NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Trifluridine–tipiracil for treating metastatic gastric or gastro-oesophageal junction cancer after 2 or more therapies [ID1507]

Committee Papers

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Trifluridine–tipiracil for treating metastatic gastric or gastro-oesophageal junction cancer after 2 or more therapies [ID1507]

Contents:

The following documents are made available to consultees and commentators:

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

1. **Company submission** from **Servier Laboratories Ltd**

2. **Clarification questions and company responses**
   a. Clarification questions and company responses
   b. Additional Clarification response A2 A3
   c. Further Clarification response EQ-5D

3. **Patient group, professional group and NHS organisation submission** from:
   a. Royal College of Physicians

4. **Expert personal perspectives** from:
   a. Elizabeth Smyth, Clinical Expert – nominated by Servier Laboratories Ltd.
   b. Wasat Mansoor, Clinical Expert – nominated by the Royal College of Physicians.
   c. Summary of Clinical Expert discussion

5. **Evidence Review Group report** prepared by **School of Health and Related Research (ScHARR)**

6. **Company and Evidence Review Group – factual accuracy check responses**

7. **Technical engagement response** from **Servier Laboratories Ltd**

8. **Technical engagement responses from experts:**
   None received.

9. **Technical engagement response from consultees and commentators:**
   None received.

10. **Evidence Review Group addendum in response to technical engagement** prepared by **School of Health and Related Research (ScHARR)**

11. Final Technical Report

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

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Trifluridine–tipiracil for treating metastatic gastric or gastro-oesophageal junction cancer after 2 or more therapies [ID1507]

Document B
Company evidence submission

June 2019

File name	Version	Contains confidential information	Date
Servier Submission Trifluridine–tipiracil [ID 1507]	V1	Yes	

Contents

Tables and figures

Tables

Figures

List of Abbreviations

AE	Adverse event
AFT	Accelerated failure time
AIC	Akaike's Information Criteria
ALT	Alanine transaminase
ANC	Absolute neutrophil count
ASCO	American Society of Clinical Oncology
AST	Aspartate transaminase
AT	As-treated
BIC	Bayesian Information Criteria
BNF	British National Formulary
BSA	Body surface area
BSC	Best supportive care
CBT	Cognitive behavioural therapy
CI	Confidence interval
CHMP	Committee for Medicinal Products for Human Use
CR	Complete response
CT	Computed tomography
CTCAE	Common Terminology Criteria for Adverse Events
DCR	Disease control rate
DNA	Deoxyribonucleic acid
DSU	Decision Support Unit
ECOG PS	Eastern Cooperative Oncology Group Performance Status
eCRF	Electronic case report form
EGFR	Epithelial growth factor receptor
EMA	European Medicines Agency
EORTC QLQ-C30	European Organisation for Research and Treatment of Cancer QOL Questionnaire – Core

EQ5D	EuroQOL five dimensions questionnaire
ESMO	European Society of Medical Oncology
FBC	Full blood count
GEE	Generalised estimating equation
GEJ	Gastro-oesophageal junction
HER2	Human epidermal growth factor receptor 2
HR	Hazard ratio
HRQoL	Health related quality of life
HTA	Health technology assessment
ICER	Incremental cost-effectiveness ratio
INMB	Incremental net monetary benefit
ITT	Intention-to-Treat
IU	International units
IXRS	Interactive-voice web-response system
LCHP	Log-cumulative hazard plot
LFT	Liver function test
mCRC	Metastatic colorectal cancer
mGC	Metastatic gastric cancer
NICE	The National Institute for Health and Care Excellence
NCCN	The National Comprehensive Cancer Network
ORR	Objective response rate
OS	Overall survival
PAS	Patient access scheme
PD	Progressive disease
PD-L1	Programmed cell death ligand 1
PFS	Progression Free Survival
PH	Proportional hazards

PR	Partial response
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PSA	Probabilistic sensitivity analysis
PSM	Parametric Survival Model
QLQ	Quality of life questionnaire
QoL	Quality of life
RCT	Randomised controlled trial
RECIST	Response Evaluation Criteria In Solid Tumours
RFT	Renal function test
RT	Radiotherapy
ROW	Rest of world
SACT	Systemic anticancer therapy
SD	Stable disease
SE	Standard error
SLR	Systematic literature review
SmPC	Summary of Product Chararcteristics
RCT	Randomised controlled trial
TNM	Tumour-node-metastasis
TPase	Thymidine phosphorylase
TR	Tumour response
TSD	Technical support document
TTD	Time to treatment discontinuation
ULN	Upper limit of normal
VEGFR	Vascular endothelial growth factor receptor

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

B.1.1 Decision problem

The submission covers the technology's full marketing authorisation application for this indication. Details of the decision problem are presented in Table 1.[1-7]

Table 1: The decision problem

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
Population	Adults with metastatic gastric (mGC) or gastro-oesophageal junction (GEJ) cancer, who have had 2 or more previous therapies	Per the final scope issued by NICE	For clarification, please note that 2 or more previous *"therapies"* means 2 prior *"regimens/lines of treatment"* not individual drugs within a chemotherapy regimen (as consistent with proposed indication)
Intervention	Fixed-dose combination of trifluridine and tipiracil hydrochloride plus best supportive care (BSC)	Per the final scope issued by NICE	
Comparator(s)	Chemotherapy (such as docetaxel or paclitaxel monotherapy or combination chemotherapy)	BSC (including but not limited to: antiemetics, blood transfusions, oesophageal stents,	BSC is reflective of clinical practice, with current guidelines, a systematic literature review and expert opinion validating the lack of an

	BSC (including but not limited to: antiemetics, blood transfusions, oesophageal stents, palliative radiotherapy and palliative surgery)	palliative radiotherapy and palliative surgery)	evidence-based active chemotherapy option in the third-line setting. The final NICE scope highlights this lack of standard treatments for advanced or metastatic oesophago-gastric cancer previously treated with **at least 2 regimens**, with the final treatment option in the oesophago-gastric cancer NICE pathway given as second-line palliative chemotherapy (that is to say one prior line of therapy), incorporating a taxane (docetaxel or paclitaxel) and combination chemotherapy regimens. For treatment of third line mGC and GEJ cancer (that is to say two prior lines of treatment) there is an unmet need, with a lack of evidence-based treatment options demonstrating benefit beyond BSC.
Outcomes	The outcome measures to be considered include: • overall survival • progression-free survival • response rate • duration of response	Per the final scope issued by NICE	

	• adverse effects of treatment • Health-related quality of life (HRQoL)		
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 time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or	Per the final scope issued by NICE.	

	outcomes between the technologies being compared. Costs will be considered from an NHS and Personal Social Services perspective The availability of any commercial arrangements for the intervention, comparator and subsequent treatment technologies will be taken into account.		

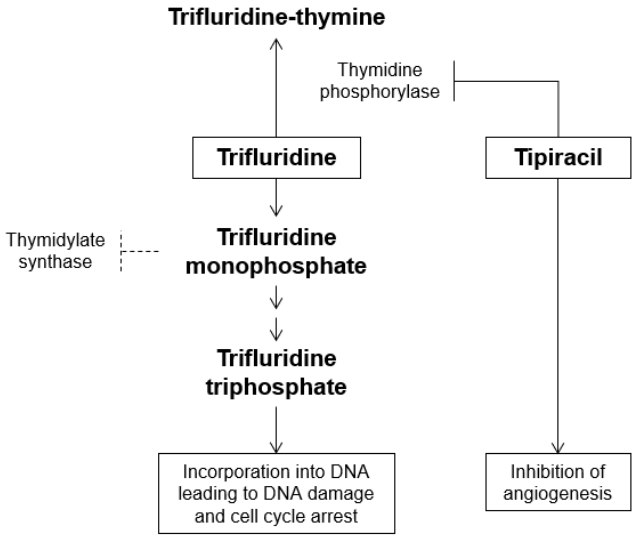
Subgroups to be considered	None specified.	No prior ramucirumab.	Ramucirumab was approved by the European Medicines Agency (EMA) in 2015 for patients with mGC who have received one prior line of therapy, that is to say second line. Within the phase III TAGS trial, it is noted that approximately two thirds of patients were not previously treated with ramucirumab (which was stratified for at randomisation). Servier is submitting analyses in the economic model pertaining to both the ITT population and the population of patients with no prior experience of ramucirumab (that is, per current UK practice for NHS patients)
Special considerations including issues related to equity or equality	No special considerations including issues related to equity or equality of life have been identified.		

B.1.2 Description of the technology being appraised

The technology being appraised is trifluridine/tipiracil in monotherapy for the treatment of adult patients with metastatic gastric or GEJ cancer, who have had 2 or more previous therapies, as described in Table 2.[5, 8, 9]

Table 2: Technology being appraised

UK approved name and brand name	Trifluridine/tipiracil (FTD/TPI; TAS-102), brand name Lonsurf®
Mechanism of action	Trifluridine/tipiracil is a novel oral cytotoxic chemotherapy comprised of an antineoplastic thymidine-based nucleoside analogue, trifluridine, and the thymidine phosphorylase (TPase) inhibitor, tipiracil hydrochloride, at a molar ratio 1:0.5 (weight ratio, 1:0.471). Trifluridine/tipiracil has a unique mechanism of action in which trifluridine is incorporated into DNA, resulting in DNA dysfunction: • Following uptake into cancer cells, trifluridine is phosphorylated by thymidine kinase, further metabolised in cells to a deoxyribonucleic acid (DNA) substrate, and incorporated directly into DNA, thereby interfering with DNA function and preventing cell proliferation. • However, trifluridine is readily metabolised by a first-pass effect following oral administration and rapidly degraded by TPase hence the inclusion of the TPase inhibitor, tipiracil hydrochloride The mechanism of action is represented in the diagram below:

	
Marketing authorisation/CE mark status	• Trifluridine/tipiracil currently holds a marketing authorisation for metastatic colorectal cancer (mCRC). • In September 2019, the marketing authorisation was extended to include the indication mGC
Indications and any restriction(s) as described in the summary of product characteristics (SmPC)	Trifluridine/tipiracil's current marketing authorisation indication for mCRC is: *Trifluridine/tipiracil is 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 chemotherapies, anti-vascular endothelial growth factor (VEGF) agents, and anti-epidermal growth factor receptor (EGFR) agents.* The additional indication for trifluridine/tipiracil is for metastatic disease: *Lonsurf is indicated as monotherapy for the treatment of adult patients with metastatic gastric cancer including adenocarcinoma of*

	the gastroesophageal junction, who have been previously treated with at least two prior systemic treatment regimens for advanced disease.
Method of administration and dosage	The recommended starting dose of trifluridine/tipiracil in adults is 35 mg/m^2/dose administered orally twice daily on days 1 to 5 and days 8 to 12 of each 28-day cycle as long as benefit is observed or until unacceptable toxicity occurs. The dosage is calculated according to body surface area (BSA). The dosage must not exceed 80 mg/dose. If doses were missed or held, the patient must not make up for missed doses. The dose may be adjusted (delayed or reduced) if the patient experiences toxicity. (Minimum dosage is 20 mg/m^2)
Additional tests or investigations	No additional tests are required for mGC patients to initiate treatment with trifluridine/tipiracil. While receiving treatment, routine tests and investigations should be undertaken.
List price and average cost of a course of treatment	Trifluridine/tipiracil is available in four presentations. The costs associated with each pack of treatment are available via the British National Formulary and are provided below for reference: – 20 x 15 mg tablets: £500.00 – 60 x 15 mg tablets: £1,500.00 – 20 x 20 mg tablets: £666.67 – 60 x 20 mg tablets: £2,000.00 The average cost per treatment as estimated by the economic model (including the PAS) is [redacted] – see Appendix J
Patient access scheme (if applicable)	There is a simple patient access scheme (PAS) discount approved as part of the previous appraisal for colorectal cancer by the Department of Health.

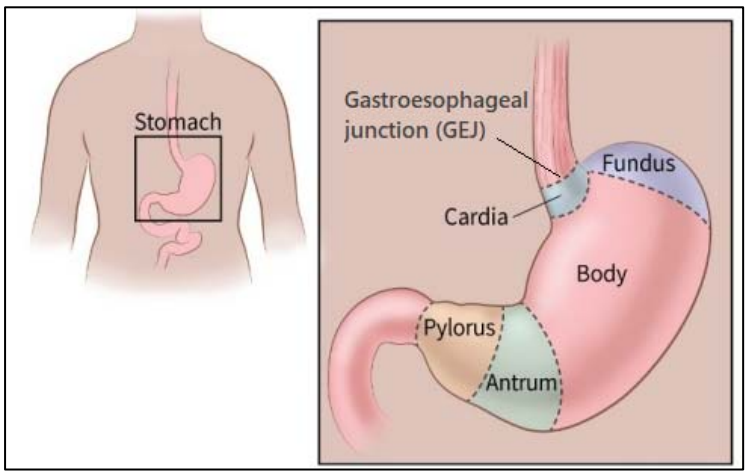
B.1.3 Health condition and position of the technology in the treatment pathway

Gastro-oesophageal cancers are malignant tumours characterised by uncontrolled cell growth in the tissues of the stomach. They are aggressive, rapidly progressing lethal diseases that have a significant impact on patients.

B.1.3.1 Anatomy and pathophysiology

Gastric cancer (including GEJ cancer) is a heterogeneous disease in which malignant cells develop in the lining of the stomach, in the region between the GEJ and pylorus.[10, 11] Please see Figure 1, below.

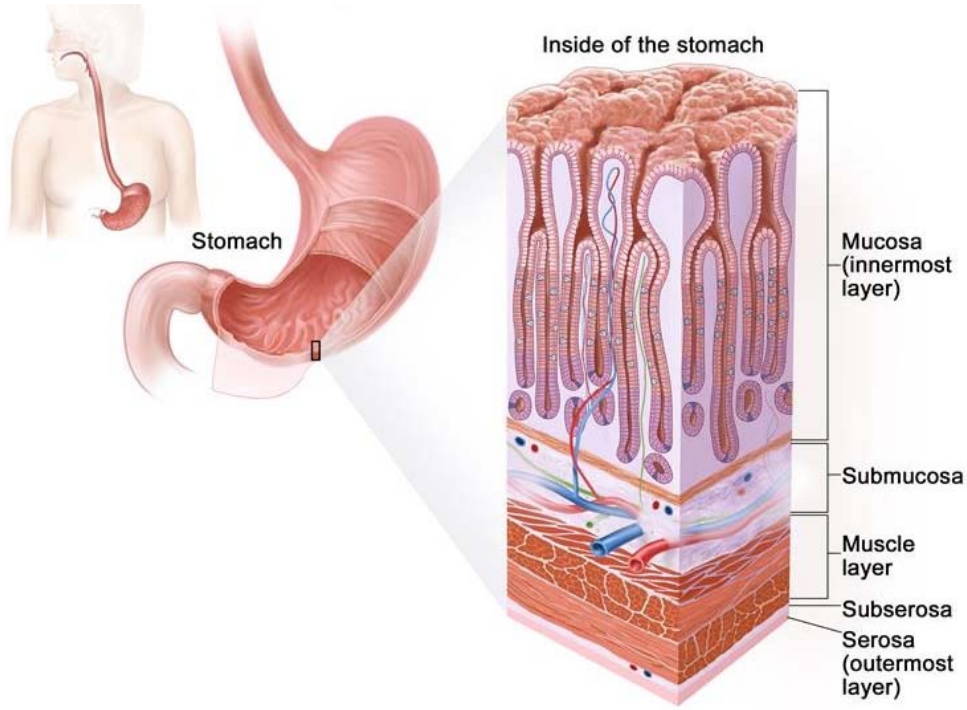
Figure 1: Gastric and gastroesophageal junction anatomy



The Lauren classification distinguishes two major subtypes of gastric cancer: diffuse (undifferentiated) and intestinal (well-differentiated).[12] These subtypes are clinically and epidemiologically distinct.

Over 90% of cases are adenocarcinomas, which develop in the cells lining the innermost mucosal layer of the stomach wall and spread through the outer layers to the muscularis (muscle layer) and serosa as they grow (Figure 2).[11, 13] As the cancer progresses, it can spread metastatically through tissue, lymph and the blood.[11] The most common sites of metastasis are the liver (48%), peritoneum (32%), lung (15%) and bone (12%).[14]

Figure 2: Layers of the stomach *(Adapted from National Cancer Institute 2018)*



The precise aetiology of gastric cancer is unknown, however there are several acknowledged risk factors associated with its development, including that it is more common in males and older patients.[15] Two other important risk factors include ethnicity, as prognosis differs between Asian and non-Asian gastric cancer populations, and infection with Helicobacter pylori (H.pylori). These and other risk factors are summarised in Table 3.[11, 13, 16]

Table 3: Risk factors

Patient Characteristics
• Male sex • Older age: sharp increase in risk over the age of 50 • Ethnicity • Type A blood • Family history of gastric cancer • Genetic predisposition (for example, carriers of *BRCA1* or *BRCA2* gene mutations) • Inherited cancer syndromes (hereditary diffuse gastric cancer, Lynch syndrome, familial adenomatous polyposis, Li-Fraumeni syndrome, Peutz-Jeghers syndrome)
Clinical Risk Factors
• Previous stomach surgery • *Helicobacter pylori* infection • Gastric diseases: chronic gastritis, pernicious anaemia, hypertrophic gastropathy, stomach ulcers • Epstein-Barr virus infection • Common variable immune deficiency

Environmental Risk Factors
• Dietary factors: diets high in salted, smoked, or preserved foods; low fruit and vegetable consumption • Obesity/over-weight • Smoking • Excessive alcohol consumption • Certain occupations: coal, metal and rubber industry workers, asbestos workers

B.1.3.2 Epidemiology

Gastroesophageal cancers are the third most frequently diagnosed cancers worldwide (with over 70% of cases occurring in Asia). They have among the highest mortality rates of all cancers, resulting in more than 1,000,000 deaths in 2018.[17] In total, there are an estimated 7,625 patients diagnosed with mGC and metastatic GEJ cancer in the UK each year.[18]

B.1.3.3 Prognosis

Although the incidence of gastric cancer has declined in recent decades (due to improved nutrition and prevention), prognosis remains poor. Early asymptomatic changes in the stomach are rarely detected and urgent investigation may not seem necessary for common, vague symptoms, the cause of which can be mistaken for other illnesses such as a stomach virus.[10, 11] Patients often only become significantly symptomatic at an advanced stage of the disease, after invasion of the muscularis propria (muscle layer) or with metastatic disease. At this point the cancer is often inoperable and current available chemotherapy options are of limited value.[10, 19]

Five-year survival is relatively good only in Japan (90%); in Europe rates vary from ~10-30%.[20] Higher survival rates in Japan are at least partially achieved by early diagnosis through endoscopic examination and consequently earlier tumour resection. If not detected early, and the patients have metastatic disease, the prognosis is poor with a median overall survival of less than a year.[10, 19] In the population relevant to this appraisal, who have been heavily pre-treated with sequential lines of chemotherapy, the median OS is approximately 4 months and survival rate at 1-year 12%.[43,4]

B.1.3.4 Diagnosis

Endoscopy determines tumour presence and anatomical location, creating an opportunity to biopsy suspicious lesions. Diagnosis involves histological classification (for example, adenocarcinoma), identification of molecular biomarkers (for example,

Human epidermal growth factor receptor 2 [HER2]) and careful tumour staging.

