in collaboration with:

ERASMUS UNIVERSITEIT ROTTERDAM
INSTITUTE OF HEALTH POLICY & MANAGEMENT

Maastricht University

Trifluridine in combination with tipiracil hydrochloride for previously treated metastatic colorectal cancer

Produced by Kleijnen Systematic Reviews (KSR) Ltd., United Kingdom (UK) in collaboration with Erasmus University Rotterdam (EUR) and Maastricht University, The Netherlands (NL)

Authors
Robert Wolff, Reviews Manager, KSR Ltd
Bram Ramaekers, Health Economist, Maastricht UMC
Anoukh van Giessen, Health Economist, Maastricht UMC
Xavier Pouwels, Health Economist, Maastricht UMC
Debra Fayter, Systematic Reviewer, KSR Ltd
Shona Lang, Systematic Reviewer, KSR Ltd
Nigel Armstrong, Health Economist, KSR Ltd
Gill Worthy, Statistician, KSR Ltd
Steven Duffy, Information Specialist, KSR Ltd
Manuela Joore, Health Economist, Maastricht UMC
Jos Kleijnen, Director, KSR Ltd, Professor of Systematic Reviews in Health Care, Maastricht University

Correspondence to Robert Wolff, Kleijnen Systematic Reviews Ltd
Unit 6, Escrick Business Park
Riccall Road, Escrick
York, UK
YO19 6FD

Date completed 09/05/2016

Source of funding: This report was commissioned by the NIHR HTA Programme as project number STA 15/121/11.

Declared competing interests of the authors
None.

Acknowledgements
None.

[redacted]

Rider on responsibility for report
The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:
Wolff R, Ramaekers BLT, Van Giessen A, Pouwels X, Fayter D, Lang S, Armstrong N, Worthy G, Duffy S, Joore MA, Kleijnen J. Trifluridine in combination with tipiracil hydrochloride for previously treated metastatic colorectal cancer: a Single Technology Assessment. York: Kleijnen Systematic Reviews Ltd, 2016.

Contributions of authors
Robert Wolff acted as project lead and systematic reviewer on this assessment, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Bram Ramaekers acted as health economic project lead, critiqued the company's economic evaluation and contributed to the writing of the report. Anoukh van Giessen, Xavier Pouwels and Nigel Armstrong acted as health economists on this assessment, critiqued the company's economic evaluation and contributed to the writing of the report. Debra Fayter and Shona Lang acted as systematic reviewers, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Gill Worthy acted as statistician, critiqued the analyses in the company's submission and contributed to the writing of the report. Steven Duffy 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

AE	Adverse event
AiC	Academic in confidence
AIC	Akaike information criterion
ARR	Absolute risk reduction
ASCO	American Society of Clinical Oncology
BRAF	Serine/threonine-protein kinase B-Raf
BSA	Body surface area
BSC	Best supportive care
CADTH	Canadian Agency for Drugs and Technologies in Health
CDF	Cancer Drugs Fund
CEA	Cost-effectiveness Analysis
CEAC	Cost effectiveness acceptability curve
CENTRAL	Cochrane Central Register of Controlled Trials
CER	Cost-effectiveness ratio
CHMP	Committee for Medicinal Products for Human Use
CI	Confidence interval
CiC	Commercial in confidence
CR	Complete response
CRC	Colorectal cancer
CRD	Centre for Reviews and Dissemination
CRO	Contract research organisation
CRUK	Cancer Research UK
CS	Company submission
CSR	Clinical study report
CT	Computed tomography
CHMP	Committee for Medicinal Products for Human Use
CTCAE	Common Terminology Criteria for Adverse Events
DARE	Database of Abstracts of Reviews of Effects
DCR	Disease control rate
DSU	Decision Support Unit
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
EED	Economic Evaluation Database
EGFR	Epidermal growth factor receptor
EMA	European Medicines Agency
EQ-5D	European Quality of Life-5 Dimensions
EREG	Epiregulin
ERG	Evidence Review Group
ESMO	European Society for Medical Oncology
EUR	Erasmus University Rotterdam
F	Female
FDA	Food and Drug Administration
FOBt	Faecal occult blood test
FOLFIRI	Chemotherapy combining folinic acid, fluorouracil and irinotecan
GP	General practitioner
HR	Hazard ratio
HRQoL	Health-related Quality of Life
HTA	Health Technology Assessment
HUI3	Health Utilities Index Mark III
ICER	Incremental Cost-effectiveness Ratio
ICTRP	International Clinical Trials Registry Platform
IRC	Independent review committee
ISPOR	International Society for Pharmacoeconomics and Outcomes Research

ITT	Intention-to-treat
IWRS	Interactive voice/web response system
JCOG	Japan Clinical Oncology Group
JSCO	Japan Society of Clinical Oncology
KRAS	Kirsten rat sarcoma viral oncogene homolog
KSR	Kleijnen Systematic Reviews
LY	Life year
LYG	Life years gained
M	Male
mCRC	Metastatic colorectal cancer
mg	Milligram
MRU	Medical resource utilisation
NA	Not applicable
NCIN	National Cancer Intelligence Network
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NL	The Netherlands
NR	Not reported
NS	Not specified
ORR	Overall response rate
OS	Overall survival
PAN	Panitumumab
PAS	Patient Access Scheme
PBO	Placebo
PD	Progressive disease
PenTAG	Peninsula Technology Assessment Group
PFS	Progression-free survival
PP	Post-progression
PPS	Post-progression survival
PR	Partial response
Pre-P	Pre-progression
PS	Performance status
PSA	Probabilistic Sensitivity Analyses
PSS	Personal Social Services
PSSRU	Personal Social Services Research Unit
QALY	Quality-adjusted Life Year
QQ	Quantile-quantile
RCT	Randomised Controlled Trial
RECIST	Response Evaluation Criteria In Solid Tumours
RFB	Regorafenib
SAE	Serious Adverse Events
SD	Stable disease
SD	Standard deviation
SE	Standard error
SIGN	Scottish Intercollegiate Guidelines Network
STA	Single Technology Appraisal
TA	Technology Appraisal
TPase	Thymidine phophorylase
T/T	Trifluridine/tipiracil
TTD	Time to treatment discontinuation
TTF	Time to treatment failure
UK	United Kingdom
VEGF	Vascular endothelial growth factor
WHO	World Health Organisation
WTP	Willingness-to-pay

Table of Contents

Table of Tables

Table of Figures

1 SUMMARY

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

The patient population described in the final scope is *"adults with metastatic colorectal cancer whose disease has progressed after standard therapies or for whom standard therapies are unsuitable"*. The final scope defined *"fixed dose combination of trifluridine and tipiracil hydrochloride"* as intervention and *"best supportive care"* as the comparator of interest. Outcomes of interest included *"overall survival, progression-free survival, response rates, adverse effects of treatment and health-related quality of life"*. The company did not offer any special considerations, including issues related to equity or equality. [redacted]

The decision problem in the company submission (CS) is in line with the final scope issued by the National Institute for Health and Care Excellence (NICE). Furthermore, the Evidence Review Group (ERG) noted that on 25 February 2016, a positive summary of opinion was issued by the European Medicines Agency (EMA). However, health-related quality of life (HRQoL) data were not collected in either of the two clinical trials presented in the CS.

1.2 Summary of clinical effectiveness evidence submitted by the company

The CS includes a systematic review of the available evidence for trifluridine/tipiracil (T/T) compared to best supportive care (BSC) for patients with advanced/metastatic colorectal cancer (mCRC) receiving treatment at the third line or beyond.

This review identified two randomised trials (phase II trial and RECOURSE). Both of these trials compared T/T to placebo with both treatment groups in the trials receiving BSC. The phase II trial included 172 participants from Japan while RECOURSE was a multinational trial including 800 participants. RECOURSE included 394 participants from Europe (nine from the United Kingdom (UK)). The company conducted analyses demonstrating that the effect of T/T did not vary according to geographical location and as a result, the trials were pooled.

Based on the pooled clinical trial results, there was an increase in median overall survival (OS) of 1.9 months (T/T: 7.3 months, BSC: 5.4 months). The pooled mean increase in OS was 2.3 months (T/T: 9.1 months, BSC: 6.8 months). Confidence intervals were not reported for the pooled analyses.

Regarding median progression-free survival (PFS), the pooled results showed an increase of 0.2 months (T/T: 1.9 months, BSC: 1.7 months). The mean PFS increase was 1.8 months (T/T: 3.7 months, BSC: 1.9 months). In the phase II trial no participant in either group had a complete response and one in the T/T group had a partial response. In RECOURSE one patient in the BSC group (placebo + BSC) had a complete response and eight in the T/T group had a partial response. A greater proportion of T/ T patients in both trials had stable disease (42.9% vs. 10.5% in the phase II trial and 42.4% vs. 15.9% in RECOURSE).

Two non-randomised trials were presented in the CS. The justification for including these was that the population was relevant to the decision problem. One study was a retrospective review of the outcomes of 55 patients with mCRC treated with T/T at a Japanese clinic. The other was a post-marketing surveillance survey presenting 370 AEs observed in 219 patients and was only reported as a poster.

No indirect or mixed treatment comparisons were presented in the CS.

The CS provides evidence from various sources to support that the submission fulfils end of life criteria. The first criterion of a short life expectancy includes the RECOURSE trial where survival was 7.7 months in the best supportive care arm. Evidence for the second criterion (an extension to life of at least three months compared to current National Health Service (NHS) treatment) is taken from the survival modelling calculations for the pooled estimate OS for both included trials (incremental survival: 3.2 months) and for RECOURSE alone (incremental survival: 3.0 months). The third criterion of a small patient population is taken from a survey of the number of patients in the UK with mCRC who would be treated at third line or beyond and from the company's estimates based on a previous technology assessment (approx. 2,600 patients) as well as expert opinion (2,490 patients).

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

The CS includes a systematic review of the available evidence for T/T compared to BSC for patients with mCRC receiving treatment at the third line or beyond. The literature searches reported in the CS were well documented and easily reproducible. A good range of databases were searched, and additional searches of conference proceedings were conducted. Searches were carried out in accordance with the NICE guide to the methods of technology appraisal Sections 5.2.2 and 5.2.4. The ERG is overall satisfied that the company identified and appraised the relevant randomised trials. The two non-randomised studies presented in the submission did not appear to have been selected systematically. We have focused our attention in this report on the two randomised trials which inform the cost effectiveness model. There is a lack of information on methods of pooling the two included randomised trials but overall it was considered acceptable from the point of view of clinical effectiveness that the trials were pooled.

The populations described in the NICE final scope, including patients with mCRC for whom standard therapies are 'unsuitable', seems approximately similar to the population described by the company, following the anticipated licence, but differs slightly from populations in the trials, which were used to inform the model. Consequently, following the licence it may be possible that patients not represented in the trials receive this medication. This includes patients *"for whom standard therapies are unsuitable"*. It remains unclear in which direction this discrepancy would influence the outcomes.

The phase II trial and RECOURSE, the two included trials identified by the company, were randomised and compared T/T to placebo with both treatment groups in the trials receiving BSC. The ERG confirmed the company's assessment that both trials were of high quality.

Following a request for clarification, the company stated that as there is no internationally accepted definition of BSC for clinical trials. Although both trials ensured consistency on medications excluded from BSC, the nature of BSC provided could vary between trial centres. The nature of BSC provided might also differ from that available in England and Wales.

RECOURSE was an international trial whereas the phase II trial was conducted solely with Japanese participants. The ERG considered that the company had provided evidence that geographical region was not a factor in effectiveness. This meant that results of the Japanese trial could be pooled with RECOURSE. However the ERG draws the committee's attention to the low proportion of UK participants in RECOURSE (9 of 800 participants). It is noted that 394 of 800 participants were from Europe. The ERG further notes that there is an under-representation of non-white, non-Asian populations across the trial (approximately 1% of RECOURSE participants are listed as 'black').

RECOURSE was powered for the primary outcome of OS so may not have had sufficient power to detect all differences between treatment groups for secondary outcomes. The included trials do not

directly assess HRQoL as specified in the NICE scope. Although there is a benefit to patients of the median increase in OS (1.9 months, pooled results) and PFS (0.2 months, pooled results), the quality of life experienced can only be inferred from effects of disease control and occurrence of adverse events.

In the phase II trial no patient in either group had a complete response and one in the T/T group had a partial response. In RECOURSE one patient in the BSC group had a complete response and eight in the T/T group had a partial response. A greater proportion of T/ T patients in both trials had stable disease (42.9% vs. 10.5% in the phase II trial and 42.4% vs. 15.9% in RECOURSE).

The occurrence of any adverse event was similar between T/T and BSC arms for both included trials. The Phase II trial found that serious adverse events [redacted]. In both trials 'treatment-related AEs' were found to be [redacted].

In both trials [redacted]. Nausea, vomiting, decreased appetite and diarrhoea were found to be [redacted] In both trials the following AEs related to myelosuppression were found to be [redacted],

In RECOURSE, more patients in the BSC arm were reported to [redacted]

It should be noted that in the RECOURSE trial all patients had to have received treatment with fluoropyrimidine, oxaliplatin, and irinotecan to be eligible. Patients were further required to have received prior chemotherapy with bevacizumab. However under NICE guidance patients in England would not be able to routinely receive bevacizumab prior to treatment with T/T. The company's interpretation in conjunction with clinical advice was that tumours in patients who had received fewer treatments were likely to be less resistant to additional therapy. This implies that the evidence for T/T presented might underestimate response in a UK population. This is an assumption, but it appears to be fair.

Regarding the CS fulfilling end of life criteria, the ERG believes that the first criterion (short life expectancy) has been met. For the second criterion (extension of life) to be met, NICE usually expects to see *"at least an additional 3 months, compared with current NHS treatment"*. As stated before, pooled estimates showed smaller differences in mean (OS: 2.3 months; PFS: 1.8 months) and median (OS: 1.9 months; PFS: 0.2 months) survival when comparing T/T to BSC (no confidence intervals available). The relevant population will be small but it should be highlighted that the figures presented might be an underestimate as they do not include Wales.

1.4 Summary of cost effectiveness submitted evidence by the company

The company developed a de novo cost effectiveness model to assess the cost effectiveness of T/T compared with BSC as third line or later treatment for patients with mCRC.

An Excel-based partitioned-survival model was constructed, consisting of the health states pre-progression, post-progression and death. Health states were selected according to the clinical pathway of care and comparable to the structure used in other late-stage cancer models. Because of the poor

prognosis of patients, a daily cycle length was applied to ensure the accuracy of survival estimates. The time horizon was 10 years effectively reflecting lifetime in this population.

In the company's base case combined data from the phase II trial and the RECOURSE trial were used to estimate OS and PFS for use in the model. PFS was also used as a proxy for time on treatment. Other parameters such as adverse events and T/T dosing were based on the RECOURSE trial only.

No HRQoL information was collected in the phase II trial or the RECOURSE study. The company conducted a systematic review to identify HRQoL studies from the published literature. In the company's base case, the health state utility values were the average of utilities obtained in the CORRECT study (identified in the systematic review) and the cetuximab NICE CS for the first line treatment of mCRC (not identified in the systematic review). Specific disutilities for adverse events were not incorporated in the model.

Categories considered for resource use and costs were: T/T costs, health state costs, post-progression treatment costs, end of life costs and adverse event costs. In the company's base case, T/T costs were calculated based on the body surface area (BSA), treatment delay and dose reductions obtained from the RECOURSE trial. Moreover, treatment delay was used to calculate the average treatment cycle length and hence also influenced pre- and post-progression medical resource utilisation (MRU). MRU included oral chemotherapy day case attendance, medical oncologist outpatient consultation, home consultation by general practitioners (GPs), community nurse specialist visit, health home visitor, district nurse visit and GP surgery visit. Post-progression treatment costs were calculated based on resource use from the RECOURSE trial. Costs of adverse events that are actively treated in the NHS are included. End-of-life care costs were taken from a published modelling study.

The company's base case incremental cost effectiveness ratio (deterministic, with PAS) was £44,032. One-way deterministic sensitivity analysis, scenario analyses and probabilistic sensitivity analyses were conducted. From the deterministic sensitivity analysis the company concluded that the most influential parameters on the model result were utility values for pre- and post-progression health states, the annual discount rate for quality-adjusted life years (QALYs) and the costs for post-progression treatment. Based on the scenario analyses, the most influential scenarios on the model results were the time horizon over which the costs and benefits of treatment are considered, and the choice of distribution from which efficacy data were fit to and extrapolated. The probabilistic sensitivity analyses indicated that at the PAS price, the probabilities of T/T being the most cost effective treatment are 0% and 77% for willingness-to-pay (WTP) thresholds of £30,000 and £50,000, respectively.

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

The economic model described in the CS is considered by the ERG to meet the NICE reference case to a reasonable extent. The ERG confirmed the company's finding that there was no existing cost-effectiveness model for T/T for the current indication. The ERG questions the sensitivity of the systematic review the company performed to identify HRQoL studies. No systematic reviews were performed for model structure and resource use, which should ideally have been performed, according to the NICE reference case.

The ERG agrees that the chosen model structure, daily cycle and the absence of a half-cycle correction are appropriate for this decision problem.

Even though pooling the effectiveness data from the RECOURSE trial and the phase II trial seems reasonable, the methods were not clearly described in the CS. After response to a clarification question

by the ERG, it appeared that individuals from both trials were naïvely combined in one dataset and compared with each other which could generate biased treatment effect estimates. In order for the ERG to assess the quality of pooling, the ERG would have liked to receive a comparison of the current meta-analysis (not stratified by trial) with a meta-analysis in which stratification by trials was performed. If the results of both meta-analyses would have been similar, the ERG would prefer the current meta-analysis to be used in the cost effectiveness model. Without this information, the ERG prefers using a more conservative assumption in its base case analysis by using RECOURSE data only. However, since there are no fundamental arguments which prevent the two trials from being pooled, besides the lack of clarity of the methodology, the ERG also presents its base case analysis based on the pooled effectiveness estimates from both trials.

Concerning the estimation of PFS and OS in the model, the ERG criticised using the Akaike information criterion (AIC) and not visual inspection of log-cumulative hazard plots to decide on using stratified or unstratified models. Based on inspection of log-cumulative hazard plots, the ERG considered it to be reasonable to use unstratified models instead of stratified models in its base case.

It was unclear to the ERG why only RECOURSE data (and not a pooled estimate from RECOURSE and the phase II trial) were used for AEs incidence rates, given that the company base case used pooled PFS and OS using evidence from both clinical trials. The ERG noted that the grade ≥3 AEs rates for the BSC arm reported in two tables of the CS and in the company's cost effectiveness model were not correct for the eight AEs. This was corrected in the ERG base case.

The ERG regards the company's arguments to estimate the health state utilities using an average of the utilities from TA176 and the CORRECT trial as incorrect or based on incorrect information. According to the ERG, the baseline utilities from the CORRECT study are the most plausible estimates for pre-progression, and the post-progression health state utilities, because it is the only study identified by the ERG in which utilities were measured using the European Quality of Life-5 Dimensions (EQ-5D) in a population that resembles the population in this appraisal (second to fourth line population with 74% ≥ third line). Therefore the ERG included utility values from the CORRECT study in its base case.

The ERG noted that the impact of AEs on HRQoL was not incorporated in the analyses, apart from the difference between the pre-progression health state utility values in the base case. Therefore, the ERG explored the estimation of a disutility for adverse events based on the occurrence of adverse events ≥ grade 3. This resulted in a disutility of 0.075 for T/T and 0.018 for BSC, calculated to one week the incremental disutility is -0.001. As these estimates do not include all AEs and heavily rely on assumptions, the ERG used a larger disutility for AEs of 0.01 per cycle for patients receiving T/T in its base case (similar assumption as in the company's base case but based on alternative justifications).

The company uses a parameterised distribution of BSA (log-normal) from RECOURSE to calculate T/T costs. The ERG notes that the population of the RECOURSE trial includes 33% of patients from Japan, which may be expected to have a lower BSA than the UK population. The CS reported that advisory board clinicians agreed with the use of a lower estimate of BSA as compared with the UK general population since mCRC patients would be expected to lose weight. According to the ERG, the non-parametrised distribution of BSA from RECOURSE is more reasonable estimate of BSA to calculate drug costs. As this most likely results in an underestimation of T/T costs, the BSA based on the UK population (which most likely results in an overestimation of T/T costs) is considered in an exploratory sensitivity analysis.

The ERG also noted that costs for adverse events were almost all estimated to equal a general medicine outpatient visit. The ERG thinks that this assumption is unrealistic and used alternative inputs in an explorative sensitivity analysis, retrieved from the NICE appraisal of TA370. Moreover, the ERG corrected the costs of a medical oncologist outpatient consultation. In addition, the ERG noted that the estimation of medical resource use was mainly based on expert opinion. Given the complete reliance on expert opinion for resource use, the ERG used an alternative source in an explorative sensitivity analysis.

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

1.6.1 Strengths

The company's submission contained a well-conducted systematic review which addressed the scope issued by NICE. Searches were carried out in line with the NICE guide to the methods of technology appraisal Sections 5.2.2 and 5.2.4. The CS and response to clarification provided sufficient details for the ERG to appraise the searches. The review identified two methodologically sound randomised controlled trials. The main trial, RECOURSE, was a large, multinational trial. The trials assessed the outcomes outlined by NICE with the exception of quality of life. Overall, the CS is well presented, transparent and in line with the final scope.

1.6.2 Weaknesses and areas of uncertainty

It should be noted that one of the outcomes defined in the scope (HRQoL) was not addressed in either of the included clinical trials (phase II trial and RECOURSE).

There is some uncertainty regarding the generalisability of the two trials as the phase II trial (172 participants) was conducted in Japan and RECOURSE (800 participants) included only nine participants from the UK (394 participants from Europe). However, analyses showed that the effect of T/T did not vary according to geographical location. Additionally, as the definition of BSC was unclear, i.e. there is currently no internationally accepted definition of BSC, it is unclear whether BSC considered in the evidence and hence in the model is representative for BSC in the UK.

The two trials included patients who had received prior chemotherapy with bevacizumab, a drug that is not included in relevant NICE guidance. It can be assumed that the evidence for T/T might underestimate response in a UK population which has received fewer treatments.

It is unclear whether all end of life criteria have been met. Some of the survival results reported in the CS do not show an improvement in life expectancy over three months when comparing T/T to BSC. Furthermore, the figures presented in support of a small patient population might be an underestimate of the relevant population.

The ERG believes incorrect search strategies for HRQoL were reported in the Appendix of the CS. The company response to the ERG clarification letter was that the reported search strategies were correct. However, the results reported in the CS suggest that separate HRQoL searches were conducted, and that four studies with HRQoL data met the inclusion criteria of the review. Without full details of the HRQoL search strategies the ERG was unable to assess their quality. The CS used unnecessary economic terms when searching NHS Economic Evaluation Database (EED; via the Cochrane Library).

Most uncertainty in the health economic model was related to the estimation of progression free survival and overall survival as well as the utility values. Additional uncertainties identified by the ERG included whether or not to use the naïve pooling provided by the company, averaging of utilities from various sources, estimation of resource use (mainly based on expert opinion) and estimation of BSA. Using mainly expert opinion for resource use (instead of empirical data) was considered by the ERG as one of

the main weaknesses. This uncertainty might have an impact on the ICER as examined in the exploratory sensitivity analyses.

1.7 Summary of exploratory and sensitivity analyses undertaken by the ERG

Compared with the company base case, the ICER increased by approximately £9,300 to £52,695 in the ERG base case (with PAS). This difference could largely be attributed to a reduction in incremental QALYs gained from 0.172 to 0.144. The difference between the results of the company and the ERG base case are mainly caused by the following changes in the model:

- Fixing errors with adverse events for BSC
- Use of RECOURSE data instead of pooled estimates
- Use of CORRECT utilities only, i.e. not averaging with utilities from the TA176 CS report.

The probability that T/T is cost effective is smaller in the ERG base case compared to the company's base case (0% versus 0% and 37% versus 77% for thresholds of £30,000 and £50,000, respectively).

Given that the pooled analyses might be preferred or might not differ substantially compared with more sophisticated pooling techniques, despite the lack of justification for the use of naïve pooling (i.e. not stratifying by trial), the ERG base case using the pooled evidence is presented as well. In these analyses, pooled evidence is used for OS, PFS, AE, BSA and dose reductions and resulted in an ICER of £49,392.

Exploratory sensitivity analyses illustrated that using the UK general population BSA estimates and an alternative source for resource use had a moderate impact on the ICER (£53,776 and £54,739 respectively). Subgroup analyses based on Kirsten rat sarcoma viral oncogene homolog (KRAS) status indicated that the ICER for the KRAS wild-type and KRAS mutant subgroups would be £53,042 and £50,721 respectively.

2 BACKGROUND

This report provides a review of the evidence submitted by Servier in support of Trifluridine/tipiracil (T/T; trade name Lonsurf®) for the treatment of metastatic colorectal cancer (mCRC) in patients whose disease has progressed after standard therapies or for whom standard therapies are unsuitable.[1]

The background section of the report by the Evidence Review Group (ERG) outlines and critiques the company's description of the underlying health problem and the company's overview of current service provision. The information is taken from Chapter 3 of the company submission (CS) with sections referenced as appropriate.[1]

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

The underlying health problem is mCRC described in the manuscript as *"disease that has spread beyond the large intestine and nearby lymph nodes"*. The company further states that *"this appraisal focuses on mCRC that is classified as Stage IV or Modified Dukes Stage D"* (Section 3.1.1 of the CS).[1]

The company highlights the role of Kirsten rat sarcoma viral oncogene homolog (KRAS) mutations which are *"generally thought to be a negative predictive marker for the treatment effect of an anti-EGFR monoclonal antibody"* (Section 3.1.1 of the CS).[1] They further state that *"KRAS should not directly affect the activity of T / T"*. To support this statement the company refers to the two main trials included in this submission and states that effectiveness has been shown in KRAS wild-type and KRAS mutant tumours.[2, 3]

The company describes the epidemiology of mCRC focusing on the incidence of mCRC in England (Section 3.1.2 of the CS).[1] Colorectal cancer (CRC) is described in the submission as the *"fourth most common cancer in the UK behind breast, lung and prostate cancer, accounting for 12% of all new cases"* (Section 3.1.2 of the CS).[4] The company notes that 26% of patients present with metastatic disease.[5]

The company states that *"approximately 55% of patients initially diagnosed with colorectal cancer Stage II or III who receive initial treatment will ultimately progress to metastatic disease"* (Section 3.1.2 of the CS).[1]

The impact of colorectal cancer on patients, carers and society is briefly considered (Section 3.2 of the CS). The company states that *"psychological distress is common in patients with CRC, with depression and anxiety being particularly common; this is exacerbated further for patients who have a stoma following surgery for their condition"*.[1] Furthermore, the company states that the main aims of treatment for mCRC are *"to relieve symptoms and to improve health-related quality of life (HRQL) and survival"*.[1]

Section 3.4 of the CS describes the life expectancy of patients with mCRC and provides estimates of the number of patients at each line of therapy.[1] The company states that *"trifluridine / tipiracil is licensed for patients who have already received standard recommended treatment for mCRC and are therefore likely to be receiving therapy at third line or later. At this stage of the disease, life expectancy is approximately 6 months"* (Section 3.4.1 of the CS).[1]

The company provides survival data based on a UK source.[6] According to Section 3.4.1 of the CS, *"one year survival is lowest for those diagnosed with stage IV disease (40% for men and 33% for women). In addition, the survival of patients with mCRC decreases with each line of therapy. Five year survival for patients with mCRC is 7% and 8% for men and women, respectively"*.[1]

ERG comment: The ERG considers the company's description of the aetiology and pathology of metastatic colorectal cancer to be appropriate. Descriptions of the disease are taken from National Institute for Health and Care Excellence (NICE) guidance. [CS references 24 and 25]. The clarification of the staging that comprises mCRC gives a more precise definition of the underlying health problem.

The reference on incidence of colorectal cancer supplied by the company was checked and found to be correctly cited and from a reputable source.[4] The reference supporting the statement that 26% of patients present with metastatic disease was found to be a broken web link. The web site is a reputable source (National Cancer Intelligence Network, NCIN) but the provenance of the figure could not be determined. The ERG notes that the CS does not include Wales in its estimates of the annual number of patients with mCRC which has implications for the budgetary impact.

The estimate regarding patients progressing to metastatic disease (*"approximately 55% of patients initially diagnosed with colorectal cancer Stage II or III who receive initial treatment will ultimately progress to metastatic disease")* was taken from a previous technology appraisal and was therefore considered to be reliable.[7]

The ERG considers the statement on the impact of colorectal cancer on patients, carers and society to be appropriate. The statement on the main aims of treatment of mCRC is based on a NICE guideline is therefore considered to be appropriate.[8]

The statement regarding the life expectancy of patients with mCRC receiving treatment at third line or later includes both of the randomised trials in the submission and appears to be appropriate.[2, 3]

The ERG identified an apparent discrepancy in survival between the data presented in Section 3.4.1 of the company submission and the survival in the RECOURSE trial. In particular, one year survival for patients with mCRC was presented as 40% and 33% for men and women, respectively, based on a UK data source.[1] The estimated one year survival in the BSC arm of the RECOURSE trial was 17.6% (Table 25 of the CS) which suggests that the survival in the trial is much lower.[1] The company was asked to explain this apparent discrepancy. In the response to request for clarification, the company stated that the Cancer Research UK (CRUK) data presented reflect all patients with mCRC irrespective of time since diagnosis of metastatic disease, number of lines of chemotherapy received etc.[9] Therefore the CRUK data are not reflective of the population in the decision problem of this appraisal (patients who have received two or more lines of chemotherapy).

The effectiveness of T/T in regard to KRAS mutations will be discussed in Section 3 of the ERG's report.

2.2 Critique of company's overview of current service provision

The company states that "*there are currently no recommended therapeutic options for patients who have failed second-line treatment*".[1]

According to the CS, *"clinical experts at the recent advisory board highlighted that trifluridine / tipiracil would be a preferred option to regorafenib based on tolerability"*.[1]

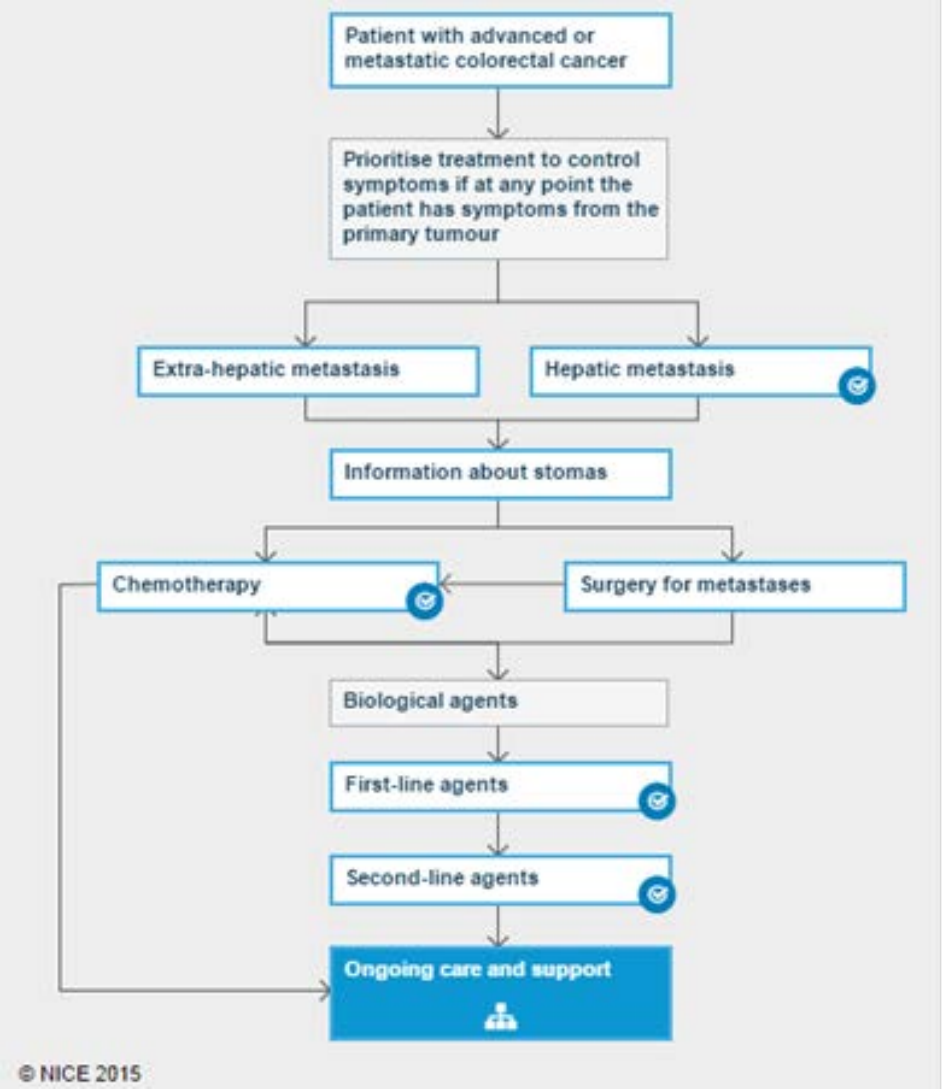
The company provides estimates of the number of mCRC patients at each line of therapy using a previous technology assessment as a basis[10] and adapted using clinical opinion (Section 3.4.2 of the CS).

Figure 8 in Section 3.4.2 of the CS provides an estimate of the number of patients with mCRC by treatment option. The company states that *"trifluridine / tipiracil would fit into the treatment pathway*

at third line or beyond. It is estimated that at this stage there would be approximately 2600 patients who may be eligible for and are motivated to receive further treatment".

The company's overview of the current clinical pathway for patients with metastatic colorectal cancer is given below. According to the CS, *"trifluridine/tipiracil provides a therapeutic option for patients with tumours that have progressed following second-line treatment and who are well enough and motivated to receive further therapeutic intervention"*.[1]

Figure 2.1: NICE clinical pathway for patients with metastatic colorectal cancer
(Based on figure 3 of the CS[1])



ERG comment: The company's description of the pathway is taken from NICE guidance which is appropriate and relevant to the decision problem.[11]

The ERG agrees with the company that *"there are currently no recommended options for patients who have failed second-line treatment"*. This is correct as regorafenib is licensed in the UK for the treatment of mCRC, however, it is not recommended by NICE due to a non-submission (TA334 – terminated appraisal). The ERG notes (as is outlined by the company) that options may be provided for patients such as repeating a previous regimen, enrolling on a clinical trial or using mitomycin C + 5FU or

capecitabine.[12] However, it should be noted that the statement that T/T *"would be a preferred option to regorafenib based on tolerability"* is based on clinical opinion alone.[13]

The ERG notes that estimates of the number of patients with mCRC by treatment option based partially on clinical opinion may be unreliable. The ERG further notes that the estimates appear to be based on England only and do not include Wales.

3 Critique of company's definition of decision problem

The company presents its response to the decision problem in Section 1.1 of the CS. This is reproduced below.

Table 3.1: Summary of the decision problem
(Based on Table 1 of the CS[1])

	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 whose disease has progressed after standard therapies or for whom standard therapies are unsuitable	Final scope	
Intervention	Fixed dose combination of trifluridine and tipiracil hydrochloride	Final scope	
Comparator(s)	Best supportive care	Final scope	
Outcomes	• overall survival • progression-free survival • response rates • adverse effects of treatment • health-related quality of life.	• overall survival • progression-free survival • response rates • adverse effects of treatment	Trifluridine/tipiracil was in-licensed by Servier Laboratories Ltd from Taiho Pharmaceutical. Health-related quality of life data were not collected in the phase III clinical trial
Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. 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 an NHS and Personal Social Services perspective.	Final Scope. The economic analysis will be presented as reported in the final scope (December 2015) and in accordance with the NICE guide to the methods of technology appraisal (2013).	

