidine	20 (100)	24 (100)

Characteristic	Regorafenib (n=20)	T/T (n=24)
Oxaliplatin	20 (100)	24 (100)
Irinotecan	20 (100)	24 (100)
Anti-VEGF antibody	20 (100)	23 (95.8)
Anti-EGFR antibody (wild KRAS or all-RAS[a])	9 (45.0)	11 (45.8)
Post-treatment use of regorafenib or T/T	7 (35.0)	10 (41.7)

Source: Table 1, Tanaka 2018[23]
ECOG PS = Eastern Cooperative Oncology Group Performance Status; EGFR = epidermal growth factor receptor; KRAS = Kirsten rat sarcoma viral oncogene homologue; T/T = trifluridine/tipiracil; VEGF = vascular endothelial growth factor

Table 3.10: Baseline characteristics of patients in Sueda 2016

Characteristic	Regorafenib (n=23)	T/T (n=14)
Age, median (range)	59 (37-83)	66 (44-80)
Sex Male	12 (52.2)	10 (71.4)
Female	7 (35.0)	9 (37.5)
ECOG PS 0	10 (43.5)	1(7.2)
1	13 (56.5)	10 (71.4)
2	0 (0.0)	3 (21.4)
Primary site of disease Colon	13 (56.5)	10 (71.4)
Rectum	10 (43.5)	4 (28.6)
KRAS exon 2 status Wild	12 (52.2)	9 (64.3)
Mutation	11 (47.8)	5 (35.7)
Metastatic site Liver	14	9
Lung	12	9
Peritoneum	2	2
Lymph node	3	7
Others	0	0
Anti-VEGF antibody	23 (100)	10 (71.4)
Anti-EGFR antibody	12 (52.2)	9 (64.3)
Post-treatment use of regorafenib or T/T	6 (26.1)	8 (57.1)

Source: Table 1, Sueda 2016[25]
ECOG PS = Eastern Cooperative Oncology Group Performance Status; EGFR = epidermal growth factor receptor; KRAS = Kirsten rat sarcoma viral oncogene homologue; T/T = trifluridine/tipiracil; VEGF = vascular endothelial growth factor

Table 3.11: Baseline characteristics of patients in Huemer 2020

Characteristic		Regorafenib third line *n*=69 (%)	T/T third line *n*=24 (%)
Age at third line start	median (range)	65 (42–85)	68 (49–81)
Tertiary cancer centre	Salzburg Wels-Grieskirchen	43 (62) 26 (38)	17 (71) 7 (29)
Sex	male female	40 (58) 29 (42)	14 (58) 10 (42)
ECOG PS at third line start	0 1 2 3 NA	21 (40) 25 (47) 6 (11) 1 (2) 16	4 (21) 9 (48) 5 (26) 1 (5) 5
Detection of metastases	synchronous metachronous	49 (71) 20 (29)	13 (54) 11 (46)
Primary tumour resected	yes no	54 (78) 15 (22)	21 (88) 3 (12)
Sidedness	left right	47 (68) 22 (32)	20 (83) 4 (17)
Metastatic pattern at third line start	liver lung peritoneum	56 (81) 49 (71) 10 (14)	16 (67) 15 (63) 6 (25)
Ascites at third line start	yes no	4 (6) 65 (94)	4 (17) 20 (83)
KRAS status	wild type mutant	(49) (51)	12 (50) 12 (50)
BRAF status	wild type mutant NA	50 (100) 0 (0) 19	18 (95) 1 (5) 5
Microsatellite status	MSS MSI NA	44 (98) 1 (2) 24	14 (93) 1 (7) 9
Subsequent therapy with regorafenib or T/T	regorafenib T/T	- 31 (45)	7 (29) -
Subsequent other systemic therapy after regorafenib and/or T/T	yes no	16 (23) 53 (77)	4 (17) 20 (83)

Source: Table 1, Huemer 2020[26]
BRAF = mutation in the B-Raf proto-oncogene; ECOG PS = Eastern Cooperative Oncology Group Performance Status; KRAS = Kirsten rat sarcoma viral oncogene homologue; MSS = microsatellite status; NA = not applicable; T/T = trifluridine/tipiracil

Table 3.12: Baseline characteristics of patients in Nakashima 2020

	Regorafenib (n (1,501))	T/T (n (3,777))
Gender (male)	928 (62)	2,250 (60)
Age, years	66 (60-73)	68 (62-75)
>65	803 (54)	2,402 (64)
Body weight, kg	56 (49-65)	56 (48-64)
Missing data	98 (7)	208 (6)
Body mass index, kg/m^2	21.6 (19.3-24.0)	21.8 (19.5-24.1)
<18.5	239 (17)	549 (15)
Missing data	102 (7)	221 (6)
Comorbidity		
Hypertension	837 (56)	2,223 (59)
Diabetes mellitus	462 (31)	1,154 (31)
Hyperlipidaemia	234 (16)	730 (19)
Hepatitis B	150 (10)	402 (11)
Hepatitis C	60 (4)	112 (3)
Peripheral neuropathy	515 (34)	1,424 (38)
Hand-foot syndrome	194 (13)	521 (14)
Anaemia	41 (3)	89 (2)
Leukopenia	41 (3)	128 (3)
Interstitial pneumonitis	7 (0)	19 (1)
Primary site of disease		
Colon	883 (59)	2,145 (57)
Rectum	406 (27)	1,039 (28)
Colon and rectum	212 (14)	593 (16)
Metastatic sites		
Liver	966 (64)	2,341 (62)
Lung	694 (46)	1,803 (48)
Lymph node	343 (23)	866 (23)
Peritoneum	431 (29)	1,125 (30)
Bone	242 (16)	576 (15)
Brain	94 (6)	207 (5)
Other metastases	148 (10)	402 (11)
Number of metastatic sites (3)	447 (30)	1,073 (28)
Previous systemic anticancer agents		
Fluorouracil	917 (61)	2,319 (61)
Capecitabine	458 (31)	1,470 (39)
Tegafur/gimeracil/oteracil	427 (28)	1,262 (33)
Tegafur/uracil	106 (7)	361 (10)
Oxaliplatin	989 (66)	2,714 (72)

	Regorafenib (n (1,501))	T/T (n (3,777))
Irinotecan	1,147 (76)	2,874 (76)
Bevacizumab (anti-VEGF antibody)	1,081 (72)	2,786 (74)
Cetuximab (anti-EGFR antibody)	232 (15)	427 (11)
Panitumumab (anti-EGFR antibody)	403 (27)	969 (26)
Aflibercept (anti-VEGF antibody)	19 (1)	40 (1)
Ranibizumab (anti-VEGF antibody)	102 (7)	269 (7)
Number of previous anticancer agents	4 (3-5)	4 (3-5)
Any previous targeted therapy	1,232 (82)	3,138 (83)
Subsequent other systemic therapy after regorafenib and/or T/T		

Source: Table 1, Nakashima 2020[1]
EGFR = epidermal growth factor receptor; T/T = trifluridine/tipiracil; VEGF = vascular endothelial growth factor

ERG comment:

There seem to be two things that stand out in an examination of the baseline characteristics for three of the observational studies, Tanaka 2018,[23] Sueda 2016,[25] and Huemer 2020,[26] imbalances between the arms in various characteristics, most notably ECOG status and anti-VEGF treatment for Sueda 2016, and the use of post-treatment regorafenib or T/T (crossover). The pattern of those imbalances is such that any consequences on the treatment effect are difficult to predict. In contrast, Nakashima 2020[1] seems to show greater balance and the baseline characteristics for patients who did not crossover, receiving BSC only as subsequent treatment, were reported separately. It should be noted that analysis regarding crossover varied by study (see Section 3.2.3.2). Three of the studies, Nakashima 2020,[1] Sueda 2016,[25] and Huemer 2020,[26] reported outcomes separately for those patients who did not crossover. Tanaka 2018,[23] in contrast, reported results for all patients regardless of whether they crossed over. It should also be noted that the reason for crossover was not reported in Nakashima 2020,[1] or Sueda 2016.[25] However, Huemer 2020,[26] reported that no patients with an ECOG score above 1 on first treatment with regorafenib or T/T crossed over and Tanaka 2018,[23] reported that crossover only occurred in patients with an ECOG PS of 0 or 1 on completion of first treatment.

3.2.4 Risk of bias assessment

3.2.4.1 Randomised trials

According to the company, both regorafenib trials were considered by the company to be methodologically robust, high-quality studies with a comprehensive approach to patient allocation, control of confounding factors, and an overall low risk of bias. Table 3.13 summarises these findings, according to the company criteria.

Table 3.13: Quality assessment of the two randomised trials

Study details	Randomisation appropriate?	Allocation concealment adequate?	Groups similar at the outset of the study in terms of prognostic factors?	Blinding to treatment allocation?	Unexpected imbalances in drop-outs between groups?	Authors measured more outcomes than they reported?	Did the analysis include an intention-to-treat analysis?
Regorafenib versus placebo							
Grothey 2013[14]	Yes	Yes	Yes	Yes	No	No	Yes
Li 2015[15]	Yes	Yes	Yes	Yes	No	No	Yes
Source: Table 4, Appendices CS[7] CS = company submission							

ERG comment:

The ERG partially agrees with the company's quality assessment of the randomised trials evaluating regorafenib versus placebo, in terms of issues concerning selection bias, attrition bias and outcome-reporting bias. However, one issue remains in relation to performance bias. In the CORRECT and CONCUR trials the median duration of treatment was 2.8 and 2.4 months in the regorafenib group and 1.8 and 1.6 months in the placebo group. It is possible that this difference could contribute to considerable performance bias in favour of regorafenib. In the clarification letter the company were asked to explain the source of this discrepancy and the likely effect on outcomes.[6] The company responded by stating that, "*In the management of mCRC (and typically all cancers), patients are treated until progression is observed. Progression signifies the cancer has become resistant to that treatment. If the patient is earlier in the treatment pathway, then progression would result in the stopping of that treatment and the start of the next line of therapy. Due to the anti-cancer activity of regorafenib it took longer for progression to take place which explains the difference in median treatment duration between arms – there is no bias that is introduced*." The ERG were satisfied by this response, which fully explained the discrepancy and consider the RCTs to be of high quality.

3.2.4.2 Comparative observational studies

Three of the observational studies were evaluated in the appendices of the CS[3] using the Downs and Black checklist. The company evaluation is summarised in Table 3.14.

Table 3.14: Quality assessment of included observational and RWE studies using Downs and Black checklist

Question	Tanaka 2018[23]	Sueda 2016[25]	Huemer 2020[26]
Is the hypothesis/aim/objective of the study clearly described?	Yes	Yes	Yes
Are the main outcomes to be measured clearly described in the introduction or methods section?	Yes	Yes	Yes
Are the characteristics of the patients included in the study clearly described?	Yes	No	Yes
Are the interventions of interest clearly described?	Yes	Yes	Yes
Are the distributions of principal confounders in each group of patients to be compared clearly described?	Yes	Yes	Yes
Are the main findings of the study clearly described?	Yes	Yes	Yes
Does the study provide estimates of the random variability in the data for the main outcomes?	Yes	Yes	Yes
Have all important adverse events that may be a consequence of the intervention been reported?	Yes	Yes	No
Have the characteristics of patients lost to follow-up been described?	No	UTD	UTD
Have actual probability values been reported (e.g., 0.035 rather than < 0.05) for the main outcomes except where the probability value is less than 0.001?	Yes	Yes	Yes
Were the subjects asked to participate in the study representative of the entire population from which they were recruited?	Yes	Yes	Yes
Were those subjects who were prepared to participate representative of the entire population from which they were recruited?	Yes	Yes	Yes
Were the staff, places, and facilities where the patients were treated representative of the treatment the majority of patients receive?	Yes	Yes	Yes
Was an attempt made to blind study subjects to the intervention they have received?	NA	NA	No
Was an attempt made to blind those measuring the main outcomes of the intervention?	No	No	No
If any of the results of the study were based on ‘data dredging’, was this made clear?	UTD	No	No

Question	Tanaka 2018[23]	Sueda 2016[25]	Huemer 2020[26]
In trials and cohort studies, do the analyses adjust for different lengths of follow-up of patients, or in case-control studies, is the time period between the intervention and outcome the same for cases and controls?	Yes	No	Yes
Were the statistical tests used to assess the main outcomes appropriate?	Yes	Yes	Yes
Was compliance with the intervention(s) reliable?	UTD	No	Yes
Were the main outcome measures used accurate (valid and reliable)?	Yes	Yes	Yes
Were the patients in different intervention groups (trials and cohort studies) or were the cases and controls (case-control studies) recruited from the same population?	Yes	Yes	Yes
Were study subjects in different intervention groups (trials and cohort studies) or were the cases and controls (case-control studies) recruited over the same period of time?	Yes	Yes	Yes
Were study subjects randomised to intervention groups?	NA	NA	No
Was the randomized intervention assignment concealed from both patients and health care staff until recruitment was complete and irrevocable?	NA	No	No
Was there adequate adjustment for confounding in the analyses from which the main findings were drawn?	Yes	No	No
Were losses of patients to follow-up taken into account?	No	UTD	Yes
Source: Table 5, Appendices of the CS[7] CS = company submission; NA = not applicable; UTD = unable to determine			

ERG comment:

The ERG mostly agrees with the company's evaluation of the three observational trials, although it should be pointed out that the ERG believes that there was not adequate adjustment for confounding for the analyses relevant to the NICE scope in Tanaka 2018[23], Sueda 2016[25] and Huemer 2020.[26] The company has not provided an overall rating of quality for the three observational trials, but the ERG evaluation is that all three studies are at very high risk of bias. This rating is largely due to very serious selection bias secondary to their non-randomised design, which is not ameliorated by statistical adjustment or matching. Therefore, any results from these three studies should be interpreted with great caution.

Nakashima 2020[1] was not included in the CS[3] and so the ERG has generated its own risk of bias evaluation, using the Downs and Black checklist, for comparability with the other three studies. Table 3.15 summarises the evaluation. Overall, as for the other three studies, Nakashima 2020[1] was classed as at very high risk of bias due to being non-randomised.

Table 3.15: Quality assessment of Nakashima 2020 [1]using Downs and Black checklist

Question	Nakashima 2020[1]
Is the hypothesis/aim/objective of the study clearly described?	Yes
Are the main outcomes to be measured clearly described in the introduction or methods section?	Yes
Are the characteristics of the patients included in the study clearly described?	Yes
Are the interventions of interest clearly described?	Yes
Are the distributions of principal confounders in each group of patients to be compared clearly described?	Yes
Are the main findings of the study clearly described?	Yes
Does the study provide estimates of the random variability in the data for the main outcomes?	Yes
Have all important adverse events that may be a consequence of the intervention been reported?	Yes
Have the characteristics of patients lost to follow-up been described?	No
Have actual probability values been reported (e.g., 0.035 rather than < 0.05) for the main outcomes except where the probability value is less than 0.001?	Yes
Were the subjects asked to participate in the study representative of the entire population from which they were recruited?	Yes
Were those subjects who were prepared to participate representative of the entire population from which they were recruited?	Yes
Were the staff, places, and facilities where the patients were treated representative of the treatment the majority of patients receive?	Yes
Was an attempt made to blind study subjects to the intervention they have received?	NA
Was an attempt made to blind those measuring the main outcomes of the intervention?	No
If any of the results of the study were based on 'data dredging', was this made clear?	UTD
In trials and cohort studies, do the analyses adjust for different lengths of follow-up of patients, or in case-control studies, is the time period between the intervention and outcome the same for cases and controls?	Yes
Were the statistical tests used to assess the main outcomes appropriate?	Yes
Was compliance with the intervention(s) reliable?	UTD

Question	Nakashima 2020[1]
Were the main outcome measures used accurate (valid and reliable)?	Yes
Were the patients in different intervention groups (trials and cohort studies) or were the cases and controls (case-control studies) recruited from the same population?	Yes
Were study subjects in different intervention groups (trials and cohort studies) or were the cases and controls (case-control studies) recruited over the same period of time?	Yes
Were study subjects randomized to intervention groups?	No
Was the randomized intervention assignment concealed from both patients and health care staff until recruitment was complete and irrevocable?	NA
Was there adequate adjustment for confounding in the analyses from which the main findings were drawn?	Possibly
Were losses of patients to follow-up taken into account?	No
NA = not applicable; UTD = unable to determine	

3.2.5 Efficacy results of the included studies

The final NICE scope lists the following outcomes that need to be covered in the TA:

- OS
- PFS
- Response rates
- HRQoL
- AEs of treatment

For the randomised data, the first four of these outcomes will be evaluated in Sections 3.2.5.1 to 3.2.5.4. Section 3.2.5.5 will summarise results for the first three outcomes from the four comparative observational studies.[1, 23, 25, 26] AEs are evaluated separately in Section 3.2.6.

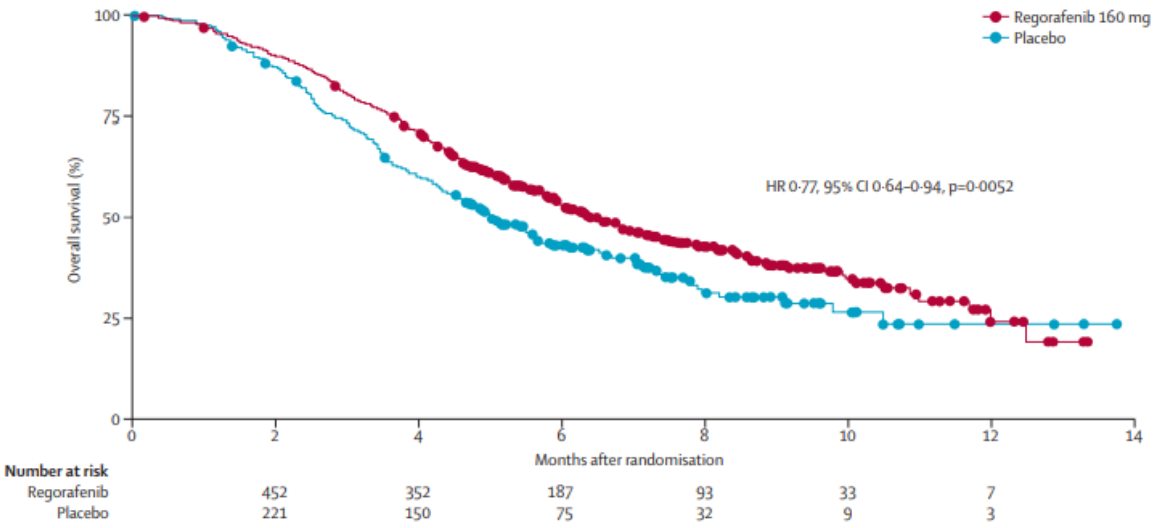
3.2.5.1 Overall survival (OS) data from the randomised trials

CORRECT

The median duration of treatment was 2.8 versus 1.8 months for regorafenib versus placebo. A total of 432 death events occurred in the ITT population (n=760), with the majority occurring in the placebo group [redacted]

In CORRECT, the primary endpoint was met. Regorafenib significantly prolonged OS compared with placebo (HR 0.77; 95% CI: 0.64, 0.94). Figure 3.1 provides the KM plot for OS in the CORRECT study.

Figure 3.1: CORRECT – Kaplan–Meier curve of overall survival (ITT)



Source: Figure 5, CS[3]
CI = confidence interval; CS = company submission; HR = hazard ratio; ITT = intention-to-treat; OS = Overall Survival

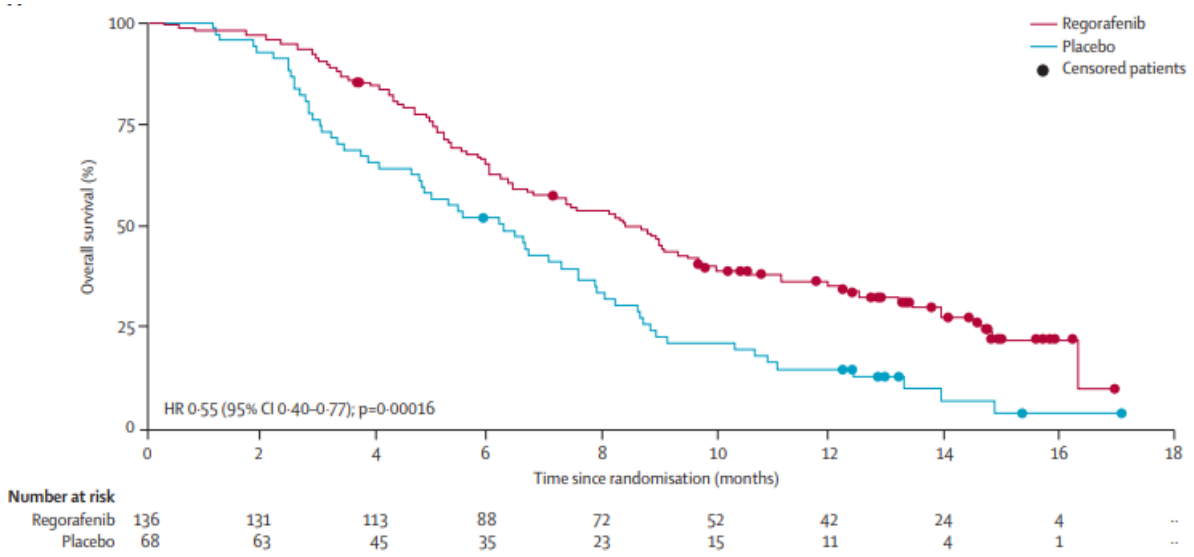
CONCUR

The median duration of treatment was 2.4 versus 1.6 months for regorafenib versus placebo. A total of 155 death events occurred in the FAS population (n=204), with the majority occurring in the placebo

group (regorafenib: 70%; placebo: 88%). Median follow-up for the OS analysis was 7.4 months (IQR: 4.3–12.2).

In CONCUR, the primary endpoint was met. Regorafenib significantly prolonged OS compared with placebo (HR 0.55; 95% CI: 0.40, 0.77). Figure 3.2 provides the KM plot for OS in the CONCUR study.

Figure 3.2: CONCUR – Kaplan–Meier curve of overall survival (FAS)

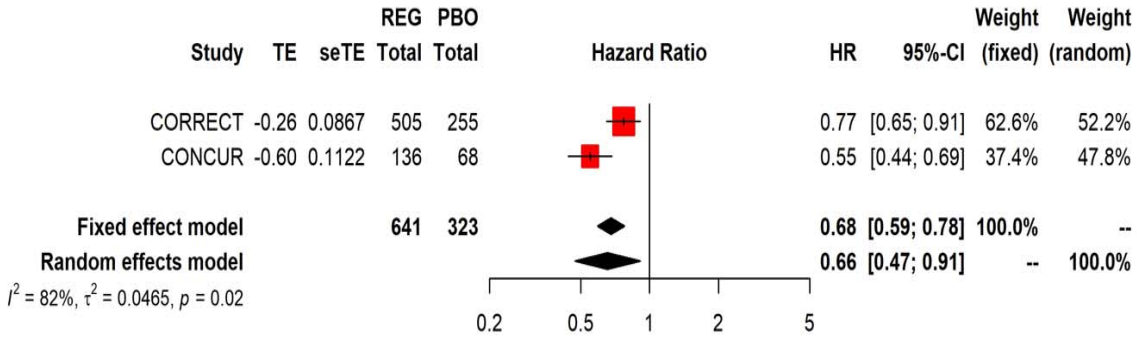


Source: Figure 7, CS[3]
CI = confidence interval; CS = company submission; FAS= full analysis set; HR = hazard ratio; OS = Overall Survival.

META-ANALYSIS of CORRECT and CONCUR

A direct meta-analysis using fixed and random effects models for OS data from CORRECT and CONCUR was performed, yielding a pooled fixed-effect HR of 0.68 (95% CI: 0.58, 0.79) and random effect HR of 0.66 (95% CI: 0.47, 0.91) (see Figure 3.3). The I^2 statistic was estimated at 82%, which suggests a high level of heterogeneity between the studies.

Figure 3.3: Direct meta-analysis – overall survival – CORRECT and CONCUR



Taken from Figure 9, CS[3]
CI = confidence interval; CS = company submission; HR = hazard ratio; PBO = placebo; REG = regorafenib; seTE = standard error of treatment effect; TE = treatment effect.

ERG comment:

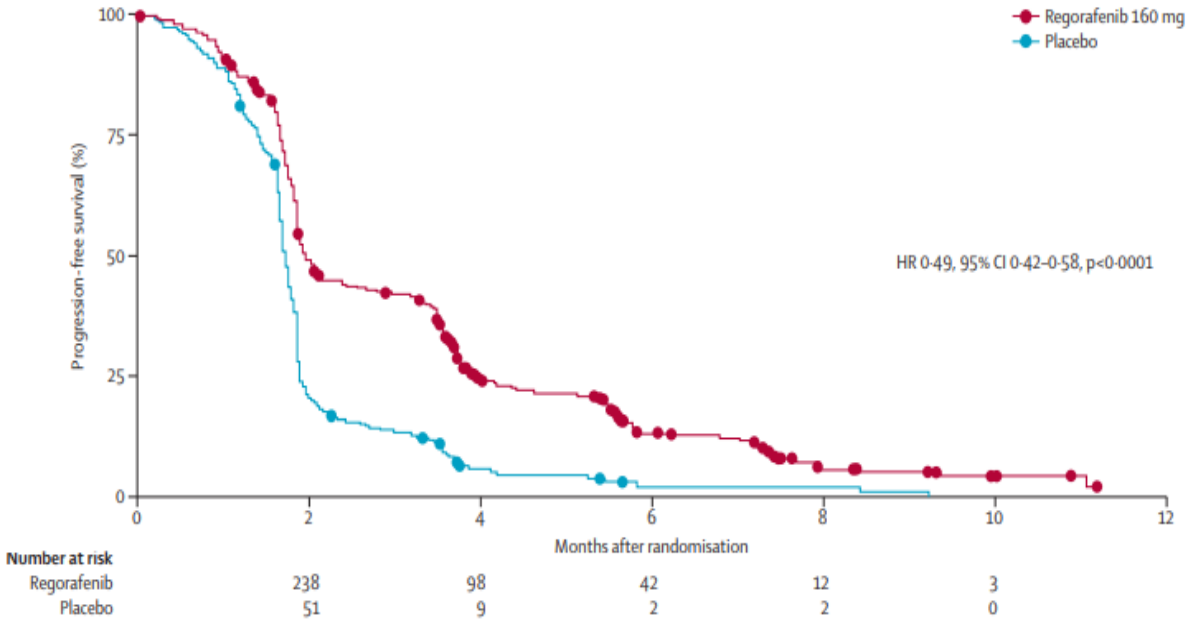
It appears as though the fixed effect HR was the data point fed through to the ITC, even though the random effects HR would have been more appropriate given the large amount of unexplained heterogeneity.

3.2.5.2 Progression-free survival (PFS) data from the randomised trials

CORRECT

PFS was significantly longer in the regorafenib group compared with placebo (HR 0.49; 95% CI: 0.42, 0.58). [redacted] of patients had experienced a PFS event. Figure 3.4 provides the KM plot for these data.

Figure 3.4: CORRECT – Kaplan–Meier curve of progression-free survival (ITT)



Source: Figure 6 in CS[3]
CI = confidence interval; CS = company submission; HR = hazard ratio; ITT = intention-to-treat; KM = Kaplan-Meier; PFS = progression-free survival
Note:
KM curves based on investigator assessment.

ERG comment:

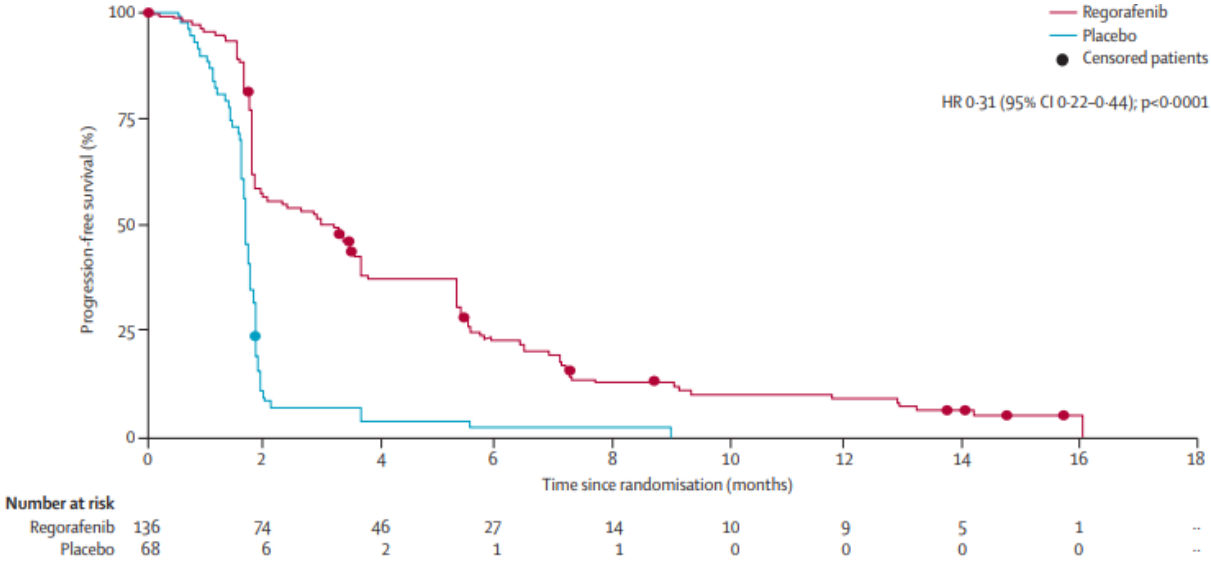
In the PFS KM curve for CORRECT there appear to be periodic increases in the negative gradient of the curve, corresponding to the periods immediately before the 2, 4, 6, 8 and 10 month follow-up points, and these are particularly marked in the regorafenib group. Is this an artefact of the timing of outcome evaluation? In the clarification letter the company has been asked to explain this phenomenon, because the shape of the curves has an impact on interpretation. The company explained that: "*In CORRECT and CONCUR patients were assessed every 8 weeks to determine if the cancer had progressed. This 8-weekly assessment matches clinical practice in the UK. The 'stepped' nature of the KM curve reflects the timepoints of clinical assessment. However, the 'step' is not perfectly vertical as there was a one-*

week window either side of 8-week timepoint where assessments could take place." The EAG was satisfied with this explanation, which clearly explained the observed phenomenon.

CONCUR

PFS was significantly longer in the regorafenib group compared with placebo (HR 0.31; 95% CI: 0.22, 0.44). [redacted] of patients had experienced a PFS event. Figure 3.5 provides the KM plot for these data.

Figure 3.5: CONCUR – Kaplan–Meier curve of progression-free survival (FAS)

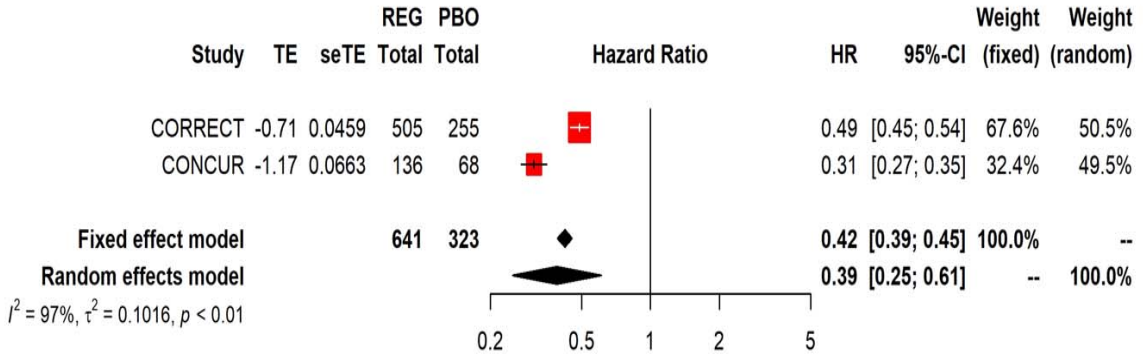


Source: Figure 8 in CS[3]
CI = confidence interval; CS = company submission; FAS = full analysis set; HR = hazard ratio; KM = Kaplan-Meier; PFS = progression-free survival
Note:
KM curves based on investigator assessment

META-ANALYSIS of CORRECT and CONCUR

A direct meta-analysis using fixed and random effects models for PFS data from CORRECT and CONCUR was performed. For PFS, the fixed effects HR was 0.42 (95% CI: 0.39, 0.45) and the random effects HR was 0.39 (95% CI: 0.25, 0.61). The I^2 statistic was estimated at 97%, which suggests a high level of heterogeneity between the studies. This is summarised in Figure 3.6.

Figure 3.6: Direct meta-analysis – progression-free survival – CORRECT and CONCUR



Source: Figure 10, CS[3]
CI = confidence interval; CS = company submission; HR = hazard ratio; PBO = placebo; REG = regorafenib; seTE = standard error of treatment effect; TE = treatment effect

ERG comment:

It appears as though the fixed effect HR was used for the ITC, even though the random effects HR would have been more appropriate given the large amount of unexplained heterogeneity.

3.2.5.3 Response rates data from the randomised trials

Response rates in CORRECT and CONCUR are summarised in Tables 3.16 and 3.17. Overall, regorafenib led to better response rates than placebo.

Table 3.16: CORRECT – tumour response (ITT)

Response	Regorafenib (n=505)	Placebo (n=255)
Best response		
CR, n (%) [95% CI]	[redacted]	[redacted]
PR, n (%) [95% CI]	[redacted]	[redacted]
SD, n (%) [95% CI]	[redacted]	[redacted]
PD, n (%) [95% CI]	[redacted]	[redacted]
Non CR/Non PD, n (%)[a] [95% CI]	[redacted]	[redacted]
ORR and DCR, n (%) [95% CI]		

	Regorafenib (n=505)	Placebo (n=255)	P-value
ORR[b]	[redacted]	[redacted]	[redacted]
DCR[c]	[redacted]	[redacted]	[redacted]

Response	Regorafenib (n=505)	Placebo (n=255)
Source: Table 10, CS[3] CI = confidence interval; CR = complete response; CS = company submission; DCR = disease control rate; ITT = intention-to-treat; ORR = overall response rate; PD = progressive disease; PR = partial response; RR = relative risk; SD = stable disease **Notes:** [a] non CR/non PD included in DCR and followed the same criteria as stable disease [b] percentage of patients with CR or PR [c] percentage of patients with CR or PR or SD according to RECIST version 1.1. Patients with SD as response performed earlier than 6 weeks after randomisation were not taken into account. Non-CR/non-PD were included in disease control rate and followed same criteria as stable disease		

Table 3.17: CONCUR – tumour response (FAS)

Response[a]	Regorafenib (n=136)	Placebo (n=68)
Best response		
CR, n (%) [95% CI]	[redacted]	[redacted]
PR, n (%) [95% CI]	[redacted]	[redacted]
SD[b], n (%) [95% CI]	[redacted]	[redacted]
PD, n (%) [95% CI]	[redacted]	[redacted]
Non CR/Non PD, n (%)[a] [95% CI]	[redacted]	[redacted]
ORR and DCR, n (%) [95% CI]		

	Regorafenib (n = 136)	Placebo (n = 68)	P-value
ORR[c]	[redacted]	[redacted]	[redacted]
DCR[d]	[redacted]	[redacted]	[redacted]

Source: table 12, CS[3]
CI = confidence interval; CR = complete response; CS = company submission; DCR = disease control rate; FAS =-full analysis set; N=number of patients; ORR = overall response rate; PD = progressive disease; PR = partial response; RR = relative risk; SD = stable disease
Notes:
[a] 95% CI by exact binomial calculation. Non-CR/non-PD included in disease control rate and followed same criteria as stable disease. PD and non-CR/non-PD assessed after at least 6 weeks. Denominator for rates (%) and 95% CIs based on FAS (all randomised patients)
[b] number of weeks without progression
[c] percentage of patients with CR or PR
[d] percentage of patients with CR or PR or SD according to RECIST version 1.1. Patients with stable disease as response performed earlier than 6 weeks after randomisation were not taken into account. Non-CR/non-PD were included in disease control rate and followed same criteria as stable disease

ERG comments: No pooling was carried out for this outcome across the CORRECT and CONCUR studies. Between-group measures of effect were not calculated by the company (although P values were

provided) making it more difficult to interpret results, but these have been calculated by the ERG where possible.

3.2.5.4 Heath-related quality of life (HRQoL) data from the randomised trials

CORRECT

Tables 3.18 and 3.19 summarise the results for European Organization for Research and Treatment of Cancer Quality of Life Questionnaire (30 items) (EORTC QLQ-C30), European Quality of Life-5 Dimensions (EQ-5D) index and visual analogue scale (VAS) score for the CORRECT study. The CS[3] provided results for each arm, but did not provide between-arm (mean difference (MD) (95% CI)) results. The ERG has calculated these for the end-of-treatment (EOT) data and added these as a final column to assist in interpretation of the data. In general, regorafenib and placebo did not differ in their effects on QoL, apart from EORTC QLQ-C30 social function, where the regorafenib group suffered a worse decrement than the placebo group

Table 3.18: CORRECT – EORTC QLQ-C30 change from baseline at Cycles 2, 3, 4 and EOT (ITT)

	Placebo (N=255)		**Regorafenib (N=505)**		**MD (95% CI) for EOT data (regorafenib – placebo)**
	n	**Mean (SD)**	**n**	**Mean (SD)**	**Calculated and added by ERG**
Physical function					
Cycle 2	[redacted]	[redacted]	[redacted]	[redacted]	
Cycle 3	[redacted]	[redacted]	[redacted]	[redacted]	
Cycle 4	[redacted]	[redacted]	[redacted]	[redacted]	
EOT	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Role function					
Cycle 2	[redacted]	[redacted]	[redacted]	[redacted]	
Cycle 3	[redacted]	[redacted]	[redacted]	[redacted]	
Cycle 4	[redacted]	[redacted]	[redacted]	[redacted]	
EOT	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Emotional function					
Cycle 2	[redacted]	[redacted]	[redacted]	[redacted]	
Cycle 3	[redacted]	[redacted]	[redacted]	[redacted]	
Cycle 4	[redacted]	[redacted]	[redacted]	[redacted]	
EOT	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Social function					
Cycle 2	[redacted]	[redacted]	[redacted]	[redacted]	
Cycle 3	[redacted]	[redacted]	[redacted]	[redacted]	
Cycle 4	[redacted]	[redacted]	[redacted]	[redacted]	
EOT	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cognitive function					

Cycle 2	[redacted]	[redacted]	[redacted]	[redacted]	
Cycle 3	[redacted]	[redacted]	[redacted]	[redacted]	
Cycle 4	[redacted]	[redacted]	[redacted]	[redacted]	
EOT	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Global health status (QoL)					
Cycle 2	[redacted]	[redacted]	[redacted]	[redacted]	
Cycle 3	[redacted]	[redacted]	[redacted]	[redacted]	
Cycle 4	[redacted]	[redacted]	[redacted]	[redacted]	
EOT	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Source: Table 85, CS appendices[7]
CS = company submission; CI = confidence interval; EORTC QLQ-C30 = European Organization for Research and Treatment of Cancer Quality of Life Questionnaire (30 items); EOT = end of treatment; ERG = Evidence Review Group; ITT = intention-to-treat; MD = mean difference; QoL = quality of life; SD = standard deviation

Table 3.19: CORRECT – EQ-5D index and VAS score changes from baseline at Cycles 2, 3, 4 and EOT (ITT)

	Change from baseline				
	Placebo (N=255)		**Regorafenib (N=505)**		**MD (95% CI) for EOT data (regorafenib – placebo)**
	n	**Mean (SD)**	**Calculated and added by ERG**	**Mean (SD)**	**Calculated and added by ERG**
EQ-5D index					
Cycle 2	[redacted]	[redacted]	[redacted]	[redacted]	
Cycle 3	[redacted]	[redacted]	[redacted]	[redacted]	
Cycle 4	[redacted]	[redacted]	[redacted]	[redacted]	
EOT	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EQ-5D VAS					
Cycle 2	[redacted]	[redacted]	[redacted]	[redacted]	
Cycle 3	[redacted]	[redacted]	[redacted]	[redacted]	
Cycle 4	[redacted]	[redacted]	[redacted]	[redacted]	
EOT	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Source: Table 86, CS appendices[7]
CS = company submission; CI = confidence interval; EQ-5D = Euroquol quality of life questionnaire; EOT = end of treatment; ERG = Evidence Review Group; ITT = intention-to-treat; MD = mean difference; QoL = quality of life; SD = standard deviation; VAS = visual analogue scale

CONCUR

Tables 3.20 and 3.21 summarise the results for EORTC QLQ-C30, EQ-5D index and VAS score for the CORRECT study. The CS[3] provided results for each arm, but did not provide between-arm (MD (95% CI)) results. The ERG has calculated these for the EOT data and added these as a final column to assist in interpretation of the data. In general, regorafenib and placebo did not differ in their effects on QoL.