Numerous investigations are recommended to guide treatment and prognosis (Table 4), with clinical and pathological staging determined using the Tumour-Node-Metastasis (TNM) system as shown on Table 5.[21] These categories are then grouped into stages from 0 to IV; mGC is defined as a Stage IV cancer.[1]

Table 4: Gastric cancer investigations (Adapted from Smyth *et al* 2016)

Procedure	Purpose
Full blood count	Assess for iron deficiency anaemia
Renal and liver function	To determine appropriate therapeutic options
Endoscopy and biopsy	Obtain tissue for diagnosis, histological classification and molecular biomarkers, for example, HER2 status
Computed tomography (CT) thorax + abdomen ± pelvis	Staging of tumour, to detect local/distant lymphadenopathy and metastatic disease or ascites
Endoscopic Ultrasound	Accurate assessment of T and N stage in potentially operable tumours Determine the proximal and distal extent of tumour
Laparoscopy ± washings	Exclude occult metastatic disease involving peritoneum / diaphragm
Positron emission tomography (PET) , if available	May improve detection of occult metastatic disease in some cases

Table 5: TNM staging of gastric cancer *as per The American Joint Committee on Cancer (AJCC)*

Primary tumour (T)		Regional lymph nodes (N)		Distant metastasis (M)	
TX	Primary tumour cannot be assessed	NX	Regional lymph node(s) cannot be assessed	M0	No distant metastasis
T0	No evidence of primary tumour	N0	No regional lymph node metastasis	M1	Distant metastasis or positive peritoneal cytology
Tis	Carcinoma in situ: intraepithelial tumour without invasion of the lamina propria	N1	Metastasis in 1–2 regional lymph nodes		
T1a	Tumour invades the lamina propria or the muscularis mucosae	N2	Metastasis in 3–6 regional lymph nodes		
T1b	Tumour invades the submucosa	N3	Metastasis in 7 or more regional lymph nodes		
T2	Tumour invades the muscularis propria	N3a	Metastasis in 7–15 regional lymph nodes		
T3	Tumour penetrates the subserosal connective tissue without invasion of the visceral peritoneum or adjacent structures*	N3b	Metastasis in 16 or more regional lymph nodes		

T4	Tumour invades the serosa (visceral peritoneum) or adjacent structures†
T4a	Tumour invades the serosa (visceral peritoneum)
T4b	Tumour invades adjacent structures†

Note: *T3 tumours also include those extending into the gastrocolic or gastrohepatic ligaments, or into the greater or lesser omentum, without perforation of the visceral peritoneum covering these structures;.†Adjacent structures include the spleen, transverse colon, liver, diaphragm, pancreas, abdominal wall, adrenal gland, kidney, small intestine and retro-peritoneum.

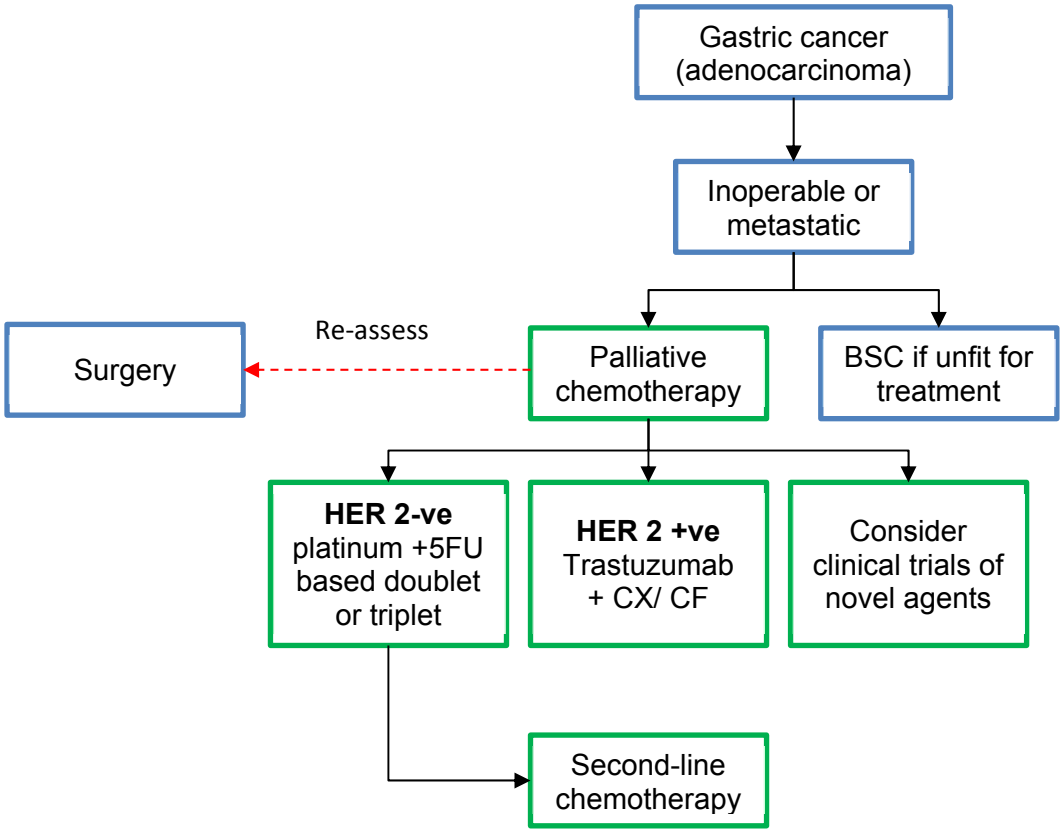
B.1.3.5 Treatment

Surgical resection is the mainstay as a potentially curative option for early-stage gastro-oesophageal cancer.[22] A minority of small tumours confined to the mucosa are appropriate for endoscopic resection.[23] For stages IB-III potentially curative surgical techniques may be an option, combined with chemotherapy to improved prognosis.[2]

Confirmed gastro-oesophageal cancer is reviewed by a multidisciplinary team to identify the optimal treatment sequencing.[6, 7]

The majority of patients are diagnosed at the metastatic stage where curative surgery is not an option.[24] These patients, including those who relapse following surgery, receive either best supportive care or a limited number of sequential lines of chemotherapy regimens as recommended by The European Society of Medical Oncology (ESMO) and NICE guidelines. However, patients who are eligible for a third-line of chemotherapy have an extremely poor prognosis and there is no standard of care. Figure 3 shows a schematic of the ESMO guidelines for inoperable or mGC.[2] Please note the pathway for inoperable advanced gastric cancer (AGC) and mGC, terms are often reported interchangeably in clinical guidance and research.[2, 6]

Figure 3: Treatment pathway overview for mGC (ESMO guidelines)



Key: BSC, Best supportive care; CF: cisplatin and 5-fluorouracil; CX: cisplatin and capecitabine; ECF: epirubicin, cisplatin and 5-fluorouracil; ECX: epirubicin, cisplatin and capecitabine; EOF: epirubicin, oxaliplatin and 5-fluorouracil; EOX: epirubicin, oxaliplatin and capecitabine; DCF: docetaxel, cisplatin and 5-fluorouracil; ESMO: European Society for Medical Oncology; FOLFIRI: folinic acid, fluorouracil and irinotecan; HER2 +ve/-ve, Human epidermal growth factor receptor 2 negative/positive.

Note: Doublet combinations of platinum and fluoropyrimidines are generally used, but triplet regimen options also include: ECF, ECX, EOF, EOX, DCF or FOLFIRI. Please note this is also the treatment pathway for inoperable advanced disease

As seen in Figure 3 above, ESMO guidelines recommend either doublet or triplet therapy as the first-line chemotherapy option.* UK expert consensus agreed that the majority of patients are treated with a doublet regimen (95%) consisting of a platinum (for example, cisplatin, oxaliplatin) and a fluropyrimidine, rather than a triplet regimen.[1-3,23]

Human epidermal growth factor receptor 2 (HER2) testing is recommended as approximately 10-15% of cases are HER2-positive; these patients are eligible to

* **Doublet regimens include:** CF (cisplatin and 5-fluorouracil) or CX (cisplatin and capecitabine). **Triplet regimens include:** ECF (epirubicin, cisplatin and 5-fluorouracil), ECX (epirubicin, cisplatin and capecitabine), EOF (epirubicin, oxaliplatin and 5-fluorouracil), EOX (epirubicin, oxaliplatin and capecitabine), DCF (docetaxel, cisplatin and 5-fluorouracil) and FOLFIRI (folinic acid, fluorouracil and irinotecan).

receive first-line treatment with HER2 targeting treatments.[2] Trastuzumab demonstrated significant improvements in response; thereby, justifying its addition to platinum/fluoropyrimidine-based doublets for the HER2-positive subtype. Patients may also be considered for clinical trials of novel agents.[2, 6, 7]

Many patients have disease progression or recurrence after first line chemotherapy and options are limited for these patients. Second-line palliative chemotherapy is guided by the initial therapy choice and performance status of the patient.[1, 6, 7] Although no standard regimen is recommended in guidelines, three cytotoxic agents have demonstrated survival benefit in second line: paclitaxel, docetaxel and irinotecan. In addition, the anti-vascular endothelial growth factor (VEGF) monoclonal antibody ramucirumab in monotherapy or combination with paclitaxel, gained marketing authorisation in 2015 for patients previously treated with chemotherapy (with one prior regimen). However, following a review of the data available, it was not considered cost-effective use of NHS resources.[5-7]

mGC patients having received two rounds of chemotherapy rarely have a performance status sufficient to be considered candidates for treatment in the third line setting. Beyond second-line, treatments have generally failed to offer meaningful benefit over BSC,[25] reflected by the absence of a standard third-line option.[1, 2, 6, 7] The ESMO guidelines state treatment options may be used sequentially in second-line and third-line settings, but evidence is lacking for a survival benefit of chemotherapy beyond second-line treatment.[2, 26] As a result, off-label chemotherapy regimens are sometimes used as a last resort to treat such patients despite a lack of evidence, or they may be enrolled in clinical trials. Patients in the third-line setting for mGC have a very poor prognosis of approximately 4 months.[4, 27, 28] Hence there is a need for improved and evidence-based therapies in third line mGC.

NICE does not currently recommend a third-line palliative chemotherapy for metastatic gastro-oesophageal cancer.[6, 7] If approved, trifluridine/tipiracil would provide a treatment option for adult patients with mGC or GEJ adenocarcinoma previously treated with at least two prior lines of chemotherapy.

B.1.4 Equality considerations

No special considerations, including issues related to equity or equality, have been identified.

B.2 Clinical effectiveness

B.2.1 Identification and selection of relevant studies

Servier Laboratories Limited commissioned BresMed Health Solutions Limited to prepare a systematic literature review (SLR). The objective was to identify evidence relating to clinical efficacy, safety and tolerability for third and later lines of treatment for advanced and/or mGC (including GEJ cancer).

Aware of ongoing relevant clinical trials relating to trifluridine/tipiracil, the commissioned search aimed to identify the wider evidence relevant to the research question, namely, trifluridine/tipiracil and any potential comparators, including those detailed in the final scope (Table 1). The full SLR methodology is presented in Appendix D: Identification, selection and synthesis of clinical evidence.

As discussed in the decision problem meeting, the BresMed search identified studies published up until the 27th June 2018, and so to supplement this, a subsequent search was prepared by a single Servier UK employee. This identified studies relevant to the research question, published between 27th June 2018 and 28th February 28th. See Appendix D).

A summary of the original BresMed methodology, highlighting any adaptations made for the supplementary search, have been detailed below.

B.2.1.1 Search strategy

The literature review included searches of the following electronic databases as standard evidence sources for clinical data as per NICE recommendations:[29]

- MEDLINE® In-Process (using Pubmed.com)
- Embase® and MEDLINE (using Embase.com)
- The Cochrane Library (using wiley.com), including the following:
  - The Cochrane Database of Systematic Reviews (CDSR)

- Database of Abstracts of Reviews of Effectiveness (DARE)
- Cochrane Central Register of Controlled Trials (CENTRAL)
- Health Technology Assessment Database (HTAD)

Electronic searching in the literature databases was not restricted by timeframe, language or interventions. However, searches were restricted by line of therapy and metastatic disease stage. In addition, conference proceedings were hand searched for the last 2 years. See Appendix D.

The only restriction on the supplementary search was the required confining of the timeframe, covering the same resources over the period 27th June 2018 to 28th February 2019. See Appendix D.

B.2.1.2 Study selection

Inclusion and exclusion criteria for clinical studies are specified in Table 6, in terms of population, interventions, comparators, outcomes, study type and other criteria.

Table 6: Eligibility criteria used in search strategy

Category	Inclusion criteria	Exclusion criteria
Population	Adult patients (≥18 years) Advanced or metastatic GC • T4 N1–3 M0, T1–4 N3 M0, and any T or N with an M1 according to TNM criteria • Stage IIIb (T3 N2 M0) and IV according to the American Joint Committee on Cancer guidelines • Unresectable Patients with gastroesophageal junction cancer	Paediatric patients Patients with early stage/newly diagnosed GC Patients with cancer/ adenocarcinoma solely of the oesophagus. Studies assessing locally-advanced GC were flagged for future reference
Line of therapy	Patients receiving third or later lines of therapy	Treatment-naïve patients or patients receiving first or second lines of therapy
Interventions	All pharmacological interventions	Non-pharmacological interventions
Comparators	Placebo	None

	Best supportive care (author defined) including active symptom control Any other pharmacological agents No comparator limit for single-arm trials A study was included if it was a clinical trial investigating at least one of the pharmacological interventions.	
Outcomes	Response rate (ORR, CR, PR, SD, PD, DCR) Overall survival Progression-free survival Mortality HRQoL Incidence of adverse events Study/treatment discontinuation	None
Study type	RCTs Non-RCTs Single arm trials† Systematic reviews*	Preclinical studies Comments, letters, editorials Case reports, case series Pharmacokinetic and economic studies Observational studies
Sample size	Studies assessing minimum of 20 patients	Patient population ≤20 sample size
Time limit	No restrictions	None
Language	No restrictions	None
Countries	No restrictions	None

Key: CR, complete response; DCR, disease control rate; GC, gastric cancer; HRQoL, health-related quality of life; ORR, overall response rate; PD, progressive disease; PR, partial response; SD, stable disease; RCT, randomised controlled trial; TNM, Tumour Node Metastasis classification of malignant tumours.

Note: A study was also included if it was a clinical trial and ≥80% of the whole study population met all inclusion criteria that is, advanced/metastatic GC (including GEJ/GOJ) patients receiving third or later lines of therapy. *Systematic reviews were utilised only for bibliography searches. † Phase 1 dosing studies excluded in supplementary search

B.2.1.3 Identification of studies

All retrieved studies were assessed against the eligibility criteria (Table 6). See Appendix D for details of the data selection and extraction process for the BresMed SLR and the supplementary search.

The initial BresMed SLR identified 9,807 publications, and subsequently the relevant data were extracted from 26 unique studies from 89 publications.

The subsequent review conducted by Servier identified 1,302 records, with 1,244 remaining once adjusted for duplicates. Screened by title and abstract for relevance to the research question, 1,121 papers failed to meet the inclusion requirements, with 123 full articles selected for further evaluation (Appendix D). Subsequently, seven records contributed towards the qualitative review of evidence, covering the period 27th June 2018 to 28th February 2019. Three of these studies provided novel data relating to third- or later-line treatment,[30-32] with the remaining four contributing an update on the BresMed SLR.[4, 33-35*] (Appendix D)

The details of the flow of studies are presented Figure 4 and Figure 5 using two separate PRISMA flow diagrams for the initial SLR and subsequent search.

B.2.1.4 Qualitative review of clinical effectiveness evidence

Overview of randomised controlled trials (RCTs)

From the BresMed SLR 10 RCTs were identified, six were excluded for reporting only sub-group data in the third-line setting, with their treatment effects presented only for the entire cohort. [36,37,38-41] The four remaining RCTs included patients solely with mGC and/or GEJ cancer treated third-line and beyond and assessed OS as the primary endpoint: [28, 42-44]

- In TAGS, trifluridine/tipiracil achieved significantly prolonged OS and PFS compared to placebo plus BSC.[4 44]
- Based on two trials, apatinib had prolonged survival and a greater response rate compared to placebo.[42, 43]
- Nivolumab significantly prolonged OS and PFS compared to placebo.[28]

The supplementary search strategy identified a further four relevant records:

* New data relating to records revealed by the commissioned SLR, such as an extend follow-up or subgroup analysis, were linked to the original reference. Full texts were obtained, along with any novel data identified, for inclusion in the qualitative review.

- In another study investigating apatinib in Asian patients receiving one or two lines of therapy, the third-line treatment effect could not be ascertained, with outcomes reported for the total population with no subgroup analysis.[31]
- JAVELIN Gastric 300 was identified by the BresMed SLR as an ongoing phase III trial assessing avelumab.[45, 46] The final publication reported it failed to meet the primary endpoint of improving OS and the secondary PFS endpoint, versus physician's choice of chemotherapy as third-line treatment.[33]
- The search identified follow-up results subsequently reported for the nivolumab RCT identified by the BresMed SLR.[28] As of data cut-off on February 2018, 2 years after last patient enrolment, the OS was reported to be maintained.[35]
- The BresMed SLR reported on the efficacy and safety findings of trifluridine/tipiracil using the Servier Clinical Study Report.[44] The article was subsequently published and identified in the search.[4]

Overview of non-RCTs and single-arm trials

Sixteen studies were included in the BresMed SLR, two non-RCTs (comparative studies)[47, 48] and 14 single-arm trials (non-comparative studies), presenting on a range of treatments assessed in third-line and beyond.[47-61 56, 62-64] The key findings highlighted by the BresMed SLR included:

- Irinotecan in combination with pemetrexed achieved an overall response rate (ORR) of 21.2%.[48] An ongoing randomised phase III study comparing irinotecan ± ramucirumab was identified (RINDBeRG trial).[65] However, only enrolling a Japanese population, it opened to accrual in February 2017 with no further updates found.[65]
- The EPOC1201 trial reported survival and safety data trifluridine/tipiracil, median OS and PFS were 8.7 and 2.9 months, respectively.[49]
- The highest response rate (ORR=45.5%) was achieved by trastuzumab deruxtecan.[61]
- The highest OS was achieved by two separate treatments. The median OS was 10.7 months following everolimus treatment[52] in one study and 10.6 months with a combination of durvalumab and tremelimumab in another.[54]
- The highest median PFS was achieved by trastuzumab deruxtecan (5.8 months).[61]

The supplementary search identified three further studies:

- A phase II single centre study reported results on apatinib plus chemotherapy versus chemotherapy alone in second-line and beyond. However, outcomes were reported for the population with no subgroup analysis for later lines.[30]

- The FOLFIRI regimen (irinotecan + folinic acid + 5-fluorouracil) as a third-line of treatment was assessed in patients progressing after ramucirumab-based second-line treatment. Based on an ORR and DCR endpoints, the authors concluded the regimen provided poor efficacy.[32]

- CheckMate 032 aimed to assess nivolumab and nivolumab plus ipilimumab. The updated search identified the full publication, presenting the primary endpoint of ORR.[34] Of 160 treated patients (59 with nivolumab 3 mg/kg, 49 with nivolumab 1 mg/kg plus ipilimumab 3 mg/kg, 52 with nivolumab 3 mg/kg plus ipilimumab 1 mg/kg), investigator-assessed objective response rates were 12% (95% CI, 5% to 23%), 24% (95% CI, 13% to 39%), and 8% (95% CI, 2% to 19%) in the three groups, respectively.[34]

Conclusion

The BresMed SLR and subsequent search highlighted the limited availability of quality evidence involving advanced and mGC (including GEJ) patients who have failed at least two lines of therapy. Several treatments (such as avelumab and everolimus) demonstrated potential within the context of early phase trials, but when subsequently studied within an RCT failed to meet the primary endpoint(s).[33, 37] Numerous potential comparators identified from the searches were only assessed using small single-arm studies. The inclusion of sub-groups of patients treated with third-line therapy and exclusively Asian cohorts were further limiting factors in ascertaining treatment effects.

Early data regarding the use of immune-checkpoint inhibitors (such as nivolumab) demonstrate potential, yet uncertainty remains regarding their positioning in the continuum of care. Nivolumab met efficacy and safety endpoints, with only all Asian patients enrolled, the Committee for Medicinal Products for Human Use (CHMP) were uncertain if the outcomes could be generalised.[66] Requesting further data to establish if the benefits outweighed risks for European patients, the manufacturer decided to withdraw the application for a European indication.[66]

Most of the study drugs identified have been experimental and have either failed to receive marketing authorisation or reimbursement (e.g. ramucirumab) in the UK setting. With the exception of the TAGS trial (trifluridine/tipiracil plus BSC vs. placebo plus BSC), studies identified by both the BresMed SLR and supplementary search were deemed outside of the scope, concluding BSC to be the most appropriate comparator.