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
Subgroups to be considered	None specified		
Special considerations including issues related to equity or equality	No special considerations, including issues related to equity or equality have been identified.		
NHS = National Health Service; NICE = National Institute for Health and Care Excellence			

3.1 Population

The patient population described in the final scope is *"adults with metastatic colorectal cancer whose disease has progressed after standard therapies or for whom standard therapies are unsuitable"*.[14]

ERG comment: The definition of the relevant population addressed in the CS is in line with the decision problem described by NICE. However, it is noteworthy to highlight some points:

- The main clinical evidence submitted by the company, the RECOURSE trial, does not include participants for whom standard therapies are unsuitable.[2] All patients had to have received treatment with fluoropyrimidine, oxaliplatin, and irinotecan to be eligible. This includes those who were refractory to treatment (disease progressed) and those who were intolerant (treatment discontinued due to toxicity or could not be re-administered for medical reasons). Furthermore, participants of the RECOURSE trial were required to have received prior chemotherapy with bevacizumab. However, under NICE guidance patients in England and Wales would not be able to routinely receive bevacizumab prior to treatment with trifluridine/tipiracil.[1]
- The company's interpretation in conjunction with clinical advice was that tumours in patients who had received fewer treatments were likely to be less resistant to additional therapy.[13] This implies that the evidence for T/T presented in the CS might underestimate response in a UK population. This is an assumption, but it appears to be fair.
- According to Table 15 of the CS, all participants of the included phase II randomised controlled trial (RCT) were recruited in Japan whereas participants of RECOURSE were from Japan, Europe, USA and Australia.[1] Potential implications for the generalisability of the trial results for patients in the UK are discussed in Section 4.2 of this report.
- In Section 1.1 of the CS, it is stated that, if approved, T/T offers an option for those patients who are *"well enough and motivated to receive further treatment"*.[1] This statement is not further explained. Section 6.2 of the CS considers the projected uptake of T/T and states that 20% of the eligible population might receive treatment in the first year of availability before reaching a steady state of approximately 40% by year three of availability.[1] These estimates appear to be based solely on clinical opinion and it is unclear how this has been elicited.
- Trial participants appeared to reflect those seen in clinical practice. Both trials include male and female participants and patients with colon and rectum cancer. Both included participants with KRAS wild-type and mutation positive status. In RECOURSE 79% of patients had been diagnosed with metastatic cancer for 18 months or more. Sixty-one per cent had received at least four prior treatment regimens.[1]

- Across the trials there is an under-representation of non-white, non-Asian populations. In RECOURSE nine patients (1%) are listed as 'black'. Although there is no evidence of any differential effects of the drug based on ethnicity, this aspect is drawn to the attention of the committee.

3.2 Intervention

The intervention is trifluridine/tipiracil. Section 2.1.4 of the CS states that *"trifluridine/tipiracil is comprised of an antineoplastic thymidine-based nucleoside analogue, trifluridine, and a thymidine phosphorylase (TPase) inhibitor, tipiracil hydrochloride, at a molar ratio 1:0.5 (weight ratio, 1:0.471)"*.[1]

According to the CS, a positive Committee for Medicinal Products for Human Use (CHMP) opinion for Lonsurf® was expected in late February 2016, with marketing authorisation in May 2016 (Section 2.2.4 of the CS).[1] The company notes that *"trifluridine/tipiracil is licensed in Japan and the US and up to December 2015 had been received by over 12,000 patients"* (Section 2.2.6 of the CS).[15]

The company stated that trifluridine/tipiracil is marketed as an oral tablet with dosing based on body surface area at a recommended starting dose of 35mg/m^2 followed by individual adjustments for safety and tolerability. An average course of treatment is 28 days with management in secondary care either as a chemotherapy day case or outpatient setting (Sections 2.3.1 and 2.4.1 of the CS).[1]

ERG comment: The CS reflects the scope which is a *"fixed-dose combination of trifluridine and tipiracil hydrochloride"*.[14]

The ERG identified that on 25 February 2016, the European Medicines Agency (EMA) issued a positive summary of opinion outlining the full indication: *"Lonsurf is indicated for the treatment of adult patients with metastatic colorectal cancer (CRC) who have been previously treated with, or are not considered candidates for, available therapies including fluoropyrimidine-, oxaliplatin- and irinotecan-based chemotherapies, anti-VEGF agents, and anti-EGFR agents. It is proposed that Lonsurf be prescribed by physicians experienced in the administration of anticancer therapy"*.[16] The number of patients receiving T/T is taken from an internal communication by the company.[15]

The included trials had a 35mg/m^2 dosage. The phase II trial allowed a reduction of 10 mg/day if necessary and RECOURSE allowed a maximum of three reductions in dose in decrements of 5 mg/m^2 (Table 15 of the CS).[1]

3.3 Comparators

The comparator is best supportive care (BSC). The scope issued by NICE recommended BSC as there are no currently recommended treatments for patients who have failed second line treatment.

For the phase II trial, *"all necessary support was provided to patients, with the exception of concomitant use of other anti-cancer drugs or other investigational drugs"*.[9] In RECOURSE, *"all necessary support was provided to patients which included permitted concomitant medications and therapies and study medication"*.[9] Specifically patients were *"not to receive other investigational anti-tumour agents or antineoplastic chemotherapy, hormonal therapy or immunotherapy. Palliative radiotherapy was not permitted while the patient was receiving study treatment"*.[9]

ERG comment: The CS is based on two placebo-controlled trials where both treatment and placebo groups received BSC. The ERG asked for clarification on the definitions of BSC used in the included trials, the guidance regarding BSC given to the centres involved in the included trials and the

applicability of the BSC to the UK setting. In their response to the request for clarification[9], the company stated that *"there is currently no internationally accepted definition of BSC for clinical trials"*. Although both trials ensured consistency on medications excluded from BSC, the nature of BSC provided could vary between trial centres. The nature of BSC provided might also differ from that available in England and Wales.

The ERG notes that, according to the CS[1], in order to obtain a positive opinion of the CHMP, the company provided additional information in the submission including a comparison to regorafenib.[1] *"Regorafenib is not recommended by NICE due to a non-submission"* and this comparison does not form part of the final scope for this CS.

3.4 Outcomes

Outcomes of interest are overall survival, progression-free survival, response rates, adverse effects of treatment and health-related quality of life.[14]

ERG comment: The two RCTs included in the clinical effectiveness part of the CS did not collect quality of life data.[2, 3] Data to populate the economic model will be discussed in the cost effectiveness section.

3.5 Other relevant factors

The company did not offer any special considerations, including issues related to equity or equality. [redacted] (Section 2.3.2 of the CS).

4 CLINICAL EFFECTIVENESS

4.1 Critique of the methods of review(s)

The company stated in Section 4.1 of the CS that *"a systematic review was conducted to retrieve relevant clinical data from the published literature regarding the efficacy and safety of trifluridine / tipiracil compared with best supportive care (BSC) for patients with advanced / metastatic colorectal cancer receiving treatment at the third line or beyond"*.[1]

ERG comment: The systematic review will be critiqued in this section of the report. It should be noted that the evidence presented in the CS compared trifluridine/tipiracil in combination with best supportive care (T/T arm) to placebo in combination with BSC (BSC arm).

4.1.1 Searches

The literature searches reported in the CS were well documented and easily reproducible. A good range of databases were searched, and additional searches of conference proceedings were conducted. Searches were carried out in accordance with the NICE guide to the methods of technology appraisal Sections 5.2.2 and 5.2.4.[17]

Description and critique of the company's search strategies

The Canadian Agency for Drugs and Technologies in Health (CADTH) evidence based checklist for the Peer Review of Electronic Search Strategies, was used to inform this critique.[18] The submission was checked against the Single Technology Appraisal (STA) specification for company/sponsor submission of evidence.[19] The ERG has presented only the major limitations of each search strategy in the main report. Further criticisms of each search strategy can be found in Appendix 1.

Clinical effectiveness

The CS states that a systematic review was conducted to retrieve relevant clinical data from the published literature regarding the efficacy and safety of trifluridine/tipiracil compared with BSC for patients with advanced/metastatic colorectal cancer receiving treatment at the third line or beyond.

Searches were conducted on 26 October 2015 in MEDLINE, MEDLINE In-Process and Other Non-Indexed Citations, Embase and the Cochrane Library (Cochrane Central Register of Controlled Trials (CENTRAL), the Cochrane Database of Systematic Reviews, the NHS Economic Evaluation Database (NHS EED), the Database of Abstracts of Reviews of Effects (DARE), and the Health Technology Assessment Database (HTA). The host provider for each database was listed; the date span of the databases searched and the specific date the searches were conducted were provided. The company additionally searched conference proceedings: American Society of Clinical Oncology (ASCO), European Society for Medical Oncology (ESMO), and International Society for Pharmacoeconomics and Outcomes Research (ISPOR). Detailed search strategies for the database searches were reported in Appendix 3. The CS did not provide full details of the conference proceedings searches. Full details of the conference proceedings searches for the utility review were provided in response[9] to the ERG request for clarification letter.[20] These searches could have been used for the clinical effectiveness review, as generic search terms for advanced and metastatic colorectal cancer were used, but it is not clear if they were.

The company translated the research question into appropriate search strategies and the ERG considered the searches to be satisfactory. Searches were clearly structured and divided into population and intervention/comparator facets, using an appropriate combination of index terms, free text and synonyms for the interventions and comparators. The search strategies included Boolean, truncation and proximity

operators. No date or language limits were used. Study design limits to identify RCTs and non-RCTs were applied. The study design filters were not referenced, so it was unclear whether the filters used were published objectively derived filters. However, the search filters appeared to be those designed by and available from the website of the Scottish Intercollegiate Guidelines Network (SIGN).[21]

The search strategies included all currently available comparators alongside the intervention, though only BSC was considered in the NICE scope. Including the comparators in the search strategy would not have affected the search results, i.e. more records were retrieved, without missing relevant T/T studies.

It is possible that the facet of search terms for 'advanced/metastatic' included in the search strategies was too restrictive, and that combining the metastatic colorectal cancer facet with T/T and the study design filters would have been sufficient.

Searches of conference proceedings were conducted. The CS reports the names of the conferences searched and which years (2013-15) in the appendix, but does not give specific details about the search methods used and exact dates searched. The CS reports that no studies were identified from the conference searches, although three conference abstracts were included (Table 14 of the CS).[1] The three conference proceedings searched were: ASCO, ESMO, and ISPOR.

A search of trials registers, such as ClinicalTrials.gov and WHO International Clinical Trials Registry Platform (ICTRP), for unpublished and ongoing trials would have been a useful addition to the literature searches.

Indirect and mixed treatment comparisons
No searches were conducted.

Non-randomised and non-controlled evidence
The same search strategies and databases used for the clinical effectiveness literature searches were used to identify non-RCT evidence. The search strategies included a study design filter for non-RCTs.

Adverse events
The same search strategies and databases used for the clinical effectiveness literature searches were used to identify adverse events data. Guidance by the Centre for Reviews and Dissemination[22] recommends that if searches have been limited by a study design filter, additional searches should be undertaken to ensure that adverse events that are long-term, rare or unanticipated are not missed. Despite the inclusion of a non-RCT search filter the ERG considered that it was possible that some relevant evidence may not have been identified as a consequence of the study design limits. Safety data were taken directly from the company's two trials (RECOURSE[2] and phase II trial[3]).

4.1.2 Inclusion criteria

Section 4.1.2 and Appendix 3 of the submission describe the methods used to select studies for inclusion in the review. The company states that *"identified studies were independently assessed by two reviewers in order to ascertain whether they met the pre-defined inclusion/exclusion criteria, and any discrepancies were resolved by a third reviewer"*.[1]

The inclusion criteria of the review are given in Table 4.1 below.

Table 4.1: Eligibility criteria used in search strategy
(Based on Table 13 of the CS[1])

	Inclusion criteria	Exclusion criteria	Comments
Population	Adult patients with advanced/ mCRC receiving treatment at third line or beyond	Patients receiving treatment at first or second line	According to NICE scope
Interventions	Trifluridine/tipiracil	-	According to NICE scope
Comparators	BSC	-	Searches were conducted to identify studies investigating all currently available comparators for trifluridine/tipiracil (to support HTA submissions in other territories); however, comparators considered relevant for the current STA were restricted to BSC according to the NICE scope†
Outcomes	Efficacy: Overall survival 1-year survival rate Progression-free survival Time to progression Response rates (complete response, partial response, stable disease) Objective response rate Disease control rate Safety: All-grade AEs of interest Grade 3 or 4 AEs of interest HRQoL	-	-
Study design	RCTs with no restriction on phase or blinding	Non-randomised, observational studies	-
Language restrictions	No restriction	-	-

† Screening of publications by title and abstract was performed to include all currently available treatments; any studies that were not relevant according to the NICE scope were then excluded upon full publication review.
AE = adverse event; BSC = best supportive care; CS = company submission; mCRC = metastatic colorectal cancer; HRQoL = health-related quality of life; HTA = health technology assessment; NICE = National Institute for Health and Care Excellence; RCT = randomised controlled trial, STA = single technology assessment

ERG comment: The methods used to select studies for the review appear to be appropriate.

The inclusion criteria for the review population are more specific than that given in the NICE scope. The final scope[14] states that the population of interest is *"adults with mCRC whose disease has*

progressed after standard therapies or for whom standard therapies are unsuitable" whereas the inclusion criteria in the CS[1] are for *"adult patients with advanced/mCRC receiving treatment at third line or beyond"*.

The CS does not provide a definition of best supportive care.[14] Following a request for clarification, the company stated that as there is no internationally accepted definition of BSC for clinical trials.[9]

A range of relevant outcomes are included in the review which includes those specified in the final scope.[14]

The review has no restrictions on study eligibility based on language which is appropriate given the multinational nature of the trials.

4.1.3 Critique of data extraction

The company states that *"relevant information was extracted into the Single Technology Appraisal (STA) template by a reviewer. A second reviewer checked the data extraction, and any inconsistencies were resolved through discussion"*.[1]

ERG comment: The methods used to extract data for the review appear to be appropriate.

4.1.4 Quality assessment

No specific mention is made in the manuscript of the involvement of two reviewers in the assessment of the quality of studies included in the review.[1]

ERG comment: It is reasonable to assume that two reviewers were involved in the assessment of the quality of the included studies given the reporting of the systematic review methods for data extraction.

4.1.5 Evidence synthesis

The company states in Section 4.9 that *"a pooled analysis using individual patient data was conducted for the Phase II and RECOURSE trials, examining OS and PFS"*.[1]

ERG comment: Justification for pooling the two included trials and a full explanation of pooling methods was not provided in the company submission.[1] The company was asked to clarify this.[20] In their response, the company stated that *"both trials were conducted in a patient population that is relevant to the decision problem for the appraisal and are consistent with the proposed marketing authorisation. (...) ...there is no evidence of a difference in efficacy based on ethnicity"*.[9] This statement was supported by a reference to a pre-specified geographic regional subgroup analysis which showed no significant differences between geographic regions in overall survival (OS) and progression-free survival (PFS). The ERG is satisfied that pooling the two trials for the clinical effectiveness section of the CS is acceptable given similarities of design, disease characteristics, intervention and outcomes. However, due to a lack of information about the statistical methods used to pool the two trials as well as any measure or test of statistical heterogeneity the ERG cannot fully comment on the statistical pooling. The forest plot provided for OS and PFS does show that the trial results appeared to be homogenous, and the pooled results are in line with the individual trial results, so it seems that the pooling was appropriate.

4.2 Critique of trials of the technology of interest, their analysis and interpretation (and any standard meta-analyses of these)

The company states in Section 4.2 that *"the systematic review of clinical evidence identified two unique RCTs of trifluridine / tipiracil versus BSC in the population of interest to this submission. (...) ...In addition, three linked abstracts were identified"*.[1]

According to the CS, 193 studies were excluded after consulting the full papers (Figure 10 of the CS).[1] Bibliographic details and reasons for exclusion were listed in Appendix 3.6 of the CS.[23]

The company identified an ongoing trial (TERRA), a study in Chinese and south East Asian patients. They stated that the trial was due for completion at the end of 2015 with a clinical study report (CSR) estimated to be available in Summer 2016.[1] The company was asked to clarify that no results were available or to provide any results.[20] In their response, the company stated that no data were currently available for this trial and that the CSR was expected in July 2016.[9]

Section 4.11 of the CS provided details and results of two non-randomised studies.[1] The company was asked to clarify how these studies were identified and selected for inclusion in the CS as the inclusion criteria for the review specified only RCTs.[20] The company replied that *"these studies were not identified via a specific search, however Servier were aware that they had been presented and as they are relevant to the decision problem it was decided to present them in section 4.11 of the company evidence submission"*.[9]

According to the CS, *"the Phase II study was the primary licensing study for trifluridine/tipiracil in Japan. It involved 172 refractory mCRC patients who had previously been treated with, or were not candidates for available therapies (Fluoropyrimidine, oxaliplatin and irinotecan). The pivotal study for trifluridine/tipiracil is the RECOURSE trial, which studied 800 end-stage mCRC patients. These patients were all refractory or intolerant to all available therapies. The results of these studies have allowed for a successful marketing authorisation application in Japan and the US and are the basis for the application within the EU"*.[1]

A comparison of the population, intervention, comparator, outcomes and study designs for the two trials is given in Table 4.2. Information to populate the table was taken from Tables 14 and 15 of the company's submission.[1]

Table 4.2: Comparison of population, intervention, comparator, outcomes and study design
(Based on Tables 14 and 15 of the CS[1])

Trial no. (acronym)	Not reported (Phase II trial, no acronym)	NCT01607957 (RECOURSE)
Population	Adult patients aged ≥20 years with histologically or cytologically confirmed unresectable metastatic colorectal adenocarcinoma with a previous treatment history of ≥2 regimens of standard chemotherapy	Adult patients aged ≥18 years with biopsy-documented adenocarcinoma of the colon or rectum who had received ≥2 prior regimens of standard chemotherapy
Intervention	Trifluridine/tipiracil + BSC	
Comparator	Placebo + BSC	
Primary Outcome	Overall survival (OS)	
Secondary Outcomes	• Progression-free survival (PFS) • Time to treatment failure (TTF) • Disease control rate (DCR)	
	• Response rate • Duration of response • Efficacy of trifluridine/tipiracil in patients with or without KRAS mutations • Adverse event profile and tolerability	• Overall response rate (ORR) • Duration of Response • Subgroup analysis by KRAS status on OS and PFS • Safety and tolerability

Trial no. (acronym)	Not reported (Phase II trial, no acronym)	NCT01607957 (RECOURSE)
Trial Design	Multi-centre, double blind, randomised (in a 2:1 ratio), placebo controlled trial	Multi-centre, double blind, randomised (in a 2:1 ratio), placebo controlled trial
BSC = best supportive care; CS = company submission; DCR = disease control rate; KRAS = Kirsten rat sarcoma viral oncogene homolog; ORR = overall response rate; OS = overall survival; PFS = progression-free survival; RCT = randomised controlled trial; TTF = time to treatment failure		

Table 4.3 provides more detail on the methodology of the two trials while Table 4.4 presents the outcome definitions used in these trials. Characteristics of participants in the two RCTs are presented in Table 4.5.

Table 4.3: Methodology of included RCTs
(Based on Table 15 of the CS[1] and CSRs[24, 25])

	Phase II trial	RECOURSE
Location	Japan	Australia, Europe, Japan, United States
Trial Design	Multi-centre, double blind, randomised (in a 2:1 ratio), placebo controlled trial	
Eligibility criteria for participants	• Previous treatment with ≥2 regimens of standard chemotherapy • Adequate bone marrow, hepatic and renal function within 7 days of enrolment	
	• ≥20 years old • ECOG PS 0-2 • Histologically or cytologically confirmed unresectable metastatic colorectal adenocarcinoma • Refractory or intolerant to a fluoropyrimidine, irinotecan, oxaliplatin • Measurable lesions as per the RECIST	• ≥18 years old • ECOG PS 0-1 • Biopsy documented adenocarcinoma of the colon or rectum • Patients were also required to have received chemotherapy with each of the following agents: fluoropyrimidine, oxaliplatin, irinotecan, bevacizumab, cetuximab or panitumumab if KRAS wild-type
Setting	Secondary care oncology, gastroenterology or general medicine outpatient departments	
Trial drugs	• 35 mg/m^2 T/T taken orally after morning and evening meals • 2 tablet doses were used in order to achieve the correct dose • T/T was taken in a 28-day cycle; a 2-week cycle of 5 days of treatment followed by a 2-day rest period and then a 14-day rest period • Placebo was matched to T/T tablets for taste, colour and size • Treatment continued until tumour progression, unacceptable toxic effects, or withdrawal of consent • No cross-over between groups after progression or toxic effects	
	• In patients who had AEs, the dose could be reduced by 10 mg/day as judged necessary • Except in cases when deemed necessary from the perspective of safety or ethics, such as the treatment of an AE, other anti-cancer drugs or other investigational drugs were not to be used concomitantly.	• Protocol allowed for a maximum of three reductions in dose in decrements of 5 mg/m^2 • Other than BSC, permitted concomitant medications and therapies and study medication, patients were not permitted to receive any other medications and therapies, including other anticancer therapies, such as chemotherapy, immunotherapy, biological response modifiers or endocrine therapy, during the study treatment period.

	Phase II trial	RECOURSE
		• Palliative radiotherapy was not permitted while the patient was receiving study treatment.
Primary Outcome	• Overall survival (OS)	
Secondary Outcomes	• Progression-free survival (PFS) • Time to treatment failure (TTF) • Disease control rate (DCR) • Duration of response	
	• Response rate • Efficacy of trifluridine/tipiracil in patients with or without KRAS mutations • Adverse event profile and tolerability	• Overall response rate (ORR) • Subgroup analysis by KRAS status on OS and PFS • Safety and tolerability
Pre-planned subgroups	• Sex (male / female) • Age (<65 years / ≥65 years) • Primary site (colon / rectum)	

	• PS (0 / 1-2)
	• Number of metastatic groups (1 / 2 / 3 / ≥4)
	  • Liver metastasis
	  • Lung metastasis
	  • Lymph node metastasis
	  • Peritoneum metastasis
	• Previous treatment
	  • Previous surgery
	  • Adjuvant chemotherapy
	  • Palliative chemotherapy
	  • Bevacizumab
	  • Cetuximab
	• KRAS mutation status
		• KRAS mutation status
		• Time since diagnosis
		• PS (0 / 1)
		• Geographic region (Japan / Rest of World
		• Number of metastatic sites
		• Number of prior regimens

AE = adverse event; BSC = best supportive care; CS = company submission; CSR = clinical study report; DCR = disease control rate; ECOG = Eastern Cooperative Oncology Group; KRAS = Kirsten rat sarcoma viral oncogene homolog; ORR = overall response rate; OS = overall survival; PFS = progression-free survival; PS = performance status; RECIST = Response Evaluation Criteria in Solid Tumours; T/T = trifluridine/tipiracil; TTF = time to treatment failure.

Table 4.4: Definition of relevant outcomes in the included RCTs

(Based on Table 15 of the CS[1])

	Phase II trial	RECOURSE
Overall survival	Time between randomisation and death from any cause or the date of last follow-up	Time (in months) between randomisation and death from any cause.
Progression-free survival	Defined as the time (in months) from randomisation to the date that the patient's condition reached progressive disease (PD). If the patient died before reaching PD, the date of death was considered the date PD was reached. For patients that had not reached PD at the point that analysis was performed,	Defined as the time (in months) from the date of randomisation until the date of the investigator-assessed radiological disease progression or death due to any cause. Patients who were alive with no radiological disease progression as of

	Phase II trial	RECOURSE
	and for patients in which the date that PD was reached was unknown, PFS time was censored at the date of the patient's final assessment prior to data cut-off. The randomisation date was used for cases in which lesion evaluation had not been performed after randomisation, and the initiation date of other (post-treatment) anti-cancer therapy was used when other anti-cancer therapy was initiated before the patient reached PD.	the analysis cut-off date were censored at the date of the last tumour assessment. Patients who received non-study cancer treatment before disease progression, or patients with clinical but not radiological evidence of progression, were censored at the date of the last radiological evaluable tumour assessment before the non-study cancer treatment was initiated.
Response rates	Based on Response Evaluation Criteria in Solid Tumours (RECIST), the tumour shrinkage effect was evaluated and the response rate was calculated. The response rate was the percentage of patients in which the best overall response was determined to be complete response (CR) or partial response (PR) in each treatment group. The determination of the antitumor effect was to be performed in accordance with RECIST Ver. 1.0. At the independent image assessment site (CRO), determination of antitumor effect was made in accordance with RECIST Ver. 1.0 as well as RECIST Ver. 1.1 as a reference.	Overall response rate (ORR): Based on investigator review of radiological images and following RECIST criteria (version 1.1, 2009). ORR was defined as the proportion of patients with objective evidence of CR or PR with no confirmatory scan required. The primary assessment of ORR was for the ITT population, restricted to patients with measurable disease (at least 1 target lesion) at baseline. At the analysis stage, the best overall response was assigned for each patient as the best response recorded from all responses recorded from the start of treatment through the treatment period (excludes assessments during follow-up). If applicable, responses recorded after radiological disease progression or after initiation of non-study anti-tumour therapy were excluded. A best response assignment of SD required that SD be maintained for at least 6 weeks from the start of treatment.
Adverse events of treatment	Assessed according to the National Cancer Institute Common Terminology Criteria for Adverse Events (version 3.0).	Standard safety monitoring and grading were performed using National Cancer Institute Common Terminology Criteria for Adverse Events Version 4.03. The evaluation of safety was based on the incidence, severity, and causality of AEs and SAEs and other safety assessments including physical examination, vital signs, ECOG performance status, 12-lead ECG, and clinical laboratory evaluations.
Health-related quality of life	Not assessed in the trial	

AE = adverse event; CR = complete response; CRO = contract research organisation; CS = company submission; ECG = electrocardiogram; ECOG = Eastern Cooperative Oncology Group; ITT = intention-to-treat; ORR = overall response rate; PD = progressive disease; PFS = progression-free survival; PR = partial response; RCT = randomised controlled trial; RECIST = Response Evaluation Criteria in Solid Tumours; SAE = serious adverse events; SD = stable disease

Table 4.5: Characteristics of participants in the included RCTs
(Based on Tables 18 and 19 of the CS[1])

	Phase II trial		RECOURSE	
	T/T (n=114)	**BSC (n=58)**	**T/T (n=534)**	**BSC (n=266)**
Age (median, range)	63 (28 – 80)	62 (39 – 79)	63.0 (27-82)	63.0 (27-82)
Gender (M/F)	64 (57%); 48 (43%)	28 (49%); 29 (51%)	326 (61.0); 208 (39.0)	165 (62.0); 101 (38.0)
Race	Asian: 114 (100%)	Asian: 59 (100%)	White: 306 (57.3); Asian: 184 (34.5); Black: 4 (0.7)	White: 155 (58.3); Asian: 94 (35.3); Black: 5 (1.9)
Geographic location (%)	Japan: 100	Japan: 100	Japan: 33.3; Europe: 50.7; USA: 12.0; Australia: 3.9	Japan: 33.1; Europe: 49.6; USA: 13.2; Australia: 4.1
ECOG PS	0: 72 (64%); 1: 37 (33%); 2 (3%)	0: 35 (61%); 1: 21 (37%); 2: 1 (2%)	0: 301 (56.4); 1: 233 (43.6)	0: 147 (55.3); 1: 119 (44.7)
Primary tumour site	Colon: 63 (56%); Rectum: 49 (44%)	Colon: 36 (63%); Rectum: 21 (37%)	Colon: 338 (63.3); Rectum: 196 (36.7)	Colon: 161 (60.5); Rectum: 105 (39.5)
Number of metastatic sites	1: 25 (22%); 2: 43 (38%); 3: 27 (24%); 4: 17 (15%)	1: 11 (19%); 2: 20 (35%); 3: 12 (21%); 4: 14 (25%)	NR	NR
Time since diagnosis of metastasis	NR	NR	<18 months: 111 (20.8); ≥18 months: 423 (79.2)	<18 months: 55 (20.7); ≥18 months: 211 (79.3)
Metastatic organ	Liver: 65 (58%); Lung: 87 (78%); Lymph: 48 (43%); Peritoneum: 11 (10%)	Liver: 38 (67%); Lung: 44 (77%); Lymph: 23 (40%); Peritoneum: 17 (30%)	NR	NR
Previous treatment and reason	Surgical history: 103 (92%); Adjuvant chemotherapy: 54 (48%)	Surgical history: 50 (88%); Adjuvant chemotherapy: 15 (26%)	NR	NR
Number of palliative chemotherapies	2: 17 (15%); ≥3: 95 (85%)	2: 13 (23%); ≥3: 44 (77%)	2: 95 (17.8); 3: 119 (22.3); ≥4: 320 (59.9)	2: 45 (16.9); 3: 54 (20.3); ≥4: 167 (62.8)
Fluoropyrimidine-based treatment	Refractory: 109 (97%); Intolerant: 3 (3%)	Refractory: 55 (96%); Intolerant: 2 (4%)	100%	100%
Oxaliplatin-based treatment	Refractory: 95 (85%); Intolerant: 17 (15%)	Refractory: 45 (79%); Intolerant: 12 (21%)	100%	100%

	Phase II trial		RECOURSE	
	T/T (n=114)	**BSC (n=58)**	**T/T (n=534)**	**BSC (n=266)**
Irinotecan-based treatment	Refractory: 106 (95%); Intolerant: 6 (5%)	Refractory: 56 (98%); Intolerant: 1 (2%)	100%	100%
Bevacizumab	87 (78%)	47 (82%)	100%	99.6%
Cetuximab	71 (63%)	36 (63%)	NR	NR
Regorafenib	NR	NR	17.0%	19.9%
Anti-EGFR (if wild-type KRAS)	NR	NR	99.6%	99.3%
KRAS mutational status	Wild-type: 54 (55%); Mutation-positive: 45 (45%)	Wild-type: 24 (48%); Mutation-positive: 26 (52%)	Wild-type: 262 (49.1); Mutation-positive: 272 (50.9)	Wild-type: 131 (49.2); Mutation-positive: 135 (50.8)

CS = company submission; ECOG = Eastern Cooperative Oncology Group; EGFR = epidermal growth factor receptor; F = female; KRAS = Kirsten rat sarcoma viral oncogene homolog; M = male; NR = not reported; PS = performance status; RCT = randomised controlled trial; T/T = Trifluridine/tipiracil

ERG comment: The ERG examined the list of excluded studies and considered all of them to have been appropriately excluded. Furthermore, the ERG is satisfied that no data from the ongoing TERRA trial could have been used to inform the CS. The ERG does not consider it appropriate to comment on two non-randomised studies in detail as they should have been excluded from the systematic review. Therefore, only the two identified RCTs (phase II trial and RECOURSE) will be discussed in this section.

As can be seen in Table 4.2, although the two studies were conducted at different phases of development they are similar in terms of population eligibility criteria, intervention and comparator, primary and secondary outcomes and trial design.

The methodology of the included studies is presented in Table 4.3 and discussed below.

Location

The phase II trial was located in Japan whereas RECOURSE was a worldwide trial. The company was asked to clarify the number of UK participants in RECOURSE and to provide baseline characteristics and results and to consider the representativeness of the two trials for a UK setting.[20] The response for request for clarification confirmed that nine patients in five centres were recruited from the UK (seven patients in T/T group and two in BSC arm).[9] Characteristics of the UK participants were provided. As the participant numbers were extremely small the company did not provide results for this subpopulation. This appears reasonable. The company cited the multivariate analysis including geographic region and the pre-specified geographical regional subgroup analysis of RECOURSE and stated *"as there is no evidence of a difference in efficacy based on ethnicity, the included patients are generalizable to the UK setting"*.[9] The ERG considers this to be reasonable but draws the attention of the committee to the lack of participants from England and Wales.

Trial design

Both trials are multi-centre, randomised with a placebo control group which is a rigorous design. More comments on the quality of the trial design will be made in the section on trial quality (below).

Eligibility criteria for participants

Both trials were in adult participants with confirmed advanced colorectal cancer previously treated with ≥ 2 regimens of standard chemotherapy. This matches the final scope which refers to *"adults with metastatic colorectal cancer whose disease has progressed after standard therapies or for whom standard therapies are unsuitable"*.[14] All patients in the phase II trial and RECOURSE had received fluoropyrimidine, irinotecan and oxaliplatin.

Furthermore, in RECOURSE patients were required to have received prior chemotherapy with bevacizumab. However under NICE guidance patients in England would not be able to routinely receive bevacizumab prior to treatment with T/T. According to the CS, *"due to recent funding changes within England, there is currently no means of obtaining bevacizumab, cetuximab or panitumumab (third or fourth line) within the NHS, apart from if a patient is included in a clinical trial or has private medical insurance. Whilst many trial patients had previously received bevacizumab, cetuximab or panitumumab, it may not be possible for future English mCRC patients to do so. There is no biological reason why trifluridine/tipiracil should not work in patients who have not received these therapies. Indeed within the Phase II study approximately 80% of patients had received bevacizumab and 60%,* [sic!] *cetuximab; meaning that not all patients had received a biological therapy, despite this the results were consistent with the RECOURSE study. Expert clinical opinion considers that patient populations who are not as highly pre-treated as the population in RECOURSE would respond better because their tumours are less resistant to treatment"*.[1, 13] Figure 19 of the CS (*"overall survival in prespecified subgroups in the Phase II trial"*) seem to support the comment, i.e. patients who have not received bevacizumab (hazard ratio (HR) 0.37, 95% confidence interval (CI) 0.16 to 0.86) or cetuximab (HR 0.41, 95% CI 0.22 to 0.76) show better OS than people who have not received these drugs (HR 0.63, 95% CI 0.42 to 0.95 and HR 0.69, 95% CI 0.44 to 1.09, respectively). The CS concludes that *"it seems patients who have not received bevacizumab or cetuximab do better, although statistically there is no interaction"* (section 4.6 of the CS).[1] The company's interpretation in conjunction with clinical advice was that tumours in patients who had received fewer treatments were likely to be less resistant to additional therapy. This implies that the evidence for T/T presented might underestimate response in a UK population. This is an assumption, but it appears to be fair.

In the phase II trial, patients with ECOG (Eastern Cooperative Oncology Group) performance status (PS) 2 were eligible whereas in RECOURSE they were ineligible (Table 4.5). The proportion of patients with ECOG PS 2 in the phase II trial was 3% so this should not make a major difference to overall results. Similar proportions of ECOG PS 0 and 1 were noted in both trials.

Setting

Both trials were conducted in secondary care oncology, gastroenterology or general medicine outpatient departments.

Trial drugs

Both trials had a similar drug regimen. The main difference was that in the phase II trial patients who had adverse events (AEs), the dose could be reduced by 10 mg/day as judged necessary whereas in RECOURSE the protocol allowed for a maximum of three reductions in dose in decrements of 5 mg/m^2. Concomitant therapies (not shown in Table 4.3) permitted were similar.

Primary outcome

Both trials had overall survival as a primary outcome which is line with the final scope.[14]

Secondary outcomes

These were similar across the trials and included progression-free survival, response rates and adverse effects of treatment as specified by the NICE scope.[14] As noted in Section 3.4 of this report neither trial assessed health-related quality of life as specified in the NICE scope.[14]

The ERG wished to examine the definitions of progression-free survival, progression and stable disease particularly given their importance in the economic model. For both trials, the ERG asked for clarification on the assessment methods e.g. how many assessors were involved and training to ensure consistency of outcome ascertainment across trial centres (Table 4.4).

- Progression-free survival was defined similarly across the two trials. In both trials if the patient died before reaching progressive disease (PD), the date of death was considered the date PD was reached.[1]
- In RECOURSE progression was defined as *"at least a 20% increase in the sum of diameters of the target lesions, taking as a reference the smallest sum on study, including the baseline sum. In addition to the relative increase of 20%, the sum must also demonstrate an absolute increase of at least 5 mm. Definitive new lesion presence also indicates progression"*.[2] The company stated that the definition of progression in the phase II trial was in the company submission but it was not. In the CSR progressive disease was defined as *"an increase of 20% or more in the maximum diameter sum of target lesions compared with the smallest maximum diameter sum (including the pre-treatment sum). However, if the maximum diameter sum is 10 mm or less, then an increase in the longest diameter sum of 20% or more is not considered PD"*.[24]
- In RECOURSE stable disease was defined as *"neither sufficient shrinkage to qualify for PR nor sufficient increase to qualify for PD, taking as a reference the smallest sum diameters while on study"*. To get a *"best response"* of *"stable disease"* response has to last for six weeks.[2] For the phase II trial, the company advised that *"the response has not reached complete response (CR) or partial response (PR) in radiologic assessments over at least six weeks since the start of study drug administration and it has been confirmed that progressed disease (PD) has not occurred"*.[9]
- In response to request for clarification, the company confirmed that for both trials training provided to each centre was consistent across all study centres. The company further stated that in order to ensure consistency across study centres all secondary efficacy endpoints in the phase II trial were subject to independent radiologic assessment.[9] Centres in RECOURSE received an imaging manual to ensure consistency and an audit plan was put in place. The ERG was satisfied with the measures in place.

Adverse events in both trials were assessed according to the National Cancer Institute Common Terminology Criteria for Adverse Events (version 3.0).[1] The company was asked to confirm if all adverse events from the included trials had been included in the submission.[20] The company replied that details of all adverse events were either in the manuscript or in the clinical study reports (CSRs).[9] The ERG examined the reports of adverse events in the two trials and provides an overview in this report.

Pre-planned subgroups

These were similar across the two trials and included variables that might be expected to impact on results, for example KRAS mutation status, age, primary site and number of prior treatment regimens.[1]

The phase II trial also included an assessment of those who had taken bevacizumab whereas in RECOURSE all patients had to have received this treatment. Thirty-five patients (22%) of the patients

in the phase II trial did not receive bevacizumab. Both those receiving bevacizumab and those who did not benefited in terms of overall survival. Those who did not receive bevacizumab, and are thus directly appropriate to the England and Wales population, represent a small percentage of the trial populations (approximately 4%).