Table 3.20: CONCUR – EORTC QLQ-C30 change from baseline at Cycles 2, 3, 4 and EOT (FAS)

	Placebo (N=68)		Regorafenib (N=136)		MD (95% CI) for EOT data (regorafenib – placebo)
	n	Mean (SD)	Calculated and added by ERG	Mean (SD)	Calculated and added by ERG
Physical function					
Cycle 2	[redacted]	[redacted]	[redacted]	[redacted]	
Cycle 3	[redacted]	[redacted]	[redacted]	[redacted]	
Cycle 4	[redacted]	[redacted]	[redacted]	[redacted]	
EOT	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Role function					
Cycle 2	[redacted]	[redacted]	[redacted]	[redacted]	
Cycle 3	[redacted]	[redacted]	[redacted]	[redacted]	
Cycle 4	[redacted]	[redacted]	[redacted]	[redacted]	
EOT	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Emotional function					
Cycle 2	[redacted]	[redacted]	[redacted]	[redacted]	
Cycle 3	[redacted]	[redacted]	[redacted]	[redacted]	
Cycle 4	[redacted]	[redacted]	[redacted]	[redacted]	
EOT	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Social function					
Cycle 2	[redacted]	[redacted]	[redacted]	[redacted]	
Cycle 3	[redacted]	[redacted]	[redacted]	[redacted]	
Cycle 4	[redacted]	[redacted]	[redacted]	[redacted]	
EOT	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cognitive function					
Cycle 2	[redacted]	[redacted]	[redacted]	[redacted]	
Cycle 3	[redacted]	[redacted]	[redacted]	[redacted]	
Cycle 4	[redacted]	[redacted]	[redacted]	[redacted]	
EOT	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Global health status (QoL)					
Cycle 2	[redacted]	[redacted]	[redacted]	[redacted]	
Cycle 3	[redacted]	[redacted]	[redacted]	[redacted]	
Cycle 4	[redacted]	[redacted]	[redacted]	[redacted]	
EOT	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Source: table 95, CS appendices[7]
CS = company submission; CI = confidence interval; EORTC QLQ-C30, European Organization for Research and Treatment of Cancer Quality of Life Questionnaire (30 items); EOT = end of treatment; ERG = Evidence Review Group; FAS = full analysis set; MD = mean difference; QoL = quality of life; SD = standard deviation

Table 3.21: CONCUR – EQ-5D index and VAS score changes from baseline at Cycles 2, 3, 4 and EOT (FAS)

	Change from baseline				
	Placebo (N=68)		Regorafenib (N=136)		MD (95% CI) for EOT data (regorafenib – placebo)
	n	Mean (SD)	Calculated and added by ERG	Mean (SD)	Calculated and added by ERG
EQ-5D index					
Cycle 2	[redacted]	[redacted]	[redacted]	[redacted]	
Cycle 3	[redacted]	[redacted]	[redacted]	[redacted]	
Cycle 4	[redacted]	[redacted]	[redacted]	[redacted]	
EOT	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
-0.21EQ-5D VAS					
Cycle 2	[redacted]	[redacted]	[redacted]	[redacted]	
Cycle 3	[redacted]	[redacted]	[redacted]	[redacted]	
Cycle 4	[redacted]	[redacted]	[redacted]	[redacted]	
EOT	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Source: table 96, CS appendices[7]
CS = company submission; CI = confidence interval; EOT = end of treatment; EQ-5D = Euroquol quality of life questionnaire; ERG = Evidence Review Group; FAS = full analysis set; MD = mean difference; SD = standard deviation; VAS = visual analogue scale

ERG comments:

No pooling across CORRECT and CONCUR was carried out for this outcome.

3.2.5.5 OS, PFS and RR data from the comparative observational studies

The four comparative observational studies[1, 23, 25, 26] have provided direct comparative data between regorafenib and T/T. Results for Nakashima 2020,[1] Tanaka 2018,[23] and Sueda 2016[25] are provided below in Table 3.22. Results for Tanaka 2018,[23] and Sueda 2016[25] are adapted from those presented in the CS[3], which only provided results for the regorafenib arm. From the largest study[1] there is a clear effect for OS favouring T/T. In addition, there is no evidence from the other two studies[23, 25] that regorafenib is more effective than T/T for OS or response rates. For PFS, there was heterogeneity between studies in terms of the direction of effect.

Table 3.22: OS, PFS and RR data from the comparative observational studies (no crossover)

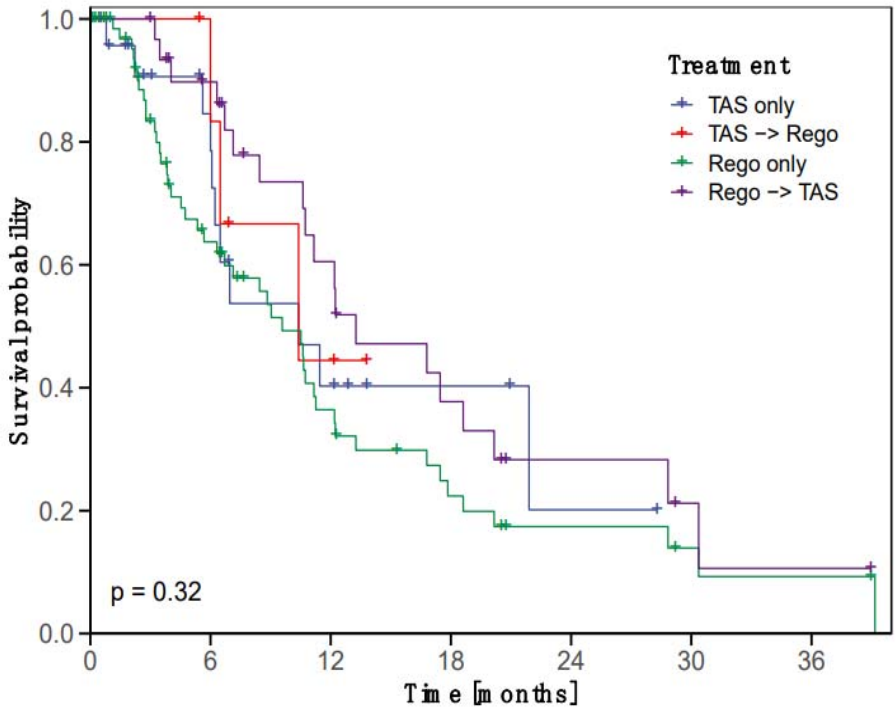
Efficacy	Nakashima 2020[1] No crossover		Tanaka 2018[23] All patients regardless of crossover		Sueda 2016[25] No crossover	
	Regorafenib (n=1,501)	T/T (n=3,777)	Regorafenib (n=20)	T/T (n=24)	Regorafenib (n=17)	T/T (n=6)

OS, median, months (95% CI)	6.4 (5.9-7.0)	10.2 (9.5-10.1)	9.1 (4.1 – 13.4)	9.3 (5.5-12.3)	4.5 (3.34-10.3)	5.3 (0.92-8.62)
	Adjusted HR 0.66 (p<0.001)					
PFS, median, months (95% CI)	-	-	2.1 (1.3 – 3.6)	3.1 (1.7-4.1)	3.0 (1.64-4.52)	2.1(0.92-6.39)
PR, %	-	-	0	0	0	0
SD, %	-	-	75.0	70.8	30.4	28.6
DCR, %	-	-	75.0	70.8	30.4	28.6

Source: Table 18, CS[3], and the original sources: Tanaka 2018[23], Sueda 2016[25]
CI = confidence interval; CS = company submission; DCR = Disease Control Rate; HR = Hazard ratio; OS = overall survival; PFS = progression-free survival; PR = partial response; RR = relative risk; SD = stable disease; T/T = trifluridine tipiracil

For Huemer 2020[26], the only outcome relevant to the NICE scope is OS. Numerical data are not reported per arm, and although the CS[3] reports an OS of 10.4 months for regorafenib, this is actually a combined value for both regorafenib and T/T groups. An informative KM graph is given in the supplement of the paper. In relation to this plot, the study reports that, "*These are adjusted survival curves using a cox proportional hazards model. Cross-over was taken into account as a time-dependent covariate beginning with the start of fourth line treatment in order to avoid immortal time bias*." This graph is shown in Figure 3.7, and does not seem to demonstrate any difference between regorafenib and T/T.

Figure 3.7: OS from initiation of third line therapy with regorafenib or TAS-102 according to therapy sequence in 93 mCRC patients



Source: Supplement S2 of Huemer 2020[26]
OS= overall survival; TAS-102 = trifluridine tipiracil; Rego = Regorafenib; mCRC = metastatic colorectal cancer

ERG comment:

Data relating to Tanaka 2018[23] and Sueda 2016[25] were unadjusted for confounding. The OS results for Huemer 2020[26] were adjusted for "immortal time bias" but there was no evidence they were adjusted for baseline covariates and so selection bias is still likely to exist. Therefore, given the imbalances in baseline characteristics (see Section 3.2.3.2), the results for these studies should be interpreted with caution. Although there was no adjustment for confounding in the patients who did not crossover, there seemed to be better balance in the baseline characteristics of Nakashima 2020.[1] However, what is unclear is what the treatment effect would be on patients who did crossover, but who might receive BSC in NHS clinical practice, particularly given evidence from two of the studies, Huemer 2020[26] and Tanaka 2018[23] of the relationship between crossover and ECOG PS.

3.2.5.6 Subgroup analyses

In both CORRECT and CONCUR, subgroup analyses of OS and PFS were pre-planned and were prespecified. Subgroup analyses were performed based on demographic information (e.g., race, sex, age group (<65 years, ≥65 years)), region (Region 1: North America, Western Europe, Israel, and Australia; Region 2; Asia; Region 3, South America, Turkey, and Eastern Europe), time from diagnosis of metastatic disease (≥18 months and <18 months), prior systemic anti-cancer therapies, historical *KRAS* mutation status, and baseline cancer characteristics of primary site of tumour (e.g., ECOG PS: 0 or 1). Subgroup analyses of AEs were also conducted.

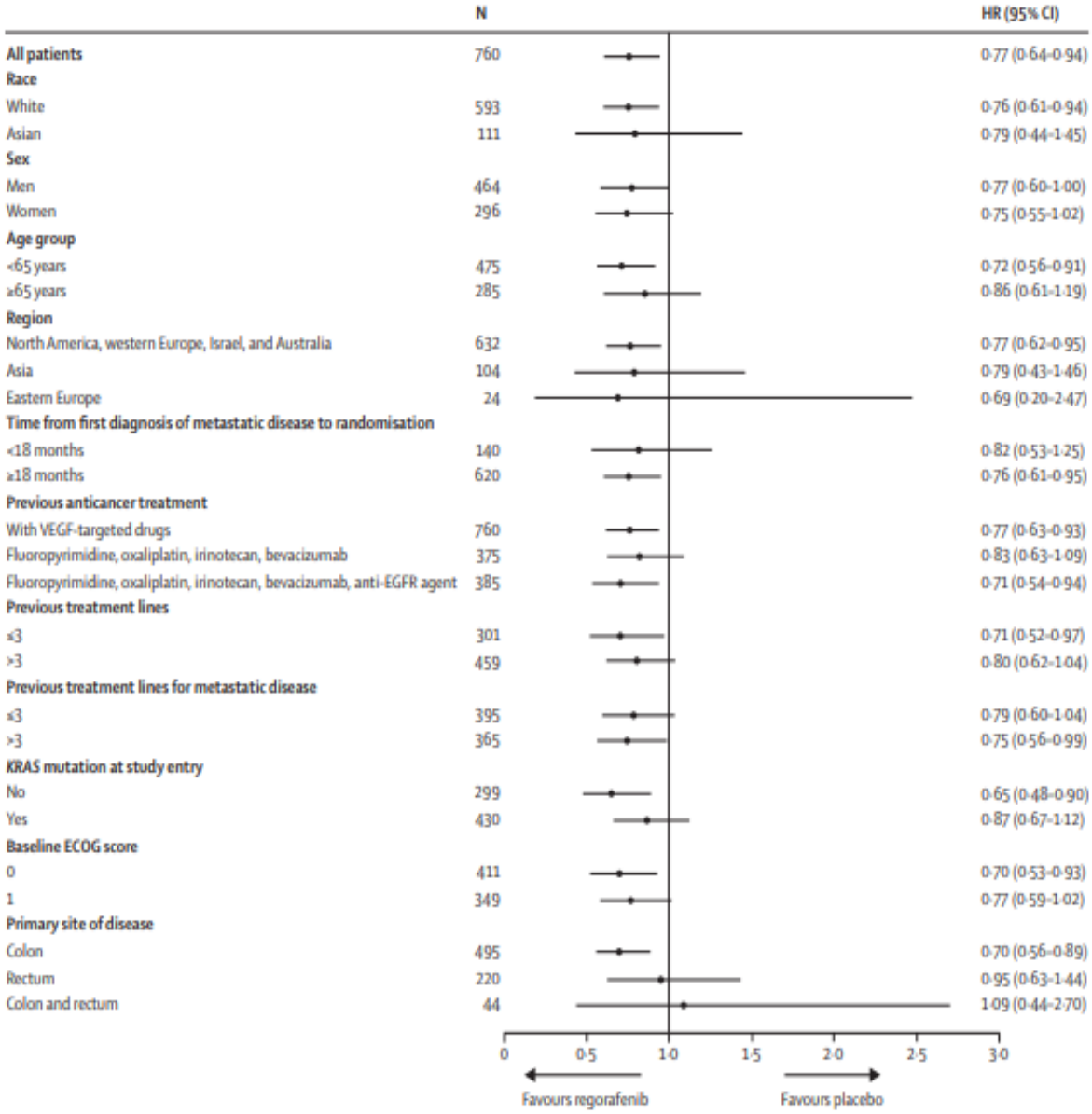
The company reported that overall, in both trials, the efficacy subgroup analyses demonstrated consistent survival benefits with regorafenib over placebo, with OS and PFS outcomes that were generally comparable with those observed in the overall populations. Subgroup analyses in both OS and PFS over the two trials are summarised in Figures 3.8 to 3.11. For TEAEs, subgroup analyses over the two trials are summarised in Tables 3.23 and 3.24.

ERG comment:

The CORRECT subgroup data appears to show that regorafenib may be less useful in terms of OS for rectal cancers, which is not specifically highlighted in the CS.[3] This is an important detail and may have implications for the applicability of regorafenib. However, it should be noted that this effect is not observed for PFS, with some evidence of improved efficacy of regorafenib in rectal cancers. In the CONCUR study no subgroup analysis was carried out for cancer location.

Other ERG comments on subgroup analyses have been covered in Section 3.2.3 on external validity.

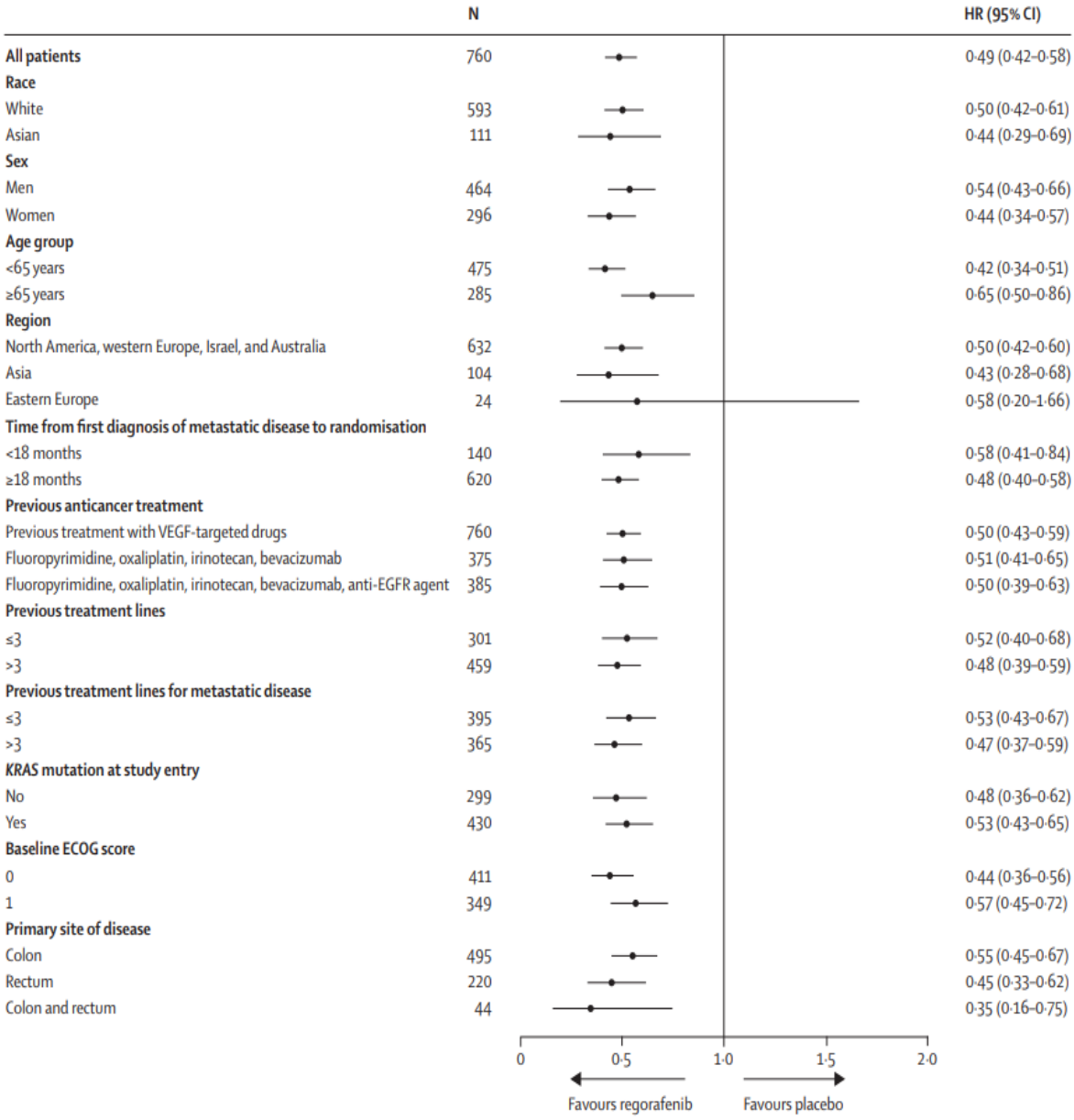
Figure 3.8: CORRECT – forest plot of subgroup analysis of OS (ITT)



Source: Figure 14, CS appendices[7]

CI = confidence interval; CS = company submission; ECOG = Eastern Cooperative Oncology Group; EGFR = epidermal growth factor receptor; KRAS = Kirsten rat sarcoma viral oncogene homologue; OS = overall survival; ITT = intention-to-treat; VEGF = vascular endothelial growth factor

Figure 3.9: CORRECT – forest plot of subgroup analysis of PFS (ITT)



Taken from Figure 15, CS appendices[7]

CI = confidence interval; CS = company submission; ECOG = Eastern Cooperative Oncology Group; EGFR = epidermal growth factor receptor; ITT = intention-to-treat; KRAS = Kirsten rat sarcoma viral oncogene homologue; PFS = progression-free survival; VEGF = vascular endothelial growth factor

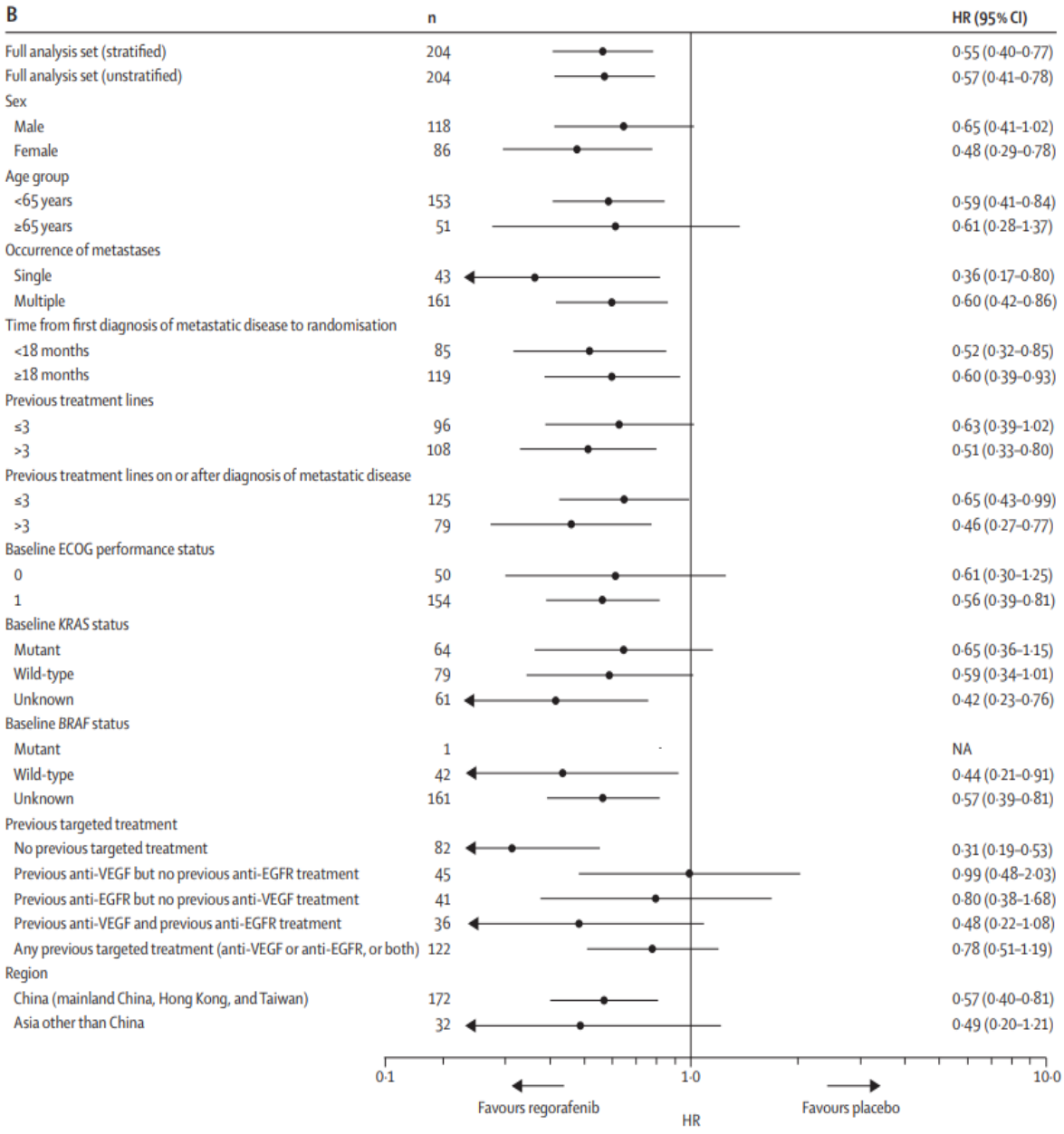
Table 3.23: CORRECT – summary of subgroup analyses of TEAEs (SAS)

Subgroup	Overall incidence, n (%)		Regorafenib/treatment-related, n (%)	
	Placebo	Regorafenib	Placebo	Regorafenib
Age				
< 65 years	157/164 (95.7)	305/307 (99.3)	94/164 (57.3)	288/307 (93.8)
≥ 65 years	88/89 (98.9)	193/193 (100.0)	60/89 (67.4)	177/193 (91.7)

Baseline BMI				
< 25 kg/m^2	103/108 (95.4)	253/254 (99.6)	59/108 (54.6)	235/254 (92.5)
25 to < 30 kg/m^2	103/108 (95.4)	253/254 (99.6)	59/108 (54.6)	235/254 (92.5)
≥ 30 kg/m^2	103/108 (95.4)	253/254 (99.6)	59/108 (54.6)	235/254 (92.5)
Sex				
Female	96/101 (95.0)	193/193 (100.0)	59/101 (58.4)	182/193 (94.3)
Male	149/152 (98.0)	305/307 (99.3)	95/152 (62.5)	283/307 (92.2)
Race				
White	195/200 (97.5)	388/389 (99.7)	122/200 (61.0)	359/389 (92.3)
Black	8/8 (100.0)	5/6 (83.3)	3/8 (37.5)	4/6 (66.7)
Asian	31/34 (91.2)	74/74 (100.0)	21/34 (61.8)	73/74 (98.6)
Other or not reported	11/11 (100.0)	31/31 (100.0)	8/11 (72.7)	29/31 (93.5)
Baseline ECOG score				
0	137/144 (95.1)	263/263 (100.0)	90/144 (62.5)	255/263 (97.0)
1	108/109 (99.1)	235/237 (99.2)	64/109 (58.7)	210/237 (88.6)
Baseline kidney function				
Moderately impaired kidney function (eGFR < 60 mL/min/1.73 m^2)	9/9 (100.0)	21/21 (100.0)	5/9 (55.6)	20/21 (95.2)
Normal/mildly impaired kidney function (eGFR ≥ 60 mL/min/1.73 m^2)	236/244 (96.7)	477/479 (99.6)	149/244 (61.1)	445/479 (92.9)
Baseline hepatic function				
Max. of baseline AST and ALT ≤ 1.5 x ULN	219/226 (96.9)	451/453 (99.6)	140/226 (61.9)	422/453 (93.2)
1.5 x ULN < max of Baseline AST and ALT ≤ 3 x ULN	22/23 (95.7)	44/44 (100.0)	11/23 (47.8)	40/44 (90.9)
3 x ULN < max of baseline AST and ALT	4/4 (100.0)	2/2 (100.0)	3/4 (75.0)	2/2 (100.0)

Source: Table 29, CS appendices[7]
ALT = alanine aminotransferase; AST = aspartate aminotransferase; BMI = body mass index; CS = company submission; ECOG = Eastern Cooperative Oncology Group; eGFR = estimated glomerular filtration rate; max = maximum; min = minutes; SAS = safety analysis set; TEAE = treatment-emergent adverse event; ULN = upper limit of normal

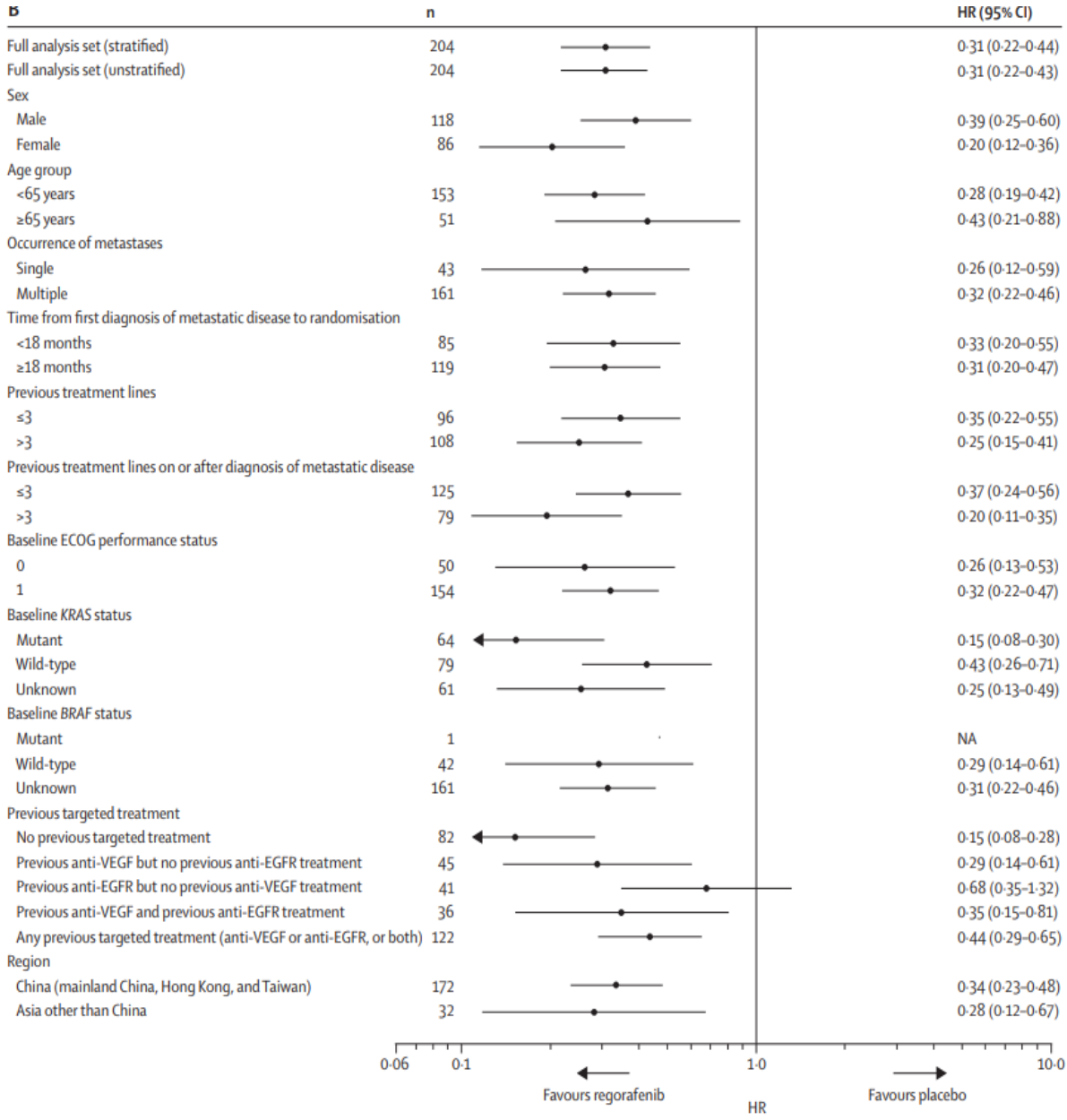
Figure 3.10: CONCUR – forest plot of subgroup analysis of OS (FAS)



Source: Figure 16, CS appendices[7]

CI = confidence interval; CS = company submission; BRAF = Mutation in the B-Raf proto-oncogene; CI = confidence interval; CS = company submission; ECOG = Eastern Cooperative Oncology Group; EGFR = epidermal growth factor receptor; FAS = full analysis set; KRAS = Kirsten rat sarcoma viral oncogene homologue; OS = overall survival; VEGF = vascular endothelial growth factor

Figure 3.11: CONCUR – forest plot of subgroup analysis of PFS (FAS)



Source: Figure 17, CS appendices[7]

CI = confidence interval; CS = company submission; BRAF = Mutation in the B-Raf proto-oncogene; ECOG = Eastern Cooperative Oncology Group; EGFR = epidermal growth factor receptor; HRs = hazard ratios; KRAS = Kirsten rat sarcoma viral oncogene homologue; VEGF = vascular endothelial growth factor

Note:

HRs and CIs were calculated using the unstratified Cox regression model for the subgroup analysis. Error bars are 95% CIs.

Table 3.24: CONCUR – summary of subgroup analyses of TEAEs (SAS)

Subgroup	Overall incidence, n (%)		Drug-related incidence, n (%)	
	Placebo	Regorafenib	Placebo	Regorafenib
Age				
< 65 years	51/58 (87.9)	95/95 (100)	24/58 (41.4)	94/95 (98.9)
≥ 65 years	9/10 (90.0)	41/41 (100)	7/10 (70.0)	38/41 (92.7)

Baseline BMI				
Missing	1/1 (100)	1/1 (100)	0/1 (0)	1/1 (100)
< 20 kg/m^2	16/18 (88.9)	20/20 (100)	7/18 (38.9)	20/20 (100)
20 to < 25 kg/m^2	31/33 (93.9)	74/74 (100)	16/33 (48.5)	71/74 (95.9)
25 to < 30 kg/m^2	12/16 (75.0)	38/38 (100)	8/16 (50.0)	38/38 (100)
≥ 30 kg/m^2	0/0 (0)	3/3 (100)	0/0 (0)	2/3 (66.7)
Sex				
Female	30/35 (85.7)	51/51 (100)	14/35 (40.0)	51/51 (100)
Male	30/33 (90.9)	85/85 (100)	17/33 (51.5)	81/85 (95.3)
Race				
Asian	60/68 (88.2)	136/136 (100)	31/68 (45.6)	132/136 (97.1)
Baseline ECOG score				
0	14/15 (93.3)	35/35 (100)	8/15 (53.3)	35/35 (100)
1	46/53 (86.8)	101/101 (100)	23/53 (43.4)	97/101 (96.0)
Baseline kidney function				
Missing	0/0 (0)	1/1 (100)	0/0 (0)	1/1 (100)
Impaired kidney function (eGFR < 60 mL/min/1.73 m^2)	0/0 (0)	2/2 (100)	0/0 (0)	2/2 (100)
Normal kidney function (eGFR ≥ 60 mL/min/1.73 m^2)	60/68 (88.2)	133/133 (100)	31/68 (45.6)	129/133 (97.0)
Baseline hepatic function				
Max. of baseline AST and ALT ≤ 1.5 x ULN	53/60 (88.3)	122/122 (100)	26/60 (43.3)	118/122 (96.7)
1.5 x ULN < max. of baseline AST and ALT ≤ 3 ULN	7/8 (87.5)	11/11 (100)	5/8 (62.5)	11/11 (100)
3 x ULN < max. of baseline AST and ALT	0/0 (0)	3/3 (100)	0/0 (0)	3/3 (100)

Source: Table 34, CS appendices[7]
ALT = alanine aminotransferase; AST = aspartate aminotransferase; BMI = body mass index; CS = company submission; ECOG = Eastern Cooperative Oncology Group; eGFR = estimated glomerular filtration rate; max = maximum; min = minutes; SAS = safety analysis set; TEAE = treatment-emergent adverse event; ULN = upper limit of normal

3.2.6 Adverse events of the included trials

3.2.6.1 Summary of adverse events

CORRECT

Table 3.25 presents an overview of the safety data up to the data cut-off date of 21 July 2011. Most patients in each group (regorafenib, 99.6%; placebo, 96.8%) experienced at least one TEAE, the majority of which were mild or moderate events. Serious AEs (SAEs) were reported at a similar rate in

both groups (regorafenib, 43.8%; placebo, 39.5%), and the incidence of treatment-emergent SAEs that were considered drug-related was slightly higher with regorafenib (11.8% versus 3.6%).

Although significantly more patients in the regorafenib arm had dose modifications because of AEs (66.6% versus 22.5%), the difference in the incidence of AEs leading to permanent treatment discontinuation was relatively small (17.6% versus 12.6%).

Table 3.25: CORRECT – overview of TEAEs (SAS)

Event	Any AE, n (%)		Drug-related AE, n (%)	
	Regorafenib (n=500)	**Placebo (n=253)**	**Regorafenib (n=500)**	**Placebo (n=253)**
TEAE	498 (99.6)	245 (96.8)	465 (93.0)	154 (60.9)
CTC Grade 1	[redacted]	[redacted]	[redacted]	[redacted]
CTC Grade 2	[redacted]	[redacted]	[redacted]	[redacted]
CTC Grade 3	[redacted]	[redacted]	[redacted]	[redacted]
CTC Grade 4	[redacted]	[redacted]	[redacted]	[redacted]
CTC Grade 5	[redacted]	[redacted]	[redacted]	[redacted]
Treatment emergent SAE	[redacted]	[redacted]	[redacted]	[redacted]
CTC Grade 1	[redacted]	[redacted]	[redacted]	[redacted]
CTC Grade 2	[redacted]	[redacted]	[redacted]	[redacted]
CTC Grade 3	[redacted]	[redacted]	[redacted]	[redacted]
CTC Grade 4	[redacted]	[redacted]	[redacted]	[redacted]
CTC Grade 5	[redacted]	[redacted]	[redacted]	[redacted]
AE leading to permanent discontinuation	[redacted]	[redacted]	[redacted]	[redacted]
AE leading to dose modification	[redacted]	[redacted]	[redacted]	[redacted]

Source: Table 19, CS[3]
AE = adverse event; CS = company submission; CTC = Common Toxicity Criteria; SAE = serious adverse event; SAS = safety analysis set; TEAE = treatment-emergent adverse event

CONCUR

Table 3.26 presents an overview of the safety data up to the data cut-off date of 29 November 2013. Most patients in each group (regorafenib, 100%; placebo, 88.2%) experienced at least one TEAE, the majority of which were mild or moderate events. Of these events, 97.1% and 45.6% in each respective treatment group were considered to be drug-related. SAEs were reported at a similar rate in both groups (regorafenib, 31.6%; placebo, 26.5%), and the incidence of treatment-emergent SAEs that were considered drug-related was higher with regorafenib (8.8% versus 3.4%).

Although more patients in the regorafenib arm had dose modifications because of AEs (71.3% versus 16.2%), the incidence of AEs leading to permanent treatment discontinuation was relatively small (14.0% versus 5.9%).

Table 3.26: CONCUR – overview of TEAEs (SAS)

Event	Any AE, n (%)		Drug-related AE, n (%)	
	Regorafenib (n = 136)	Placebo (n = 68)	Regorafenib (n = 136)	Placebo (n = 68)
TEAE	136 (100.0)	60 (88.2)	132 (97.1)	31 (45.6)
CTC Grade 1	[redacted]	[redacted]	[redacted]	[redacted]
CTC Grade 2	[redacted]	[redacted]	[redacted]	[redacted]
CTC Grade 3	[redacted]	[redacted]	[redacted]	[redacted]
CTC Grade 4	[redacted]	[redacted]	[redacted]	[redacted]
CTC Grade 5	[redacted]	[redacted]	[redacted]	[redacted]
Treatment-emergent SAE	[redacted]	[redacted]	[redacted]	[redacted]
CTC Grade 1	[redacted]	[redacted]	[redacted]	[redacted]
CTC Grade 2	[redacted]	[redacted]	[redacted]	[redacted]
CTC Grade 3	[redacted]	[redacted]	[redacted]	[redacted]
CTC Grade 4	[redacted]	[redacted]	[redacted]	[redacted]
CTC Grade 5	[redacted]	[redacted]	[redacted]	[redacted]
AE leading to permanent discontinuation	[redacted]	[redacted]	[redacted]	[redacted]
AE leading to dose modification	[redacted]	[redacted]	[redacted]	[redacted]

Source: Table 21, CS[3]
AE = adverse event; CS = company submission; CTC = Common Toxicity Criteria; SAE = serious adverse event; SAS = safety analysis set; TEAE = treatment-emergent adverse event
Note:
For patients experiencing the same AE several times, the AE has been counted only once by the worst severity grade

3.2.6.2 TEAEs

CORRECT

Table 3.27 presents TEAEs that occurred in ≥ 5% of patients in either treatment group. Grade 3 or 4 TEAEs occurred at a higher rate in the regorafenib group than in the placebo group (54% versus 14%). The most frequent Grade ≥ 3 regorafenib-related AEs (affecting ≥ 5% of patients) were hand–foot skin reaction (17%), fatigue (< 10%), diarrhoea (< 8%), hypertension (7%), and rash or desquamation (6%).

Table 3.27: CORRECT – TEAEs occurring in ≥ 5% of patients in either group from start of treatment to 30 days after end of treatment (SAS)

	Regorafenib (n=500)			Placebo (n=253)		
	Any Grade	Grade 3	Grade 4	Any Grade	Grade 3	Grade 4
Any event , n (%)	465 (93)	253 (51)	17 (3)	154 (61)	31 (12)	4 (2)
Clinical AE, n (%)						
Fatigue	237 (47)	46 (9)	2 (< 1)	71 (28)	12 (5)	1 (< 1)
Hand–foot skin reaction	233 (47)	83 (17)	0 (0)	19 (8)	1 (< 1)	0 (0)
Diarrhoea	169 (34)	35 (7)	1 (< 1)	21 (8)	2 (1)	0 (0)

	Regorafenib (n=500)			Placebo (n=253)		
	Any Grade	Grade 3	Grade 4	Any Grade	Grade 3	Grade 4
Anorexia	152 (30)	16 (3)	0 (0)	39 (15)	7 (3)	0 (0)
Voice changes	147 (29)	1 (< 1)	0 (0)	14 (6)	0 (0)	0 (0)
Hypertension	139 (28)	36 (7)	0 (0)	15 (6)	2 (1)	0 (0)
Oral mucositis	136 (27)	15 (3)	0 (0)	9 (4)	0 (0)	0 (0)
Rash or desquamation	130 (26)	29 (6)	0 (0)	10 (4)	0 (0)	0 (0)
Nausea	72 (14)	2 (< 1)	0 (0)	28 (11)	0 (0)	0 (0)
Weight loss	69 (14)	0 (0)	0 (0)	6 (2)	0 (0)	0 (0)
Fever	52 (10)	4 (1)	0 (0)	7 (3)	0 (0)	0 (0)
Constipation	42 (8)	0 (0)	0 (0)	12 (5)	0 (0)	0 (0)
Dry skin	39 (8)	0 (0)	0 (0)	7 (3)	0 (0)	0 (0)
Alopecia	36 (7)	0 (0)	0 (0)	1 (< 1)	0 (0)	0 (0)
Taste alteration	35 (7)	0 (0)	0 (0)	5 (2)	0 (0)	0 (0)
Vomiting	38 (8)	3 (1)	0 (0)	13 (5)	0 (0)	0 (0)
Sensory neuropathy	34 (7)	2 (< 1)	0 (0)	9 (4)	0 (0)	0 (0)
Nose bleed	36 (7)	0 (0)	0 (0)	5 (2)	0 (0)	0 (0)
Dyspnoea	28 (6)	1 (< 1)	0 (0)	4 (2)	0 (0)	0 (0)
Muscle pain	28 (6)	2 (< 1)	0 (0)	7 (3)	1 (< 1)	0 (0)
Headache	26 (5)	3 (1)	0 (0)	8 (3)	0 (0)	0 (0)
Pain, abdomen	25 (5)	1 (< 1)	0 (0)	10 (4)	0 (0)	0 (0)
Laboratory abnormalities, n (%)						
Thrombocytopenia	63 (13)	13 (3)	1 (< 1)	5 (2)	1 (< 1)	0 (0)
Hyperbilirubinemia	45 (9)	10 (2)	0 (0)	4 (2)	2 (1)	0 (0)
Proteinuria	35 (7)	7 (1)	0 (0)	4 (2)	1 (< 1)	0 (0)
Anaemia	33 (7)	12 (2)	2 (< 1)	6 (2)	0 (0)	0 (0)
Hypophosphatemia	25 (5)	19 (4)	0 (0)	1 (< 1)	1 (< 1)	0 (0)

Source: Table 20, CS[3]
AEs = adverse events; CS = company submission; SAS = safety analysis set; TEAE = treatment-related adverse events
Notes:
Data cut-off date 21 July 2011

CONCUR

Table 3.28 presents TEAEs that occurred in ≥ 10% of patients in either treatment group. The most frequent AEs of Grade ≥ 3 associated with regorafenib were hand–foot skin reaction (16%), hypertension (11%), hyperbilirubinemia, hypophosphatemia, and alanine aminotransferase (ALT) concentration increases (7% each).