Figure 5: PRISMA diagram for BresMed SLR

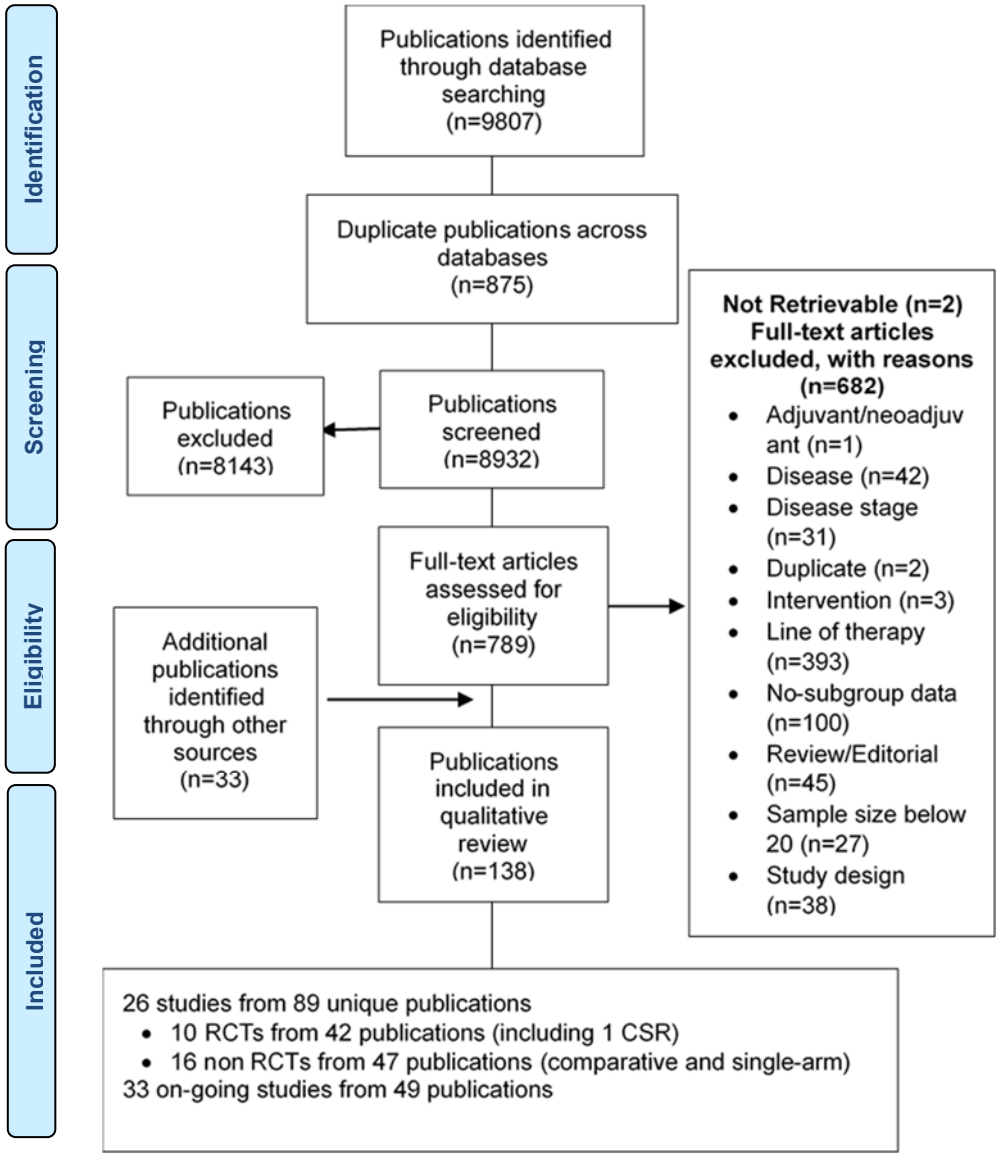
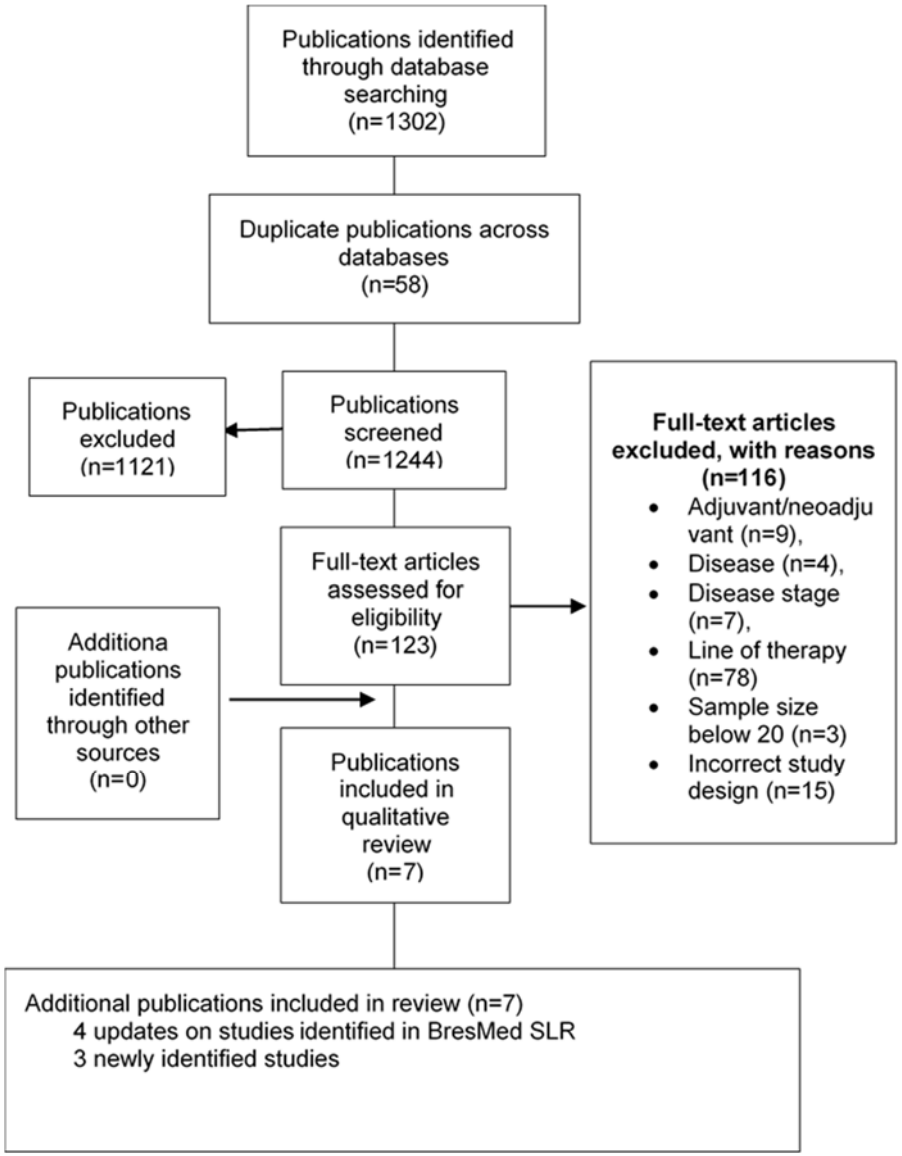


Figure 4: PRISMA diagram for supplementary search



B.2.2 List of relevant clinical effectiveness evidence

Table 7 below gives a summary of the TAGS trial.[4]

Table 7: Clinical effectiveness evidence

Study	NCT02500043 Trifluridine/tipiracil versus placebo in patients with heavily pre-treated mGC (TAGS): a randomised, double-blind, placebo-controlled, phase III trial Shitara K et al. Lancet Oncol 2018.				
Study design	Phase III, randomised, double-blind, placebo-controlled study at 110 sites in 18 countries to evaluate the efficacy and safety of trifluridine/tipiracil versus placebo in patients with previously treated mGC				
Population	507 patients were enrolled and randomly assigned, 337 to the trifluridine/tipiracil group and 170 to the placebo group. All patients were aged 18 or older with histologically confirmed, non-resectable, metastatic gastric adenocarcinoma (including adenocarcinoma of the GEJ) who had undergone one or two previous chemotherapy regimens (and had experienced radiological disease progression) that contained fluoropyrimidine, platinum agents, and taxanes or irinotecan.				
Intervention(s)	Oral trifluridine/tipiracil (35 mg/m² twice daily on days 1–5 and days 8–12 every 28 days) plus BSC				
Comparator(s)	Placebo plus BSC				
Indicate if trial supports application for marketing authorisation	Yes	✓	**Indicate if trial used in the economic model**	Yes	✓
	No			No	
Rationale for use/non-use in the model	Pivotal phase III RCT.				
Reported outcomes specified in the decision problem	**Overall survival:** A 2.1-month improvement in median OS (5.7 vs 3.6 months) and a 31% reduction in risk of death (HR, 0.69; 95% CI, 0.56–0.85; p=0.0003) was reported for patients on trifluridine/tipiracil vs placebo. Patients with no prior ramucirumab had a median OS of 6.0 months and 3.3 months, respectively (HR: 0.66; 95% CI: 0.51–0.86) **PFS:** The risk of disease progression was significantly lowered by 43% in the trifluridine/tipiracil group compared to the placebo group (HR 0.57, 95% CI: 0.47–0.70, two-sided p<0.0001) **DCR:** DCR was 44.1% of patients in the trifluridine/tipiracil group, compared to 14.5% of patients in the placebo group (p<0.0001)				

	ORR: The ORR was higher in the trifluridine/tipiracil group, with a rate of 4.5% (one patient achieved CR and 12 patients achieved PR) compared with 2.1% in the placebo group (3 patients achieved PR). **HRQoL:** HRQoL remained stable in both treatment groups with no clinically relevant changes from baseline, suggesting that HRQoL was maintained during treatment with trifluridine/tipiracil **Safety:** Trifluridine/tipiracil showed a predictable and manageable safety profile, consistent with that seen previously in patients with mCRC.
All other reported outcomes	**Time to ECOG PS ≥ 2:** The median time to ECOG PS deterioration of two or more was significantly longer in the trifluridine/tipiracil group compared with the placebo group (HR 0.69, 95% CI: 0.56–0.85, two-sided p=0.0005), with a median of 4.3 months (95% CI: 3.7–4.7) and 2.3 months (95% CI: 2.0–2.8) in the trifluridine/tipiracil and placebo groups, respectively

Key: BD: Twice daily; CI, Confidence interval; DCR, disease control rate; ECOG PS, Eastern cooperative oncology group performance status; HR, Hazard ratio; mGC, metastatic gastric cancer; ORR, Objective response rate; OS, Overall survival; QOL, quality of life; RCT, Randomised controlled trial.

Please note that EPOC1201,[49] a phase II study of trifluridine/tipiracil in patients with pre-treated advanced gastric cancer is not included in Sections 2.2 to 2.6 or used to populate the economic model. This study met its primary endpoint, showing positive efficacy and an acceptable toxicity profile. Although the results of this study support this submission, it is not included due to it being a phase II study conducted only in Japanese patients, with a sub-group receiving a different dose than was evaluated in the phase III TAGS trial, and a sub-group having received only 1 previous line of therapy. For further information about this study, please see Appendix L: Summary of EPOC1201 study.

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

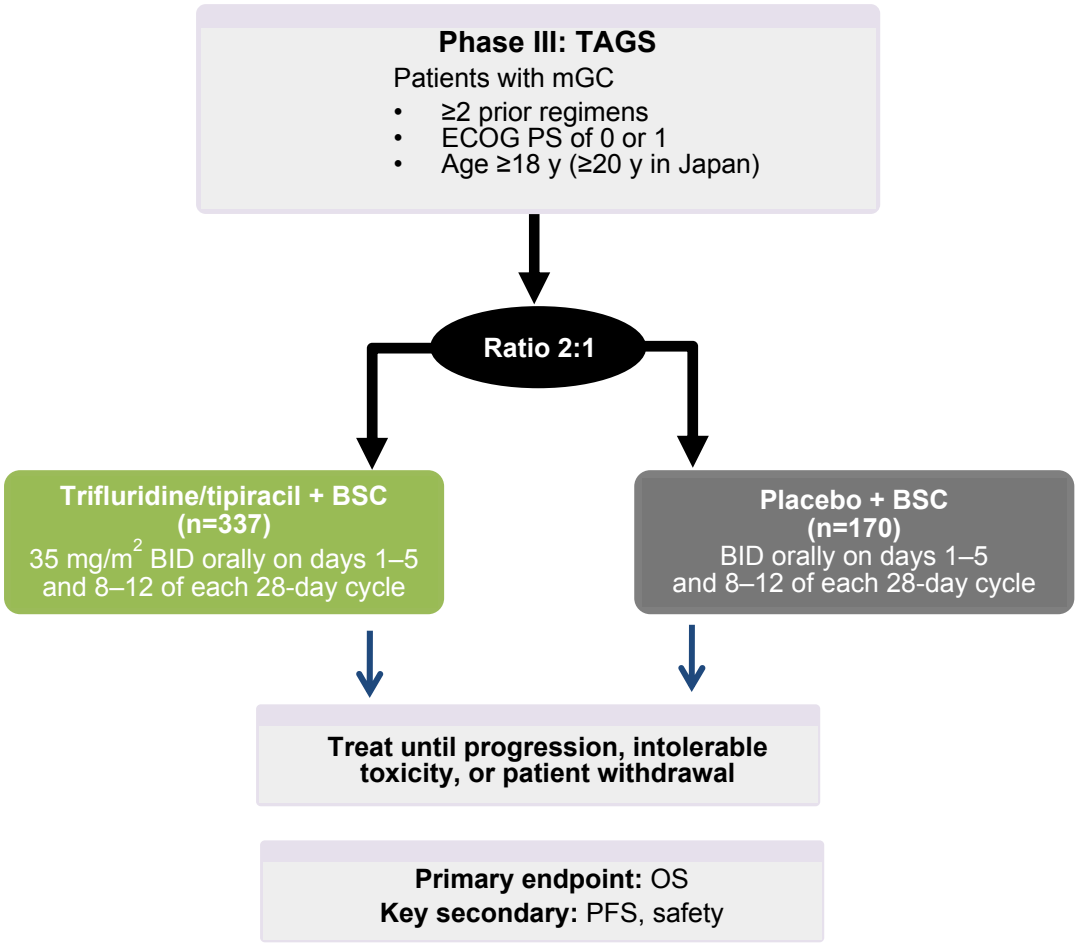
B.2.3.1 Study design

The TAGS trial was a multinational, double-blind, parallel, randomised, phase III study which investigated the efficacy and safety of oral trifluridine/tipiracil 35 mg/m^2 plus BSC versus placebo plus BSC in patients with mGC who received at least two prior regimens for advanced disease. Eligible patients were centrally randomised (2:1) to receive trifluridine/tipiracil plus BSC (experimental arm) or placebo plus BSC (control arm) and stratified by:[4, 44, 67]

- Region of the world (rest of world (ROW e.g. Europe and US) vs Japan)
- Eastern Cooperative Oncology Group Performance Status (ECOG PS) (0 vs 1)
- Prior treatment with ramicurimab (yes vs no)

The study design for the phase III TAGS trial is presented below in Figure 6.[4]

Figure 6: Study design



B.2.3.2 Inclusion and exclusion criteria

Eligibility criteria for patients enrolled into the phase III TAGS trial are described below in Table 8.[67]

Table 8: TAGS eligibility criteria

Criteria	Description
Inclusion criteria	1. Written informed consent

	2. Histologically confirmed non-resectable, metastatic gastric adenocarcinoma (including adenocarcinoma of the GEJ) as defined by the AJCC staging classification 3. Previously received at least two prior regimens (at least one cycle per regimen) for advanced disease and were refractory to or unable to tolerate their most recent prior therapy: a. Prior regimen(s) must have included a fluoropyrimidine, platinum, and either a taxane- and/or irinotecan-containing regimen; patients whose tumours were HER2+ must have received prior anti-HER2+ therapy if available b. Patients had progressed based on imaging during or within three months of the last administration of their most recent prior regimen c. Patients who had withdrawn from their most recent prior regimen due to unacceptable toxicity warranting discontinuation of treatment and precluding retreatment with the same agent prior to progression of disease were also eligible to enter the study d. Patients who received postoperative adjuvant chemotherapy or chemo-radiotherapy, and had recurrence during or within six months of completion of the adjuvant chemotherapy were allowed to count the adjuvant therapy as one prior regimen for advanced disease. Patients who received pre- and post-operative adjuvant chemotherapy, and had recurrence during or within 6 months of completion of the adjuvant chemotherapy were allowed to count the adjuvant therapy as 1 prior regimen only if the same regimen was administered both pre- and post-operatively 4. Had measurable or non-measurable disease as defined by RECIST 1.1 criteria

	5. Able to take medications orally (study treatment was not administered via a feeding tube) 6. Aged 18 years or older (20 years or older for patients in Japan) 7. ECOG PS of 0 or 1 at time of randomisation 8. Adequate organ function as defined by the following criteria: a. ANC of ≥ 1,500/mm^3 (that is, ≥ 1.5 × 10^9/L by IU) b. Platelet count ≥ 100,000/mm^3 (IU: ≥ 100 × 10^9/L) c. Haemoglobin value of ≥ 9.0 g/dL prior to randomisation based on measurements obtained 2 weeks or more after last transfusion received d. AST and ALT ≤ 3.0 × ULN; if liver function abnormalities were due to underlying liver metastasis, AST and ALT ≤ 5 × ULN e. Total serum bilirubin of ≤ 1.5 × ULN (except for Grade 1 hyperbilirubinemia due solely to a medical diagnosis of Gilbert's syndrome) f. Serum creatinine ≤ 1.5 mg/dL 9. Was willing and able to comply with scheduled visits, treatment plans, laboratory tests, and other study procedures. 10. Negative pregnancy test (urine or serum) within 7 days prior to starting the study drug. Both males and females agreed to use effective birth control during the study (prior to the first dose and for 6 months after the last dose) if conception was possible during this interval
Exclusion criteria	1. Had a serious illness or medical condition(s) including, but not limited to, the following: a. Concurrently active malignancies excluding malignancies that were disease-free for more than 5 years or carcinoma-in-situ deemed cured by adequate treatment

	b. Known brain metastasis or leptomeningeal metastasis c. Active infection (that is, body temperature ≥ 38°C due to infection) including active or unresolved pneumonia/pneumonitis d. Intestinal obstruction, pulmonary fibrosis, renal failure, liver failure, or cerebrovascular disorder e. Uncontrolled diabetes f. Myocardial infarction within 12 months prior to randomisation, severe/unstable angina, symptomatic congestive heart failure New York Heart Association class III or IV g. Gastrointestinal haemorrhage (Grade ≥ 3) within 2 weeks prior to randomisation h. Known HIV, AIDS-related illness, or chronic or acute hepatitis B or hepatitis C i. Patients with autoimmune disorders or history of organ transplantation who required immunosuppressive therapy j. Psychiatric disease that may have increased the risk associated with study participation or study drug administration, or may have interfered with the interpretation of study results 2. Had any of the following within the specified time frame prior to randomisation: a. Major surgery within prior 4 weeks b. Any anticancer therapy within prior 3 weeks c. Extended field radiation within prior 4 weeks or limited field radiation within prior 2 weeks d. Any investigational drug/device received within prior 4 weeks 3. Had previously received trifluridine/tipiracil

	4. Had unresolved toxicity of greater than or equal to CTCAE Grade 2 attributed to any prior therapies (excluding anaemia, alopecia, skin pigmentation, and platinum-induced neurotoxicity) 5. Was a pregnant or lactating female 6. Was inappropriate for entry into this study in the judgment of the Investigator 7. Had known or assumed hypersensitivity to trifluridine/tipiracil or any of its ingredients

B.2.3.3 Randomisation

Patients were enrolled by study investigators. Eligible patients were randomised (2:1) to trifluridine/tipiracil plus BSC or placebo plus BSC via a dynamic allocation method (biased coin) with an interactive-voice web-response system (IXRS). Almac (Craigavon, UK) operated the IXRS and created the algorithm that generated the individual patient allocation when the study site accessed the system. The company had no other role in the trial. Once a patient's eligibility was confirmed and the criteria for randomisation were met, study-site personnel logged on to the IXRS to allocate patients to treatment. The IXRS randomly assigned study medication (trifluridine/tipiracil or placebo) by assigning a kit number to that patient. Randomisation was stratified by region (Japan vs rest of world), ECOG performance status (0 vs 1), and previous treatment with ramucirumab (yes vs no). Patients, investigators and study-site personnel, those assessing outcomes, and those analysing the data were masked to treatment assignment. Tablets of identical appearance were used to maintain masking. Only personnel from the contract research organisations involved in drug labelling and distribution (Fisher Clinical Services [Allentown, PA, USA] and Bell Medical Solutions [Tokyo, Japan]) and IXRS activities (Almac) were aware of treatment assignment.[4, 44, 67]

B.2.3.4 Intervention for each group

Patients received either oral trifluridine/tipiracil 35 mg/m² twice daily plus BSC or placebo twice daily plus BSC on days 1–5 and days 8–12 of each 28-day treatment cycle. For the trifluridine/tipiracil treatment group:[4, 44, 67]

- Maximum dose allowed was 80 mg/dose
- Dosing could be delayed for grade 3 or worse non-haematological adverse events (except for grade 3 nausea, vomiting, or diarrhoea that responded to supportive care) until the adverse event resolved to grade 0 or 1. Treatment was then resumed, with a dose reduction of 5 mg/m².
- Dosing could be delayed if patients had neutropenia (that is, <0·5 × 10^9 neutrophils per L) or thrombocytopenia (that is, <50 × 10^9 platelets per L), until counts returned to at least 1·5 × 10^9 neutrophils per L or 75 × 10^9 platelets per L. Additionally, all patients had to have the aforementioned counts to be eligible to start subsequent cycles. Treatment was resumed at the same dose level, except in cases of grade 4 neutropenia or thrombocytopenia requiring a dosing delay of more than a week, in which case the dose was reduced by 5 mg/m².
- The minimum dose allowed was 20 mg/m² (representing a maximum of three dose reductions of 5 mg/m²), and the maximum delay allowed to the start of the next treatment cycle was 28 days.
- Treatment continued until disease progression, intolerable toxicity, or patient withdrawal.