RECOURSE conducted a subgroup analysis of participants from Japan compared to participants from the rest of the world. This was used to show the applicability of the phase II trial conducted solely in Japan as results were found to be similar. The company stated *"as RECOURSE included Japanese patients, it was possible to observe whether all patients responded to trifluridine/tipiracil in a similar manner; as would be expected from the known pharmacology of the compound. In patients treated with trifluridine/tipiracil, outcomes and response for pre-specified regional subgroups were similar, with non-significant tests for interaction. Hence, it is possible to generalise the results of both studies to Western populations"* (Section 4.6 of the CS).[1] The ERG believes this to be reasonable.

Sample size calculations and analysis methods

According to the CS, for the phase II trial *"a sample size of 162 patients with a one-sided significance level of 10% was necessary to verify superiority in overall survival (OS) with a power of 80%, with an expected HR of 0.67. Median OS was anticipated to be 9.0 months in the trifluridine/tipiracil group and 6.0 months in the placebo group. A clinically relevant HR was estimated as 0.70. Patients continued to receive the study treatment (with group assignments remaining concealed) until the primary analysis of OS was done. The efficacy analysis was done in the intention-to-treat (ITT) population, and the safety analyses in the per-protocol population, when the number of deaths in the trial reached 121. The Kaplan-Meier method was used to estimate survival distribution. A stratified log-rank test was used and adjusted by the allocation factor, for comparisons between the two groups, and a Cox proportional hazards model to estimate HRs, the two-tailed 80% CIs corresponding to the significance level, and 95% CIs"*.[1]

For RECOURSE, *"the study was designed to have 90% power to detect a HR for death of 0.75 (a 25% reduction in risk) in the trifluridine/tipiracil group compared with the placebo group, with a one-sided type I error rate of 0.025. Given the treatment assignment ratio of 2:1 (trifluridine/tipiracil: placebo), 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. OS (the primary endpoint) and radiologically confirmed PFS were analysed in the ITT population with the use of a two-sided, stratified log-rank test, with the HR and two-sided 95% confidence intervals based on a stratified Cox model and the associated Kaplan-Meier survival estimates. The primary analysis of OS includes follow-up data (including death events) obtained up to the date of the 571st death observed in the study. Patients having a documented survival status (alive or dead) after this date were censored at the cut-off date, but are they included in an updated analysis, which is used in the economic analysis. The median survival times were determined from the Kaplan-Meier curves. Rates of objective response and disease control were compared with the use of Fisher's exact test in the subgroup of the ITT population that had measurable disease at baseline"*.[1]

ERG comment: The sample size calculations in both trials were based on the primary endpoint of OS only, therefore neither trial was powered for secondary outcomes. Both trials used one-sided significance levels in the sample size calculation although in RECOURSE that was equivalent to the standard two-sided 95% CI which was reported in the results. In Phase II they used a larger significance level of a one-sided 10% level (equivalent to a two-sided 80% CI) without justifying this choice. However the 95% CIs were reported in the submission which use a stricter significance level and correspond with the RECOURSE results. Both trials reached their recruitment targets for numbers of

participants and deaths so both appear to be adequately powered for OS. Both trials also used appropriate statistical analysis methods for all outcomes.

Quality Assessment

Table 21 of the company submission presents the quality assessment results of the included trials. It is reproduced in Table 4.6. ERG comments can be found below the table.

Table 4.6: Quality assessment of the included RCTs

(Based on Table 21 of the CS[1])

	Phase II	RECOURSE
Was the randomisation carried out appropriately?	**Yes** Following confirmation of eligibility as a subject for randomisation, on the basis of probability theory minimising methods, patients were assigned by the registration centre to the two treatment groups (trifluridine/tipiracil group and placebo group) at a ratio of 2:1. So as to ensure balance between the therapy groups, subjects were to be stratified at the time of randomisation according to the following stratification factors: • Performance Status: 0 vs. 1/2 At the registration centre, on the basis of a random assignment table, a drug number including the appropriate drug that was distributed to each implementing medical institution was assigned. The drug number was recorded in the raw data of each patient. The assignment was a dynamic allocation and thus caution was taken that the drug numbers were conferred randomly. Note that in cases in which the investigational drug of a drug number assigned to a patient was not used, other patients were not to use it, including the same patient in a later study period. For details of the random assignment and drug number assignment, the "Registration manual" was referred to. Rationale for setting of allocation adjustment factors; 'PS (0, 1/2)' is a general prognosis factor in cancer clinical trials and it was established considering the difference in efficacy and safety evaluations due to differences in the patient's condition.	**Yes** Once patient confirmation of eligibility and the criteria for randomisation had been met, patients were centrally randomised in a 2:1 ratio to trifluridine/tipiracil or placebo via an IWRS based on a dynamic allocation method (biased coin). The IWRS assigned kit numbers corresponding to the patient's treatment assignment and informed the study site user of the kit number that had been assigned to the patient for the dispensing of study drug. If a patient was mistakenly given a kit(s) of study medication that was not the kit assigned by the IWRS, resulting in the patient being initiated in the alternate arm from which they were assigned at randomisation, the patient continued to receive this treatment for the rest of the study. Study medication administration was to begin within 3 days following randomisation.
Was the concealment of treatment allocation adequate?	**Yes** This study was blinded for all the concerned parties of implementing medical institutions (such as patients, investigator	**Yes** This was a double-blind study. Trifluridine/tipiracil tablets of each strength, 15-mg or 20-mg, and the corresponding placebo tablets, 15-mg and 20-mg, were identical in appearance and

	Phase II	RECOURSE
	or sub-investigators, and study research staff) as well as the sponsor. The investigator or a sub-investigator was to prescribe to the patient an investigational drug of the investigational drug number assigned by the registration centres. In cases where information was necessary on the treatment group to which a patient was assigned in order to manage symptoms of the patient during an emergency resulting from, for example, a serious adverse event during the course of the study, the investigator was to contact a specific management service. Unblinding of the study was to be made after the events specified in the "Statistical analysis implementation period" were reached. The investigational drug assignment manager was to confirm that closing out of all applicable cases was completed by the sponsor. In addition, prior to the unblinding, the investigational drug assignment manager was to confirm the sealed status of the collected investigational drug and confirm that the keycode for emergency unblinding was appropriately stored and managed.	were packaged in identical containers. During the conduct of the study, the treatment assignment was unknown to all patients, investigators, and ancillary study personnel at each study site. During the conduct of the study, assigned treatment was unknown to the study team at Taiho Oncology, Inc. and Taiho Pharmaceutical Co., Ltd. except for pre-specified personnel involved in pharmacovigilance reporting activities and clinical trial material management. Among the CROs who assisted in the conduct of the study, treatment assignment was unknown except for personnel involved in drug labelling and distribution. Unblinding of the study treatment by the investigator was not to occur unless needed to manage a patient's medical condition. In an emergency, when specific knowledge of the patient's treatment assignment was needed to manage a patient's medical condition, the investigator could unblind the patient by calling the IWRS to obtain the patient's treatment assignment. If unblinding occurred, the investigator was not to disclose the unblinding information.
Were the groups similar at the outset of the study in terms of prognostic factors?	**No** There were some slight differences in some of the subgroups; namely sex, metastatic site, number of prior chemotherapy regimens and KRAS status.	**Yes** The groups were balanced in terms of KRAS status, time since diagnosis of 1st metastasis, region, BRAF status, age, race, gender, primary tumour site, ECOG score, number of prior regimens, and number of metastatic sites.
Were the care providers, participants and outcome assessors blind to the treatment allocation?	**Yes** See above regarding concealment of treatment allocation	**Yes** See above regarding concealment of treatment allocation
Were there any expected imbalances in drop-outs between groups?	**No** Please see patient disposition	**No** Please see patient disposition

	Phase II	RECOURSE
Is there any evidence to suggest that the authors measured more outcomes than they reported?	**No**	**No**
Did the analysis include an intention to treat analysis?	**Yes**	**Yes**

BRAF = serine/threonine-protein kinase B-Raf; ECOG = Eastern Cooperative Oncology Group; IWRS = interactive voice/web response system; KRAS = Kirsten rat sarcoma viral oncogene homolog; PS = performance status

ERG comments: Randomisation and concealment of treatment allocation were carried out appropriately in both trials. Patients in the phase II trial were stratified on ECOG performance status (0 vs. 1/2) whereas stratification for RECOURSE was based on KRAS mutation status.

In terms of prognostic factors, participants in RECOURSE were balanced between treatment groups. The phase II trialists noticed some slight differences in terms of sex, metastatic site, number of prior chemotherapy regimens and KRAS status. The ERG notes that these differences did not appear to bias the trial in favour of T/T.

Procedures for blinding of patients, care providers and outcome assessors appear to be appropriate.

Drop-out: The ERG found no evidence of differential dropout between treatment groups in the two trials and an ITT analysis was included in both trials. In the phase II trial, two patients did not receive the allocated treatment (1 T/T, 1 BSC – reasons supplied) and one had a protocol violation. These patients were omitted from the efficacy analysis but the latter was included in the safety analysis. In RECOURSE two patients did not receive the allocated treatment (1 T/T, 1 BSC). Six patients were lost to follow-up (three in each group) and one patient on T/T dropped out (Figure 12).[1] All patients were included in efficacy analyses with the exception of two who had not received treatment.[1]

Measurement of more outcomes than reported: The ERG agrees with the assessment in the CS.

Results of trials and pooled analyses

Table 4.7 details the results of the two included trials and the pooled analysis for the primary and secondary outcomes. A comparison of discontinuation rates in the two trials is given in Table 4.8. Adverse events in the phase II trial and RECOURSE are reported in Table 4.9 (all grades) and Table 4.10 (grade ≥3).

Table 4.7: Results of the included RCTs

(Based on Figure 27, Tables 22, 23 and 29 as well as Sections 4.7.1, 4.9 and 5.7.2 of the CS[1])

Outcome	Phase II	RECOURSE	Pooled Analysis*
Outcomes in the final scope[14]			
Number of deaths	T/T: 75 (67%) BSC: 48 (84.2%)	*Original analysis (574 deaths)* T/T: 364 (68.2%) BSC: 210 (78.9%) *Updated analysis (712 deaths)* T/T: NR BSC: NR	T/T: 538 (83.3%) BSC: 297 (92%)
Overall survival (OS)	*Median* T/T: 9.0 months (95% CI 7.3 to 11.3) BSC: 6.6 months (95% CI 4.9 to 8.0)	*Original analysis (574 deaths, median)* T/T: 7.1 months (95% CI 6.5 to 7.8) BSC: 5.3 months (95% CI 4.6 to 6.0) *Updated analysis (712 deaths, median)* T/T: 7.2 months (95% CI 6.6 to 7.8) BSC: 5.2 months (95% CI	*Median* T/T: 7.3 months BSC: 5.4 months *Mean* T/T: 9.1 months BSC: 6.8 months

HR OS	0.56 (95% CI 0.39 to 0.81)	4.6 to 5.9 *Original analysis (574 deaths)* 0.68 (95% CI 0.58 to 0.81) *Updated analysis (712 deaths)* 0.69 (95% CI 0.59 to 0.81)	0.67 (95% CI 0.58 to 0.78)
Progression-free survival (PFS)	*Median (IRC)* T/T: 2 months (95% CI 1.9 to 2.8) BSC: 1 month (95% CI 1.0 to 1.0)	*Median* T/T: 2 months (95% CI 1.9 to 2.1) BSC: 1.7 months (95% CI 1.7 to 1.8)	*Median* T/T: 1.9 months BSC: 1.7 months *Mean* T/T: 3.7 months BSC: 1.9 months
HR PFS	*IRC* 0.41 (95% CI 0.28 to 0.59)	0.48 (95% CI 0.41 to 0.57)	0.46 (95% CI 0.40 to 0.53)
Response rates	*IRC* *CR:* 0 in both groups *PR:* T/T: 1; BSC: 0 *SD:* T/T: 48 (-42.9%); BSC: 6 (-10.5%) *Progression of disease:* T/T: 53 (-47.3%); BSC: 44 (-77.2%)	*CR:* T/T: 0; BSC: 1 (0.4%) *PR:* T/T: 8; BSC: 0 *SD:* T/T: 213 (-42.4%); BSC 41 (-15.9%) *Progression of disease:* T/T: 260 (-51.8%); BSC 195 (-75.6%)	NA
Adverse effects of treatment	See tables 4.9 and 4.10		NA
Health-related quality of life	NR	NR	NA
Outcomes not defined in the final scope[14]			
Median time to treatment failure	*IRC* T/T: 1.9 months; BSC: 1 month	T/T: 1.9 months; BSC: 1.7 months	NA
HR time to treatment failure	*IRC* 0.40 (95% CI 0.28 to 0.56)	0.50 (95% CI 0.42 to 0.58)	NA
Disease control rate (CR + PR + SD; n (%))	T/T: 49 (-43.8%); BSC 6 (-10.5%)	T/T: 221 (44%); BSC: 42 (16.3%)	NA
Treatment discontinuation	See table 4.8		NA

* Using the updated RECOURSE analysis of 712 deaths (8 October 2014)

BSC = best supportive care; CI = confidence interval; CR = complete response; CS = company submission; HR = hazard ratio; IRC = independent review committee; NA = not applicable; NR = not reported; OS = overall survival; PFS = progression-free survival; PR = partial response; RCT = randomised controlled trial; SD = stable disease; T/T = trifluridine/tipiracil

Table 4.8: Comparison of discontinuation rates in the included RCTs

(Based on Table 11 and Figure 12 of the CS[1] and the CSRs[24, 25]. Numbers extracted from the CS. Where a discrepancy has been identified, the information from the CSR has been extracted as well.)

	Phase II trial						RECOURSE					
	Trifluridine/tipiracil			BSC			Trifluridine/tipiracil			BSC		
	n	Number of events	%	n	Number of events	%	n	Number of events	%	n	Number of events	%
Discontinued treatment (any reason)	114	109	95.6	58	57	98.3	534	496	93	266	263	>99
Discontinued treatment due to AE/SAE	114	4 [redacted]	3.5 [redacted]	58	1	1.7	534	18 [redacted]	4	266	4	2
Discontinued treatment due to death	114	NR	NR	58	0	0	534	7	1	266	4	2

AE = adverse event; CS = company submission; CSR = clinical study report; SAE = serious adverse event

Table 4.9: Comparison of adverse events in the RECOURSE trial and phase II trial (all grades)

(Based on Tables 41 and 43 of the CS[1], the CSRs[24, 25] and Mayer et al. 2015[2]. Numbers extracted from the CS. Where the information was not reported in the CS or a discrepancy has been identified, relevant information from the CSR and/or Mayer et al. 2015 have been extracted as well.)

	Phase II						RECOURSE					
All grades AE	**Trifluridine/tipiracil**			**BSC**			**Trifluridine/tipiracil**			**BSC**		
	n	**Number of events**	**%**	**n**	**Number of events**	**n**	**n**	**Number of events**	**%**	**n**	**Number of events**	**%**
Any event	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	533	524	98.3	265	247	93.2
Any SAE	113	41[*] (21 patients)	18.6	57	8 (5 patients)	8.8	533	158	29.6	265	89	33.6
Any treatment-related AE	113	109[‡*]	96.5	57	40[‡]	70.2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Nausea	113	73	64.6	57	16	28.1	533	258[†]	48.4	265	63	23.8
Vomiting	113	38	33.6	57	14	24.6	533	148[†]	27.8	265	38	14.3
Decreased appetite	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	533	208[†]	39.0	265	78	29.4
Diarrhoea	113	43	38.1	57	12	21.1	533	170[†]	31.9	265	33	12.5
Abdominal pain[†]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	533	113[†]	21.2	265	49	18.5
Gastrointestinal disorders	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Neutropenia	113	81	71.7	57	1	1.8	528[§] [redacted]	358 [redacted] Mayer: 353	67.8 [redacted] Mayer: 67	263 [redacted]	2 [redacted]	0.8 [redacted]
Leucopenia	113	86	76.1	57	2	3.5	528[§]	407	77.1	263	12	4.6

	Phase II						RECOURSE					
All grades AE	**Trifluridine/tipiracil**			**BSC**			**Trifluridine/tipiracil**			**BSC**		
	n	**Number of events**	**%**	**n**	**Number of events**	**n**	**n**	**Number of events**	**%**	**n**	**Number of events**	**%**
Anaemia	113	82	72.6	57	9	15.8	528[§]	404	76.5	263	87	33.1
Thrombocytopenia	113	44	38.9	57	1	1.8	528[§]	223	42.2	263	21	8.0
Treatment related death	113	0[*]	0	57	0	0						
Death due to AE												

[*] Page 112 of CS

[ǂ] per patient

[†] Adverse events of any grade that are listed as most common occurred in 10% or more of patients in the trifluridine/tipiracil group and in a greater percentage in that group than in the BSC group.

[#] Diarrhoea and/or nausea and/or vomiting

[§] The denominator for the percentage of patients with laboratory abnormalities is the number of patients with at least one post baseline measurement during treatment.

AE = adverse event; CS = company submission; CSR = clinical study report; SAE = serious adverse event

Table 4.10: Comparison of adverse events in the RECOURSE trial and phase II trial (grade ≥3)

(Based on Tables 42 and 44 of the CS[1], the CSRs[24, 25] and Mayer et al. 2015[2]. Numbers extracted from the CS. Where the information was not reported in the CS or a discrepancy to the CSR has been identified, relevant information from the CSR and/or Mayer et al. 2015 have been extracted as well.)

	Phase II						**RECOURSE**					
Grade ≥3 AE	**Trifluridine/tipiracil**			**BSC**			**Trifluridine/tipiracil**			**BSC**		
	n	**Number of events**	**%**	**n**	**Number of events**	**%**	n	**Number of events**	**%**	n	**Number of events**	**%**
Any event							533	370	69.4	265	137	51.7
Any treatment-related AE							[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Nausea†	113	5	4.4	57	0	0.0	533	10	1.9	265	3	1.1
Vomiting†	113	4	3.5	57	0	0.0	533	11	2.1	265	1	0.4
Decreased appetite†							533	19	3.6	265	13	4.9
Diarrhoea†	113	7	6.2	57	0	0.0	533	16	3.0	265	1	0.4
Abdominal pain†							533	13	2.4	265	10	3.8
Neutropenia§	113	57	50.4	57	0	0.0	528	200	37.9	263	2 Mayer: 0	0.8 Mayer: 0
Leucopenia§	113	32	28.3	57	0	0.0	528	113	21.4	263	12 Mayer: 0	4.6 Mayer: 0
Anaemia§	113	19	16.8	57	3	5.3	528	96	18.2	263	87 Mayer: 8	33.1 Mayer: 3

	Phase II						RECOURSE					
Grade ≥3 AE	**Trifluridine/tipiracil**			**BSC**			**Trifluridine/tipiracil**			**BSC**		
	n	**Number of events**	**%**	**n**	**Number of events**	**%**	n	**Number of events**	**%**	n	**Number of events**	**%**
Thrombocytopenia[§]	113	5	4.4	57	0	0.0	528	27	5.1	263	21	8

[†] Adverse events of any grade that are listed as most common occurred in 10% or more of patients in the trifluridine/tipiracil group and in a greater percentage in that group than in the BSC group.

[§] The denominator for the percentage of patients with laboratory abnormalities is the number of patients with at least one post baseline measurement during treatment.

[ǂ] per patient

AE = adverse event; CS = company submission; CSR = clinical study report; NR = not reported; SAE = serious adverse event

ERG comments: Results are reported for the original analysis of RECOURSE (574 deaths, 24 January 2014) and the updated analysis (712 deaths, 8 October 2014). The pooled results use the updated data from RECOURSE. This appears reasonable.

Overall survival

Based on the updated analysis of 712 deaths in RECOURSE an increase in median overall survival of two months in the T/T group was observed (T/T: 7.2 months, BSC: 5.2 months). This was statistically significant (HR 0.69; 95% CI 0.59 to 0.81). The phase II trial showed an increase in median overall survival of 2.4 months (T/T: 9.0 months, BSC: 6.6 months). This was statistically significant (HR 0.56; 95% CI 0.39 to 0.81). In the pooled analysis, there was an increase in survival of 1.9 months (T/T: 7.3 months, BSC: 5.4 months). This was statistically significant (HR 0.67; 95% CI 0.58 to 0.78). The pooled mean increase in survival is 2.3 months (T/T: 9.1 months, BSC: 6.8 months). It is noted that, based on the trial data, the increase in survival for T/T compared to BSC is less than that specified in end of life care (minimum of three months, see Section 7).

Progression-free survival

Median PFS was similar in RECOURSE and in the pooled results. The pooled results showed an increase of 0.2 months (T/T: 1.9 months, BSC: 1.7 months). This was statistically significant (0.46; 95% CI 0.40 to 0.53; $p < 0.0001$). The mean PFS increase was 1.8 months (T/T: 3.7 months, BSC: 1.9 months).

Response rates

In the phase II trial no patient in either group had a complete response and one in the T/T group had a partial response. In RECOURSE one patient in the BSC group had a complete response and eight in the T/T group had a partial response. A greater proportion of T/ T patients in both trials had stable disease (42.9% vs. 10.5% in the phase II trial and 42.4% vs. 15.9% in RECOURSE).

Adverse effects of treatment

Rates of discontinuation (for any reason, due to adverse events (AEs), due to serious AE (SAEs) or due to death) were found to be broadly similar between T/T and BSC arms in the phase II trial and RECOURSE (summarised in Table 4.8). In both trials, one discrepancy was noted between the company submission[1] and the respective clinical study report[24,25]; this was the number of patients who discontinued due to AE. This appears to be a minor difference which should not influence the overall result.

All adverse events were graded according to National Cancer Institute Common Terminology Criteria for Adverse Events, version 4.03 in the RECOURSE trial, whilst the Common Terminology Criteria for Adverse Events (CTCAE Ver. 3.0 Japanese translation, Japan Clinical Oncology Group (JCOG)/ Japan Society of Clinical Oncology (JSCO) version) was used for the phase II trial. The ERG compared the rates of all major adverse events and in particular noted those associated with myelosuppression which the Food and Drug Administration (FDA) considered an important side effect of this drug[26, 27] and gastrointestinal side effects which are considered important by the Committee for Medicinal Products for Human Use (CHMP).[16]

Any AE or SAE were similar between T/T and BSC arms for the RECOURSE trial.[25] The phase II trial did not report data for any AE, however numbers were reported in Table 12.2.1-1 (p. 217) of the clinical study report (CSR) for the phase II trial and were [redacted].[24] [redacted]. The phase II trial reported

numbers for 'adverse drug reactions' in the text of the company submission (p. 112).[1] The definition of 'adverse drug reaction' was "*those that were determined to have a positive relationship with the investigational drug"* (Section 9.5.3.2.6 of the CSR); which would be consistent with 'treatment-related AE' reported in the CSR of RECOURSE (Table 35 of the CSR).[25] In both trials 'treatment-related AEs' were found to be [redacted].

In both trials the following gastrointestinal related AE were found to be [redacted] (see Table 4.9 for details). Results for abdominal pain were similar in both arms for the RECOURSE trial as reported in the CS[1] or the CSR (Table 37)[25]; and for the phase II trial (results identified in the CSR[24]). Gastrointestinal disorders were recorded as a class in the CSR for the phase II trial (Table 12.2.3.1-1) and RECOURSE (Table 52) and therefore are reported here. In both trials [redacted].

In both trials the following AEs related to myelosuppression were found to be [redacted], see Table 4.9. The results were inconsistently reported between the submission and the clinical study reports and the publication of RECOURSE.[2] These discrepancies may be due to differences between using number of events and number of patients as the numerator; however it did not change the overall direction of the results.

In the CS, only the phase II trial reported treatment-related deaths and found none occurred. Results for this AE (Table 35 of the CSR) were identified in the CSR of the RECOURSE trial[25]; only one death was reported for the T/T arm. 'Death due to AE' was not reported within the CS but was identified in the CSR for both trials. In RECOURSE, more patients in the BSC arm were reported to [redacted].[24]

Adverse events which were of a higher severity (≥3 grade) are shown in Table 4.10. Results for any AE were found to be higher in the T/T arm of the RECOURSE trial (69.4% vs. 51.7%). in addition any treatment related AE (Table 35 of the CSR) were found to [redacted] (see Table 4.10 for details). Corresponding events related to myelosuppression [redacted].

Overall, more treatment related adverse events occurred in the T/T treatment arm rather than BSC. [redacted].

Health-related quality of life

Neither of the two included trials assessed health-related quality of life.

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

The company submission did not present an indirect comparison and/or multiple treatment comparison.

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

The company submission did not present an indirect comparison and/or multiple treatment comparison.

4.5 Additional work on clinical effectiveness undertaken by the ERG

No additional work on clinical effectiveness was undertaken by the ERG.

4.6 Conclusions of the clinical effectiveness section

The company's submission includes a systematic review of the available evidence for T/T compared to BSC for patients with advanced/metastatic colorectal cancer receiving treatment at the third line or beyond. Although some issues were highlighted in searching for studies of adverse events for the systematic review, the ERG is overall satisfied that the company identified and appraised the relevant randomised trials. The two non-randomised studies in the adverse effects section of the submission did not appear to have been selected systematically. We have focused our attention in this report on the two randomised trials which inform the cost effectiveness model.

There is a lack of information on methods of pooling the two included randomised trials but overall it was considered acceptable from the point of view of clinical effectiveness that the trials were pooled.

The two included trials (phase II trial and RECOURSE) were randomised and compared T/T to placebo with both treatment groups in the trials receiving best supportive care. Our evaluation of the quality confirmed the company's assessment that both trials were of high quality.

RECOURSE was an international trial whereas the phase II trial was conducted solely with Japanese participants. The ERG considered that the company had provided evidence that geographical region was not a factor in effectiveness. This meant that results of the Japanese trial could be pooled with RECOURSE. However the ERG draws the committee's attention to the low proportion of UK participants in RECOURSE (9 of 800 patients). However 394 of 800 were from Europe. The ERG further notes that there is an under-representation of non-white, non-Asian populations across the trial (approximately 1% of RECOURSE participants are listed as 'black').

Considering further the issue of applicability of the trials, the population in RECOURSE is a more treated population than might be expected in practice in England and Wales. Patients were required to have received chemotherapy with fluoropyrimidine, oxaliplatin, irinotecan and bevacuzimab. They were also required to have received cetuximab or panitumumab if KRAS wild-type. Bevacuzimab is not currently available in England and Wales. A small number in the phase II trial had not received bevacuzimab (22%) but the phase II trial included fewer participants than RECOURSE. Those who did not receive bevacizumab, and are thus appropriate to the England and Wales population, represent a small percentage of the trial populations (approximately 4%). The company states that T/T might be expected to work better in a less treated population based on clinical advice. This appears to be reasonable but is drawn to the attention of the committee.

The scope issued by NICE recommended comparing T/T to best supportive care (BSC) as there are no currently recommended treatments for patients who have failed second line treatment. The CS is based on two placebo-controlled trials where both treatment and placebo groups received BSC. The ERG asked for clarification on the definitions of BSC used in the included trials, the guidance regarding BSC given to the centres involved in the included trials and the applicability of the BSC to the UK setting.

The company clarified that there is no internationally accepted definition of BSC for clinical trials. Although both trials ensured consistency on medications excluded from BSC, the nature of BSC provided could potentially vary between trial centres. The nature of BSC provided might also differ from that provided in England and Wales given that a very small number of participants were from centres in England and Wales.

In relation to outcomes, the ERG notes that the company provided two analyses of overall survival for the RECOURSE trial, an original (24 January 2014, 574 deaths) and an updated analysis (8 October 2014, 712 deaths). This updated, post-hoc analysis was requested during the CHMP review and the ERG considers it appropriate to present this analysis in the submission to maximise the data available. The ERG notes that the pooled analysis for overall survival was based on the updated analysis of RECOURSE.

In the pooled analysis there was an increase in median overall survival of 1.9 months (T/T: 7.3 months, BSC: 5.4 months, no CIs reported). The pooled mean increase in overall survival is 2.3 months (T/T: 9.1 months, BSC: 6.8 months, no CIs reported). It is noted that, based on the trial data, the increase in survival is less than that specified in end of life care (minimum of three months).

The main trial, RECOURSE, was powered for the outcome of overall survival so may not have had sufficient power to detect all differences between treatment groups for secondary outcomes. The included trials do not directly assess health-related quality of life as specified in the NICE scope. Although there is a benefit to patients of the mean increase in overall survival of 2.3 months (pooled result) the quality of life experienced can only be inferred from effects of disease control and occurrence of adverse events. A significant benefit of T/T for progression-free survival has been shown although this is modest. In terms of disease control, a greater proportion of T/ T patients in both trials had stable disease (42.9% vs. 10.5% in the phase II trial and 42.4% vs. 15.9% in RECOURSE). However numbers achieving partial response or complete response were very small overall.

Rates of adverse events and serious adverse events were similar between T/T and BSC for the RECOURSE trial.[25] The phase II trial was found to be similar between treatment arms for adverse events but SAE were found to be higher in the T/T arm (18.6% vs. 8.8%).[24] The phase II trial reported numbers for 'adverse drug reactions'.[1] The definition was found to be consistent with 'treatment-related AE' reported in the CSR of RECOURSE.[25] In both trials 'treatment-related AEs' were found to be higher in the T/T arms than in the BSC arms (85.7% vs. 54.7% and 96.5% vs. 70.2%, respectively).

In RECOURSE, more patients in the BSC arm were reported to die from AE than in the T/T arm (11.3% vs. 3.2%)[25], whilst in the phase II trial only one case of death due to AE was reported in the T/T treatment arm (Table 12.3.1.1-1 of the CSR).[24]

We compared the rates of all major adverse events and in particular noted those associated with myelosuppression which the Food and Drug Administration (FDA) considered an important side effect of this drug[26, 27] and gastrointestinal side effects which are considered important by the Committee for Medicinal Products for Human Use (CHMP).[16]

Rates of discontinuation (for any reason, due to adverse events (AEs), due to serious AE (SAEs) or due to death) were found to be broadly similar between T/T and BSC in the phase II trial and RECOURSE.

5 COST EFFECTIVENESS

5.1 ERG comment on company's review of cost effectiveness and health-related quality of life evidence

5.1.1 Objective and searches of cost effectiveness review

A systematic review of the published literature was conducted by the company to identify cost effectiveness studies assessing the treatment of patients with mCRC with T/T compared with BSC as third line or later treatment.

Cost effectiveness

The CS states that a systematic review of the published literature was conducted to identify cost effectiveness studies assessing the treatment of patients with mCRC with trifluridine/tipiracil compared to BSC as third line or later treatment.

The searches were conducted on 26 October 2015 in the same databases searched for the clinical effectiveness searches: MEDLINE, MEDLINE In-Process and Other Non-Indexed Citations, Embase and the Cochrane Library (CENTRAL, the Cochrane Database of Systematic Reviews, NHS EED, DARE, and HTA). The host provider for each database was listed; the date span of the databases searched and the specific date the searches were conducted were provided. The company additionally searched conference proceedings: American Society of Clinical Oncology (ASCO), European Society for Medical Oncology (ESMO), and International Society for Pharmacoeconomics and Outcomes Research (ISPOR). Detailed search strategies for the database searches were reported in appendix 6 of the CS.[23] The CS did not provide details of the conference proceedings searches. Full details of the conference proceedings searches for the utility review were provided in response[9] to the ERG clarification letter.[20] These searches could have been used for the cost effectiveness review, as generic search terms for advanced and metastatic colorectal cancer were used, but it is not clear if they were.

The company translated the research question into appropriate search strategies and the ERG considered the searches to be satisfactory. Searches were clearly structured and divided into population, intervention/comparator and cost-effectiveness facets. The search strategies included Boolean, truncation and proximity operators. No date or language limits were included. It was not clear whether a validated study design filter was used for the cost effectiveness facet of search terms.

The searches for cost effectiveness were quite precise, and may have retrieved additional studies with a more sensitive search strategy, i.e. searching for 'economic evaluation OR models', rather than 'economic evaluation AND models'.

All databases included in the Cochrane Library were searched, when only NHS EED and HTA include relevant studies. Further, the search strategy used in the Cochrane Library contained a study design search filter limiting the results to economic evaluations. The ERG considered this to be overly restrictive and unnecessary as the Cochrane databases are pre-filtered resources, i.e. the database of relevance to this search, NHS EED, only contains economic evaluations.

A search of other economic resources, such as the cost effectiveness analysis (CEA) Registry and ScHARRHUD, for cost-utility analyses might have been a useful addition to the literature searches.

Resource identification, measurement and valuation

Searches were not conducted for healthcare resource use identification. Resource costs were identified from two recent NICE technology appraisals, TA242[7] and ID794.[28]

5.1.2 Inclusion/exclusion criteria used in the study selection

Screening of publications by title and abstract was performed to include all currently available treatments; any studies which were not relevant according to the NICE scope were then excluded upon full publication review. Table 5.1 presents the eligibility criteria used for the review.

Table 5.1: Inclusion/exclusion criteria used in the study selection
(Based on Appendix 6 of the CS[1])

	Inclusion criteria
Population	Adult patients with advanced/metastatic CRC receiving treatment at third line or beyond
Interventions	T/T
Comparators	BSC
Outcomes	ICERs Range of ICERs as per sensitivity analyses Assumptions underpinning model structures Key cost drivers Sources of clinical, cost and quality of life inputs Discounting of costs and health outcomes Model summary and structure
Study design	Cost-utility analyses
Language restrictions	None

BSC = best supportive care; CRC = colorectal cancer; ICER = incremental cost-effectiveness ratio; NICE = National Institute for Health and Care Excellence; STA = Single Technology Appraisal

ERG comment: The in- and exclusion criteria seem appropriate for the objective of this review.

5.1.3 Included/excluded studies in the cost effectiveness review

In total, 890 potentially relevant studies were identified of which zero remained after exclusion of duplicates (85 excluded), reviewing title and abstracts (719 excluded) and full paper reviewing (86 excluded). No additional relevant publications were identified via hand searching.

ERG comment: The rationales for excluding studies after full paper reviewing seem appropriate (see Table 2 of Appendix 6 of the CS[1]) given the defined in- and exclusion criteria.

5.1.4 Conclusions of the cost effectiveness review

There were no relevant studies identified in the literature that assess the treatment of patients with mCRC with T/T compared with BSC as third line or later treatment.

ERG comment: The ERG agrees with the conclusions from the company that none of the selected studies were relevant for the decision problem given the in- and exclusion criteria defined by the company.

5.1.5 Objective and searches of health-related quality of life review

No health-related quality of life (HRQoL) data were collected in either the phase II trial[3] or the RECOURSE trial.[2] Therefore, the company conducted a systematic review to identify HRQoL studies from the published literature relevant to the decision problem.

The CS states that a systematic review was conducted to identify HRQoL studies from the published literature relevant to the decision problem; in particular, studies reporting European Quality of Life-5 Dimensions (EQ-5D) health state utility values (in line with the NICE preferred method) relating to patients with advanced/mCRC receiving third line treatment or beyond were considered eligible for inclusion.

The search strategies reported in Appendix 10 of the CS were identical to those reported in Appendix 6 for the cost effectiveness review[23], and the database search results reported here did not correspond with those reported in Section 5.4.2 and Figure 35 (flow chart) of the CS.[1] The ERG asked for clarification that the correct search strategies for identifying HRQoL studies had been reported.[20] In response the company stated that although the captions for MEDLINE and Embase were incorrect, the 'search strategies themselves were correct'.[9] The captions for the MEDLINE and Embase search strategies provided were actually identical to those already reported in the CS.[1] The search strategies reported in Appendix 10 were designed to identify cost effectiveness studies, not HRQoL studies. Without full details of the HRQoL search strategies the ERG was unable to assess their quality.

The company reported additionally searching conference proceedings: ASCO, ESMO and ISPOR. The CS did not provide full details of the conference proceedings searches. Full details of the conference proceedings searches for the utility review were provided in response[9] to the request for clarification.[20]

A search of other economic resources, such as the CEA Registry and ScHARRHUD, for cost-utility analyses might have provided additional useful HRQoL data.

The list of excluded studies reported in Table 7 (Section 10.7 of the CS) were identical to those excluded studies reported for the cost effectiveness review in Table 2 (Section 6.7). In response to the request for clarification[20] asking if the list of excluded studies was correct, the company reported that the list was correct.[9]

5.1.6 Inclusion/exclusion criteria used in the study selection

In the CS,[1] it is stated that studies reporting EQ-5D health state utility values (in line with the NICE preferred method) relating to patients with advanced/mCRC receiving third line treatment or beyond were considered eligible for inclusion.

ERG comment: The in- and exclusion criteria seem appropriate.

5.1.7 Included/excluded studies in the health-related quality of life review

The company identified a total of 547 papers through the electronic searches. After removal of 83 duplicates and exclusion of 436 papers after title and abstract review, 28 full papers were reviewed. Full paper reviewing resulted in four relevant papers for final inclusion (see Figure 35 of the CS[1]).