Table 3.28: CONCUR – TEAEs occurring at any Grade in ≥ 10% of patients, or at Grade ≥ 3 in any patients in either group, from the start of treatment to 30 days after the end of treatment (SAS)

	Regorafenib (n=136)				Placebo (n=68)			
AE, n (%)[a]	**Grade 1–2**	**Grade 3**	**Grade 4**	**Grade 5**	**Grade 1–2**	**Grade 3**	**Grade 4**	**Grade 5**
Any event	58 (43)	67 (49)	5 (4)	2 (1)	21 (31)	9 (13)	1 (1)	0 (0)
Hand–foot skin reaction	78 (57)	22 (16%)	NA	NA	3 (4)	0 (0)	NA	NA
Hyperbilirubinemia	41 (30)	6 (4)	3 (2)	NA	4 (6)	1 (1)	0 (0)	NA
Alanine aminotransferase conc. increased	23 (17)	9 (7)	0 (0)	NA	5 (7)	0 (0)	0 (0)	NA
Aspartate aminotransferase conc. increased	24 (18)	7 (5)	1 (1)	NA	6 (9)	0 (0)	0 (0)	NA
Hypertension	16 (12)	15 (11)	0 (0)	0 (0)	1 (1)	2 (3)	0 (0)	0 (0)
Hoarseness	27 (20)	1 (1%)	NA	NA	0 (0)	0 (0)	NA	NA
Diarrhoea	23 (17)	1 (1)	0 (0)	0 (0)	1 (1)	1 (1)	0 (0)	0 (0)
Fatigue	19 (14)	4 (3)	NA	NA	4 (6)	1 (1)	NA	NA
Thrombocytopenia	9 (7)	3 (2)	1 (1)	N/A	1 (1)	0 (0)	0 (0)	NA
Hypophosphatemia	4 (3)	9 (7)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Proteinuria	11 (8)	2 (1)	NA	NA	0 (0)	1 (1)	NA	NA
Maculopapular rash	6 (4)	6 (4)	NA	NA	1 (1)	0 (0)	NA	NA
Leucopenia	8 (6)	3 (2)	0 (0)	NA	0 (0)	0 (0)	0 (0)	NA
Anorexia	9 (7)	1 (1)	0 (0)	0 (0)	3 (4)	0 (0)	0 (0)	0 (0)
Lipase conc. increased	3 (2)	6 (4)	0 (0)	NA	3 (4)	1 (1)	0 (0)	NA
Neutropenia	4 (3)	3 (2)	0 (0)	NA	0 (0)	0 (0)	0 (0)	NA
Myalgia	6 (4)	1 (1)	NA	NA	0 (0)	0 (0)	NA	NA
Abdominal pain	5 (4)	1 (1)	NA	NA	3 (4)	0 (0)	NA	NA
Anaemia	3 (2)	1 (1)	1 (1)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Other investigations[b]	3 (2)	1 (1)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Other skin/subcutaneous tissue disorders	3 (2)	1 (1)	0 (0)	0 (0)	1 (1)	0 (0)	0 (0)	0 (0)
Alkaline phosphatase conc. increased	3 (2)	0 (0)	0 (0)	NA	0 (0)	1 (1)	0 (0)	NA
Hypoalbuminemia	2 (1)	1 (1)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Hypokalaemia	2 (1)	1 (1)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Visceral arterial ischaemia	0 (0)	1 (1)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
γ glutamyl transferase conc. increased	1 (1)	1 (1)	0 (0)	NA	0 (0)	0 (0)	0 (0)	NA
Pharyngitis	1 (1)	1 (1)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)

	Regorafenib (n=136)				Placebo (n=68)			
AE, n (%)[a]	**Grade 1–2**	**Grade 3**	**Grade 4**	**Grade 5**	**Grade 1–2**	**Grade 3**	**Grade 4**	**Grade 5**
Atrial fibrillation	1 (1)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (1)	0 (0)
Cardiac arrest	NA	NA	0 (0)	1 (1)	NA	NA	0 (0)	0 (0)
Oesophageal varices haemorrhage	0 (0)	1 (1)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Death not otherwise specified	NA	NA	NA	1 (1)	NA	NA	NA	0 (0)
Serum amylase conc. increased	1 (1)	0 (0)	0 (0)	N/A	0 (0)	1 (1)	0 (0)	NA
Wound infection	0 (0)	1 (1)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Flank pain	0 (0)	1 (1)	N/A	N/A	0 (0)	0 (0)	N/A	N/A
Vaginal fistula	0 (0)	1 (1)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Conduction disorder	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (1%)	0 (0)	0 (0)
Heart failure	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (1)	0 (0)
Acute kidney injury	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (1)	0 (0)
Other vascular disorders	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (1)	0 (0)	0 (0)

Source: Table 22, CS[3]
AE = adverse event; CS = company submission; NA = not applicable; SAS = safety analysis set; TEAE = treatment-related adverse events.
Notes:
Data in each column show the number of patients experiencing that grade as their worst severity of the relevant AE.
[a] For patients with more than one AE, only the highest grade of the most severe event is shown
[b] Laboratory or diagnostic tests or clinical assessments

3.2.6.4 AEs leading to discontinuations and dose modifications

CORRECT

Rates of permanent discontinuations due to TEAEs were as follows: regorafenib [redacted]%; placebo [redacted]%. Overall, more patients receiving regorafenib had AEs that led to dose modifications than those on placebo (66.6% versus 22.5%). Of these, dose reductions occurred in 38% versus 3%, respectively, and dose interruption in 61% versus 22%, respectively. The most frequent AEs necessitating dose modification were dermatological, gastrointestinal, constitutional, and metabolic or laboratory events.

CONCUR

Rates of permanent discontinuations due to TEAEs were as follows: regorafenib, 14.0%; placebo, 5.9%. Overall, more patients receiving regorafenib had AEs that led to dose modifications than those on placebo (75.0% versus 22.1%). Of these, dose reductions occurred in 39.7% and 0% of patients, and dose interruptions occurred in 62.5% and 16.2% of patients, respectively.

3.2.6.5 Serious adverse effects and deaths

CORRECT

There was a similar incidence of SAEs between groups (regorafenib, 43.8%; placebo, 39.5%). The differences regarding Grade 4 and 5 treatment-emergent SAEs between the two treatment groups was small and clinically not relevant. Grade 3 treatment-emergent SAEs were reported at a higher incidence

in the regorafenib group than in placebo (18.2% versus 13.8%), although the incidences of Grade 3 and 4 SAEs were similar between groups.

Overall, there were 110 deaths (regorafenib, 13.8% (n=69); placebo, 16.2% (n=41)) reported during the study (i.e., up to within 30 days of last dose). The most common reason for death was PD (regorafenib, 11.6%; placebo, 13.8%); other reasons were reported as due to an AE not associated with clinical disease progression (1.6% versus 1.2%), unknown cause (0.4% versus 0.4%), and other cause (0.8% versus 0.2%). In the regorafenib group, the AEs not associated with disease progression that contributed to death were: pneumonia (n=2), gastrointestinal bleeding (n=2), intestinal obstruction (n=1), pulmonary haemorrhage (n=1), seizure (n=1) and sudden death (n=1). In the placebo group, these AEs were pneumonia (n=2) and sudden death (n=1). Occurrence of thromboembolism did not differ between groups (2% for both groups).

CONCUR

There was a similar incidence of SAEs between groups (regorafenib, 31.6%; placebo, 26.5%). The differences regarding Grade 4 and 5 treatment-emergent SAEs between the two treatment groups was small and clinically not relevant. Grade 3 treatment-emergent SAEs were reported at the same incidence in both groups (11.8%). Most treatment-emergent SAEs were not related to study drug treatment.

Overall, there were 19 deaths reported during the study (regorafenib, n=12 (8.8%); placebo, n=7 (8.8%)). One additional patient in the placebo group died during the follow-up period. Fourteen of these cases were due to progression of underlying disease (regorafenib, n=8; placebo, n=6). In the regorafenib group, the deaths of two (1%) patients were deemed to be drug-related within 30 days after the last dose. Brief narratives of the two patients are as follows:

> The first patient was a 65-year-old woman who stopped regorafenib treatment during her first cycle as a result of a non-serious Grade 2 increase in bilirubin. One week after stopping treatment, she collapsed at home and had a cardiac arrest.
>
> The second patient was a 67-year-old man who received regorafenib for 2 days. On the next day, he had a Grade 4 cardiac arrest, resulting in admission to hospital and death.

3.2.6.7 Adverse events of special interest (AEOSI)

CORRECT

Liver dysfunction

The incidence of liver dysfunction was as follows: regorafenib, [redacted]%; placebo, [redacted]%. More patients in the regorafenib group than placebo had AEs of liver dysfunction that resulted in fatal outcomes: [redacted] patients versus [redacted] patients. Three events of liver dysfunction that had a fatal outcome in the regorafenib group were assessed by the treating investigator as related to study drug; for two of these patients, the cause of death as stated by the investigator was disease progression.

Cardiac ischaemia/infarction and bleeding events

The incidence of cardiac ischaemia/infarction was as follows: regorafenib, [redacted]%; placebo, [redacted]%. In most of the cases, patients had existing cardiovascular risk factors. In both groups, there was [redacted] death reported, which was not assessed as drug related. Regarding bleeding events, the incidence was higher in the regorafenib group than in the placebo group (regorafenib, [redacted]%; placebo, [redacted]%); however, the majority of events were Grade 1 nose bleeds in the regorafenib arm. The incidence of serious bleeding

was as follows: regorafenib, [redacted]%; placebo, [redacted]%. In total, [redacted]_([redacted]%) patients in the regorafenib group had a bleeding event resulting in death, while no deaths due to bleeding were reported in the placebo group.

Hand–foot skin reaction (palmar–plantar erythrodysesthesia)

The incidence of hand–foot skin reaction was higher in the regorafenib group than in the placebo group (regorafenib, [redacted]%; placebo, [redacted]%). [redacted]% of patients in the regorafenib group had Grade 3 events, compared with [redacted]% in the placebo group. The incidence of drug-related hand–foot skin reaction events was notably higher in the regorafenib group compared with the placebo group ([redacted]% versus [redacted]%). The incidence of SAEs was as follows: [redacted]% versus [redacted]%, respectively. Permanent discontinuation of treatment due to hand–foot skin reaction was as follows: regorafenib [redacted]%; placebo [redacted]%, with most reactions being managed by dose reductions or interruptions.

Rash

The incidence of rash was higher in the regorafenib group than in the placebo group (regorafenib, [redacted]%; placebo, [redacted]%). The majority of these events were Grade 1 in both treatment groups. The incidence of Grade 3 events was as follows: regorafenib [redacted]%; placebo [redacted]%. Rash TEAEs could usually be managed by dose reductions or interruptions, and these TEAEs led to permanent discontinuation of treatment in [redacted] ([redacted]%) regorafenib-treated patients and [redacted] placebo-treated patients.

Renal failure

The incidence of renal failure was as follows: regorafenib [redacted]%; placebo [redacted]%. Most events were Grade 3, with [redacted] Grade 4 AE (in the placebo group) and [redacted] Grade 5 AE (in the regorafenib group) reported. TEAEs of renal failure that were assessed as related to treatment were reported at the same incidence in regorafenib-treated patients and placebo-treated patients ([redacted]% in each group).

Proteinuria

The incidence of proteinuria was as follows: regorafenib, [redacted]%; placebo, [redacted]%. Grade 3 events were reported for [redacted]% and [redacted]%, respectively, in each group. No Grade 4 or Grade 5 events were reported.

CONCUR

Acute liver failure

There were no reports of hepatic failure, hepatic necrosis or Grade 2–4 drug-related hepatobiliary disorder AEs in either treatment group. Overall, treatment-emergent hepatobiliary/pancreas SAEs (any Grade) were reported with the following incidence: regorafenib [redacted]%; placebo, [redacted]%. In the regorafenib group, there was one Grade 4 AE of bile duct stenosis that led to permanent study discontinuation. There were no deaths in either treatment group that resulted from liver dysfunction. No cases of significant transaminase increase, or severe drug-induced liver injury have been identified from the ongoing hepatotoxicity monitoring from this study.

Acute cardiac failure and Grade ≥ 3 bleeding events

There was one report of heart failure in the placebo group and one report of cardiac arrest in the regorafenib group. There were no reports of cardiac ischaemia/infarction-related AEs. Incidence of cardiac disorders was as follows: regorafenib, [redacted]%; placebo, [redacted]%).

Regarding Grade ≥3 bleeding events, the overall incidence was higher in the regorafenib group compared with the placebo group ([redacted]% versus [redacted]%); however, the majority of the bleeding events in the regorafenib group were Grade 1 or 2 anaemia. Serious bleeding AEs were only reported in the regorafenib group ([redacted] events). No Grade 5 bleeding events were reported, and there were [redacted] haemorrhage/bleeding events that were the cause of permanent discontinuation of study medication in either treatment group.

Hand–foot skin reaction (palmar–plantar erythrodysesthesia)

The incidence of hand–foot skin reaction was higher in the regorafenib group than placebo (regorafenib, [redacted]%; placebo, [redacted]%). The majority of these events were Grade 1 or 2, although Grade 3 events were reported in the regorafenib group ([redacted]% versus [redacted]%, respectively). There were [redacted] SAEs of hand–foot skin reaction in either treatment group. These events led to the permanent discontinuation of treatment in only one (0.7%) regorafenib-treated patient and no placebo-treated patients.

Rash

The incidence of rash was [redacted]; placebo, [redacted]. The majority of events were Grade 1. No TEAEs related to rash led to permanent discontinuation in either treatment group. SAEs of rash were as follows: regorafenib, [redacted]%; placebo, [redacted]%, and all were reported as Grade 3 events and were considered related to the study drug.

Acute renal failure (any grade) or severe proteinuria (Grade ≥ 3)

The incidence of proteinuria was [redacted].

ERG comment:

No pooling across CORRECT and CONCUR was carried out for this outcome and there was no attempt to compare to T/T. This directly conflicts with the main stated rationale for the submission, which was to evaluate a drug with a different adverse effects profile to T/T.

3.2.6.7 Additional data from REDOs and REARRANGE

REDOs and REARRANGE were not included in the company's SLR as they do not fit the decision problem, but because they both involve regorafenib the ERG asked the company to include safety data from these studies in the submission. The data that follows was submitted by the company as part of the clarification response.[6]

REDOS

ReDOS was a Phase II dose-escalation study of regorafenib in patients with mCRC which evaluated two dosing strategies. The primary endpoint was the proportion of patients in each arm starting Cycle 3 of treatment.

The AEs from REDOS are presented in Table 3.29.

Table 3.29: REDOS – adverse events

	Dose escalation group (n=54)				Standard dose group (n=62)			
Adverse event	**Grade 1-2**	**Grade 3**	**Grade 4**	**Grade 5**	**Grade 1-2**	**Grade 3**	**Grade 4**	**Grade 5**
Fatigue	42(72%)	7(13%)	0	0	44(71%)	11(18%)	0	0
Hand foot skin reaction	27(50%)	8(15%)	0	0	33(53%)	10(16%)	0	0
Hypertension	34(63%)	4(7%)	0	0	29(47%)	9(15%)	0	0
Nausea	23(43%)	0	0	0	31(50%)	0	0	0
Diarrhoea	23(43%)	1(2%)	0	0	25(40%)	2(3%)	0	0
Anorexia	14(26%)	1(2%)	0	0	16(26%)	2(3%)	0	0
Rash maculopapular	10(19%)	0	0	0	16(26%)	3(5%)	0	0
Vomiting	13(24%)	0	0	0	14(23%)	1(2%)	0	0
Blood bilirubin increased	7(13%)	2(4%)	0	0	13(21%)	5(8%)	0	0
Anaemia	12(22%)	1(2%)	0	0	12(19%)	1(2%)	0	0
Aspartate aminotransferase increased	8(15%)	1(2%)	0	0	12(19%)	4(6%)	0	0
Alkaline phosphatase increased	6(11%)	3(6%)	0	0	10(16%)	1(2%)	1(2%)	0
Abdominal pain	1(2%)	9(17%)	0	0	5(8%)	4(6%)	0	0
Dyspnoea	5(9%)	1(2%)	1(2%)	0	8(13%)	4(6%)	0	0
Alanine aminotransferase increased	8(15%)	0	0	0	8(13%)	1(2%)	0	0
Hoarseness	8(15%)	0	0	0	8(13%)	0	0	0
Weight loss	4(7%)	1(2%)	0	0	10(16%)	1(2%)	0	0
Hyponatremia	0	2(4%)	1(2%)	0	7(11%)	4(6%)	1(2%)	0
Platelet count decreased	7(13%)	0	0	0	8(13%)	0	0	0
Mucositis oral	4(7%)	1(2%)	0	0	8(13%)	1(2%)	0	0
Hypoalbuminemia	5(9%)	1(2%)	0	0	7(11%)	0	0	0

	Dose escalation group (n=54)				Standard dose group (n=62)			
Adverse event	**Grade 1-2**	**Grade 3**	**Grade 4**	**Grade 5**	**Grade 1-2**	**Grade 3**	**Grade 4**	**Grade 5**
Peripheral sensory neuropathy	6(11%)	0	0	0	6(10%)	0	0	0
Lymphocyte count decreased	1(2%)	4(7%)	0	0	6(10%)	0	0	0
Hypocalcaemia	6(11%)	0	0	0	3(5%)	1(2%)	0	0
Hypokalaemia	3(6%)	1(2%)	0	0	5(8%)	0	1(2%)	0
Generalised muscle weakness	5(9%)	1(2%)	0	0	2(3%)	1(2%)	0	0
Myalgia	0	1(2%)	0	0	6(10%)	2(3%)	0	0
Pain	5(9%)	0	0	0	3(5%)	1(2%)	0	0
Dehydration	1(2%)	0	0	0	2(3%)	5(8%)	0	0
Investigations, other (specified)	3(6%)	0	0	0	4(6%)	1(2%)	0	0
Back pain	1(2%)	1(2%)	0	0	5(8%)	0	0	0
Dry skin	1(2%)	1(2%)	0	0	3(5%)	0	0	0
Neoplasms: benign, malignant, unspecified, other (specified)	0	0	0	2(4%)	0	0	0	2(3%)
Colonic obstruction	0	3(6%)	0	0	0	0	0	0
Hyperglycaemia	1(2%)	1(2%)	0	0	0	1(2%)	0	0
Hyperkalaemia	1(2%)	0	0	0	1(2%)	1(2%)	0	0
Sinus tachycardia	0	1(2%)	0	0	1(2%)	1(2%)	0	0
Ascites	1(2%)	1(2%)	0	0	0	0	0	0
Chest wall pain	0	1(2%)	0	0	1(2%)	0	0	0
Death not otherwise specified	0	0	0	1(2%)	0	0	0	1(2%)
Encephalopathy	0	0	0	0	0	2(3%)	0	0
Respiratory failure	0	0	1(2%)	0	0	0	0	1(2%)
Sepsis	0	0	1(2%)	0	0	0	1(2%)	0
Thromboembolic event	1(2%)	1(2%)	0	0	0	0	0	0

	Dose escalation group (n=54)				Standard dose group (n=62)			
Adverse event	**Grade 1-2**	**Grade 3**	**Grade 4**	**Grade 5**	**Grade 1-2**	**Grade 3**	**Grade 4**	**Grade 5**
Abdominal infection	0	1(2%)	0	0	0	0	0	0
Adult Respiratory Distress Syndrome	0	0	1(2%)	0	0	0	0	0
Atelectasis	0	0	0	0	0	1(2%)	0	0
Colitis	0	0	0	0	0	0	0	0
Confusion	0	1(2%)	0	0	0	1(2%)	0	0
Hepatic failure	0	0	0	0	0	0	1(2%)	0
Hepatobiliary disorders	0	1(2%)	0	0	0	0	0	0
Hypoxia	0	0	0	0	0	1(2%)	0	0
Infections and infestations, other	0	0	0	0	0	1(2%)	0	0
Increased international normalised ratio	0	0	0	0	0	1(2%)	0	0
Lower gastrointestinal haemorrhage	0	0	0	0	0	1(2%)	0	0
Lung infection	0	0	1(2%)1(2%)	0	0	0	0	0
Myocardial infarction	0	0	0	0	0	0	0	0
Rectal fistula	0	1(2%)	0	0	0	0	0	1(2%)
Rectal obstruction	0	1(2%)	0	0	0	0	0	0
Syncope	0	1(2%)	0	0	0	0	0	0
Urinary retention	0	1(2%)	0	0	0	0	0	0

Source: Bekaii-Saab TS et al. Lancet Oncol 2019;20:1070-82 (http://dx.doi.org/10.1016/)[29], in response to request for clarification document[6]
All Grade 3, 4 and 5 events are shown as well as Grade 1 and 2 occurrence of these events. For other Grade 1-2 events, only events occurring in at least 10% of patients are included.

REARRANGE

REARRANGE investigated the impact of initial flexible dosing on regorafenib tolerability.

Refractory mCRC patients were randomised 1:1:1 to standard dose 160 mg/day 3 weeks on 1 week off (SD), reduced dose 120 mg/day 3 weeks on 1 week off (RD) or intermittent dose 160 mg 1 week on 1 week off (ID). Patients in RD or ID escalated to SD after cycle 1 if no limiting toxicity occurred. Primary endpoint was percentage of patients with Grade 3-4 TRAE in each arm.[30]

From July 2016 to September 2017, 299 patients were randomised. Safety population set was: 100 SD, 98 RD, 99 ID. Median number of prior lines and age were 4 and 64 years. Percentage of patients with Grade 3-4 AE were: 60 SD, 56 RD, 55 ID.[30]

3.2.6.8 Comparative observational studies

The comparative observational study Nakashima 2020[1] presented a range of AEs across regorafenib and T/T, which are summarised in Table 3.30.

Table 3.30: AEs in regorafenib and T/T in the non-randomised studies

Adverse effects	Nakashima 2020[1]	
	Regorafenib (n=1,501)	T/T (n=3,777)
Any AEs	777(52%)	1,622(43%)
Hand-foot syndrome	257(17%)	182(5%)
Peripheral neuropathy	114(8%)	290(8%)
Hypertension	287(19%)	446(12%)
Nausea	127(8%)	371(10%)
Diarrhoea	116(8%)	249(7%)
Oral mucositis	119(8%)	167(4%)
Rash/desquamation	73(5%)	56(1%)
Fever	44(3%)	117(3%)
Hepatotoxicity	20(1%)	9(0%)
Fatigue	14(1%)	31(1%)
Leukopenia	33(2%)	597(16%)
Interstitial pneumonitis	8(1%)	12(0%)
AEs = adverse events; T/T = trifluridine/tipiracil		

Tanaka 2018 and Sueda 2016 both reported an array of AEs that were similar to each other for regorafenib and T/T, which are presented in Table 3.31.

Table 3.31: AEs in regorafenib and T/T in the non-randomised studies

Adverse effects	Tanaka 2018		Sueda 2016	
	Regorafenib (n=20)	T/T (n=24)	Regorafenib (n = 23)	T/T (n=14)
Any adverse event	100	92	100	92.9
Hand-foot skin reaction	70	0	60.9	7.1
Hypertension	45	0	47.8	0
Leukopenia	0	50		
Neutropenia	0	54	0	14.3
Anaemia	0	63	0	0
Source: Tanaka 2018[23], Sueda 2016[25] AE = adverse event; T/T = trifluridine/tipiracil				

ERG comment:

These results suggest a greater burden from regorafenib than T/T in terms of hand-foot skin reactions and hypertension, but from T/T in haematological events, including leukopenia neutropenia and possibly anaemia.

3.2.7 Included studies: supporting evidence

Not applicable.

3.2.8 Ongoing studies

Not applicable.

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

CORRECT and CONCUR compared regorafenib to placebo, but the company's choice of comparator was T/T (see Sections 2.1 and 2.3). In order to allow regorafenib to be compared to T/T, RCTs evaluating T/T versus placebo were also evaluated, so that a NMA between regorafenib and T/T could be undertaken. Note that the term ITC has sometimes been used given that no head-to-head comparison was included in the NMA.

The company considered the RECOURSE, TERRA and Yoshino studies to be the only randomised studies identified by the clinical SLR to explore the effectiveness and safety of T/T for the treatment of adult patients with mCRC receiving ≥ third line therapy.

Therefore, five RCTs were evaluated in total: two evaluating regorafenib versus placebo (CORRECT, CONCUR), and three evaluating T/T versus placebo (RECOURSE, TERRA and Yoshino).

Two of the five randomised trials included in the ITC which evaluated regorafenib versus placebo - CORRECT[14] and CONCUR[15] - have already been critiqued in Section 3.2. This Section will focus on the remaining RCTs which evaluated T/T versus placebo: RECOURCE,[16] TERRA[17] and Yoshino 2012[18] (Table 3.32).

Very little information is given in the CS[3] about the three RCTs evaluating T/T versus placebo, and so the information summarised in Table 3.33 has been taken from the primary sources.[16-18].

Information on the statistical analyses used for the main analyses for the RECOURSE, TERRA and Yoshini 2012 trials were not available in the CS[3] and so these have been taken from the primary sources. Table 3.34 summarises this information.

Baseline data for the RECOURSE, TERRA and Yoshini 2012 trials were derived from the appendices of the CS[7]. Table 3.35 summarises these data. Nine baseline characteristics were considered by the company and a number of differences were highlighted, including:

- There were higher proportions of patients < 65 years in CONCUR and TERRA compared with the rest of the studies.
- There was a higher proportion of males in the intervention arm compared with the placebo arm in CONCUR and Yoshino 2012.
- All patients in CORRECT and most patients in RECOURSE received the targeted biological treatment bevacizumab, while in CONCUR, TERRA, and Yoshino 2012 a large proportion of patients had not received a prior targeted biological treatment.

- CONCUR, TERRA and Yoshino 2012 only included Asian patients, while CORRECT and RECOURSE recruited patients from across the world.
- Patients in CONCUR and TERRA had a shorter median time since the diagnosis of first metastases compared with patients in CORRECT.
- CONCUR and TERRA had a smaller proportion of patients in ECOG PS 0 relative to the other studies.

The company acknowledged that as "*…bevacizumab is not approved by NICE for the treatment of mCRC, the populations of CORRECT and RECOURSE may not represent the population that would potentially be treated for mCRC in the UK. CONCUR, TERRA, and Yoshino 2012 may better represent the UK population*." (page 18)[7]. Furthermore, previous treatments lines were not reported in a standardised way across the studies, and it was not feasible to compare and assess prior lines of treatment on or after diagnosis of metastases. Differences in patients baseline characteristics were considered important by the company only if the characteristic was identified as a potential treatment effect modifier (see following Section 3.4).

Table 3.32: Studies included in the ITC analyses

Study	Year	Trial design	Population	Treatment	OS data availability	PFS data availability
CORRECT	2013	RCT	ITT	Regorafenib, placebo	PLD	PLD
CONCUR	2015	RCT	ITT	Regorafenib, placebo	PLD	PLD
RECOURSE	2015	RCT	ITT	Trifluridine/tipiracil, placebo	KM curve	KM curve
Yoshino 2012	2012	RCT	ITT	Trifluridine/tipiracil, placebo	KM curve	KM curve
TERRA	2017	RCT	ITT	Trifluridine/tipiracil, placebo	KM curve	KM curve

Source: Table 6 of Appendix D[7]
ITC = indirect treatment comparison; ITT = Intention To Treat; KM = Kaplan–Meier; OS = overall survival; PFS = progression-free survival; PLD = patient level data; RCT = randomised controlled trial

Table 3.33: Study methodology for RECOURSE, TERRA and Yoshino 2012

Trial number (acronym)	RECOURSE (NCT01607957) Mayer 2015[16]	TERRA (NCT01955837) Xu 2018[17]	Yoshino 2012[18]
Location	Japan, USA, Europe, Australia	China, the Republic of Korea, and Thailand	Japan
Trial design	Multicentre, double-blind phase III placebo-controlled RCT	Multicentre, double-blind, placebo-controlled, phase III RCT	Multicentre, double-blind, placebo-controlled phase-2 RCT
Eligibility criteria for participants	**Inclusion criteria:** Adenocarcinoma of the colon or rectum; received at least two previous courses of standard chemotherapies; knowledge of KRAS status; received chemotherapy with	**Inclusion criteria**: Patients ≥ 18 years old with histologically or cytologically confirmed adenocarcinoma of the colon or rectum and known KRAS status who	**Inclusion criteria:** 20 years or older; histologically or cytologically confirmed unresectable metastatic colorectal adenocarcinoma; previous

Trial number (acronym)	**RECOURSE** (NCT01607957) Mayer 2015[16]	**TERRA** (NCT01955837) Xu 2018[17]	**Yoshino 2012[18]**
	each of the following agents: a fluoropyrimidine, oxaliplatin, irinotecan, bevacizumab, and (if KRAS wild type tumours) cetuximab or panitumumab; 18 years of age or older; adequate bone-marrow, liver, and renal function; ECOG PS of 0 or 1	were refractory or intolerant to two or more prior chemotherapy regimens	treatment history of two or more regimens of standard chemotherapy; refractory or intolerant to a fluoropyrimidine, irinotecan, and oxaliplatin; able to take oral drugs; measurable lesions; ECOG PS of between 0 and 2. Adequate bone-marrow, hepatic, and renal functions. **Exclusion criteria:** Serious comorbidities
Settings and locations where the data were collected	Japan, USA, Europe, Australia	30 sites in China, the Republic of Korea, and Thailand	Japan
Study periods and trial drugs	Trifluridine/tipiracil (T/T) versus placebo. Study drug or placebo was administered twice daily, after food, 5 days a week, with 2 days of rest, for 2 weeks, followed by a 14-day rest period, thus completing one treatment cycle. The regimen was repeated every 4 weeks.	T/T versus placebo. Participants were randomly assigned to receive T/T (twice per day orally; 5 days on and 2 days off for 2 weeks, followed by 14 days off per cycle) or placebo.	T/T versus placebo. 35 mg/m² of the study drug taken orally twice a day after food. Taken in a 28-day cycle: 2-week cycle of 5 days of treatment followed by a 2-day rest period, and then a 14-day rest period. In patients who had adverse events, the dose could be reduced by 10 mg/day. Treatment continued until tumour progression, unacceptable toxic effects, or withdrawal of consent. Patients were not allowed to crossover.
Concomitant medication	Not reported (NR)	NR	Nr
Primary outcomes	Overall survival (OS)	OS	OS
Other outcomes used in the economic model/specified in the scope	Progression-free survival (PFS) Response rate (RR) Disease control rate (DCR) Adverse events (AEs)	PFS Time to treatment failure Overall response rate (ORR) DCR Duration of response (DOR)	PFS Objective response DCR DOR Time to treatment failure AEs

Trial number (acronym)	**RECOURSE** (NCT01607957) Mayer 2015[16]	**TERRA** (NCT01955837) Xu 2018[17]	**Yoshino 2012[18]**
		AEs	
Pre-planned subgroups	KRAS status, time from diagnosis, geographic region, sex, age, ECOG PS, primary tumour site, prior study drug use, number of prior regimens, number of metastatic sites.	KRAS mutational status and number of prior treatment regimens	KRAS mutational status

Source: Yoshino 2012[18], Mayer 2015[16] and Xu 2018[17]
AE = adverse event; DCR = disease control rate; DOR = duration of response; ECOG PS = Eastern Cooperative Oncology Group Performance Status; KRAS = Kirsten rat sarcoma viral oncogene homologue; NR = not reported; ORR = overall response rate; OS = overall survival; PFS = progression-free survival; RCT = randomised controlled trial; RR = response rate; T/T = trifluridine/tipiracil; USA = United States of America

Table 3.34: Summary of statistical analyses for the primary analysis in the RECOURSE, TERRA and Yoshini trials

Trial	**RECOURSE** (NCT01607957) Mayer 2015[16]	**TERRA** (NCT01955837) Xu 2018[17]	**Yoshino 2012[18]**
Hypothesis objective	Not stated	Not stated	Not stated
Statistical analysis	Multivariate Cox regression analysis was performed to examine the effect of all prespecified factors (prognostic and predictive) on the overall survival effect of the study drug	The hazard ratio (HR) estimate and corresponding 95% confidence interval (CI) were determined using a Cox proportional hazards model, including treatment and stratification factors (KRAS status and country).	Cox proportional hazards model
Analysis sets	Intention-to-treat (ITT)	ITT – used for all outcomes except those below	ITT – efficacy outcomes
		Tumour response (TR) - duration of response (DOR), disease control rate (DCR), and overall response rate (ORR)	Per-protocol – adverse events (AEs)
Sample size, power calculation	The study was designed to have 90% power to detect a HR for death of 0.75 (a 25% reduction in risk) in the study drug group as compared with the	The trial was designed to detect with 90% power a HR for overall survival (OS) of 0.67 (33% risk reduction) in the trifluridine/tipiracil (T/T) arm versus the placebo arm, with a one-sided type I error of 0.025. A variable accrual	A sample size of 162 patients with a one-sided significance level of 10% was deemed necessary to verify superiority in overall survival with a power of 80%, with an expected HR of 0·67.

Trial	RECOURSE (NCT01607957) Mayer 2015[16]	TERRA (NCT01955837) Xu 2018[17]	Yoshino 2012[18]
	placebo group, with a one-sided type I error rate of 0.025. Given the treatment assignment ratio of 2:1 it was calculated that 800 patients had to be enrolled in the study, and at least 571 events (deaths) would be required for the primary analysis	period of 18 months and an approximately 10% loss to survival follow-up rate were assumed. Using a treatment allocation of 2:1 (T/T to placebo), a target of 288 events (deaths) was required for the primary analysis, and the target number of patients was set at 400.	
Data management, patient withdrawals	Not stated	Not stated	Not stated

Source: Yoshino 2012[18], Mayer 2015[16], (Xu 2018[17])
AE = adverse event; CI = confidence interval; DCR = disease control rate; HR = hazard ratio; ITT = intention-to-treat; KRAS = Kirsten rat sarcoma viral oncogene homologue; ORR = overall response rate; OS = overall survival; TR = tumour response; T/T = trifluridine/tipiracil

Table 3.35: Baseline characteristics of patients in the RECOURSE, TERRA and Yoshino trials

	CORRECT	CONCUR	RECOURSE	TERRA	Yoshino 2012
Sample size	Regorafenib (N=505) Placebo (N=255)	Regorafenib (N=136) Placebo (N=68)	T/T (N=534) Placebo (N=266)	TT (N=271) Placebo (N=135)	TT (N=112) Placebo (N=57)
Age (years, median)	Regorafenib: 61 Placebo: 61	Regorafenib: 57.5 Placebo: 55.5	T/T: 63 Placebo: 63	T/T: 58 Placebo: 56	T/T: 63 Placebo: 62
Race	Regorafenib: White 78%, Asian 15%, Black 1% Placebo: White 79%, Asian 14%, Black 3%	Regorafenib: White 0%, Asian 100%, Black 0% Placebo: White 0%, Asian 100%, Black 0%	T/T: White 57%, Asian 34%, Black <1% Placebo: White 58%, Asian 35%, Black 2%	T/T: White 0%, Asian 100%, Black 0% Placebo: White 0%, Asian 100%, Black 0%	T/T: White 0%, Asian 100%, Black 0% Placebo: White 0%, Asian 100%, Black 0%
No prior targeted biological treatment (%)	Regorafenib: 0% Placebo: 0%	Regorafenib: 41% Placebo: 38%	T/T: 0% Placebo: <1%	T/T: 55% Placebo: 49%	T/T: 22% Placebo: 18%
Sex (% female)	Regorafenib: 38% Placebo: 40%	Regorafenib: 38% Placebo: 51%	T/T: 39% Placebo: 38%	T/T: 37% Placebo: 38%	T/T: 43% Placebo: 51%
KRAS mutation (%)	Regorafenib: 54% Placebo: 62%	Regorafenib: 34% Placebo: 26%	T/T: 51% Placebo: 51%	T/T: 37% Placebo: 37%	T/T: 55% Placebo: 52%
Primary site of disease – colon (%)	Regorafenib: 64% Placebo: 68%	Regorafenib: 58% Placebo: 71%	T/T: 63% Placebo: 61%	T/T: 57% Placebo: 63%	T/T: 56% Placebo: 63%
Previous treatment lines on or after diagnosis of metastases (four or more, %)	Regorafenib: 49% Placebo: 47%	Regorafenib: 38% Placebo: 40%	T/T: 60% Placebo: 63%	T/T: 50% Placebo: 55%	T/T: NR Placebo: NR

	CORRECT	CONCUR	RECOURSE	TERRA	Yoshino 2012
Time from diagnosis of first metastases (<18 months, %)	Regorafenib: 18% Placebo: 19%	Regorafenib: 39% Placebo: 47%	T/T: 21% Placebo: 21%	T/T: 49% Placebo: 39%	T/T: NR Placebo: NR
ECOG PS 0 (%)	Regorafenib: 52% Placebo: 57%	Regorafenib: 26% Placebo: 22%	T/T: 56% Placebo: 55%	T/T: 24% Placebo: 22%	T/T: 64% Placebo: 61%

Source: Tables 9 to 17, CS appendices[7]
CS = company submission; ECOG PS = Eastern Cooperative Oncology Group Performance Status; NR = not reported; KRAS = Kirsten rat sarcoma viral oncogene homologue; T/T: trifluridine/tipiracil

3.3.1 Risk of bias

Risk of bias was assessed according to the minimum criteria for assessment proposed by NICE[31]. According to the company assessment, the three T/T trials were also regarded as methodologically strong reporting low risk of bias. Table 3.36 summarises the updated findings of the assessment provided in the response to clarification letter (reference).

Table 3.36: Quality assessment of the five RCT trials based on the NICE checklist

Study details	Randomisation appropriate?	Allocation concealment adequate?	Groups similar at the outset of the study in terms of prognostic factors?	Blinding to treatment allocation?	Unexpected imbalances in drop-outs between groups?	Authors measured more outcomes than they reported?	Did the analysis include an intention-to-treat analysis?
Trifluridine/tipiracil versus placebo							
Mayer 2015[16]	Yes	Yes	Yes	Yes	No	No	Yes
Xu 2018[17]	Yes	Yes	Yes	Yes	No	No	Yes
Yoshino 2012[18]	Yes	Yes	Yes	Yes	No	Yes	Yes
Based on Table 22, appendices CS[7] and Table A36.1 of the response to clarification letter (reference) CS = company submission; NICE = National Institute for Health and Care Excellence; RCT = randomised controlled trial							

3.3.2 Outcomes

Overall survival (OS)

The OS results from RECOURSE[16], TERRA[17] and Yoshini 2012[18], comparing T/T and placebo are provided in the appendices of the CS[7], and summarised in Table 3.37. The pooled OS result for these three trials was given as HR: 0.68 (0.62-0.76). Details of the pooling process are not provided in the CS[3].

Table 3.37: Hazard ratios (HRs) for OS reported across studies

Study Name	Treatment	N	HR (95% CI)
CORRECT	Regorafenib 160 mg	505	0.77 (0.64, 0.94)
	Placebo	255	-
CONCUR	Regorafenib 160 mg	136	0.55 (0.40, 0.77)
	Placebo	68	-
RECOURSE[16]	Trifluridine/tipiracil	534	0.68 (0.58, 0.81)
	Placebo	266	-
TERRA[17]	Trifluridine/tipiracil	271	0.79 (0.62, 0.99)
	Placebo	135	-
Yoshino 2012[18]	Trifluridine/tipiracil	135	0.56 (0.39, 0.81)
	Placebo	157	-

Study Name	Treatment	N	HR (95% CI)
Source: Table 7 of Appendix D[7] CI = confidence interval; HR = hazard ratio; OS = overall survival			

Progression-free survival (PFS)

The PFS results from RECOURSE, TERRA and Yoshini 2012 comparing T/T to placebo are provided in the appendices of the CS[7], and summarised in Table 3.38. The pooled PFS result for these three trials was given as HR: 0.45 (0.42-0.48). Details of the pooling process are not provided in the CS.[3]

Table 3.38: Hazard ratios (HRs) for PFS reported across studies

Study Name	Treatment	N	HR (95% CI)
CORRECT	Regorafenib 160 mg	505	0.49 (0.42, 0.58)
	Placebo	255	-
CONCUR	Regorafenib 160 mg	136	0.31 (0.22, 0.44)
	Placebo	68	-
RECOURSE	Trifluridine/tipiracil	534	0.48 (0.41, 0.57)
	Placebo	266	-
TERRA	Trifluridine/tipiracil	271	0.43 (0.34, 0.54)
	Placebo	135	-
Yoshino 2012	Trifluridine/tipiracil	135	0.41 (0.28, 0.59)
	Placebo	157	-
Source: Based on Table 8 of Appendix D[7] CI = confidence interval; HR = hazard ratio; PFS = progression-free survival			

Adverse events (AEs)

The CS proposes that the use of regorafenib would provide a chemotherapy-free treatment with a different AE profile from that of T/T. As such, the ERG requested[32] that the company would present an indirect treatment comparison analysis focusing on TEAEs and illustrating how the profiles of the two treatments might differ.

According to the company the incidence of Grade ≥ 3 TEAEs in CORRECT and CONCUR were similar to those reported in RESOURSE and TERRA; also noting that Grade ≥ 3 TEAEs data were not available for the study by Yoshino 2012. Regarding TRAEs, the company stated that the difference in the AEs profile lies in the different type of AEs related to each of the treatments (Table 3.39). "*Trifluridine/tipiracil is associated with higher haematological AEs such as Grade ≥ 3 neutropenia (33–50%), leukopenia (21–28%) and anaemia (17–18%), whereas regorafenib is associated with higher Grade ≥ 3 hand–foot skin reactions (16–17%) and hypertension (7–11%).*" (page 81)[6]. The rationale is that patients who have a record of not tolerating the AE profile of T/T could be treated with regorafenib instead. The company has also executed an additional NMA on TEAEs discussed in the following Section (3.4).

Table 3.39: Grade ≥ 3 TRAEs in > 2% in regorafenib and T/T studies included in the ITC.