B.2.3.5 Changes to protocol

The original study protocol was issued on the 30th Jun 2015. At the time of data cut-off for primary analysis, there were 1 administrative and 2 substantial global amendments. In addition, there were 3 country-specific amendments for Japan and 2 country-specific amendments for Germany. See Appendix M: Protocol amendments.

B.2.3.6 Setting and location

This trial was conducted in 110 academic hospitals located in 17 countries (Belarus, Belgium, Czech Republic, France, Germany, Ireland, Israel, Italy, Japan, Poland, Portugal, Romania, Russia, Spain, Turkey, the UK, and the USA). There were 7 trial sites in the UK. A full list of each trial site and number of patients recruited can be found in Appendix N.[4, 68]

B.2.3.7 Primary and secondary outcomes

The primary outcome in the TAGS trial was OS*, defined as the time from the date of randomisation until the date of death.[4, 44]

The key secondary endpoints of the trial were:[44]

- Progression Free Survival (PFS), defined as the time from the date of randomisation until the date of the investigator-assessed radiological disease progression or death due to any cause
- Safety and tolerability, based on assessment of adverse events†

Efficacy was assessed in the intention-to-treat population, while safety was evaluated in all patients who received at least one dose of treatment.[4, 21]

Other secondary endpoints included:[4, 21]

- Objective response rate (ORR), defined as the proportion of patients with complete response (CR) or partial response (PR)‡
- DCR defined as the proportion of patients with CR, PR, or stable disease (SD)
- Time to deterioration of the ECOG PS, defined as the time from randomisation until the first date on which an ECOG PS score of 2 or higher was observed

* After discontinuation of treatment, all patients were followed up for survival every 4 weeks until death or loss to follow-up, or until the targeted number of events (deaths) was met.

† Adverse events were graded according to the US National Cancer Institute's Common Terminology Criteria for Adverse Events (version 4.03) and recorded from the first dose of study drug (that is, day 1, cycle 1) until 30 days after the last dose of study drug. Haematology and serum chemistry measurements were done within the 7 days before day 1 of cycle 1, on day 15 of cycle 1, within 24 h before the start of study treatment for every cycle from cycle 2 onwards, at the end of the treatment visit (if applicable), and at the 30-day safety follow-up visit. Urinalysis was done within the 7 days before day 1 of cycle 1 and thereafter as clinically indicated.

‡ Tumour assessments by CT of the chest and abdomen (and pelvis if clinically indicated) were done for all patients within the 28 days before day 1 of cycle 1 and every 8 weeks during study treatment until radiologically confirmed disease progression. For patients who discontinued treatment for reasons other than radiologically confirmed disease progression, tumour assessments were done every 8 weeks until the development of radiological progression or the start of new anticancer treatment, whichever occurred first. On-site tumour assessments were done by investigators or local radiologists according to RECIST (version 1.1).

- HRQoL, evaluated by European Organisation for Research and Treatment of Cancer (EORTC) Quality of Life Questionnaire (QLQ-C30) and the Quality of Life (QoL) Questionnaire-Gastric-specific module (QLQ-STO22), which is a module specific to patients with gastric cancer.

B.2.3.8 Baseline characteristics

Between Feb 24th 2016, and Jan 5th 2018, 507 patients were enrolled and randomly assigned to the trifluridine/tipiracil group (n=337) or the placebo group (n=170). 503 patients received at least one dose of study treatment, 335 in the trifluridine/tipiracil group and 168 in the placebo group (safety analysis population). The patient baseline characteristics can be found below in Table 9.[4]

Table 9: Patient baseline characteristics

	Trifluridine/tipiracil (n=337)	**Placebo (n=170)**
Age (years)		
Median (IQR)	64.0 (24–89)	62.5 (32–82)
<65	183 (54%)	96 (56%)
≥65	154 (46%)	74 (44%)
Sex		
Male	252 (75%)	117 (69%)
Female	85 (25%)	53 (31%)
Ethnicity		
White	244 (72%)	113 (66%)
Asian	51 (15%)	29 (17%)
Other	4 (1%)	4 (2%)
Not available	38 (11%)	24 (14%)
Region		
USA	21 (6%)	5 (3%)
Europe*	270 (80%)	138 (81%)
Japan	46 (14%)	27 (16%)
ECOG performance status		
0	123 (36%)	68 (40%)
1	214 (64%)	102 (60%)
Primary site		
Gastric	239 (71%)	121 (71%)
GEJ	98 (29%)	47 (28%)
Both	0	2 (1)
Measurable disease	306 (91%)	150 (88%)
Histology		
Diffused	53 (16%)	21 (12%)
Intestinal	103 (31%)	52 (31%)

	Trifluridine/tipiracil (n=337)	Placebo (n=170)
Mixed	14 (4%)	8 (5%)
Unknown	132 (39%)	69 (41%)
Not available	35 (10%)	20 (12%)
HER2 status		
Positive	67 (20%)	27 (16%)
Negative	207 (61%)	106 (62%)
Not assessed	62 (18%)	37 (22%)
No. of metastatic sites		
1–2	155 (46%)	72 (42%)
≥3	182 (54%)	98 (58%)
Peritoneal metastases	87 (26%)	53 (31%)
Previous gastrectomy	147 (44%)	74 (44%)
No. of prior regimens		
2	126 (37%)	64 (38%)
3	134 (40%)	60 (35%)
≥4	77 (23%)	46 (27%)
Prior systemic cancer therapeutic agents	337 (100%)	170 (100%)
Platinum	336 (>99%[a])	170 (100%)
Fluoropyrimidine	311 (92%)	148 (87%)
Taxane[†]	183 (54%)	98 (58%)
Irinotecan[‡]	114 (34%)	55 (32%)
Ramucirumab	60 (18%)	24 (14%)
Anti-HER2 therapy	25 (7%)	7 (4%)
Immunotherapy (anti–PD-1/PD-L1)	77 (23%)	41 (24%)
Other		

Key: ECOG PS: Eastern Cooperative Oncology Group performance status; HER2: human epidermal growth factor receptor 2; PD-1: programmed death-1; PD-L1: programmed death-ligand 1

Note: Data are n (%) unless noted otherwise. *Please note that Europe refers to Belarus, Belgium, Czech Republic, France, Germany, Ireland, Israel, Italy, Poland, Portugal, Romania, Russia, Spain, Turkey, and the UK; †One patient did not receive a fluoropyrimidine; ‡All patients received irinotecan or taxane or both.

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

B.2.4.1 Study populations

The safety and efficacy study populations were defined in the protocol as follows:[67]

- **Intent-to-Treat (ITT) population:** This population included all randomised patients and is the primary population for all efficacy parameters. All analyses used in this population were based on the treatment assigned following randomisation.

- **As-Treated (AT) population:** This population included all patients who took part of any dose of the study treatment. This population was used for safety analyses. All analyses using this population were based on the treatment received.

- **Tumour Response (TR) evaluable population:** This population included all patients in the ITT population with measurable disease (at least one target lesion) at baseline and with at least one tumour evaluation while on treatment (except for early disease progression/cancer-related death). All analyses using this population were based on the treatment assigned by IXRS.

B.2.4.2 Study endpoints

- **Primary Efficacy Endpoint – OS:** Survival was the primary endpoint of this study and was defined as the time from the date of randomisation to the death date. In the absence of death confirmation or for patients alive as of the OS cut-off date, survival time was censored at the date of last study follow-up, or the cut-off date, whichever was earlier. The OS cut-off date used for the primary analysis was based on the date of the 384th death in the study.[67]

- **Secondary Efficacy Endpoint – PFS:** PFS was defined as the time 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 disease progression as of the analysis cut-off date were censored at the date of the last tumour assessment. Patients who received non-study cancer treatments before disease progression, or patients with clinical but not radiological evidence of progression were censored at the date of the last evaluable tumour assessment before the non-study cancer treatment was initiated.[67]

- **Secondary Efficacy Endpoint – ORR:** The assessment of ORR was based on investigator review of the images. ORR was defined as the proportion of patients with objective evidence of CR or PR. At the analysis stage, the best overall response was assigned for each patient as the best response recorded from all responses recorded after study randomisation. If applicable, responses recorded after disease progression or initiation of non-study cancer treatment were excluded. A patient's best response assignment of SD needed to be maintained for at least 6 weeks after study randomisation. Per RECIST 1.1, responses of PR or CR in studies with survival as the primary endpoint did not have a minimum time requirement to maintain the response.[67]

- **Secondary Efficacy Endpoint –DCR:** DCR was assessed in parallel to that of ORR, with DCR defined as the proportion of patients with objective evidence of CR, PR, or SD.[67]

- **Time to Deterioration of ECOG Performance Status:** The time to deterioration of ECOG performance status was defined as the time from randomisation to the first date on which an ECOG performance status score of 2 or higher was observed.[67]

- **QOL:** QoL was evaluated via 2 questionnaires at baseline[*], every 4 weeks and at treatment end, using:[†]

    1. **The European Organisation for Research and Treatment of Cancer QOL Questionnaire – Core (EORTC QLQ-C30)**

    The EORTC QoL questionnaire is an integrated system for assessing the HRQoL of cancer patients participating in international clinical trials. The core questionnaire, the QLQ-C30, incorporates 5 functional scales (physical, role, cognitive, emotional, and social), 3 symptom scales

[*] Patients completed the EORTC – QLQ-C30 and QLQ-STO22 questionnaires within 7 days prior to randomisation, prior to dose administration on Day 1 of Cycles ≥2, and at the 30-day safety follow-up if not performed within the prior 4 weeks.

† The number of patients completing the questionnaires decreased with each visit, as treatment discontinuation reduced the sample size. Only results at time points for which at least 10% of the initial cohort completed the questionnaires were considered valid for analysis. The 10% cut point corresponded to 6 cycles in the trifluridine/tipiracil group and 3 cycles in the placebo group.

(fatigue, pain, and nausea and vomiting), a global health status scale, and a number of single items assessing additional symptoms commonly reported by cancer patients (dyspnoea, loss of appetite, insomnia, constipation and diarrhoea) and perceived financial impact of the disease.[44, 69]

2. Gastric Specific Module (QLQ-ST022)

The gastric cancer module (QLQ-STO22) is intended for use among gastric cancer patients varying in disease stage and treatment modality. This 22-item instrument is used alongside the 30-item QLQ-C30 core questionnaire, resulting in a total of 52 items. The international field validation study demonstrated good sensitivity to changes in health status and reported that the instrument had good reliability for all the subscales.[44]

B.2.4.3 Statistical Methods used

OS and radiologically confirmed progression-free survival were analysed in the ITT population with a one-sided stratified log-rank test, with the HR and two-sided 95% CIs based on a prespecified stratified Cox model and associated Kaplan-Meier survival estimates. The Cox proportional hazards assumption was not verified for the model. One-sided and two-sided p values were presented for OS. The median follow-up time for survival was calculated with the reverse Kaplan-Meier method. A prespecified multivariate subgroup analysis was done with a Cox proportional hazards model that included the three stratification factors and potential prognostic or predictive factors: age (<65 years *vs* ≥65 years), ethnicity (white *vs* Asian *vs* other), sex (male *vs* female), number of previous chemotherapy regimens (two *vs* three or more), previous therapy (yes *vs* no for each of ramucirumab, irinotecan, and taxane), previous gastrectomy (yes *vs* no), GEJ involvement (yes *vs* no), peritoneal, liver, or lung metastases (yes *vs* no), the number of metastatic sites (one or two *vs* three or more), measurable disease (yes *vs* no), histology subtype (intestinal *vs* diffuse, mixed, or unknown), and HER2 status (negative *vs* positive or not assessed). All comparisons for secondary efficacy endpoints were made at the two-sided 0·05 significance level. Treatment comparisons for ORR and DCR were done with Fisher's exact test in the population with assessable tumour responses— that is, all patients in the ITT

population with measurable disease (at least one target lesion) at baseline who underwent at least one tumour assessment while on treatment (except for early disease progression or cancer-related death). Safety analyses were summarised with descriptive statistics in the safety analysis population, which included all patients who received at least one dose of study treatment. Sensitivity analyses were performed for OS from which patients who did not have documented refractory mGC were excluded, all major protocol violations were excluded or adjusted for, stratification was based on the case report form designation rather than the IXRS, sites with high accrual (>25 patients) were excluded, the date of all collected events (deaths) and survival status as of April 30, 2018, were used, or the as-treated (safety) population and treatment allocation were used. Sensitivity analyses were carried out for progression-free survival in which clinical progression was considered a progression-free survival event in addition to the presence of radiological evidence of progression; clinical progression was a progression-free survival event that also counted initiation of non-study anti-cancer therapy as an event date rather than as the date used to censor subsequent response assessment; all deaths and response assessments (without censoring missed visits) were included and radiological evidence of progression, clinical progression, initiation of non-study anticancer therapy, or death up to the survival cut-off date were counted as events; the time to first, second, and third radiological tumour assessments from the date of randomisation were used; and sites with high accrual (>25 patients) were excluded for all the previously mentioned analyses. All statistical analyses were done with SAS software (version 9.4).[4, 44]

B.2.4.4 Justification of sample size

The study was designed to detect with 90% power a HR for death of 0.70 (30% risk reduction) in the trifluridine/tipiracil group compared with the placebo group with an overall one-sided type 1 error of 0·025. A variable accrual period of 18 months and a 5% per year loss to survival follow-up rate was assumed. Using a treatment allocation of 2:1 (trifluridine/tipiracil: placebo) with 500 patients, 384 deaths were targeted for the final OS analysis. The primary analysis of OS included survival data obtained through the date of the 384th death observed in the study (27th March 2018).[44]

B.2.4.5 Interim analysis

An independent data-monitoring committee periodically assessed cumulative efficacy and safety data. After a planned interim analysis after 220 events, the committee decided that the study should continue until the targeted 384 events had occurred. The Lan-DeMets alpha spending approach was used with O'Brien-Fleming stopping boundaries to guide the efficacy assessment in the interim and final OS analyses. The associated OS boundaries were one-sided p values of 0·0031 for the interim analysis and 0·0215 for the final analysis. The OS futility boundary for the interim analysis was pre-fixed at an HR for overall survival of 0·95 or more.[4]

For a summary of the statistical analysis in the TAGS trial, please see below Table 10.[4, 44, 67]

Table 10: Summary of statistical analyses

Trial number (acronym)	NCT02500043 (TAGS trial)
Hypothesis objective	Trifluridine/tipiracil and BSC improves OS compared to placebo and BSC in patients with non-resectable, metastatic gastric adenocarcinoma (including adenocarcinoma of the GEJ) who had undergone one or two previous chemotherapy regimens
Statistical analysis	The cumulative efficacy and safety data was periodically assessed. After a planned interim analysis after 220 events, the committee decided that the study should continue until the targeted 384 events had occurred. OS and radiologically confirmed PFS were analysed in the ITT population with a one-sided stratified log-rank test, with the HR and two-sided 95% CIs based on a prespecified stratified Cox model and associated Kaplan-Meier survival estimates.
Sample size, power calculation	The study was designed to detect with 90% power a HR for death of 0.70 (30% risk reduction) in the trifluridine/tipiracil group compared with the placebo group with an overall one-sided type 1 error of 0·025. Using a treatment allocation of 2:1

	(trifluridine/tipiracil: placebo) with 500 patients, 384 deaths were targeted for the final OS analysis.
Data management, patient withdrawals	Electronic case report forms (eCRFS) were provided with a detailed completion guide. All eCRFs were to be fully completed to ensure accurate data interpretation.