No additional relevant publications were identified via hand searching. A full list of studies excluded on the basis of full publication review is available in Appendix 10 of the CS along with a rationale for exclusion.

ERG comment: The rationales for excluding studies after full paper reviewing seem appropriate (see Table 7 of Appendix 10 of the CS[1]).

5.1.8 Conclusions of the health-related quality of life review

The company concluded that there were two HRQoL studies[29-31] that may meet the requirements of the NICE reference case. However, assessment of consistency with the NICE reference case and quality

assessment were hampered by limited reporting of details regarding methods of elicitation and valuation, the patient recruitment process, eligibility criteria and response rates (see Tables 58 and 59 of the CS[1]).

ERG comment: The ERG agrees with the conclusions from the company that two out of the four included studies[29-31] might potentially be consistent with the NICE reference case. Nevertheless, it is unclear whether these studies meet the requirements of the NICE reference case on all aspects. Moreover, the company was unclear why the study by Siena et al. (i.e. the CORRECT study)[29, 30] was preferred as the source for HRQoL data above the study by Chang et al.[31] which might potentially be consistent with the NICE reference case. This was clarified by the company in the clarification letter[9] by stating that Chang et al.[31] *"did not provide health-state specific utility values for use in the model"* and that is was *"only abstracts and did not present utility values by progression status"*.[9] The ERG thinks this is reasonable.

Additionally, the ERG identified relevant studies for the estimation of health state utilities (see Section 5.2.8) that were not in the list of excluded papers after full reading, and therefore presumably not identified in the systematic review by the company. As a result, the sensitivity of the systematic review may be questioned, and other potentially relevant studies may be overlooked. This lack of sensitivity might be because the company did not specifically search for relevant studies on health-related quality of life, but instead used the search for relevant cost effectiveness studies to identify model inputs for health-related quality of life.

5.2 Summary and critique of company's submitted economic evaluation by the ERG

Table 5.2: Summary of the company's economic evaluation (with signposts to CS)

	Approach	Source / Justification	Signpost (location in CS)
Model	A partitioned-survival model was constructed to evaluate the cost-effectiveness of T/T compared with BSC in adult patients with mCRC who have been previously treated with, or are not considered candidates for, available therapies.		5.2.2 (pg. 130)
States and events	The model was based on disease progression, consisting of the health states pre-progression, post-progression and death.	Health states were selected according to the clinical pathway of care and comparable to the structure used in other late-stage cancer models.	5.2.2 (pg. 130)
Comparators	Best supportive care.	As there is currently no recommended treatment for patients in the population covered by the anticipated T/T licence, the company selected BSC as the comparator.	5.2.3 (pg. 131)

	Approach	Source / Justification	Signpost (location in CS)
Population	Adult patients with mCRC who have been previously treated with, or are not considered candidates for, available therapies including fluoropyrimidine-, oxaliplatin- and irinotecan- based chemotherapy, anti-VEGF biological therapies, and anti-EGFR therapies.	The population in the analysis is similar to the population in the scope but slightly different from the populations in the phase II trial and RECOURSE study that were used to inform input parameters.	5.2.1 (pg. 129-130)
Treatment effectiveness	The intervention was defined by the company is an orally administered combination of trifluridine, a thymidine-based nucleic acid analogue, and a thymidine phosphorylase inhibitor, tipiracil hydrochloride. It is administered at a dose of 35mg/m2 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.	The intervention defined in the NICE final scope was 'fixed dose combination of trifluridine and tipiracil hydrochloride'.	5.2.3 (pg. 131)
Adverse events	The company incorporated costs of adverse events if they were actively treated in the NHS, as verified with clinical and medical oncologists.	RECOURSE trial	5.5.4 (pg. 161-163)
Health related Quality of Life	Health related quality of life information was not collected in the phase II study and the RECOURSE trial. Estimates for health state utilities were based on literature and assumptions. Disutilities for adverse events were not explicitly modelled, and based on assumption.	Health state utilities for pre and post progression were based on the average of values reported in the CORRECT study[30] and the company submission of TA176[32].	5.4 (pg. 148 - 155)
Resource utilisation and costs	Drug costs were estimated from the RECOURSE trial, taking into account dosage (based on BSA), dose reduction, treatment delay, and time on treatment. The weighted average cost in the third cycle was [redacted] at list price. MRU costs were based on expert opinion and included oral chemotherapy day case attendance and health home visitor for patients	TA794[28] for mCRC, RECOURSE trial, and expert opinion. Unit costs for the regularly scheduled follow-up procedures were determined using the NHS Reference Costs, 2014-15. End-of-life care costs were taken from a modelling study by Round et al.[33]	5.5 (pg. 155 - 165)

	Approach	Source / Justification	Signpost (location in CS)
	treated with T/T (£203). Patients receiving BSC had a medical oncologist outpatient consultation and a health home visitor (£182). For all patients GP home consultation, community nurse specialist visits, district nurse visits, and GP surgery visits were included in post-progression (£193). The RECOURSE trial data was used to estimate the average cost of post-progression treatment per patient, which was £1,549 for T/T and £1,487 for BSC, but were incorporated on average for all patients (£1,528). End-of-life care costs included health care, social care and charity care. The total end-of-life care cost of £6,910 was applied in the model as a lump sum upon death for both arms. The company incorporated costs of adverse events if they were actively treated in the NHS. These events were included at rates observed from the RECOURSE trial resulting in £923 for T/T and £426 for BSC.		
Discount rates	3.5 % for utilities and costs	According to NICE reference case	5.2.2 (pg. 131)
Sub groups	Subgroup analysis is not considered in the de novo analysis, given the size of the patient population and that, in RECOURSE, T/T was associated with a clinically relevant prolongation in OS in all treatment subgroups.		5.9 (pg. 188)
Sensitivity analysis	Deterministic and probabilistic sensitivity analyses were conducted. The model was mainly sensitive to changes in health related quality of life inputs and survival estimates.		5.8 (pg. 175 – 188)

BSA = body surface area; BSC = best supportive care; EGFR = epidermal growth factor receptor; GP = general practitioner; mCRC = metastatic colorectal cancer; MRU = medical resource utilisation; NHS =

	Approach	Source / Justification	Signpost (location in CS)
National Health Service; NICE = National Institute for Health and Care Excellence; OS = overall survival; T/T = trifluridine/tipiracil; VEGF = vascular endothelial growth factor			

5.2.1 NICE reference case checklist (TABLE ONLY)

Table 5.3: NICE reference case checklist

Elements of the economic evaluation	Reference Case	Included in submission	Comment on whether de novo evaluation meets requirements of NICE reference case
Population	As per NICE scope	Y	Population in the CS is per NICE scope, but may differ slightly from population in trials on which evaluation is based (see 5.2.3).
Comparator(s)	Therapies routinely used in the NHS, including technologies regarded as current best practice	Y	T/T is evaluated against best supportive care.
Type of economic evaluation	Cost-effectiveness analysis	Y	
Perspective on costs	NHS and Personal Social Services (PSS)	Y	PSS costs are not reported.
Perspective on outcomes	All health effects on individuals	Y	
Time horizon	Sufficient to capture differences in costs and outcomes	Y	Time horizon of 10 years is effectively lifetime as <1% of patients are still alive (5.2.5).
Synthesis of evidence in outcomes	Systematic review	Partly	Ideally, a dedicated systematic review would also have been performed to inform the model structure, quality of life and resource use.
Measure of health effects	Quality adjusted life years (QALYs)	Y	
Source of data for measurement HRQoL	Described using a standardised and validated instrument	Y	HRQoL data were not collected in the phase II and the phase III clinical trial.

Elements of the economic evaluation	Reference Case	Included in submission	Comment on whether de novo evaluation meets requirements of NICE reference case
Source of preference data for valuation of changes in HRQoL	Time-trade off or standard gamble	Y	
Discount rate	An annual rate of 3.5% on both costs and health effects	Y	
Equity weighting	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	Y	
Probabilistic modelling	Probabilistic modelling	Y	BSA was included in the PSA as a stochastic parameter.
Sensitivity analysis		Y	A range of sensitivity analyses were performed.

BSA = body surface area; CS = company submission; HRQoL = health-related quality of life; NHS = National Health Service; NICE = National Institute for Health and Care Excellence; PSA = probabilistic sensitivity analysis; PSS = personal social services; QALY = quality-adjusted life year; T/T = trifluridine/tipiracil

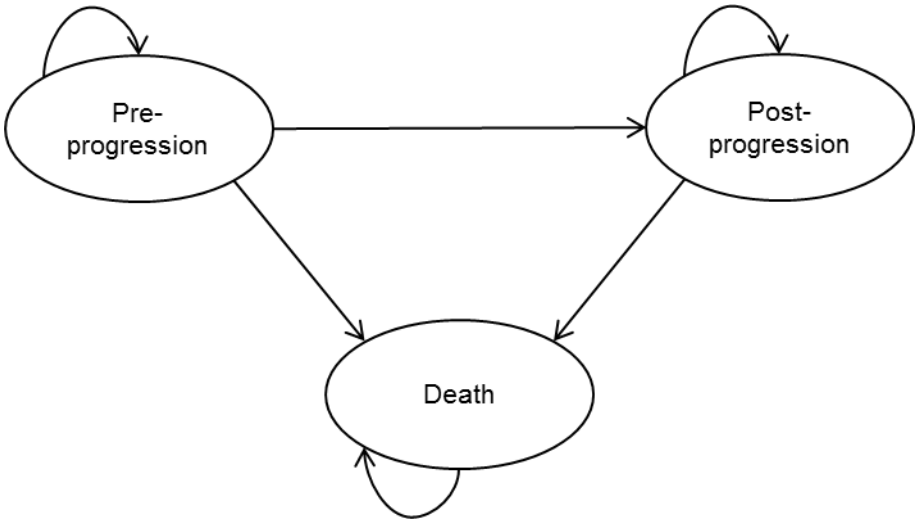
5.2.2 Model structure

An excel-based partitioned-survival model was constructed to evaluate the cost effectiveness of T/T compared with BSC in adult patients with mCRC who have been previously treated with, or are not considered candidates for, available therapies. The model was based on disease progression, consisting of the health states pre-progression, post-progression and death (Figure 5.1). Health states were selected according to the clinical pathway of care and comparable to the structure used in other late-stage cancer models.

All patients enter the model in the pre-progression state. Patients may transition between health states based on PFS curves that were fitted to the clinical trial data. Patients that have progressed to the post-progression state are not permitted to transition back to the pre-progression state. Patients may transition to the death state from any health state. The model structure is identical for patients treated receiving T/T or BSC.

Because of the poor prognosis of patients, a daily cycle length was applied to ensure the accuracy of survival estimates. A longer cycle length was considered to be inappropriate due to the kinks in the curve caused by the frequency of progression assessment in the clinical trials. Consequently, a half-cycle correction was not deemed to be required.

Figure 5.1: Model structure
(Based on Figure 29 of the CS[1])



ERG comment: Ideally, following the NICE reference case, a systematic approach, including a review, should have been performed to inform the model structure. Nevertheless, the ERG agrees that the chosen model structure, daily cycle and the absence of a half-cycle correction are appropriate for this decision problem.

5.2.3 Population

The company reported that following the anticipated licence, T/T was indicated for the treatment of adult patients with mCRC who have been previously treated with, or are not considered candidates for, available therapies including fluoropyrimidine-, oxaliplatin- and irinotecan-based chemotherapy, anti-VEGF biological therapies, and anti-EGFR therapies.[34] The company considered this population to be reflective of the population discussed in the decision problem and the scope, as well as in the clinical trials from which efficacy data are derived to inform the model (see Table 5.4). In line with the licence, T/T is expected to be used from the third line onwards.

Table 5.4: Populations

NICE final scope	Company (following anticipated licence)	Phase II RCT	RECOURSE
Adults with metastatic colorectal cancer whose disease has progressed after standard therapies or for whom standard therapies are unsuitable.	Adult patients with mCRC who have been previously treated with, or are not considered candidates for, available therapies including fluoro-pyrimidine-, oxaliplatin- and irinotecan- based chemotherapy, anti-VEGF biological therapies, and anti-EGFR therapies.	Adult patients aged ≥20 years with histologically or cytologically confirmed unresectable metastatic colorectal adenocarcinoma with a previous treatment history of ≥2 regimens of standard chemo-therapy [3]	Adult patients aged ≥18 years with biopsy-documented adenocarcinoma of the colon or rectum who had received ≥2 prior regimens of standard chemo-therapy [2]
EGFR = epidermal growth factor receptor; NICE = National Institute for Health and Care Excellence; RCT = randomised controlled trial; VEGF = vascular endothelial growth factor			

ERG comment: The ERG notes that the populations described in the NICE final scope[14], including patients with mCRC for whom standard therapies are 'unsuitable', seems approximately similar to the

population described by the company, following the anticipated licence, but differs slightly from populations in the trials, which were used to inform the model (Table 5.4). Consequently, following the licence it may be possible that patients not represented in the trial receive this medication. This includes patients *"for whom standard therapies are unsuitable"*. It remains unclear in which direction this discrepancy would influence the outcomes.

5.2.4 Interventions and comparators

The intervention defined in the NICE final scope was *"fixed dose combination of trifluridine and tipiracil hydrochloride"*.[14] The intervention was defined by the company as an orally administered combination of trifluridine, a thymidine-based nucleic acid analogue, and a thymidine phosphorylase inhibitor, tipiracil hydrochloride. It is administered at a dose of 35 mg/m^2 twice daily, five days a week, with two days of rest, for two weeks, followed by a 14-day rest period. This treatment cycle is repeated every four weeks.[34] Following the anticipated licence and the RECOURSE trial protocol, T/T treatment is continued until determination of RECIST-defined disease progression, clinical progression, the development of severe adverse events, withdrawal from the study, death, or a decision by the treating physician that discontinuation would be in the patient's best interest.[2, 35]

As there is currently no recommended treatment for patients in the population covered by the anticipated T/T licence, the company selected BSC as the comparator, in line with the phase II trial and RECOURSE.[2, 3]

ERG comment: The ERG agrees with the selected intervention and comparator. The ERG asked the company to provide the definition of BSC in the trials. The company responded that BSC was defined as follows[9]:

- Phase II trial: All necessary support was provided to patients, with the exception of concomitant use of other anti-cancer drugs or other investigational drugs.
- RECOURSE: All necessary support was provided to patients which included permitted concomitant medications and therapies and study medication. All patients received the best supportive care available but were not to receive other investigational antitumour agents or antineoplastic chemotherapy, hormonal therapy, or immunotherapy. Palliative radiotherapy was not permitted while the patient was receiving study treatment. If used concomitantly with study medication, antiviral drugs that are human thymidine kinase substrates (e.g. stavudine, zidovudine, telbivudine) were to be used with caution because such drugs may theoretically compete with the effect of trifluridine/tipiracil, i.e. trifluridine, for activation via thymidine kinases.

Based on these definitions it is uncertain whether BSC as provided in the trial is representative for the UK.

5.2.5 Perspective, time horizon and discounting

The economic evaluation used the perspective of the National Health Service (NHS). Utilities and costs were discounted at 3.5% over a time horizon of 10 years. The company justified the time horizon of 10 years as being effectively lifetime as less than 1% of patients are still alive (Table 48 of the CS).[1]

ERG comment: The ERG agrees with the chosen discounting rates and agrees that 10 years is effectively a lifetime horizon in this population.

5.2.6 Treatment effectiveness and extrapolation

Data sources and pooling

Overall survival (OS) and progression-free survival (PFS) estimates were obtained from RECOURSE[2] and the phase II trial[3]. The definitions of these endpoints in each trial are provided in Table 15 of the CS[1] (Table 5.5). RECOURSE is an international randomised controlled phase III trial performed in Europe, Australia, the United States and Japan while the phase II trial included only Japanese patients. Both trials used a 2:1 randomisation scheme of T/T+BSC versus placebo+BSC. Trial data were considered mature with 89% and 72.9% of the patients being deceased in RECOURSE and the phase II trial, respectively.[1] Updated OS data from RECOURSE were available, which means that OS data are based on the last known alive date instead of being capped at the 571th death as provided in the publication of the trial (original data).[2]

Table 5.5: Definition of OS and PFS in RECOURSE and the phase II clinical trial
(Based on Table 15 of the CS[1])

Outcomes	Definition in phase II trial	Definition in RECOURSE
Primary outcome: Overall survival (OS)	Time between randomisation and death from any cause or the date of last follow-up	Time (in months) between randomisation and death from any cause
Secondary outcome: Progression-free survival (PFS)	Defined as the time (in months) from randomisation to the **date that the patient's condition reached progressive disease (PD)**. If the patient died before reaching PD, the date of death was considered the date PD was reached. […]	Defined as the time (in months) from the date of randomisation until the **date of the investigator-assessed radiological disease progression** or death due to any cause. […]
CS = company submission; OS = overall survival; PD = progressive disease; PFS = progression-free survival		

In the company base case analysis, effectiveness data from both trials have been pooled (updated RECOURSE data + phase II clinical trial). According to the company, pooling provided a "*meaningful increase in the number of placebo-treated patients*".[1] No detail on the pooling procedure was provided in the cost effectiveness assessment part of the CS.[1] Effectiveness data from RECOURSE only (original and updated data) and from the phase II clinical trial only were used in sensitivity analyses. Results of those analyses are provided in Section 5.2.11 of the current report.

Transition probabilities between health states were based on the area under the curve (i.e. partitioned survival model) from OS and PFS survival curves. The OS curve estimated the proportion of patients which were 'alive' and the PFS curve estimated the proportion of patient which remained in the 'pre-progression' health state, at any point in time. The proportion of patients with progression was estimated by the difference between 'alive' and 'pre-progression' patients. The proportion of deceased patients was estimated by '1-proportion of patients still alive'.

ERG comment: As can be seen in Table 5.5, the definitions for PFS were not identical in both trials, which could have led to different assessment of progression between trials. Furthermore, the trial populations are slightly different. These two factors may have led to heterogeneity between the trials, but did not completely hamper pooling. For a more extensive discussion on reasons to pool the data from both trials, the ERG refers to Section 4.15 of the current report.

Even though pooling the trials seems reasonable, the methods were not clearly described in the CS. The ERG asked clarification on how pooling was performed and the company referred to the meta-analysis presented in Section 4.9 of the CS[1], without providing additional details. As a result, the ERG was unable to critically assess whether the pooling procedure was reasonable (see Section 4.15 of this report). In order for the ERG to critically assess the pooling, the ERG would have liked to receive a comparison of the current meta-analysis (not stratified by trial) with a meta-analysis in which stratification by trials was performed. If the results of both meta-analyses would have been similar, the ERG would prefer the current meta-analysis to be used in the cost effectiveness model. Without this information, the ERG prefers using a more conservative assumption in its base case analysis by using RECOURSE data only. However, since there are no fundamental arguments which prevent the two trials from being pooled, besides the lack of clarity of the methodology, the ERG also presents its base case analysis based on the pooled effectiveness estimates from both trials.

PFS and OS were the only pooled data while other estimates, such as adverse event rates, time on treatment and dose reductions were based on RECOURSE only. The ERG did not understand the rationale behind this choice and asked for pooled estimates for these other estimates (i.e. adverse event rates, time on treatment and dose reductions). The company provided an updated model containing pooled estimates for adverse event rates, time on treatment and dose reductions with its response to the ERG clarification letter. The ERG used this updated model in its analyses.

Model selection for progression-free survival and overall survival
Different stratified by treatment and unstratified parametric survival models were compared to select survival models to represent OS and PFS in the cost effectiveness analysis. In the stratified models, two curve fits were produced for T/T and BSC separately while unstratified models contained a covariate representing the treatment arm. The following candidate survival models were examined:

- Log-logistic (stratified and unstratified)
- Generalised gamma (stratified and unstratified)
- Log-normal (stratified and unstratified)
- Weibull (stratified and unstratified)
- Gompertz (stratified and unstratified)
- Exponential (unstratified)
- Extreme value (stratified and unstratified)

The most suitable survival model was chosen based on the Akaike Information Criterion (AIC) goodness of fit statistics and visual examination. Goodness of fit statistics for PFS and OS survival models are presented in Table 5.6. The curve fits of the different candidate survival models are provided in Appendix 7 of the CS.[1]

Table 5.6: Progression-free survival and overall survival – goodness of fit statistics
(Based on Tables 49 and 50 of the CS[1])

Model	AIC (PFS)	Goodness of fit ranking (PFS)	AIC (OS)	Goodness of fit ranking (OS)
Stratified log-logistic	**9,331**	**1**	**10,898**	**2**
Stratified generalised gamma	9,352	2	10,901	4
Stratified log-normal	9,356	3	10,905	6

Model	AIC (PFS)	Goodness of fit ranking (PFS)	AIC (OS)	Goodness of fit ranking (OS)
Log-logistic	9,385	4	10,896	1
Generalised gamma	9,403	5	10,899	3
Log-normal	9,407	6	10,903	5
Stratified Weibull	9,589	7	10,958	8
Weibull	9,607	8	10,957	7
Stratified Gompertz	9,754	9	11,041	10
Gompertz	9,759	10	11,040	9
Exponential	9,773	11	11,079	13
Extreme value	9,855	12	11,063	12
Stratified extreme value	9,857	13	11,060	11

AIC = Akaike Information Criterion; CS = company submission; OS = overall survival; PFS = progression-free survival

For PFS, the stratified log-logistic model provided the lowest AIC and had a good visual fit. Therefore, it was chosen to represent PFS in the base case analysis (Figure 5.2). For OS, the unstratified log-logistic model had the best AIC estimate. However, the stratified log-logistic model was chosen to represent OS in order to be consistent with the selected model for PFS. Moreover, the stratified log-logistic model provided a good visual fit to the OS Kaplan-Meier curve (Figure 5.3) and was the second best-fitting model according to the AIC (with two AIC points difference with the unstratified log-logistic model). Another argument of the company to use stratified models was the uneven randomisation in both trials (2:1).[1] The chosen survival models for the base case analysis are bold printed in Table 5.6 above. The influence of using alternative survival models was investigated in sensitivity analyses. Results of these sensitivity analyses are presented in Section 5.2.11 of the current report.

Figure 5.2: Stratified log-logistic survival curve for PFS (two years)
(Based on Figure 30 of the CS[1])

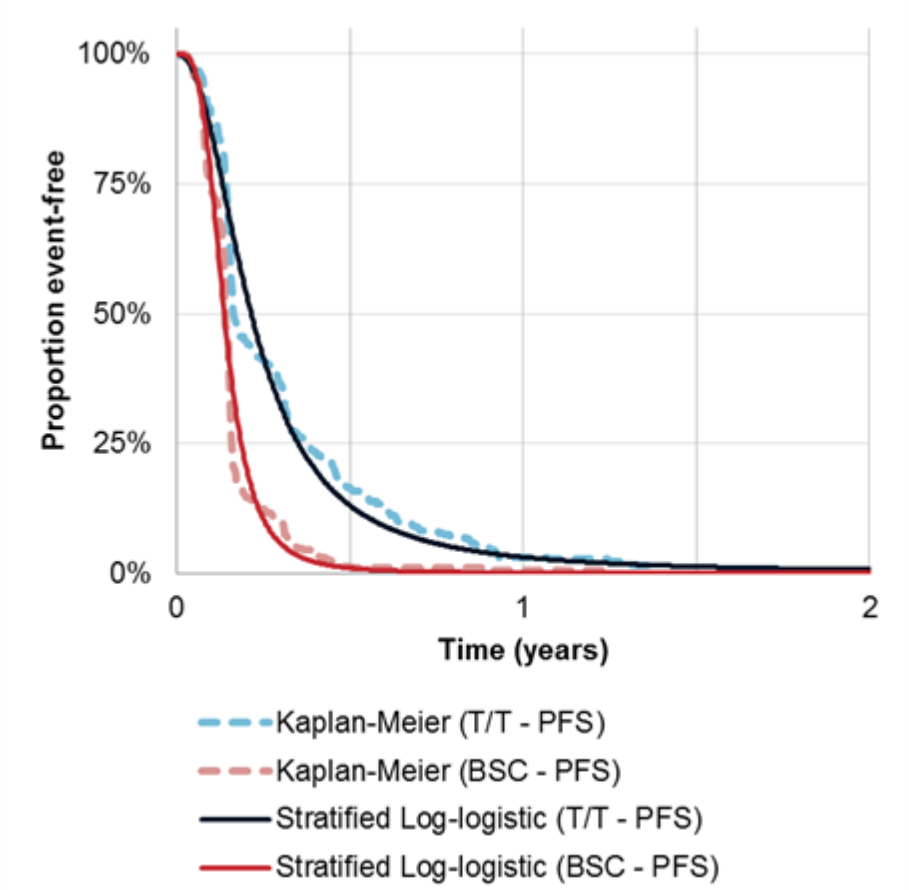
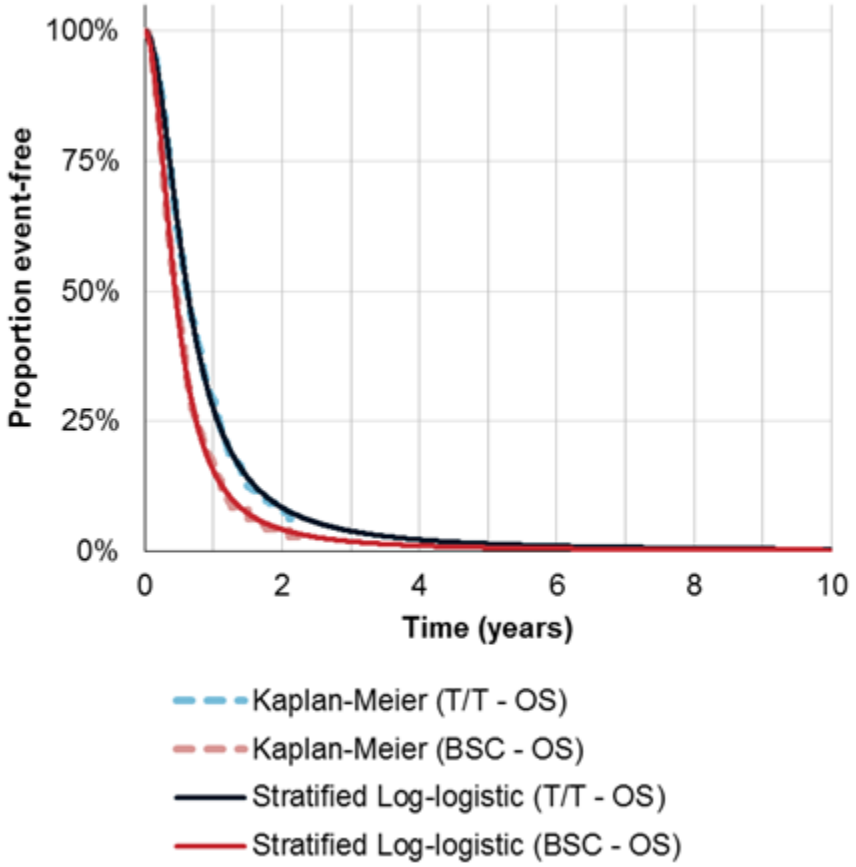


Figure 5.3: Stratified log-logistic survival curve for OS (10 years)
(Based on Figure 32 of the CS[1])



ERG comment: The following issues concerning survival model selection are raised by the ERG: log-cumulative hazard or quantile-quantile (QQ) plots were not used to decide on using stratified or unstratified models, uneven randomisation as an argument for the selection of a stratified model, AIC calculations for stratified models were unclear.

Log-cumulative hazard or QQ plots were not used to decide on using stratified or unstratified models
The use of stratified or unstratified model should be based on a visual examination of log-cumulative or QQ plots, as recommended by the NICE Decision Support Unit (DSU) on survival analysis.[36] This step was missing in the model selection process described in the CS. Therefore, the ERG asked the company to provide these plots for all survival models presented in the CS. In its response to the ERG clarification letter, the company provided the log-cumulative hazard plots for the PFS and OS of the pooled, RECOURSE and phase II population respectively.[9] The QQ plots of the different survival models were not presented. The ERG examined the log-cumulative hazard plots from RECOURSE data only because pooling was not deemed suitable in the current assessment based on above-mentioned arguments. The log-cumulative hazard plots, for the updated RECOURSE data are displayed in Figures 5.4 (OS for the RECOURSE population ('Updated OS')) and 5.5 (PFS for the RECOURSE population).

Figure 5.4: Log-cumulative hazard plot for OS – RECOURSE population
(Based on Figure 3 of the response to request for clarification[9])

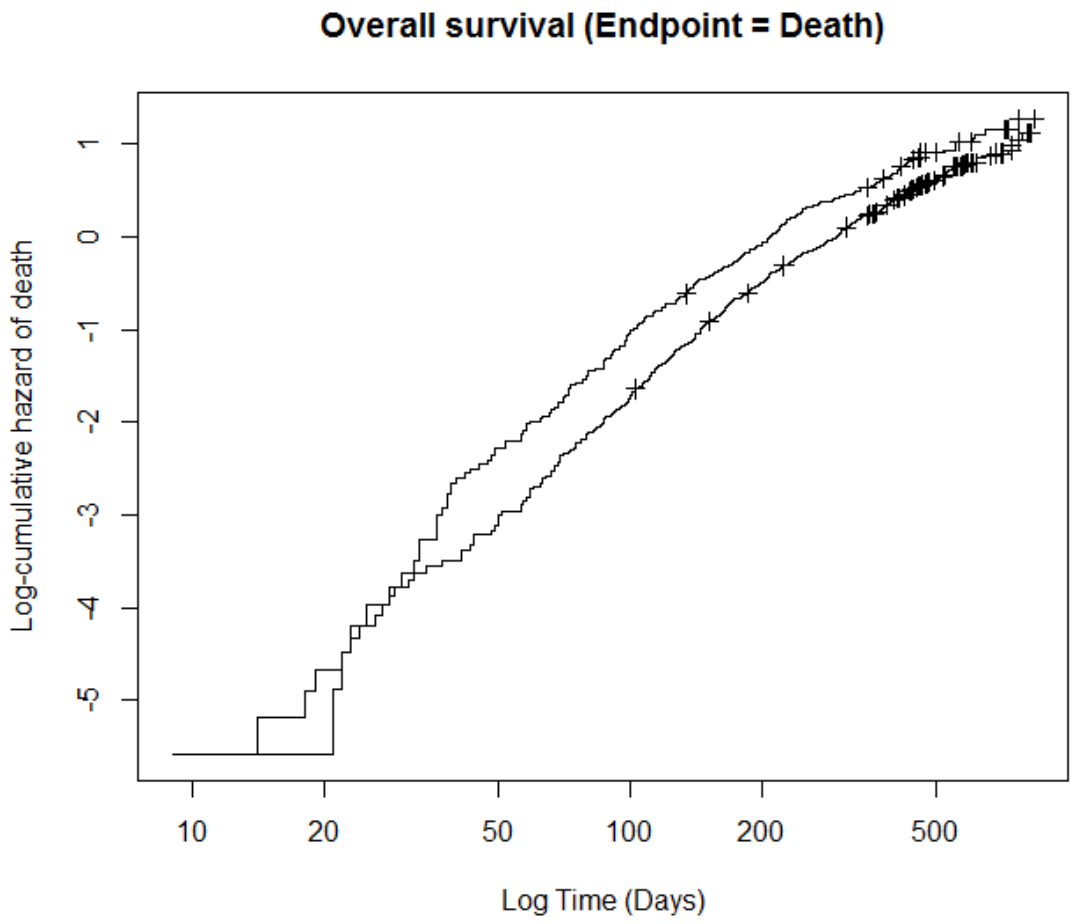
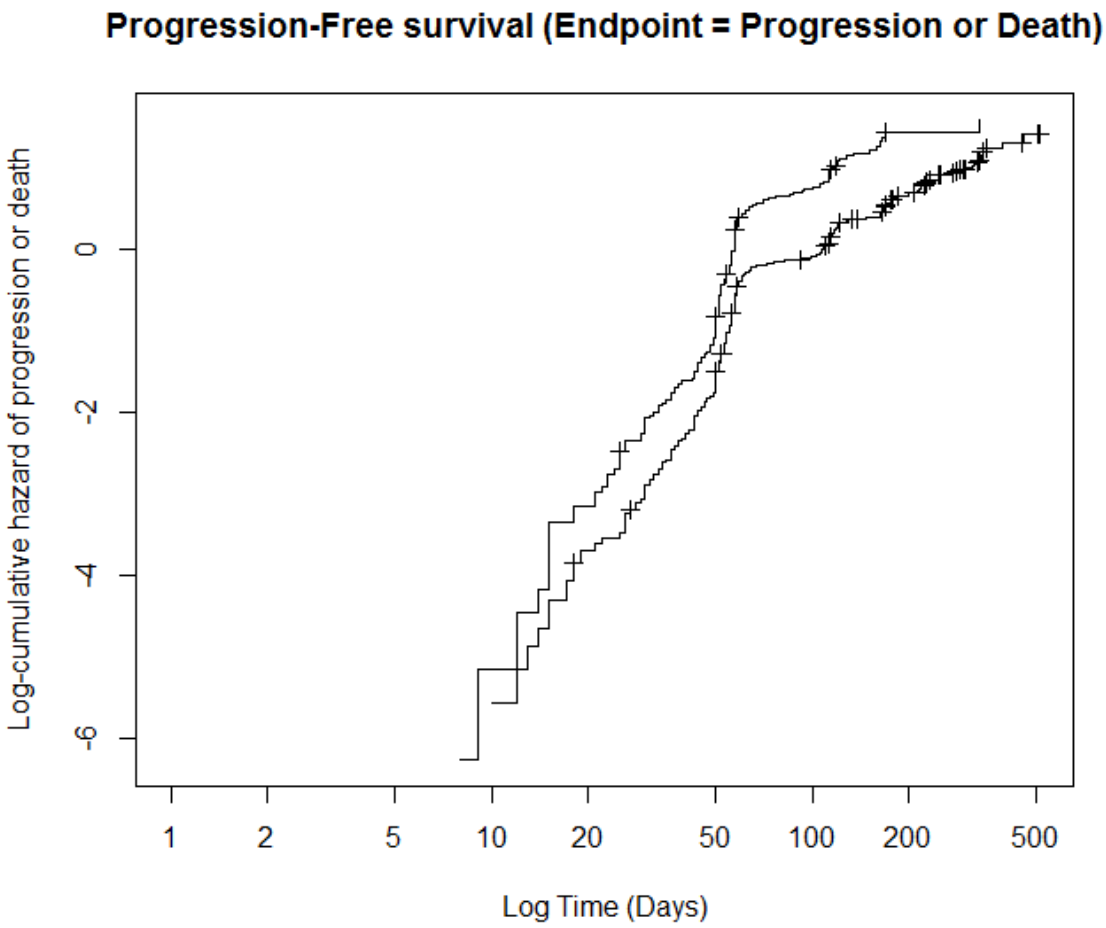


Figure 5.5: Log-cumulative hazard plot for PFS – RECOURSE population
(Based on Figure 6 of the response to request for clarification[9])



Since log-cumulative hazard plots (Figures 5.4 and 5.5) for the RECOURSE population were reasonably parallel, the ERG preferred using unstratified survival models in its base case analysis.

Uneven randomisation as an argument for the selection of a stratified model
Furthermore, uneven randomisation was an argument for the selection of stratified models instead of unstratified models. This was however unclear to the ERG and clarification was asked on this point. The company responded with the following: *"Unequal randomisation (in this case 2:1) implies that unstratified parametric survival models will inherently utilise a relatively larger proportion of patients in the larger patient group (in this case, patients receiving trifluridine/tipiracil) compared with the smaller patients group (in this case, patients receiving placebo) in the estimate of the associated parametric curve parameters."*[9] Because stratified models were deemed suitable, this argument was not taken into account during model selection by the ERG.

AIC calculations for stratified model were unclear
It was unclear to the ERG how the AIC were calculated for stratified models since they presumably led to two curve fits. Comparing AIC from unstratified and stratified survival models consequently leads to a penalty for stratified models since unstratified models contain a covariate that stratified model do not contain. For these reasons, the ERG asked the company to clarify how unique AIC for stratified models were obtained. In its response to the clarification, the company stated that *"AIC scores were obtained for the stratified models using the same methodology as per the unstratified models"*.[9] Pragmatically,

the same R function was used to calculate the AIC of stratified and unstratified models. Calculations seemed to be performed correctly according to the ERG.

In order to select the survival models to represent PFS and OS in its base case cost effectiveness analysis, the ERG followed the algorithm provided by the DSU on survival analysis.[36] First, based on the examination of the log-cumulative hazard curves of the RECOURSE population, the ERG does not agree with the choice of stratified model for OS and PFS and preferred using unstratified models since the curves in the plots (Figures 5.4 and 5.5) were reasonably parallel. Second, based on the AIC and visual examination, the ERG thinks that the most appropriate model for both OS and PFS would be the unstratified log-logistic models. These models were used in the ERG base case analysis. Results of this analysis are provided in Chapter 6 of the current report.