AE Grade ≥ 3	Regorafenib		Trifluridine/tipiracil (T/T)		
	CORRECT	CONCUR	RECOURSE	TERRA	Yoshino 2012
Abdominal pain			2.4%		

AE Grade ≥ 3	Regorafenib		Trifluridine/tipiracil (T/T)		
	CORRECT	CONCUR	RECOURSE	TERRA	Yoshino 2012
Anaemia	**2.8%**		**18.0%**	**17.7%**	**16.8%**
Anorexia	3.2%				4.4%
Asthenia			3.4%		
Decreased appetite			3.6%		
Diarrhoea	7.2%		3.0%		6.2%
Fatigue	9.6%	2.9%	3.9%		6.2%
Febrile neutropenia		2.2%	3.8%		4.4%
Hand–foot skin reaction	**16.6%**	**16.2%**			
Hypertension	7.2%	11.0%			
Leukopenia		**2.2%**	**21.2%**	**20.7%**	**28.3%**
Lymphopenia				**14.4%**	**9.7%**
Mucositis	3.0%				
Nausea			1.9%		4.4%
Neutropenia		**2.2%**	**37.5%**	**33.2%**	**50.4%**
Rash	5.8%	4.4%			
Thrombocytopenia	2.8%	2.9%	5.1%	3.0%	4.4%
Vomiting			2.1%		3.5%
Hyperbilirubinemia	2.0%	6.6%	8.4%	7.0%	
Hypophosphatemia	3.8%	6.6%	7.9%		
Increase in ALT level		6.6%	1.9%	1.1%	
Increase in AST level		5.9%	4.3%	3.7%	
Increase in lipase level	3.2%	4.4%			

Based on Table A28.1 of the response to request for clarification from the ERG[6]
AE = adverse event; ALT = alanine aminotransferase; AST = aspartate aminotransferase; ERG = Evidence Review Group; ITC = indirect treatment comparison; T/T = trifluridine/tipiracil; TRAEs = treatment related adverse events

ERG comments:

- In the Yoshino 2012 study there were gender differences, with a greater proportion of males in the T/T group. However, a subgroup analysis in Yoshino 2012 shows that this difference would have favoured the placebo group, and therefore may have reduced the observed measure of effect between T/T and placebo in terms of the main outcomes. Therefore, this does not threaten the validity of the overall conclusion that T/T was better than placebo.
- The ERG executed a risk of bias assessment using the same tool. For T/T versus placebo, Xu 2018[17] and Mayer 2015[16] did not clarify the use of allocation concealment. This raises the risk of selection bias in those two trials, reducing confidence in the pooled estimates for T/T versus placebo and therefore also reducing confidence in the overall ITC estimates.
- In their response[6] the company in discussing regorafenib's safety profiles refers to both TEAEs and TRAEs. It is not clear whether a distinction is made between the two. They start by comparing the

rates of Grade ≥ 3 TEAEs in RECOURSE and TERRA to the rates of Grade ≥ 3 TEAEs in CORRECT and CONCUR, but then go on comparing the types of TRAEs that were experienced in these studies.

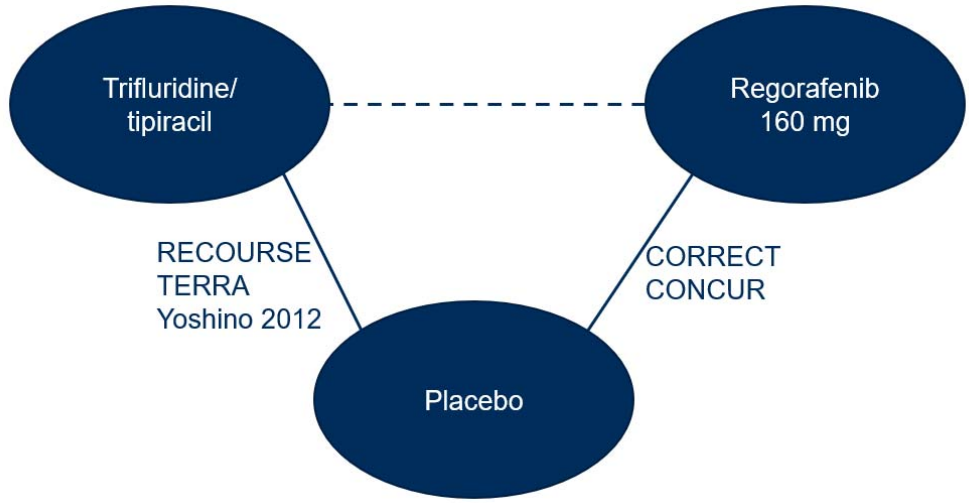
- The company maintains that regorafenib would provide a different option to T/T due to their somewhat different AE profile. To this effect the company has provided a list of Grade ≥ 3 TRAEs experienced by patients receiving both treatments. It does appear to be the case that there is a far lower risk of haematological AEs with regorafenib. In contrast, hand–foot skin reaction appears to occur with regorafenib only. These results do seem to be consistent with the comparative observational evidence presented in Section 3.2.6.8, which also shows that hypertension is also more common with regorafenib.

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

A fixed-effect NMA was carried out, in order to generate the target OS and PFS results - regorafenib versus T/T – from the available pooled OS and PFS results for regorafenib versus placebo and T/T versus placebo. Estimation of the effect of regorafenib versus T/T from the pooled effects for regorafenib versus placebo and T/T versus placebo was based on an ITC because there were no closed loops, as shown in Figure 3.12. After the request of the ERG[6] random-effect models were also fitted.

According to the company the five studies were found to fulfil *"...the basic assumptions of homogeneity, similarity, and consistency, with evidence for the existence of treatment effect modification being relatively weak*" (page 35)[7]. The identification of potential treatment effect modifiers was based on the results of subgroup analysis and further corroboration from clinical expert opinion. Although the three characteristics of prior targeted treatment, age and gender were initially identified as potential treatment effect modifiers the company's investigation concluded that the evidence was weak. The results of the investigation are reported in Section B.3.1.8.1 of Appendix D[7].

Figure 3.12: Network diagram for OS.



Taken from Figure 11 in CS[3]

CS = company submission; OS = overall survival

3.4.1 Efficacy NMA results

The results of the fixed-effects NMA for OS are presented in Table 3.40. The point estimates for the HR of regorafenib versus placebo and T/T versus placebo were exactly the same (0.68) with a small difference in the 95% credible intervals (CrIs) (0.59, 0.78 versus 0.62, 0.76). The indirect comparison

of the two treatments showed a very similar OS effectiveness of both treatments OS HR: 0.99 (95% CrI; 0.84, 1.17). Regarding PFS the fixed-effects NMA results illustrated a slightly lower HR for regorafenib versus placebo HR: 0.42 (95% CrI; 0.39, 0.45) than T/T versus placebo HR: 0.45 (95% CrI; 0.42, 0.48), as presented in Table 3.41. The indirect comparison of the two treatments suggested similar PFS effectiveness as shown by the HR and the 95% CrI that includes one, PFS HR: 0.93 (95% CrI; 0.85, 1.03).

Table 3.40: Results of the fixed effects NMA of OS

Comparison	HR (95% CrI)
Regorafenib versus placebo (direct pooled evidence)	0.68 (0.59, 0.78)
Trifluridine/tipiracil (T/T) versus placebo (direct pooled evidence)	0.68 (0.62, 0.76)
Regorafenib versus T/T (indirect estimate)	0.99 (0.84, 1.17)

Based on Table 14, CS[3]
CrI = credible interval; CS = company submission; HR = hazard ratio; NMA = network meta-analysis; OS = overall survival; T/T = trifluridine/tipiracil

Table 3.41: Results of the fixed effects NMA of PFS

Comparison	HR (95% CrI)
Regorafenib versus placebo (direct pooled evidence)	0.42 (0.39, 0.45)
Trifluridine/tipiracil (T/T) versus placebo (direct pooled evidence)	0.45 (0.42, 0.48)
Regorafenib versus T/T (indirect estimate)	0.93 (0.85, 1.03)

Based on Table 15, CS[7]
CrI = credible interval; CS = company submission; HR = hazard ratio; NMA = network meta-analysis; PFS = progression-free survival; T/T = trifluridine/tipiracil

3.4.2 Sensitivity analyses

Two different approaches were implemented by the company to examine the fitness of the fixed-effects OS and PFS NMAs. One approach used the anchored matching-adjusted indirect comparison (MAIC) method for weighting specific baseline characteristics that were investigated as possible effect modifiers. In the other approach, some of the five studies of the NMA were excluded, at a time, and the NMA was run again to test for differences between studies as identified in the assessment of study heterogeneity.

Although the company found that all evidence for effect modification were weak, they fitted sensitivity analyses using the method of anchored MAIC (as described in Appendix D, Section B.3.1.8.2 of the CS[7]). Within the anchored MAIC process, the pooled population of CORRECT and CONCUR was matched to the population of pooled RECOURSE, TERRA and Yoshino 2012 based on the proportion of:

- male patients,
- patients aged < 65 years and,
- patients without any prior biological treatment.

The fitness of the MAIC process was judged based on the approximation of the effective sample size (ESS), meaning the comparison of the number of matched patients to the original number of patients. The fitness of the NMA was assessed by comparing the OS and PFS results of the NMA to the anchored MAIC results. The matching results for the patients' characteristics used in the MAIC are presented in Table 3.42 and the derived weight in Figure 3.13. The ESS after matching ([redacted]) was similar

to the original ITT population in CORRECT + CONCUR (n=964). As shown in the histogram most re-scaled weights were close to 1. The results of the anchored MAIC process as well as a comparison to NMA results are presented in Table 3.43 for OS and Table 3.44 for PFS. The HRs and accompanying 95% CIs produced by the anchored MAIC are similar to the NMA results and the differences remain not statistically significant (CI include 1), for both OS and PFS outcomes. The company supports that the NMA results are the most robust analysis because the treatment effect modification evidence was weak which was supported by the similarity of the results after adjusting for them.

Table 3.42: Matching CORRECT + CONCUR to RECOURSE + TERRA + Yoshino 2012

	Prior to matching		After matching
	CORRECT + CONCUR (ITT)	**RECOURSE + TERRA + Yoshino 2012 (ITT)**	**CORRECT + CONCUR (ITT)**
N/ESS	N=964	N=1,375	ESS=[redacted]
% male patients	60.37%	61.90%	61.90%
% patients < 65	65.15%	38.11%	38.11%
% no prior targeted biologic treatment	8.51%	18.11%	18.11%

Based on Table 19 of Appendix D of the CS[7]
CS = company submission; ESS = effective sample size; N= number of patients; ITT = intention-to-treat

Figure 3.13: Histogram of rescaled weights (pooled CORRECT and CONCUR versus pooled RECOURSE, TERRA and Yoshino)

[redacted]

Regorafenib versus placebo	[redacted]	CORRECT + CONCUR
Trifluridine/tipiracil (T/T) versus placebo	[redacted]	RECOURSE + TERRA + Yoshino 2012
Regorafenib versus T/T	[redacted]	Fixed-effects NMA
Regorafenib versus placebo (adjusted)	[redacted]	CORRECT + CONCUR (weighted)
Regorafenib versus T/T (adjusted)	[redacted]	Anchored MAIC

Based on Table 21 of Appendix D of the CS[7]

CI = confidence interval; CS = company submission; HR = hazard ratio; MAIC = matching-adjusted indirect comparison; NMA = network meta-analysis; PFS = progression-free survival; T/T = trifluridine/tipiracil

The ERG suggested that more characteristics should be included in the anchored MAIC process, highlighting the potential presence of more treatment modifiers. The proposed characteristics to be weighted were ECOG PS, previous treatment lines, KRAS status and time from diagnosis. The company executed additional weighting adjusting for the proportion of patients:

- with ECOG PS 0
- with KRAS mutation
- whose time from diagnosis of first metastases >= 18 months and
- whose previous treatment line <4

The Yoshiro 2012 study was not included in this additional analysis because data were not available regarding time from diagnosis of metastases and previous treatment lines. The ESS remained high (n=851), but the derived weights were not reported. Nevertheless, the results of the additional anchored MAIC showed similar results to the fixed-effects NMA; regorafenib versus T/T OS HR: 0.95 (95% CI; 0.77, 1.18) and PFS HR: 1.00 (95% CI; 0.81, 1.22).

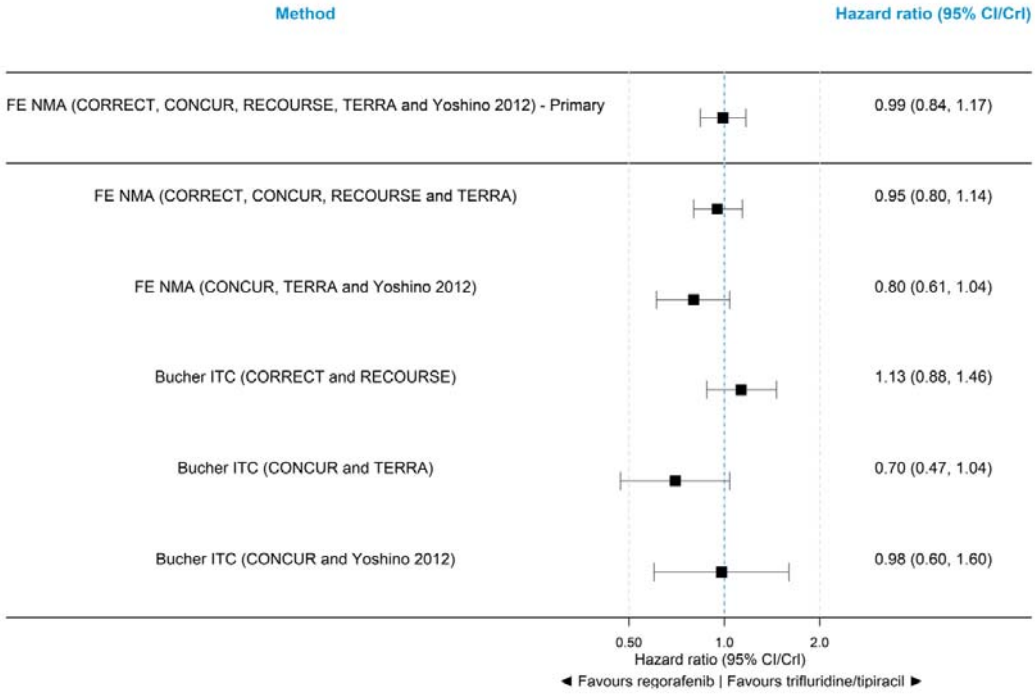
A second approach to sensitivity analysis (SA) was also executed via including some of the studies at a time as presented in Table 3.45. The results of the SA are presented in Figure 3.14 for OS and Figure 3.15 for PFS. Overall, the SA results are consistent to those of the base-case fixed-effects NMA. The only difference in the results was observed in SA three, including only CONCUR, TERRA and Yoshino 2012, where the PFS HR was significant 0.73 (95% CI: 0.62, 0.86). These studies only included Asian patients with some of them previously treated with anti-VEGF treatments, thus making them "…*more representative of treatment in the UK setting*" (page 70)[3].

Table 3.45: ITC sensitivity analyses (SAs)

	Regorafenib		Trifluridine/tipiracil (T/T)			
	CORRECT	CONCUR	RECOURSE	Yoshino 2012	TERRA	Brief Explanation
Base-case	x	x	x	x	x	
SA1	x	x	x		x	Removal of Yoshino 2012 as the only phase II trial
SA2	x		x			All patients had been received prior anti-VEGF in these two trials
SA3		x		x	x	These three studies included Asian only patients. Treatment most closely aligned with UK clinical practice
SA4		x			x	Phase III trials most closely aligned to UK clinical practice
SA5		x		x		Mainly for completeness i.e., Asian-only study and complements SA4

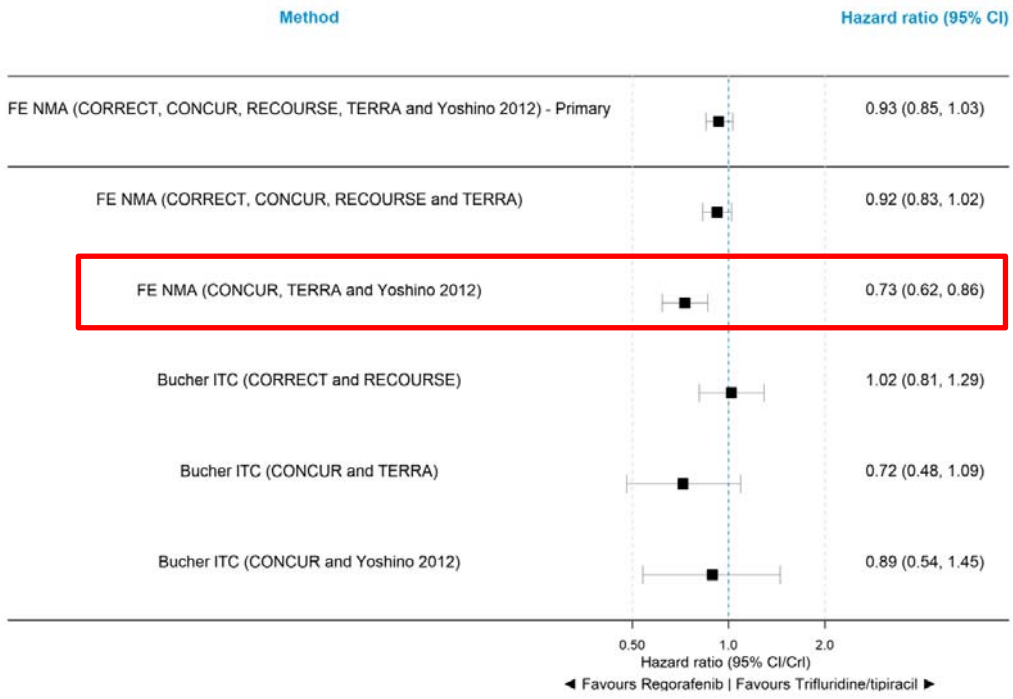
Based on Table 16 of Document B of the CS[3]
CS = company submission; ITC = indirect treatment comparison; SA = sensitivity analysis; T/T = trifluridine/tipiracil; UK = United Kingdom; VEGF = Vascular endothelial growth factor

Figure 3.14: Overview of sensitivity analysis results for OS



CI = confidence interval; CrI = credible interval; FE = fixed effects; ITC = indirect treatment comparison; NMA = network meta-analysis; OS = overall survival

Figure 3.15: Overview of sensitivity analysis results for PFS



CI = confidence interval; CrI = credible interval; FE = fixed effects; ITC = indirect treatment comparison; NMA = network meta-analysis; PFS = progression-free survival

3.4.3 Adverse events (AEs) NMA results

After the suggestion of the ERG, the company included an additional NMA on AEs. The focus was on TEAEs, as requested by the ERG. Four studies were included in this NMA, as the data for Yoshino 2012 were not available (Table 3.46). The results of the fixed-effects NMA for Grade 3 and 4 TEAEs are presented in Table 3.47, showing similar odds of experiencing Grade 3 or 4 TEAEs for the two treatments, OR: 0.90 (95% CrI; 0.55, 1.47).

Table 3.46: Data included in fixed effects NMA in treatment emergent Grade 3 or 4 AEs

Study Name	Treatment	Safety N	Number of treatment related Grade 3 or 4 AEs: n (%)
CORRECT	Regorafenib	500	270 (54.0)
	Placebo	253	35 (13.8)
CONCUR	Regorafenib	136	74 (54.4)
	Placebo	68	10 (14.7)
RECOURSE	Trifluridine/tipiracil	533	261 (49.0)
	Placebo	265	27 (10.2)
TERRA	Trifluridine/tipiracil	271	124 (45.8)
	Placebo	135	14 (10.4)
Based on Table A28.2 of the company's response to request for clarification from the ERG[6] AEs = adverse events; ERG = Evidence Review Group; N = total number of patients; n = number of patients with an event; NMA = network meta-analysis			

Table 3.47: Results of the fixed effect NMA in treatment emergent Grade 3 or 4 AEs

Comparison	OR (95% CrI) - NMA
Regorafenib versus placebo	7.32 (5.19, 10.44)
Trifluridine/tipiracil (T/T) versus placebo	8.11 (5.74, 11.65)
Regorafenib versus T/T	0.90 (0.55, 1.47)
Based on Table A28.3 of the company's response to request for clarification from the ERG[6] AEs = adverse events; CrI = credible interval; ERG = Evidence Review Group; NMA = network meta-analysis; OR = odds ratio; T/T = trifluridine/tipiracil	

Two additional AE NMAs were executed, one including all TEAEs and one focusing on the discontinuation of treatments due to AEs. The data included in the analysis are presented in Tables 3.48 and 3.50 and the results in Tables 3.49 and 3.51. The results suggest there are higher odds of patients treated with regorafenib compared to T/T experiencing TEAEs (OR: 1.94 (95% CrI; 1.20, 3.17)), while discontinuation of treatment due to AEs was similar (OR: 1.10 (95% CrI; 0.53, 2.24)). It should be noted that CrIs include one in both results. The company noted that the TEAEs analysis included Grade 1 and 2 AEs which as mild and moderate AEs would not be expected to impact on QoL or costs.

Table 3.48: Data included in fixed effects NMA in all TEAEs

Study Name	Treatment	Safety N	Number of all treatment emergent AEs: n (%)
CORRECT	Regorafenib 160 mg	500	465 (93.0)
	Placebo	253	154 (60.9)
CONCUR	Regorafenib 160 mg	136	132 (97.1)
	Placebo	68	31 (45.6)
RECOURSE	Trifluridine/tipiracil	533	458 (85.9)
	Placebo	265	146 (55.1)
TERRA	Trifluridine/tipiracil	271	244 (90.0)
	Placebo	135	70 (51.9)
Based on Table A28.4 of the company's response to request for clarification from the ERG[6] AEs = adverse events; ERG = Evidence Review Group; mg = milligrams; N = total number of patients; n = number of patients with an event; NMA = network meta-analysis; TEAEs = treatment emergent adverse events			

Table 3.49: Results of the fixed effect NMA in all TEAEs

Comparison	OR (95% CrI) - NMA
Regorafenib versus placebo	11.42 (7.78 to 17.10)
Trifluridine/tipiracil (T/T) versus placebo	5.90 (4.43 to 7.89)
Regorafenib versus T/T	1.94 (1.20 to 3.17)
Based on Table A28.5 of the company's response to request for clarification from the ERG[6] CrI = credible interval; ERG = Evidence Review Group; NMA = network meta-analysis; OR = odds ratio; TEAEs = treatment emergent adverse events	

Table 3.50: Data included in fixed effects NMA of discontinuations due to AEs

Study Name	Treatment	Safety N	Discontinuation due to AEs: n (%)
CORRECT	Regorafenib 160 mg	500	85 (17.0)
	Placebo	253	30 (11.9)
CONCUR	Regorafenib 160 mg	136	19 (14.0)
	Placebo	68	4 (5.9)
RECOURSE	Trifluridine/tipiracil (T/T)	533	19 (3.6)
	Placebo	265	4 (1.5)
TERRA	T/T	271	24 (8.9)
	Placebo	135	11 (8.1)
Yoshino 2012	T/T	113	1 (0.9)
	Placebo	57	4 (7.0)

Based on Table A28.6 of the company's response to request for clarification from the ERG[6]

AEs = adverse events; ERG = Evidence Review Group; mg = milligrams; N = total number of patients; n = number of patients with an event; NMA = network meta-analysis; T/T = trifluridine/tipiracil

Table 3.51: Results of the fixed effect NMA in discontinuation due to AEs

Comparison	OR (95% CrI) - NMA
Regorafenib versus placebo	1.66 (1.11 to 2.56)
Trifluridine/tipiracil (T/T) versus placebo	1.51 (0.86 to 2.78)
Regorafenib versus T/T	1.10 (0.53 to 2.24)

Based on Table A28.7 of the company's response to request for clarification from the ERG[6]

AEs = adverse events; CrI = credible interval; ERG = Evidence Review Group; NMA = network meta-analysis; OR = odds ratio; T/T = trifluridine/tipiracil

ERG comment:

As with the efficacy NMAs, there seems to be little difference between regorafenib and T/T in terms of Grade 3 or 4 AEs and discontinuation due to AEs. This is notwithstanding the clear differences in specific AEs, such as hand-foot skin reaction, hypertension, and various haematological events, as described in Sections 3.2.6.8 and 3.3.2. There does seem to be a higher rate of any AEs with regorafenib, although the clinical significance of this is unclear.

3.4.4 Uncertainties in the ITCs

There are certain limitations in the ITC analyses that the company acknowledges. The 95% CrIs produced by the fixed-effects efficacy NMAs were narrower than the CIs in the clinical trials. The company attributes this difference to the type of models used i.e., fixed versus random-effects. Indeed, when the additional NMAs were executed using random-effects models the 95% CrIs were wider.

Regarding the important issue of study heterogeneity, differences were identified in the phase of the study (phases II, III), the race of patients (Asian populations versus non-Asian) and the use of prior targeted biologic treatments i.e., anti-VEGF treatments (bevacizumab). The latter was identified as a potential treatment effect modifier thus possibly introducing bias in the analysis. This characteristic was investigated in series on SAs (see Section 3.4.2), which found that this difference did not affect the robustness of the primary analysis results.

Further limitations of the NMA analysis regarding OS and PFS outcomes as well as safety outcomes have been identified by the ERG and are described in the following comments section.

ERG comments:

The ERG expressed concerns[32] on the comparability of the populations in the studies included in the ITC OS and PFS analyses regarding race, region, the median age of patients, previous targeted biological treatments, ECOG status and number of prior treatment lines. The company acknowledged the highlighted differences but maintains that these characteristics were not identified as potential treatment modifiers and as such, they would have very little impact in the results.

- The ERG noted in the clarification letter[32] that the only substantial difference in the SA results was for SA three which included only CONCUR, TERRA and Yoshino 2012, where the PFS HR was significant 0.73 (95% CI; 0.62, 0.86). These studies only included Asian patients with some of them instead of 100% previously treated with anti-VEGF treatments. The company maintained that "*Although we believe the prior treatments received in these three trials make them more representative of clinical practice in England (specifically a significant proportion of patients had not received anti-VEGF), we consider the benefit favouring regorafenib in the sensitivity analysis referred to in the question to be a chance-effect*" (page 98)[6]. The ERG considers that it makes some sense to group these trials together given the relative homogeneity of race and region and prior anti-VEGF treatment, although the nature of any treatment effect modification of these characteristics is not clear (see Section 3.2.1).
- The company was also asked to include the further characteristics of race and region to be included in the anchored MAIC analysis, to which they responded that "*Race and region were not identified as potential treatment effect modifiers and therefore the inclusion of these variables may lead to over matching*." (page 87)[6]. In their comparison of baseline characteristics, they acknowledge that both CORRECT and RECOURSE recruited patients from across the world while CONCUR, TERRA and Yoshino 2012 only included Asian patients, but in their 'Summary of treatment effect modifier investigations' in Appendix D of the CS[7], they make no note on race and region. As discussed in Section 3.2.3, it is unclear what the effect of race or region might be on outcome.
- Recognising the potential significance of previous anti-VEGF treatments, the ERG requested that a subgroup ITC analysis should be executed including only patients with no such history, as this would be more closely aligned to UK practise. The company responded that such an analysis could not be performed due to studies not reporting the necessary data and limited sample sizes when they did. It is also important to note that the relationship between treatment effect and prior targeted treatment might not be straightforward: it might vary by outcome with PFS showing that no previous anti-VEGF treatment counterintuitively reduces the treatment effect in CONCUR (see Section 3.2.3). Also, it might vary by intervention with the T/T trial with 100% prior targeted treatment, RECOURSE, having a higher treatment effect (lower HR) for OS than one of the trials with a mixture of prior treatment experience, TERRA.
- To address and explore the studies' differences in terms of race and previous treatment with an anti-VEGF, the ERG suggested executing additional SA including only CORRECT or CONCUR in the NMA. The company responded that such comparisons had already been done in the existing ITC SA. In addition, they stated that CORRECT could not be compared to TERRA and Yoshino 2012, since they included only Asian patients, while 50% and 20% of patients had received prior anti-VEGF, respectively.
- The company did provide additional anchored MAIC analyses for CORRECT versus RECOURSE, adjusting for the percentage of male patients, percentage of patients <65, percentage of White

patients, percentage area not Asia; CONCUR versus TERRA and CONCUR versus Yoshino 2012, adjusting for percentage of male patients, percentage of patients <65 and percentage with no prior targeted biological treatment. Overall, the above MAIC analysis showed minimal impact in the results of OS and PFS of regorafenib versus T/T and most of the results remained not significant (95% CIs including 1).

- The ERG suggested that a random-effects model might be more appropriate, and the company indeed provided this additional analysis. The point estimates were similar but the 95% CrIs were wider. Unfortunately, measures of heterogeneity were not made available (I^2, Q) so the fitness of the models cannot be statistically appraised.
- The NMA was conducted in R and the ERG requested that the company would report the data frames, the R packages and the code that was used, in order to replicate the models and appraise the decision of the model parameters. The company did provide the data frames and the packages used, but not the code without offering a justification. They provided a brief description of the parameters of the model, but without the code it is clear the analysis cannot be replicated. In addition, they provided a JAGS model script[33] that was used to perform the continuous endpoint (PFS and OS) meta-analyses, but did not report why it was used and how it was connected to the NMA models.

In conclusion, the ERG notes that the five main SAs did not change the fixed effect NMA estimates significantly. Neither did either of the MAICs. Although it is unclear if these SAs accounted for all sources of inconsistency or heterogeneity in the network, the ERG considers that it is unlikely that any further SAs varying which RCTs are included would be informative, notwithstanding the potential inclusion of comparative observational studies. Four retrospective observational studies directly comparing regorafenib to T/T were identified: Nakashima 2020,[1] Tanaka 2018,[23] Huemer 2020[26] and Sueda 2016[25] (see Sections 3.2.1.2, and 3.2.5.5). These studies provide further evidence on efficacy, as well as safety and could potentially be used in a further NMA combining evidence from RCTs and observational studies, notwithstanding the issues identified regarding selection bias and the effect of decisions regarding crossover.[2] It should be noted that the study by Nakashima 2020[1], which has a large sample size of N=2,529 (for OS, PFS and RR outcomes), offers head-to-head efficacy evidence. Given the difference in treatment effect on OS between the NMA and the Nakashima 2020 study, as well as the questionable comparability between the RCTs, there is substantial uncertainty in the effectiveness of regorafenib versus. T/T, which the ERG therefore identifies as a key issue.

3.5 Conclusions of the clinical effectiveness section

The literature searches overall were adequate to identify most of the relevant clinical information from evidence related to regorafenib or T/T for the treatment of mCRC in the third- or later-line setting. However, the ERG was concerned that the structure of the MEDLINE/Embase search, particularly the inclusion of three drugs outside the remit of this review and the additional of a facet for non-response was overly restrictive and may have led to the exclusion of some relevant papers. The company reran the searches at clarification finding an additional 21 unique relevant papers after screening: whilst the company stated that no new trials were identified, the ERG considered that the comparative observational study by Nakashima 2020 should be included.[1]

The main clinical effectiveness evidence of the effectiveness of regorafenib consisted of two RCTs of regorafenib versus placebo, CORRECT and CONCUR, which provided estimates of outcomes including OS, PFS, response rates, HRQoL and AEs for regorafenib plus BSC versus BSC only. Both RCTs showed that regorafenib was more effective than BSC only in terms of HRs for OS and PFS that were less than 1 including the 95% CI. For response in terms ORR and DCR, regorafenib was also

superior with a P value less than 0.05 for both in CONCUR and for DCR only in CORRECT. In general, regorafenib and placebo did not differ in their effects on QoL measured by EORTC QLQ-C30 or EQ-5D.

Although the ERG had little concern regarding internal validity i.e., selection bias between arms within each trial, the two trials did produce quite different results. with better results in favour of regorafenib for the CONCUR trial in both OS and PFS. There are three main differences between cohorts that could explain this: prior anti-VEGF treatment in the form of bevacizumab, race, and number of prior treatments. One of the main limitations of the RCTs was that there were no UK patients: indeed, although CORRECT included both Asian and non-Asian patients, CONCUR only included Asian patients. A further and issue on the generalisability to the UK population, with possibly inverse effects, was prior treatment with an anti-VEGF treatment (bevacizumab) which is not licenced in the UK practice. The entire population of CORRECT and 39.7% of the population in CONCUR were previously treated with this agent. Patients in CONCUR generally receive fewer lines of therapy from diagnosis of metastatic disease: 38% to 40% versus 47% to 49% received at least four (>3) prior lines in CONCUR versus CORRECT, depending on arm. Subgroup analyses were conducted to investigate the effect of each of these in each trial where feasible: only CORRECT could be used for race and only CONCUR for prior anti-VEGF treatment because all patients in CORRECT received it. Generally, the ERG holds the view that low powering of subgroup analyses does not remove the need to look for possible subgroup differences. Because of the lack of statistical power there is a need to interpret point estimates more broadly, especially where the magnitude of between-subgroup difference is large, with the onus on being vigilant for possible type II errors. However, although in CORRECT, Asian race seemed to improve PFS a little, the direction of effect was small and uncertain, and the direction of effect was opposite for OS. Prior anti-VEGF treatment in CONCUR did lead to a substantial increase in the point estimate for the HR for OS, but size of the increase seemed to depend on the effect of previous anti-EGFR treatment, which was also true for PFS. For number of lines of previous treatment, the subgroup analysis results were counter to what would be expected if this were an explanation for the better results for regorafenib in CONCUR i.e., CONCUR had fewer patients with >3 previous lines, but the HR point estimate was lower for >3 lines for OS and PFS in both trials. Because it is uncertain as to the size and direction of subgroup differences and the disparity between treatment experience and race in the trials and NHS clinical practice, this is a key issue. This is also the justification for the sensitivity analyses in the NMA used to compare regorafenib with T/T, with various combinations of both regorafenib and T/T trials.

To compare the efficacy of regorafenib with T/T the company conducted a fixed effects NMA, which, given the lack of inclusion of head-head trials, was in the form of an ITC. This efficacy analysis included CORRECT and CONCUR, along with three the RCTs, RECOURSE[16], TERRA[17] and Yoshini 2012,[18] on T/T versus placebo. The efficacy outcomes explored in the ITC/NMA were OS and PFS. No difference in clinical effectiveness was observed between regorafenib and T/T for the outcome of OS or PFS, with HRs below 1, but 95% CIs that overlapped 1. Also, even if not due to sampling error, the benefit, being small (HR for PFS 0.93 (0.85, 1.03) and OS 0.99 (0.84, 1.17)) may not be of clinical importance. However, a set of limitations of the NMA was identified relating to the heterogeneity of the included studies in characteristics that are potential treatment effect modifiers, as highlighted by the considerable differences in the populations' baseline characteristics and, between RCTs of the same comparison, differences in treatment effect. Because of this heterogeneity five sensitivity analyses were carried out with various combinations of trials based on homogeneity of design (removing a phase II study), prior anti-VEGF treatment and race. A MAIC was also conducted to adjust for potential treatment effect modifiers, which included prior targeted treatment, but not race or number of prior lines

of therapy. The only substantial difference in the SA results was for the one which included only CONCUR, TERRA and Yoshino 2012, which were all Asian only and had a mix of prior targeted treatment (as opposed to 100% prior treatment), and only for PFS, where the HR was significant 0.73 (95% CI; 0.62, 0.86). This combination does make some sense in the relative homogeneity, although the role of these characteristics as treatment effect modifiers is not clear and probably not straightforward, as discovered by examining the regorafenib RCTs, as well as in comparisons between the T/T RCTs. The ERG considers that, although there is a potential lack of comparability between the RCTs in the NMA and potential lack of applicability to NHS clinical practice, there is probably little to be gained by further SAs with the RCTs.

Given that the NMA contained no head-to-head comparison RCTs of regorafenib versus T/T, the ERG has also chosen to include the four comparative observational studies that directly compared regorafenib and T/T. Three of the comparative observational studies[23, 25, 26] produced results that were similar to the NMA. Very importantly, for the largest one, Nakashima 2020[1], there was a strong effect favouring T/T for OS (HR: 0.66, P<0.001). There is likely to be selection bias in all of these observational studies, although there did seem to be better balance in the baseline characteristics for Nakashima 2020. They were all also conducted in a context where some patients, apparently those with lower ECOG PS, could crossover to receive subsequent treatment with either regorafenib or T/T. All but Tanaka 2018 presented results separately for those who did not crossover, but it is unclear what the treatment effect would be on patients who crossed over, but who would receive BSC in NHS clinical practice. Given the difference in treatment effect on OS between the NMA and the Nakashima 2020 study, as well as the questionable comparability between the RCTs, there is substantial uncertainty in the effectiveness of regorafenib versus T/T, which the ERG therefore identifies as a key issue. The ERG therefore suggests that a further NMA combining evidence from RCTs, and observational studies could be conducted, notwithstanding the risk of selection bias and the effect of excluding patients who did not crossover.

Regarding the safety of regorafenib, as reported in the RCTs, there was a slightly greater rate of permanent treatment discontinuation in the regorafenib arms compared to placebo in both CORRECT (17.6% versus 12.6%) and CONCUR (14.0% versus 5.9%). However, regorafenib did not lead to a greater number of deaths than placebo, and SAEs did not differ greatly between regorafenib and placebo arms. The alternative safety profile of regorafenib to T/T was one of the chief justifications of the submission, but no comparative AE evidence was presented. The ERG considered the comparative observational studies and found evidence clearly suggests a greater adverse effects burden from regorafenib than T/T, as reported from the largest identified observational study[1]. There was also a clear distinction by type of AE with hand-foot skin reactions and hypertension occurring almost solely with regorafenib as opposed to haematological events such as leukopenia, anaemia and neutropenia occurring almost solely with T/T. Following request by the ERG, a comparison was made with T/T in the RCTs, which largely confirmed the findings in the observational studies. The RCTs included in the efficacy NMAs (apart from Yoshino 2012), were also used in three AEs NMAs related to AEs, one focusing on Grade 3 or 4 TRAEs, one on the discontinuation of treatments due to AEs and one on TEAEs of all Grades. The first two analyses showed similar outcomes for both treatments while the third one favoured T/T.

4. COST EFFECTIVENESS

4.1 ERG comment on company's review of cost effectiveness evidence

This Section pertains mainly to the review of cost effectiveness analysis (CEA) studies. However, the search Section (4.1.1) also contains summaries and critiques of other searches related to cost effectiveness (CE) presented in the CS. Therefore, the following Section includes searches for the CEA review, measurement, and evaluation of health effects as well as for cost and healthcare resource identification, measurement and valuation.

4.1.1 Searches performed for cost effectiveness section

The following paragraphs contain summaries and critiques of all searches related to CE presented in the CS.

Table 4.1: Data sources for the cost effectiveness (CE) systematic review (as reported in CS)

Resource	Host/Source	Date Ranges (to date of last search)	Dates searched
Electronic databases			
MEDLINE	Embase.com	2010-2022/02/22	5.3.21 Updated 22.2.22
Embase	Embase.com	2010-2022/02/22	5.3.21 Updated 22.2.22
MEDLINE In-Process	Pubmed.com	2010-2022/02/22	19.3.21 Updated 22.2.22
EconLit	Ebsco.com	2010-2022/02/22	19.3.21 Updated 22.2.22
HTAD	CRD	2010-2018/03/31	7.4.21
NHS EED	CRD	2010-2015/03/31	7.4.21
Conferences			
ASCO		2019-2022	2022/02
ESMO		2019-2022	2022/02
DDW		2019-2022	2022/02
ISPOR (Annual and European)		2019-2022	2022/02
Additional searches			
NICE		All years	
SMC		All years	
AWMSG		All years	
Handsearching	Bibliographies of key systematic review and meta-analysis articles were		

Resource	Host/Source	Date Ranges (to date of last search)	Dates searched
	screened to fully evaluate the relevant economic studies		

ASCO = American Society of Clinical Oncology; AWMSG = All Wales Medicines Strategy Group; CE = cost effectiveness; CRD = Centre for Reviews and Dissemination; CS = company submission; ESMO = European Society for Medical Oncology; DDW = Digestive Disease Week; HTAD = Health Technology Assessment Database; ISPOR = International Society for Pharmacoeconomics and Outcomes Research; NHS EED = National Health Service Economic Evaluation Database; NICE = National Institute for Health and Care Excellence; SMC = Scottish Medicines Consortium

ERG comment:

- The CS and response to clarification provided sufficient details for the ERG to appraise the literature searches.
- Searches covered a good range of resources, including databases, Health Technology Assessment (HTA) agency websites and conference proceedings. Strategies contained a good mix of free text and subject headings.
- The MEDLINE search was conducted using Embase.com as described in the clinical effectiveness Section, therefore the same limitations will apply (see Section 3.1.1).
- The MEDLINE In-Process search via PubMed contained the same search limit error as previously reported in Section 3.1.1. However, this omission appears to have been corrected in the update searches.
- The CE Embase/MEDLINE search followed the same structure as the clinical searches, including the three drugs not relevant to this review and the additional facet relating to non-response etc. Therefore, the same concerns regarding the restrictive nature of this search exist as in Section 3.1.1. Given the more pragmatic nature of costs searches and the additional sources searches this is less likely to have impacted on the overall recall of results and the ERG did not request that these searches be rerun.
- The EconLit search appeared to contain an error in the use of Boolean logic:
  ("colon cancer" OR "colon carcinoma" OR "colorectal cancer" OR "rectum cancer" OR "rectum adenoma") OR ((cancer OR carcinoma OR adenoma OR adenocarcinoma OR tumor OR tumour OR neoplasm OR malignant OR malignancy) AND (colorectal OR "colo-rectal" OR "colonrectal" OR "colon rectal" OR" colon-rectal" OR colon OR rectum OR rectal OR pararectal OR bowel OR sigmoid)) ***AND*** *(crc OR mcrc OR "m-crc")*[7]

  The AND appears to have been used in error, instead of OR as in previous searches, but given the other resources searched this is unlikely to have affected the overall recall of results.

The CS reported that the SLRs looking into HRQoL studies (see Appendix H), and cost and healthcare resource identification (see Appendix I), were performed using the same methods, and that overall methods were only reported in Appendix G. Please see Appendix 1 for those searches unique to HRQoL and resource use, for additional searches of HTA organisations and conference proceedings see Table 4.1 above.

4.1.2 Inclusion/exclusion criteria

In- and exclusion criteria for the review on CE studies, utilities and costs and resource use are presented in Table 4.2.