B.2.5 Quality assessment of the relevant clinical effectiveness evidence

TAGS is a robust international, multi-centre phase III trial. An assessment of TAGS was conducted using the revised Cochrane Collaborations Risk of Bias Tool (RoB-2). Based on this analysis, the study was determined to be 'Low risk'.[70] The complete quality assessment is included in Appendix D1.3. A tabulated summary of the quality assessment results are presented in Table 11 below.[4, 44, 67]

Table 11: Quality assessment results for TAGS trial

Trial number (acronym)	**TAGS trial (NCT02500043)**
Was randomisation carried out appropriately?	Yes Patients were enrolled by study investigators. Eligible patients were randomised (2:1) to trifluridine/tipiracil plus BSC or placebo plus BSC via a dynamic allocation method (biased coin) with an interactive-voice web-response system (IXRS). Almac (Craigavon, UK) operated the IXRS and created the algorithm that generated the individual patient allocation when the study site accessed the system. The company had no other role in the trial. Once a patient's eligibility was confirmed and the criteria for randomisation were met, study-site personnel logged on to the IXRS to allocate patients to treatment. The IXRS randomly assigned study medication (trifluridine/tipiracil or placebo) by assigning a kit number to that patient. Randomisation was stratified by region (Japan vs rest of world), ECOG performance status (0 vs 1), and previous treatment with ramucirumab (yes vs no).
Was the concealment of treatment allocation adequate?	Yes Patients, investigators and study-site personnel, those assessing outcomes, and those analysing the data were masked to treatment assignment. Tablets of identical appearance were used to maintain masking. Only personnel from the contract research organisations involved in drug labelling and distribution (Fisher Clinical Services

	[Allentown, PA, USA] and Bell Medical Solutions [Tokyo, Japan]) and IXRS activities (Almac) were aware of treatment assignment.
Were the groups similar at the outset of the study in terms of prognostic factors?	Yes Baseline demographic and disease characteristics were generally balanced between the two treatment arms.
Were the care providers, participants and outcome assessors blind to treatment allocation?	Yes Please see above regarding concealment of treatment allocation.
Were there any unexpected imbalances in drop-outs between groups?	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes
Adapted from Systematic reviews: CRD's guidance for undertaking reviews in health care (University of York Centre for Reviews and Dissemination)	

Considerations for UK Clinical practice

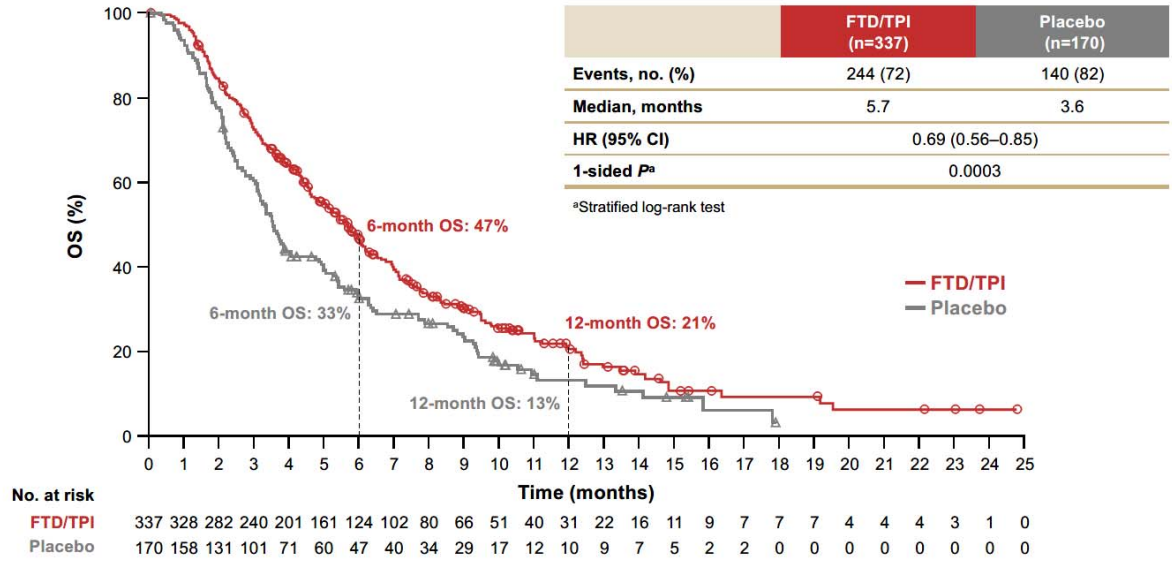
At the time of writing, ramucirumab was the most recently licensed treatment for patients who had previously received one line of treatment. It is not available to most of the UK population as it was not deemed cost-effective by NICE.[5] This population of "*no prior ramucirumab*" was pre-stratified for in the TAGS trial and had a clinically significant improvement in OS.[4] Servier has submitted this as an additional population within its economic model. This is discussed further in section B 2.1.3.

B.2.6 Clinical effectiveness results of the relevant trials

B.2.6.1 Primary endpoint: Overall survival

In the TAGS trial, the primary endpoint of OS was met, with the risk of death statistically significantly lower by 31% in the trifluridine/tipiracil group compared to the placebo group (HR: 0.69; 95% CI: 0.56–0.85, one-sided p=0.0003, two-sided p=0.0006).[4] The data demonstrates that nearly half of all patients in the trifluridine/tipiracil group were alive at 6 months (47% in the trifluridine/tipiracil group versus 33% in the placebo group) with over 20% alive at one year (21% in the trifluridine/tipiracil group versus 13% in the placebo group). See Figure 7 for the Kaplan Meier analysis of OS.[4]

Figure 7: Phase III TAGS trial: OS primary endpoint in all patients



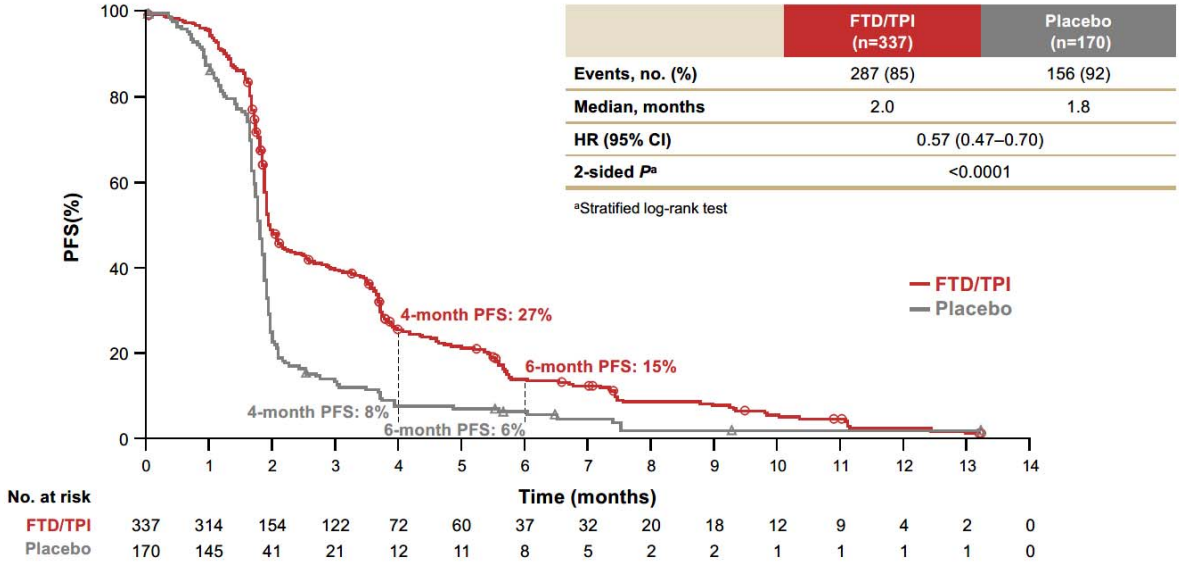
Key: CI: confidence interval; FTD/TPI: trifluridine/tipiracil; HR: hazard ratio; OS: overall survival.

The median OS was 5.7 months in the trifluridine/tipiracil group (95% CI: 4.8-6.2), compared to 3.6 months in the placebo group (95% CI: 3.1-4.1). This translates into a statistically significant and clinically meaningful benefit of 2.1 months in median OS.[4]

B.2.6.2 Secondary endpoint: Progression Free Survival

The risk of disease progression was significantly lowered by 43% in the trifluridine/tipiracil group compared to the placebo group (HR 0.57, 95% CI: 0.47–0.70, two-sided p<0.0001). See Figure 8 of Kaplan-Meier analysis of PFS.[4]

Figure 8: Phase III TAGS trial: PFS in all patients



	FTD/TPI (n=337)	Placebo (n=170)
Events, no. (%)	287 (85)	156 (92)
Median, months	2.0	1.8
HR (95% CI)	0.57 (0.47–0.70)	
2-sided *P*[a]	<0.0001	

[a]Stratified log-rank test

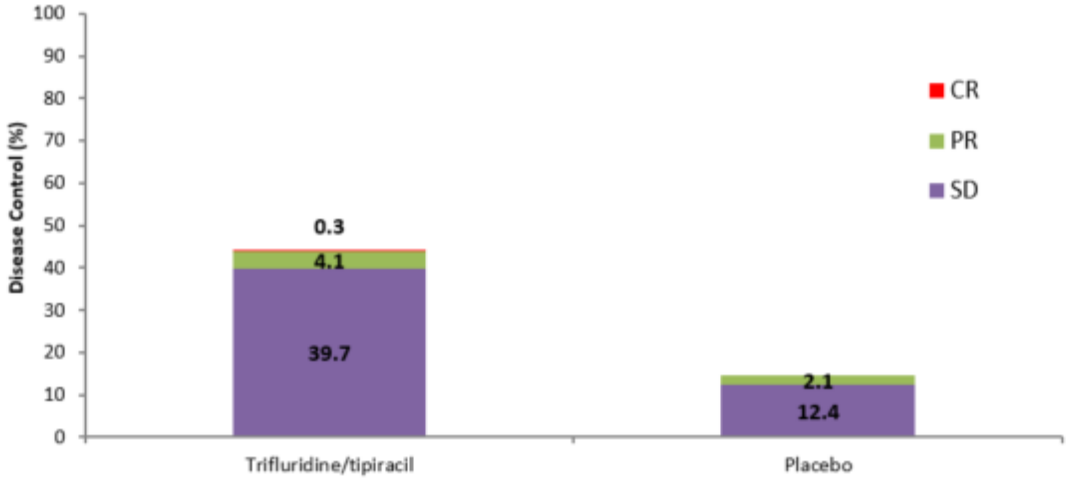
At 4 months, the PFS rate in the trifluridine/tipiracil group was over three times that of those in the placebo group (27% versus 8%). Similarly, at 6 months, the percentage of patient's progression free was 15% for the trifluridines/tipiracil group compared with 6% for the placebo group. Although median PFS was similar for the two treatment groups (2.0 versus 1.8 months), the percentage of patients with PFS was consistently higher for the trifluridine/tipiracil group than for the placebo group.[4]

B.2.6.3 Secondary endpoint: DCR/ORR

The assessment of DCR and ORR was restricted to patients with measurable disease (at least one target lesion) at baseline from the ITT population in the TAGS trial and with at least one post-baseline evaluation (tumour response population). The tumour response was evaluable in 290 (86.1%) patients out of 337 in the trifluridine/tipiracil group and 145 (85.3%) patients out of 170 in the placebo group.[4]

- DCR was 44.1% of patients in the trifluridine/tipiracil group, compared to 14.5% of patients in the placebo group (p<0.0001), therefore the trifluridine/tipiracil group demonstrated a significant three-fold increase in the proportion of patients with tumour shrinkage or SD when directly compared to placebo.[4]

Figure 9: Phase III TAGS trial: DCR in tumour response evaluable population



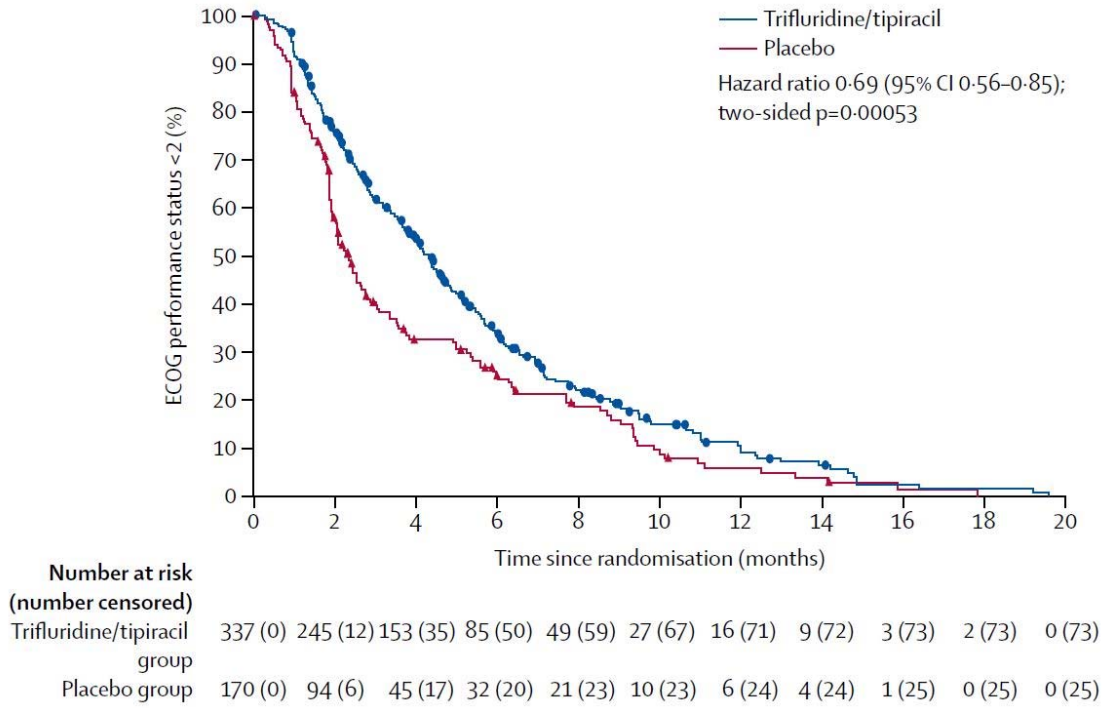
Key: CR, Complete response; PR, Partial response; SD, Stable disease.

Improvements in DCR were primarily due to differences in the proportion of patients with SD in the trifluridine/tipiracil group compared with the placebo group (39.7% versus 12.4%). The ORR was higher in the trifluridine/tipiracil group, with a rate of 4.5% (one patient achieved CR and 12 patients achieved PR) compared with 2.1% in the placebo group (3 patients achieved PR).[4]

B.2.6.4 ECOG PS deterioration

The median time to ECOG PS deterioration of two or more was significantly longer in the trifluridine/tipiracil group compared with the placebo group (HR 0.69, 95% CI: 0.56–0.85, two-sided p=0.0005), with a median of 4.3 months (95% CI: 3.7–4.7) and 2.3 months (95% CI: 2.0–2.8) in the trifluridine/tipiracil and placebo groups, respectively. Figure 10 depicts time to deterioration of the ECOG PS score to 2 or higher.[4]

Figure 10: Phase III TAGS trial: time to deterioration of ECOG PS score to 2 or higher



B.2.6.5 Health-related quality of life

HRQoL was balanced at baseline in both treatment groups with no differences greater than 10 points on either questionnaire. A mean change from baseline of ≥10 points was considered clinically relevant.[44]

- Overall, for the treatment period in the TAGS trial, HRQoL remained stable in both treatment groups with no clinically relevant changes from baseline, suggesting that HRQoL was maintained during treatment with trifluridine/tipiracil:[*44]
    - There was no clinically relevant difference in the mean change from baseline for the global health status in either group. For most other items in both scales, mean changes from baseline remained under the 10-point threshold, with the following exceptions:
        - Amelioration of hair loss (negative changes are associated with improvement):

[*] The overall compliance rate was 84% for both questionnaires

	Mean change from baseline >10					
	Cycle 1	Cycle 2	Cycle 3	Cycle 4	Cycle 5	Cycle 6
Trifluridine/tipiracil	-10.6	-11.0	-10.5	-10.4	-11.9	-12.0
Placebo	-14.1	-12.2	–	–	–	–

- Minor deteriorations > 10 in the placebo group at Cycle 1 or 2 for role functioning, fatigue, pain and appetite loss*

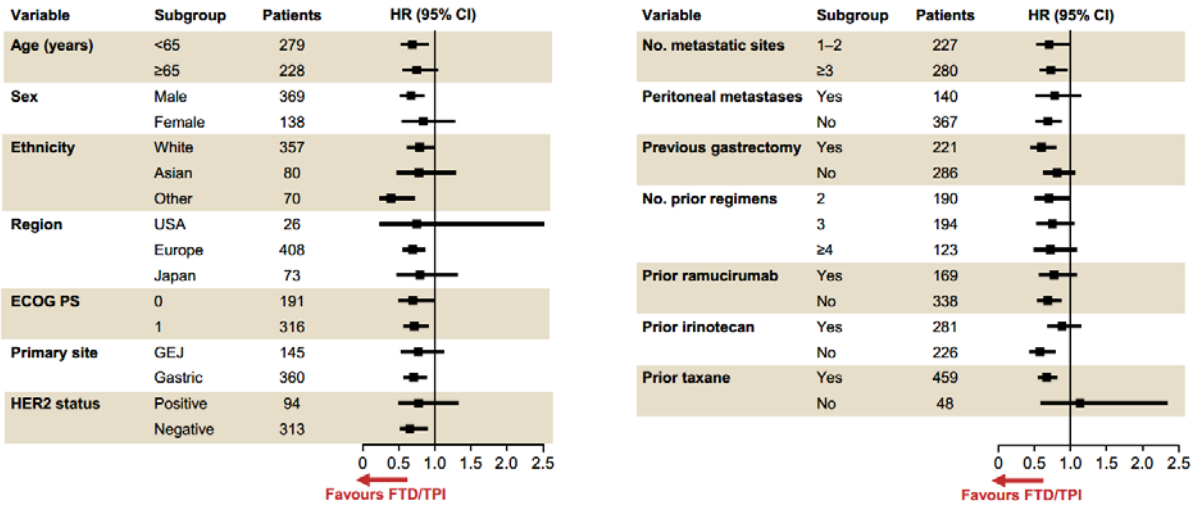
No clinically relevant difference between treatment groups in changes from baseline were observed, except for pain: 11.3 (Cycle 2) in favour of trifluridine/tipiracil and role functioning: 10.0 (Cycle 3) in favour of placebo.[44]

B.2.7 Subgroup analysis

Prespecified subgroup analyses of OS were conducted according to baseline demographics and disease characteristics and demonstrated that patient benefits from treatment with trifluridine/tipiracil were consistent across subgroups. Although ECOG PS, age, number of previous chemotherapy regimens (two versus three), number of metastatic sites, and HER2 status were prognostic of improved OS, the magnitude of the trifluridine/tipiracil treatment effect was maintained after adjustment for these factors (adjusted HR: 0.69; 95% CI: 0.56–0.85).[4, 71]

* Deteriorations at Cycle 1 or 2 for placebo included: Role functioning (negative changes are associated with deterioration): mean change from baseline for placebo: -10.4 (Cycle 1) and -12.1 (Cycle 2); Fatigue (positive changes are associated with deterioration): mean change from baseline for placebo: 11.1 (Cycle 2); Pain (positive changes are associated with deterioration): mean change from baseline for placebo: 12.9 (Cycle 2); Appetite loss (positive changes are associated with deterioration): mean change from baseline for placebo: 13.9 (Cycle 1).

Figure 11: Overall survival by sub-group analysis in TAGS trial



Key: CI: confidence interval; ECOG PS: Eastern Cooperative Oncology Group performance status; HER2: human epidermal growth factor receptor 2; HR: hazard ratio.

Sub-groups of note include:[2, 4, 44]

- **Patients with gastrectomy:** This sub-group included 147 patients in the trifluridine/tipiracil group and 74 patients in the placebo group. Results demonstrated that trifluridine/tipiracil prolonged survival compared to placebo regardless of prior gastrectomy in patients with mGC (median OS: 6.0 months and 3.4 months in the trifluridine/tipiracil group and placebo group respectively [HR: 0.57; 95% CI: 0.41–0.79]) and is therefore an effective treatment option in this patient population
- **Prior ramucirumab (yes vs no):** Patients who had not received prior ramucirumab had a median OS of 6.0 months in the trifluridine/tipiracil group, and 3.3 months in the placebo group respectively (HR: 0.66; 95% CI: 0.51–0.86). This is relevant to the UK population as ramucirumab is not reimbursed.
- **Geographic region (Japan versus ROW [EU/US]):** Patients in Japan had a median OS in the trifluridine/tipiracil (n=46) and placebo (n=27) groups of 6.3 months and 5.9 months, respectively (HR: 0.77; 95% CI: 0.46–1.30) compared to patients in ROW (EU/US) who had a median OS of 5.4 months in the trifluridine/tipiracil (n=291) and 3.3 months in the placebo group (n=143) (HR: 0.68; 95% CI: 0.54–0.85).

B.2.8 Meta-analysis

Only one phase III, randomised, double -blind controlled trial of trifluridine/tipiracil with a relevant comparator (BSC) has been conducted: TAGS. TAGS is a high-quality phase III trial, which included patients from 7 UK trial centres and which we believe is representative of UK clinical practice. A second trial was conducted (EPOC1201), which was a multi-centre phase II single-arm study of trifluridine/tipiracil. This trial was conducted only in Japanese centres and included 35 patients, 29 of whom received the same drug dosage (35 mg/m^2 BID PO D1-5,8-12 q4 weeks) as in the TAGS trial, the other 6 patients received a higher dose. Of the 29 patients, 5 had only received 1 line of chemotherapy prior to initiating trifluridine/tipiracil (an eligibility criterion for the TAGS trial was 2 or more prior lines of therapy).[4, 49]

Given the differences in trial design of the two studies and the differing characteristics of enrolled patients at baseline, it was deemed inappropriate to pool the patient data. Therefore, a meta-analysis has not been conducted for this submission. ,.

B.2.9 Indirect and mixed treatment comparisons

Not applicable for this appraisal.