5.2.7 Adverse events

The company's cost effectiveness model includes all 'common' adverse events (AEs) based on AEs incidence rates from the RECOURSE trial. 'Common' was defined as AEs that occurred in 10% or more of the patients receiving T/T and which occurred in a higher proportion of patients receiving T/T than in patients receiving BSC. The incidence rates of AEs from the RECOURSE trial are listed in Table 5.7. The bold-printed percentages are the ones that are explicitly used in the model to calculate AEs treatment costs. More details on the costing procedure of AEs are provided in Section 5.2.9 of the current report. No distinction was made between AEs occurring before or after progression.

Table 5.7: Adverse events rates with absolute risk reduction (ARR) from RECOURSE
(Based on Tables 43, 44 and 57 of the CS[1])

	Trifluridine/tipiracil		BSC		ARR % (any grade)	ARR % (grade ≥3 AEs)
	% of events (any grade)	% of grade ≥ 3 AEs	% of events (any grade)	% of grade ≥ 3 AEs		
Any event	98.3	69.4	93.2	51.7	-5.1	-17.7
Any serious event	NA	29.6	NA	33.6	NA	3.9
Nausea†	48.4	**1.9**	23.8	**1.1**	-24.6	**-0.7**
Vomiting†	**27.8**	**2.1**	**14.3**	**0.4**	**-13.4**	**-1.7**
Decreased appetite†	39.0	**3.6**	29.4	**4.9**	-9.6	**1.3**
Fatigue†	35.3	**3.9**	23.4	**5.7**	-11.9	**1.7**
Diarrhoea†	**31.9**	**3.0**	**12.5**	**0.4**	**-19.4**	**-2.6**
Abdominal pain†	21.2	**2.4**	18.5	**3.8**	-2.7	**1.3**
Fever†	**18.6**	**1.3**	**14.0**	**0.4**	**-4.6**	**-0.9**
Asthenia†	18.2	**3.4**	11.3	**3.0**	-6.9	**-0.4**
Febrile neutropenia**	**3.8**	**3.8**	**0.0**	**0.0**	-3.8	**-3.8**
Stomatitis**	8.1	**0.4**	6.4	**0.0**	-1.7	**-0.4**
Hand-foot syndrome**	2.3	0.0	2.3	0.0	0.0	0.0
Cardiac ischaemia** ‡	**0.4**	**0.2**	**0.4**	**0.4**	**0.0**	**0.2**
Neutropenia§	67.8	**37.9**	0.8	**0.8**	-67.0	**-37.1**
Leucopenia§	77.1	**21.4**	4.6	**4.6**	-72.5	**-16.8**
Anaemia§	76.5	**18.2**	33.1	**33.1**	-43.4	**14.9**

	Trifluridine/tipiracil		BSC		ARR % (any grade)	ARR % (grade ≥3 AEs)
	% of events (any grade)	% of grade ≥ 3 AEs	% of events (any grade)	% of grade ≥ 3 AEs		
Thrombocytopenia[§]	42.2	**5.1**	8.0	**8.0**	-34.3	**2.9**
Increase in alanine aminotransferase level[§]	24.0	1.9	26.6	26.6	2.7	24.7
Increase in aspartate aminotransferase level[§]	21.9	4.4	34.7	34.7	12.8	30.3
Increase in total bilirubin[§]	35.4	8.6	26.3	26.3	-9.0	17.8
Increase alkaline phosphatase level[§]	39.0	8.0	45.0	45.0	6.1	37.1
Increase in creatinine level[§]	13.5	0.9	12.2	12.2	-1.3	11.2

Trial data from Mayer et al. 2015[2]. Calculations not possible when absolute risk in placebo group = 0.

All adverse events were graded according to National Cancer Institute Common Terminology Criteria for Adverse Events, version 4.03.

[†] Adverse events of any grade that are listed as most common occurred in 10% or more of patients in the trifluridine/tipiracil group and in a greater percentage in that group than in the placebo group.

[**] Events associated with fluoropyrimidine treatment.

[‡] Events included acute myocardial infarction, angina pectoris, and myocardial ischaemia.

[§] The denominator for the percentage of patients with laboratory abnormalities is the number of patients with at least one post baseline measurement during treatment.

Bold-printed percentages are the ones that are explicitly used in the model to calculate AEs treatment costs.

AE = adverse event; ARR = absolute risk reduction; BSC = best supportive care; CS = company submission; NA = not applicable

ERG comment: It was unclear to the ERG why only RECOURSE data (and not a pooled estimate from RECOURSE and the phase II trial) were used for AEs incidence rates, especially because PFS and OS in the company base case analysis were based on pooled evidence of both clinical trials. In its clarification letter, the ERG asked for a pooled analysis of AEs incidence rates, based on both trials.[20] The company provided new AEs incidence rates based on both trials. Adverse events were included in this analysis based *"upon the most frequently observed adverse events (defined as occurring with a frequency of at least 3% in the safety population) in the Phase II trial, as reported in the publication by Yoshino et al. (2009). The rates presented in this publication have been selected for inclusion using the same criteria as per the adverse events from the RECOURSE study, which were taken from the publication by Mayer et al. (2015)."* The pooled AEs incidence rates and reasons for exclusion of specific AEs from the costing procedure are presented in Table 5.8.

Table 5.8: Adverse events rates with absolute risk reduction (ARR) from RECOURSE
(Based on Table 4 of the response to request for clarification[9])

Grade 1 or 2 adverse events	**Trifluridine/tipiracil**	**BSC**	**Excluded?**
Diarrhoea	43/113 (38%)	12/57 (21%)	
Febrile neutropenia	5/113 (4%)	0	
Vomiting	38/113 (34%)	14/57 (25%)	
Grade 3 or 4 adverse events	**Trifluridine/tipiracil**	**BSC**	**Excluded?**
Neutropenia	57/113 (50%)	0	
Leucopenia	32/113 (28%)	0	
Anaemia	19/113 (17%)	3/57 (5%)	
Lymphopenia	11/113 (10%)	2/57 (4%)	Yes[a]
Thrombocytopenia	5/113 (4%)	0	
Fatigue	7/113 (6%)	2/57 (4%)	
Diarrhoea	7/113 (6%)	0	
Nausea	5/113 (4%)	0	
Anorexia	5/113 (4%)	2/57 (4%)	Yes[b]
Febrile neutropenia	5/113 (4%)	0	
Vomiting	4/113 (4%)	0	
Reasons for exclusion:			
a: <1% of patients in both arms of the RECOURSE trial experienced Grade ≥3 lymphopenia b: Anorexia is not explicitly reported in the RECOURSE trial – the most similar adverse events would be Grade ≥3 "Weight Decreased" or "Decreased Appetite". "Decreased Appetite" is already included within the model, and "Weight Decreased" only occurred in 1 trifluridine/tipiracil patient (and 0 BSC patients).			
ARR = absolute risk reduction			

The updated version of the cost effectiveness model, provided with the response to the ERG clarification letter, included the pooled AEs incidence rates.[9] Results of this analysis are presented in Section 5.2.11 of the current report.

Since the ERG decided not to use pooled estimates in its base case, the ERG used AEs incidence rates from RECOURSE only. However, the ERG would like to note that the grade ≥ 3 AEs rates for the BSC arm reported in Tables 44 and 57 of the CS, and in the company's cost effectiveness model, are not correct for the following AEs:

- Neutropenia
- Leukopenia
- Anaemia
- Thrombocytopenia
- Increase in alanine aminotransferase level
- Increase in aspartate aminotransferase level
- Increase in total bilirubin
- Increase alkaline phosphatase level
- Increase in creatine level

The ERG corrected these rates, by using the rates reported in the RECOURSE publication (Table 2).[2] The corrected AEs rates are given in italics in Table 5.9 besides the other AEs rates used in the ERG base case analysis. Results of the ERG base case are presented in Section 6 of the current report.

Table 5.9: Adverse events rates used in the ERG base case analysis with ARR from RECOURSE (Based on Tables 43, 44 and 57 of the CS[1] and Table 2 of RECOURSE[2])

	Trifluridine/tipiracil		BSC		ARR % (any grade)	ARR % (grade ≥3 AEs)
	% of events (any grade)	% of grade ≥ 3 AEs	% of events (any grade)	% of grade ≥ 3 AEs		
Any event	98.3	69.4	93.2	51.7	-5.1	-17.7
Any serious event	NA	29.6	NA	33.6	NA	3.9
Nausea†	48.4	**1.9**	23.8	**1.1**	-24.6	**-0.7**
Vomiting†	**27.8**	**2.1**	**14.3**	**0.4**	**-13.4**	**-1.7**
Decreased appetite†	39.0	**3.6**	29.4	**4.9**	-9.6	**1.3**
Fatigue†	35.3	**3.9**	23.4	**5.7**	-11.9	**1.7**
Diarrhoea†	**31.9**	**3.0**	**12.5**	**0.4**	**-19.4**	**-2.6**
Abdominal pain†	21.2	**2.4**	18.5	**3.8**	-2.7	**1.3**
Fever†	**18.6**	**1.3**	**14.0**	**0.4**	**-4.6**	**-0.9**
Asthenia†	18.2	**3.4**	11.3	**3.0**	-6.9	**-0.4**
Febrile neutropenia**	**3.8**	**3.8**	**0.0**	**0.0**	-3.8	**-3.8**
Stomatitis**	8.1	**0.4**	6.4	**0.0**	-1.7	**-0.4**
Hand-foot syndrome**	2.3	0.0	2.3	0.0	0.0	0.0
Cardiac ischaemia** ‡	**0.4**	**0.2**	**0.4**	**0.4**	**0.0**	**0.2**
Neutropenia§	67.8	**37.9**	0.8	***0.0***	-67.0	***-37.9***
Leucopenia§	77.1	**21.4**	4.6	***0.0***	-72.5	***-21.4***
Anaemia§	76.5	**18.2**	33.1	***0.0***	-43.4	***-18.2***
Thrombocytopenia§	42.2	**5.1**	8.0	***0.0***	-34.3	***-5.1***
Increase in alanine aminotransferase level§	24.0	1.9	26.6	*0.0*	2.7	*-1.9*
Increase in aspartate aminotransferase level§	21.9	4.4	34.7	*0.1*	12.8	*-4.3*
Increase in total bilirubin§	35.4	8.6	26.3	*0.1*	-9.0	*-8.5*
Increase alkaline phosphatase level§	39.0	8.0	45.0	*0.1*	6.1	*-7.9*
Increase in creatinine level§	13.5	0.9	12.2	*0.0*	-1.3	*-0.9*

Trial data from Mayer et al. 2015[2]. Calculations not possible when absolute risk in BSC group = 0.

All adverse events were graded according to National Cancer Institute Common Terminology Criteria for Adverse Events, version 4.03.

	Trifluridine/tipiracil		BSC		ARR % (any grade)	ARR % (grade ≥3 AEs)
	% of events (any grade)	% of grade ≥ 3 AEs	% of events (any grade)	% of grade ≥ 3 AEs		
[†] Adverse events of any grade that are listed as most common occurred in 10% or more of patients in the trifluridine/tipiracil group and in a greater percentage in that group than in the BSC group.						
[**] Events associated with fluoropyrimidine treatment.						
[‡] Events included acute myocardial infarction, angina pectoris, and myocardial ischaemia.						
[§] The denominator for the percentage of patients with laboratory abnormalities is the number of patients with at least one post baseline measurement during treatment.						
Bold-printed percentages are the ones that are explicitly used in the model to calculate AEs treatment costs. The corrected numbers are printed in Italic.						
AE = adverse event; ARR = absolute risk reduction; CS = company submission; ERG = Evidence Review Group						

5.2.8 Health-related quality of life

No health-related quality of life information was collected in the phase II trial or the RECOURSE study. The company conducted a systematic review to identify health-related quality of life studies from the published literature. Four studies were included: Chan et al.[37], Mittmann et al.[38], Chang et al.[31], and Siena et al.[29]. In Chan et al. and Mittmann et al. the Health Utilities Index Mark III (HUI3) instrument was used to determine utilities. This is not in line with the NICE reference case, and for that reason these studies were not used by the company. It was stated that the abstracts from Chang et al. and Siena et al. *"may meet the NICE requirement"*. Siena et al. was a publication based on data from the CORRECT study of regorafenib monotherapy for previously treated metastatic CRC.[30]

In the base case analyses the health state utility values were the average of utilities obtained in the CORRECT study (not from the abstract by Siena et al.[29], but as published in Grothey et al.[30]) and the cetuximab NICE CS for the first-line treatment of mCRC, TA 176[32] (see Table 5.10). The justification for using the CORRECT study as a source of utilities was that this study was conducted at the same disease stage. The justification for using an average of the above-mentioned two sources in the base case is that these are the *"two most appropriate sources"*.

Table 5.10: Summary of utility values for cost effectiveness analysis
(Based on Table 60 of the CS[1])

State	Base case Utility value mean (SE)*	Regorafinib CORRECT study Utility value mean (SE)	Cetuximab NICE CS Utility value (TA176) mean (SE)
Pre-progression – on treatment	0.73 (0.01)	0.73 (NR)	0.73 (NR)§
Pre-progression – BSC	0.74 (0.02)	0.74 (NR)	0.73 (NR)§
Post-progression – T/T	0.64 (0.01)	0.59 (NR)	0.68 (NR)#
Post-progression – BSC	0.64 (0.02)	0.59 (NR)	0.68 (NR)#
Dead	0ß	0ß	0ß

* Average of CORRECT study and the cetuximab NICE company submission for the first-line treatment of mCRC, TA176; § Second line; # Third line; ß NICE reference case.
BSC = best supportive care; CS = company submission; NICE, National Institute for Health and Care Excellence; CS, company submission; SE = standard error; TA = technology appraisal; T/T = trifluridine/tipiracil

In sensitivity analyses the utilities from the CORRECT study and the TA176 were used as health state utility values.

Disutilities for AEs were not incorporated in the model. This is justified in the CS by stating a lack of evidence to estimate disutilities, and by the argument that small changes in health-related quality of life attributable to AEs are already incorporated in the chosen estimates for the health state utilities.

ERG comment: The ERG comments regarding health-related quality of life focus on: the estimation of health state utilities, and not incorporating the impact of adverse events on health-related quality of life in the analysis.

Health state utilities
The ERG has doubts whether the CS for TA176[32] is an appropriate source for health state utilities. The health state utility used for pre-progression (0.73) taken from the TA176 CS report was derived with the HUI3 instrument from the study of Mittmann et al.[38], as became apparent in the Merck Serono response on the ERG's clarification questions[39]. This study by Mittmann et al. was excluded by the company from their systematic review because the method is not in line with the NICE reference case. Moreover, the 0.73 value was mentioned in the TA176 CS report, but as described in the ID794 assessment report[39], another value (0.77 from Bennett et al.[40]) was used in the model. The 0.68 value for post-progression was determined in a population of patients with chemo refractory wild type KRAS metastatic colorectal cancer using EQ-5D and a Quality-Adjusted Time Without Symptoms or Toxicity (Q-TWIST) approach and taken from a poster by Wang et al.[41], The ERG was unable to access the poster but the online abstract does not mention any utility values. Another publication by the same authors (and the same year) does not mention a utility value of 0.68; instead values of 0.63 (panitumumab) and 0.64 (best supportive care) are mentioned for patients with relapse.[42]

The ERG asked the company to clarify why the base case model inputs for health state utilities are based on an average of utilities from the CORRECT study[30] and the TA176 CS report[32]. The company answered that TA176 was selected as *"an appropriate source for an upper bound of health state utilities, given that the utility used for patients in pre-progression was taken from patients on second-*

line treatment".[9] The lower bound estimate was taken from the CORRECT study, because the toxicity profile of regorafenib *"may be deemed worse than the 'acceptable toxicity profile' of trifluridine/tipiracil given the increased incidence of Grade ≥3 hypertension and hand-foot syndrome associated with regorafenib treatment"*.[9]The ERG thinks this latter argument is incorrect because the health state utilities in the BSC group were very similar to the utilities in the regorafenib group (0.74 and 0.73 pre-progression and 0.59 and 0.59 post-progression, respectively). Moreover, the quoted pre-progression utilities were determined at baseline.[30]

The ERG also asked the company to justify why other NICE appraisals that may contain relevant information (e.g. TA118[43], TA212[44], TA307[45] and ID794[39, 46, 47]) were not used. The company responded that utility values in TA307 were commercial in confidence, and that in TA212 the same values as in TA176 were used. The company considered the utility values from TA118 and ID794 for pre-progression inappropriate, as these values are higher than the values in TA176.[9] The ERG agree that the utilities used in TA118 are less relevant, but for other reasons than stated by the company: non NICE reference methods were used (direct time trade-off [48] and Q-TWIST[49]), and utilities were obtained in an adjuvant population. The ERG thinks that in TA176 and ID794 potentially relevant information can be found.

In summary, according to the ERG, the arguments to estimate the health state utilities based on an average of the utilities mentioned in the CS report of TA176 and the CORRECT trial are incorrect. Therefore, the ERG prepared an overview of health state utilities used or presented in the above-mentioned appraisals, as well as more recent or other publications from the authors or studies included in these appraisals (CS or ERG report), see Table 5.11. According to the ERG there is paucity of robust evidence on health related quality of life in metastatic colorectal cancer, especially beyond first line. In this light the omission to collect health related quality of life information in the phase II trial and the RECOURSE study is particularly problematic. When disregarding the studies not using the NICE reference case methodology[38, 41, 42], the utilities for pre progression range from 0.68[50] for chemotherapy refractory patients to 0.77[40] for second line. The post-progression health state utilities range from 0.59[30] from the CORRECT study to 0.66[51] or 0.64[52] for a Finnish end stage or palliative population, respectively. According to the ERG, the baseline utilities from the CORRECT study are the most plausible estimates for pre-progression and the post-progression utilities because it is the only study identified by the ERG in which utilities were measured using the EQ-5D in a population that resembles the population in this appraisal (second to fourth line population with 74% ≥ third line). Therefore the ERG included utility values from the CORRECT study in its base case.

Table 5.11: Overview of utility values from the literature

Source	Population of metastatic colorectal cancer	UK	Instrument	Pre progression			Post progression		
				Mean	Mean	SD (N)			SD (N)
Grothey 2013[30] (CORRECT) this submission	26% 1st / 2nd line 26% 3rd line 48% 4th line	Worldwide including UK	EQ-5D UK tariff	Regorafenib* Placebo*	0.73 0.74	0.25 (500) 0.27 (253)	Regorafinib Placebo	0.59 0.59	0.31 (500) 0.34 (253)
Bennett 2011[40] (NCT0339183) TA176 model; ID794	2nd line	Worldwide including UK	EQ-5D UK tariff	PAN* FOLFIRI*	0.77 0.76	0.23 (263) 0.25 (267)			
Wang 2011[42] (NCT00113763) TA176, ID794	Chemo refractory wild-type KRAS	Worldwide including UK	EQ-5D UK tariff & Q-TWIST	No toxicity PAN No toxicity BSC Toxicity PAN Toxicity BSC	0.77 0.66 0.60 0.44	NR (104) NR (103) NR (37) NR (13)	PAN BSC	0.63 0.64	NR (68) NR (63)
Farkkila 2013[52]	All lines	Finland	EQ-5D UK tariff	Non palliative	0.82	0.20 (108)	Palliative	0.64	0.31 (41)
Farkkila 2014[51]	End stage§	Finland	EQ-5D UK tariff	Mean 0.66, SD 0.30, N 57					
Stein 2014[53]	All lines, no brain metastasis	UK, Netherlands	EQ-5D UK tariff		0.74	0.23 (42)		0.73	0.29 (32)
Odom 2011[50] (NCT0339183)	Chemo refractory	Worldwide including UK	EQ-5D UK tariff	PAN* BSC*	0.72 0.68	0.24 (188) 0.25 (175)			
Koukakis 2016[54] (NCT00113763)	3rd / 4th line RAS wild type			PAN# BSC#	0.78 0.73	NR (62) NR (60)			

* Baseline values; § no chemo- or radiotherapy or within 6 months before death; [3]Median values instead of mean

BSC = best supportive care; EQ-5D = European Quality of Life-5 Dimensions; FOLFIRI = chemotherapy combining folinic acid, fluorouracil and irinotecan; KRAS = Kirsten rat sarcoma viral oncogene homolog; NR = not reported; PAN = panitumumab; SD = standard deviation; TA = technology appraisal; UK = United Kingdom

Impact of adverse events on health related quality of life

The ERG noted that the impact of AEs on health-related quality of life was not incorporated in the analyses, apart from the difference between the pre-progression health state utility values in the base case. Patients receiving T/T had more grade >2 adverse events in general, and for instance more neutropenia, leukopenia, anaemia, and gastro intestinal events than placebo in the RECOURSE trial, see Tables 44 and 45 of the CS.[1] Therefore, the ERG questions the justification that the 0.01 utility difference between the utility scores 0.73 (pre-progression on treatment) and 0.74 (pre-progression BSC) captures the difference in AEs impact on quality of life. Therefore, the ERG asked the company to incorporate the impact of adverse events on health-related quality of life in the economic analysis.[20] The company responded that it was not feasible to explicitly model the impact of adverse events on health-related quality of life because they did not have a detailed insight into the two sources they used to estimate utilities (CORRECT study[30] and TA176[32]). Moreover, the company argued that the utilities they used already incorporated the impact of adverse events.[9]According to the ERG, these arguments are incorrect, for the following reasons:

1. Regarding the first argument (not feasible to explicitly model the impact of adverse events), the incidence of adverse events is known from the phase II study and RECOURSE, and for instance from the recent NICE diagnostic assessment report by Freeman et al.[55], a review on the impact of common adverse events on health-related quality of life in colorectal cancer is available. This information was also used in the ID794 assessment report.[39]
2. Regarding the second argument (already incorporated the impact of adverse events), as the 0.73 and 0.74 utility values used are the baseline utility values measured in the CORRECT trial, any difference between those values is probably due to randomness and cannot be due to differential impact of treatment related adverse events.

The ERG explored the estimation of a disutility for adverse events based on the RECOURSE occurrence of adverse events ≥ grade 3 as reported in Table 5.9. The ERG based the disutilities for adverse events on the ones reported in Freeman et al.[55] and the ID794 assessment report[39] and, similar to these two appraisals, assumed a disutility duration of one week. Disutilities for thrombocytopenia, nausea, decreased appetite, hand-foot syndrome and vomiting were not reported in these sources and assumed to be the same as for fatigue. For fever, febrile neutropenia and cardiac ischemia the same disutility as for neutropenia was assumed. This resulted in a disutility of 0.075 for T/T and 0.018 for BSC, calculated to one week the incremental disutility is -0.001. As these estimates do not include all AEs and heavily rely on assumptions, in the base case the ERG used a larger disutility for AEs of 0.01 per cycle for patients receiving T/T (similar to the company's base case, i.e. 0.74 (on T/T) - 0.73 (on BSC), but based on alternative justifications).

5.2.9 Resources and costs

The company based its resource use and costs on the company submission of a recent NICE technology appraisal in mCRC (ID794).[28] Additional resource use was based on published literature and expert opinion.

Drug costs

T/T is available in 15 mg or 20 mg tablets, in pack sizes of 20 and 60. Unit costs of these pack sizes were presented in at the list price (Table 5.12). Dosage was based on BSA, where pack size could cater for all doses (Table 62 of the CS).

Table 5.12: Unit costs of treatment
(Based on Table 61 of the CS[1])

Treatment	Unit dose (mg)	Pack size	Unit cost	Source
Trifluridine/tipiracil	15	20	£500	Servier
		60	£1,500	
	20	20	£667	
		60	£2,000	
CS = company submission; mg = milligram				

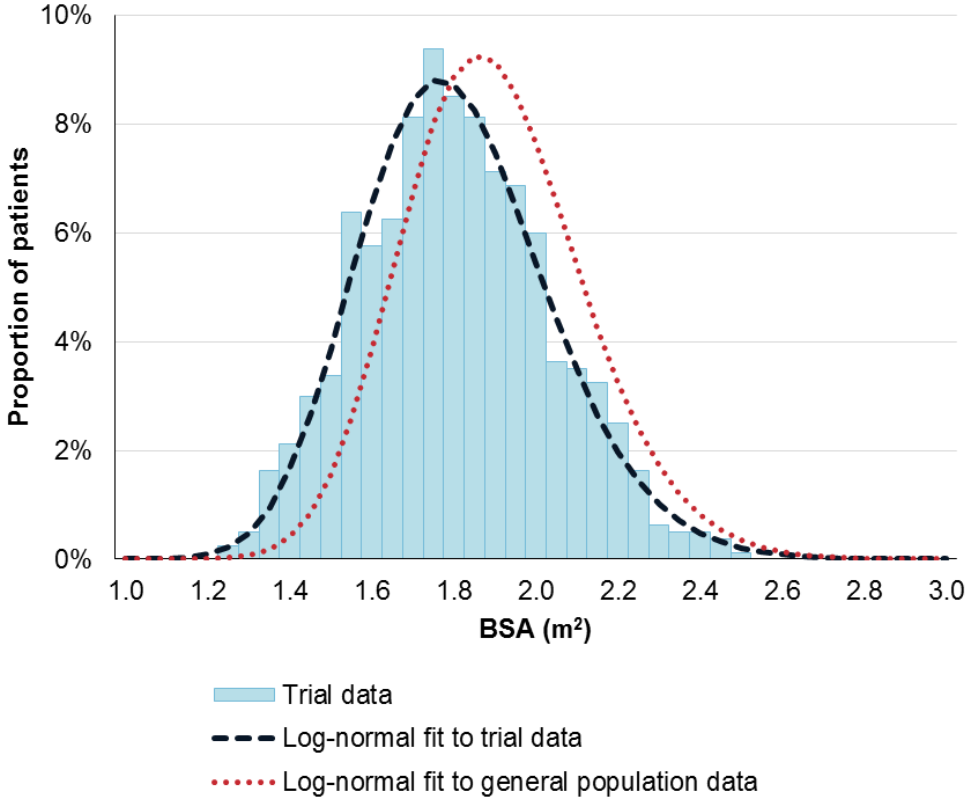
The RECOURSE trial data were used to calculate the BSA distribution in the population. In order to calculate T/T dosing, patients were categorised into 10 groups, each group having an assigned dosage. The distribution of BSA used in the model base case was derived from a log-normal fit to the distribution of BSA in the RECOURSE trial, which the company reports was done *"to produce a more realistic estimate of the distribution of patient BSA"*. The CS reports that *"clinicians at the advisory board indicated that patients with mCRC would be expected to lose weight, given their disease status, and therefore agreed with the use of a lower estimate of BSA compared with the general population particularly at the line of treatment relevant to the decision problem"*.[13] Distributions of the BSA are presented in Table 5.13 and Figure 5.6.

Table 5.13: T/T based on BSA
(Based on Tables 55 and 62 of the CS[1])

BSA (m^2)	Distribution of BSA				
	Dosage (mg; 2x daily)	Cost per cycle (list price)	RECOURSE data	Log-normal fit to RECOURSE data	Log-normal fit to general population data*
< 1.07	35	£1,167	0.00%	0.00%	0.00%
1.07 - 1.22	40	£1,333	0.13%	0.19%	0.01%
1.23 - 1.37	45	£1,500	2.38%	2.15%	0.39%
1.38 - 1.52	50	£1,667	9.25%	9.55%	3.58%
1.53 - 1.68	55	£1,833	19.88%	22.47%	14.70%
1.69 - 1.83	60	£2,000	27.00%	25.97%	25.26%
1.84 - 1.98	65	£2,167	21.38%	20.57%	26.14%
1.99 - 2.14	70	£2,333	12.63%	12.13%	18.35%
2.15 - 2.29	75	£2,500	5.75%	4.72%	7.82%
≥2.30	80	£2,667	1.63%	2.25%	3.75%
Weighted average cost per cycle (list price)			[redacted]	[redacted]	[redacted]

* General population data applies to Health Survey for England data sourced by Porter et al. 2015.[56]
BSA = body surface area; CS = company submission; mg = milligram; T/T = trifluridine/tipiracil

Figure 5.6: Distribution of body surface area
(Based on Figure 33 of the CS[1])



BSA = body surface area: CS = company submission

The distribution of patients' BSA was used to calculate the weighted average cost per patient in the first treatment cycle. From cycle 2 onwards, this price was then adjusted according to the proportion of patients who experienced a dose reduction in the RECOURSE trial.[2] To all prices, the confidential discount of [redacted] was then applied.

Dose reduction

In the RECOURSE trial, 53 (9.9%) patients receiving T/T treatment had a single dose reduction, 18 (3.4%) had two reductions, and two (0.4%) had three reductions.[2] To account for these dose reductions, the proportion of patients receiving each dose for a given treatment cycle was adjusted in the subsequent treatment cycles. In the first cycle, all patients were expected to receive the T/T dose based on BSA in the first treatment cycle. Subsequently, patients from each dosing group with a dose reduction were moved to the dosing group (see BSA categories in Table 5.13) below for the next treatment cycle. This means that 9.9%, 3.4% and 0.4% of the patients receiving T/T were moved to the dosing group below their current group in the second, third and fourth cycle respectively. After the fourth cycle, it was assumed that all patients remained on their current dose until discontinuation of treatment. The proportion of patients receiving each dose of T/T per cycle is (based on the log-normal fit to RECOURSE data) shown in Table 5.14 and presented in Figure 34 of the CS.[1]

Table 5.14: Proportion of patients receiving T/T
(Based on Table 56 of the CS[1])

BSA (m^2)	Dosage (mg; 2x daily)	Cycle 1	Cycle 2	Cycle 3	Cycle 4+
< 1.07	35	0.00%	0.02%	0.04%	0.04%
1.07 - 1.22	40	0.19%	0.38%	0.47%	0.48%
1.23 - 1.37	45	2.15%	2.88%	3.15%	3.18%
1.38 - 1.52	50	9.55%	10.83%	11.24%	11.28%
1.53 - 1.68	55	22.47%	22.82%	22.91%	22.91%
1.69 - 1.83	60	25.97%	25.44%	25.25%	25.22%
1.84 - 1.98	65	20.57%	19.73%	19.45%	19.42%
1.99 - 2.14	70	12.13%	11.40%	11.16%	11.14%
2.15 - 2.29	75	4.72%	4.47%	4.39%	4.38%
≥2.30	80	2.25%	2.03%	1.96%	1.95%
Weighted average cost per cycle (list price)		[redacted]	[redacted]	[redacted]	[redacted]

BSA = body surface area; CS = company submission; mg = milligram

Treatment delay
The incorporation of treatment delays into the model allowed additional medical resource use for patients who experience a delay in treatment. As the additional medical resource use applies to all patients, regardless of treatment received, the average delay in treatment initiation was calculated for both T/T and BSC patients (Table 5.15). This resulted in an applied cycle length of 30.72 days for T/T and 29.40 days for BSC.

Table 5.15: Average delay in treatment initiation
(Based on Table 54 of the CS[1])

	Trifluridine/tipiracil	BSC
Total number of cycles	1828	598
Total number of delayed cycles	752	228
Average delay in treatment initiation for delayed patients	6.61 days	3.67 days
Average delay in treatment initiation for all patients (A)	2.72 days	1.40 days
Protocol treatment cycle length (B)	28 days	28 days
Applied treatment cycle length in model (A+B)	30.72 days	29.40 days

BSC = best supportive care; CS = company submission

Time on treatment
Treatment with T/T is 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. Not all of these factors were included in the estimation of time on treatment due to lack of available data. The company expected their estimated time on treatment to be an overestimation of the observed time on treatment and hence used PFS as a proxy for time on treatment.

Medical resource use
The company identified medical resource use items following consultation with clinical experts, due to a lack of published literature on the medical resource use of patients in this setting. An overview of

medical resource use costs can be found in Table 5.16. Medical resource use cost per health state were £203 for T/T and £182 for BSC in pre-progression, and £193 in post-progression in both arms. All other resource costs (including social care for patients toward the end of life) were assumed to be captured in the end-of-life care cost applied for all patients upon death.

Table 5.16: Summary of medical resource use
(Based on Tables 64 and 65 of the CS[1])

MRU item	Occurrence per treatment cycle[†]			Unit cost (£)	Reference
	Pre-P		PP		
	T/T	BSC			
Oral chemotherapy day case attendance[*]	1			192.32	NHS reference costs 2014-15: Day case and Regular Day/Night; SB11Z; Deliver exclusively oral chemotherapy
Medical oncologist outpatient consultation		1		170.85	NHS reference costs 2014-15: 370; Medical Oncology - Outpatient, consultant led
GP home consultation			0.25	96.92	PSSRU 2013: GP - per out of surgery visit lasting 23 minutes (without qualifications) - inflated using PSSRU 2015 inflation indices
Community nurse specialist visit			1	44.00	PSSRU 2015: Nurse Specialist (Community) Cost per hour (without qualifications) - 10.4 (contact assumed to last 1 hour)
Health home visitor	0.25	0.25	1	44.00	PSSRU 2015: Health Visitor Cost per hour (without qualifications) - 10.3 (contact assumed to last 1 hour)
District nurse visit			1	44.00	PSSRU 2015: Health Visitor Cost per hour (without qualifications) - 10.1 (contact assumed to last 1 hour)
GP surgery visit			1	37.00	PSSRU 2015: GP consultation (Per patient contact lasting 11.7 minutes, without qualifications) - 10.2
Average MRU	£203	£182	£193		

[*] Patients who experience a delay in treatment initiation incur the cost of an additional oral chemotherapy day case attendance.
[†] MRU items are incurred according to an average unadjusted treatment cycle (i.e. 28 days). Adjustments for delays in treatment initiation are captured by the repeat chemotherapy day case attendance.
BSC = best supportive care; CS = company submission; GP = general practitioner; MRU = medical resource utilisation; NHS = National Health Service; PP = post-progression; Pre-P = pre-progression; PSSRU = Personal Social Services Research Unit; T/T = trifluridine/tipiracil

Post-progression treatment costs
Following treatment discontinuation in post-progression, 42% of the RECOURSE trial patients received non-study anti-tumour treatments.[2] The RECOURSE trial data was used to estimate the average cost of post-progression treatment per patient, which was £1,549 for T/T and £1,487 for BSC (Appendix 11 of the CS).[1] Clinical experts confirmed that prior treatment with T/T is not expected to have an effect on the choice of treatments available following progression at this line of therapy. Therefore, the average

cost per patient for all patients post-treatment was used in both arms of the model (£1,528). A sensitivity analysis was performed with different costs of post-progression treatment per patient of £1,549 for T/T and £1,487 for BSC (Table 69 of the CS).[1]

End of life

End of life care costs were taken from a modelling study by Round et al, which estimates the cost of caring for people at the end of life.[33] Costs for end of life from this source take into account health care (£4,854), social care (£1,489) and charity care (£470), and excludes the cost of informal care as per the NICE reference case.[17] The total end of life care cost of £6,910 was applied in the model as a lump sum upon death for both arms.

Adverse events

The company incorporated costs of adverse events if they were actively treated in the NHS, as verified with clinical and medical oncologists. The adverse events incorporated in the CS model are presented in Table 5.17. Incorporating these adverse events at their unit costs to the rates observed from the RECOURSE clinical trial yielded a cost of AEs of £923 for T/T and £426 for BSC (table 68 of the CS).[1] These costs are applied one time, at the start of the model.