Table 4.2: Eligibility criteria for the SLRs

	Inclusion criteria	Exclusion criteria
Patient population	Patients with relapsed/refractory metastatic colorectal cancer (mCRC) previously treated with standard therapies, e.g., fluorouracil, capecitabine, oxaliplatin, irinotecan or cetuximab monotherapy or combination therapy	Healthy volunteers Paediatric population Treatment-naïve patients with mCRC Early-stage mCRC Disease other than mCRC
Intervention	Cost-effectiveness (CE) studies: regorafenib nivolumab/ipilimumab encorafenib trifluridine/tipiracil (T/T) Utility and cost/resource use studies: no restrictions	CE studies: interventions not available in the list. Utility and cost/resource use studies: no restrictions
Comparator	No restrictions	No restrictions
Outcomes(s) 1 (Published economic evaluations)	Not specified	Not specified
Outcomes(s) 2 (Utility studies)	Not specified	Not specified
Outcomes(s) 3 (Cost/resource use studies)	Not specified	Not specified
Study design 1 (Cost effectiveness analysis studies)	Full economic evaluations Cost–consequence Cost-minimisation CE Cost–utility Cost–benefit Budget impact	In vitro studies Preclinical studies Reviews, comments, letters and editorials Case reports, case series Clinical studies reporting only efficacy and safety data
Study design 2 (Utility studies)	Studies reporting utility values (regardless of treatment) Utility data such as EQ-5D®, SF-6D® Disutilities	In vitro studies Preclinical studies Reviews, comments, letters or editorials Case reports, case series Clinical studies reporting only efficacy and safety data
Study design 3	Cost and resource use studies: Cost studies	In vitro studies Preclinical studies

	Inclusion criteria	Exclusion criteria
(Cost/resource use studies)	Resource use studies Economic evaluations reporting costs or resource use	Reviews, comments, letters and editorials Case reports, case series Clinical studies reporting only efficacy and safety data
Source: CS Appendix G, Table 53, CS Appendix H, Table 61 and CS Appendix I, Table 70. CS = company submission; CE = cost effectiveness; EQ-5D = European Quality of Life-5 Dimensions; mCRC = metastatic colorectal cancer; SLRs = systematic literature reviews; T/T = trifluridine/tipiracil		

ERG comment:

The ERG agrees that the eligibility criteria are suitable to fulfil the company's objective to identify CE studies. The rationales for excluding CE studies after full paper reviewing are considered appropriate given the defined in- and exclusion criteria.

4.1.3 Conclusions of the CE review

The CS provides an overview of the included CE, utility and resource use and costs studies, but no specific conclusion was formulated.

ERG comment:

The review was performed adequately.

4.1.4 Review of HRQoL and costs and healthcare resources

The critique of the searches is reported in an Appendix in Section 8.

4.2 Summary and critique of company's submitted economic evaluation by the ERG

4.2.1 NICE reference case checklist

Table 4.3: NICE reference case checklist

Element of HTA	Reference case	ERG comment on CS
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	In line with reference case
Perspective on costs	NHS and PSS	In line with reference case
Type of economic evaluation	Cost utility analysis with fully incremental analysis	In line with reference case
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	In line with reference case
Synthesis of evidence on health effects	Based on systematic review	In line with reference case
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of HRQoL in adults.	In line with reference case

Element of HTA	Reference case	ERG comment on CS
Source of data for measurement of HRQoL	Reported directly by patients and/or carers	In line with reference case
Source of preference data for valuation of changes in HRQoL	Representative sample of the UK population	In line with reference case
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	In line with reference case
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	In line with reference case
Discounting	The same annual rate for both costs and health effects (currently 3.5%)	In line with reference case

CS = company submission; ERG = Evidence Review Group; EQ-5D = European Quality of Life-5 Dimensions; HRQoL = health-related quality of life; HTA = Health Technology Assessment; NHS = National Health Service; NICE = National Institute for Health and Care Excellence; PSS = Personal Social Services; QALY = quality-adjusted life year; UK = United Kingdom

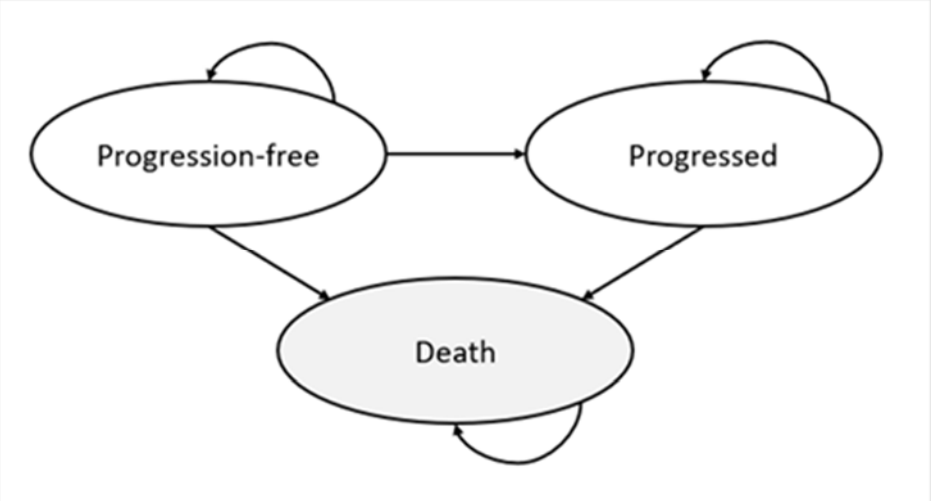
4.2.2 Model structure

A partitioned survival model (PSM) was developed in Microsoft Excel including three health states: a progression-free (PF) state, a progressed disease (PD) state, and death (see Figure 4.1). Model health states were selected in accordance with the clinical pathway of care and have also been used in previous later-line mCRC NICE technology appraisals[5, 34-36].

The allocation of patients into health states was directly based on PFS and OS curves that were fitted to the clinical trial data. All patients started the model in the PF state and remain in this state until disease progression or death. The company stated in the CS that the PF state captures the benefits from an active treatment whilst the disease is controlled prior to progression, leading to relatively higher QoL, while the PD state captures the relatively poor QoL following disease progression and prior to death.

A lifetime horizon (i.e., 10 years) with a weekly cycle length (including half-cycle correction) was applied to ensure all costs and QALYs were captured.

Figure 4.1: Model structure



Source: Based on Figure 14 of the CS

CS = company submission

ERG comment:

The main concern of the ERG relates to the use of a PSM without exploring a state transition model (STM) alongside it. The NICE Decision Support Unit (DSU) TSD 19 recommends the use of STMs alongside PSMs to verify the plausibility of PSM extrapolations and to explore key clinical uncertainties in the extrapolation period. In response to clarification question B2, the company stated that although the PSM approach did not explicitly assume structural dependence between PFS and OS, it did implicitly capture the dependence between PFS, and OS as observed in CORRECT and CONCUR. In addition, the company argued that the use of a PSM approach was consistent with prior NICE mCRC appraisals (e.g., TA405[5] and TA668[34], and given the relative maturity of the CORRECT and CONCUR data, they preferred to utilise this data directly using a partitioned survival approach rather than a STM reliant on an assumed structural dependence. The ERG considers the company's arguments to be reasonable and agrees on the appropriateness of the used PSM approach.

4.2.3 Population

The population considered in the CS was adults with mCRC who have failed on first-line chemotherapy/first-line biologic and who are being considered for ≥ third-line treatment. Specifically, patients for whom treatment with T/T is being considered. This population reflects the populations from the CORRECT and CONCUR trial, which was narrower than the anticipated license for regorafenib and the population in the final NICE scope.

ERG comment:

The main concerns of the ERG relate to a) the modelled population being narrower than the NICE scope, and b) part of the treatment population in both the CORRECT and CONCUR trials received anti-VEGF treatment.

a) The modelled population is more precisely defined than the final NICE scope. The ERG considers that if the population is defined essentially according to the comparator T/T, then only patients who

might be considered for T/T and no other treatment should be considered for regorafenib. More details regarding this issue are provided in Section 2.1 of the ERG report.

b) As indicated in the CS, anti-VEGF is not indicated for patients which fall into the scope of this submission in the UK. However, a considerable part of the treatment population in both the CORRECT and CONCUR trial received anti-VEGF treatment. The ERG requested a subgroup analysis for those with no prior anti-VEGF treatment in clarification question B1. The company responded that a subgroup analysis for the anti-VEGF naïve population versus T/T was not feasible but provided a post-hoc subgroup analysis by anti-VEGF treatment based on the CONCUR trial. The OS HR for patients without anti-VEGF was lower (HR: 0.470; 95% CI: 0.309, 0.714) compared to patients who received prior anti-VEGF (HR: 0.726; 95% CI: 0.430, 1.224) and these results are, according to the company, supportive of a greater potential to benefit in patients who have not received prior anti-VEGF treatment. Given that the post-hoc nature of this subgroup analysis and the fact that it did not include a (indirect) comparison to T/T, it remains unclear to the ERG what the potential impact of prior anti-VEGF treatment is on the CE results.

4.2.4 Interventions and comparators

The intervention considered in the CS was regorafenib. Consistent with the license, regorafenib was implemented in the economic model at a recommended dose of 160 mg (4 x 40 mg tablets) once daily for 3 weeks followed by 1 week off therapy.

The modelled comparators were T/T and BSC. The NICE scope listed the following comparators: single-agent irinotecan (after FOLFOX), FOLFIRI (after either FOLFOX or CAPOX), FOLFOX (after either FOLFIRI or CAPOX), raltitrexed (if 5-FU/FA are not suitable), T/T and BSC. The company justified the selection of comparators by stating that the listed treatments were available before regorafenib was licensed in mCRC, and therefore fall under the definition of "available therapies" in the license wording. T/T is orally administered at a dose of 35 mg/m2 twice daily, 5 days a week, with 2 days of rest, for 2 weeks, followed by a 14-day rest period. As per the trial protocols and licence in mCRC, treatment with regorafenib and T/T was continued until disease progression, clinical progression, the development of severe AEs, withdrawal from the study, death, or a decision by the treating physician that discontinuation would be in the patient's best interest.

ERG comment:

The main concerns of the ERG relate to a) the exclusion of single-agent irinotecan, FOLFIRI, FOLFOX, and raltitrexed as comparators, and b) the modelling of BSC as a comparator.

a) The ERG questioned whether the exclusion of single-agent irinotecan, FOLFIRI, FOLFOX, and raltitrexed was appropriate, given that these treatments were listed as comparators in the NICE scope. However, excluding FOLFOX and FOLFIRI based on their marketing authorisation seems reasonable, and other types of chemotherapy, irinotecan and raltitrexed, would probably also be excluded as they would be considered 'available therapies'. Therefore, the ERG considers the assumption that T/T is the only comparator, given that the company essentially defined the population to be only those patients who would be eligible for T/T, to be reasonable.

b) Next to T/T, the economic model also included BSC as a comparator. In the CS the company stated that this was mainly done because BSC was the comparator in the pivotal trials for regorafenib, rather than it being considered directly relevant to the requested position. In response to clarification question B4, the company confirmed that BSC is not considered to be a comparator: "*we are seeking a recommendation for regorafenib as a treatment option alongside trifluridine/tipiracil (i.e., the*

position for regorafenib in clinical guidelines). This submission was made in response to physician's requests for an alternative to trifluridine/tipiracil. Therefore, BSC isn't a comparator".

4.2.5 Perspective, time horizon and discounting

The analysis is performed from the NHS and Personal Social Services (PSS) perspective. Discount rates of 3.5% are applied to both costs and benefits. The model cycle length is 1 week with a lifetime time horizon (10 years) and a half-cycle correction is applied.

ERG comment:

The approach is in concordance with the NICE reference case.

4.2.6 Treatment effectiveness and extrapolation

The main sources of evidence on treatment effectiveness used for regorafenib and BSC were the CORRECT[14] (NCT01103323) and CONCUR[15] (NCT01584830) trials. Both were phase III, multicentre, randomised, double-blind, placebo-controlled trials in adult patients with stage IV mCRC that had progressed after standard treatments, evaluating regorafenib + BSC versus placebo + BSC. The data cut-offs used in CS base-case analyses were from January 2011 and November 2013 for the CORRECT and CONCUR trials, respectively. Pooled data of the CORRECT and CONCUR trials were used in the company's base-case to model OS, PFS, and ToT. As per CS base-case, the company chose to pool the data, because neither of the two trials were completely generalisable to the UK setting

The relative effectiveness of regorafenib compared to T/T was estimated via an ITC based on three phase III multicentre, randomised, double-blind, placebo-controlled trials: the RECOURSE[16] (T/T, n=534; placebo n= 266), TERRA[17] (T/T, n=271; placebo, n= 135) and Yoshino 2012[18] (T/T, n=135; placebo n= 157) trials.

In the company's base-case, PFS and time on treatment (ToT) were modelled using KM data when available and applying parametric survival models when KM data was not available anymore (full parametric modelling was explored for PFS and ToT in scenario analyses). OS was estimated by fitting parametric survival models to the KM data for regorafenib. For the modelling of OS in the company's base-case (and PFS and ToT modelling in scenario analyses), seven standard parametric survival models were considered (exponential, Weibull, Gompertz, log-normal, log-logistic, gamma, and generalised gamma). These models were assessed with regards to their statistical fit (based on Akaike information criterion (AIC) and Bayesian information criterion (BIC)) as well as their visual fit to the KM curves. In addition, for OS, the clinical plausibility of the extrapolated part of the model was justified with clinical expert opinion.

The process of selecting the most appropriate survival models for OS, PFS and ToT is summarised in Table 4.4. HRs estimated from the ITC were used to model OS (HR 0.99) and PFS (HR 0.93) for T/T. The company assumed the HR used for the modelling of PFS to be a proxy for the modelling of T/T ToT due to a lack of data availability.

Table 4.4: Criteria for choice of survival models.

	OS	PFS	ToT
General considerations	All No treatment waning was applied in the model (CS, Table 24). T/T No models were fitted to data for T/T. OS was modelled by applying the ITC HRs versus regorafenib to the regorafenib OS model.	All No treatment waning was applied in the model (CS, Table 24). T/T No models were fitted to data for T/T. PFS was modelled by applying the ITC HRs versus regorafenib to the regorafenib PFS model.	All The model assumed that ToT on average would be similar but slightly shorter than PFS, as expected in practice. T/T No models were fitted to data for T/T. ITC could not be performed for ToT given the limited available data. Hence, the company assumed the HR of ToT to be equal to the HR of PFS.
Fit to the observed data based on AIC and BIC	Regorafenib The AIC and BIC indicate that the log-logistic distribution has the best fit. The best following options are the generalised gamma (with 5-point difference on AIC and 9-point difference BIC) and the log-normal (with 8-point difference on AIC and BIC) (clarification response Table B8.4) BSC The AIC and BIC indicate that the log-normal distribution has the best fit. The best following options are the generalised gamma (with 5-point difference on AIC and 2 points difference on BIC) and log-logistic (with 6-point difference on AIC and BIC) (clarification response Table B8.4).	Regorafenib The AIC and BIC indicate that the log-logistic distribution has the best fit. The best following options are the log-normal (with 3-point difference on AIC and BIC) and the generalised gamma (with 3-point difference on AIC and 8 points difference on BIC) (clarification response Table B8.5). BSC The AIC and BIC indicate that the log-logistic distribution has the best fit. The best following options are the log-normal (with 52-point difference on AIC and BIC) and the generalised gamma (with 55-point difference on AIC and 51-point difference on BIC) (clarification response Table B8.5).	Regorafenib The AIC and BIC indicate that the log-logistic distribution has the best fit. The best following options are the log-normal (with 6-point difference on AIC and BIC) and the generalised gamma (with 6-point difference on AIC and 10 points difference on BIC) (clarification response Table B8.6). BSC The AIC and BIC were not updated in the clarification response. As per CS, the AIC and BIC indicate that the log-logistic distribution has the best fit. The best following options are the generalised gamma (with 39-point difference on AIC and 36-point difference on BIC) and the log-normal (with 42-point difference on

	OS	PFS	ToT
	T/T CS and clarification response did not include this information for T/T.	T/T CS and clarification response did not include this information for T/T.	AIC and BIC) (CS Appendix N, Table 98). T/T CS and clarification response did not include this information for T/T.
Fit to the observed data based on visual comparison with the KM curves	Regorafenib Log-normal distribution fits the KM data best, followed by the log-logistic and generalised gamma (CS, Figure 15). Long-term extrapolations (CS, Figure 16) were not explicitly discussed. BSC The log-normal, log-logistic, and generalised gamma were considered to have the best visual fit to KM (CS, Figure 17). Long-term extrapolations (CS, Figure 18) were not explicitly discussed. T/T No models were fitted to data for trifluridine/tipiracil.	Regorafenib Visual fit was illustrated in Figure 24 of Appendix N, and clarification response Figure B8.9. Given its subjective nature, it was not explicitly discussed. BSC Visual fit was illustrated in Figure 25 of Appendix N, and clarification response Figure B8.10. Given its subjective nature, it was not explicitly discussed. T/T No models were fitted to data for trifluridine/tipiracil.	Regorafenib Visual fit was illustrated in Figure 26 of Appendix N, and clarification response Figure B8.11. Given its subjective nature, it was not explicitly discussed. BSC Visual fit was illustrated in Figure 27 of Appendix N, and clarification response Figure B8.12. Given its subjective nature, it was not explicitly discussed. T/T No models were fitted to data for trifluridine/tipiracil.
Clinical plausibility of the extrapolation based on comparison with historical data	Regorafenib Clinical input considered all models clinically plausible and recommended alignment with TA405. Median OS from TA405 (7.4 months) was close to predicted median OS of all distributions, which ranged from [redacted] to [redacted] months (CS, Table 26). Median OS from trial	Regorafenib Not explicitly discussed, but clinical input considered all models clinically plausible and recommended alignment with TA405. BSC	Regorafenib Not explicitly discussed, but clinical input considered all models clinically plausible and recommended alignment with TA405. BSC

	OS	PFS	ToT
	data (6.9) was slightly lower (clarification response part 2, Q8f). BSC Not explicitly discussed, but clinical input considered all models clinically plausible and recommended alignment with TA405. T/T Clinical input considered all models clinically plausible and recommended alignment with TA405. Median OS from TA405 (7.4 months) was close to predicted median OS of all distributions, which ranged from [redacted] to [redacted] months (CS, Table 26). Mean OS of log-normal ([redacted]) and log-logistic ([redacted]) were the closest to the mean OS from TA405 (11.1 months) (CS, Table 26).	Not explicitly discussed, but clinical input considered all models clinically plausible and recommended alignment with TA405. T/T Not explicitly discussed.	Not explicitly discussed, but clinical input considered all models clinically plausible and recommended alignment with TA405. T/T Not explicitly discussed.
Clinical plausibility of the extrapolation based on trial data	Regorafenib Median OS, and 6 months and 1 year OS were relatively similar among all distributions. 5-year OS for the log-normal (1.5%) and log-logistic (1.98%) models were higher than trial data (0%). BSC Median OS, and 6 months, 1 year and 5-year OS were relatively similar among all models. 5-year OS for the log-normal (0.25%) and log-logistic (0.82%) and	Regorafenib As per the CS, the exponential extrapolation was closest to the KM data. Clarification response QB8f, showed that median PFS, and 6 months, 1 year and 5-year PFS were relatively similar among all distributions. BSC As per the CS, the exponential extrapolation was closest to the KM data. Clarification response QB8f, showed that	Regorafenib Clarification response QB8f, showed that median ToT, and 6 months, 1 year and 5-year ToT were relatively similar among all models. However, log-normal (4.25%) and log-logistic (4.63%) were closest to the 1-year ToT trial data (4.68%). BSC Clarification response QB8f, showed that median ToT, and 6 months, 1 year and 5-year ToT were relatively similar among

	OS	PFS	ToT
	generalised gamma (0.48%) model were higher than trial data (0%). T/T Not explicitly discussed.	median PFS, and 6 months, 1 year and 5-year PFS were relatively similar among all models. However, almost all 6 months PFS (2.08%) and 1 year OS (0.22%) values lower for parametric models than trial data. T/T Not explicitly discussed.	all models. However, almost all 6 months ToT (3.07%) values lower for parametric models than trial data. T/T Not explicitly discussed.
Clinical plausibility of the extrapolation based on clinical expert opinion	Regorafenib Experts considered all models plausible. Clinical experts recommended the log-logistic as it was in line with TA405[5] BSC As per the second part of the clarification response (Q8b), clinical experts agreed that all models resulted in plausible survival predictions. T/T Modelled OS results were considered similar to the T/T results reported in TA405[5] (CS, Table 26). For mean OS, log-logistic and log-normal were considered to be most in line with TA405[5]	Regorafenib As per the second part of the clarification response (Q8b), clinical experts agreed that all models resulted in plausible survival predictions. BSC As per the second part of the clarification response (Q8b), clinical experts agreed that all models resulted in plausible survival predictions. T/T Not explicitly discussed.	Regorafenib As per the second part of the clarification response (Q8b), clinical experts agreed that all models resulted in plausible survival predictions. BSC As per the second part of the clarification response (Q8b), clinical experts agreed that all models resulted in plausible survival predictions. T/T Not explicitly discussed.
Base-case approach	Regorafenib Log-logistic. BSC	Regorafenib For the CS base-case KM data were used directly followed by an exponential extrapolation for the remainder of the	Regorafenib For the CS base-case KM data were used directly followed by a log-logistic extrapolation for the remainder of the

	OS	PFS	ToT
	Log-logistic. T/T To model T/T, the resulting ITC HRs versus regorafenib were applied to the regorafenib OS.	model. Parametric survival models were explored in scenario analyses. BSC For the CS base-case KM data were used directly followed by an exponential extrapolation for the remainder of the model. Parametric survival models were explored in scenario analyses. T/T To model T/T, the resulting ITC HRs versus regorafenib were applied to the regorafenib PFS.	model. Parametric survival models were explored in scenario analyses. BSC For the CS base-case KM data were used directly followed by a log-logistic extrapolation for the remainder of the model. Parametric survival models were explored in scenario analyses. T/T HR of ToT to be equal to the HR of PFS.

Source: CS Section B.3.3 and response to clarification letter

AIC = Akaike information criterion; BIC = Bayesian information criterion; BSC = best supportive care; CS = company submission; HRs = hazard ratios; ITC = indirect treatment comparison; KM = Kaplan-Meier; OS = overall survival; PFS = progression-free survival; PPP = platinum pre-treated population; ToT = time on treatment; T/T = trifluridine/tipiracil

ERG comment:

The main concerns of the ERG relate to: a) using KM data for the modelling of PFS and ToT, b) HR estimates for OS being different between the ITC conducted by the company and the direct comparison from a large observational study, c) missing details of TSD 14 criteria regarding the choice of parametric models, d) treatment waning and e) applying the PFS HR as a proxy for the modelling of ToT in the T/T arm.

a) The company modelled PFS and ToT for both regorafenib and BSC by directly using the KM data and applying a parametric survival model for the remainder of the model if KM data were no longer available. In line with NICE DSU TSD 14, which states that "*parametric models are likely to represent the preferred method for incorporating survival data into health economic models in the majority of cases*", the ERG prefers using fully parametric models. The 'stepped' nature of KM curves, resulting from protocol-driven follow-up, may introduce overfitting to the trial data and could affect the representativeness of the survival analyses results to UK clinical practice. This is particularly applicable in this case, as progression was only assessed every 8 weeks and median survival was short. Upon request for justification of using KM data, the company responded that KM data should be used as it is mature, complete and that it would reflect clinical practice as disease progression in clinical practice is also assessed every 8 weeks. Although the ERG acknowledges that the company's arguments for directly using KM data may be valid, in line with NICE DSU TSD 14[37], full parametric models are used in the ERG base-case to model PFS and ToT. Based on the log-cumulative hazard plots, statistical and visual fit to the observed data and expert opinion provided by the company, the ERG implemented log-logistic models for the modelling of PFS (for both regorafenib and BSC) and ToT (for regorafenib).
b) For the modelling of OS in the T/T arm, the company conducted an ITC and applied a HR of 0.99 for regorafenib versus T/T. The ERG identified a large observational cohort study including data from 269 hospitals in Japan, which directly compared the effectiveness of regorafenib (n=1,501) and T/T (n=3,777). Contrary to the ITC conducted by the company, the results of this study demonstrated an OS benefit for T/T compared to regorafenib (HR 0.66 (95% CI: 0.587 to 0.742)). Although the ERG acknowledges the limitations of observational studies, the comparability between the RCTs in the NMA was also questionable. Therefore, the ERG explored a scenario analysis in which the observed HR of T/T versus regorafenib was converted to the HR of regorafenib versus T/T (1.515 (95% CI: 1.348 to 1.704)) and applied to the economic model for the modelling of OS in the T/T arm. A scenario analysis applying the new HR to the OS of T/T resulted in negative incremental net monetary benefits (iNMBs) (willingness-to-pay (WTP) thresholds of 30,000 and 51,000 per QALY gained) for regorafenib versus T/T.
c) The company stated that for the choice of extrapolating OS, PFS, and ToT three criteria were taken into account: statistical fit (through AIC and BIC statistics), visual inspection and clinical validity. The arguments used for the choice of parametric models by the company were not based on all criteria described in the NICE DSU TSD 14 guidance[37] and additional information was therefore requested by the ERG. While the company provided additional details, as summarised in Table 4.4, some information is still lacking (mainly clinical plausibility of the extrapolation based on historical data, trial data and expert opinion). Further, the ERG requested the company to rank the parametric models according to their visual fit, but this was not provided. Despite some remaining lack of detail regarding the NICE DSU TSD 14 criteria, the ERG largely agreed with the company's choices of full parametric models provided for OS (in the company's base-case), PFS and ToT (in the company's scenario analyses). Nevertheless, in contrast to the company's approach of

modelling the OS of BSC using a log-logistic model, based on the provided TSD 14 assessment, the ERG preferred using the log-normal in its base-case.

d) The company did not include treatment waning in their base-case. The ERG requested justification and additional evidence (HR plots) supporting this assumption, as well as a scenario analysis exploring treatment waning in the economic model. The company did not provide the requested scenario analysis as it argued that assuming treatment waning was inappropriate due to the direct implementation of KM curves, the maturity of the data and the short duration of survival. Although the company ideally should have provided the requested HR plots and scenario analyses to fully address the uncertainty around treatment waning in this submission, the ERG agrees that based on the company's arguments this is likely a minor issue.

e) Due to insufficient publicly available ToT data for T/T, the company applied the PFS HR as a proxy for the modelling of ToT in the T/T arm. This assumption can only hold if disease progression and AE profiles are similar between regorafenib and T/T. Upon request for justification, the company provided the stopping rules of regorafenib and T/T and a NMA to demonstrate that reasons for treatment discontinuation were similar for regorafenib and T/T. While the ERG considers using the PFS HR (0.93) as a proxy for the modelling of ToT in the T/T arm to be questionable, it is likely a conservative assumption as ToT, and consequently also treatment costs are lower in the T/T arm relative to the regorafenib arm.

4.2.7 Adverse events (AEs)

The main sources of evidence on AEs were data from the CORRECT and CONCUR trials for regorafenib and BSC. The main publications from RECOURSE, TERRA, and Yoshino 2012 provided evidence utilised for T/T AE rates per cycle. The model considered AEs of Grade 3 and higher that occurred in at least 2% of patients in any treatment arm. The CS stated that "*this cut-off was chosen to ensure that infrequent, but costly or severe AEs are also considered in the model*."[3]

The average AE rate per cycle per treatment across different trials was calculated by combining the observed AEs and the number of patients. The AE rate per treatment cycle was calculated by dividing the average AE rate by the weighted average treatment duration and converted to a weekly probability (CS Table 29). To provide an average AE cost per patient per week, the weekly probabilities were combined with costs per AE. The average AE costs per patient per week were applied to each model cycle until progression (CS Table 36).

ERG comment:

The main concerns of the ERG relate to: a) the exclusion of Grade 1 and 2 AEs, and b) results from AE NMAs not being included in the CEA.

a) The company exclusively included Grade 3+ AEs that occurred in at least 2% of the population. Due to the company's positioning of regorafenib as being *"a chemotherapy-free alternative therapy with a different adverse event profile*" (page 25, CS), the ERG requested a scenario analysis including lower grade AEs. The company, however, did not include this scenario analysis in its response. The company suggested that AEs are not a driver of CE, and that Grade 1 and 2 AEs are not expected to impact costs or QoL. Time constraints and Grade 1 and 2 AEs not generally being modelled in oncology were also arguments to justify excluding the requested scenario analysis. Given the company's positioning of regorafenib as stated above, the ERG would prefer more evidence (e.g., an updated economic model and scenario analysis including Grade 1 and 2 AEs) to support the claim that Grade 1 and 2 AEs would not be impactful.

b) Upon the ERG's request, the company provided NMAs for all TEAEs, Grade 3+ TEAEs, and treatment discontinuation due to AEs in response to clarification question A28. The ERG further requested that all CEAs were updated with the results of the AE-related NMAs carried out. The company did not update the CE model as they did not deem the inclusion of these NMAs to be suitable. For the discontinuations due to AEs NMA, the company's results included odds ratios which could not be used for adjusting survival data. A CEA using NMA results for TEAEs was not conducted, as these mainly included Grade 1 and Grade 2 AEs, which the company expected not to have an impact on costs and QoL. To incorporate the results of the NMA for Grade 3+ AEs, the company suggested that one must assume the AEs observed for regorafenib are common for T/T (and vice versa), with the only difference being in the proportion of patients. The company cited Table A28.1 of the response to the clarification letter[6] as failing to support these assumptions and suggested the odds ratios used are not suitable for adjusting survival data. The ERG acknowledges the limitations related to the applying odds ratios for adjustment of the survival data and highlights that the odds of experiencing Grade 3 or 4 AEs and the odds of discontinuation of treatment due to AEs were similar between regorafenib and T/T. However, the NMA for all TEAEs suggested higher odds of patients treated with regorafenib compared to T/T, and not incorporating these results likely biased the CE results in favour of regorafenib.

4.2.8 Health-related quality of life (HRQoL)

Utility values were estimated for the following health states: progression-free health state, post-progression health state. In addition, utility decrements were included to capture the impact of AE.

HRQoL data identified in the review

A SLR identified alternative utility values for patients receiving ≥ third line treatment for mCRC and to validate those in CORRECT and CONCUR (CS Table 28). The SLR identified 13 studies, with three being relevant to UK population[5, 34, 38] (one unspecified country)[39]. None of the utility values identified in the SLR were used in the model.

Health state utility values (HSUV)

Utility values were derived through pooling EQ-5D-3L data collected in the COLLECT and CONCUR trials and were captured for patients on treatment and for patients at the end of treatment. Despite differences between regorafenib and BSC EQ-5D-3L data, the model assumed no treatment dependent utilities.

None of the trials informing the survival analysis of T/T captured EQ-5D data. According to the CS, PRECONNECT was the only study to capture QoL data for T/T[40], which matched the pooled utility values for pre- and post-progression from CORRECT and CONCUR. As such, T/T utility was assumed to be equal to regorafenib and BSC.

Pre-progression and post-progression utility were assumed to be equal to the pooled on-treatment and pooled end of treatment utilities, respectively. A summary of all utility values used in the CEA is provided in Table 4.5.

Disutility values

AE disutilities were applied in the company's base-case for regorafenib, T/T, and BSC. In the CS, it was stated that, whilst AE disutility was likely to have been captured within the observed utilities in the trials, most AEs were transient in nature. As such, it is uncertain whether any disutility from the experienced AE would be present on the day the EQ-5D was administered[3].

AE disutilities were derived from past TAs (CS Table 30). To calculate the average AE disutility per treatment, disutilities were combined with the pooled weekly AE probabilities. The model assumed an AE duration of one week (in line with TA405)[5]. Therefore, the average AE disutility was directly subtracted from the pre-progression utility to generate treatment-specific utilities.

Table 4.5 shows the utilities, with the applied AE disutilities in the company base-case:

Table 4.5: Utility values for cost-effectiveness analysis (CEA)

	Regorafenib	Trifluridine/tipiracil	BSC
Progression-free			
PFS utility (standard error)*	0.72 (0.005)	0.72 (0.005)	0.72 (0.005)
AE disutility applied	-0.00361	-0.0077	-0.00124
Final PFS utility applied in the model	0.716	0.712	0.719
Progressed			
Final PPS utility applied in the model (standard error)*	0.59 (0.014)	0.59 (0.014)	0.59 (0.014)

Source: CORRECT[14], CONCUR[15], Sabater 2019[40]; based on CS Table 31[3]
AE = adverse event; BSC = best supportive care; CEA – cost-effectiveness analysis; CS = company submission; PFS = progression-free survival; PPS = post-progression survival; T/T = trifluridine/tipiracil
*Health state utilities were assumed to be the same for regorafenib, T/T, and BSC

ERG comment:

The main concerns of the ERG relate to: a) using end-of-treatment utility as a proxy for the modelling of post-progression survival (PPS) utility; and b) differences in estimated utility values between CORRECT and CONCUR trials.

a) Uncertainty exists as to the plausibility of using the pooled end-of-treatment utility as a proxy for the PPS health state. The ERG requested justification for using these proxy utility values, and for the plausibility of using a relatively low PPS utility value, when compared with those identified in the literature. In the company response[6], the PPS utility was justified due to being within the range of values identified in the literature review and being equal to the PPS value in TA405. In addition, the company stated that most patients would stop treatment upon disease progression. An updated economic model and scenario analyses were requested, informing PPS utility as derived from other relevant TAs, with justification as to how these compare to the initial end-of-treatment utility. The analyses were not provided by the company as the utility values used were suggested to be appropriate and closely aligned to TA405. The company did explore the impact of using the highest PFS utility (0.810) and lowest PPS utility (0.5) from other relevant TAs, which had minimal impact on the CE results. The company suggested that this shows that PFS and PPS utilities are not a driver of CE, stating that this is expected given the comparable efficacy of regorafenib and T/T. Although the ERG questions the plausibility of using the pooled end-of-treatment utility from CONCUR and CORRECT as a proxy for the PPS health state utility, it considers the impact of this assumption to be likely minor.
b) Health state utility values (HSUV) in the economic model were derived through pooling of the EQ-5D-3L index scores from CORRECT and CONCUR. However, slight differences exist between the

utility values captured in the CORRECT and CONCUR trials, and therefore, the ERG requested an updated model and scenario analyses including utility values for each trial separately. The company did not provide these analyses, justifying this due to pooling being stated as the most appropriate method and suggesting the model was not sensitive to the analysis in which the difference in pre- and post-progression utility values were widened. Whilst the ERG would like to have seen the impact of the scenario analyses, it agrees that the impact is likely to have been marginal and considers the issue to be minor.

4.2.9 Resources and costs

The cost categories included in the model were treatment costs for regorafenib and T/T, health-state unit costs, adverse reaction unit costs, and miscellaneous unit costs, including end-of-life and subsequent treatment costs.

Unit prices were based on the NHS reference costs 2019-20, British National Formulary (BNF) and the Personal Social Services Research Unit (PSSRU) 2021. Past NICE technology appraisals were used when a corresponding NHS reference cost was not available. NHS reference costs were inflated to 2021 value using the PSSRU price index.

Resource use and costs data identified in the review

According to the CS, the SLR identified a total of four studies that reported relevant healthcare resource use. Out of these, the resource costs from NICE TA405[5]; have been utilised in the model, as it evaluated T/T.

Treatment acquisition costs

All treatment costs were based on the BNF 2021, as per NICE reference case which is presented in Table 4.6. Regorafenib was given in 28-day treatment cycles, which included 160 mg (4 x 40 mg tablets) daily dose for 3 weeks followed by one rest week and was offered at a price (including PAS) of [redacted] per cycle. According to the expected treatment use in UK clinical practice by clinical experts, the full costs of the ongoing cycle were applied in the model even if the patient stopped treatment at any point of the (28-day) cycle.

T/T was given in 28-day treatment cycles consisting of 20 doses; 2 weeks of active treatment at 35 mg/m^2 twice daily for 5 days, two rest days per week, and then 2 weeks of rest. T/T dosing was based on patients' BSA (CS, Table 32). The calculated average dose assumed that 55.8% of mCRC patients were male, and the UK cancer patient's BSA distribution from Sacco was applied.[41] The average dose in the CS base-case was 1.48 x 15 mg and 2.00 x 20 mg tablets, with a cost of £2,071.00 per cycle. As per regorafenib, all costs were incurred at the start of a cycle, and wastage was applied when patients discontinued treatment.

BSC treatment costs were assumed to be £0, as the company argued that they would have been captured by BSC HRU costs, which is in line with past mCRC appraisals.

Table 4.6: Treatment costs (with regorafenib PAS)

Name	Formulation	Price per tablet	Number of tablets per dose	Doses per treatment cycle*	Cost per treatment cycle	Final cost per treatment cycle
Regorafenib	40 mg per tablet	£44.57	4.00	[redacted]	£3,744.00	[redacted]

Name	Formulation	Price per tablet	Number of tablets per dose	Doses per treatment cycle*	Cost per treatment cycle	Final cost per treatment cycle
T/T	15 mg tablet	£25.00	1.48	20.00	£741.27	£2,071.38
	20 mg tablet	£33.33	2.00	20.00	£1,330.11	

Source: CS model, cost per dose tab
CS = company submission; mg = milligrams; PAS = Patient Access Scheme; T/T = trifluridine/tipiracil
* 28-day treatment cycle [1]

The model included a RDI for regorafenib, which also included cycle delay, based on the weighted average of the mean dose patients received in CORRECT ([redacted]) and CONCUR ([redacted]). A scenario analysis was applied in which RDI was based on the number of pills dispensed.

For T/T, the dose reduction and cycle delays were modelled separately using data from TA405, as no RDI measure was reported. According to the company, this different modelling strategy of combining dose reductions and cycle delays approximated how RDI was assessed for regorafenib, and it would be more reflective of the clinical practice. The dose of regorafenib would be generally reduced if toxicities develop, whereas T/T would be delayed. Hence, it was assumed in the CS base-case, that all T/T dose reductions were already applied during the first dose and continued for the full course of treatment; in practice, the dose would decline gradually. A scenario analysis assuming equal RDI between regorafenib and T/T was explored to address the uncertainty around T/T RDI estimates.

Administration costs

No administration costs were included in the model since all the active treatments were oral and any routine visit cost was assumed to be covered by the HRU costs.

Health state costs

Health resource use estimates used in the model were obtained from the four studies[5, 34, 38, 42] identified by the SLR. According to the CS, despite having different active treatments, the chosen HRU rates were comparable across the studies. Nonetheless, the CS base-case only applied the ERG preferred rates from TA405 that combined both active treatment rates and used lower BSC rates. As per CS, HRU resources were validated by a clinical advisory board who agreed on using the ERG-preferred values from TA405, which is summarised in CS Table 35. Oral chemotherapy, CT scan, GP home visits, nurse specialist visits and other costs were included in the model. The predicted resource use is summarised in Table 4.7, while the costs associated with the PF and PD states are listed in Table 4.8.

Table 4.7: Summary of predicted resource use

Treatment	Active treatment	AE costs	Administration costs	mCRC management costs	End-of-life costs	Total
Regorafenib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
T/T	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Source: CS Appendix I, Table 75
AE = adverse event; BSC = best supportive care; CS = company submission; mCRC = metastatic colorectal cancer; T/T = trifluridine/tipiracil

Table 4.8: Summary of total costs by health state

Treatment	Progression-free	Progressed
Regorafenib	[redacted]	[redacted]
T/T	[redacted]	[redacted]
BSC	[redacted]	[redacted]
Source: CS Appendix I, Table 74 BSC = best supportive care; CS = company submission; T/T = trifluridine/tipiracil		

Adverse event (AE) costs

AE costs were applied each treatment cycle in the model and were calculated based on the pooled weekly AE probabilities and the cost per AE as reported in table 4.9.

Table 4.9: Weekly AE costs per treatment

Treatment	Costs	Source
Regorafenib	£19.18	Pooled CORRECT and CONCUR AE probabilities, regorafenib arm
T/T	£39.95	Pooled RECOURSE, TERRA, and Yoshino AE probabilities, trifluridine/tipiracil arm
BSC	£3.10	Pooled CORRECT and CONCUR AE probabilities, BSC arm
Source: CS, Table 37. AE = adverse event; BSC = best supportive care; CS = company submission; T/T = trifluridine/tipiracil		

End-of-life costs

A one-off end-of-life cost (£6,832.17) was applied to all patients transitioning to the death state in the CS base-case. This cost was derived from the Round 2015[43] modelling study on people with CRC and only included healthcare and social care costs.

Subsequent treatment

The company stated that, according to the clinical experts, the proportion of patients receiving regorafenib or T/T who might receive further treatment is <10%; therefore, no post-progression treatment was assumed in the base-case. Subsequent treatment costs from NICE TA405[5] were used in a scenario analysis to explore their impact on the CE estimates.

ERG comment:

The main concerns of the ERG related to a) the modelling of missed doses and dose reductions for regorafenib and T/T, b) subsequent treatment assumption, c) exclusion of relevant studies informing HRU rates that were identified in the SLR and d) lack of justification for BSC costs.

a) The company used different approaches to model missed doses and dose reductions for regorafenib ([redacted] RDI) and T/T (97.4% RDI and 2.72 days cycle delay). For regorafenib, missed doses and dose reductions were modelled as a single RDI measure, which was calculated by the weighted average of the mean dose as reported in CONCUR and CORRECT. The company stated that for T/T no single RDI measure has been reported, and dose reductions and cycle delays were therefore modelled separately instead, with data on the number of dose reductions and cycle delays from TA405[5]. In response to clarification question B18a, the company stated that for both regorafenib

and T/T, it is recommended to manage specific AEs through dose reduction and/or delays. However, regorafenib tends to have more reductions, while T/T have more delays, and the different implementation of RDI was driven by data availability. However, the ERG identified a large observational study by Nakashima 2020[1] directly comparing regorafenib to T/T, which reported comparable (54% and 48% for regorafenib and T/T respectively) dose reductions. Although the ERG acknowledges the limitations of observational studies, the comparability between the RCTs in the NMA was also questionable. The real-world data from the observational study suggest comparable dose reductions for regorafenib and T/T, which contradicts the company's statement that regorafenib tends to have more reductions, while T/T have more delays. Therefore, the ERG assumed the T/T RDI in its base-case to be equal to the RDI of regorafenib.

b) Although post-progression treatment was given to a substantial number of patients in the CORRECT (regorafenib 26%, BSC 30%) and CONCUR (regorafenib 31%, BSC 43%) trials, the company stated that, according to clinical experts, the proportion of patients receiving regorafenib or T/T who might receive further treatment was <10%. As a result, the company assumed no post-progression treatment in the company base-case. Nonetheless, a scenario analysis used the subsequent costs reported in TA405 inflated to 2021 (£1,633.18) as a one-off cost to both regorafenib and T/T patients upon progression. More details regarding this issue are provided in Section 2.1 of the ERG report. The ERG considers that the uncertainty surrounding subsequent treatment use remains and cannot be reduced without the data on subsequent treatment use in the T/T trials.

c) The ERG questioned that despite identifying four studies reporting relevant HRU rates in the SLR, the HRU rates used in the company's base-case were solely based on TA405. According to the company, this study "represents a good middle ground". As per CS Table 34, monthly HRU rates of the different identified studies were compared, showing different percentages and one extra category (medical oncologist OP visit) that were not considered in TA405. In response to clarification question B20, the company stated that the preferred HRU rates from TA405 were simply presented and then validated by the clinical experts as appropriate for both T/T and regorafenib, and the values from Bullement 2018[42], Hoyle 2013[38], and TA668[34] were not discussed. The company also stated that clinical experts assumed that HRU did not differ between regorafenib and T/T. Furthermore, it was stated that the medical oncologist visits were not part of the ERG preferences in TA405, and that the model is not sensitive to the inclusion/exclusion of this cost, nor will it impact the iNMB of regorafenib versus T/T. Although the ERG agrees it's likely a minor issue, it would like to see an updated economic model and scenario analyses exploring HRU rates informed by Bullement 2018,[42] Hoyle 2013,[38] and TA668[34].

d) In the CS base-case, treatment costs for BSC were assumed to be £0, as BSC would have been captured by BSC HRU costs. However, the CORRECT and CONCUR trials included several concomitant medications and treatments as BSC. As per clarification response to question B21a, the company stated that they were not able to perform a costing exercise for BSC in the UK and instead they performed a pragmatic scenario analysis with an assumed BSC cost of £50 per 28-day treatment cycle, which they considered on the high scale. Including these costs had a negligible effect on the CE results. The ERG is satisfied with the analysis provided by the company and agrees that the company's base-case approach assuming zero costs to BSC is conservative.