B.2.10 Adverse reactions

Trifluridine/tipiracil demonstrated a predictable and manageable safety profile, consistent with prior data.[49] Safety was assessed in the TAGS trial among the AR population, which included the 503 patients who received treatment (trifluridine/tipiracil group n=335; placebo group n=168).[4, 44]

The overall incidence of adverse events was 97.3% for the trifluridine/tipiracil group and 93.5% for the placebo treatment group. A summary of the adverse events (including grade 1–2 events that were reported in 10% or more of patients and grade 3–5 events that were reported in 2% or more of patients) can be found in Table 12.[4, 44]

Table 12: Total adverse events and adverse events for which grade 1–2 events were reported in 10% or more, or grade 3–5 were reported in 2% of patients in either treatment group

	Trifluridine/tipiracil group(n=335)				Placebo group (n=168)			
	Grade 1–2	Grade 3	Grade 4	Grade 5	Grade 1–2	Grade 3	Grade 4	Grade 5
Any AE of any cause*	59 (18%)	172 (51%)	51 (15%)	44 (13%)	60 (36%)	64 (38%)	14 (8%)	19 (11%)
Any treatment-related AE	95 (28%)	136 (41%)	39 (12%)	1 (<1%)†	73 (43%)	21 (13%)	0	1 (1%)‡
Most common adverse events of any causes								
Nausea	114 (34%)	10 (3%)	0	0	48 (29%)	5 (3%)	0	0
Anaemia	86 (26%)	63 (19%)	1 (<1%)	0	19 (11%)	12 (7%)	1 (1%)	0
Decreased appetite	86 (26%)	28 (8%)	1 (<1%)	0	41 (24%)	9 (5%)	2 (1%)	0
Vomiting	71 (21%)	10 (3%)	2 (1%)	0	31 (18%)	3 (2%)	0	0
Diarrhoea	67 (20%)	8 (2%)	1 (<1%)	0	21 (13%)	3 (2%)	0	0
Fatigue	66 (20%)	23 (7%)	0	0	25 (15%)	10 (6%)	0	0
Neutropenia	62 (19%)	85 (25%)	29 (9%)	0	7 (4%)	0	0	0
Asthenia	49 (15%)	14 (4%)	2 (1%)	0	29 (17%)	11 (7%)	0	0
Thrombocytopenia	49 (15%)	7 (2%)	4 (1%)	0	8 (5%)	0	0	0

	Trifluridine/tipiracil group(n=335)				Placebo group (n=168)			
	Grade 1–2	Grade 3	Grade 4	Grade 5	Grade 1–2	Grade 3	Grade 4	Grade 5
Leukopenia	47 (14%)	28 (8%)	3 (1%)	0	3 (2%)	0	0	0
Abdominal Pain	41 (12%)	14 (4%)	0	0	16 (10%)	15 (9%)	0	0
Constipation	41 (12%)	3 (1%)	1 (<1%)	0	16 (10%)	4 (2%)	0	0
Back Pain	23 (7%)	2 (1%)	0	0	7 (4%)	4 (2%)	0	0
↑ blood [alkaline phosphatase]	21 (6%)	9 (3%)	0	0	9 (5%)	5 (3%)	0	0
Dyspnoea	18 (5%)	6 (2%)	0	0	11 (7%)	4 (2%)	2 (1%)	0
Dysphagia	13 (4%)	6 (2%)	1 (<1%)	0	4 (2%)	4 (2%)	0	0
Ascites	7 (2%)	12 (4%)	0	0	5 (3%)	10 (6%)	0	1 (1%)
General deterioration of physical health	1 (<1%)	4 (1%)	1 (<1%)	17 (5%)	2 (1%)	3 (2%)	1 (1%)	11 (7%)
Hyponatraemia	1 (<1%)	4 (1%)	0	0	1 (1%)	7 (4%)	0	0
↑ y-[glutamyltransferase]	1 (<1%)	2 (1%)	1 (<1%)	0	0	4 (2%)	1 (1%)	0

Note: Data are n (%) and are presented for all treated patients. Adverse events were defined according to the Common Terminology Criteria for Adverse Events; *Adverse event data were missing for accidental overdose (n=1 [<1%]) and drug misuse (n=1 [<1%]) in the trifluridine/tipiracil group and encephalopathy (n=1 [1%]) in the placebo group; †Attributed to cardiopulmonary arrest; ‡Attributed to toxic hepatitis.

A summary of serious adverse events reported in >1% of patients in either treatment group can be found in Table 13.[44]

Table 13: Summary of Serious Adverse Events by System Organ Class, Preferred Term and Treatment Group (Preferred Terms Reported for > 1.0% of Patients in Either Treatment Group, As-treated Population)

System Organ Class Preferred Term	Trifluridine/tipiracil (N=335) n (%)	Placebo (N=168) n (%)
Number of patients with at least 1 serious adverse event	143 (42.7)	70 (41.7)
Blood and lymphatic system disorders	**25 (7.5)**	**4 (2.4)**
Anaemia	13 (3.9)	4 (2.4)
Pancytopenia	7 (2.1)	0
Febrile neutropenia	4 (1.2)	0
Neutropenia	4 (1.2)	0
Gastrointestinal disorders	**55 (16.4)**	**31 (18.5)**
Vomiting	9 (2.7)	1 (0.6)
Abdominal pain	8 (2.4)	6 (3.6)
Diarrhoea	6 (1.8)	0
Dysphagia	6 (1.8)	2 (1.2)
Gastrointestinal haemorrhage	4 (1.2)	1 (0.6)
Intestinal obstruction	4 (1.2)	3 (1.8)
Ascites	3 (0.9)	7 (4.2)
Gastric haemorrhage	3 (0.9)	3 (1.8)
Small intestinal obstruction	3 (0.9)	2 (1.2)
Upper gastrointestinal haemorrhage	2 (0.6)	2 (1.2)
General disorders and administration site conditions	**28 (8.4)**	**21 (12.5)**
General physical health deterioration	21 (6.3)	15 (8.9)
Asthenia	1 (0.3)	3 (1.8)
Infections and infestations	**20 (6.0)**	**9 (5.4)**
Neutropenic sepsis	4 (1.2)	0
Pneumonia	4 (1.2)	2 (1.2)
Metabolism and nutrition disorders	**18 (5.4)**	**7 (4.2)**
Decreased appetite	11 (3.3)	4 (2.4)
Musculoskeletal and connective tissue disorders	**1 (0.3)**	**3(1.8)**
Back pain	0	3 (1.8)
Neoplasms benign, malignant and unspecified (including cysts and polyps)	**8 (2.4)**	**4 (2.4)**

System Organ Class Preferred Term	Trifluridine/tipiracil (N=335) n (%)	Placebo (N=168) n (%)
Malignant ascites	1 (0.3)	2 (1.2)
Respiratory, thoracic and mediastinal disorders	**15 (4.5)**	**4 (2.4)**
Pleural effusion	5 (1.5)	1 (0.6)
Pulmonary embolism	5 (1.5)	2 (1.2)
Dyspnoea	4 (1.2)	2 (1.2)
Note: At each level of summation (overall, system organ class, preferred term), patients were only counted once at the highest toxicity grade		

Action taken due to adverse events

Dosing adjustments for trifluridine/tipiracil based on the event of haematological and/or non-haematological toxicities are detailed in section 4.2 of the Summary of Product Characteristics.[4, 8] A summary of action taken due to any adverse event in the TAGS trial can be seen in Table 14 and most common adverse events leading to modification can be found in Table 15.[71]

Table 14: Action taken due to any adverse event

Action taken due to AEs (any grade)	Trifluridine/tipiracil (n=335) %	Placebo (n=168) %
Dosing modification (dosing delay or dose reduction)	58	22
Dose reduction	11	1
Treatment discontinuation	13	17
G-CSF treatment for neutropenia	16	2

Table 15: Most common adverse events leading to dosing modification

Most common AEs leading to dosing modification	Trifluridine/tipiracil (n=335) %		Placebo (n=168) %	
	Any Grade	≥ Grade 3	Any Grade	≥ Grade 3
Neutropenia and/or decreased neutrophil count	37	25	1	0
Anaemia and/or decreased haemoglobin level	9	4	2	2
Leukopenia and/or decreased white blood cell count	6	3	0	0

Deaths

As of the cut-off date for non-survival data (31 Mar 2018), a total of 104 patients died while on treatment or within 30 days of last dose; the majority of on-study deaths in both groups were due to disease progression. Forty-five (13.4%) patients in the trifluridine/tipiracil group and 19 (11.3%) patients in the placebo group experienced adverse events resulting in death. The most frequently reported adverse event resulting in death in both treatment groups was general physical health deterioration. One fatal adverse event in each treatment group was considered related to study treatment.[44] The patient in the trifluridine/tipiracil group experienced a Grade 5 cardio-respiratory arrest and died at home. The investigator stated the *"most probable cause of death would be due to gastric cancer; however, a possible relationship to study medication cannot be excluded."*[44]

In total, 17 deaths occurred on-study or within 30 days of last dose of study therapy for which an adverse event was identified by the investigator as the primary cause of death (15 patients in the trifluridine/tipiracil group [including 1 patient whose cause of death was "Other"] and 2 patients in the placebo group). For the 15 fatal serious adverse events for patients in the trifluridine/tipiracil group, 11 were due to disease progression and 4 were due to Grade 5 adverse events. Fourteen of the 15 fatal serious adverse events were assessed by the investigator as not related to trifluridine/tipiracil.[44]

B.2.11 Ongoing studies

No expected relevant Servier trial results will be available within 12 months following this appraisal.

B.2.12 Innovation

Following second-line treatment, there are no recommended treatment options for use in the third-line setting in UK NHS practice. For patients who are well enough and wishing to pursue further treatment, trifluridine/tipiracil provides patients with a reasonably well-tolerated and clinically-effective option. Due to its oral route of administration, trifluridine/tipiracil provides patients with an option to continue receiving treatment in the community setting.

B.2.13 Interpretation of clinical effectiveness and safety evidence

In total, there are an estimated 7,625 patients diagnosed with mGC and metastatic GEJ cancer in the UK each year. These patients have a short survival time which further decreases at each line of therapy, with a significant correlation found between the survival rate and the number of treatment lines received (P<0.001).[27] It is estimated that only approximately 250 patients in the UK will reach third-line and they will have a median overall survival of approximately 4 months.[43,4 72]

Guidelines for the management of mGC have been published by various international and national bodies, including ESMO, USA NCCN, French intergroup (comprising seven medical societies), and NICE.[1, 2, 6, 7, 73] Other available guidelines focus on Asian populations, however due to biological differences between gastric cancer in Asian versus non-Asian patients relevance of these guidelines to European patients is not relevant.[74, 75] There are no specific recommendations for the treatment of patients beyond second line in most guidelines, with the exception of NCCN guidelines which recommend the use of pembrolizumab for third line mGC only in patients tested positive for PD-L1, due to the US Food and Drug Administration (FDA) approval of pembrolizumab for mGC in the US.[1] However, this accelerated approval was based on a phase II multi-cohort study (KEYNOTE-059) on the condition that further studies were provided.[76] Subsequently, pembrolizumab failed a phase III study (KEYNOTE-061) in patients with second line and beyond advanced gastric cancer or GEJ adenocarcinoma expressing PD-L1, where the median OS did not significantly improve compared to paclitaxel.[4] Importantly, the use of pembrolizumab for PD-L1 positive patients with third line and beyond mGC is not authorised by the EMA and there are currently no active submissions. According to the ESMO and NCCN guidelines, treatment options used in second line mGC may be used for third line treatment and the choice of third line therapy is largely based on previous therapeutic strategy, the patient's tolerability to certain treatments and their PS.[1, 2] However, the guidelines acknowledge that there is no clear evidence for a benefit beyond second-line. See Table 16 for a summary of guideline recommendations.

Table 16: Summary of 3L+ treatment recommendations in international and national guidelines

Guidelines	3L+ treatment
ESMO clinical practice guidelines	Treatment options may be used sequentially in 2L and 3L, but there is no clear evidence for a benefit beyond 2L treatment
NCCN clinical practice guidelines	Therapy options include pembrolizumab for PD-L1-positive patients and regimens recommended for 2L that were not previously used
French intergroup clinical practice guidelines	*No recommendations made*
NICE oesophagogastric cancer assessment and management	*No recommendations made*

For patients with 3L+ mGC in Europe, there is a lack of evidence from a European population for the effectiveness of currently used cytotoxic treatments and there is an absence of treatment options which offer a survival advantage and/or improved QoL.[4, 77]

Furthermore, progress in therapy has been minimal in recent years and multiple phase III trials in mGC have not met their primary endpoints.[4, 33, 78, 79] Immuno-oncology therapies that have demonstrated OS benefits in many other tumour types have failed to demonstrate significant OS improvements in mGC, due to the aggressive and rapidly progressing nature of the disease.

The trials which have demonstrated clinical benefit in 3L+ mGC have been conducted exclusively in Asian populations, including studies investigating apatinib (China) and nivolumab (Japan, Korea and Taiwan).[28, 80]

Table 17. Phase III trials in pre-treated mGC that have demonstrated clinical benefit in Asian populations

Agent (trial name)	Trial location	Line of treatment	Trial arms	Patient population/ treatment setting
Apatinib (NCT01512745, NCT00970138)	China	3L+	• Apatinib • Placebo	Advanced or metastatic stomach or GEJ adenocarcinoma
Nivolumab (ATTRACTION-2)	Japan, Korea, Taiwan	3L+	• Nivolumab • Placebo	Advanced gastric cancer or GEJ cancer

Key: 3L+, third-line plus; BSC, best supportive care; GC, gastric cancer; GEJ, gastroesophageal junction

Due to biological differences between gastric cancer in Asian versus non-Asian patients efficacy of these treatments in European patients is uncertain, as recognised by ESMO and JSMO, who developed a Pan-Asian adapted ESMO clinical practice guideline for the management of patients with mGC. The EMA's Committee for Medicinal Products for Human Use (CHMP) highlighted that nivolumab led to an incremental survival benefit of only one month, and it was not clear whether this benefit demonstrated in an Asian patient population would be seen in European patients as these populations are known to be affected differently by gastric cancer.[66, 74] The CHMP stated that in the absence of further data, it is not possible to establish that the benefits of nivolumab outweigh its risks for European patients with mGC. Subsequently, Bristol-Myers Squibb withdrew its application for the use of nivolumab for the treatment of advanced 3L+ mGC in Europe.[66]

Therefore, there is a recognised need for additional therapies in 3L+ mGC with evidence from randomised controlled trials.

Trifluridine/tipiracil has demonstrated a significant improvement in OS and a manageable safety profile in this heavily pre-treated population.[4] The evidence for the efficacy and safety of trifluridine/tipiracil comes from one of the largest phase III international multi-centre trials in patients with previously treated mGC: TAGS. This robust, multinational, double-blind, parallel, randomised, phase III study investigated the efficacy and safety of oral trifluridine/tipiracil 35 mg/m^2 plus BSC versus placebo plus BSC in patients with mGC who had received **at least two** prior regimens for advanced disease. Eligible patients were centrally randomised (2:1) to receive trifluridine/tipiracil plus BSC (experimental arm) or placebo plus BSC (control arm) with

relevant pre-stratification criteria. Baseline demographic and disease characteristics were generally balanced between the two treatment arms.[4]

This is currently the only completed global phase III trial in third-line (or greater) mGC (with over 80% of patients from Europe[*]) to successfully assess survival benefit as the primary objective.[4]

The primary endpoint of OS was met with a clinically-meaningful reduction in the risk of death by 31% when directly compared to placebo, which translated to nearly half the patients being alive at 6 months (47% vs 33%) and over 20% alive at 1 year (21% vs 13%). The median OS was improved by 2.1 months (almost a 60% extension in survival) from 3.6 to 5.7 months. This effect of improved OS was robust and observed consistently for trifluridine/tipiracil across all randomisation strata (Region [Japan, ROW]; ECOG performance status [0 vs 1 at baseline]; and prior treatment with ramucirumab [yes vs no]) and 47 of 49 pre-specified subgroups.[4, 44]

Of note, treatment with trifluridine/tipiracil plus BSC demonstrated a reduction in the risk of death versus placebo plus BSC in both the prior treatment with ramucirumab (HR: 0.76) and no prior treatment with ramucirumab (HR: 0.66; Median OS 6.0 months) groups that was consistent with the effect seen in the ITT population (HR: 0.69). Median OS times for patients who had not received prior ramucirumab were 6.0 months for the trifluridine/tipiracil group and 3.3 months for the placebo group (HR = 0.66; 95% CI: 0.506, 0.855) – an extension of 2.7 months. This is relevant for the UK as ramucirumab is not reimbursed for NHS patients.[44]

The key secondary endpoint of PFS echoed the OS results with a reduction in the risk of disease progression or death by 43% when directly compared to placebo, which translated into a tripling of PFS rates at 4 months (27% vs 8%) and more than doubling at 6 months (15% vs 6%).[4]

Other important outcomes included:[4, 44]

[*] Comprised of the EU, Belarus, Israel, Russian Federation, and Turkey

- A statistically significant prolongation of time to deterioration of ECOG performance status ≥2 was observed for the trifluridine/tipiracil group compared to the placebo group.

- HRQoL as measured by EORTC QLQ-C30 and QLQ-STO22, which remained stable in both treatment groups with no clinically relevant change from baseline, indicating that QoL was maintained during treatment with trifluridine/tipiracil.

- A predictable, well-established and manageable safety profile, with low rates of AE-related treatment discontinuations comparable with placebo plus BSC (13% vs 17%). In addition, Grade 3 or higher non-haematological AEs were reported in less than 10% of patients in the trifluridine/tipiracil group.

The TAGS population is generalisable to the UK, with 270 patients (80%) from Europe, including patients from 7 UK trial centres. In addition, patients who had not received prior ramucirumab (which is not available to most UK patients) represented 82% (n=277) of the trifluridine/tipiracil treatment group. This was pre-stratified for at randomisation.[68] The patient baseline characteristics were assessed by an expert panel who agreed that they were consistent with a UK population.[21]

Thus, trifluridine/tipiracil, an oral treatment option, demonstrated a statistically significant improvement in OS by 2.1 months (an almost 60% extension in median survival) in the ITT population and 2.7 months in patients that did not receive ramucirumab (an 82% extension in median survival). This significant survival gain was achieved whilst maintaining HRQoL for patients receiving end-of-life treatment, for a disease where there has been very little positive news. Evidence for end-of-life criteria is provided in Table 18.[44, 68,81]

Table 18: End of life criteria

Criterion	Data available	Reference in submission (section and page number)
The treatment is indicated for patients with a short life expectancy, normally less than 24 months	Most patients with 3L mGC do not survive for 4 months (Shitara et al, mOS 3.6 months, Kang et al mOS 4.1 months)[43,4] survival rate at 1-year ~12%.[43,4]	

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	The median OS in the ITT population was improved by 2.1 months (almost 60% extension in median survival) from 3.6 months to 5.7 months. • With a clinically-meaningful reduction in the risk of death by 31% when directly compared to placebo, which translates into nearly half of all patients being alive at 6 months (47% vs 33%) and over 20% alive at 1 year (21% vs 13%). Median OS times for patients who had not received prior ramucirumab (pre-stratifed criteria) were 6.0 months for the trifluridine/tipiracil group and 3.3 months for the placebo group (HR = 0.66; 95% CI: 0.506, 0.855). This represented an improvement in OS of 2.7 months which represented an extension of survival of 82% Although the improvement in OS is not the '*normally*' considered 3 months to meet end-of life criteria, Servier believe it should be considered in relation to the very poor prognosis of this population of approx. 4 months (3.6 months in TAGS trial). In this population an extension of 2.1 months in survival represents an almost 60% extension in life expectancy. In the patients who had not received ramucirumab, the observed 2.7 months increase represented an 82% extension. A decision in favour of end-of life criteria would be in line with NICE's previous decisions to consider end-of-life not just in the context of absolutes but in relation to the relative gain. For example in appraisal TA476, the appraisal committee stated that it "*recognised that this survival gain should be considered in the context of the very poor prognosis for metastatic pancreatic cancer. The committee noted that the survival gain was below what is normally considered appropriate for the extension-to-life criterion to be met (that is, it was less than 3 months). However, it agreed that the survival gain was particularly important relative to the average survival of people with this condition, and*	

	therefore this criterion could be accepted as met in this circumstance." In this case, the estimated life expectancy was 6 months, and extension to life of 2.4 months, so that paclitaxel as albumin-bound nanoparticles (nab-paclitaxel) with gemcitabine in untreated metastatic pancreatic cancer led to a relative increase in life expectancy of 40%	

B.3 Cost effectiveness

B.3.1 Published cost-effectiveness studies

Details of the systematic literature review undertaken to identify published cost-effectiveness studies relevant to the technology appraisal are provided within Appendix G: Published cost-effectiveness studies.