Table 5.17: Adverse events included in the model

(Based on Table 67 of the CS[1])

Adverse event	Actively treated		Cost of treatment		Reference (see notes for sources)	
	All grades	Grade ≥3	Grade ≤2	Grade ≥3	Grade ≤2	Grade ≥3
Nausea		✓		£158.43		**a**
Vomiting	✓	✓	£158.43	£158.43	**a**	**a**
Decreased appetite		✓		£158.43		**a**
Fatigue		✓		£158.43		**a**
Diarrhoea	✓	✓	£158.43	£158.43	**a**	**a**
Abdominal pain		✓		£139.52		**b**
Fever	✓	✓	£158.43	£158.43	**a**	**a**
Asthenia		✓		£158.43		**a**
Febrile neutropenia	✓	✓	£2,583.98	£2,583.98	**c**	**c**
Stomatitis		✓		£158.43		**a**
Hand-foot syndrome		✓		£158.43		**a**
Cardiac ischaemia	✓	✓	£158.43	£158.43	**a**	**a**
Neutropenia		✓		£1,227.95		**d**
Leucopenia		✓		£158.43		**a**
Anaemia		✓		£799.00		**e**
Thrombocytopenia		✓		£643.48		**f**

Adverse event	Actively treated		Cost of treatment		Reference (see notes for sources)	
	All grades	Grade ≥3	Grade ≤2	Grade ≥3	Grade ≤2	Grade ≥3
References: a NHS Reference costs 14-15: Outpatient visit, general medicine[57]; b NHS Reference costs 14-15: Outpatient visit, pain management[57]; c NICE DSU report[58]; d NHS Reference costs 14-15: Average non-elective inpatient stay[57]; e PENTAG ERG Report for cetuximab[39]; f NHS Reference costs 14-15: Weighted cost of thrombocytopenia based on complications and comorbidities score.[57] CS = company submission; DSU = Decision Support Unit; NHS = National Health Service; NICE = National Institute for Health and Care Excellence; PenTAG = Peninsula Technology Assessment Group						

Table 5.18: Health states and associated costs per treatment cycle
(Based on Table 66 of the CS[1])

Health state	Items	Value		CS Reference
		T/T	BSC	
Pre-progression	Technology¥	[redacted]	£0	Table 63
	MRU*	£203	£182	Table 65
Progressed	Technology	£0		Table 63
	MRU	£193		Table 65
Non-health state costs applied as a lump sum	Adverse events†	£923	£426	Table 68
	End of life‡	£6,910		Section 5.5
	Post-progression treatment$^{\Delta}$	£1,528		Table 69

* additional chemotherapy day case attendance applies for patients experiencing delays.
† applied for all patients in the first model cycle.
‡ applied upon death.
$^{\Delta}$ applied upon progression.
¥ based on average BSA in RECOURSE of 1.78 m^2.
BSC = best supportive care; CS = company submission; MRU = medical resource utilisation; T/T = trifluridine/tipiracil

ERG comment: Following the NICE reference case[17], *"evidence should be presented to demonstrate that resource use and cost data have been identified systematically"*. Hence, a more systematic approach, including a review, would have been desirable to inform model parameters on resources use and costs. After a request in the clarification letter, the company explained that a review of NICE technology appraisals and the associated assessment reports in mCRC was undertaken and these data were presented at advisory boards and face to face meetings. However, a review with broader search objectives and strategy (e.g. including other interventions than T/T only) would potentially identify cost effectiveness studies relevant for informing the model produced by the company (e.g. model structure, health state utility, resource use and BSC parameters). For instance, the studies by Goldstein et al.,[59] Starling et al.,[60] Shiroiwa et al.[61] and Hoyle et al.[62] which were identified by the company but eventually excluded (see Table 2 in Appendix 6 of the CS[1]), might have been relevant for informing the model. In particular regarding resource use and costs.

The ERG has the following specific issues with the modelling of resources use and costs:

- estimation of BSA to calculate drug costs,
- estimation of dose reductions,
- estimation of treatment delay,
- estimation of time on treatment,
- assuming equal post-progression costs for T/T and BSC,
- estimation of medical resource use,
- calculation of end-of-life costs,
- calculation of adverse event costs.

These issues are discussed below and addressed in the ERG's additional analyses.

Estimation of BSA to calculate drug costs
The CS reported that advisory board clinicians agreed with the use of a lower estimate of BSA (following from the log-normal distribution fitted to the RECOURSE data) as compared with the general UK population since mCRC patients would be expected to lose weight. The ERG, however, notes that the population of the RECOURSE trial includes 33% of patients from Japan, which may be expected to have a lower BSA than the UK population.[13]

The company reports that the non-parameterised distribution of BSA from RECOURSE was also explored, as well as the application of a log-normal fit of BSA from general population data, which were explored as scenario analyses. The results of these scenario analyses were initially not reported, but were provided after requesting this in the clarification letter.[9] According to the ERG, the non-parametrised distribution of BSA from RECOURSE is a reasonable estimate of BSA to calculate drug costs. As this most likely results in an underestimation of T/T costs, the BSA based on the UK population (which most likely results in an overestimation of T/T costs) is considered in an exploratory sensitivity analysis.

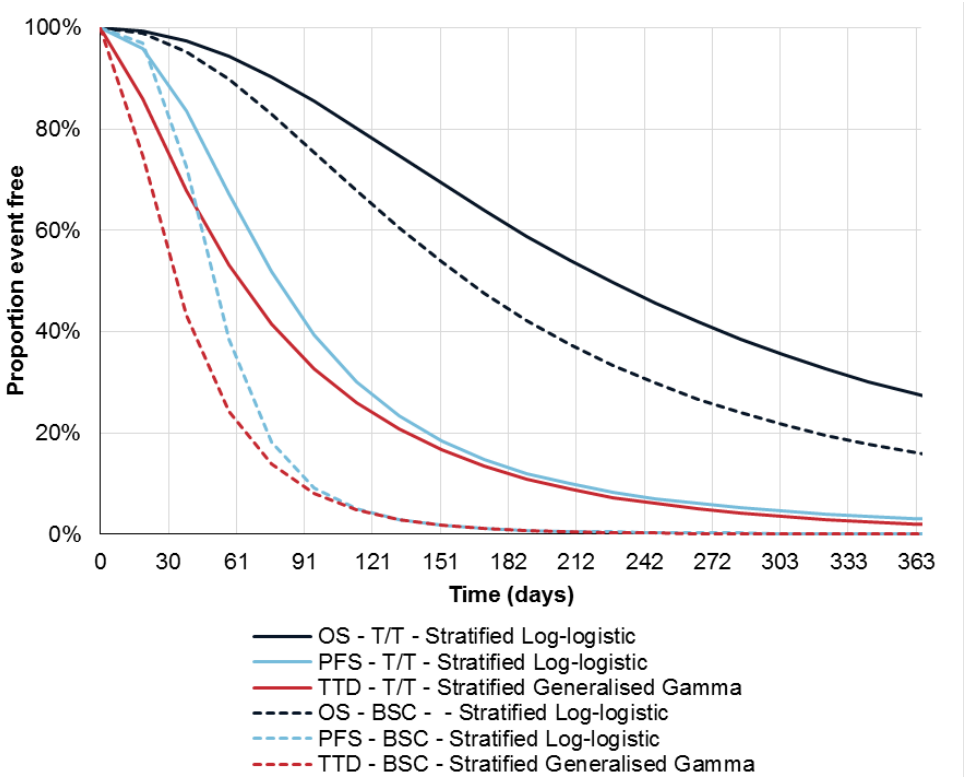
Estimation of dose reductions
Dose reductions for T/T were estimated based on the RECOURSE trial. Although the assumption that in case of a dose reduction patients were moved to the dosing group below their current group can be questioned, the impact of the assumption is probably small (informally explored by the ERG).

Estimation of treatment delay
The company applied a cycle length of 30.72 days for T/T and 29.40 days for BSC in the model to account for treatment delay, as observed in RECOURSE. This leads to slightly more medical resource use in BSC over the time horizon of the model. The estimate of 29.40 days was calculated based on BSC treatment (see company's response on clarification question B8[9]), and is thus not representative for clinical practice. In its base case the ERG applied the same cycle length for T/T and BSC.

Estimation of time on treatment
The ERG asked the company to clarify why PFS was used to approximate time on treatment, while it seems that empirical data was available to estimate this. The company responded: "*...time on treatment was not explicitly reported in either of the clinical trials from which efficacy data were derived, (...) but data are available regarding the start and end time of treatment for patients within both studies, from which an estimate of TTD (time to treatment discontinuation) may be derived.*"[9]

The provided additional analyses based on the assumption that all remaining patients experience the event of treatment discontinuation at the end time of treatment (i.e. no patients have been censored at this time, due to available data). The company tested different survival curves to represent time to treatment discontinuation (TTD). Since the stratified generalised gamma provided the best AIC estimate, it was chosen to represent TTD in the cost-effectiveness model provided in the response to the ERG clarification letter (Figure 5.7).[9]

Figure 5.7: Estimation of OS, PFS and TTD used in the economic model
(Based on Figure 9 of the response to the request for clarification[9])



BSC = best supportive care; OS = overall survival; PFS = progression free survival; T/T = trifluridine/tipiracil; TTD = time to treatment discontinuation

Given that not all relevant factors were included in the estimation of time on treatment (as stated by the company, see above) and the assumption that all patients experience the event of treatment discontinuation at the end time of treatment (i.e. no patients have been censored at this time, due to available data), the ERG regards the company's approach to use PFS as proxy as reasonable. Hence, this was used in the ERG base case. The ERG used time on treatment in an explorative sensitivity analysis.

Assuming equal post-progression costs for T/T and BSC

The ERG asked the company to clarify why the cost of post-progression treatment was assumed to be the same for both groups of patients. The company stated that *"clinical expert opinion at the advisory board held in January 2016 suggested that the costs would be approximately equal following progression given that patients would be expected to be eligible for the same treatment following progression and that patient prognosis following progression at this late stage of disease is similarly poor across treatment groups. Analysis of the data demonstrated that costs between trifluridine/tipiracil versus BSC patients were approximately equal (£1,549 versus £1,487)".*[9] As empirical data are

available for both treatments, the ERG would prefer to use the empirical estimates instead of assuming equal costs for both treatments. Hence, treatment specific post-progression costs were incorporated in the ERG base case.

Estimation of medical resource use

The estimation of medical resource use was based on expert opinion, while empirical evidence could have been collected in the phase II trial and RECOURSE. Given the complete reliance on expert opinion for resource use, the ERG used an alternative source in an explorative sensitivity analysis. Accordingly, it was assumed that there were no medical oncologist outpatient consults for BSC and costs of computed tomography (CT) scans were included for T/T (assuming one scan per three cycles costing £112 each).[62]

The ERG noted a small error in the costs of a medical oncologist outpatient consultation (the ERG could not replicate the cost estimate reported in the CS). This was recalculated by the ERG using the weighted average of WF01A, WF01B, WF01C and WF01D from NHS reference costs 2014-15: £168.40, instead of £170.85.[57] This was corrected in the ERG's analyses.

Calculation of end-of life costs

Considering the end-of-life costs calculated based on Round et al.[33], the ERG notes that charity care costs (£470), consisting of hospice inpatient days and hospice outpatient visits, neither falls within NHS nor PSS cost. The paper by Round et al. reports that *"charities also provide care through other means, often paid for in part by local authorities and the health service – these costs will have been captured where possible in the social care element of spending"* (p.902). Hence, only the reported health care (NHS, £4,854) and social care (PSS, £1,489) costs in this study are relevant. These costs are included as end-of-life costs in the ERG base case.

Calculation of adverse event costs

The ERG noted that several adverse events in Table 57 in the CS (an overview of adverse events observed in the RECOURSE trial) are missing from Table 67 (an overview of adverse events for which costs are incorporated in the model). The ERG asked the company to include all adverse events reported in Table 57 in an updated version of Table 67 and to include these adverse events in the model analyses, which the company did in a sensitivity analysis.

Table 5.19: Adverse events included in the model

(Based on Table 67 of the CS[1] and Table 5 of the response to request for clarification[9])

Adverse event	Actively treated		Cost of treatment		Reference (see notes for sources)	
	All grades	Grade ≥3	Grade ≤2	Grade ≥3	Grade ≤2	Grade ≥3
Nausea		✓		£158.43		**a**
Vomiting	✓	✓	£158.43	£158.43	**a**	**a**
Decreased appetite		✓		£158.43		**a**
Fatigue		✓		£158.43		**a**
Diarrhoea	✓	✓	£158.43	£158.43	**a**	**a**
Abdominal pain		✓		£139.52		**b**
Fever	✓	✓	£158.43	£158.43	**a**	**a**
Asthenia		✓		£158.43		**a**

Adverse event	Actively treated		Cost of treatment		Reference (see notes for sources)	
	All grades	Grade ≥3	Grade ≤2	Grade ≥3	Grade ≤2	Grade ≥3
Febrile neutropenia	✓	✓	£2,583.98	£2,583.98	**c**	**c**
Stomatitis		✓		£158.43		**a**
Hand-foot syndrome		✓		£158.43		**a**
Cardiac ischaemia	✓	✓	£158.43	£158.43	**a**	**a**
Neutropenia		✓		£1,227.95		**d**
Leucopenia		✓		£158.43		**a**
Anaemia		✓		£799.00		**e**
Thrombocytopenia		✓		£643.48		**f**
Increase in alanine aminotransferase level		✓		£158.43		**a**
Increase in aspartate aminotransferase level		✓		£158.43		**a**
Increase in total bilirubin		✓		£158.43		**a**
Increase alkaline phosphatase level		✓		£158.43		**a**
Increase in creatine level		✓		£158.43		**a**

References: a NHS Reference costs 14-15: Outpatient visit, general medicine[57]; b NHS Reference costs 14-15: Outpatient visit, pain management[57]; c NICE DSU report[58]; d NHS Reference costs 14-15: Average non-elective inpatient stay[57]; e PENTAG ERG Report for cetuximab[39]; f NHS Reference costs 14-15: Weighted cost of thrombocytopenia based on complications and comorbidities score.[57]

CS = company submission; DSU = Decision Support Unit; NHS = National Health Service; NICE = National Institute for Health and Care Excellence; PenTAG = Peninsula Technology Assessment Group

The ERG also noted that costs for adverse events were almost all estimated to equal a general medicine outpatient visit. The ERG thinks that this assumption is unrealistic and used alternative inputs (see Table 5.20), retrieved from the NICE appraisal of bortezomib TA370.[63, 64]

Table 5.20: Alternative inputs for the costs of adverse events

Adverse event	ERG estimate	Source
Neutropenia Grade 3-5*	£167.28	NHS reference costs 2013-2014; HRG code: XD25Z
Thrombocytopenia Grade 3*	£570.97	NHS reference costs 2013-2014; NEI_S; weighted average of HRG codes: SA12G, H, J, and K
Thrombocytopenia Grade 4-5*	£2,191.65	NHS reference costs 2013-2014; NEI_L; weighted average of HRG codes: SA12G, H, J, and K
Anaemia Grade 3*	£516.66	NHS reference costs 2013-2014; NEI_S; weighted average of HRG codes: SA04G, H, J, K and L
Anaemia Grade 4-5*	£1,853.10	NHS reference costs 2013-2014; NEI_L; weighted average of HRG codes: SA04G, H, J, K and L
Leukopenia Grade 3-5*	£167.28	Costs assumed to be equal to neutropenia

Adverse event	ERG estimate	Source
Fatigue Grade 3-5*	£12.00	NICE ERG report abiraterone (TA259), table 24, p. 64.
Diarrhoea Grade 3*	£572.80	NHS reference costs 2013-2014; NEI_S; HRG code: PF26B
Febrile neutropenia Grade 3#	£999.20	NHS reference costs 2013-2014; NEI-S; weighted average of PM45A, B, C and D; Febrile Neutropenia with Malignancy; Short Stay
Febrile neutropenia Grade 4/5#	£5,379.59	NHS reference costs 2013-2014; NEI_L; Weighted average of PM45A, B, C and D; Febrile Neutropenia with Malignancy; Long stay
Diarrhoea Grade 4/5#	£579.21	NHS reference costs 2013-2014; NEI_S; Weighted average of PF26A&B; Other Gastrointestinal Disorders with CC Score 1+; Short Stay

* Retrieved from table 6.21 of assessment report TA370[63]; # Retrieved from table 61 CS TA370 [64]; TA259 [65]
CS = company submission; ERG = Evidence Review Group; NHS = National Health Service; TA = technology appraisal

5.2.10 Cost effectiveness results

At the list price, T/T is associated with an incremental cost effectiveness ratio (ICER) of [redacted] per additional QALY gained (see Table 5.21). At the commercial in confidence patient access scheme (PAS) price, T/T is associated with an ICER of £44,032 per additional QALY gained.

Table 5.21: Base-case results without and with patient access scheme
(Based on Tables 72 and 73 of the CS[1])

	Total			Incremental			
Technologies	**costs (£)**	**QALYs**	**LYG**	**costs (£)**	**QALYs**	**LYG**	**CER (£) (QALYs)**
BSC	10,286	0.42	0.66				
T/T without PAS	[redacted]	0.59	0.92	[redacted]	0.17	0.27	[redacted]
T/T with PAS	16,386	0.59	0.92	7,574	0.17	0.27	44,032

BSC = best supportive care; CER = cost-effectiveness ratio; LYG = life years gained; PAS = patient access scheme; QALYs = quality-adjusted life years; T/T = trifluridine/tipiracil

The company also provided disaggregated model results: QALYs, life years (LYs) and costs per health state (Tables 5.22 and 5.23). The cost difference of £7,574 is predominantly accrued in the pre-progression state.

Table 5.22: Summary of QALY and life year gain by health state
(Based on Tables 75 and 76 of the CS[1])

Health state	QALY T/T	QALY BSC	Increment	Absolute increment	% absolute increment
Pre-progression	0.22	0.12	0.10	0.10	61%
Post-progression	0.37	0.30	0.07	0.07	39%
Total	0.59	0.42	0.17	0.17	100%
Health state	**LY T/T**	**LY BSC**	**Increment**	**Absolute increment**	**% absolute increment**
Pre-progression	0.30	0.16	0.15	0.15	55%
Post-progression	0.62	0.50	0.12	0.12	45%
Total	0.92	0.66	0.27	0.27	100%

BSC = best supportive care; LY = life year; QALY = quality-adjusted life year; T/T = trifluridine/tipiracil

Table 5.23: Summary of costs by health state and category – PAS price
(Based on Table 78 of the CS[1])

Health state	Costs T/T (£)	Costs BSC (£)	Increment (£)	Absolute increment (£)	% absolute increment
Pre-progression	**8,325**	**869**	**7,456**	**7,456**	**100%**
Drug costs	6,550	0	6,550	6,550	88%
Monitoring	852	443	409	409	5%
Adverse events	923	426	497	497	7%
Post-progression	**2,860**	**2,672**	**188**	**188**	**100%**
Drug costs	1,511	1,519	-8	8	4%
Monitoring	1,348	1,152	196	196	96%
Total	**17,859**	**10,286**	**7,574**	**7,574**	**100%**
Drug costs	8,062	1,519	6,542	6,542	85%
Monitoring	2,200	1,595	605	605	8%
Adverse events	923	426	497	497	6%
End of life*	6,675	6,745	-71	71	1%

* End-of-life care costs apply for all patients irrespective of progression status.
BSC = best supportive care; CS = company submission; PAS = Patients Access Scheme; T/T = trifluridine/tipiracil

ERG comment: In response to questions posed by the ERG, the company carried out updated analyses. These analyses differ from the original base case with respect to the use of pooled estimates for adverse events rates, time on treatment and dose reductions instead of RECOURSE data only, and the incorporation of costs for adverse events that were previously missing. Moreover, the company corrected an error in the number of AE for BSC. However an error in AE for T/T was induced (both errors were corrected in the ERG base case). In the updated analysis T/T is associated with an ICER of [redacted] per additional QALY gained. At the commercial in confidence PAS price, T/T is associated with an ICER of £42,674 per additional QALY gained (deterministic results, Table 5.24).

Table 5.24: Updated results with and without patient access scheme ([redacted]) – deterministic

Technologies	Total			Incremental			
	Costs (£)	QALYs	LYG	Costs (£)	QALYs	LYG	ICER (£)
BSC	10,116	0.42	0.66				
T/T without PAS	[redacted]	0.59	0.92	[redacted]	0.17	0.27	[redacted]
T/T with PAS	17,456	0.59	0.92	7,340	0.17	0.27	42,674

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

The ERG noted that only the deterministic results were provided, while according to the NICE Methods Guide[17] probabilistic methods provide the best estimates of mean costs and outcomes in non-linear decision models. In response to the ERG's clarification question the company provided the probabilistic results for all analyses (base case outcomes and sensitivity analyses). In the updated probabilistic analysis T/T is associated with an ICER of [redacted] per additional QALY gained (Table 5.25). At the commercial in confidence PAS price, T/T is associated with an ICER of £44,057 per additional QALY gained (probabilistic results).

Table 5.25: Updated results with and without patient access scheme ([redacted]) – probabilistic

Technologies	Total			Incremental			
	Costs (£)	QALYs	LYG	Costs (£)	QALYs	LYG	ICER (£)
BSC	10,205	0.42	0.66				
T/T without PAS	[redacted]	0.59	0.92	[redacted]	0.17	0.26	[redacted]
T/T with PAS	17,424	0.59	0.92	7,219	0.17	0.26	44,057

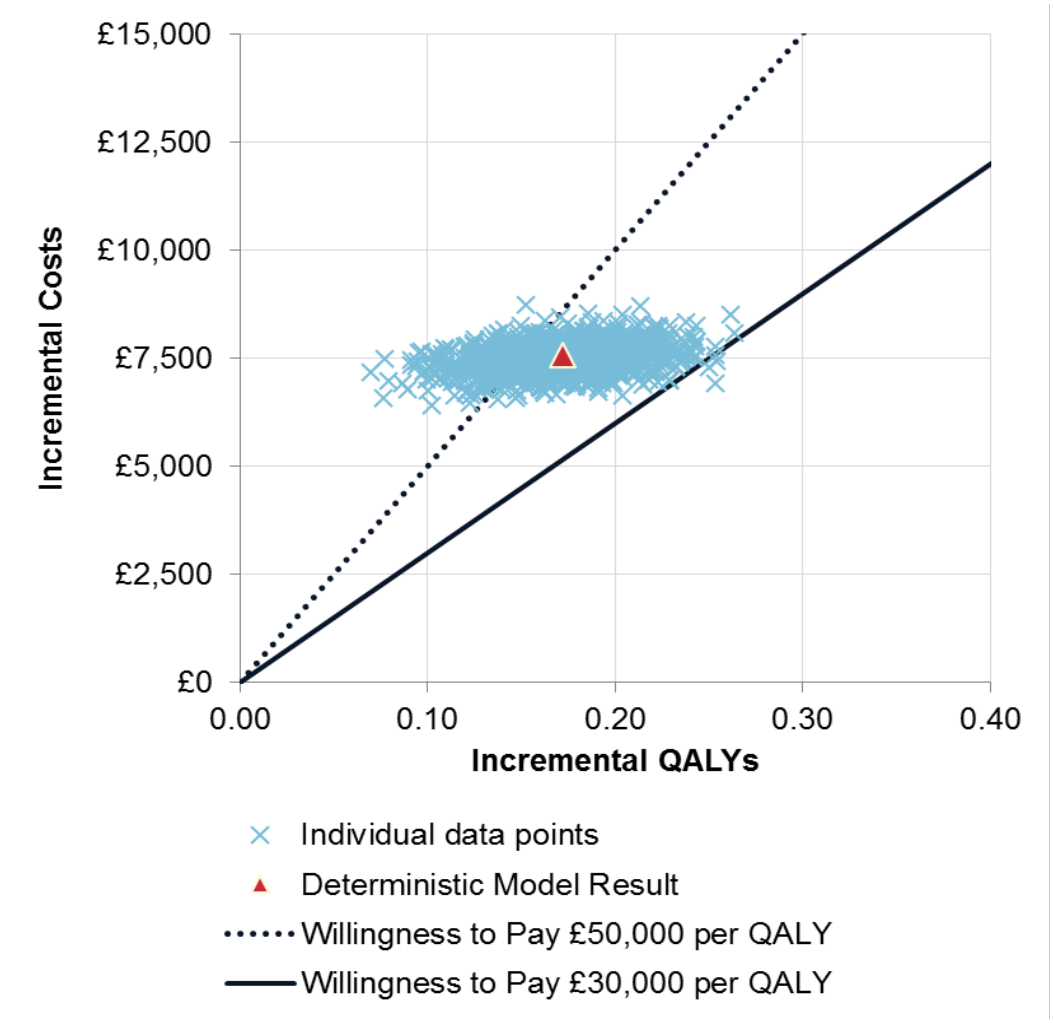
BSC = best supportive care; ICER = incremental cost-effectiveness ratio; LYG = life years gained; QALYs = quality-adjusted life years; T/T = trifluridine/tipiracil

5.2.11 Sensitivity analyses

Probabilistic sensitivity analysis

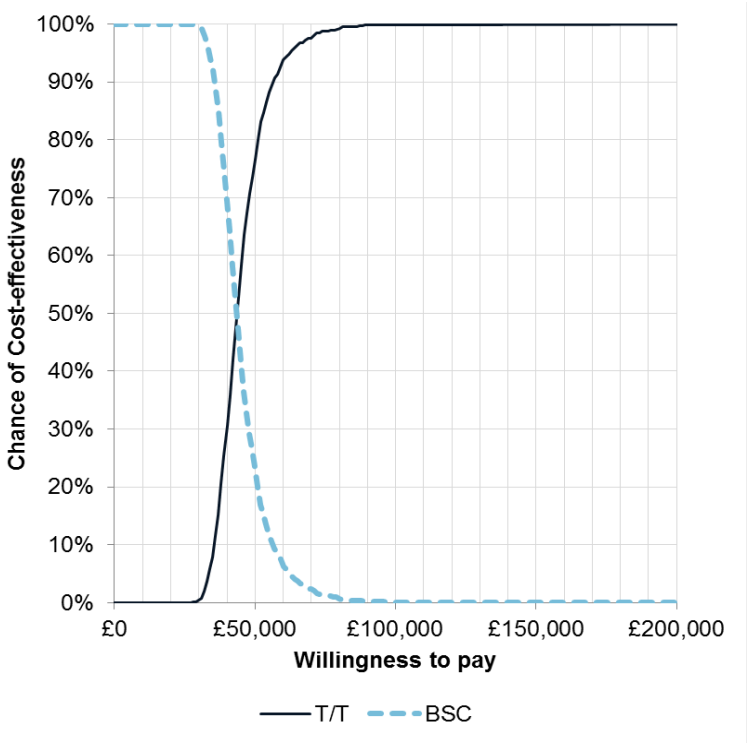
The company carried out a probabilistic sensitivity analysis (PSA) with 1,000 draws and used these simulation results to inform PSA scatterplots and cost effectiveness acceptability curves (CEAC). It is stated that *"the PSA scatterplots demonstrate an even spread of points in regards to the deterministic model result, with the majority of uncertainty shown in the estimation of the QALY gain as expected. This is likely driven by the variability in the utility values chosen, due to the lack of information regarding the uncertainty in these estimates"*.[1] The CEACs show that at the list price, the probabilities of T/T being the most cost effective treatment are 0% and 36% for willingness-to-pay (WTP) thresholds of £30,000 and £50,000, respectively (Figures 5.8 and 5.9). At the PAS price, the probabilities of T/T being the most cost effective treatment are 0% and 77% for WTP thresholds of £30,000 and £50,000, respectively.

Figure 5.8: Probabilistic sensitivity analysis scatter plot – PAS price
(Based on Figure 38 of the CS[1])



CS = company submission; PAS = patient access scheme; QALY = quality-adjusted life year

Figure 5.9: Cost effectiveness acceptability curve – PAS price
(Based on Figure 40 of the CS[1])



BSC = best supportive care; CS = company submission; PAS = patient access scheme; T/T = trifluridine/tipiracil

ERG comment: In the PSA, the minimum and maximum of multiple parameters was assumed to be +/- 20% of the mean, and a triangular distribution was used, also when information seemed to be available to estimate variance (see Table 70 of the CS[1]). This was the case for parameters estimated based on RECOURSE data (treatment delay, dosing, resource use), or expert opinion (resource use). The ERG asked the company to use the empirical data (either from RECOURSE or expert opinions) if possible to estimate the variance for input parameters (e.g. for treatment delay per patient per cycle and post-progression costs) and provide the estimated distributions.[20] In response, the company provided standard errors to estimate a distribution for post-progression costs in the PSA based on empirical data, but not for treatment delay (or other model inputs) due to a time constraint.[9] It turned out that the bounds for post-progression costs produced by the empirical data were smaller than the bounds produced using +/- 20% of the mean. The company provided an adjusted model with a setting to use the empirically derived distribution, but did not use this setting in the updated results.

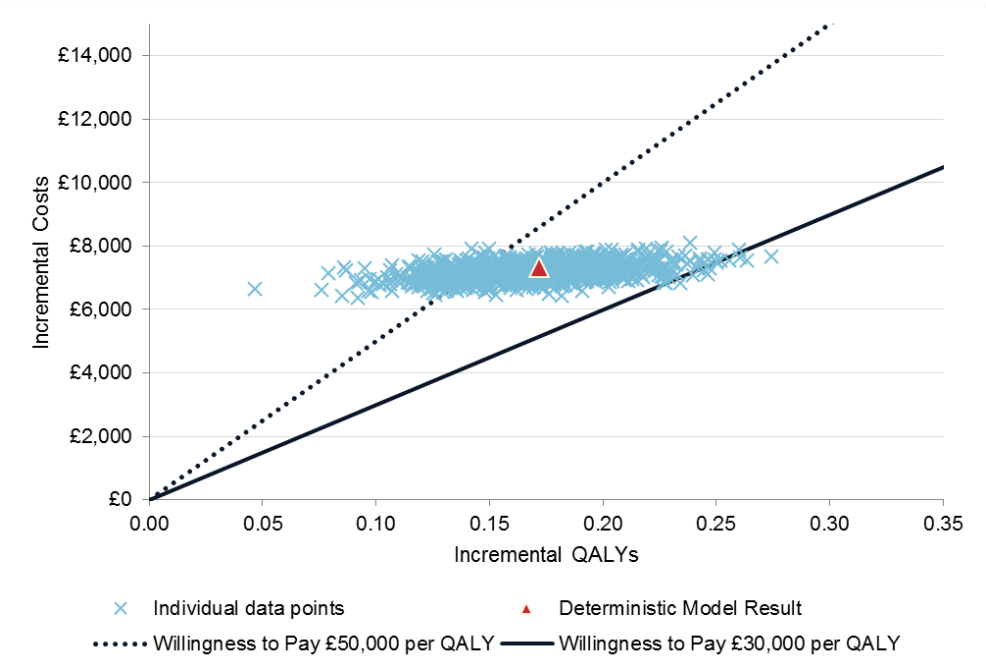
BSA (to calculate treatment dosage and hence costs) was included in the PSA, which is incorrect as variance in BSA is an indication of patient variability and not of parameter uncertainty. In its additional analysis the ERG set BSA as fixed in the PSA.

The PSA was presented correctly. However, the ERG thinks the argument that the PSA scatterplots *"demonstrate the majority of uncertainty shown in the estimation of the QALY gain as driven by the variability in the utility values chosen, due to the lack of information regarding the uncertainty in these estimates"* is somewhat flawed. The choice of scale for the axes of the scatterplot influences the visual inspection of the spread. The use of non-symmetrical scales (regarding the QALY threshold), easily biases this visual inspection. In this case, symmetrical scales based on a threshold of 30,000/QALY

would have produced a slightly more symmetrical scatter, hence suggesting that uncertainty in costs and QALYs is less different.

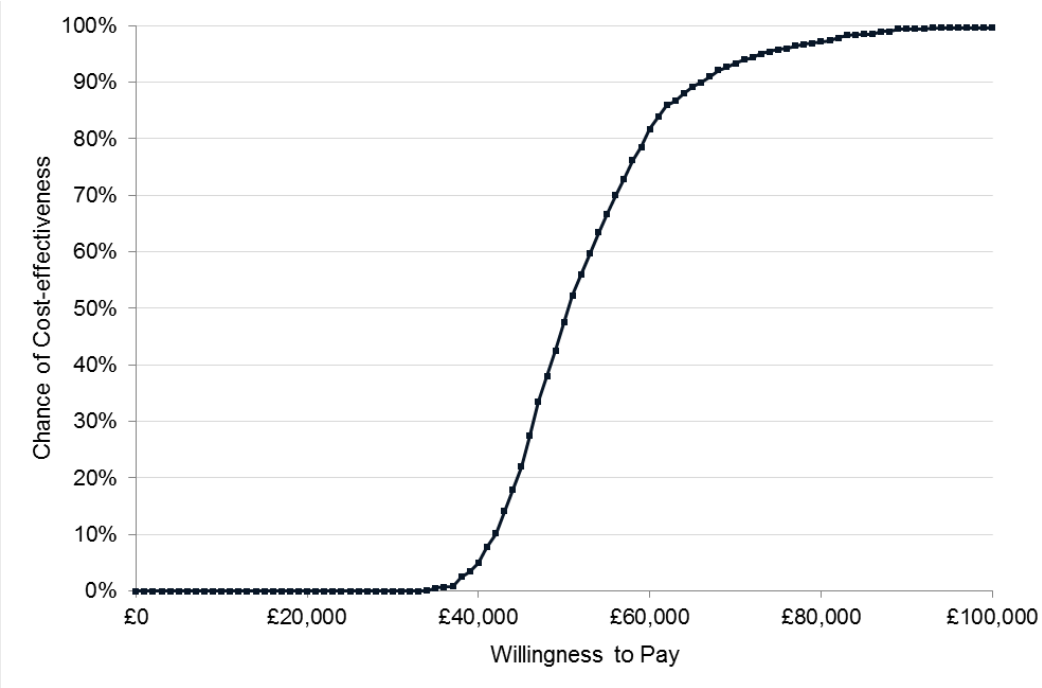
In response to clarification questions the company provided a PSA scatterplot and CEAC of the updated analysis (Figures 5.10 and 5.11).

Figure 5.10: Updated probabilistic sensitivity analysis scatter plot – PAS price
(Based on Figure 13 of the response to the request for clarification[9])



PAS = patient access scheme; QALY = quality-adjusted life year

Figure 5.11: Updated cost effectiveness acceptability curve – PAS price
(Based on Figure 14 of the response to the request for clarification[9])

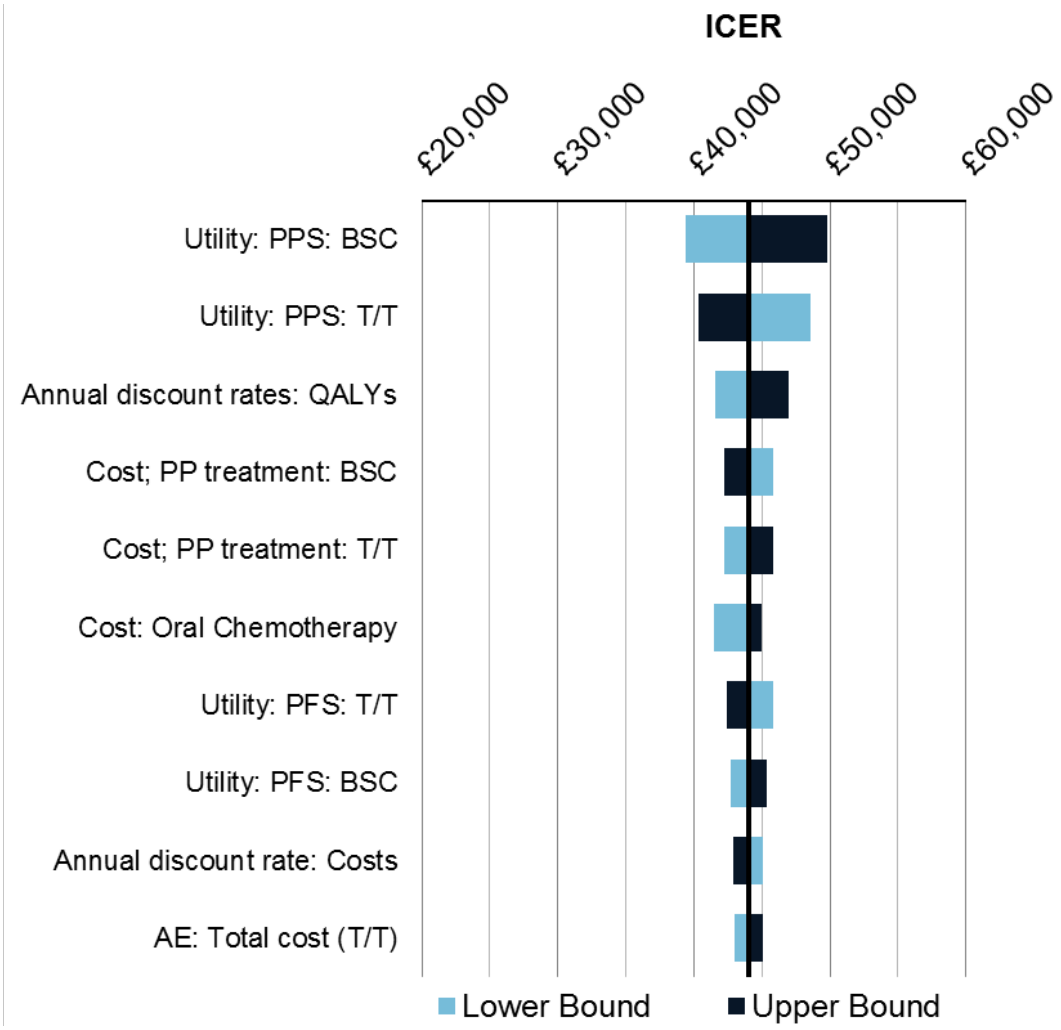


PAS = patient access scheme

Deterministic sensitivity analyses

The company performed deterministic sensitivity analyses (Figure 5.12) and presented the 10 most influential ones in tornado diagrams (with list price and with PAS).