4.2.10 Severity

Due to the severity of the disease, patients suffering from ≥ third line mCRC experience a substantial QALY shortfall, compared to the general population. The expected total QALYs for the general population, with age = 60 year and 56% males, was estimated to be 12.36. The remaining QALYs for patients with mCRC were estimated using the base-case model results, this was [redacted] for T/T and

[redacted] for BSC. Consequently, the proportional shortfall for ≥ third line mCRC patients is [redacted] and [redacted] for patients currently treated with T/T and BSC, respectively, justifying a 1.7 x QALY weight for both comparisons. This QALY weight will be applied indirectly in the base-case by using a higher WTP threshold of £51,000.

ERG comment:

The ERG reproduced the shortfall analysis reported in CS Section B.3.6. The reported proportional QALY shortfall (CS Table 40) and the 1.7 x QALY weight were successfully reproduced. Moreover, in response to clarification question B26, the company provided the estimated absolute shortfall of 11.79 QALYs.

5. COST EFFECTIVENESS RESULTS

5.1 Company's cost effectiveness (CE) results

The probabilistic CS base-case results indicated that regorafenib was more effective ([redacted]) and [redacted] ([redacted]) compared to T/T, [redacted] (iNMB [redacted]). Compared to BSC, patients treated with regorafenib accrued an additional [redacted] QALYs at an additional cost of [redacted], which resulted in an ICER of [redacted] per QALY gained (iNMB [redacted]) (Table 5.1). At a WTP threshold of £51,000, corresponding to a QALY weight of 1.7, [redacted].

Table 5.1: Company's probabilistic base-case results

Treatment	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER	iNMB[a]
Regorafenib	[redacted]	[redacted]	-	-	-	-
T/T	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Source: CS Table 45
BSC = best supportive care; CS = company submission; ICER = incremental cost-effectiveness ratio, iNMB = incremental net monetary benefit; T/T = trifluridine/tipiracil; QALY = quality adjusted life year; WTP = willingness-to-pay
[1] £30,000/QALY WTP threshold

Overall, the model is set to affect QALYs:

- Compared to T/T, regorafenib increased the progression-free life-years by [redacted]. This resulted in a QALY increase of [redacted].
- Compared to BSC, regorafenib increased progression-free and progressed disease life-years by [redacted] ([redacted] years pre-progression and [redacted] years post-progression). This resulted in an overall QALY increase of [redacted].

Overall, the model is set to affect costs:

- Compared to T/T, regorafenib [redacted] ([redacted]). This cost comprised [redacted] of the incremental costs.
- Compared to BSC, regorafenib mainly increased the treatment costs by [redacted] and increased pre-progression care by [redacted], together comprising [redacted] of the incremental costs.

ERG comment:

The ERG's main concern linked to the company's CE results is related to information about restricted mean survival time (RMST). Upon request by the ERG to provide a table with information about the RMST, the company responded that they were unable to do so due to time pressure.

5.2 Company's sensitivity analyses

The company performed and presented the results of probabilistic sensitivity analyses (PSA), deterministic sensitivity analyses (DSA) as well as scenario analyses.

The parameters that had the greatest effect on the ICER based on the company's DSA compared to T/T were:

- The HR used for OS
- The HR used for PFS
- The RDI of regorafenib

The parameters that had the greatest effect on the ICER based on the company's DSA compared to BSC were:

- The RDI of regorafenib
- The proportion of patients in the regorafenib arm which received daily chemotherapy
- The cost of AE management in the regorafenib arm.

The company further conducted several scenario analyses, of which most were only presented for the comparison with T/T. The following were most influential:

- Using only the CORRECT and RECOURSE trials to inform the ITC [redacted] the iNMB to [redacted].
- Using only the CONCUR and Yoshino trials to inform the ITC [redacted] the iNMB to [redacted].

ERG comment:

Not all scenario analyses that were provided for the comparison with T/T were also provided for the comparison with BSC. In its clarification letter, the ERG requested the company to also conduct all scenario analyses versus BSC, but the company did not provide these stating that BSC was not considered a relevant comparator. In addition, the ERG requested the company to conduct all scenario analyses with the fully parametric models applied for PFS and ToT, which were also not provided by the company due to time constraints.

5.3 Model validation and face validity check

5.3.1 Face validity assessment

The clinical validity of the model and assumptions were validated by UK clinical experts during various smaller online and offline one-to-one interactions and an advisory board involving nine clinical experts. In addition, an external health economic validation meeting was organised to confirm whether the modelling approach was appropriate for the decision problem. According to the company, both the clinical advisory board and the external health economic validation meeting supported the modelling approach (no reference to the advisory board notes was made in CS Section B3.14.1).

5.3.2 Technical verification

To verify the results of the CE model, internal quality control procedures were undertaken by the model developers to ensure that the mathematical calculations were performed correctly and were consistent with the model specifications. Moreover, health economists not involved in the development of the model reviewed the model for coding errors, inconsistencies and the plausibility of inputs and results.

The company did also subject the model to a checklist of known modelling errors, and the assumptions have been questioned. This involved checks on the selection and results of different modelling options,

calculation spot checks, cross checks against source data and extreme value scenarios to check if the model behaved logically. The validation identified no major issues with the computational accuracy of the model. Several small inaccuracies were identified and rectified by the company.

5.3.3 Comparisons with other technology appraisals

The company stated that TA405 served as an important anchor point as it is the key comparator (T/T) for this appraisal. CS Table 26 provides a cross comparison of mean and median OS for T/T in the current assessment and TA405 while this is provided for health state utilities and resource use in CS Tables 28, 34 and 35.

5.3.4 Comparison with external data used to develop the economic model

As part of the validation process, model outcomes for regorafenib and BSC were compared to the pooled clinical trial data (CORRECT and CONCUR). CS Figures 26 and 27 seem to support that PFS and OS estimated with the model is consistent with the data used to develop the economic model.

5.3.5 Comparison with external data not used to develop the economic model

Comparisons with external data not used to develop the economic model are not discussed in CS section B.3.14.3.

ERG comment:

The main concerns of the ERG relate to a) differences between probabilistic results likely due to the lack of a fixed random seed in the economic model; b) the technical verification; c) cross validation and d) face validity assessment.

a) The results of the PSA are slightly different when running the same analysis multiple times (without changing model settings). This is likely due to the lack of a fixed random seed in the model PSA, which results in slightly different random draws each time the model runs.
b) Although the company did not specifically mention all tests/checks used to verify the technical implementation of the model (in CS Section B3.14.2), the company did respond to clarification question B29 that the "*model was validated using the Lumanity checklist: a quality control procedure developed using publicly available checklists from Drummond and Philips. The Lumanity checklist includes all checks listed in the published TechVER checklist*" and provided example of tests/checks that were performed. Moreover, the company stated that on "*completion of the QC, the model was updated to correct any identified problems*". Therefore, the ERG is reassured that the technical validity is appropriately verified.
c) Clarification question B30 asked the company to provide a cross validations, i.e., comparisons with other relevant NICE TAs focussed on similar, potentially relevant, diseases (e.g., related NICE recommendations and NICE pathways listed in the final scope, including those mentioned in CS Table 24) and elaborate on the identified differences regarding. The company did not provide further information in response to this clarification question.
d) Clarification question B31 asked the company to report on the face validity assessment (mentioned in CS Section B.3.14.1) of the model structure, model assumptions, model inputs, intermediate outcomes as well as final outcomes in more detail (including what aspects were assessed and what were the considerations as well as conclusions). The company did not provide further information in response to this clarification question other than mentioning that "*The model has a structure which is common in oncology and is entirely appropriate for comparing T/T and regorafenib. The validity of the model structure was confirmed by clinical experts and matches the model structure used in TA405*".

6. EVIDENCE REVIEW GROUP'S ADDITIONAL ANALYSES

6.1 Exploratory and sensitivity analyses undertaken by the ERG

Table 6.1 summarises the key issues related to the CE categorised according to the sources of uncertainty as defined by Grimm 2020[44]:

- Transparency (e.g., lack of clarity in presentation, description, or justification)
- Methods (e.g., violation of best research practices, existing guidelines, or the reference case)
- Imprecision (e.g., particularly wide CIs, small sample sizes, or immaturity of data)
- Bias and indirectness (e.g., there is a mismatch between the decision problem and evidence used to inform it in terms of population, intervention/comparator and/or outcomes considered)
- Unavailability (e.g., lack of data or insight)

Identifying the source of uncertainty can help determine what course of action can be taken (i.e., whether additional clarifications, evidence and/or analyses might help to resolve the key issue). Moreover, Table 6.1 lists suggested alternative approaches, expected effects on the CE, whether it is reflected in the ERG base-case as well as additional evidence or analyses that might help to resolve the key issues.

Based on all considerations in the preceding Sections of this ERG report, the ERG defined a new base-case. This base-case included multiple adjustments to the original base-case presented in the previous Sections. These adjustments made by the ERG form the ERG base-case and were subdivided into three categories (derived from Kaltenthaler 2016)[45]:

- Fixing errors (FE) (correcting the model where the company's submitted model was unequivocally wrong)
- Fixing violations (FV) (correcting the model where the ERG considered that the NICE reference case, scope or best practice had not been adhered to)
- Matters of judgement (MJ) (amending the model where the ERG considers that reasonable alternative assumptions are preferred)

6.1.1 ERG base-case

Adjustments made by the ERG, to derive the ERG base-case (using the CS base-case as starting point) are listed below. Table 6.2 shows how individual adjustments impact the results plus the combined effect of all abovementioned adjustments simultaneously, resulting in the ERG base-case. The 'FE' adjustments were combined, and the other ERG analyses were performed also incorporating these 'FE' adjustments given the ERG considered that the 'FE' adjustments corrected unequivocally wrong issues.

Fixing errors (FE)

There were no errors identified by the ERG.

Fixing violations (FV)

There were no violations identified by the ERG.

Matters of judgement (MJ)

1. Alternative parametric survival curve for OS in BSC arm (Section 4.2.6): Instead of a log-logistic curve, the ERG implemented a log-normal curve for the modelling of OS in the BSC arm.

2. Implementation of parametric survival curves for PFS in regorafenib and BSC arms (Section 4.2.6): In line with the company's scenario analyses, the ERG implemented a log-logistic curve for the modelling of PFS in the regorafenib and BSC arms.
3. Implementation of a parametric survival curve for ToT in the regorafenib arm (Section 4.2.6): In line with the company's scenario analysis, the ERG implemented a log-logistic curve for the modelling of ToT in the regorafenib arm.
4. Equal RDI for regorafenib and T/T (Section 4.2.9): In line with the company's scenario analysis, the ERG applied the pooled RDI of [redacted] to both regorafenib and T/T.

6.1.2 ERG exploratory scenario analyses

The ERG performed the following exploratory scenario analyses to explore the impact of alternative assumptions conditional on the ERG base-case.

Exploratory scenario analyses

5. Alternative HR for OS in the T/T arm (Section 4.2.6): instead of an OS HR of 0.99, the ERG implemented an OS HR of 1.515.

6.1.3 ERG subgroup analyses

No subgroup analyses were performed by the ERG.

Table 6.1: Overview of key issues related to the CE (conditional on fixing errors highlighted in Section 5.1)

Key issue	Section	Source of uncertainty	Alternative approaches	Expected impact on CE[a]	Resolved in ERG base-case[b]	Required additional evidence or analyses
Implementation of parametric survival curves instead of KM curves for PFS and ToT.	4.2.6	Methods	Implement parametric survival curves for modelling of PFS and ToT.	+/-	Yes	NA
Low grade AEs may also be relevant, but the company did not consider these in their economic model.	4.2.7	Bias and indirectness	A scenario analysis including Grade 1 and 2 AEs.	+/-	No	A scenario analysis including Grade 1 and 2 AEs
The company assumed different RDIs for regorafenib and T/T, but comparable dose reductions were reported in the literature.	4.2.9	Bias and indirectness	Assuming equal RDIs for regorafenib and T/T.	+	No, the ERG presented a different assumption in their base-case.	NA
The company did not comply with three requests related to their base-case analysis and scenario analyses.	5.2	Unavailability	Provide the requested analyses.	+/-	No	Provide the requested analyses
The results of the PSA are slightly different when running the same analysis multiple times, likely due to the lack of a fixed random seed in the model PSA.	5.3	Imprecision	Implement a fixed random seed to the model PSA	+/-	No	Implement a fixed random seed to the model PSA

AE = adverse event; CE – cost effectiveness; ERG = Evidence Review Group; KM = Kaplan-Meier; NA = not applicable; PFS = progression-free survival; PSA = probabilistic sensitivity analysis; RDI = relative dose intensity; ToT = time on treatment; T/T = trifluridine/tipiracil

[a] Likely conservative assumptions (of the intervention versus all comparators) are indicated by '-'; while '+/-' indicates that the bias introduced by the issue is unclear to the ERG and '+' indicates that the ERG believes this issue likely induces bias in favour of the intervention versus at least one comparator

[b] Explored

6.2 Impact on the ICER of additional clinical and economic analyses undertaken by the ERG

In Section 6.1 the ERG base-case was presented, which was based on various changes compared to the company base-case. Table 6.2 shows how individual changes impact the results plus the combined effect of all changes simultaneously. The exploratory scenario analyses are presented in Table 6.3. These are all conditional on the ERG base-case. The submitted model file contains technical details on the analyses performed by the ERG (e.g., the "ERG" sheet provides an overview of the cells that were altered for each adjustment).

Table 6.2: Deterministic/probabilistic ERG base-case

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER[1] (£/QALY)	iNMB[2] (£30,000)	iNMB[3] (£51,000)
CS base-case (Deterministic)							
Regorafenib	[redacted]	[redacted]					
T/T	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Matter of judgement (1-Log-normal for OS BSC instead of log-logistic)							
Regorafenib	[redacted]	[redacted]					
T/T	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Matter of judgement (2-Implementation of parametric survival curves for PFS)							
Regorafenib	[redacted]	[redacted]					
T/T	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Matter of judgement (3-Implementation of parametric survival curves for ToT)							
Regorafenib	[redacted]	[redacted]					
T/T	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Matter of judgement (4-Equal RDIs for regorafenib and T/T)							
Regorafenib	[redacted]	[redacted]					
T/T	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Deterministic ERG base-case							
Regorafenib	[redacted]	[redacted]					
T/T	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Probabilistic ERG base-case							
Regorafenib	[redacted]	[redacted]					
T/T	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

BSC = best supportive care; CS = company submission; ERG = Evidence Review Group; ICER = incremental cost-effectiveness ratio; iNMB = incremental net monetary benefit; OS = overall survival; PFS = progression-free survival;

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER[1] (£/QALY)	iNMB[2] (£30,000)	iNMB[3] (£51,000)
QALY = quality-adjusted life year; RDI = relative dose intensity; ToT = time on treatment; T/T = trifluridine/tipiracil; QALY = quality adjusted life year; WTP = willingness-to-pay [1]ICER versus regorafenib [2]iNMB for WTP of £30,000 per QALY [3]iNMB for WTP of £51,000 per QALY							

Table 6.3: Probabilistic scenario analyses (conditional on ERG base-case)

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER[1] (£/QALY)	iNMB[2] (£30,000)	iNMB[3] (£51,000)
ERG base-case							
Regorafenib	[redacted]	[redacted]					
T/T	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Scenario analysis (5-OS HR versus T/T from Nakashima 2020)							
Regorafenib	[redacted]	[redacted]					
T/T	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BSC = best supportive care; ERG = Evidence Review Group; HR = hazard ratio; ICER = incremental cost-effectiveness ratio; iNMB = incremental net monetary benefit; OS = overall survival; T/T = trifluridine/tipiracil; QALY = quality adjusted life year; WTP = willingness-to-pay [1]ICER versus regorafenib [2]iNMB for WTP of £30,000 per QALY [3]iNMB for WTP of £51,000 per QALY							

6.3 ERG's preferred assumptions

The CS base-case probabilistic iNMBs of regorafenib versus T/T were [redacted] (WTP £30,000 per QALY gained) and [redacted] (WTP £51,000 per QALY gained), respectively. For regorafenib versus BSC, these were [redacted] (WTP £30,000 per QALY) and [redacted], respectively. The estimated ERG base-case iNMBs for regorafenib versus T/T (probabilistic), based on the ERG preferred assumptions highlighted in Section 6.1, were [redacted] (WTP £30,000 per QALY) and [redacted] (WTP £51,000 per QALY). For regorafenib versus BSC, these were [redacted] (WTP £30,000 per QALY) and [redacted] (WTP £51,000 per QALY). The most influential adjustment was the assumption of equal RDIs for regorafenib and T/T. The scenario analysis using an alternative HR for OS in the T/T arm had a large impact on the iNMB of regorafenib versus T/T.

6.4 Conclusions of the cost effectiveness (CE) section

The company's CE model complied with the NICE reference case. The most prominent issues highlighted by the ERG were 1) the use of KM curves to inform the survival analyses of PFS and ToT for regorafenib and BSC, 2) differences in the OS HR for the modelling of T/T between the ITC conducted by the company and a direct comparison identified in the literature, 3) not including low grade AEs in the economic model despite their potential relevance to this submission, and 4) the assumption of different RDIs for regorafenib and T/T.

Firstly, PFS and ToT for regorafenib and BSC were modelled by directly using the KM data and applying a parametric survival model for the remainder of the model if KM data were no longer available. The company argued that KM data should be used as it is mature, complete and it reflects

clinical practice. However, this is not in line with NICE DSU TSD 14, which states that "parametric models are likely to represent the preferred method for incorporating survival data into health economic models in the majority of cases". The ERG decided that the 'stepped' nature of KM curves, resulting from protocol-driven follow-up, may introduce overfitting to the trial data and could affect the representativeness of the survival analyses results to UK clinical practice. Although the company's arguments for directly using KM data may be valid, in line with NICE DSU TSD 14, the ERG implemented full parametric models to inform the survival analyses of PFS and ToT in its base-case.

Secondly, based on an ITC conducted by the company, a HR of 0.99 for regorafenib versus T/T was applied to model OS in the T/T arm. However, the ERG identified a large observational cohort study[1] in which regorafenib (n=1,501) was directly compared to T/T (n=3,777) in patients with mCRC based on inpatient and outpatient medical care from 269 hospitals in different regions throughout Japan. This study reported OS results that contradict the OS HR of regorafenib versus T/T estimated from the company's ITC. Although the ERG acknowledges the limitations of observational studies, the comparability between the RCTs in the NMA was also questionable. The ERG therefore explored a scenario analysis in which the HR of regorafenib versus T/T from the observational study was applied for the modelling of OS in the T/T arm, which had a substantial impact on the CE results.

Thirdly, the economic model exclusively included Grade 3+ AEs that occurred in at least 2% of the population. Despite the company's positioning of regorafenib as being *"a chemotherapy-free alternative therapy with a different adverse event profile*", low grade AEs were not included in the economic model. The company argued that AEs are not a driver of CE, and that Grade 1 and 2 AEs are not expected to impact costs or QoL. Time constraints and Grade 1 and 2 AEs not generally being modelled in oncology were also arguments to justify excluding the requested scenario analysis. Given the company's positioning of regorafenib as stated above, the ERG would prefer more evidence (e.g., an updated economic model and scenario analysis including Grade 1 and 2 AEs) to support the claim that Grade 1 and 2 AEs would not be impactful.

Finally, different approaches were used to model missed doses and dose reductions for regorafenib and T/T, resulting in different RDIs. The company stated that for regorafenib, missed doses and dose reductions were modelled as a single RDI measure, whereas for T/T no single RDI measure has been reported, and dose reductions and cycle delays were therefore modelled separately instead based on TA405[5]. The company further argued that patients on regorafenib tend to have more dose reductions, while patients on T/T have more dose delays. However, the ERG identified a large observational study cohort study directly comparing regorafenib to T/T, which reported comparable dose reduction. Although the ERG acknowledges the limitations of observational studies, the real-world data from this observational study contradict the company's statement that regorafenib tends to have more reductions, while T/T have more delays. The ERG therefore assumed the T/T RDI to be equal to the RDI of regorafenib in its base-case.

The CS base-case probabilistic iNMBs of regorafenib versus T/T were [redacted] (WTP £30,000 per QALY gained) and [redacted] (WTP £51,000 per QALY gained), respectively. For regorafenib versus BSC, these were [redacted] (WTP £30,000 per QALY) and [redacted], respectively. The estimated ERG base-case iNMBs for regorafenib versus T/T (probabilistic), based on the ERG preferred assumptions highlighted in Section 6.1, were [redacted] (WTP £30,000 per QALY) and [redacted] (WTP £51,000 per QALY). For regorafenib versus BSC, these were [redacted] (WTP £30,000 per QALY) and [redacted] (WTP £51,000 per QALY). The most influential adjustment was the assumption of equal RDIs for regorafenib and T/T. The scenario analysis using an alternative HR for OS in the T/T arm had a large impact on the iNMB of regorafenib versus T/T.

In conclusion, there remains uncertainty about the effectiveness and CE of regorafenib, which can be at least partly resolved by the company by conducting further analyses (e.g., incorporation of low-grade AEs into the economic model). Moreover, the contradicting estimations of the relative effectiveness of regorafenib versus T/T in terms of OS (based on the ITC conducted by the company and the direct comparison from the literature) adds substantial uncertainty to the effectiveness and CE of regorafenib. Therefore, the ERG believes that the CS nor the ERG report contains an unbiased CE estimation of regorafenib compared with relevant comparators.

7. REFERENCES

[1] Nakashima M, Takeuchi M, Kawakami K. Effectiveness and safety of regorafenib vs. trifluridine/tipiracil in unresectable colorectal cancer: a retrospective cohort study. *Clin Colorectal Cancer* 2020;19(4):e208-25.

[2] Welton NJ, Phillippo DM, Owen R, Jones HJ, Dias S, Bujkiewicz S, et al. *NICE DSU Report: CHTE2020 Sources and synthesis of evidence; update to evidence synthesis methods [Internet]*. Sheffield: Decision Support Unit, ScHARR, 2020 [accessed 12.7.22] Available from: https://www.sheffield.ac.uk/sites/default/files/2022-02/CHTE-2020_final_20April2020_final.pdf

[3] Bayer Plc. *Regorafenib for treating metastatic colorectal cancer (ID4002): Document B. Submission to National Institute of Health and Care Excellence. Single technology appraisal (STA)*, 2022 [accessed 26.5.22]. 178p.

[4] National Institute for Health and Care Excellence. *Regorafenib for treating metastatic colorectal cancer: final scope [Internet]*. London: NICE, 2022 [accessed 1.6.22] Available from: https://www.nice.org.uk/guidance/gid-ta10924/documents/final-scope

[5] National Institute for Health and Care Excellence. *Trifluridine–tipiracil for previously treated metastatic colorectal cancer: NICE Technology appraisal guidance 405 [Internet]*. London: NICE, 2016 [accessed 26.5.22] Available from: www.nice.org.uk/guidance/ta405

[6] Bayer Plc. *Regorafenib for treating metastatic colorectal cancer [ID4002]: Response to request for clarification from the ERG (v2:complete)*, 2022 [accessed 7.7.22]. 192p.

[7] Bayer Plc. *Regorafenib for treating metastatic colorectal cancer (ID4002): Appendices C-P. Submission to National Institute of Health and Care Excellence. Single technology appraisal (STA)*, 2022 [accessed 26.5.22]. 176p.

[8] Latimer NR, Abrams KR. *NICE DSU Technical Support Document 16: Adjusting survival time estimates in the presence of treatment switching*. Sheffield: Decision Support Unit, ScHARR, 2014. 57p. Available from: http://nicedsu.org.uk/wp-content/uploads/2016/03/TSD16_Treatment_Switching.pdf

[9] McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS Peer Review of Electronic Search Strategies: 2015 Guideline statement. *J Clin Epidemiol* 2016;75:40-6.

[10] Sampson M, McGowan J, Cogo E, Grimshaw J, Moher D, Lefebvre C. An evidence-based practice guideline for the peer review of electronic search strategies. *J Clin Epidemiol* 2009;62(9):944-52.

[11] Higgins JPT, Green S, eds. *Cochrane handbook for systematic reviews of interventions [Internet]*. Version 5.1.0 [updated March 2011]: The Cochrane Collaboration, 2011 [accessed 23.3.11]. Available from: http://handbook.cochrane.org/

[12] National Library of Medicine (NLM). PubMed user guide: status subsets [Internet]. Bethesda, MD: NLM, 2022 [accessed 13.7.22]. Available from: https://pubmed.ncbi.nlm.nih.gov/help/#publication-types

[13] Banno M, Tsujimoto Y, Kataoka Y. Using the Cochrane Central Register of Controlled Trials to identify clinical trial registration is insufficient: a cross-sectional study. *BMC Med Res Methodol* 2020;20:200.

[14] Grothey A, Van Cutsem E, Sobrero A, Siena S, Falcone A, Ychou M, et al. Regorafenib monotherapy for previously treated metastatic colorectal cancer (CORRECT): an international, multicentre, randomised, placebo-controlled, phase 3 trial. *Lancet* 2013;381(9863):303-12.

[15] Li J, Qin S, Xu R, Yau TC, 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. *Lancet Oncol* 2015;16(6):619-29.

[16] Mayer RJ, Van Cutsem E, Falcone A, Yoshino T, Garcia-Carbonero R, Mizunuma N, et al. Randomized trial of TAS-102 for refractory metastatic colorectal cancer. *N Engl J Med* 2015;372(20):1909-19.

[17] Xu J, Kim TW, Shen L, Sriuranpong V, Pan H, Xu R, et al. Results of a randomized, double-blind, placebo-controlled, phase III trial of trifluridine/tipiracil (TAS-102) monotherapy in Asian patients with previously treated metastatic colorectal cancer: the TERRA study. *J Clin Oncol* 2018;36(4):350-8.

[18] Yoshino T, Mizunuma N, Yamazaki K, Nishina T, Komatsu Y, Baba H, et al. TAS-102 monotherapy for pretreated metastatic colorectal cancer: a double-blind, randomised, placebo-controlled phase 2 trial. *Lancet Oncol* 2012;13(10):993-1001.

[19] Adenis A, de la Fouchardiere C, Paule B, Burtin P, Tougeron D, Wallet J, et al. Survival, safety, and prognostic factors for outcome with Regorafenib in patients with metastatic colorectal cancer refractory to standard therapies: results from a multicenter study (REBECCA) nested within a compassionate use program. *BMC Cancer* 2016;16:412.

[20] Ducreux M, Petersen LN, Öhler L, Bergamo F, Metges JP, de Groot JW, et al. Safety and effectiveness of regorafenib in patients with metastatic colorectal cancer in routine clinical practice in the prospective, observational CORRELATE study. *Eur J Cancer* 2019;123:146-54.

[21] Schulz R, Janssen J, Strauss UP, Langen M, Dworschak KU, Fiala-Buskies S, et al. Clinical efficacy and safety of regorafenib (REG) in the treatment of metastatic colorectal cancer (mCRC) in daily practice in Germany: final results of the prospective multicentre non-interventional RECORA study. *J Clin Oncol* 2018;36(4 Suppl):748.

[22] Novakova-Jiresova A, Kopeckova K, Boublikova L, Chloupkova R, Melichar B, Petruzelka L, et al. Regorafenib for metastatic colorectal cancer: an analysis of a registry-based cohort of 555 patients. *Cancer Manag Res* 2020;12:5365-72.

[23] Tanaka A, Sadahiro S, Suzuki T, Okada K, Saito G, Miyakita H. Retrospective study of regorafenib and trifluridinetipiracil efficacy as a thirdline or later chemotherapy regimen for refractory metastatic colorectal cancer. *Oncol Lett* 2018;16(5):6589-97.

[24] Banzi M, Romagnani A, Berselli A, Tamberi S, Frassoldati A, Luppi G, et al. Retrospective study on real-world use of regorafenib in colorectal cancer in the emiliaromagna region CoreGI01 preliminary results. *Tumori* 2019;105(6):20.

[25] Sueda T, Sakai D, Kudo T, Sugiura T, Takahashi H, Haraguchi N, et al. Efficacy and safety of regorafenib or TAS102 in patients with metastatic colorectal cancer refractory to standard therapies. *Anticancer Res.* 2016;36(8):4299-306.

[26] Huemer F, Piringer G, Schlintl V, Hackl H, Rinnerthaler G, Thaler J, et al. Hospitalizations and clinical outcome in metastatic colorectal cancer during regorafenib or TAS102 therapy. *Cancers (Basel)* 2020;12(10):44207.

[27] Hirano G, Makiyama A, Makiyama C, Esaki T, Oda H, Uchino K, et al. Reduced dose of salvage-line regorafenib monotherapy for metastatic colorectal cancer in Japan. *Anticancer Res.* 2015;35(1):371-8.

[28] Zengin G, Özyurt N, Beypinar I, Cinkir HY, Bilgin B, Gumusay O, et al. Efficiency and safety of regorafenib in metastatic colorectal cancer mCRC Real life experience from Turkey. *J Clin Oncol* 2018;36(15).

[29] Bekaii-Saab TS, Ou FS, Ahn DH, Boland PM, Ciombor KK, Heying EN, et al. Regorafenib dose-optimisation in patients with refractory metastatic colorectal cancer (ReDOS): a randomised, multicentre, open-label, phase 2 study. *Lancet Oncol* 2019;20(8):1070-82.

[30] Argiles G, Margalef NM, Valladares-Ayerbes M, Prado JVD, Grávalos C, Alfonso PG, et al. Results of REARRANGE trial: A randomized phase 2 study comparing different dosing approaches for regorafenib (REG) during the first cycle of treatment in patients (pts) with metastatic colorectal cancer (mCRC). *Ann. Oncol.* 2019;30:aa135.

[31] National Institute for Health and Care Excellence. *User guide for company evidence submission template*, 2015 (Updated Apr 2017) [accessed 3.3.21] Available from: https://www.nice.org.uk/article/pmg24/chapter/4-Clinical-effectiveness#quality-assessment-of-the-relevant-randomised-controlled-trials

[32] National Institute for Health and Care Excellence. *Regorafenib for treating metastatic colorectal cancer [ID4002]: Clarification letter*. London: NICE, 2022. 35p.

[33] Bayer Plc. *Regorafenib [ID4002]: A37 fixed effects consistency model [text file provided by company]*, 2022 [accessed 4.7.22]

[34] National Institute for Health and Care Excellence. *Encorafenib plus cetuximab for previously treated BRAF V600E mutation-positive metastatic colorectal cancer. NICE technology appraisal guidance 668 [Internet]* London: NICE, 2021 [accessed 26.5.22] Available from: www.nice.org.uk/guidance/ta668

[35] National Institute for Health and Care Excellence. *Aflibercept in combination with irinotecan and fluorouracil-based therapy for treating metastatic colorectal cancer that has progressed following prior oxaliplatin-based chemotherapy. NICE technology appraisal guidance 307 [Internet]*. London: NICE, 2014 [accessed 26.5.22] Available from: www.nice.org.uk/guidance/ta307

[36] National Institute for Health and Care Excellence. *Cetuximab, bevacizumab and panitumumab for the treatment of metastatic colorectal cancer after first-line chemotherapy: Cetuximab (monotherapy or combination chemotherapy), bevacizumab (in combination with non-oxaliplatin chemotherapy) and panitumumab (monotherapy) for the treatment of metastatic colorectal cancer after first-line chemotherapy. NICE technology appraisal guidance 242 [Internet]*. London: NICE, 2012 [accessed 26.5.22] Available from: https://www.nice.org.uk/Guidance/TA242

[37] Latimer N. *NICE DSU Technical Support Document 14: Undertaking survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data.* Sheffield: Decision Support Unit, ScHARR, 2011. 52p. Available from:

http://nicedsu.org.uk/wp-content/uploads/2016/03/NICE-DSU-TSD-Survival-analysis.updated-March-2013.v2.pdf

[38] Hoyle M, Peters J, Crathorne L, Jones-Hughes T, Cooper C, Napier M, et al. Cost-effectiveness of cetuximab, cetuximab plus irinotecan, and panitumumab for third and further lines of treatment for KRAS wild-type patients with metastatic colorectal cancer. *Value Health* 2013;16(2):288-96.

[39] Koukakis R, Gatta F, Hechmati G, Siena SCP. Skin toxicity and quality of life during treatment with panitumumab for RAS wild-type metastatic colorectal carcinoma: results from three randomised clinical trials. *Qual Life Res* 2016;25(10):2645-56.

[40] Sabater J, Ralph L, Batteson R, McCarthy G, Wex J, Vera SRM, et al. Validation of cost-effectiveness of trifluridine/tipiracil versus best supportive care and regorafenib for previously treated metastatic colorectal cancer in the UK using phase IIIb PRECONNECT early access clinical trial data in the real world setting. *J Clin Oncol* 2019;37(4 Suppl):639.

[41] Sacco JJ, Botten J, Macbeth F, Bagust A, Clark P. The average body surface area of adult cancer patients in the UK: a multicentre retrospective study. *PLoS One* 2010;5(1):e8933.

[42] Bullement A, Underhill S, Fougeray R, Hatswell AJ. Cost-effectiveness of trifluridine/tipiracil for previously treated metastatic colorectal cancer in England and Wales. *Clin Colorectal Cancer* 2018;17(1):e143-51.

[43] Round J, Jones L, Morris S. Estimating the cost of caring for people with cancer at the end of life: a modelling study. *Palliat Med* 2015;29(10):899-907.

[44] Grimm SE, Pouwels X, Ramaekers BLT, Wijnen B, Knies S, Grutters J, et al. Development and Validation of the TRansparent Uncertainty ASsessmenT (TRUST) Tool for assessing uncertainties in health economic decision models. *Pharmacoeconomics* 2020;38(2):205-16.

[45] Kaltenthaler E, Carroll C, Hill-McManus D, Scope A, Holmes M, Rice S, et al. The use of exploratory analyses within the National Institute for Health and Care Excellence single technology appraisal process: an evaluation and qualitative analysis. *Health Technol Assess* 2016;20(26):1-48.

8. APPENDIX

The CS reported that the SLRs looking into health-related quality of life studies (Appendix H), and cost and healthcare resource identification (Appendix I), were performed using the same methods, and that overall methods were only reported in appendix G. Therefore, the following sections will only review those searches unique to HRQoL and Resource Use, for additional searches of HTA organisations and conference proceedings please the table above.

Table 8.1: Data sources for HRQoL (as reported in CS)

Resource	Host/Source	Date Ranges	Dates searched
Electronic databases			
MEDLINE	Embase.com	From inception	5.3.21 Updated 22.2.22
Embase	Embase.com	From inception	5.3.21 Updated 22.2.22
MEDLINE In-Process	Pubmed.com	From inception	19.3.21 Updated 22.2.22
EconLit	Ebsco.com	From inception	19.3.21 Updated 22.2.22
HTAD	CRD	From inception-2018/03/31	7.4.21
NHS EED	CRD	From inception-2015/03/31	7.4.21
Additional searches			
Hand-searching	Bibliographies of key systematic review and meta-analysis articles were screened to fully evaluate the relevant economic studies		
HTAD = Health Technology Assessment Database; NHS EED = National Health Service Economic Evaluation Database; NICE = National Institute for Health and Care Excellence;			

ERG comment:

- The CS and response to clarification provided sufficient details for the ERG to appraise the literature searches.
- Searches covered a good range of resources, including databases, HTA agency websites and conference proceedings (see Table 8.1). Strategies contained a good mix of free text and subject headings.
- The MEDLINE search was conducted using Embase.com as described in the clinical effectiveness section, therefore the same limitations will apply (See section 3.1.1).

- The MEDLINE In-Process search via PubMed contained the same search limit error as previously reported in section 3.1.1. However, this omission appears to have been corrected in the update searches.

Table 8.2: Data sources for Cost and healthcare resource identification, measurement and valuation (as reported in CS)

Resource	Host/Source	Date Ranges	Dates searched
Electronic databases			
MEDLINE	Embase.com	2010-2022/02/22	5.3.21 Updated 22.2.22
Embase	Embase.com	2010-2022/02/22	5.3.21 Updated 22.2.22
MEDLINE In-Process	Pubmed.com	2010-2022/02/22	19.3.21 Updated 22.2.22
EconLit	Ebsco.com	2010-2022/02/22	19.3.21 Updated 22.2.22
HTAD	CRD	2010-2018/03/31	7.4.21
NHS EED	CRD	2010-2015/03/31	7.4.21
Additional searches			
Hand-searching	Bibliographies of key systematic review and meta-analysis articles were screened to fully evaluate the relevant economic studies		
HTAD = Health Technology Assessment Database; NHS EED = National Health Service Economic Evaluation Database; NICE = National Institute for Health and Care Excellence;			

ERG comment:

- The CS and response to clarification provided sufficient details for the ERG to appraise the literature searches.
- Searches covered a good range of resources, including databases, HTA agency websites and conference proceedings (See Table 8.2). Strategies contained a good mix of free text and subject headings.
- The MEDLINE search was conducted using Embase.com as described in the clinical effectiveness section, therefore the same limitations will apply (See section 3.1.1).
- The MEDLINE In-Process search via PubMed contained the same search limit error as previously reported in section 3.1.1. However, this omission appears to have been corrected in the update searches.

Single Technology Appraisal

Regorafenib for previously treated metastatic colorectal cancer [ID4002]

EAG report – factual accuracy check and confidential information check

"Data owners may be asked to check that confidential information is correctly marked in documents created by others in the evaluation before release." (Section 5.4.9, NICE health technology evaluations: the manual).

You are asked to check the EAG report to ensure there are no factual inaccuracies or errors in the marking of confidential information contained within it. The document should act as a method of detailing any inaccuracies found and how they should be corrected.

If you do identify any factual inaccuracies or errors in the marking of confidential information, you must inform NICE by the end of **8 August 2022** using the below comments table.

All factual errors will be highlighted in a report and presented to the Appraisal Committee and will subsequently be published on the NICE website with the committee papers.

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

Issue 1 The quality of the RCTs is questioned

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
The report casts doubt on the quality of the RCTs. However, evidence supporting that the RCTs are of high quality and at low risk of bias has been presented in the CS and further supported at the clarification stage where the ERGs concern regarding performance bias was fully explained and accepted. The ERG concluded that they were "*satisfied by this response* [regarding potential performance bias], *which fully explained the discrepancy and consider the RCTs to be of high quality*". In light of the ERGs conclusion and the evidence supporting the RCTs are high quality and have a low risk of bias, any statements which conclude there is doubt over the quality of the RCTs are not accurate	<u>Proposed amendment 1</u> Removal of the word 'ostensibly' from table 1.5 page 16 Removal of references to the quality of the studies being in doubt <u>Proposed amendment 2</u> Removal of the word 'partially' from page 68	The RCTs are of high quality and low risk of bias. **Regorafenib RCTs - Evidence for the 5 RCTs being of high quality and low risk of bias** On page 68 the ERG present the quality assessment of the regorafenib trials which shows the trials to be methodologically robust with a comprehensive approach to patient allocation, control of confounding factors, and an overall low risk of bias". The ERGs concern regarding potential performance bias was addressed at the clarification stage and is discussed on page 68 - the ERG commented that they "*were satisfied by this response, which fully explained the discrepancy and consider the RCTs to be of high quality*"	Not a factual inaccuracy. Our comment in table 1.5 on page 16 of the report refers to the quality of all the RCTs included in the NMA and not only that of CORRECT and CONCUR. More specifically, the comment refers partly to our independent risk of bias assessment results for the three additional RCTs included in the NMA (RECOURSE, TERRA and Yoshino 2012) which is reported in section 3.3 page 112 (see second ERG comment). It is also important to note that quality might not be simply a function of a risk of bias assessment, but also the appropriateness of the evidence for pooling and decision making: each of the RCTs might provide estimates of treatment

		T/T - Evidence for the 5 RCTs being of high quality and low risk of bias: Table 3.36 of the ERG report also shows the T/T trials to be of high quality with low risk of bias.	effect that are at lower risk of bias, but none of those estimates are of the treatment effect of interest i.e. regorafenib vs. T/T. Pooling RCTs, each at low risk of bias, does not imply that the pooled estimate is also at low risk of bias: that risk depends also on the comparability of all of those RCTs, which was discussed in detail in the ERG report (See Sections 3.3 and 3.4).