Four relevant cost-effectiveness studies were identified, none of which included trifluridine/tipiracil. Each of the studies was conducted to establish the cost-effectiveness of apatinib (YN968D1, LSK BioPharma) or nivolumab (Opdivo®, Bristol-Myers Squibb) in a Chinese population. Each of the four analyses considered a three-state model based on progression status or death.

B.3.2 Economic analysis

No pre-existing cost-effectiveness analyses of trifluridine/tipiracil + BSC for the treatment of mGC were identified by the systematic literature review. Therefore, a *de novo* cost-effectiveness model was constructed to inform this submission.

B.3.2.1 Patient population

Trifluridine/tipiracil has been studied within the TAGS clinical trial for the treatment of adult patients with mGC including adenocarcinoma of the GEJ, treated with at least two prior systemic treatment regimens for advanced disease. Within the TAGS trial, prior treatment failure was based on exposure to at least two of the following agents:

- Any anticancer therapy within prior 3 weeks
- Extended field radiation within prior 4 weeks or limited field radiation within prior 2 weeks

- Any investigational drug/device received within prior 4 weeks
- Major surgery within prior 4 weeks (the surgical incision should be fully healed prior to study drug administration)

Approximately 34% of patients in the TAGS trial had previously received ramucirumab (Cyramza®, Eli Lilly and Company Limited) which is not recommended for the treatment of first- or second-line mGC by NICE (NICE TA378). Within the TAGS trial, patients were stratified according to prior treatment with ramucirumab ("yes" or "no") at randomisation (amongst other stratification factors, see Section B.2.3 for further details). Subgroup analyses for patients with and without prior treatment with ramucirumab were pre-specified within the TAGS study protocol.

The economic evaluation conducted to determine the cost-effectiveness of trifluridine/tipiracil in patients with mGC and GEJ cancer analysed two key populations:

- Only patients with no prior ramucirumab experience (the "no prior ramucirumab" population)
- All patients regardless of prior ramucirumab treatment (the "intention-to-treat" [ITT] population)

The base-case analysis presented in this submission is based on the population of patients with no prior ramucirumab experience, as this population is more reflective of the UK patient population who would be eligible to be treated with trifluridine/tipiracil should it be recommended. The corresponding cost-effectiveness results for the ITT population are presented within Section B.3.9.

B.3.2.2 Model structure

A partitioned-survival ("area under the curve") cost-effectiveness model was constructed in Microsoft Excel® consisting of three health states: "progression free", "progressed disease" and "dead".

The model structure was chosen due to the following:

- It is similar to those adopted to inform previous submissions in late-stage cancers, notably including the previous submission of trifluridine/tipiracil + BSC

in mCRC and the submission of ramucirumab for treating mGC or GEJ adenocarcinoma after chemotherapy[18, 82, 83]

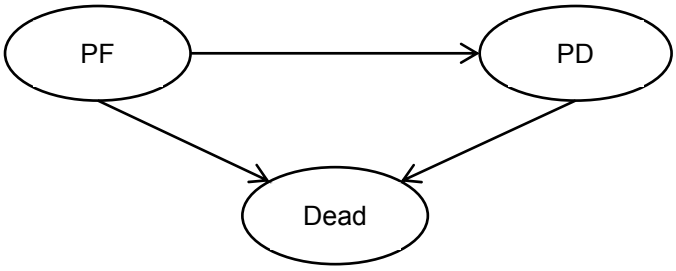
- The partitioned-survival structure allows for a clear application of the primary endpoint of the TAGS clinical trial (OS), as well as the secondary endpoint of PFS. In allowing health state occupancy to be determined in an intuitive manner, it is possible to consider a wide range of statistical extrapolation techniques that may be clearly interpreted, including the option to use the Kaplan-Meier curve directly

- The 3-state structure is aligned with the late-line positioning of trifluridine/tipiracil (that is, limited post-progression treatment is given in practice, and there would be relatively little to gain from further sub-dividing this health state)

- The partitioned-survival structure provides an advantage over the traditional Markov state-transition model as the hazard of death (or experiencing progression) within a given health state may be modelled as non-constant over time

The model schematic and associated permitted transitions are presented in Figure 12.

Figure 12: Model structure



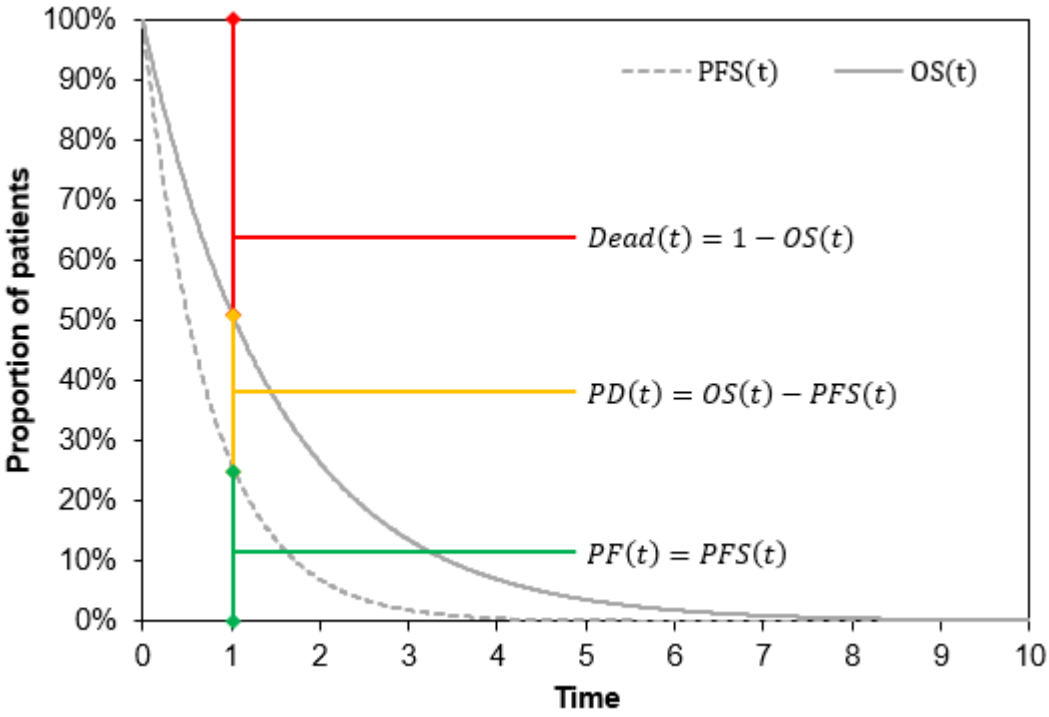
Key: PF, progression free; PD, progressed disease.

All patients enter the model in the "progression free" health state and remain in this state until disease progression or death. The probability of patients transitioning between the alive health states is inferred via extrapolated PFS and OS curves that are fitted to the clinical trial data; however, transition probabilities are not explicitly calculated for all possible transitions. Patients are unable to transition from the

"progressed disease" to the "progression free" health state as disease progression is irreversible. Patients are able to transition to the "dead" health state from any other health state, and once in the "dead" state cannot transition to any other health state (that is, "dead" is an absorbing state).

An illustration of the application of the extrapolated survival curves within the economic model is shown in Figure 13. The area underneath the PFS curve denotes the population of patients whom have not yet experienced disease progression, and the area between the OS and PFS curves denotes the population of patients with progressed disease. The remaining proportion of patients (above the OS curve, bounded by 100%) represents those patients who have died.

Figure 13: Illustration of survival model application



Key: OS, overall survival; PD, progressed disease; PF, progression free; PFS, progression-free survival.

For patients receiving trifluridine/tipiracil, the duration of treatment was determined via the use of treatment discontinuation data from the TAGS trial. Given that trifluridine/tipiracil is not administered beyond disease progression, the cost of treatment is restricted to patients in the "progression free" health state, though some patients in this health state may have discontinued treatment due to toxicity or

withdrawal of consent. Further information regarding the duration of treatment exposure may be found in Section B.3.5.

The model operates on a cycle length of 7 days to ensure sufficient accuracy of survival estimates without compromising model run time or file size. Due to the relatively short model cycle length, a half-cycle correction was not deemed necessary.[84] A time horizon of 10 years is adopted within the model, which was long enough to reflect the lifetime of mGC patients at the third-line treatment setting and beyond.

At the time of writing, NICE have published final recommendations for three technology appraisals in the treatment of mGC, however none of these appraisals were conducted in the "two or more prior therapies" (that is, third-line and beyond) population. The previous mGC technology appraisals are:

- TA191: Capecitabine for the treatment of advanced gastric cancer (recommended, first-line)
- TA208: Trastuzumab for the treatment of HER2-positive mGC (recommended, first-line)
- TA378: Ramucirumab for treating advanced gastric cancer or GEJ adenocarcinoma previously treated with chemotherapy (not recommended, second-line)

A summary of the key features of these appraisals are described alongside the corresponding features of this appraisal in Table 19.

Table 19: Key features of the economic analysis

Factor	Previous appraisals			Current appraisal	
	TA191	TA208	TA378	Chosen values	Justification
Time horizon	<1 year	Lifetime (8 years)	Lifetime (7.23 years)	Lifetime (10 years)	Time horizon long enough to reflect the lifetime of patients. The majority of patients are expected to have died before this time (for example, OS in the pivotal trial at the maximum follow-up time was ~6.5% at ~2.1 years)
Model structure	Cost minimisation	3-state AUC	3-state AUC	3-state AUC	Reflects progressive nature of condition appropriately, and has been accepted in previous NICE submissions
Treatment waning effect?	Not applicable	None described outside of preferred extrapolation methods	None described outside of preferred extrapolation methods	None described outside of preferred extrapolation methods	Assumptions pertaining to treatment waning effect are implicit within the selected extrapolation methods
Source of utilities	Not applicable	Trial utilities	Trial utilities with external AE decrements	Trial utilities with external AE decrements	Per NICE reference case (external sources only considered to address data gaps)
Source of costs	NHS reference costs, BNF and published literature	NICE TA179, NHS reference costs, BNF, PSSRU, expert opinion and published literature	NICE TA208, NHS reference costs, BNF, eMit, PSSRU and published literature	NICE TA378, NHS reference costs, BNF, eMit, PSSRU and published literature	These reflect resource utilisation and costs accepted in previous NICE submissions still considered appropriate following clinical expert consultation
Discount of 3.5% for utilities and costs	Annual rate of 3.5% on health effects. Costs were not discounted since the time horizon was <1 year	✓	✓	✓	NICE reference case
Perspective (NHS/PSS)	✓	✓	✓	✓	NICE reference case

Key: AUC, area-under-the-curve; BNF, British National Formulary; eMit, electronic Marketing information tool; NHS, National Health Service; NICE, National Institute for Health and Care Excellence; PSSRU, Personal Social Services Research Unit; TA, technology appraisal.

B 3.2.3 Intervention technology and comparators

The intervention considered within the cost-effectiveness analysis is trifluridine/tipiracil (Lonsurf®) + BSC. Trifluridine/tipiracil is administered *per os* (orally) at a dose of 35 mg/m² twice daily on days 1 to 5 and 8 to 12 of each 28-day treatment cycle.[8] This dose was administered within the TAGS trial, is aligned with the current marketing authorisation for the treatment of mCRC, and is representative of the marketing authorisation for mGC.

Within the TAGS trial, both the intervention (trifluridine/tipiracil) and comparator (placebo) drugs were administered in combination with BSC. The definition of BSC differs across clinical trial protocols, and is highly dependent on the context to which it applies. In principle however, BSC is provided to alleviate disease-related symptoms and maximise HRQoL for patients, without attempting to modify the disease course (that is, BSC is administered with palliative intent). As BSC does not constitute any interventional treatment beyond standard medical management, the costs associated with BSC are captured within the cost-effectiveness analysis via the use of health-state medical resource use costs.

Treatment with either trifluridine/tipiracil or placebo was continued until one of the following occurred: disease progression, death, unacceptable levels of toxicity or withdrawal of consent. Given that no cost was associated with placebo (for the purposes of the economic model), only treatment duration data for patients treated with trifluridine/tipiracil were included. Further information regarding the duration of treatment with trifluridine/tipiracil is provided within Section B.3.5.

There are no treatments recommended by NICE for patients with mGC who have failed at least two prior lines of treatment. As such, the comparator within the cost-effectiveness analysis is BSC (which is expected to be representative of current practice within the NHS). The comparator within the economic model is consistent with the comparator used in the pivotal phase III TAGS trial, and expected to reflect current UK NHS practice.

Off-label chemotherapy is rarely used for the “third-line and beyond” population in UK clinical practice. At an advisory board held by Servier in March 2019, the consensus from the attending 12 clinicians was that chemotherapy should not be considered a

comparator for the purpose of this appraisal. There are no clinical guidelines that recommend a specific treatment in this patient population.

Owing to the fact that in clinical practice patients would not receive placebo, henceforth within the dossier the term 'BSC' may be used in lieu of 'placebo + BSC'.

B.3.3 Clinical parameters and variables

Clinical data from the pivotal phase III TAGS trial were used to inform the economic model. The TAGS trial was a multicentre, double-blind, randomised, controlled study in which patients were assigned in a 2:1 ratio to receive trifluridine/tipiracil or placebo, along with BSC. The clinical data used to inform the economic model are described in turn within the sections below.

B.3.3.1 Patient characteristics

An overview of baseline patient characteristics within the TAGS trial is presented in Table 20. Body surface area (BSA) estimates were used within the model to inform the dosing of trifluridine/tipiracil - further details may be found in Section B.3.5.

Table 20: Baseline patient characteristics (ITT population)

Characteristic	T/T + BSC	PBO + BSC
Number	337	170
Mean age (years)	61.9	62.7
Female	31%	25%
BSA	1.75	1.75
ECOG PS 0	36%	40%
ECOG PS 1	64%	60%
2 prior lines of chemotherapy	37%	38%
3 prior lines of chemotherapy	40%	35%
4+ prior lines of chemotherapy	23%	27%
Key: BSA, body surface area; BSC, best supportive care; ECOG, Eastern Cooperative Oncology Group; ITT, intention-to-treat; PBO, placebo; PS, performance status; T/T, trifluridine/tipiracil.		

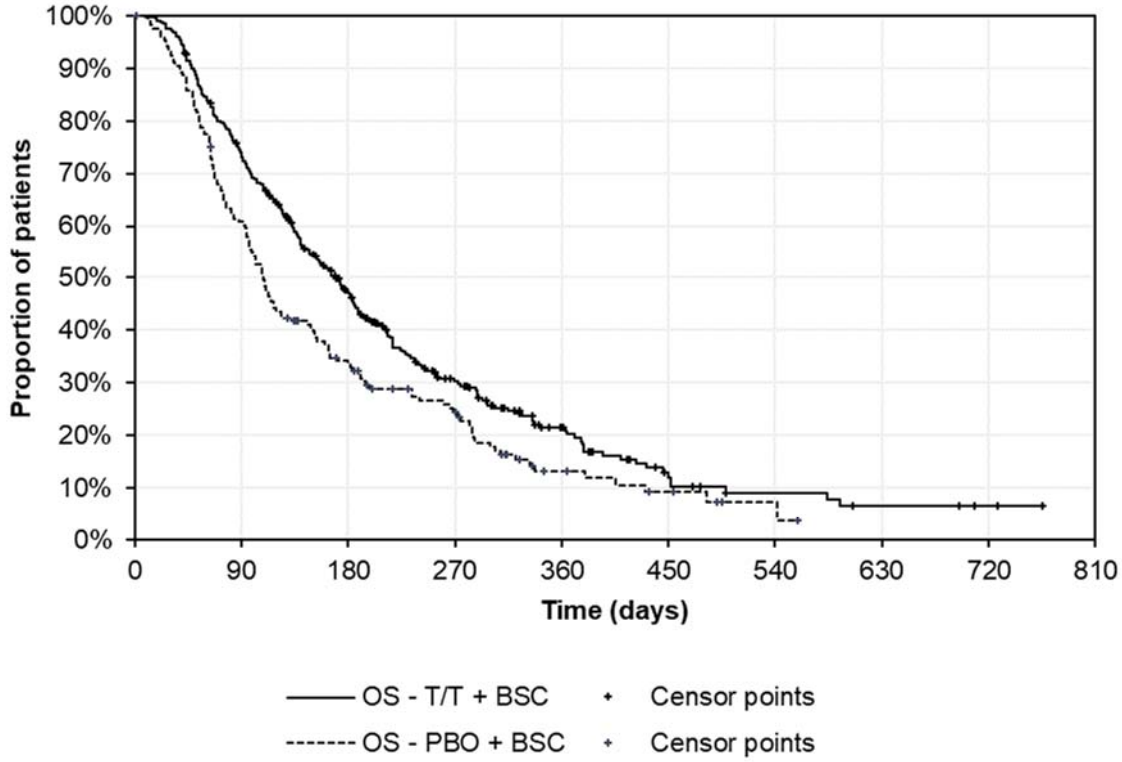
B.3.3.2 Efficacy

Efficacy data from the TAGS trial were used to inform transitions between the health states within the economic model. Parametric survival models (PSMs) were fitted to OS and PFS data.

Overall survival

The primary endpoint of the TAGS trial was OS (defined as the time from randomisation until death).[4] Trifluridine/tipiracil was shown within TAGS to significantly improve OS versus placebo (HR = 0.69, 95% CI: 0.56-0.85), with a 58.3% increase in median OS (+2.1 months, 5.7 versus 3.6 months for trifluridine/tipiracil versus placebo). A summary of the available OS data from the TAGS trial is provided in Figure 14.

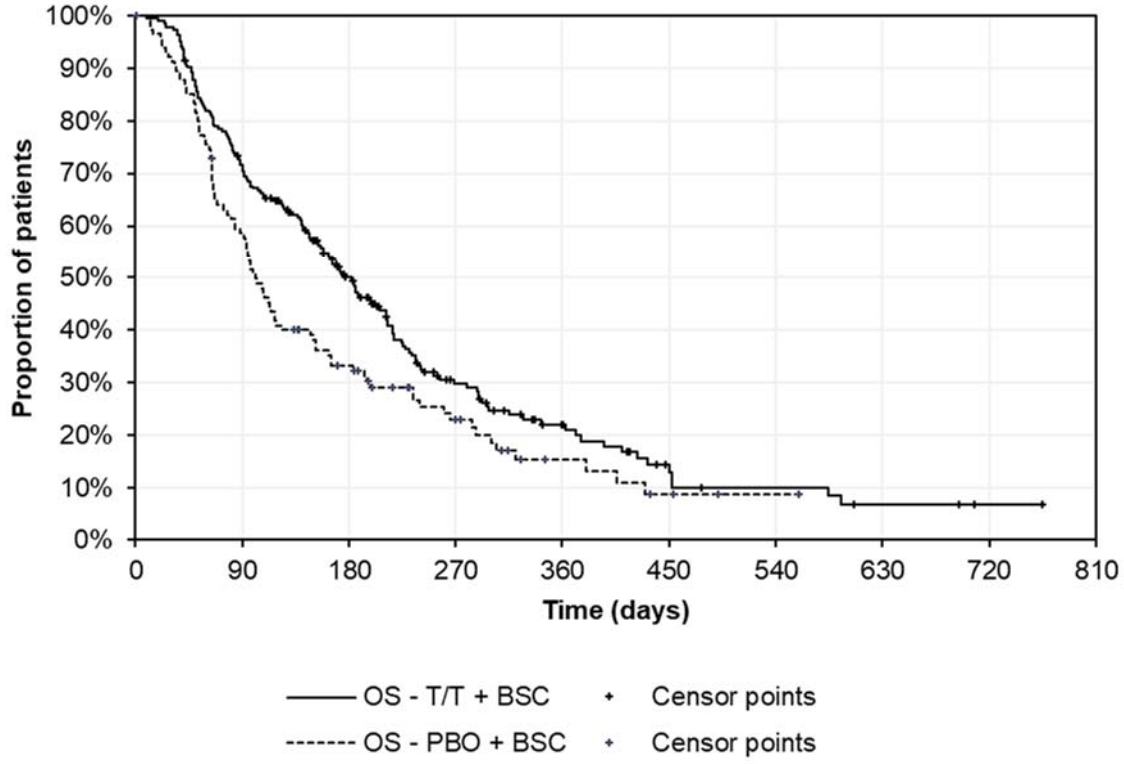
Figure 14: TAGS: Overall survival (ITT)



Key: BSC, best supportive care; OS, overall survival; T/T, trifluridine/tipiracil.