Figure 5.12: One-way sensitivity analysis: Tornado diagram – PAS price
(Based on Figure 42 of the CS[1])



AE = adverse event; BSC = best supportive care; ICER = incremental cost-effectiveness ratio; PAS = patient access scheme; PFS = progression-free survival; PP = post-progression; PPS = post-progression survival; QALY = quality-adjusted life year; T/T = trifluridine/tipiracil

ERG comment: In response to clarification questions the company provided the probabilistic results of the updated scenario analyses.[9] The ICERs with the PAS price range from £38,128 per QALY gained for the analysis based on the phase II study population, to £57,576 per QALY gained when using a stratified log logistic model for OS and PFS (Table 5.26).

Table 5.26: Scenario analysis results for the updated analysis - probabilistic
(Based on Table 23 of the response to the request for clarification[9])

Input	Base case	Scenario	ICER (List price)	ICER (PAS price)
Updated			[redacted]	£44,057
Time horizon	10 years	2 years	[redacted]	£56,629
		4 years	[redacted]	£49,674
		6 years	[redacted]	£47,019
		8 years	[redacted]	£45,686
Patient population	Pooled	RECOURSE	[redacted]	£49,661
		Phase II	[redacted]	£38,128
Comparator	BSC	RFB	[redacted]	83% T/T dominates
Subgroup	Updated OS	Original OS	[redacted]	£47,369
OS and PFS curve choice	Stratified log-logistic	Generalised Gamma	[redacted]	£52,234
		Log-logistic	[redacted]	£48,644
		Log-normal	[redacted]	£49,618
		Stratified Generalised Gamma	[redacted]	£57,576
		Stratified Log-normal	[redacted]	£45,848
Resource use	Total cost	+20% of total cost	[redacted]	£46,491
		-20% of total cost	[redacted]	£45,381
Utility source	Pooled sources	Cetuximab NICE submission	[redacted]	£46,487
		CORRECT study	[redacted]	£47,972
		CORRECT study – BSC utility used for all patients	[redacted]	£45,590
Discounting (Costs, LYs, QALYs)	3.5%, 0%, 3.5%	0%, 0%, 0%	[redacted]	£44,779
		6%, 6%, 6%	[redacted]	£46,999
PP treatment cost by treatment arm	Equal costs	Unequal costs	[redacted]	£48,181
KRAS status	All patients	Wild type	[redacted]	£45,919
		Mutant type	[redacted]	£51,881
BSA from RECOURSE	Not used	Used	[redacted]	£47,216
Revised TTD estimate	Used	Not used	[redacted]	£45,623
Derived SE for PP treatment cost	Not used	Used	[redacted]	£47,216
RECOURSE only AEs	Not used	Used	[redacted]	£47,216
Additional AEs	Used	Not used	[redacted]	£45,623

AE = adverse event; BSC = best supportive care; BSA = body surface area; ICER = incremental cost-effectiveness ratio; LY = life year; KRAS = Kirsten rat sarcoma viral oncogene homolog; NICE = National Institute for Health and Care Excellence; OS = overall survival; PAS = patient access scheme; PFS = progression-free survival; PP = post-progression; QALY = quality-adjusted life year; RFB = regorafenib; SE = standard error; T/T = trifluridine/tipiracil; TTD = time to treatment discontinuation

5.2.12 Subgroup analyses

T/T provided a clinically significant prolongation of OS in all treatment subgroups. Therefore, the company did not perform any subgroup analyses.

ERG comment: Treatment might be effective in all subgroups, but it does not guarantee cost effectiveness in all subgroups. Therefore, the ERG requested subgroup analyses based on the different

subgroups described in RECOURSE and the phase II trial in its clarification letter. The ERG asked for subgroup analyses, based on:

- Tumour status:
  - wild-type KRAS
  - mutant KRAS
- The time between first diagnosis of metastases and randomisation:
  - <18 months
  - ≥18 months
- Geographic region:
  - Europe only
  - United States, Europe and Australia
- Age:
  - <65 year
  - ≥65 year
- Number of prior regimens:
  - 2 and 3
  - ≥4
- ECOG PS:
  - 0
  - 1
- Number of metastatic sites:
  - 1-2
  - 3
- Liver metastases:
  - yes
  - no

NICE, however, decided not to request all these subgroup analyses to be performed by the company. The only analyses requested by NICE was the subgroup analysis based on tumour status (wild-type KRAS, mutant KRAS). The company provided results for these analyses in their response to the clarification letter. Results, based on the cost effectiveness model provided with the clarification letter[9], indicated that the company's probabilistic ICER is £51,881 for the subgroup with mutant KRAS status while it is £45,919 in the subgroup with wild-type KRAS status.

5.2.13 Model validation and face validity check

Face validity

In Section 5.10 of the CS, the company states that "*the de novo cost-effectiveness analysis was validated using a range of experts and methods, detailed in Table 82"* (Table 5.28).[1] No further details were provided concerning the face validity assessment of the model.

Table 5.27: Validation of the de novo cost effectiveness analysis
(Based on Table 82 of the CS[1])

Validation performed by	Nature of validation	Date	Aspects covered
Prof. Martin Hoyle	Full technical review	December 2015	Cost effectiveness model and section 5 of the CS.
Advisory board of health economic (and clinical) experts	Review	January 2016	Complete cost effectiveness model and submission
BresMed	Quality-control check	January 2016	Cost effectiveness model
CS = company submission			

Internal validity
Section 5.10 of the CS contains an overview of persons involved in the validation of the cost effectiveness model (Table 5.27), but no details were provided concerning how the internal validity of the model was assessed.

Cross-validation
No cross-validation of the model results was undertaken, presumably because the review of cost effectiveness studies did not identify any cost effectiveness studies relevant for the current decision problem.

External validity

Comparison with pooled trial data
The company compared the clinical outcomes (OS and PFS) obtained from the model with estimates obtained from the pooled trial data to assess whether the model accurately estimates PFS and OS. Mean PFS estimates from the model were equal to the mean PFS estimates from pooled trial data. Mean OS from the model are however longer than the mean OS obtained from the pooled trial data (for both treatment arms). The difference in OS between T/T and BSC is also larger when mean OS from the cost-effectiveness model are used (3.2 months) instead of the pooled trial data (2.3 months). Differences between modelled PFS and OS estimates and estimates from the pooled trial data are presented in Table 5.28.

Table 5.28: Summary of model results when compared with clinical data
(Based on Table 74 of the CS[1])

Outcome	Clinical trial results (pooled data)	Model result
Overall survival	*Median:* BSC: 5.4 months T/T: 7.3 months Increment: 1.9 months *Mean:* BSC: 6.8 months T/T: 9.1 months Increment: 2.3 months	*Median:* BSC: 5.3 months T/T: 7.4 months Increment: 2.1 months *Mean:* BSC: 7.9 months T/T: 11.1 months Increment: 3.2 months

Outcome	Clinical trial results (pooled data)	Model result
Progression-free survival	*Median:* BSC: 1.7 months T/T: 1.9 months Increment: 0.2 months *Mean:* BSC: 1.9 months T/T: 3.7 months Increment: 1.8 months	*Median:* BSC: 1.6 months T/T: 2.6 months Increment: 1 months *Mean:* BSC: 1.9 months T/T: 3.7 months Increment: 1.8 months
BSC = best supportive care; CS = company submission; T/T, trifluridine/tipiracil		

Comparison with cancer research UK data (CRUK)

Model outcomes were also compared with the CRUK survival estimates for Stage 4 bowel cancer. The five year survival from CRUK was compared with the two year survival of the model. This comparison was deemed suitable by the company because patients in the model already survived *35.2 months on average (i.e. approximately 3 years)* before inclusion in the trial.[1] The five year survival of CRUK was 7-8% and was considered consistent with the two year survival estimated in the model, which was 4% for the BSC group (table 51 of CS[1]).

ERG comment: Assumptions incorporated in the cost effectiveness model were clearly described in the CS. Furthermore, the economic model provided in Excel was transparent. Re-running the model confirmed the outcomes provided by the company in the CS.[1]

Face validity

Since no details were provided on face validation steps undergone during model development, the ERG asked for clarification concerning the validation efforts described in Table 5.28. In its response to the clarification letter, the company explained that the model was entirely reviewed by Professor Hoyle and that he acknowledged that the model was "*appropriate to the NICE decision problem*". Furthermore, *"The model [was] also fully reviewed by health economic and clinical experts at an advisory board. The findings of the group were that the model was appropriate to the NICE decision problem."*[9] However, no further details were provided on the different steps undergone to assess face validity of the cost-effectiveness model. The ERG was not able to judge whether the face validity of the submitted model was appropriately addressed by the company.

Internal validity

In addition, the company explained in its response to the clarification letter that *"Professor Hoyle was provided with the complete model and conducted a systematic assessment. As part of this assessment he undertook the following: validation of model inputs, parameters, results and sensitivity analyses. In addition he checked the economic model by constructing an independent simplified model"*.[9] This simplified model provided similar results to the submitted model, which eliminated the existence of major errors in the submitted cost-effectiveness model. The ERG agrees with the efforts provided to ensure internal validity.

Cross-validation

Cross-validation was not performed due to the absence of other cost effectiveness assessment for T/T versus BSC in the third treatment line of mCRC. However, a study from Goldstein et al.[59] concerning

the cost effectiveness of regorafenib was performed in the same treatment line as the current decision problem. The ERG asked the company to compare the model structure, utility estimates, resource use estimates, adverse events and outcomes between the BSC arms of the current assessment and Goldstein et al.[59] study. Despites the use of similar utility estimates, outcomes of the studies could not properly be compared because resource use estimates and total LY for the BSC arm were not described in Goldstein et al.[59]

Furthermore, the ERG asked for a comparison of the model structure, utility estimates, resource use estimates, adverse events and outcomes between the BSC arms of ID 794[28] and the current assessment. The company acknowledged the similarities in model structure and AEs profiles between the assessments, but outcomes of the studies were not deemed comparable because patients considered in the assessments are at different disease stages.

Cross-validation is consequently not thoroughly investigated in the current assessment due to the absence of comparable studies with the current assessment. The ERG agrees the impossibility to present a thorough cross-validation of the current assessment with previous studies.

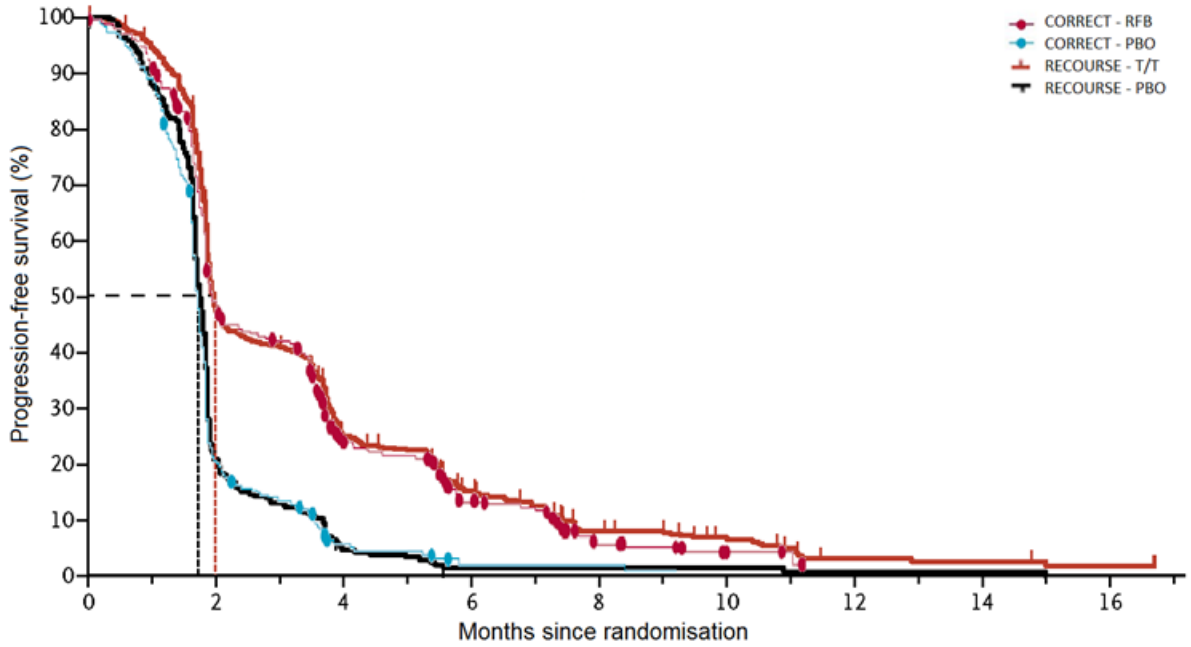
External validity

The CS contains a comparison of the survival estimates from CRUK and the current assessment. However, the ERG did not consider this comparison to be adequate because the populations from the current assessment and the CRUK were not considered comparable. The ERG consequently asked the company to explain why the external validity of the survival estimates of the model could be assessed through a comparison with data from CRUK. The company responded that they agreed that the CRUK data was not representative of the population from the current decision problem because of the following reasons: *"the data [from the CRUK] and in particular those for mCRC (stage IV) are limited by the fact that they apply to all patients with mCRC irrespective of time since diagnosis of metastatic disease, number of lines of chemotherapy received etc. Therefore the CRUK data are not reflective of the population defined by the decision problem for this appraisal."*[9] This is further justified by the fact that *"The decision problem defines a patient population diagnosed with mCRC who would have received two or more previous lines of chemotherapy (i.e. they have received NICE recommended standard therapies for mCRC and their disease has progressed or when they received the therapy they were found to be intolerant to it). Patients at this line of therapy have much lower survival than those receiving first or second line therapy."*[9] Both parties agreed that a comparison with CRUK data is not suitable for the current decision problem.

The ERG also requested a comparison of survival estimates with a study of Jonker et al.[66] However, the company was not able to conduct this comparison because the study of Jonker et al. focused on *"mCRC patients with high epiregulin (EREG) gene expression plus KRAS wild-type status"* [9], a subgroup which was not considered in the current assessment. Therefore, the results of Jonker et al. and the present assessment would unlikely be comparable, according to the company.

As an alternative, the company provided a comparison of the survival data from the CORRECT and the RECOURSE trials (Figure 5.13). As can be seen, survival curves for the placebo group (BSC) from CORRECT and RECOURSE are almost similar. However, this is not a comparison of the model results with external sources.

Figure 5.13: PFS from the RECOUSE and CORRECT studies – For T/T, PBO and RFB
(Based on Figure 10 of the response to the request for clarification[9])



PBO = placebo; PFS = progression-free survival; RFB = regorafenib; T/T = trifluridine/tipiracil

The ERG was not able to assess whether face validity was properly addressed during model development. Internal validity was correctly assessed through an entire review of the cost effectiveness model. Cross-validation could not be properly performed but trial results seemed comparable to another trial performed in the same treatment line. In conclusion, the ERG think that validation efforts of the cost effectiveness model could have been more intense but were limited by the absence of comparable assessments.

5.3 Exploratory and sensitivity analyses undertaken by the ERG

Based on all considerations from Section 5.2, the ERG defined a new base case (see Table 6.1). This base case included multiple adjustments to the original base case by the company presented in the CS.[1] These adjustments were subdivided into three categories (derived from Kaltenthaler et al. 2016[67]):

1. Fixing errors (correcting the model were the company's submitted model was unequivocally wrong)
2. Fixing violations (correcting the model where the ERG considered that the NICE reference case, scope or best practice had not been adhered to)
3. Matters of judgement (amending the model were the ERG considers that reasonable alternative assumptions are preferred)

The combination of these corrections/amendments resulted in the ERG base case. Additionally, several explorative sensitivity analyses were performed based on the ERG base case to test uncertainties within the model.

Fixing errors

The ERG identified one error in the model submitted by the Company:

1. the following adverse events rates for BSC (grade ≥ 3) were incorrect in the model (and in table 44 of the CS[67]):

- Neutropenia
- Leukopenia
- Anaemia
- Thrombocytopenia
- Increase in alanine aminotransferase level
- Increase in aspartate aminotransferase level
- Increase in total bilirubin
- Increase alkaline phosphatase level
- Increase in creatinine level

These adverse events were corrected to be in line with the published literature,[2] see Section 5.2.7 for more details.

Fixing violations
The following violations were fixed in the ERG base case to be in line with best practices and the NICE reference case.

2. Keep BSA fixed in PSA (see Section 5.2.11)
3. Correct end-of-life costs to be consistent with the NHS and PSS perspective (see Section 5.2.9)
4. Correct medical oncologist outpatient consultation costs to be consistent with the NHS reference prices (see Section 5.2.9)

Matters of judgement
5. BSA based on observed trial data (parametric estimation; see Section 5.2.9)
6. Updated costs of adverse events (see Section 5.2.9)
7. Use treatment specific post progression treatment costs (see Section 5.2.9)
8. Equal treatment delay (see Section 5.2.9)
9. Use RECOURSE data instead of pooled estimates (see Section 5.2.6)
10. Use unstratified time-to-event models for PFS and OS (see Section 5.2.6)
11. Use utilities derived from the CORRECT study (including AE disutility of 0.01 for being on TT; see Section 5.2.8)

The company and ERG base cases (with PAS) are presented in Table 5.30. Compared with the company base case, the ICER increased by approximately £9,300 to £52,695 in the ERG base case. This difference could largely be attributed to a reduction in incremental QALYs from 0.172 to 0.144. The difference between the results of the company and the ERG base case are mainly caused by the following changes in the model:

- Fixing errors with adverse events for BSC
- Use of RECOURSE data instead of pooled estimates
- Use of CORRECT utilities[30] only (i.e. not averaging with utilities from the TA176 CS report[32]).

Giving that the pooled analyses might be preferred or might not differ substantially compared with more sophisticated pooling techniques, despite the lack of justification for/use of naïve pooling (i.e. not stratifying by trial), Table 5.29 presents ERG base case using the pooled evidence. In this analyses, pooled evidence is used for OS, PFS, AE, BSA and dose reductions.

Table 5.29: Company and ERG base case (with PAS) – probabilistic results

	T/T		BSC				
	QALYs	Costs	QALYs	Costs	ΔCosts	ΔQALY	ICER
Company base case*	0.593	£17,783	0.420	£10,299	0.172	£7,484	£43,427
ERG base case	0.542	£17,167	0.398	£9,605	0.144	£7,562	£52,695
ERG base case pooled	0.561	£17,197	0.407	£9,584	0.154	£7,613	£49,392

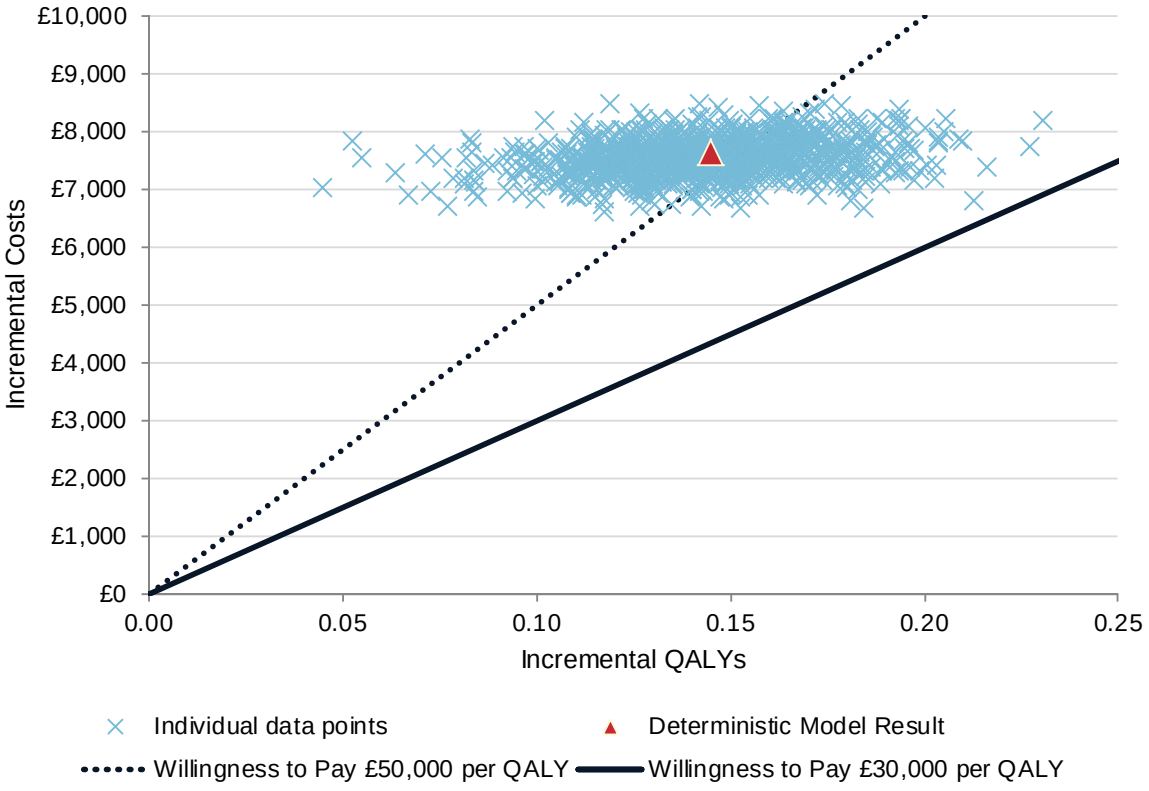
* Calculated by the ERG

BSC = best supportive care; ERG = Evidence Review Group; ICER = incremental cost-effectiveness ratio; QALY = quality-adjusted life year; T/T = trifluridine/tipiracil

5.3.1 Probabilistic sensitivity analyses (ERG base case)

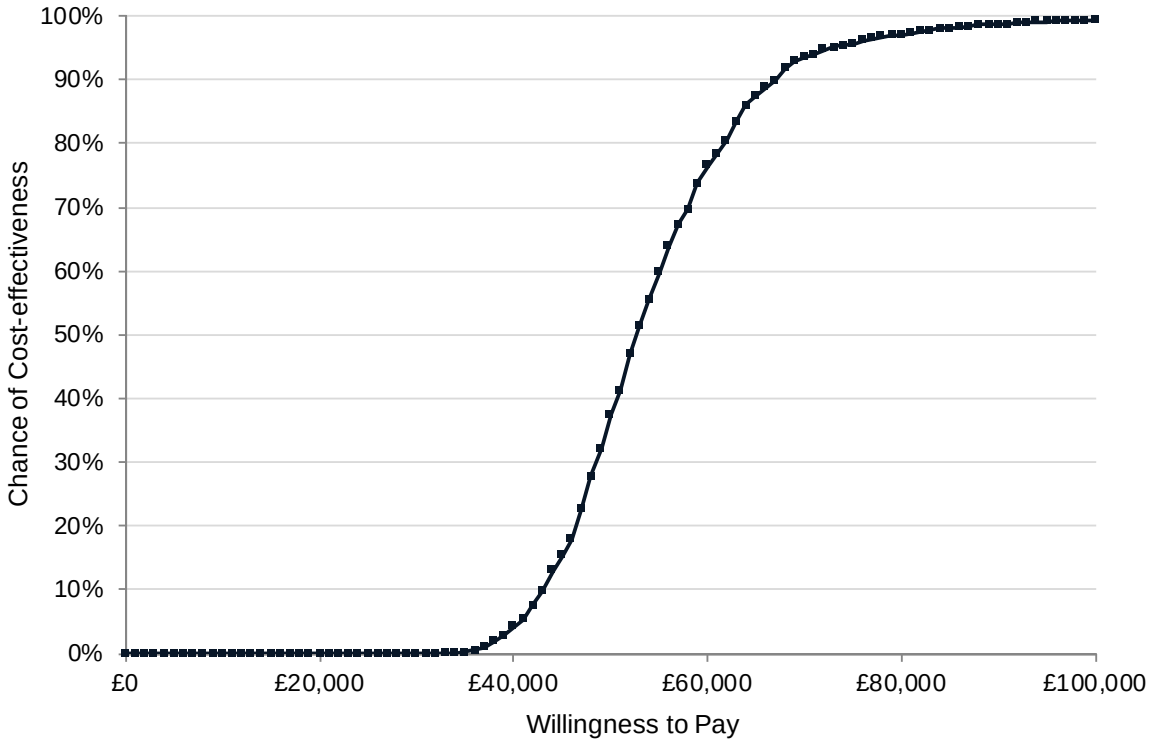
A PSA was performed to capture the uncertainty in the estimation of input parameters in the new ERG base case. Figure 5.14 presents the cost effectiveness plane and Figure 5.15 shows the cost effectiveness acceptability curves (CEACs). The probability that T/T is cost effective is smaller in the ERG base case compared to the company's base case (0% versus 0% and 37% versus 77% for thresholds of £30,000 and £50,000, respectively).

Figure 5.14: Cost effectiveness plane for all treatment options (QALYs; ERG base case)



ERG = Evidence Review Group; QALY = quality-adjusted life year

Figure 5.15: Cost effectiveness acceptability curves (ERG base case)



ERG = Evidence Review Group

5.3.2 Additional exploratory and subgroup analyses performed by the ERG base case

Additional exploratory sensitivity analyses were performed by the ERG to examine the potential impact of various alternative assumptions on the cost effectiveness estimates. These analyses were performed based on the ERG base case and illustrated that using the UK general population BSA estimates and an alternative source for resource use had a moderate impact on the results. These two analyses increased the ERG base case ICER of £52,695 to £53,776 and £54,739, respectively (Table 6.2).

Subgroup analyses based on KRAS status (Table 6.3) indicated that the ICER for the KRAS wild-type and KRAS mutant subgroups would be £53,042 and £50,721 respectively.

5.4 Conclusions of the cost effectiveness section

The economic model described in the CS is considered by the ERG to meet the NICE reference case to a reasonable extent. Reviewing the overall evidence, the ERG confirmed that there was no existing cost effectiveness model for T/T for the current indication.

In terms of population, there is uncertainty regarding the generalisability of the RECOURSE trial population to the population for whom T/T is considered in the UK. More specifically, following the licence it may be possible that patients not represented in the trial receive this medication. Additionally, as the definition of BSC was unclear, i.e. there is currently no internationally accepted definition of BSC, it is unclear whether BSC considered in the evidence, and hence in the model, is representative for BSC in the UK.

The company model follows a logical structure with respect to the nature of the disease. One of the main strengths of the CS (including the economic model) is the clarity and transparency. The cost effectiveness results were generally robust under the one-way sensitivity and scenario analyses

conducted. The model was most sensitive to changes in utility scores and selection of OS and PFS curves. Major uncertainties identified by the ERG were: whether or not to use the naïve pooling provided by the company, averaging of utilities from various sources, estimation of resource use (mainly based on expert opinion) and estimation of BSA.

The company base case ICER (probabilistic) was £43,427 (with PAS). The ERG had a total of 11 adjustments/corrections which lead to the ERG base case ICER of £52,695 (with PAS). This included fixing errors, fixing violations and matters of judgement. The most influential adjustments/corrections were 1) fixing errors with adverse events for BSC; 2) use of RECOURSE data instead of pooled estimates and; 3) use of CORRECT utilities[30] only. Fixing errors concerning adverse events rates was an issue that was unequivocally wrong in the economic model submitted by the company. Moreover, the ERG preference to use the data from the RECOURSE trial only, instead of the pooled evidence (including the phase II trial) was mainly due to the lack of justification for/use of naïve pooling by the company (i.e. not stratifying by trial) and the potential bias incurred by this adjustment was unknown (both the direction and magnitude). Nevertheless, as this is a matter of judgement and the pooled analysis might be preferred or might not differ substantially compared with more sophisticated pooling techniques, the ERG presented a pooled base case (based on pooled data of the phase II and RECOURSE trials) wherein the ICER decreased with £3,303 to £49,392. Finally, the ERG preferred to use the utilities from the CORRECT study[30] only, instead of averaging these with utility values from the CS of TA176.[32] The ERG doubts whether TA176[32] is an appropriate source for health state utilities for the present decision problem.

Exploratory sensitivity analyses illustrated that using the UK general population BSA estimates and an alternative source for resource use had a moderate impact on the ICER (£53,776 and £54,739, respectively). Subgroup analyses based on KRAS status indicated that the ICER for the KRAS wild-type and KRAS mutant subgroups would be £53,042 and £50,721, respectively.

6 IMPACT ON THE ICER OF ADDITIONAL CLINICAL AND ECONOMIC ANALYSES UNDERTAKEN BY THE ERG

In Section 5.3 the ERG base case was presented, which was based on various changes compared to the company base case. Tables 6.1 and 6.2 show how each individual change impacts the ICER plus the combined effect of all changes simultaneously. The analyses numbers in Table 6.1 correspond to the analyses numbers reported in Section 5.3. Moreover, the exploratory sensitivity and subgroup analyses are presented in Tables 6.2 and 6.3 (both conditional on the ERG base case). Appendix 3 and the economic model sent by the ERG contains technical details on the analyses performed by the ERG.

Table 6.1: ERG base case, incorporating corrections and amendments identified by the ERG (with PAS) – probabilistic results

	T/T		BSC				
	QALYs	**Costs**	**QALYs**	**Costs**	**ΔCosts**	**ΔQALY**	**ICER**
Company base case*	0.593	£17,783	0.420	£10,299	0.172	£7,484	£43,427
1-4 Fixing errors and violations	0.593	£17,494	0.421	£9,679	0.172	£7,815	£45,335
5 BSA based on observed trial data	0.593	£17,634	0.422	£10,116	0.170	£7,517	£44,120
6 Updated costs of adverse events	0.592	£18,479	0.420	£10,892	0.172	£7,587	£43,986
7 Use treatment specific post progression treatment costs	0.593	£17,642	0.422	£10,120	0.171	£7,523	£43,997
8 Equal treatment delay	0.592	£17,772	0.422	£10,241	0.170	£7,531	£44,271
9 Use RECOURSE data instead of pooled estimates	0.573	£17,320	0.416	£10,139	0.157	£7,181	£45,784
10 Use unstratified time-to-event models	0.588	£17,257	0.427	£10,259	0.161	£6,999	£43,446
11 Use CORRECT utilities	0.568	£17,754	0.401	£10,262	0.167	£7,493	£44,851
ERG base case	**0.542**	**£17,167**	**0.398**	**£9,605**	**0.144**	**£7,562**	**£52,695**
ERG base case (pooled)	**0.561**	**£17,197**	**0.407**	**£9,584**	**0.154**	**£7,613**	**£49,392**

* Calculated by the ERG

BSC = best supportive care; ERG = Evidence Review Group; ICER = incremental cost-effectiveness ratio; PAS = patient access scheme; QALY = quality-adjusted life year; T/T = trifluridine/tipiracil

Table 6.2: Exploratory sensitivity analyses based on ERG base case (with PAS) – probabilistic results

	T/T		BSC				
	QALYs	**Costs**	**QALYs**	**Costs**	**ΔCosts**	**ΔQALY**	**ICER**
ERG base case	0.542	£17,167	0.398	£9,605	0.144	£7,562	£52,695
Incorporating costs of additional AE	0.542	£17,340	0.397	£9,715	0.145	£7,625	£52,545
Use time on treatment instead of PFS	0.544	£17,510	0.398	£9,913	0.146	£7,597	£52,146
Alternative source for medical resource use (Hoyle et al. 2013[62]; table 4)	0.544	£17,162	0.397	£9,097	0.147	£8,065	£54,739
Alternative AE disutility for being on TT	0.545	£17,169	0.398	£9,616	0.147	£7,553	£51,358
Use BSA from the UK	0.543	£17,556	0.397	£9,733	0.145	£7,823	£53,776

* Calculated by the ERG

AE = adverse event; BSC = best supportive care; ERG = Evidence Review Group; ICER = incremental cost-effectiveness ratio; PAS = patient access scheme; PFS = progression-free survival; QALY = quality-adjusted life year; T/T = trifluridine/tipiracil; UK = United Kingdom

Table 6.3: Subgroup analyses based on ERG base case (with PAS) – probabilistic results

	T/T		BSC				
	QALYs	**Costs**	**QALYs**	**Costs**	**ΔCosts**	**ΔQALY**	**ICER**
KRAS wild-type	0.544	£17,281	0.398	£9,509	0.147	£7,771	£53,042
KRAS mutant	0.542	£16,925	0.397	£9,581	0.145	£7,344	£50,721

* Calculated by the ERG

BSC = best supportive care; ERG = Evidence Review Group; ICER = incremental cost-effectiveness ratio; KRAS = Kirsten rat sarcoma viral oncogene homolog; PAS = patient access scheme; QALY = quality-adjusted life year; T/T = trifluridine/tipiracil

7 END OF LIFE

According to Section 4.13.1 of the CS, T/T fulfils the criteria for end of life care.[1] The relevant table from the submission is reproduced below.

Table 7.1: Summary of the decision problem
(Based on Table 47 of the CS[1])

<table>
<tr><th>Criterion</th><th>Data available</th></tr>
<tr><td>The treatment is indicated for patients with a short life expectancy, normally less than 24 months</td><td>1. Final appraisal determination NICE TA242[7], section 4.4.19.
“For metastatic colorectal cancer that has progressed after first-line treatment, the Committee agreed that the technologies fulfil the first criterion related to life expectancy, because estimates of life expectancy from people randomised to best supportive care in the second-line setting were less than 12 months”
2. Hoyle et al. 2013[62]
Describes the cost-effectiveness analysis of cetuximab, cetuximab plus irinotecan, and panitumumab for third and further lines of treatment for KRAS wild-type patients with mCRC. This reports a mean OS for BSC of 0.51 years (6.2 months)
3. Mean OS (RECOURSE)[2]
The mean OS in the BSC arm was 0.64 years (7.7 months)
4. Mean OS pooled analysis (RECOURSE and Yoshino)[2, 3]
The mean OS in the BSC arm was 0.66 years (7.9 months)</td></tr>
<tr><td>There is sufficient evidence to indicate that the treatment offers an extension to life, normally of at least an additional 3 months, compared with current NHS treatment</td><td>The estimates of OS are based on mature survival data. The proportion of patients who had died in the RECOURSE and phase II trials were 89.0% and 72.9%, respectively.
1. Mean OS - Pooled analysis
<table>
<tr><th></th><th>Days</th><th>Months</th></tr>
<tr><td>Trifluridine/tipiracil</td><td>338</td><td>11.1</td></tr>
<tr><td>BSC</td><td>240</td><td>7.9</td></tr>
<tr><td>Incremental</td><td>98</td><td>3.2</td></tr>
</table>
2. Mean OS (RECOURSE)
<table>
<tr><th></th><th>Days</th><th>Months</th></tr>
<tr><td>Trifluridine/tipiracil</td><td>326</td><td>10.7</td></tr>
<tr><td>BSC</td><td>234</td><td>7.7</td></tr>
<tr><td>Incremental</td><td>92</td><td>3.0</td></tr>
</table></td></tr>
<tr><td>The treatment is licensed or otherwise indicated for small patient populations</td><td>1. Section 3.4.2 and section 6.1[1]
Based on the epidemiological data that are available for mCRC and expert clinical opinion, it is estimated that approximately 2,600 patients may receive further active therapy at third line or beyond (i.e. where trifluridine/tipiracil may be considered). Currently, this treatment comprises capecitabine, chemotherapy re-challenge or clinical trials
2. Market research
Pharmacor (Decision Resources Group) determined that the number of patients in the UK with mCRC (KRAS wild-type and KRAS mutation-</td></tr>
</table>

Criterion	Data available
	positive) who would be treated at third line or beyond was 2,490. Further details of the survey are available in appendix 5.[23]
BSC = Best supportive care; KRAS = Kirsten rat sarcoma viral oncogene homolog; mCRC = Metastatic colorectal cancer; NHS = National Health Service; NICE = National Institute for Health and Care Excellence; OS = overall survival; TA = Technology Appraisal	

ERG comment: The company provided evidence from various sources to support that the submission fulfils end of life criteria.

1. The first criterion of a short life expectancy includes the RECOURSE trial where survival was 7.7 months in the best supportive care arm. The ERG considers this criterion to have been met.
2. Evidence for the second criterion (an extension to life of at least three months compared to current NHS treatment) is taken from the pooled estimate of the included trials (phase II trial and RECOURSE) and for RECOURSE alone. If the more relevant figure from the RECOURSE trial is used the criterion is just met as overall incremental survival is three months exactly. The ERG notes that the pooled mean result using the actual trial data shows a mean increase in overall survival of 2.3 months (T/T: 9.1 months; BSC: 6.8 months).
3. The third criterion of a small patient population is taken from a survey by Pharmacor (see Appendix 5 of the CS for details[23]) of the number of patients in the UK with mCRC (KRAS wild-type and KRAS mutation-positive) who would be treated at third line or beyond and from the company's estimates based on a previous technology assessment[10] and expert opinion. The ERG agrees that the population to be treated is likely to be small but it is noted that the figure of 2,600 patients to be treated might be an underestimate given that the CS does not include Wales in its estimates of the incidence of mCRC.