Issue 2 **It is concluded that the differences between the regorafenib and T/T trials calls into question the ITC result of comparable efficacy**

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
Differences between the regorafenib trials were noted in several patient characteristics which were suggested might affect comparability or external validity. The potential effect of these differences on the comparison between regorafenib and T/T has been comprehensively explored and the conclusion of	In table 1.5 the statement "However, there is doubt as to the quality of these RCTs and their comparability" could be changed to "However, there are differences between the trials which may affect comparability. Exploration of these differences did not support the differences having any	In the CS and clarification questions, differences raised by the ERG were explored by means of subgroup analysis, mixed adjusted indirect comparisons and inclusion/exclusion of different RCTs from the analyses: - race/ethnicity was explored in	Not a factual inaccuracy. Regarding the reference to 'quality' please see previous comment in Issue 1. Regarding the comparability issues, the ERG has highlighted a number of sources of potential problems which are detailed

comparable efficacy was found to be robust. In light of this it does not seem accurate for 'comparability' of the trials to be a key issue.	meaningful impact on the comparability of the studies"	clarification question (CQ) A18, A19, A29 - prior treatment with anti-VEGF was explored in CQ A22, A23 - adjustment for ECOG performance status, previous number of treatment lines, KRAS status, time from diagnosis was explored in CQ A31 - adjustment by prior treatment, age, gender was explored in the company submission Overall, the results of these analyses support that the differences between the trials did not affect the conclusion of comparable efficacy and that the trials were suitable for indirect comparison.	in our comments in Sections 3.3 and 3.4. The company's views on these issues, as reported in both the CS documents and the response to clarification questions, have been taken into consideration and are critiqued in detail in our report.

Issue 3 The observational study by Nakashima is discussed as providing a credible alternative estimate of relative efficacy

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
The 5 RCTs provide a high quality and low bias evidence base from which to compare regorafenib and T/T. Using these trials in an ITC supports the conclusion of comparable efficacy. Despite a comprehensive set of sensitivity and scenario analyses demonstrating comparability, and 3 RWE studies similarly suggesting comparability, Nakashima, which is described as being at high risk of bias, is put forward as calling into question the conclusions of ITC and other RWE. It is not correct to consider Nakashima as providing a credible alternate estimate of relative efficacy. This is a single study which was described by the ERG as being at "high risk of bias".	Throughout the report, the study by Nakashima should be described in the context of being subject to high-risk of bias. Nakashima should not be suggested as providing a credible alternative to the extremely comprehensive set of analyses from the CS and clarification questions which demonstrate comparable efficacy. The relative HR from Nakashima, being significantly different to that of the ITC, should not be described as calling into question the results from the ITC. We suggest the removal of the suggestion to incorporate Nakashima into the ITC alongside the 5 RCTs	**Nakashima is a retrospective observational study and does not provide a credible estimate of relative efficacy** 1) the estimate of relative efficacy between regorafenib and T/T in this study is a OS hazard ratio of 0.66. The size of the benefit over regorafenib is of the same order as the benefit of T/T over placebo from its RCTs (Mayer 2015, Xu 2017, Yoshino 2012). This is not credible. 2) If the registration studies for T/T are to be believed (and they should be) the HR of 0.66 vs regorafenib from Nakashima implies either that: a) regorafenib's efficacy is no different to placebo which is in direct contrast to the efficacy observed of regorafenib versus placebo in its registration studies, or	Not a factual inaccuracy. The Nakashima 2020 study was considered to be the largest, single-arm observational study and of best quality e.g., Table 1.5., page 16 or Table 3.15., page 71). The risk of bias assessment results for the study are described in detail in Table 3.15 and in our commentary on page 71. The HR of 0.66 vs regorafenib from Nakashima implies a large treatment effect in favour of T/T (and not regorafenib) for OS. The ERG acknowledged the risk of selection bias. However, the ERG also noted that there appeared to be greater balance in baseline characteristics than in the other observational studies (Section 3.2.3.2, page 67).

		b) T/T is twice as effective as observed in its registration studies i.e. if regorafenib is effective against placebo (as per CORRECT and CONCUR) and T/T is as effective against regorafenib as it was against placebo this result indicates an efficacy twice that observed in its registration studies 3) Patients who formed the evidence base for the Nakashima study were <u>selected</u> for regorafenib or T/T and were not randomised to either treatment. Consequently, whereas unknown confounders have a potential to be equally distributed between groups in RCTs this is not the case in observational studies. 4) The ERG report considers the Nakashima study to be at high-risk of bias	Therefore, the conclusions regarding Nakashima 2020 plus the questions of comparability of the RCTs causes considerable uncertainty in the effectiveness of regorafenib versus T/T (e.g., Section 3.4.4, page 122). The use and the potential benefit as well as the limitations in using non-randomized (observational) studies in evidence synthesis is described in the Technical Support Document (TSD) TSD 20.[1]

		All that the Nakashima study shows is that patients who were selected for T/T by clinicians had a longer OS that patients selected for regorafenib. There is no randomisation for covariates (known or unknown) and the study does not provide a credible estimate of relative efficacy and should not be considered as providing such. The inclusion of Nakashima into the ITC would increase rather than decrease uncertainty. The ITC will not be improved by including highly biased evidence from Nakashima into a low bias network.	

Issue 4 Nakashima is suggested as being better than the 3 other RWE studies on the basis of more closely matching baseline characteristics

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
The baseline characteristics between the treatment groups from the 4 RWE studies is compared and discussed. The	The report could more strongly conclude that none of the RWE studies, either alone	On page 53 the ERG state that "*because the allocation to groups was non-random it is likely that there exist differences*	Not a factual inaccuracy. The Nakashima 2020 study was considered to have

ERG conclude that Nakashima appears to have "*far better comparability between arms*" (page 53) A conclusion of Nakashima providing a more reliable indication of comparative efficacy appears to be reached. None of the RWE provides a credible estimate of relative efficacy and all the RWE studies have a high risk of bias. As they are non-randomised they differ not only in baseline characteristics but also in covariates which are not measured and any conclusion that one or more RWE studies provides a credible estimate of relative efficacy is inaccurate.	or combined, provide a credible estimate of relative efficacy. The suggestion to incorporate the RWE data into the RCT network (table 1.5, page 16) should be either removed, or modified to indicate this would increase uncertainty as it would involve including data at high-risk of bias into a low-risk of bias RCT network.	*in non-measured covariates that may have impaired internal validity*" This statement is important as the ERG conclude on page 62 that the efficacy difference between CORRECT and CONCUR may be due to unknown covariates. On the basis of patients being randomised in CORRECT and CONCUR, there is a tendency for unknown covariates to balance across groups. In the retrospective setting of the RWE patients were selected (not randomised) to regorafenib or T/T and these studies are at much greater risk of unknown covariates driving results. The RWE data as a whole does not provide data that provides a reliable indication of comparative efficacy.	balanced baseline characteristics (see also Table 3.12). The Nakashima 2020 study was also the largest observational study (N=5,278) compared with the other three, hence with the largest power to detect any difference between regorafenib and T/T and the greatest generalisability. The company's concerns around the existence of unknown covariates have also been recognised and reported by the ERG (see p. 53). However, it is also clear that there were differences in treatment effect between the RCTs for both regorafenib and T/T vs. placebo, which were difficult to explain by the observed covariates, as discussed in detail in Sections 3.3 and 3.4.

Issue 5 The company did an analysis on TRAEs as opposed to TEAEs as requested by the ERG

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
We apologise as it was a labelling error by Bayer that has resulted in this error I.e. on page 119 the ERG state that "*the recommendation of the ERG was to execute an NMA focusing on TEAEs: nevertheless, the company focused on TRAEs instead*" The tables in the CQ were incorrectly labelled as TRAEs when they should have been labelled as TEAEs.	Remove the reference to the company performing an analysis on TRAEs	The tables in the CQ were incorrectly labelled by Bayer. We apologise for this error.	It is not clear which tables of the CQ were incorrectly labeled in the description of the problem. The ERG assumes that tables A28.2 and A28.3 are the ones to be corrected. A correction has been made in Section 3.4.3, on page 119, and the titles of Tables 3.46 and 3.47 have been altered. If Table A28.1 was also incorrectly labeled further corrections should be made regarding Table 3.39 of the report and several statements on pages 111-113.

Issue 6 Reporting errors

Description of problem	Description of proposed amendment	Justification for amendment	ERGEEG comment
ERG report – page 19, 20 and 156	There was a possible copy-paste error for the probabilistic results.	Based on the similarity of deterministic and probabilistic	Amended.

The probabilistic ERG base-case results in Table 1.11, Table 6.2, and the text do not match the results reported in the ERG model. In addition, the total T/T costs differ substantially between the deterministic and probabilistic analysis, which does not seem plausible		model results for all other analyses there may have been a typographical error.	
ERG report – page 26 The percentage of patients who received systematic anticancer therapy in the placebo group in CONCUR is incorrect	Please could you amend the sentence (the change has been underlined) to read: 'In response to clarification request the company reproduced data from Appendix D, which showed systemic anticancer therapy use to be: 29.8% versus 25.9% in CORRECT <u>42.6</u>% versus 30.9% in CONCUR for placebo versus regorafenib respectively'	The amendment will correct the percentage of patients who received systematic anticancer therapy in the placebo group in CONCUR	'42.9' has now been corrected to '42.6' on page 26 of the report.
Page 37 – "both were phase II randomised…"	Change 'II' to 'III'	CORRECT and CONCUR were phase III studies	Typo corrected to 'III' on page 37.
Table 3.4 page 37	Amend '36' to '136' in respect of patient numbers	Typo	Typo corrected on Table 3.4, page 39.

ERG report – page 36 onwards The table number references in text	The table number references in text are out of sync from page 36 onwards. For example, please could you change the sentence below on page 36 from: 'These are summarised in Table 3.2 above.' To 'These are summarised in Table 3.3 above.'	The amendment will align the table numbering throughout the report	Corrected throughout Section 3, where the issue was localised.

Location of incorrect marking	**Description of incorrect marking**	**Amended marking**	**ERGEEG comment**
ERG report – page 148/149 section 5.1 [redacted]	Could you please: [redacted]	[redacted]	Amended.
Page 98 – ERG comment	The statement "with a different adverse effect profile to T/T" does not need to be marked as CIC		Amended.

References

[1] Welton NJ, Phillippo DM, Owen R, Jones HJ, Dias S, Bujkiewicz S, et al. *NICE DSU Report: CHTE2020 Sources and synthesis of evidence; update to evidence synthesis methods [Internet]*. Sheffield: Decision Support Unit, ScHARR, 2020 [accessed 12.7.22] Available from: https://www.sheffield.ac.uk/sites/default/files/2022-02/CHTE-2020_final_20April2020_final.pdf

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Regorafenib for treating metastatic colorectal cancer [ID4002]

Technical engagement response form

As a stakeholder you have been invited to comment on the External Assessment Report (EAR) for this evaluation.

Your comments and feedback on the key issues below are really valued. The EAR and stakeholders' responses are used by the committee to help it make decisions at the committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

Information on completing this form

We are asking for your views on key issues in the EAR that are likely to be discussed by the committee. The key issues in the EAR reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the EAR.

You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise.

If you would like to comment on issues in the EAR that have not been identified as key issues, you can do so in the 'Additional issues' section.

If you are the company involved in this evaluation, please complete the 'Summary of changes to the company's cost-effectiveness estimates(s)' section if your response includes changes to your cost-effectiveness evidence.

Please do not embed documents (such as PDFs or tables) because this may lead to the information being mislaid or make the response unreadable. Please type information directly into the form.

Do not include medical information about yourself or another person that could identify you or the other person.

We are committed to meeting the requirements of copyright legislation. If you want to include journal articles in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.

Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, and all information submitted under 'depersonalised data' in pink. If confidential information is submitted, please also send a second version of your comments with that information redacted. See the NICE health technology evaluation guidance development manual (sections 5.4.1 to 5.4.10) for more information.

The deadline for comments is the end of day on **21st September 2022**. Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

NICE National Institute for Health and Care Excellence

About you

Table 1 About you

Your name	[redacted]
Organisation name: stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder, please leave blank)	Bayer Plc
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	**Current Situation** • Bayer does not have direct or indirect links with, or funding from, manufacturers, distributors or sellers of smoking products but Bayer provides pesticides for crops, which would therefore include tobacco crops. • Bayer is a member of the Cooperation Centre for Scientific Research Relative to Tobacco (CORESTA) (http://www.coresta.org/) within the scope of recommendations of pesticides used for protection of tobacco plants. • It is also a member of country and EU business federations such as the Confederation of British Industry (CBI) and 'Business Europe', which include tobacco companies. **Past Situation** In 2006, Bayer and its subsidiary Icon Genetics piloted a new process for producing biotech drugs in tobacco plants. Icon Genetics was acquired by Nomad Bioscience GmbH from Bayer in 2012.

Key issues for engagement

All: Please use the table below to respond to the key issues raised in the EAR.

Key issue	Does this response contain new evidence, data or analyses?	Response
Key issue 1: Potential ambiguity of the population in the decision problem, which has implications for the comparators The company considers regorafenib suitable for people whose disease has progressed following first-line chemotherapy/first-line biologic and who are being considered for ≥ 3rd-line treatment. This definition (≥ 3rd-line treatment) suggests the	No	**Cost-effectiveness against BSC** Although we do not believe BSC is a comparator (please see below), a limited set of cost-effectiveness results against BSC were included in our submission in the interests of transparency and to support the conclusions against T/T i.e. T/T showed itself to be cost-effective against BSC and in our analyses we have shown regorafenib to be cost-effective against T/T. To 'close the loop', showing cost-effectiveness of regorafenib against BSC was necessary, if only from a triangulation perspective. The basecase results from the CS are presented below (Table 1.1 and Table 1.2). In addition, the ERG report states the following in respect of BSC: "*In the company and ERG base-case, regorafenib would not be cost effective versus BSC with a willingness-to-pay (WTP) of £30,000 per quality adjusted life year (QALY), but it would be if WTP was £51,000 i.e., with the 1.7 x QALY weight applied*"

appropriate comparator could be either trifluridine/tipiracil or best supportive care. The comparator influences cost-effectiveness estimates.

Table 1.1 (Table 43 from CS: base-case results)

Technology	Total costs (£)	Total LYG	Total QALYs	Incremental costs regorafenib (£)	Incremental LYG regorafenib	Incremental QALYs regorafenib	ICER regorafenib versus Tx (£/QALY)	ICER incremental (£/QALY)
Regorafenib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
T/T	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Key: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years; T/T, trifluridine/tipiracil.

Table 1.2 (Table 44 from CS: Net health benefit)

Technologies	Total costs (£)	Total QALYs	Incremental costs regorafenib (£)	Incremental QALYs regorafenib	NHB regorafenib versus Tx at £20,000	NHB regorafenib versus Tx at £30,000	NHB regorafenib versus Tx at £51,000
Regorafenib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
T/T	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Key: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years; T/T, trifluridine/tipiracil.

		Regorafenib for treating metastatic colorectal cancer The submission was made in response to physician requests for an alternative to T/T. Prior to the recommendation of T/T, patients who failed second-line treatments would receive BSC alone. On its recommendation by NICE, T/T replaced BSC (TA405) and formed a new last active line of treatment in mCRC i.e. BSC was effectively moved one treatment-line later. **≥3rd line treatment (referred to as 'Subsequent or alternative therapy' in NG151)** Patients who are fit enough to receive active treatment In clinical practice, if a patient is 'fit' enough to receive ≥3rd line therapy they will do so. For these patients the only recommended current option is T/T. Regorafenib, if recommended, will be an alternative to T/T in this patient group who are fit enough to receive active treatment, making T/T the comparator. Patients who are NOT fit enough to receive active treatment Patients who are not fit enough to receive active treatment receive BSC as there is no other choice. Logically, as these patients are not fit enough then active treatment is not a clinical option and therefore not a comparator to BSC. **Patients who fail treatment with T/T** With each successive treatment, fewer patients remain fit enough for the next treatment option. After failure of ≥3rd line treatment (currently T/T) the vast majority of patients are no longer fit enough for active treatment i.e. active treatment is no longer an option. As this patient group is no longer fit enough to receive active treatment there are no options in this treatment space and active treatment (whatever that may be) is not a comparator to BSC.

		We acknowledge that there will be a handful of patients who will be fit enough but this is a very small minority. The draft Resource Template commented that there may be " *a very small amount of use as a fourth line therapy for very fit patients after trifluridine/tipiracil*."
Key issue 2: Potential difference between subsequent treatment use in the RCTs and NHS clinical practice with unknown effect of subsequent treatment The proportion of people receiving post-progression treatment after regorafenib in CORRECT and CONCUR trials were 21% and 31%, respectively. Whereas, the company has suggested that in UK clinical practice <10% of people would be fit enough for post-progression active treatment.	No	With each successive treatment, fewer patients remain fit enough for the next treatment option. After failure of ≥3rd line treatment (currently T/T) the vast majority of patients are no longer fit enough for active treatment i.e. active treatment is no longer an option. The advice we have been given by clinical experts is that in the UK <10% of patients would receive active treatment after T/T (or regorafenib if recommended). In this context, 'active treatment' is rechallenge with prior therapies which have already failed but to which initial treatment response (disease control and tolerability) was favourable. However, as resistance to prior therapy is already established in the cancer, any benefit of rechallenge is likely to be very marginal. In the RCTs for regorafenib a higher proportion of patients received post-progression treatment than would be expected in the UK. Clinicians have indicated this is likely due to incentives leading to overtreatment in trial centres in other countries i.e. patients are likely to have received treatment that would not be considered in the UK. In both the CORRECT and CONCUR studies, a slightly greater proportion of patients in the BSC arm received further active therapy after failure of trial treatment (see table 2.1). It is therefore reasonable to assume that any benefit of this treatment would advantage the BSC arm and bias against the regorafenib arm. However, as active treatment is expected to be minimally effective, and the absolute difference between the arms in terms of the proportion receiving active treatment is small, any effect is anticipated to be minor.

<table>
<tr>
<td></td>
<td></td>
<td>
Table 2.1. Post-progression treatment in regorafenib trials
<table>
<tr><th></th><th>Regorafenib</th><th>BSC</th></tr>
<tr><td>CORRECT</td><td>26%</td><td>30%</td></tr>
<tr><td>CONCUR</td><td>31%</td><td>43%</td></tr>
</table>

Slightly higher levels of post-progression treatment were also received in the T/T trials (table 2.2). However, for the same reasons as for regorafenib, we do not expect any impact on the generalisability of the T/T trials to the NHS.

Table 2.2. Post-progression treatment in the T/T trials
<table>
<tr><th></th><th>T/T</th><th>BSC</th></tr>
<tr><td>RECOURSE</td><td>42%</td><td>42%</td></tr>
<tr><td>TERRA</td><td>37.6%</td><td>45.2%</td></tr>
<tr><td>Yoshino</td><td>43%</td><td>46%</td></tr>
</table>
</td>
</tr>
<tr>
<td>Key issue 3: Lack of external validity of and comparability between the regorafenib and trifluridine/tipiracil RCTs
It is uncertain as to the size and direction of subgroup differences in terms of treatment</td>
<td>No</td>
<td>The EAG notes that the treatment arms in the trials were well-balanced and they had little concern regarding internal validity of the regorafenib trials. However, the EAG note 3 main concerns regarding external validity i.e. comparability between trials and with NHS clinical practice:

1. A large proportion of patients in CORRECT and CONCUR received prior anti-VEGF treatment in the form of bevacizumab. Bevacizumab is not recommended in the UK
2. Number of prior treatments received
3. The CONCUR trial was conducted exclusively in Asian patients which does not match UK ethnicity.</td>
</tr>
</table>

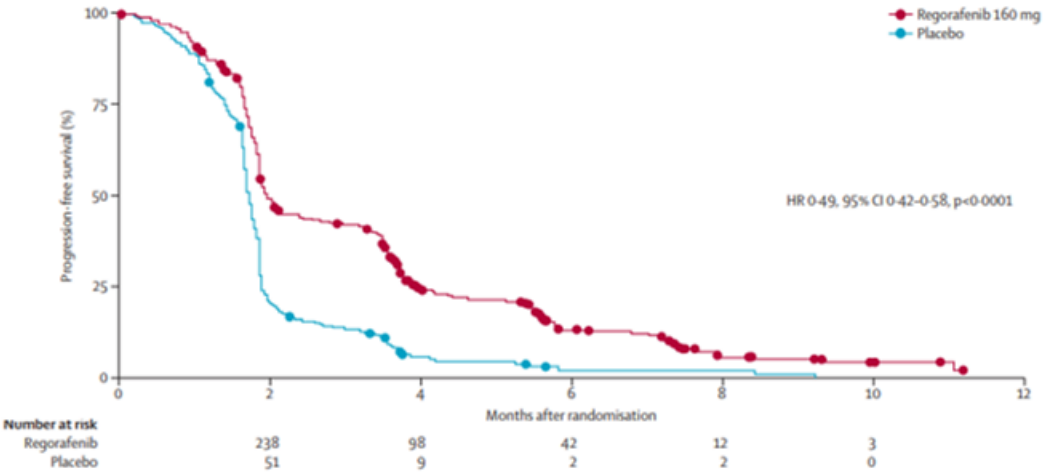
experience and race in the CORRECT and CONCUR trials. There is also a disparity between the trials and NHS clinical practice.		Anti-VEGF In our response to the clarification questions we presented subgroup results which showed a better response in anti-VEGF naïve patients *(HR: 0.470; 95% CI: 0.309, 0.714)* versus anti-VEGF experienced patients *(HR: 0.726; 95% CI: 0.430, 1.224).* This would indicate that more benefit might be realised in the NHS compared to the trials as patients in the UK are anti-VEGF naive. However, as we have acknowledged throughout the submission and clarification questions, the trials were powered at the ITT level and we are not 'claiming' a more favourable effect in the NHS. Number of prior treatments Subgroups were formed based on number of prior therapies. The results in these subgroups were counterintuitive i.e. there was a numerically better response in patients who had received a greater number of prior therapies - the reverse would be expected. The results were not significant (confidence intervals overlapped). We consider that focusing on this result and then questioning external validity is to overinterpret the data - the trials were not stratified on number of prior treatments and powering was at the ITT level. The results of both trials indicate a benefit of regorafenib irrespective of number prior treatments and we see no reason why the results are not generalisable to the NHS. Race There was no indication in any analyses of an interaction between race and efficacy and therefore no reason to expect response to treatment to differ in the NHS. Clinicians have similarly indicated that they expect no difference in efficacy according to race/ethnicity. Summary Based on the above there is: • a potential better effect of treatment in the NHS versus the trials (based on the anti-VEGF subgroup analyses),

		• an unknown effect based on efficacy according to the number of prior treatments; and • a neutral effect based on differences in race. These results do not indicate a risk of poorer treatment response in the NHS versus the clinical trials. We do note however, the EAGs general concern related to powering i.e. the trials were not powered to detect differences between subgroups - therefore lack of statistical difference is not proof of there being no difference. Whilst we acknowledge, and agree for the need to be vigilant to type II error, there is no suggestion within the data, or based on clinically plausible mechanisms, that the results of CORRECT and CONCUR are not generalisable to the NHS. The differences between the regorafenib trials and NHS clinical practice pertain equally to the T/T trials. We consider the regorafenib trials and the T/T trials to be generalisable to the NHS. Extensive analysis indicate the treatments are comparable in efficacy. We acknowledge the uncertainties raised by the EAG, but these are essentially unresolvable without a head-to-head study conducted in the UK.
Key issue 4: Difference in the treatment effect of regorafenib versus trifluridine/tipiracil depending on evidence source The network meta-analysis (NMA) of randomised controlled trials (RCTs) for the comparison of	Yes	A conclusion of comparable efficacy is supported by the NMA of the 5 RCTs (Regorafenib - CORRECT and CONCUR; T/T – RECOURSE, Yoshino, TERRA) and a comprehensive set of sensitivity analyses. As indicated by the EAG “there is probably little to be gained by further SAs with the RCTs”. Based on the RCTs a conclusion of comparable efficacy is supported. However, a retrospective observational study using a nationwide database in Japan has shown a OS HR favouring T/T over regorafenib (Nakashima 2020). This study has been ‘elevated’ by the EAG and is suggested to provide a credible indication of relative efficacy. <u>Nakashima</u>

regorafenib with trifluridine/tipiracil (T/T) indicates that, depending on which trials are included, there is little difference in either overall survival (OS) or progression-free survival (PFS). However, a large comparative observational study by Nakashima 2020 showed an advantage to T/T in OS, with a hazard ratio (HR) of 1.515 in contrast with the indirect treatment comparison (ITC) base-case, HR = 0.99.		Nakashima uses a Japanese nationwide database to retrospectively compare the outcomes of treatment in patients prescribed regorafenib or T/T. The paper presents OS results for patients who received 1) regorafenib alone, 2) T/T alone, 3) regorafenib followed by T/T, and 4) T/T followed by regorafenib. A key flaw in the analyses is the absence of data on ECOG performance status, which if available would have provided information on the baseline health of the patients. As ECOG PS is an important prognostic factor this is a significant limitation (Shitara 2011, Steinberg 1992). Dividing the patients into four groups instead of two creates bias in an already biased (because of the lack of randomization) dataset, as patients who progress or switch are essentially removed from the treatment comparison. Patients who receive subsequent treatment are likely healthier than those who don't. The analyses of regorafenib and T/T therefore compares two groups of 'sicker' patients. The lack of reporting of key variables such as ECOG performance status inhibits an understanding of whether the populations are comparably 'sick' and inhibits meaningful adjustment for confounders. The assignment to the regorafenib or T/T monotherapy group, based on a future event observed in the study (i.e. switching to regorafenib or T/T) introduces bias (Lee 2014, Latimer 2016) that cannot be corrected by balancing patient characteristics at baseline with propensity-scoring methods. In essence, the analysis starts with selection bias due to the non-randomized nature of treatment assignment and then makes it worse. There is a consensus that excluding switchers (or here, putting them in their own group) is prone to selection bias (Latimer 2016) and is an approach that should be avoided. <u>The result from the Nakashima study is so extreme as to not be credible</u> The primary analysis is that of T/T monotherapy vs regorafenib monotherapy, for which they report significant differences in favour of T/T monotherapy, with almost twice the median OS, and an HR of 0.67 (0.66 with a propensity score adjustment).

		1) The size of the benefit of T/T over regorafenib is of the same order as the benefit of T/T over placebo from its RCTs (Mayer 2015, Xu 2017, Yoshino 2012). 2) If the registration studies for T/T are to be believed the HR of 0.66 vs regorafenib from Nakashima implies either that: a) regorafenib's efficacy is no different to placebo which is in direct contrast to the efficacy observed for regorafenib versus placebo in its registration studies, or b) T/T is twice as effective as observed in its registration studies i.e. if regorafenib is effective against placebo (as per CORRECT and CONCUR) and T/T is as effective against regorafenib as it was against placebo this result indicates an efficacy twice that observed in its registration studies <u>Moriwaki 2018</u> As noted above, the study by Nakashima did not adjust for ECOG PS and therefore there is no way to know if patients prescribed regorafenib are comparable to those prescribed T/T - this introduces potential bias. Moriwaki presents results for T/T vs regorafenib, also using retrospective observational data from Japan, however with significant differences: - patients were only included for analysis if they were able/eligible to receive T/T or regorafenib - propensity matching was done on ECOG performance status The results of this analysis are in line with those of the submitted ITC and contrast with those of Nakashima, finding similar OS in the regorafenib (7.9 months) and T/T groups (7.4 months). There was no significant difference between the two groups (unadjusted HR of T/T to regorafenib, 1.03;95%CI, 0.85-1.26; p=.75. In the propensity score adjusted analysis for OS, similar results were observed between the two group (adjusted OS HR 0.96 (95%CI 0.78-1.18; p=.69).

		Summary All that the Nakashima study can be taken as indicating is that patients who were selected for treatment with T/T had a longer OS than patients who were selected for treatment with regorafenib. In the absence of data on important variables such as ECOG status at baseline it is not possible to determine comparability between the patient groups. In a different retrospective study also conducted in Japan, that did adjust for ECOG status, a conclusion of comparable efficacy was reached. The study by Moriwaki is similarly at high-risk of bias but it does strongly show that real-world data cannot be used to compare regorafenib and T/T and that depending on the adjustments made very different conclusions can be reached. The best data to assess the comparability of regorafenib and T/T is from the RCTs, as done in our submission. **References:** Latimer NR et al. Treatment switching: statistical and decision-making challenges and approaches. International Journal of Technology Assessment in Health Care, 32:3 (2016), 160-66 Lee Y et al. Analysis of clinical trials by treatment actually received: is it really an option: Stat Med. 1991;10:1595-1605 Moriwaki T et al. Propensity score analysis of regorafenib versus trifluridine/tipiracil in patients with metastatic colorectal cancer refractory to standard chemotherapy (REGOTAS): a Japanese society for cancer of the colon and rectum multicentre observational study. The Oncologist 2018;23:7-15 Shitara K et al. Prognostic factors for metastatic colorectal cancer patients undergoing irinotecan-based second-line chemotherapy. Gastrointest Cancer Res 2011 Sep-Dec,4(5-6): 168-172 Steinberg J et al. Prognostic factors in patients with metastatic colorectal cancer receiving 5-fluorouracil and folinic acid. European Journal of Cancer 28(11), 1992, 1817–1820)

Key issue 5: Treatment effectiveness and extrapolation The company implemented Kaplan-Meier (KM) curves instead of parametric survival models for the survival analyses of progression-free survival (PFS) and time on treatment (ToT) for regorafenib and best supportive care (BSC), leading to potential overfitting of modelled survival outcomes.	No	In CORRECT and CONCUR, patients were assessed every 8 weeks (+/- 1 week) to determine if their cancer had progressed. The clinical assessment of progression, at defined intervals, is why the KM curve is 'stepped' in nature, rather than continuous – see KM curve below. **Figure 5.1 (figure 6 from CS)** **Figure 6: CORRECT – Kaplan–Meier curve of progression-free survival (ITT)**  **Key:** CI, confidence interval; HR, hazard ratio; ITT, intention-to-treat. **Note:** Kaplan–Meier curves based on investigator assessment The 'stepped' nature of the KM curve reflects the timepoints of clinical assessment. The 'steps' are not perfectly vertical as there was a one-week window either side of the 8-week timepoint where assessments could take place. The assessment of progression on an 8-weekly basis (give or take a week) mirrors assessment in NHS clinical practice. In our submission, for PFS and ToT we used the KM curves for this

reason. When KM data was no longer available we extrapolated from the last data point onwards using the exponential function.

However, it is clear that actual disease progression does not 'conveniently' align with clinical assessment, but will occur as on a daily basis. This exact timepoint of progression could only be ascertained with daily clinical assessment which does not occur in practice. It is precisely because progression, in practice, is 'diagnosed' at intervals determined by clinic visits, and that subsequent treatment decisions are made at those same visits, that the KM curve best reflects clinical practice i.e. it preserves the stepwise nature of clinical practice. This is why ToT is also stepped in nature.

Nonetheless, we appreciate the EAGs concern i.e. if some NHS practice has a different schedule for clinic assessment then direct use of the KM data as we have done could be described as 'overfitting'. We have therefore provided cost-effectiveness results using full parametisation (see table 5.1 and table 5.2). The results indicate that the decision to use KM data directly or to use 'smoothed' functions has no meaningful impact on cost-effectiveness.

Table 5.1: deterministic NMB versus T/T (other inputs per CS)

KM curve*	Weibull	Log-normal	Log-logistic	Exp.	Gen. Gamma	Gompertz	Gamma
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Key: Exp., exponential, gen., generalized

Note: *Company base case

<table>
<tr>
<td></td>
<td></td>
<td>Table 5.2: deterministic NMB versus T/T (other inputs per ERG preference as indicated in section 6.1 of ERG report)
<table>
<tr><th>KM curve</th><th>Weibull</th><th>Log-normal</th><th>Log-logistic*</th><th>Exp.</th><th>Gen. Gamma</th><th>Gompertz</th><th>Gamma</th></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td colspan="8">Key: Exp., exponential, gen., generalized
Note: *EAG base case</td></tr>
</table>
</td>
</tr>
<tr>
<td>Key issue 6: Adverse events
The company positions regorafenib as being a chemotherapy-free alternative therapy that has a different adverse event (AE) profile compared to trifluridine/tipiracil (T/T). However, minor to moderate AEs (grade 1 and 2) were not considered in their economic model. Only > grade 3 AEs (severe) were considered.</td>
<td>No</td>
<td>We believe we have been conservative by including AE disutility separately in the economic model – this is because the impact of AEs on a patient's quality of life will be, to some extent, already 'captured' in the patients EQ5D responses i.e. if the patient is experiencing an AE (any grade) at the time of the EQ5D response then its impact on QOL will be included. In this regard, the explicit addition of AE disutilities may have an element of double-counting.

Grade 1 (mild) and Grade 2 (moderate) adverse events are expected to have a negligible impact on costs or quality of life and were therefore not included in the economic model – this is an approach that is consistent with other modelling in oncology.

Nonetheless, we have endeavoured to incorporate grade 1 / 2 adverse events into the model to ascertain potential impact of cost-effectiveness.

Approach to scenario analysis

We used the trial publications to determine the incidence of grade 1 and 2 AEs. However, data were only reported for either grade 3+ AEs or all AEs, so it was not possible to assess the impact of grade 1 and grade 2 AEs individually. Instead, grade 1-2 AEs were considered as a single group. Details on how grade 1-2 AE information was included in the model are available below.</td>
</tr>
</table>

		Adverse events (and frequencies) to include in the economic model 1) **Data on adverse events was extracted from the publications for the 5 RCTs**: - CORRECT (Grothey 2013 Table 2) - CONCUR (Li 2015 Table 2) - RECOURSE (Mayer 2015 Table 2) - Yoshino 2012 (Table 2) - TERRA (Xu 2017 Table 2) 2) **A common AE reporting definition was applied across the publications** i.e. the criteria used in the RECOURSE study were applied to the other 4 studies. RECOURSE provided the only definition that could be applied to the other studies (see explanation below). This was the only way to remove the inherent bias of including more adverse events in the model for one treatment solely on the basis of different reporting thresholds. <u>Explanation</u> There were different reporting criteria for adverse events between the 5 publications. Notably, reporting differed according to the frequency threshold and the requirement for adverse events to be experienced in a greater proportion of patients in the intervention arm than the placebo arm: - Yoshino reported adverse events with a frequency of at least 3% - TERRA reported adverse events that are listed as most common occurred in at least 10% of patients, or were grade 3+ in either arm - RECOURSE reported adverse events of any grade that are listed as most common occurred in 10% or more of patients in the T/T group and in a greater percentage in that group than in the placebo group - CORRECT reported adverse events occurring in ≥5% of patients in either group

- CONCUR reported adverse events occurring at any grade in at least 10% of patients, or at grade 3 or higher in any patients in either group

3) **the combined frequency of grade 1/2 adverse events was calculated as follows i.e. (grade 1/2 AEs) = (all AEs) – (grade 3+ AEs)**

Application of steps 1 – 3 provided the following adverse events and frequencies which were included in the model

Table 6.1 – grade 1/2 adverse events included in the model

	Regorafenib		T/T		
	CORRECT (n=500)	**CONCUR (n=136)**	**RECOURSE (n=533)**	**TERRA (n=271)**	**Yoshino (n=113)**
Adverse event					
Anaemia			(308/528) = 58.3%	(161/271) = 59.4%	(63/113) = 55.6%
Asthenia			(79/533) = 14.8%		
Abdominal pain			(100/533) = 18.8%		
Anorexia	(136/500) = 27.2%				(65/113) = 57.5%
Decreased appetite			(189/533) = 35.5%	(65/271) =24.0%	
Diarrhoea	(133/500) = 26.6%	(23/136) =16.9%	(154/533) = 28.9%	(38/271) = 14.0%	(36/113) = 31.9%
Fatigue	(189/500) = 37.8%	(19/136) =14%	(167/533) = 31.3%	(51/271) = 18.8%	(59/113) = 52.2%
Fever	(48/500) = 9.6%		(92/533) = 17.3%		
Hand-foot skin reaction	(150/500) = 30%	(78/136) =57.4%			
Hoarseness		(27/136) =19.9%			
Hypertension	(103/500) = 20.6%	(16/136) =11.7%			
Nausea	(70/500) = 14.0%		(248/533) = 46.5%	(96/271) = 35.4%	(68/113) = 60.2%
Oral Mucositis	(121/500) = 24.2%				
Rash or desquamation	(101/500) = 20.2%				
Stomatitis			41/533 = 7.7%		

Weight loss	(69/500) = 13.8%					
Vomiting				(137/533) = 25.7%	(48/271) = 17.7%	(34/113) = 30.1%
Voice changes	(131/500) = 26.2%					
Laboratory abnormalities						
Leukopenia				(294/528) = 55.7%	(134/271)= 49.4%	(54/113) = 47.8%
Lymphopenia					(107/271) =39.5%	(28/113) = 24.8%
Neutropenia				(153/528) = 29.0%	(92/271) = 33.9%	(24/113) = 21.2%
Thrombocytopenia	(49/500) = 9.8%			(196/528) = 37.1%	(88/271) = 32.5%	(39/113) = 34.5%
Increased alanine aminotransferase (ALT)			(23/136) =16.9%	(116/526) = 22.1%	(80/271) = 29.5%	
Increase in total bilirubin				(144/526) = 27.4%		
Hyperbilirubinaemia	(35/500) =7.0 %		(41/136) =30.2%			
Increase in alkaline phosphatase				(163/526) = 31.0%	(80/271) = 29.5%	
Increase in creatinine				(66/527) = 12.5%		
Hyponatraemia					(69/271) = 25.5%	
Hypocalcaemia					(74/271) = 27.3%	
Increase in AST (aspartate transaminase) levels			(24/136) =17.7%	132/524 = 25.2%		
Hypokalemia					(29/271) = 10.7%	

Cost of treating Adverse events

Grade 1 adverse events are mild and are not treated. Grade 2 adverse events we understand to be infrequently treated, and if they are it is with inexpensive medicines. For example laxatives for constipation and analgesics for pain/discomfort. In respect of low grade haematological abnormalities or elevated liver enzymes, ongoing and existing monitoring will be employed.

As a scenario analysis we have included a cost of £5 per adverse event. This cost represents a pragmatic approach and is informed by the cost of laxatives (senna 7.5mg tab 60 = £2.03) and simple analgesics (paracetamol 500mg tab 100 = £2.34). As we expect only a proportion of patients to receive medication a flat cost of £5 seems conservative.

Disutilities associated with grade 1 / 2 adverse events

We do not consider that grade 1 - 2 adverse events are associated with a significant decrease in quality of life, and certainly not a disutility that could be discriminated using the EQ5D.

In a scenario analysis we have included a disutility per AE of 0.01. This utility decrement is arbitrarily taken as 10% of 0.1 i.e. the typical decrement for a grade 3+ AE in the economic model. We consider this to be an extreme scenario as these AEs are defined as mild/moderate and many are laboratory abnormalities of which the patient will be unaware.

The results of the inclusion of grade 1 - 2 adverse events are shown in the tables below. As can be seen there is a negligible impact on cost-effectiveness i.e. a less than £[redacted] change in NMB when both costs and disutilities are included.

Table 6.1 – Model results with grade 1-2 AEs included, assuming £5 per AE (no disutility applied)

	Total costs	Total QALYs	ICER	NMB
Regorafenib	[redacted]	[redacted]	[redacted]	[redacted]
T/T	[redacted]	[redacted]	[redacted]	[redacted]

<table>
<tr>
<td></td>
<td></td>
<td>Table 6.2 – Model results with grade 1 -2 AEs included, assuming £5 per AE and a disutility per AE of 0.01
<table>
<tr><th></th><th>Total costs</th><th>Total QALYs</th><th>ICER</th><th>NMB</th></tr>
<tr><td>Regorafenib</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td></tr>
<tr><td>T/T</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td></tr>
</table>
</td>
</tr>
<tr>
<td>Key issue 7: Resource use and costs
The company assumed different relative dose intensity (RDIs) for regorafenib and trifluridine/tipiracil (T/T). However, a large observational study by Nakashima 2020, directly comparing regorafenib to T/T, reported comparable dose reductions (54% and 48% respectively).</td>
<td>No</td>
<td>We recognise that RDI is an uncertainty in our modelling. We therefore presented a scenario in the CS using an equal RDI for both regorafenib and T/T – this analysis [redacted].</td>
</tr>
<tr>
<td>Key issue 8: Company’s sensitivity analyses
Uncertainty due to the company’s failure to provide the following results:</td>
<td>No</td>
<td>Please find below the results of the analyses we were unable to complete at the clarification question stage of the appraisal. We apologise that we were unable to complete these at the time. We have not conducted scenario analyses against BSC as we do not consider BSC to be a comparator.

The results of the RMST are shown in Table 8.1. We were only able to calculate the observed RMST for regorafenib, as we do not have access to survival data for T/T that was pooled across T/T studies. We therefore estimated the RMST for T/T using the PFS and OS from the model instead. As RMST cut-offs we</td>
</tr>
</table>

1) observed and modelled overall survival (OS) and progression-free survival (PFS) using restricted mean survival time (RMST)
2) scenario analyses for the best supportive care (BSC) comparison that were conducted for trifluridine/tipiracil (T/T) comparison
3) scenario analyses with the fully parametric models applied for PFS and time on treatment (ToT).

used 6 months, 12 months, and the last observed timepoint for regorafenib, i.e. 17 months for OS and 16 months for PFS.

The RMST showed that the modelled results for regorafenib are in line with the observed trial data, with the observed RMST life-years (LY) being closely aligned to the estimated RMST LY from the model (Table 8.1). In addition, the RMST showed that 27.2% of the total LY for regorafenib and 26.7% of the LY for T/T were modelled after the last OS observation, indicating that the majority of the survival in the model is informed by observed data. Although 66% of the LY gained by regorafenib were after the last observation, this only amounted to 0.09 LYs, considering the equivalent survival between regorafenib and T/T. Altogether, this analysis indicates that there is little uncertainty around the survival estimates in the model, as the observed RMST is well aligned with the modelled RMST, and the majority of the survival in the model is informed by observed data.