8 OVERALL CONCLUSIONS

8.1 Statement of principal findings

The CS was based on two randomised trials (phase II trial and RECOURSE) of trifluridine/tipiracil (T/T) compared to best supportive care (BSC) alone for patients with advanced/metastatic colorectal cancer (mCRC) receiving treatment at the third line or beyond. No indirect or mixed treatment comparisons were presented The ERG agreed that the randomised trials were appropriately selected using systematic review methods and were both of high quality. Although both trials ensured consistency on medications excluded from BSC, the nature of BSC provided could vary between trial centres. The nature of BSC provided might also differ from that provided in England and Wales and this is drawn to the attention of the committee.

The phase II trial included 172 participants from Japan while RECOURSE was a multinational trial including 800 participants. RECOURSE included 394 participants from Europe (nine from the United Kingdom (UK)). The company conducted analyses demonstrating that the effect of T/T did not vary according to geographical location and as a result, the trials were pooled. There is a lack of information on methods of pooling the two included randomised trials but overall it was considered acceptable from the point of view of clinical effectiveness that the trials were pooled.

The ERG further notes that there is an under-representation of non-white, non-Asian populations across the trial (approximately 1% of RECOURSE are listed as 'black'). Considering further the issue of applicability of the trials, the population in RECOURSE is a more treated population than might be expected in practice in England and Wales. Patients were required to have received chemotherapy with fluoropyrimidine, oxaliplatin, irinotecan and bevacuzimab. Bevacuzimab is not currently available in England and Wales. A small number in the phase II trial had not received bevacuzimab (22%) but the phase II trial included fewer participants than RECOURSE. Those who did not receive bevacizumab, and are thus appropriate to the England and Wales population, represent a small percentage of the trial populations (approximately 4%). The company states that T/T might be expected to work better in a less treated population based on clinical advice. This appears to be reasonable.

The included trials do not directly assess health-related quality of life as specified in the NICE scope. Although based on the pooled result there is a benefit to patients of the median increase in overall survival of 2.3 months (T/T: 9.1 months, BSC: 6.8 months), the quality of life experienced can only be inferred from effects of disease control and occurrence of adverse events. Regarding median progression-free survival (PFS), the pooled results showed an increase of 0.2 months (T/T: 1.9 months, BSC: 1.7 months). In terms of disease control, a greater proportion of T/ T patients in both trials had stable disease (42.9% vs. 10.5% in the phase II trial and 42.4% vs. 15.9% in RECOURSE). However numbers achieving partial response or complete response were very small overall. Rates of adverse events and serious adverse events were similar between T/T and BSC for the RECOURSE trial.[25] In both trials 'treatment-related AEs' were found to be [redacted]. In RECOURSE, more patients in the BSC arm were reported to [redacted][25], [redacted].[24]

The CS provides evidence from various sources to support that the submission fulfils end of life criteria. The first criterion of a short life expectancy includes the RECOURSE trial where survival was 7.7 months in the best supportive care arm. Evidence for the second criterion (an extension to life of at least three months compared to current National Health Service (NHS) treatment) is taken from the

survival modelling calculations for the pooled estimate OS for both included trials (incremental survival: 3.2 months) and for RECOURSE alone (incremental survival: 3.0 months). The third criterion of a small patient population is taken from a survey of the number of patients in the UK with mCRC who would be treated at third line or beyond and from the company's estimates based on a previous technology assessment (approx. 2,600 patients) as well as expert opinion (2,490 patients).

The company base case ICER (probabilistic) was £43,427 (with PAS). The ERG had a total of 11 adjustments/corrections which lead to the ERG base case ICER of £52,695 (with PAS). This included fixing errors, fixing violations and matters of judgement. The most influential adjustments/corrections were 1) fixing errors with adverse events for BSC; 2) use of RECOURSE data instead of pooled estimates and; 3) use of CORRECT utilities[30] only. Fixing error concerning adverse events rates was an issue that was unequivocally wrong in the economic model submitted by the company. Moreover, the ERG preference to use the data from the RECOURSE trial only, instead of the pooled evidence (including the phase II trial) was mainly due the lack of justification for/use of naïve pooling (i.e. not stratifying by trial) and the potential bias incurred by this adjustment was unknown (both the direction and magnitude). Nevertheless, as this is a matter of judgement and the pooled analysis might be preferred or might not differ substantially compared with more sophisticated pooling techniques, the ERG presented a pooled base case (based on pooled data of the phase II and RECOURSE trials) wherein the ICER decreased by £3,303 to £49,392. Finally, the ERG preferred to use the utilities from the CORRECT study[30] only, instead of averaging these with utility values from the CS of TA176.[32] The ERG doubts whether TA176[32] is an appropriate source for health state utilities for the present decision problem.

Exploratory sensitivity analyses illustrated that using the UK general population BSA estimates and an alternative source for resource use had a moderate impact on the ICER (£53,776 and £54,739, respectively). Subgroup analyses based on KRAS status indicated that the ICER for the KRAS wild-type and KRAS mutant subgroups would be £53,042 and £50,721 respectively.

8.2 Strengths and limitations of the assessment

The company's submission contained a well-conducted systematic review which addressed the scope issued by NICE. Searches were carried out in line with the NICE guide to the methods of technology appraisal Sections 5.2.2 and 5.2.4. The CS and response to clarification provided sufficient details for the ERG to appraise the searches. The review identified two methodologically sound randomised controlled trials. The main trial, RECOURSE, was a large, multinational trial. The trials assessed the outcomes outlined by NICE with the exception of quality of life. Overall, the CS is well presented, transparent and in line with the final scope.

Considering the population, there is uncertainty regarding the generalisability of the RECOURSE trial population to the population for whom T/T is considered in the UK. More specifically, following the licence it may be possible that patients not represented in the trial receive this medication. Additionally, as the definition of BSC was unclear, i.e. there is currently no internationally accepted definition of BSC, it is unclear whether BSC considered in the evidence and hence in the model is representative for BSC in the UK.

The ERG believes incorrect search strategies for HRQoL were reported in the Appendix of the CS. The company response to the ERG clarification letter was that the reported search strategies were correct. However, the results reported in the CS suggest that separate HRQoL searches were conducted, and that four studies with HRQoL data met the inclusion criteria of the review. Without full details of the HRQoL search strategies the ERG was unable to assess their quality.

Most uncertainty in the health economic model was related to the estimation of progression free survival and overall survival as well as the utility values. Additional uncertainties identified by the ERG included whether or not to use the naïve pooling provided by the company, averaging of utilities from various sources, estimation of resource use (mainly based on expert opinion) and estimation of BSA. Using mainly expert opinion for resource use (instead of empirical data) was considered by the ERG as one of the main weaknesses is. This uncertainty might have an impact on the ICER as examined in the exploratory sensitivity analyses.

8.3 Suggested research priorities

Given the paucity of robust evidence on health-related quality of life in metastatic colorectal cancer, especially beyond first line, further research is warranted in this area. Additionally, the estimation of resource use (mainly based on expert opinion) was an area of uncertainty in the cost effectiveness model.

9 REFERENCES

[1] Servier Laboratories. *Colorectal cancer (metastatic) - trifluridine with tipiracil hydrochloride, after standard therapy [ID876]. Company evidence submission to National Institute of Health and Care Excellence. Single technology appraisal (STA)*: Servier Laboratories, 2016

[2] 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.

[3] 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.

[4] Cancer Research UK. Bowel cancer incidence by sex and UK region. (Last reviewed: 17 February 2016) [Internet]. London: Cancer Research UK, [accessed 10.3.16]. Available from: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer/incidence#heading-Zero

[5] National Cancer Intelligence Network (NCIN). Stage Breakdown by CCG 2013 [Internet]. London: National Cancer Intelligence Network (NCIN), 2013 [accessed 11.4.16]. Available from: http://www.ncin.org.uk/publications/survival_by_stage

[6] Cancer Research UK. Bowel cancer survival statistics [Internet]. London: Cancer Research UK, [accessed 10.3.16]. Available from: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer/survival

[7] 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: National Institute for Health and Care Excellence, 2012 [accessed 11.3.16] Available from: https://www.nice.org.uk/guidance/ta242

[8] National Institute for Health and Care Excellence. *Colorectal cancer: diagnosis and management. NICE clinical guideline 131 [Internet]*. London: National Institute for Health and Care Excellence, 2011 [accessed 11.3.16] Available from: https://www.nice.org.uk/guidance/cg131

[9] Servier Laboratories. *Colorectal cancer (metastatic) - trifluridine with tipiracil hydrochloride, after standard therapy [ID876]. Response to request for clarification from the ERG*: Servier Laboratories, 2016

[10] Hind D, Tappenden P, Tumur I, Eggington S, Sutcliffe P, Ryan A. The use of irinotecan, oxaliplatin and raltitrexed for the treatment of advanced colorectal cancer: systematic review and economic evaluation. *Health Technol Assess* 2008;12(15):iii-ix, xi-162.

[11] National Institute for Health and Care Excellence. NICE Pathways: Staging colorectal cancer [Internet]. London: NICE, Updated December 2015 [accessed 3.3.16]. Available from: http://pathways.nice.org.uk/pathways/colorectal-cancer/staging-colorectal-cancer

[12] Beating Bowel Cancer. Treatment options for metastatic colorectal cancer in England [Internet]. Teddington: Beating Bowel Cancer, 2015. Available from: https://www.beatingbowelcancer.org/sites/default/files/page_files/BeatingBowelCancerTreatmentOptionsMetastasesV2%20Mar2015.pdf

[13] Servier Laboratories. *TAS-102 Advisory Board Executive Summary. Data on File*: Servier Laboratories, 2016

[14] National Institute for Health and Care Excellence. *Trifluridine in combination with tipiracil hydrochloride for previously treated metastatic colorectal cancer. Final scope (pre-referral) [Internet]*. London: National Institute for Health and Care Excellence, 2015 [accessed 9.3.16] Available from: https://www.nice.org.uk/guidance/GID-TA10023/documents/final-scope

[15] Servier Laboratories. *Internal communication - patient exposure from marketing experience. Data on File*: Servier Laboratories, 2015

[16] Committee for Medicinal Products for Human Use (CHMP), European Medicines Agency. *Lonsurf: trifluridine / tipiracil. Summary of opinion (initial authorisation) [Internet]*. London: European Medicines Agency, 2016 [accessed 7.3.16] Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Summary_of_opinion_-_Initial_authorisation/human/003897/WC500202369.pdf

[17] National Institute for Health and Care Excellence. *Guide to the methods of technology appraisal 2013 [Internet]*. London: NICE, 2013 [accessed 3.3.16] Available from: http://publications.nice.org.uk/pmg9

[18] Canadian Agency for Drugs and Technologies in Health. *CADTH peer review checklist for search strategies [Internet]*. Ottawa: CADTH, 2013 [accessed 3.3.16] Available from: http://www.cadth.ca/en/resources/finding-evidence-is

[19] National Institute for Health and Care Excellence. *Single Technology Appraisal: company evidence submission template [Internet]*. London: NICE, 2015 [accessed 3.3.16] Available from: http://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisals/specification-for-company-submission-of-evidence-2015-version.docx

[20] National Institute for Health and Care Excellence. *Colorectal cancer (metastatic) - trifluridine with tipiracil hydrochloride, after standard therapy [ID876]. Clarification letter*. London: National Institute for Health and Care Excellence, 2016

[21] Scottish Intercollegiate Guidelines Network. Search filters [Internet]. Edinburgh: Scottish Intercollegiate Guidelines Network, 2015 [accessed 3.3.16]. Available from: http://www.sign.ac.uk/methodology/filters.html

[22] Centre for Reviews and Dissemination. *Systematic Reviews: CRD's guidance for undertaking reviews in health care [Internet]*. York: University of York, 2009 [accessed 3.3.16] Available from: http://www.york.ac.uk/inst/crd/SysRev/!SSL!/WebHelp/SysRev3.htm

[23] Servier Laboratories. *Colorectal cancer (metastatic) - trifluridine with tipiracil hydrochloride, after standard therapy [ID876]: appendices. Single technology appraisal (STA)*: Servier Laboratories, 2016

[24] Taiho Pharmaceutical Company Ltd. *Placebo-controlled, multicenter, double-blind, randomized, Phase II study of TAS-102 in patients with unresectable advanced or recurrent colorectal cancer who have had two or more chemotherapy regimens and who are refractory or intolerant to fluoropyrimidine, lrinotecan, and oxlliplllltin. (Cinical Study Report: J003 - 10040030). 31 August 2011. Data on File*: Taiho Pharmaceutical Company Ltd., 2011

[25] Taiho Pharmaceutical Company Ltd. *Randomised, double-blind, Phase 3 study of TAS-102 plus best supportive care (BSC) versus placebo plus BSC in patients with metastatic colorectal cancer refractory to standard chemotherapies. (Clinical Study Report: TPU-TAS-102-301). 26 August 2014. Data on File*: Taiho Pharmaceutical Company Ltd., 2014

[26] US Food and Drug Administration. *Lonsurf (trifluridine and tipiracil). NDA 207981. Summary Review [Internet]*: US Food and Drug Administration, 2015 [accessed 7.3.16] Available from: http://www.accessdata.fda.gov/drugsatfda_docs/nda/2015/207981Orig1s000SumR.pdf

[27] US Food and Drug Administration. *Lonsurf (trifluridine and tipiracil). NDA 207981. Medical Review [Internet]*: US Food and Drug Administration, 2015 [accessed 7.3.16] Available from: http://www.accessdata.fda.gov/drugsatfda_docs/nda/2015/207981Orig1s000MedR.pdf

[28] National Institute for Health and Care Excellence. *Colorectal cancer (metastatic) - cetuximab (review TA176) and panitumumab (part review TA240) (1st line). ID794 [Internet]*. London: National Institute for Health and Care Excellence, 2016 [accessed 11.3.16] Available from: https://www.nice.org.uk/guidance/indevelopment/gid-tag470

[29] Siena S, Grothey A, Sobrero A, Falcone A, Ychou M, Lenz HJ, et al. Effects of regorafenib therapy on health-related quality of life in patients with metastatic colorectal cancer in the phase III CORRECT study. Paper presented at European Cancer Congress; 27 Sep-1 Oct 2013; Amsterdam: The Netherlands. *Eur J Cancer* 2013;49(27):S482.

[30] 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.

[31] Chang J, Ngai C, Odom D, Radder C, Kappeler C, Xu R-H, 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. Paper presented at 2015 Gastrointestinal Cancers Symposium; 15-17 Jan 2015; San Francisco: United States. *J Clin Oncol* 2015;33(15 Suppl 1):697.

[32] Merck Serono Ltd. *Single technology appraisal submission: erbitux (cetuximab) for the first-line treatment of metastatic colorectal cancer [Internet]*: Merck Serono Ltd, 2008 [accessed 11.3.16]

Available from: https://www.nice.org.uk/guidance/TA176/documents/colorectal-cancer-first-line-cetuximab-merckserono2

[33] 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.

[34] Servier Laboratories. *Summary of Product Characteristics. Lonsurf®*, 2015

[35] Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer* 2009;45(2):228-47.

[36] Latimer N. *NICE DSU technical support document 14: survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data. Report by the Decision Support Unit [Internet]*. Sheffield: School of Health and Related Research (ScHARR), 2013 [accessed 10.3.16] Available from: http://www.nicedsu.org.uk/

[37] Chan KK, Tu D, O'Callaghan CJ, Au H-J, Leighl NB, Brundage MD, et al. A mapping algorithm of health preferences from EORTC QLQ C30 to health utility index mark 3 (HUI3) in advanced colorectal cancer. Paper presented at 2014 Gastrointestinal Cancers Symposium; 16-18 Jan 2014; San Francisco: United States. *J Clin Oncol* 2014;32(Suppl 3):A547.

[38] Mittmann N, Au HJ, Tu D, O'Callaghan CJ, Isogai PK, Karapetis CS, et al. Prospective cost-effectiveness analysis of cetuximab in metastatic colorectal cancer: evaluation of National Cancer Institute of Canada Clinical Trials Group CO.17 trial. *J Natl Cancer Inst* 2009;101(17):1182-92.

[39] Huxley N, Crathorne L, Varley-Campbell J, Tikhonova I, Snowsill T, Briscoe S, et al. *The clinical effectiveness and cost-effectiveness of cetuximab (review of TA176) and panitumumab (partial review of TA240) for previously untreated metastatic colorectal cancer: a systematic review and economic evaluation [Internet]*. Exeter: Peninsula Technology Assessment Group (PenTAG), University of Exeter, 2015 [accessed 11.3.16] Available from: https://www.nice.org.uk/guidance/GID-TAG470/documents/colorectal-cancer-metastatic-cetuximab-review-ta176-and-panitumumab-part-review-ta240-1st-line-id794-assessment-report2

[40] Bennett L, Zhao Z, Barber B, Zhou X, Peeters M, Zhang J, et al. Health-related quality of life in patients with metastatic colorectal cancer treated with panitumumab in first- or second-line treatment. *J Clin Oncol* 2011;29(15 Suppl):e19500.

[41] Wang J, Zhao Z, Sherrill B, Peeters M, Wiezorek J, Barber B. A Q-TWiST analysis comparing panitumumab plus best supportive care (BSC) with BSC alone in patients with wild-type KRAS metastatic colorectal cancer. Paper presented at 16th Annual International Meeting of the International Society for Pharmacoeconomics and Outcomes Research, ISPOR; 21-25 May 2011; Baltimore: United States. *Value Health* 2011;14(3):A170.

[42] Wang J, Zhao Z, Barber B, Peeters M, Wiezorek J. A Q-TWiST analysis comparing panitumumab plus best supportive care (BSC) with BSC alone in patients with wild-type KRAS metastatic colorectal cancer. *Br J Cancer* 2011;104(12):1848-53.

[43] National Institute for Health and Care Excellence. *Bevacizumab and cetuximab for the treatment of metastatic colorectal cancer. NICE technology appraisal guidance 118 [Internet]*. London: National Institute for Health and Care Excellence, 2007 [accessed 11.3.16] Available from: https://www.nice.org.uk/guidance/ta118

[44] National Institute for Health and Care Excellence. *Bevacizumab in combination with oxaliplatin and either fluorouracil plus folinic acid or capecitabine for the treatment of metastatic colorectal cancer. NICE technology appraisal guidance 212 [Internet]*. London: National Institute for Health and Care Excellence, 2010 [accessed 11.3.16] Available from: https://www.nice.org.uk/guidance/ta212

[45] 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: National Institute for Health and Care Excellence, 2014 [accessed 11.3.16] Available from: https://www.nice.org.uk/guidance/ta307

[46] Amgen. *Panitumumab for the first-line treatment of metastatic colorectal cancer [ID794]. Amgen evidence submission. Multiple technology appraisal [Internet]*: Amgen, 2015 [accessed 11.3.16] Available from: https://www.nice.org.uk/guidance/GID-TAG470/documents/committee-papers

[47] Merck Serono. *Cetuximab (review of TA176) and panitumumab (partial review of TA240) for the first line treatment of metastatic colorectal cancer. Merck Serono evidence submission. Multiple technology appraisal [Internet]*: Merck Serono, 2015 [accessed 11.3.16] Available from: https://www.nice.org.uk/guidance/GID-TAG470/documents/committee-papers

[48] Smith RD, Hall J, Gurney H, Harnett PR. A cost-utility approach to the use of 5-fluorouracil and levamisole as adjuvant chemotherapy for Dukes' C colonic carcinoma. *Med J Aust* 1993;158(5):319-22.

[49] Brown ML, Nayfield SG, Shibley LM. Adjuvant therapy for stage III colon cancer: economics return to research and cost-effectiveness of treatment. *J Natl Cancer Inst* 1994;86(6):424-30.

[50] Odom D, Barber B, Bennett L, Peeters M, Zhao Z, Kaye J, et al. Health-related quality of life and colorectal specific symptoms in patients with chemotherapy refractory metastatic disease treated with panitumumab. *Int J Colorectal Dis* 2011;26(2):173-81.

[51] Färkkilä N, Torvinen S, Roine RP, Sintonen H, Hänninen J, Taari K, et al. Health-related quality of life among breast, prostate, and colorectal cancer patients with end-stage disease. *Qual Life Res* 2014;23(4):1387-94.

[52] Färkkilä N, Sintonen H, Saarto T, Järvinen H, Hänninen J, Taari K, et al. Health-related quality of life in colorectal cancer. *Colorectal Dis* 2013;15(5):e215-22.

[53] Stein D, Joulain F, Naoshy S, Iqbal U, Muszbek N, Payne KA, et al. Assessing health state utility values in patients with metastatic colorectal cancer: a utility study in the United Kingdom and the Netherlands. *Int J Colorectal Dis* 2014;29(10):1203-10.

[54] Koukakis R, Gatta F, Hechmatt G, 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;Apr 15:[Epub ahead of print].

[55] Freeman K, Connock M, Cummins E, Gurung T, Taylor-Phillips S, Court R, et al. *Fluorouracil plasma monitoring: the My5-FU assay for guiding dose adjustment in patients receiving fluorouracil chemotherapy by continuous infusion. Diagnostic Assessment Report commissioned by the NIHR HTA Programme on behalf of the National Institute for Health and Care Excellence [Internet]*. Coventry: Warwick Evidence, 2014 [accessed 26.4.16] Available from: https://www.nice.org.uk/guidance/DG16/documents/fluorouracil-chemotherapy-the-my5fu-assay-for-guiding-dose-adjustment-dar2

[56] Porter J, Latimer N, Lee D, Hatswell AJ. Vial sizes of pharmaceuticals for infusion - the potential for cost reductions and reduced wastage by optimising fill volumes. PRM29. Paper presented at ISPOR 18th Annual European Congress; 7-11 Nov 2015; Milan: Italy. 2015.

[57] Department of Health. *Reference costs 2014-2015 [Internet]*. London: Department of Health, 2015 [accessed 3.3.16] Available from: https://www.gov.uk/government/publications/nhs-reference-costs-2014-to-2015

[58] Morgan A, Sutton A, Wailoo A. *NICE DSU report: the risk and costs of febrile neutopenia in patients with non small cell lung cancer treated with docetaxel. Report by the NICE Decision Support Unit [Internet]*. Sheffield: School of Health and Related Research (ScHARR), 2007 [accessed 29.4.16] Available from: http://www.nicedsu.org.uk/PDFs%20of%20reports/Erlotinib%20DSU%20final%20report1.pdf

[59] Goldstein DA, Ahmad BB, Chen Q, Ayer T, Howard DH, Lipscomb J, et al. Cost-effectiveness analysis of regorafenib for metastatic colorectal cancer. *J Clin Oncol* 2015;33(32):3727-32.

[60] Starling N, Tilden D, White J, Cunningham D. Cost-effectiveness analysis of cetuximab/irinotecan vs active/best supportive care for the treatment of metastatic colorectal cancer patients who have failed previous chemotherapy treatment. *Br J Cancer* 2007;96(2):206-12.

[61] Shiroiwa T, Motoo Y, 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.

[62] 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.

[63] Wolff R, Ramaekers B, van Dongen-Leunis A, Lang S, Luyendijk M, Zaim R, et al. *Bortezomib for previously untreated mantle cell lymphoma: a Single Technology Appraisal [Internet]*. York: Kleijnen Systematic Reviews Ltd, 2015 [accessed 29.4.16] Available from: https://www.nice.org.uk/guidance/TA370/documents/committee-papers

[64] Janssen. *Lymphoma (mantle cell, untreated) – bortezomib [ID724]. Company evidence submission. Single technology appraisal [Internet]*: Janssen, 2015 [accessed 29.4.16] Available from: https://www.nice.org.uk/guidance/TA370/documents/committee-papers

[65] Connock M, Cummins E, Shyangdan D, Hall B, Grove A, Clarke A. *Abiraterone acetate for the treatment of metastatic, castrate-resistant prostate cancer following previous cytotoxic chemotherapy: a Single Technology Appraisal [Internet]*. Warwick: Warwick Evidence, 2011 [accessed 29.4.16] Available from: https://www.nice.org.uk/guidance/TA259/documents/prostate-cancer-metastatic-castration-resistant-abiraterone-following-cytoxic-therapy-evidence-review-group-report2

[66] Jonker DJ, C. K, Harbison C, O'Callaghan CJ, Tu D, Simes RJ, et al. High epiregulin (EREG) gene expression plus K-ras wild-type (WT) status as predictors of cetuximab benefit in the treatment of advanced colorectal cancer (ACRC): results from NCIC CTG CO.17-A phase III trial of cetuximab versus best supportive care (BSC). Paper presented at ASCO 2009, 45th Annual Meeting of American Society Of Clinical Oncology; 29 May-2 Jun 2009; Orlando: United States. *J Clin Oncol* 2009;27(15 Suppl):4016.

[67] 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.

Appendix 1: Further critique of searches in the company submission

Clinical effectiveness

- CAS Registry numbers for the interventions were not included in the search strategies.
- There was no animal/human limit included in either the MEDLINE or Embase search strategy. This would probably have had little impact on the results because of the number of facets already combined in the strategy, and particularly the inclusion of both the precise 'advanced/metastatic' facet and 'RCT/observational studies' filter.
- The RCT search filter includes 'Review of reported cases.pt.' and 'Review, multicase.pt.': neither term identifies any records; neither term is included in the SIGN RCT filter[21] from which this is derived; and neither term is actually a publication type (pt) in MEDLINE (Ovid).
- Reporting the exact date span of the database searches would have been more transparent than using 'to present' for MEDLINE. This would allow others to replicate the search more accurately. In the list of databases given in the main CS for each of the 3 searches conducted, the date span was given as '1980 to present' for Embase, but it was then reported more specifically with the search strategies in the appendices: Embase 1980 to 2015 Week 43; Searched on 26th October 2015.
- The Cochrane Library database issue numbers were not reported. Further, the results from the Cochrane Library search would have been better reported per database rather than as a total.
- The company did not supply website addresses or details of the search strategy or search terms used for the conference searches. There are a number of ASCO and ISPOR meetings each year, and it was not clear which were searched. It would not be possible to reproduce the conference proceedings searches reported in the CS.
- There were no searches for unpublished and ongoing trials via Trials registers, e.g. ClinicalTrials.gov and ICTRP.

Cost effectiveness

- In the MEDLINE search strategy it appears that search line #26 was inadvertently combined with search line #25. Search line #25 comprises search terms for economic evaluation, whilst the facet which includes line #26 was comprised of search terms for 'models': these facets were then combined using Boolean AND. Search line #26 consisted of a set of acronyms for economic analyses (CEA, CBA, CUA, etc.) and should have been included in that facet of search terms (search line #24). In the Embase search strategy the corresponding search lines were line #33 (economic evaluation) and #32 (economic analyses acronyms).
- There were redundant search terms where hyphenated phrases have been replicated: the databases searched do not recognise hyphens, and so the same results are achieved with or without hyphens. e.g., 'cost benefit analysis' retrieves the same as 'cost-benefit analysis'.
- The Cochrane Library database issue number (NHS EED and HTA) were not reported. Further, the results from the Cochrane Library search would have been better reported per database rather than as a total.
- The cost-effectiveness facet of terms used in the Cochrane Library was inappropriate. NHS EED only consists of economic evaluations, and so this facet of terms was redundant.

Measurement and valuation of health effects

- Appendix 10 refers to the search strategy for section 5.4.3. This should be section 5.4.2.

Appendix 2: Summary list of cost effectiveness evaluation

Question(s)	*Response (Y, N or NS)*	*Comments*
Is there a clear statement of the decision problem?	Y	In the executive summary
Is the objective of the evaluation and model specified and consistent with the stated decision problem?	Y	
Is the primary decision-maker specified?	Y	
Is the perspective of the model stated clearly?	Y	
Are the model inputs consistent with the stated perspective?	N	Some of the end of life costs are not consistent with the perspective
Has the scope of the model been stated and justified?	Y	
Are the outcomes of the model consistent with the perspective, scope and overall objective of the model?	Y	
Is the structure of the model consistent with a coherent theory of the health condition under evaluation?	Y	
Are the sources of data used to develop the structure of the model specified?	Y	
Are the causal relationships described by the model structure justified appropriately?	Y	
Are the structural assumptions transparent and justified?	Y	
Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	Y	
Is there a clear definition of the options under evaluation?	N	A clear definition of BSC is missing
Have all feasible and practical options been evaluated?	Y	
Is there justification for the exclusion of feasible options?	Y	Regorafenib, the only other licensed product in the same disease stage as T/T, is not considered in the base case as it is not recommended for use in the NHS (by NICE or the CDF).
Is the chosen model type appropriate given the decision problem and specified causal relationships within the model?	Y	
Is the time horizon of the model sufficient to reflect all important differences between options?	Y	

Question(s)	*Response (Y, N or NS)*	*Comments*
Are the time horizon of the model, the duration of treatment and the duration of treatment effect described and justified?	Y	
Do the disease states (state transition model) or the pathways (decision tree model) reflect the underlying biological process of the disease in question and the impact of interventions?	Y	
Is the cycle length defined and justified in terms of the natural history of disease?	Y	
Are the data identification methods transparent and appropriate given the objectives of the model?	Partly	Unclear how health state utility values, not identified in the systematic review, were selected.
Where choices have been made between data sources, are these justified appropriately?	Partly	See above. In addition, it is unclear why the study by Siena et al (i.e. the CORRECT study)[29, 30] was preferred as the source for HRQoL data above the study by Chang et al.[31] which might potentially be consistent with the NICE reference case.
Has particular attention been paid to identifying data for the important parameters in the model?	Partly	Systematic search have been performed to identify relevant cost-effectiveness and health-related quality of life studies. However, a broader search objective and strategy (e.g. including other interventions than T/T only in the cost effectiveness review) would potentially identify cost-effectiveness studies relevant for informing the model produced by the company. For instance, the studies by Goldstein et al.,[59] Starling et al.,[60], Shiroiwa et al.,[61] and Hoyle et al.,[62] which were identified by the company but eventually excluded (see Table 2 of Appendix 6 of the CS[1]), might have been relevant for informing the model.
Has the quality of the data been assessed appropriately?	Partly	It is unclear how the quality of the data from ID794[28] is assessed.
Where expert opinion has been used, are the methods described and justified?	N	Methods for estimating resource use based on expert opinion were not described.
Is the data modelling methodology based on justifiable statistical and epidemiological techniques?	Partly	The selection of a stratified or non-stratified time-to-event model based on AIC is methodologically incorrect.
Is the choice of baseline data described and justified?	Y	
Are transition probabilities calculated appropriately?	Y	
Has a half-cycle correction been applied to both cost and outcome?	N	No half-cycle correction is required given the short (daily) cycle length.

Question(s)	*Response (Y, N or NS)*	*Comments*
If not, has this omission been justified?	Y	
If relative treatment effects have been derived from trial data, have they been synthesised using appropriate techniques?	Unclear	Pooling methods are not described
Have the methods and assumptions used to extrapolate short-term results to final outcomes been documented and justified?	Y	
Have alternative extrapolation assumptions been explored through sensitivity analysis?	Y	
Have assumptions regarding the continuing effect of treatment once treatment is complete been documented and justified?	Y	"It is noted that the long-term plausibility of the log-logistic distribution should be justified given that the curves typically predict long tails, which may not be clinically justified in some disease areas. However, Kaplan-Meier data are mature (with approximately 10% (T/T) and 5% (BSC) of patients still alive at the end of each curve); therefore, even if this is the case, OS would not be vastly over-predicted."
Have alternative assumptions regarding the continuing effect of treatment been explored through sensitivity analysis?	N	
Are the costs incorporated into the model justified?	Partly	Rationale / justification for assumptions / expert opinion regarding resource use are unclear.
Has the source for all costs been described?	Y	
Have discount rates been described and justified given the target decision-maker?	Y	
Are the utilities incorporated into the model appropriate?	N	Unclear why the utilities identified in the literature review were averaged with utilities from an alternative sources (not identified in the literature review) which does not seem to be applicable.
Is the source for the utility weights referenced?	Y	
Are the methods of derivation for the utility weights justified?	Y	
Have all data incorporated into the model been described and referenced in sufficient detail?	Y	
Has the use of mutually inconsistent data been justified (i.e. are	Y	

Question(s)	*Response (Y, N or NS)*	*Comments*
assumptions and choices appropriate)?		
Is the process of data incorporation transparent?	Y	
If data have been incorporated as distributions, has the choice of distribution for each parameter been described and justified?	N	Triangular distributions are not justified (particularly for post-progression treatment costs)
If data have been incorporated as distributions, is it clear that second order uncertainty is reflected?	N	BSA is incorporated in the probabilistic sensitivity analyses, this is more likely a reflection of first order uncertainty (i.e. variability). Moreover, reference prices, which are typically fixed are varied in the probabilistic sensitivity analyses.
Have the four principal types of uncertainty been addressed?	Partly	Patient heterogeneity was not considered.
If not, has the omission of particular forms of uncertainty been justified?	N	The justification provided: "Subgroup analysis is not considered in the de novo analysis, given the size of the patient population and that, in RECOURSE, trifluridine/tipiracil was associated with a clinically relevant prolongation in OS in all treatment subgroups" is flawed since the finding that T/T is associated with clinically relevant prolongation in OS in most treatment subgroups does not indicate that it is cost-effective in all subgroups.
Have methodological uncertainties been addressed by running alternative versions of the model with different methodological assumptions?	Y	
Is there evidence that structural uncertainties have been addressed via sensitivity analysis?	N	
Has heterogeneity been dealt with by running the model separately for different subgroups?	N	
Are the methods of assessment of parameter uncertainty appropriate?	Partly	BSA and reference prices are incorporated in the probabilistic sensitivity analyses.
If data are incorporated as point estimates, are the ranges used for sensitivity analysis stated clearly and justified?	N	Arbitrary ranges of +/- 20% of the mean are used.
Is there evidence that the mathematical logic of the model has been tested thoroughly before use?	Partly	Although the cost-effectiveness analysis was validated (see table 82 of the CS[1]), a detailed description of the validation process is missing.
Are any counterintuitive results from the model explained and justified?	N	Higher post-progression drug costs for BSC compared with T/T (see table 78 of the CS[1]) seems counterintuitive given that the post-progression drug costs are equal for both comparators and T/T

Question(s)	*Response (Y, N or NS)*	*Comments*
		has more life year in the post-progression health state. After inspecting the model, the ERG noticed that this difference was driven by the discounting of costs.
If the model has been calibrated against independent data, have any differences been explained and justified?	N	The differences between the model estimates and the data from Cancer Research UK have not been explained and justified.
Have the results of the model been compared with those of previous models and any differences in results explained?	N	Despite, the model results, in particular for BSC, could be cross validated with other economic models considering ≥3rd line treatment for mCRC. BSC cross validation might have been possible using Goldstein et al.,[59] Starling et al.,[60], Shiroiwa et al.,[61] and/or Hoyle et al.[62]
AIC = Akaike information criterion; BSC = best supportive care; BSA = body surface area; CDF = Cancer Drugs Fund; CS = company submission; ERG = Evidence Review Group; HRQoL = health-related quality of life; mCRC = metastatic colorectal cancer; N = No; NHS = National Health Service; NICE = National Institute for Health and Care Excellence; NS = not specified; OS = overall survival; T/T = trifluridine/tipiracil; UK = United Kingdom; Y = Yes		

Appendix 3: Details and deterministic ICER of ERG analyses (for validation purposes)

#		Adjusted cell(s)	Deter-ministic ICER
	Fixing errors		
1	AE in updated model (BSC)	Adverse EventsQ39:R47	£45,808
	Fixing violations		
2	Keep BSA fixed in PSA	ParametersO32:O33	£44,032
3	Correct EOL costs	ListsI54	£44,059
4	Correct Medical oncologist outpatient consultation costs	CostsF98	£44,066
1-4	Fixing errors + violations combined		£45,870
	Matters of judgement		
5	BSA based on observed trial data	DosingJ18	£44,194
6	Update costs of adverse events	Adverse EventsI30:J42 & Adverse EventsAC21:AF42	£44,658
7	Use treatment specific post progression treatment costs	CostsF80	£44,385
8	Equal treatment delay (using TT value)	Survival and ProgressionI42 & Survival and ProgressionI35	£44,407
9	Use RECOURSE data instead of pooled estimates	ControlsG15	£45,748
10	Use unstratified time-to-event models for PFS and OS	Survival and ProgressionI18 & Survival and ProgressionI21	£43,935
11	Use CORRECT utilities (including AE disutility of 0.01 for being on TT)	UtilitiesF13	£45,509
	ERG base case		£52,648
	ERG Pooled analyses	CostsF56:58 & DosingJ19 & Adverse EventsM17 & ControlsG15	£49,963
	Exploratory sensitivity analyses (conditional on ERG base case)		
	Incorporating costs of additional AE	Adverse EventsM18	£52,545
	Use time on treatment instead of PFS	TTDG13	£52,967
	Alternative source for medical resource use (Hoyle et al 2013[62]; Table 4)	Resource useI18 & CostsF97	£56,709
	Alternative AE disutility for being on TT (see ERG report)	'Adverse EventsAH22:AI46 & UtilitiesD22 & PF - IntS14	£52,090
	Use BSA from the UK	DosingJ18	£54,442