Table 8.1. RMST results of OS and PFS for regorafenib and T/T

	Observed LY	Modelled LY		
	Restricted mean survival time (RMST)	Estimated (RMST)	Estimated (lifetime time horizon)	Proportion beyond observed data
OS - RMST period / truncation point: 6 months				
Regorafenib	[redacted]	[redacted]	[redacted]	[redacted]
T/T	[redacted]	[redacted]	[redacted]	[redacted]
Increment	[redacted]	[redacted]	[redacted]	[redacted]
OS - RMST period / truncation point: 12 months				
Regorafenib	[redacted]	[redacted]	[redacted]	[redacted]
T/T	[redacted]	[redacted]	[redacted]	[redacted]
Increment	[redacted]	[redacted]	[redacted]	[redacted]
OS - RMST period / truncation point: 17 months (Last observation)				
Regorafenib	[redacted]	[redacted]	[redacted]	[redacted]

T/T	[redacted]	[redacted]	[redacted]	[redacted]	
Increment	[redacted]	[redacted]	[redacted]	[redacted]	
PFS - RMST period / truncation point: 6 months					
Regorafenib	[redacted]	[redacted]	[redacted]	[redacted]	
T/T	[redacted]	[redacted]	[redacted]	[redacted]	
Increment	[redacted]	[redacted]	[redacted]	[redacted]	
PFS - RMST period / truncation point: 12 months					
Regorafenib	[redacted]	[redacted]	[redacted]	[redacted]	
T/T	[redacted]	[redacted]	[redacted]	[redacted]	
Increment	[redacted]	[redacted]	[redacted]	[redacted]	
PFS - RMST period / truncation point: 16 months (Last observation)					
Regorafenib	[redacted]	[redacted]	[redacted]	[redacted]	
T/T	[redacted]	[redacted]	[redacted]	[redacted]	
Increment	[redacted]	[redacted]	[redacted]	[redacted]	

The results of the scenarios with parametric PFS and ToT input are shown in Table 8.2. In line with the results shown in Issue 5, using parametric PFS and ToT input only had a limited impact on all model results, leading to minor differences in the total costs and QALYs in all scenarios. In addition, as with the base case, regorafenib continued to [redacted].

Table 8.2: Scenario results with parametric PFS and ToT input as the base case

		Regorafenib		Trifluridine/tipiracil			
#	**Scenario Name**	**Total costs**	**Total QALYs**	**Total costs**	**Total QALYs**	**ICER**	**NMB**
-	**Base case – Parametric PFS and ToT**	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
1	CONCUR efficacy data only	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
2	CORRECT efficacy data only	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
3	Weibull OS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
4	Log-normal OS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
5	Exponential OS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
6	Generalized gamma OS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
7	Gompertz OS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
8	Gamma OS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
9	KM data for PFS and TOT	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
10	NMA without Yoshino	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
11	CORRECT vs RECOURSE ITC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
12	CONCUR vs TERRA and Yoshino ITC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
13	CONCUR vs TERRA ITC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
14	CONCUR vs Yoshino ITC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
15	Assume equal efficacy	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
16	Apply RDI on pack count	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
17	Equal T/T RDI	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
18	Include post-progression treatment costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
19	Exclude AE disutilities	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

		20 | 5 year time horizon | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] 21 | No discounting | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] | [redacted] **Key:** AE, adverse event; Dom, dominated; ITC, indirect treatment comparison; NMA: network meta-analysis; OS, overall survival; PFS, progression-free survival, RDI, relative dose intensity; [redacted]; ToT, time on treatment; T/T, trifluridine/tipiracil. [redacted]
Key issue 9: Lack of a fixed random seed in model PSA The results of the probabilistic sensitivity analysis (PSA) are slightly different when running the same analysis multiple times (without changing model settings), likely due to the lack of a fixed random seed in the model PSA.	No	We can confirm that the small variability in results in different PSA runs is due to the random seed in the model not being fixed. We have provided a model with a fixed random number seed. This ensures the same sequence of random numbers are used for each probabilistic analysis.

Table 2 Key issues

Additional issues

All: Please use the table below to respond to additional issues in the EAR that have not been identified as key issues. Please do **not** use this table to repeat issues or comments that have been raised at an earlier point in this evaluation (for example, at the clarification stage).

Table 3 Additional issues from the EAR

Issue from the EAR	Relevant section(s) and/or page(s)	Does this response contain new evidence, data or analyses?	Response
Additional issue 1: Insert additional issue	Please indicate the section(s) of the EAR that discuss this issue	Yes/No	Please include your response, including any new evidence, data or analyses, and a description of why you think this is an important issue for decision making
Additional issue 2: Insert additional issue	Please indicate the section(s) of the EAR that discuss this issue	Yes/No	Please include your response, including any new evidence, data or analyses, and a description of why you think this is an important issue for decision making
Additional issue N: Insert additional issue			[INSERT / DELETE ROWS AS REQUIRED]

Summary of changes to the company's cost-effectiveness estimate(s)

Company only: If you have made changes to the base-case cost-effectiveness estimate(s) in response to technical engagement, please complete the table below to summarise these changes. Please also provide sensitivity analyses around the revised base case. If there are sensitivity analyses around the original base case which remain relevant, please re-run these around the revised base case.

Table 4 Changes to the company's cost-effectiveness estimate

Key issue(s) in the EAR that the change relates to	Company's base case before technical engagement	Change(s) made in response to technical engagement	Impact on the company's base-case incremental cost-effectiveness ratio (ICER)
Insert key issue number and title as described in the EAR	Briefly describe the company's original preferred assumption or analysis	Briefly describe the change(s) made in response to the EAR	Please provide the ICER resulting from the change described (on its own), and the change from the company's original base-case ICER.
Insert key issue number and title as described in the EAR	…	…	[INSERT / DELETE ROWS AS REQUIRED]
Company's base case following technical engagement (or revised base case)	Incremental QALYs: [QQQ]	Incremental costs: [£££]	Please provide company revised base-case ICER

Sensitivity analyses around revised base case

PLEASE DESCRIBE HERE

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Regorafenib for treating metastatic colorectal cancer [ID4002]

Patient expert statement and technical engagement response form

Thank you for agreeing to give us your views on this treatment and its possible use in the NHS.

Your comments and feedback on the key issues below are really valued. You can provide a unique perspective on conditions and their treatment that is not typically available from other sources. The external assessment report (EAR) and stakeholder responses are used by the committee to help it make decisions at the committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

Information on completing this form

In part 1 we are asking you about living with metastatic colorectal cancer or caring for a patient with metastatic colorectal cancer.

The text boxes will expand as you type.

In part 2 we are asking for your views on key issues in the EAR that are likely to be discussed by the committee. The key issues in the EAR reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the EAR (listed in Section 1.1 with more explanation in Sections 1.3, 1.4 and 1.5).

A patient perspective could help either:

- resolve any uncertainty that has been identified OR

Patient expert statement

- provide missing or additional information that could help committee reach a collaborative decision in the face of uncertainty that cannot be resolved.

You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise. We have given guidance on the issues in which we expect this to be the case and advice on what you could consider when giving your response.

In part 3 we are asking you to provide 5 summary sentences on the main points contained in this document.

Help with completing this form

If you have any questions or need help with completing this form please email the public involvement (PIP) team at pip@nice.org.uk (please include the ID number of your appraisal in any correspondence to the PIP team).

Please use this questionnaire with our hints and tips for patient experts. You can also refer to the Patient Organisation submission guide. **You do not have to answer every question** – they are prompts to guide you. There is also an opportunity to raise issues that are important to patients that you think have been missed and want to bring to the attention of the committee.

Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable. Please type information directly into the form.

We are committed to meeting the requirements of copyright legislation. If you want to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

Your response should not be longer than 15 pages.

Please note, **part 1** can be completed at any time. We advise that **part 2** is completed after the expert engagement teleconference (if you are attending or have attended). At this teleconference we will discuss some of the key issues, answer any specific questions you may have about the form, and explain the type of information the committee would find useful.

The deadline for your response is **5pm on 21st September 2022.** Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

Part 1: Living with this condition or caring for a patient with metastatic colorectal cancer

Table 1 About you, metastatic colorectal cancer, current treatments and equality

1. Your name	[redacted]
2. Are you (please tick all that apply)	X☐ A patient with metastatic colorectal cancer ? ☐ A patient with experience of the treatment being evaluated? ☐ A carer of a patient with metastatic colorectal cancer ? ☐ A patient organisation employee or volunteer? ☐ Other (please specify):
3. Name of your nominating organisation	
4. Has your nominating organisation provided a submission? (please tick all options that apply)	☐ No (please review all the questions and provide answers when possible) ☐ Yes, my nominating organisation has provided a submission ☐ I agree with it and **do not wish to** complete a patient expert statement ☐ Yes, I authored / was a contributor to my nominating organisations submission ☐ I agree with it and **do not wish to** complete this statement X☐ I agree with it and **will be** completing

5. How did you gather the information included in your statement? (please tick all that apply)	**X**☐ I am drawing from personal experience X☐ I have other relevant knowledge or experience (for example, I am drawing on others' experiences). Please specify what other experience: People with mCRC ☐ I have completed part 2 of the statement **after attending** the expert engagement teleconference ☐ I have completed part 2 of the statement **but was not able to attend** the expert engagement teleconference X☐ I have not completed part 2 of the statement

6. What is your experience of living with metastatic colorectal cancer? **If you are a carer (for someone with metastatic colorectal cancer) please share your experience of caring for them**	I was diagnosed in 2016 with two primary bowel cancers that required extensive surgery and adjuvant chemotherapy for 6 months, a regimen of Oxaliplatin and Capecitabine. I was diagnosed with Lynch syndrome being the cause of my cancer diagnosis. Three months after completing adjuvant chemotherapy and returning to work as a Clinical Nurse Specialist I was diagnosed with metastatic bowel cancer that had spread to my liver (multiple tumours), adrenal gland, local and distant lymph nodes, skull bones, spine and pelvis bones. This came with a poor prognosis and was advised that any chemotherapy treatment would have little impact. I commenced FOLFIRI and Panitumumab. I had a positive response and continued on treatment for 2.5 years. I suffered with severe toxicity issues and side effects that prevented me from working, leaving my home and had a huge impact both physically and mentally on my quality of life. However, due to NHS funding issues I wasn't allowed to have a break off my treatment as I would lose the funding for drug that was keeping my cancer stable. Eventually my cancer progressed. As I had MSI-H, Immunotherapy was an option but unavailable from the NHS. I was awarded compassionate use from the drug company for Nivolumab. After two years of treatment I had a complete metabolic response and stopped treatment. I have remained in remission for 1 year and currently having 12 weekly surveillance scans. I remain unable to work due to ongoing effects of treatments that have caused multiple health issues.

7a. What do you think of the current treatments and care available for metastatic colorectal cancer on the NHS? **7b. How do your views on these current treatments compare to those of other people that you may be aware of?**	Current treatments for metastatic colorectal cancer from the NHS are very limited. The treatments options available appear to be received with a negative outset and poor efficacy for the majority of people. Having MSI-H gave me alternative options but this is for the minority of the population. Many people I speak with are funding drugs privately with a postcode lottery of costs that vary hugely. It feel's like a constant battle to try and get treatments for people particularly the younger population who are seeking options. People endeavour access to many drug options that are approved and readily available in Europe and USA. There is an inconsistency of treatment pathways which drive people to try and access some larger cancer centres as the options and re challenge of treatments appear to be more accessible at bigger centres. Holistic care for metastatic colorectal cancer in my experience is underfunded. I have not received any support (psychological, emotional, well-being, financial advice) since being diagnosed with metastatic colorectal cancer from my local team. I have been my own advocate, challenged my treatment plan and researched treatment. From my experience, people diagnosed with other cancers have significantly more options and support in every aspect of care.
8. If there are disadvantages for patients of current NHS treatments for metastatic colorectal cancer (for example, how they are given or taken, side effects of treatment, and any others) please describe these	Side effects from platinum based drugs can be severe and cause long term neuropathy issues and liver damage which I have been diagnosed with. Side effects from anti - EGFR can cause severe skin toxicity issues which I experienced. This had a huge impact on my quality of life, prevented me from leaving my house, unable to go outside in day light. Having steroid treatment for skin toxicity caused me to have adrenal insufficiency which made me extremely unwell.

9a. If there are advantages of regorafenib over current treatments on the NHS please describe these. For example, the effect on your quality of life, your ability to continue work, education, self-care, and care for others? **9b. If you have stated more than one advantage, which one(s) do you consider to be the most important, and why?** **9c. Does regorafenib help to overcome or address any of the listed disadvantages of current treatment that you have described in question 8? If so, please describe these**	Providing more options for metastatic colorectal cancer can only be an advantage to many people given that the current options are so limited. Side effects with Regorafenib report to be different to the comparator and therefore could open up options of one drug at third line, if one drug is not tolerated. People want option's when other treatments have failed to increase overall survival.
10. If there are disadvantages of regorafenib over current treatments on the NHS please describe these. For example, are there any risks with regorafenib? If you are concerned about any potential side effects you have heard about, please describe them and explain why	
11. Are there any groups of patients who might benefit more from regorafenib or any who may benefit less? If so, please describe them and explain why Consider, for example, if patients also have other health conditions (for example difficulties with mobility, dexterity or cognitive impairments) that affect the suitability of different treatments	Advantage of oral treatment that does not require regular hospital attendances will be an advantage to many people particularly the more old and frail.

12. Are there any potential equality issues that should be taken into account when considering metastatic colorectal cancer and regorafenib? Please explain if you think any groups of people with this condition are particularly disadvantaged Equality legislation includes people of a particular age, disability, gender reassignment, marriage and civil partnership, pregnancy and maternity, race, religion or belief, sex, and sexual orientation or people with any other shared characteristics More information on how NICE deals with equalities issues can be found in the NICE equality scheme Find more general information about the Equality Act and equalities issues here.	
13. Are there any other issues that you would like the committee to consider?	

Part 2: Technical engagement questions for patient experts

Issues arising from technical engagement

The issues raised in the EAR are listed in table 2. We welcome your comments on the issues, but you do not have to provide a response to every issue, such as the ones that are technical, that is, cost effectiveness-related issues. We have added a comment to the issues where we consider a patient perspective would be most relevant and valuable. If you think an issue that is important to patients has been missed in the EAR, please let us know in the space provided at the end of this section.

For information: the patient organisation that nominated you has also been sent a technical engagement response form (a separate document) which asks for comments on each of the key issues that have been raised in the EAR, the patient organisation responses will also be considered by the committee.

Table 2 Issues arising from technical engagement

Key issue 1: Potential ambiguity of the population in the decision problem, which has implications for the comparators The company considers regorafenib suitable for people whose disease has progressed following first-line treatment and who are being considered for ≥ 3rd-line treatment. This definition (≥ 3rd-line treatment) suggests the appropriate comparator could be either trifluridine/tipiracil or best supportive care. The comparator influences cost-effectiveness estimates.	

Key issue 2: Potential difference between subsequent treatment use in the RCTs and NHS clinical practice with unknown effect of subsequent treatment In clinical trials, more people were given subsequent treatment after regorafenib, than the company estimates. The proportion of people receiving post-progression treatment after regorafenib in CORRECT and CONCUR trials were 21% and 31%, respectively. Whereas, the company has suggested that in UK clinical practice <10% of people would be fit enough for subsequent treatment.	

Key issue 3: Lack of external validity of and comparability between the regorafenib and trifluridine/tipiracil RCTs The differences in prior treatment, and race of people enrolled in the relevant clinical trials makes it difficult to compare different trials. Similarly, comparison with NHS practice is difficult. We consider patient perspectives may particularly help to address this issue.	

Key issue 4: Difference in the treatment effect of regorafenib versus trifluridine/tipiracil depending on evidence source Randomised controlled trials (RCTs) show there is little difference between regorafenib and trifluridine/tipiracil (T/T) in either overall survival (OS) or progression-free survival (PFS). The combined trials comparing regorafenib with T/T (network meta-analysis) reports a hazard ratio (HR) of 0.99. However, a large comparative observational study by Nakashima 2020 showed an advantage to T/T (HR = 1.515).	

Key issue 5: Treatment effectiveness and extrapolation Using the appropriate model ensures treatment benefit is not overestimated. The company implemented Kaplan-Meier (KM) curves instead of parametric survival models for the survival analyses of progression-free survival (PFS) and time on treatment (ToT) for regorafenib and best supportive care (BSC), leading to potential overfitting of modelled survival outcomes.	

Key issue 6: Adverse events This can affect the cost-effectiveness estimates. The company positions regorafenib as being a chemotherapy-free alternative therapy that has a different adverse event (AE) profile compared to trifluridine/tipiracil (T/T). However, minor to moderate AEs (grade 1 and 2) were not considered in their economic model. Only > grade 3 AEs (severe) were considered. We consider patient perspectives may particularly help to address this issue.	

Key issue 7: Resource use and costs The company assumed different relative dose intensity (RDIs) for regorafenib and trifluridine/tipiracil (T/T). However, a large observational study by Nakashima 2020, directly comparing regorafenib to T/T, reported comparable dose reductions (54% and 48% respectively).	

Key issue 8: Company's sensitivity analyses Sensitivity analyses can tell us which factors affect the cost-effectiveness estimates the most. Without these there remain factors which we are uncertain about. When asked, the company did not provide the following: 1) observed and modelled overall survival (OS) and progression-free survival (PFS) using restricted mean survival time (RMST) 2) scenario analyses for the best supportive care (BSC) comparison that were conducted for trifluridine/tipiracil (T/T) comparison 3) scenario analyses with the fully parametric models applied for PFS and time on treatment (ToT).	

Key issue 9: Lack of a fixed random seed in model PSA Due to settings in the company's economic model, the cost-effectiveness estimates change each time the model is run. The results of the probabilistic sensitivity analysis (PSA) are slightly different when running the same analysis multiple times (without changing model settings), likely due to the lack of a fixed random seed in the model PSA.	
Are there any important issues that have been missed in EAR?	

Part 3: Key messages

In up to 5 sentences, please summarise the key messages of your statement:

- Click or tap here to enter text.
- Click or tap here to enter text.
- Click or tap here to enter text.
- Click or tap here to enter text.
- Click or tap here to enter text.

Thank you for your time.

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see NICE's privacy notice.

NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Regorafenib for treating metastatic colorectal cancer [ID4002]

Technical engagement response form

As a stakeholder you have been invited to comment on the External Assessment Report (EAR) for this evaluation.

Your comments and feedback on the key issues below are really valued. The EAR and stakeholders' responses are used by the committee to help it make decisions at the committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

Information on completing this form

We are asking for your views on key issues in the EAR that are likely to be discussed by the committee. The key issues in the EAR reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the EAR.

You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise.

If you would like to comment on issues in the EAR that have not been identified as key issues, you can do so in the 'Additional issues' section.

If you are the company involved in this evaluation, please complete the 'Summary of changes to the company's cost-effectiveness estimates(s)' section if your response includes changes to your cost-effectiveness evidence.

Please do not embed documents (such as PDFs or tables) because this may lead to the information being mislaid or make the response unreadable. Please type information directly into the form.

Do not include medical information about yourself or another person that could identify you or the other person.

We are committed to meeting the requirements of copyright legislation. If you want to include journal articles in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.

Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, and all information submitted under 'depersonalised data' in pink. If confidential information is submitted, please also send a second version of your comments with that information redacted. See the NICE health technology evaluation guidance development manual (sections 5.4.1 to 5.4.10) for more information.

The deadline for comments is the end of day on **21st September 2022**. Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

NICE National Institute for Health and Care Excellence

About you

Table 1 About you

Your name	
Organisation name: stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder, please leave blank)	Servier Laboratories
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	**None**

Key issues for engagement

All: Please use the table below to respond to the key issues raised in the EAR.

Table 2 Key issues

Key issue	Does this response contain new evidence, data or analyses?	Response
Key issue 1: Potential ambiguity of the population in the decision problem, which has implications for the comparators The company considers regorafenib suitable for people whose disease has progressed following first-line chemotherapy/first-line biologic and who are being considered for ≥ 3rd-line treatment. This definition (≥ 3rd-line treatment) suggests the appropriate comparator could be either trifluridine/tipiracil or best supportive care. The comparator influences cost-effectiveness estimates.	Yes/No	Servier recommend to include patients who might not get T/T e.g., those not fit enough to receive it and who might only be eligible for BSC. Due to the different adverse event profile of T/T and Regorafenib there may be some pts who cant take T/T but would still be suitable for Regorafenib and therefore the BSC becomes the comparator

Key issue 2: Potential difference between subsequent treatment use in the RCTs and NHS clinical practice with unknown effect of subsequent treatment The proportion of people receiving post-progression treatment after regorafenib in CORRECT and CONCUR trials were 21% and 31%, respectively. Whereas, the company has suggested that in UK clinical practice <10% of people would be fit enough for post-progression active treatment.	Yes/No	Please provide your response to this key issue, including any new evidence, data or analyses
Key issue 3: Lack of external validity of and comparability between the regorafenib and trifluridine/tipiracil RCTs It is uncertain as to the size and direction of subgroup differences in terms of treatment experience and race in the CORRECT and CONCUR trials. There is also a disparity between the trials and NHS clinical practice.	Yes/No	Please provide your response to this key issue, including any new evidence, data or analyses
Key issue 4: Difference in the treatment effect of regorafenib versus trifluridine/tipiracil depending on evidence source The network meta-analysis (NMA) of randomised controlled trials	Yes	Nakashima shows better outcomes with trifluridine/tipiracil as well as the publication by Patel et al, 2021. (Trifluridine /Tipiracil and Regorafenib in patients with metastatic colorectal Cancer. The Oncologist; 26;e2161-2169) This retrospective cohort study found that patients treated with trifluridine tipiracil had a significantly better overall response rate and a better disease control rate. There was also better overall survival with trifluridine tipiracil. Servier is concerned that

(RCTs) for the comparison of regorafenib with trifluridine/tipiracil (T/T) indicates that, depending on which trials are included, there is little difference in either overall survival (OS) or progression-free survival (PFS). However, a large comparative observational study by Nakashima 2020 showed an advantage to T/T in OS, with a hazard ratio (HR) of 1.515 in contrast with the indirect treatment comparison (ITC) base-case, HR = 0.99.		this has not been factored in to the economic model. Servier agrees with the EAG that although not routinely done, the RCTs and observational comparative studies could be combined in a NMA, using methods as described in Technical Support Document (TSD) TSD 20 Further studies for consideration in this analysis are Ogato et al 2020 https://www.frontiersin.org/articles/10.3389/fonc.2022.867546/full Meng-Che Hsieh et al 2022 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7292354/, Ageo et al, 2022 (Clinical Colorectal Cancer, Vol. 21, No. 2, 132–140), and Patel et al, 2021. (Trifluridine /Tipiracil and Regorafenib in patients with metastatic colorectal Cancer. The Oncologist; 26;e2161-2169) These observational studies all show an advantage to T/T in OS or PFS and therefore should be taken in to consideration to some extent.
Key issue 5: Treatment effectiveness and extrapolation The company implemented Kaplan-Meier (KM) curves instead of parametric survival models for the survival analyses of progression-free survival (PFS) and time on treatment (ToT) for regorafenib and best supportive care (BSC), leading to potential overfitting of modelled survival outcomes.	Yes/No	Please provide your response to this key issue, including any new evidence, data or analyses
Key issue 6: Adverse events The company positions regorafenib as being a chemotherapy-free	Yes	Servier provides a published retrospective analysis by Huemer et al, 2020. (*Cancers*, 12, 2812; doi:10.3390/cancers12102812) to investigate hospitlisation frequency during treatment with T/T and regorafenib. It found that treatment with regorafenib was an independent risk factor for hospitalisation (HR 1.95), and

alternative therapy that has a different adverse event (AE) profile compared to trifluridine/tipiracil (T/T). However, minor to moderate AEs (grade 1 and 2) were not considered in their economic model. Only > grade 3 AEs (severe) were considered.		hospitalisations due to gastrointestinal toxicity was only found with regorafenib. Servier are unsure if this has currently been accounted for within the economic model although a copy of the model has been requested Servier believe that there are reasons why Grade 1- 2 AEs should have been considered, in particular the Hand foot skin reaction which is clinically relevant and its impact on QOL and discontinuation. No data on the economic cost of treating regorafenib-related HFSR have been reported. However, an analysis of the cost of managing cutaneous toxicities associated with sorafenib and sunitinib for cancer at a single US dermatology department found that sorafenib-related HFSR was the most costly cutaneous toxicity to manage, accounting for a median medication cost of $968 per patient (Borovicka JH, Calahan C, Gandhi M et al.. Economic burden of dermatologic adverse events induced by molecularly targeted cancer agents. *Arch Dermatol* 2011; 147: 1403–1409) The fact that 47% of patients had this at any grade needs to be considered within the model and associated costs factored in.
Key issue 7: Resource use and costs The company assumed different relative dose intensity (RDIs) for regorafenib and trifluridine/tipiracil (T/T). However, a large observational study by Nakashima 2020, directly comparing regorafenib to T/T, reported	No	Servier support that in a large observational study by Nakashima 2020, directly comparing regorafenib to T/T, there were comparable dose reductions, although the dose of T/T was slightly lower (54% and 48% respectively).

comparable dose reductions (54% and 48% respectively).		
Key issue 8: Company's sensitivity analyses Uncertainty due to the company's failure to provide the following results: 1) observed and modelled overall survival (OS) and progression-free survival (PFS) using restricted mean survival time (RMST) 2) scenario analyses for the best supportive care (BSC) comparison that were conducted for trifluridine/tipiracil (T/T) comparison 3) scenario analyses with the fully parametric models applied for PFS and time on treatment (ToT).	Yes/No	
Key issue 9: Lack of a fixed random seed in model PSA The results of the probabilistic sensitivity analysis (PSA) are slightly different when running the	Yes/No	

same analysis multiple times (without changing model settings), likely due to the lack of a fixed random seed in the model PSA.		

Additional issues

All: Please use the table below to respond to additional issues in the EAR that have not been identified as key issues. Please do **not** use this table to repeat issues or comments that have been raised at an earlier point in this evaluation (for example, at the clarification stage).

Table 3 Additional issues from the EAR

Issue from the EAR	Relevant section(s) and/or page(s)	Does this response contain new evidence, data or analyses?	Response
Additional issue 1: Cost of T/T	Cost of T/T	Yes	Trifluridine/tipiracil is available in 15 mg and 20 mg tablets for £500.00 and £666.67 per 20 tablets (NHS list price), respectively. This is what has been used in the company model. [redacted]
Additional issue 2: Insert additional issue	Please indicate the section(s) of the EAR that discuss this issue	Yes/No	Please include your response, including any new evidence, data or analyses, and a description of why you think this is an important issue for decision making
Additional issue N: Insert additional issue			[INSERT / DELETE ROWS AS REQUIRED]

NICE National Institute for Health and Care Excellence

Summary of changes to the company's cost-effectiveness estimate(s)

Company only: If you have made changes to the base-case cost-effectiveness estimate(s) in response to technical engagement, please complete the table below to summarise these changes. Please also provide sensitivity analyses around the revised base case. If there are sensitivity analyses around the original base case which remain relevant, please re-run these around the revised base case.

Table 4 Changes to the company's cost-effectiveness estimate

Key issue(s) in the EAR that the change relates to	Company's base case before technical engagement	Change(s) made in response to technical engagement	Impact on the company's base-case incremental cost-effectiveness ratio (ICER)
Insert key issue number and title as described in the EAR	Briefly describe the company's original preferred assumption or analysis	Briefly describe the change(s) made in response to the EAR	Please provide the ICER resulting from the change described (on its own), and the change from the company's original base-case ICER.
Insert key issue number and title as described in the EAR	…	…	[INSERT / DELETE ROWS AS REQUIRED]
Company's base case following technical engagement (or revised base case)	Incremental QALYs: [QQQ]	Incremental costs: [£££]	Please provide company revised base-case ICER

Sensitivity analyses around revised base case
PLEASE DESCRIBE HERE

in collaboration with:

Maastricht University

Regorafenib for treating metastatic colorectal cancer (ID4002)

Technical engagement response critique

Produced by Kleijnen Systematic Reviews (KSR) Ltd, in collaboration with Maastricht University Medical Centre (UMC)

Authors

Nigel Armstrong, Health Economics Manager, KSR Ltd, United Kingdom (UK)
Willem Witlox, Health Economist, Maastricht UMC, the Netherlands
Mark Perry, Reviews Manager, KSR Ltd, UK
Bram Ramaekers, Health Economist, Maastricht UMC, the Netherlands
Evangelos Danopoulos, Systematic Reviewer, KSR Ltd, UK
Bradley Sugden, Health Economist, Maastricht UMC, the Netherlands
Pawel Posadzki, Reviews Manager, KSR Ltd, UK
Andrea Fernandez Coves, Health Economist, Maastricht UMC, the Netherlands
Teebah Abu-Zahra, Health Economist, Maastricht UMC, the Netherlands
Thomas Otten, Health Economist, Maastricht UMC, the Netherlands
Caro Noake, Information Specialist, KSR Ltd, UK
Manuela Joore, Health Economist, Maastricht UMC, the Netherlands
Jos Kleijnen, Founder and Owner, KSR Ltd, UK

Correspondence to Nigel Armstrong, Kleijnen Systematic Reviews Ltd
Unit 6, Escrick Business Park
Riccall Road, Escrick
York, YO19 6FD
United Kingdom

Abbreviations

AIC	Akaike information criterion
AE	Adverse event
AEOSI	Adverse events of special interest
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
BIC	Bayesian information criterion
BMI	Body mass index
BNF	British National Formulary
BRAF	Mutation in the B-Raf proto-oncogene
BSC	Best supportive care
CADTH	Canadian Agency for Drugs and Technologies in Health
CE	Cost effectiveness
CEA	Cost effectiveness analysis
CENTRAL	Cochrane Central Register of Controlled Trials
CI	Confidence interval
CR	Complete response
CrI	Credible interval
CRD	Centre for Reviews and Dissemination
CS	Company submission
CSR	Clinical study report
CT	Computerised tomography
CTC	Common Toxicity Criteria
CTCAE	Common Terminology Criteria for Adverse Events
DCR	Disease control rate
DOR	Duration of response
DSA	Deterministic sensitivity analyses
DSU	Decision Support Unit
eMIT	Electronic market information tool
ECOG	Eastern Cooperative Oncology Group
ECOG PS	Eastern Cooperative Oncology Group Performance Status
eGFR	Estimated glomerular filtration rate
EGFR	Epidermal growth factor receptor
EORTC	European Organisation for Research and Treatment of Cancer
EORTC QLQ C30	European Organization for Research and Treatment of Cancer Quality of Life Questionnaire (30 items)
EOT	End of treatment
EQ-5D	European Quality of Life-5 Dimensions
ERG	Evidence Review Group
ESS	Effective sample size
FAS	Full analysis set
FE	Fixing errors
FOIB	Fluoropyrimidine, oxaliplatin, irinotecan, bevacizumab
FV	Fixing violations
G-CSF	Granulocyte-colony stimulating factor
HR	Hazard ratio
HRQoL	Health-related quality of life
H_0	Null hypothesis
H_1	Alternative hypothesis
HSUV	Health state utility values
HTA	Health Technology Assessment
IC	Indirect comparison
ICER	Incremental cost-effectiveness ratio

ICTRP	International Clinical Trials Registry Platform
iNMB	Incremental net monetary benefit
INR	International normalised ratio
ITC	Indirect treatment comparison
ITT	Intention-to-treat
IVRS	Interactive Voice Response System
KM	Kaplan-Meier
KSR	Kleijnen Systematic Reviews Ltd
KRAS	Kirsten rat sarcoma viral oncogene homologue
MAIC	Matching-adjusted indirect comparison
mCRC	Metastatic colorectal cancer
MD	Mean difference
MedDRA	Medical Dictionary for Regulatory Activities
MeSH	Medical subject headings
MJ	Matters of judgement
MSS	Microsatellite status
N	Number of patients
NA	Not applicable
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
NMA	Network meta-analysis
NR	Not reported
NYHA	New York Heart Association
Od	Once a day
ORR	Overall response rate
OS	Overall survival
PAS	Patient Access Scheme
PBO	Placebo
PD	Progressed disease
PF	Progression-free
PFS	Progression-free survival
PLD	Patient level data
PK	Pharmacokinetic
Po	Per os (oral)
PPS	Post-progression survival
PR	Partial response
PRESS	Peer Review of Electronic Search Strategies
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-analyses
PRO	Patient reported outcome
PSA	Probabilistic sensitivity analysis
PSM	Partitioned survival model
PSS	Personal Social Services
PSSRU	Personal Social Services Research Unit
PTT	Partial thromboplastin time
QALY	Quality-adjusted life year
QLQ-C30	Quality of Life Questionnaire
QoL	Quality of life
RCT	Randomised controlled trial
RDI	Relative dose intensity
RECIST	Response Evaluation Criteria in Solid Tumours
RMST	Restricted mean survival time
RR	Relative risk; risk ratio
RWE	Real-world evidence
SA	Sensitivity analysis

SAE	Serious adverse events
SAS	Safety analysis set
SD	Stable disease
SD	Standard deviation
SeTE	Standard error of treatment effect
SLR	Systematic literature review
SmPC	Summary of product characteristics
STA	Single Technology Appraisal
STM	State transition model
T/T	Trifluridine/tipiracil
TA	Technology Assessment
TAS-102	Trifluridine/tipiracil
TEAE	Treatment emergent adverse events
TKI	Tyrosine kinase inhibitor
ToT	Time on treatment
tpCR	Total pathological complete response
TRAEs	Treatment related adverse events
TSD	Technical Support Document
ULN	Upper limit of normal
UK	United Kingdom
UMC	University Medical Centre
US	United States
USA	United States of America
UTD	Unable to determine
VAS	Visual analogue scale
VEGF	Vascular endothelial growth factor
WHO	World Health Organization
WT	Wild type
WTP	Willingness-to-pay

Table of Contents

Key issue 1: Potential ambiguity of the population in the decision problem, which has implications for the comparators

The company continue to argue that the only comparator is trifluridine/tipiracil (T/T) and therefore BSC is not a comparator, although they also defend their inclusion of a cost effectiveness analysis vs. T/T: *"To 'close the loop', showing cost-effectiveness of regorafenib against BSC was necessary, if only from a triangulation perspective."* (p.4)

The basis of the argument that T/T should be the only comparator is that it is the only active treatment at 3rd line, having been recommended by NICE in TA405 and that it replaced BSC for patients fit enough for active treatment, thus consigning BSC as only appropriate for those not fit enough either at this line or post-T/T.

ERG comment:

As stated in the ERG report, unless the NICE recommendation is to include patients who might not get T/T e.g., those not fit enough to receive it and who might only be eligible for BSC, the population is those who otherwise would receive T/T.[1]

Key issue 2: Potential difference between subsequent treatment use in the RCTs and NHS clinical practice with unknown effect of subsequent treatment

The company have represented data on subsequent therapy use in the two regorafenib RCTs, CORRECT and CONCUR, adding similar data for the three T/T RCTs, RECOURSE, TERRA and Yoshino. They argue that any adjustment would favour regorafenib and be minimal because of higher subsequent therapy use in the BSC arm and minimal effectiveness of subsequent therapy respectively. They also state that the generalisability of the T/T would not be affected for the same reasons.

ERG comment:

The ERG agrees with the likely direction of effect of any adjustment for the regorafenib trials, and for two of the T/T trials given that there is no difference between the arms of the RECOURSE trial. The ERG disagrees that the main issue relating to the T/T trials is generalisability: it is potential bias in the treatment effect of T/T vs. BSC and thus, through any indirect treatment comparison (ITC), between regorafenib and T/T. However, it does appear that the difference between intervention and BSC arm is higher in the regorafenib trials (4% and 12%) than in the T/T trials (0%, 7.6%, 3%), which means that any adjustment is likely to increase the treatment effect more for regorafenib vs. BSC than T/T vs. BSC. This implies that adjusting for subsequent therapy would probably favour regorafenib by an amount that is uncertain.

Key issue 3: Lack of external validity of and comparability between the regorafenib and trifluridine/tipiracil RCTs

No new evidence was presented. The company continue to argue that they performed sufficient ITC sensitivity analyses to indicate that regorafenib and T/T are of comparable efficacy and that any uncertainties would only be resolved by a head-to head study in the UK.

ERG comment:

As stated in the ERG report, further sensitivity analyses for the ITC with different combinations of trials could be conducted, although considering the lack of difference between the combinations already examined this is unlikely to be particularly informative.[1]

Key issue 4: Difference in the treatment effect of regorafenib versus trifluridine/tipiracil depending on evidence source

The company argue that the ITC of RCTs provides the best evidence, and that the Nakashima study design is flawed given the selection of only patients for the comparison of the monotherapies regorafenib and T/T who did not crossover between regorafenib and T/T, lack of ECOG performance status (PS) data, and that the treatment effect estimated in that study is implausible high (hazard ratio (HR) 0.67 in the propensity score matched analysis). They also provided an additional study, Moriwaki 2018, also a retrospective observational study in Japan comparing regorafenib to T/T, which did not suffer the same study design problem and employed propensity score matching "*on ECOG"* PS. This study showed a HR that was close to 1 with a 95% confidence interval that overlapped 1.

ERG comment:

As stated in the ERG report, the ERG agrees with the company that there is a serious risk of bias with any observational study, including Nakashima 2020.[1] In addition, it is unclear how patients were selected for monotherapy and that, as stated in the ERG report, this might have been based on ECOG status with those with poorer prognosis not being found to be eligible for subsequent treatment. However, the problem with this argument is that it does not imply any kind of bias in the treatment effect between the two treatments since there is no reason to believe that patients with poorer prognosis were more likely to be selected for regorafenib than T/T monotherapy. Of course, it could be an issued of generalisability in that the treatment effect might favour T/T only with worse prognosis and/or higher ECOG PS. However, there was no evidence that this was the case, at least between PS 0 and 1 in the CORRECT or CONCUR studies, as shown in Figures 3.8 to 3.11. It is also worth noting in Nakashima 2020 that in the patients who crossed over, those who received T/T first also did better than those who received regorafenib first. The authors of the Nakashima 2020 also claimed that they *"…could not find any reasons why crossover could not be performed after monotherapy."* (p. e214) It is also important to note that in the Moriwaki 2018 study, crossover was permitted, and the rate was higher in the regorafenib group than in the TFTD group (60% vs. 40%; $p < .001$). Also, ECOG PS was found to be similar at the time of discontinuation of regorafenib and T/T.

In conclusion, the ERG is not convinced that the observational studies can be disregarded and, in particular, the Nakashima study, because of its comparison without crossover, relatively balanced baseline characteristics and size, should be considered. This is notwithstanding the limitations of Nakashima, but also those of the RCTs, particularly their comparability.

Key issue 5: Treatment effectiveness and extrapolation

The company argues that the assessment of progression on an 8-weekly basis (give or take a week) mirrors assessment in NHS clinical practice and hence used the KM data to inform PFS and ToT. When KM data was no longer available, the company extrapolated from the last data point onwards using the exponential function. The company further argues that using KM data is precisely because progression, in practice, is 'diagnosed' at intervals determined by clinic visits, and that subsequent treatment decisions are made at those same visits, that the KM curve best reflects clinical practice i.e. it preserves the stepwise nature of clinical practice. Finally, the company provided cost-effectiveness results using full parametrisation, acknowledging that if NHS practice has a different schedule for clinic assessment, then direct use of the KM data could be described as 'overfitting'.

ERG comment:

The ERG appreciates the provided full parametrisation results and acknowledges that the company's arguments for directly using KM data may be valid. However, in line with NICE DSU TSD 14, the ERG's perspective as described in the ERG report remains unchanged.[1]

Key issue 6: Adverse events

The company argues that including AE disutility separately in the economic model was a conservative approach because the impact of AEs on a patient's quality of life will be, to some extent, already 'captured' in the patients EQ5D responses and explicit addition of AE disutilities may therefore have an element of double-counting. Although the company expects the impact of grade 1 and 2 AEs on costs and quality of life to be negligible, they have endeavoured to incorporate grade 1 / 2 adverse events into the model to ascertain potential impact of cost-effectiveness.

Data on AEs was extracted from CORRECT, CONCUR, RECOURSE, Yoshino 2012 and TERRA. As there were different reporting criteria for AEs between the 5 publications, a common AE reporting definition was applied across the publications, i.e. the criteria used in the RECOURSE study were applied to the other 4 studies. The combined frequency of grade 1/2 AEs was then calculated by subtracting grade 3+ AEs from all AEs.

In the end, the company conducted two scenario analyses: 1) assuming a cost of £5 per AE (informed by the cost of laxatives (senna 7.5mg tab 60 = £2.03) and simple analgesics (paracetamol 500mg tab 100 = £2.34) without applying disutilities, and 2) assuming a cost of £5 per AE and a disutility of 0.01 per AE (taken as 10% of 0.1 i.e. the typical decrement for a grade 3+ AE in the economic model). Both scenario analyses had a minor impact on the cost-effectiveness results.

ERG comment:

The company fulfilled the ERG's request of providing more evidence (i.e. an updated economic model and scenario analysis including grade 1 and 2 AEs) to support the claim that grade 1 and 2 AEs would not have a meaningful impact on the cost-effectiveness results. Although the scenario analyses conducted by the company requires several assumptions, the ERG agrees that the impact of adding grade 1 and 2 AE costs and disutilities to the economic model is likely negligible.

Key issue 7: Resource use and costs

No compelling new arguments/evidence provided. Hence the ERG perspective as described in the ERG report remains unchanged.[1]

Key issue 8: Company's sensitivity analyses

The company provided analyses in response to request 1.) observed and modelled overall survival and progression-free survival using restricted mean survival time, and 3.) scenario analyses with the fully parametric models applied for PFS and time on treatment. The company did not provide analyses in response to request 2.) scenario analyses for the best supportive care comparison that were conducted for T/T, arguing that BSC was not a comparator.

ERG comment:

The ERG is satisfied with the additional analyses conducted by the company, which demonstrated that the observed trial data and modelled results for regorafenib are well aligned, and showed the impact of the company's scenario analyses conditional on using parametric survival models rather than KM data to model PFS and ToT. Some of the uncertainty remains unquantified for the comparison with BSC, as a substantial number of scenario analyses were not conducted for this comparator.

Key issue 9: Lack of a fixed random seed in model PSA

The company confirms that the small variability in results in different PSA runs is due to the random seed in the model not being fixed. A model with a fixed random number seed was provided to ensure the same sequence of random numbers are used for each probabilistic analysis.

ERG comment:

The ERG is satisfied with the company's response and confirms that the fixed random number seed was correctly implemented.

1. REFERENCES

[1] Armstrong N, Witlox W, Perry M, Ramaekers B, Danopoulos E, Sugden B, et al. *Regorafenib for treating metastatic colorectal cancer (ID4002): a Single Technology Assessment*. York: Kleijnen Systematic Reviews Ltd, 2022