For the population of patients whom have not previously received treatment with ramucirumab, the corresponding OS Kaplan-Meier curve is presented in Figure 15.

Figure 15: TAGS: Overall survival (no prior ramucirumab)



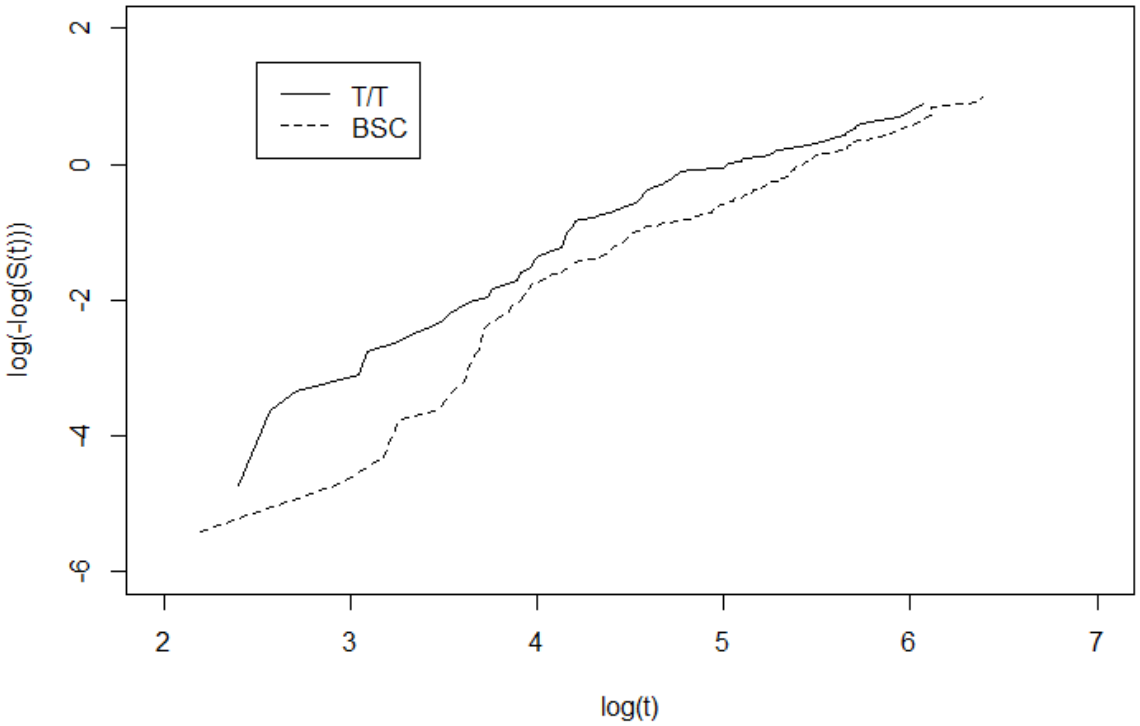
Key: BSC, best supportive care; OS, overall survival; T/T, trifluridine/tipiracil.

A range of PSMs were fitted to the OS data from the TAGS trial. These were the exponential, Weibull, Gompertz, log-normal, log-logistic and generalised gamma PSMs. Candidate PSMs were selected based on guidance from the NICE Decision Support Unit (DSU) Technical Support Document (TSD) 14.[85] To assess the appropriateness of each of these PSMs, a series of hazard-based plots were produced, and are described in turn below.

A log-cumulative hazard plot (LCHP) was produced to assess the appropriateness of fitting PSMs that assume proportional hazards (PH), as well as the use of the exponential and Weibull PSMs specifically. The LCHP is presented in Figure 16.

The gradient of each of the curves in the LCHP appears to decrease over time, indicating non-linearity and therefore PSMs that assume PH may be inappropriate for consideration. The non-constant gradient of the curves indicates that both the Weibull and exponential PSMs are unlikely to provide a good fit to the data. However, for completeness, these PSMs were not discounted from consideration, as the interpretation of the LCHP is subjective.

Figure 16: Log-cumulative hazard plot – Overall survival (TAGS, no prior ramucirumab)

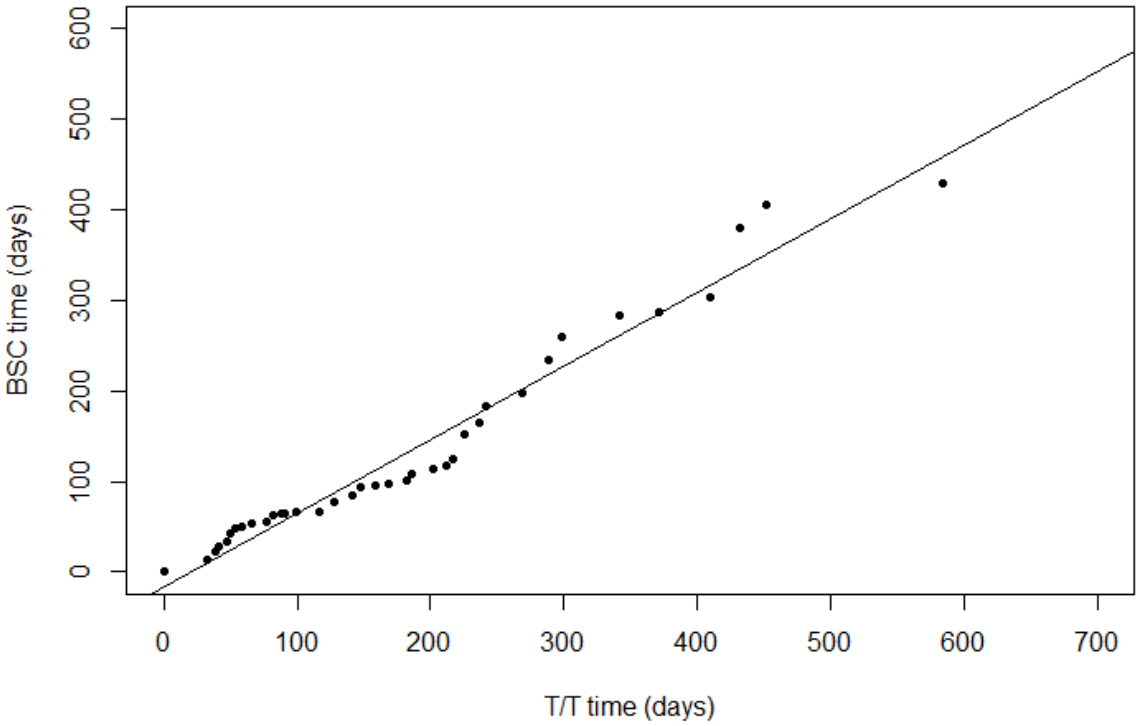


Key: BSC, best supportive care; S(t), survivor function; t, time; T/T, trifluridine/tipiracil.

Note: Approximately straight lines indicate that the survivor function is Weibull. If the gradient is approximately equal to 1, the survivor function is exponential.

To assess the appropriateness of accelerated failure time (AFT) models, a quantile-quantile plot was produced (Figure 17). The quantile-quantile plot demonstrates that the percentiles of the corresponding OS times for each treatment arm follow an approximately linear pattern when plotted against each other. This indicates that the treatment effect is approximately constant over time, and therefore AFT models fitted with a covariate for treatment assignment are appropriate to consider.

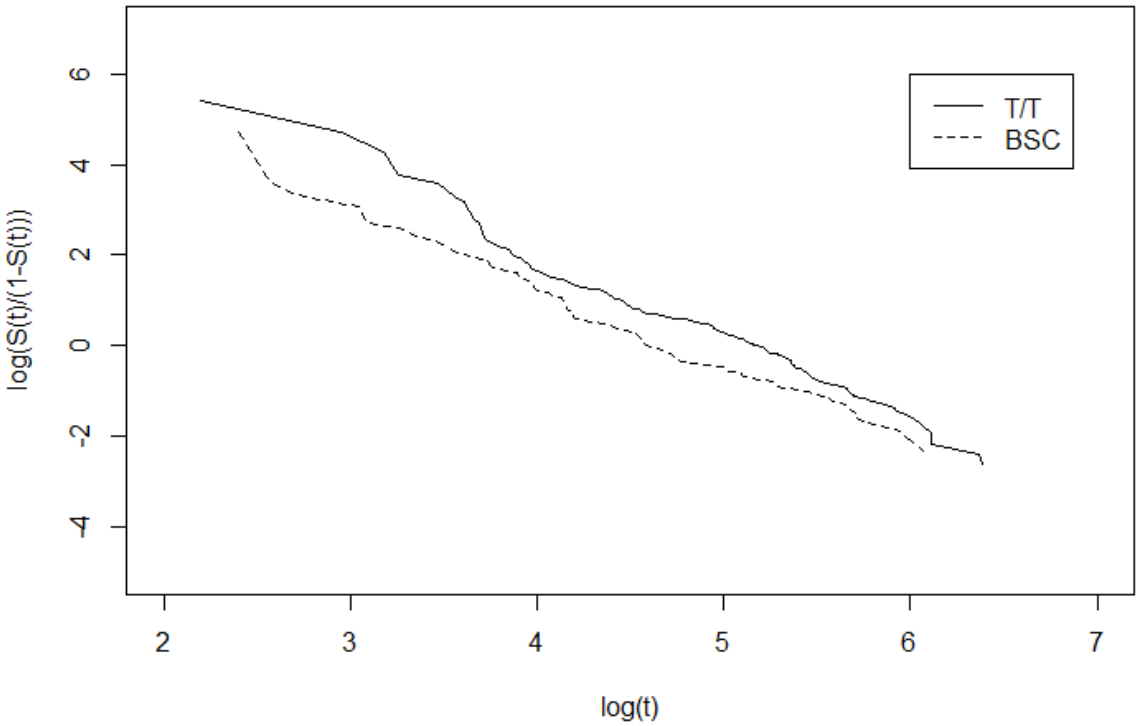
Figure 17: Quantile-quantile plot – Overall survival (TAGS, no prior ramucirumab)



Key: BSC, best supportive care; T/T, trifluridine/tipiracil.
Note: Straight line indicates non-violation of accelerated failure time (AFT) assumption

To assess the suitability of a log-logistic PSM, the logit function of survival (that is, the log-odds of the survivor function) may be plotted against the log of time – if the result yields approximately straight lines, log-logistic PSMs may provide a good fit to the data. As shown in Figure 18, the curves are approximately linear, and therefore log-logistic PSMs may provide a good fit to the data.

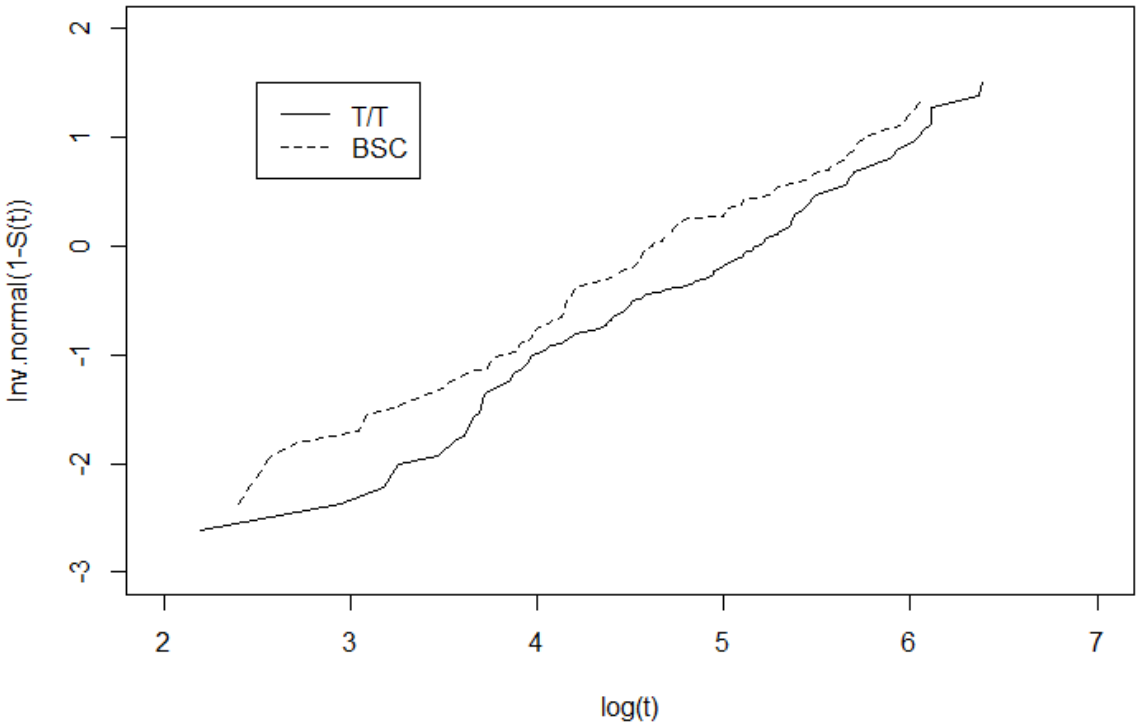
Figure 18: Logit survival plot – Overall survival (TAGS, no prior ramucirumab)



Key: BSC, best supportive care; S(t), survivor function; t, time; T/T, trifluridine/tipiracil.
Note: Approximately straight lines indicate that the survivor function is log-logistic.

To assess the suitability of a log-normal PSM, the inverse Normal cumulative distribution function applied to the probability of death over time may be plotted against the log of time. Like the logit survival plot, if the result yields approximately straight lines, log-normal PSMs may provide a good fit to the data. As shown in Figure 18, the curves are approximately linear, and therefore log-normal PSMs may provide a good fit to the data.

Figure 19: Inverse normal survival plot – Overall survival (TAGS, no prior ramucirumab)

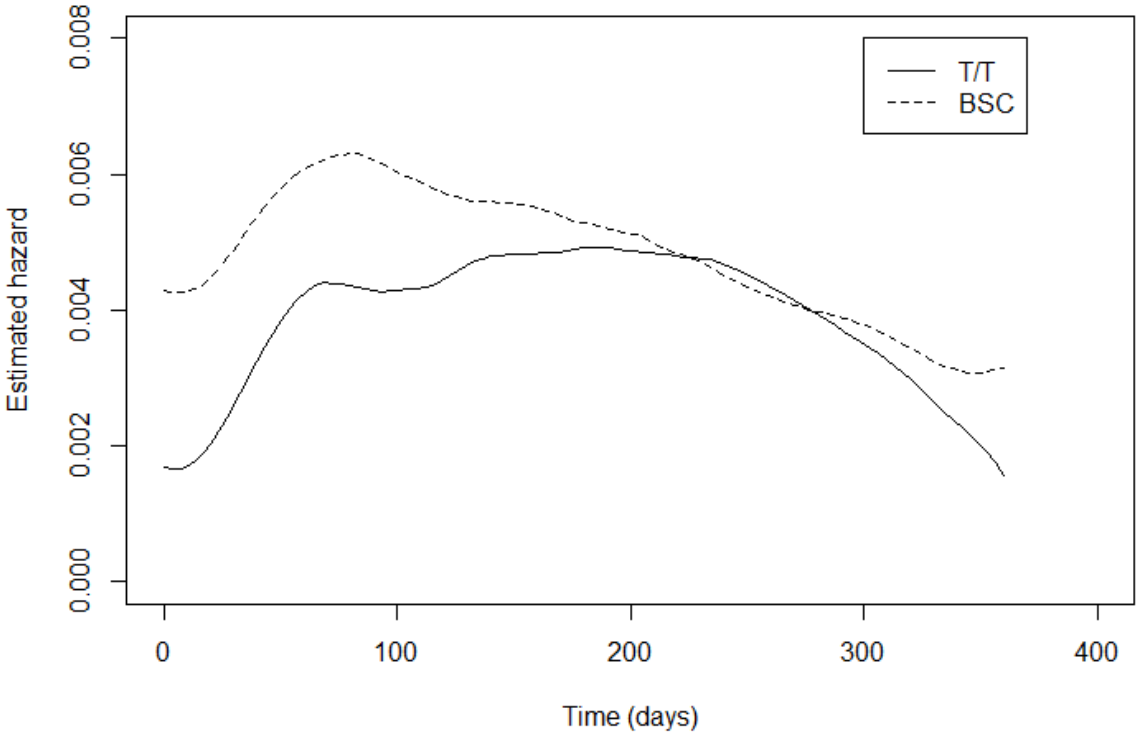


Key: BSC, best supportive care; Inv.norm, Inverse normal S(t), survivor function; t, time; T/T, trifluridine/tipiracil.
Note: Approximately straight lines indicate that the survivor function is log-normal.

The final assessment of the survivor data undertaken comprises a smoothed hazard plot. These plots were produced using the R package *'muhaz'*, which provides an estimated, smoothed hazard function for each treatment arm which may be used to infer which PSMs are likely to yield better fits than others. A maximum time point of 360 days was set when producing the smoothed hazard plots, as hazard estimates are subject to substantial uncertainty when the number of patients at risk is small.

The plots for both treatment arms (shown in Figure 20) demonstrate that the hazard of death does not appear to be constant over time, nor does it appear to be monotonic (that is, either consistently increasing or decreasing). This provides further evidence to suggest that the exponential and Weibull models may not provide a good fit to the data, and in addition the Gompertz (which assumes monotonic hazards) may also provide a poor fit to the data.

Figure 20: Smoothed hazard plots – Overall survival (TAGS, no prior ramucirumab)



Note: Turning points indicate the need for parametric survival models that are able to reflect non-monotonic hazard functions. A maximum time point of 360 days was selected to calculate the smoothed hazard estimation within the muhaz package.

Based on the diagnostic plots, it was determined that while the evidence suggested the log-normal and log-logistic models may be more likely to provide better fits to the observed data versus the exponential, Weibull, and Gompertz models, no specific parameterisations were ruled out. Furthermore, there is no specific diagnostic plot that was produced to determine the suitability of the generalised gamma curve (as there is no simplistic representation of this distribution that may be compared using such means).

Consequently, a total of 12 distinct OS extrapolations were available for use in each treatment arm within the economic model – 6 fitted with a covariate for treatment assignment, and 6 fitted independently by treatment arm.

To determine the most appropriate PSMs for use in the base-case analysis, guidance from NICE technical support document 14 was followed.[85] Following an inspection of the Kaplan-Meier curve for OS, and the assessment of the underlying hazard function, the following features of the fitted models were considered:

- Visual assessment: *does the parametric model provide a reasonable fit versus the Kaplan-Meier curve (within the time period over which data are available)?*

- Statistical goodness-of-fit: *does the parametric model yield an improved fit to the data relative to another model when considering its complexity (again, within the time period over which data are available)?*

- Long-term plausibility: *does the extrapolated portion of the model yield clinically realistic estimates of survival (beyond the time period over which data are available)?*

The statistical goodness-of-fit of all fitted PSMs is provided in Table 21. Based on the AIC and BIC scores, the top 6 models were visually compared in order to select the base-case extrapolation (shown in Figure 21 and Figure 22 for T/T and BSC, respectively). Visual fit within the observed period may be assessed on a per-model basis via the figures presented in Appendix O: Survival analysis.

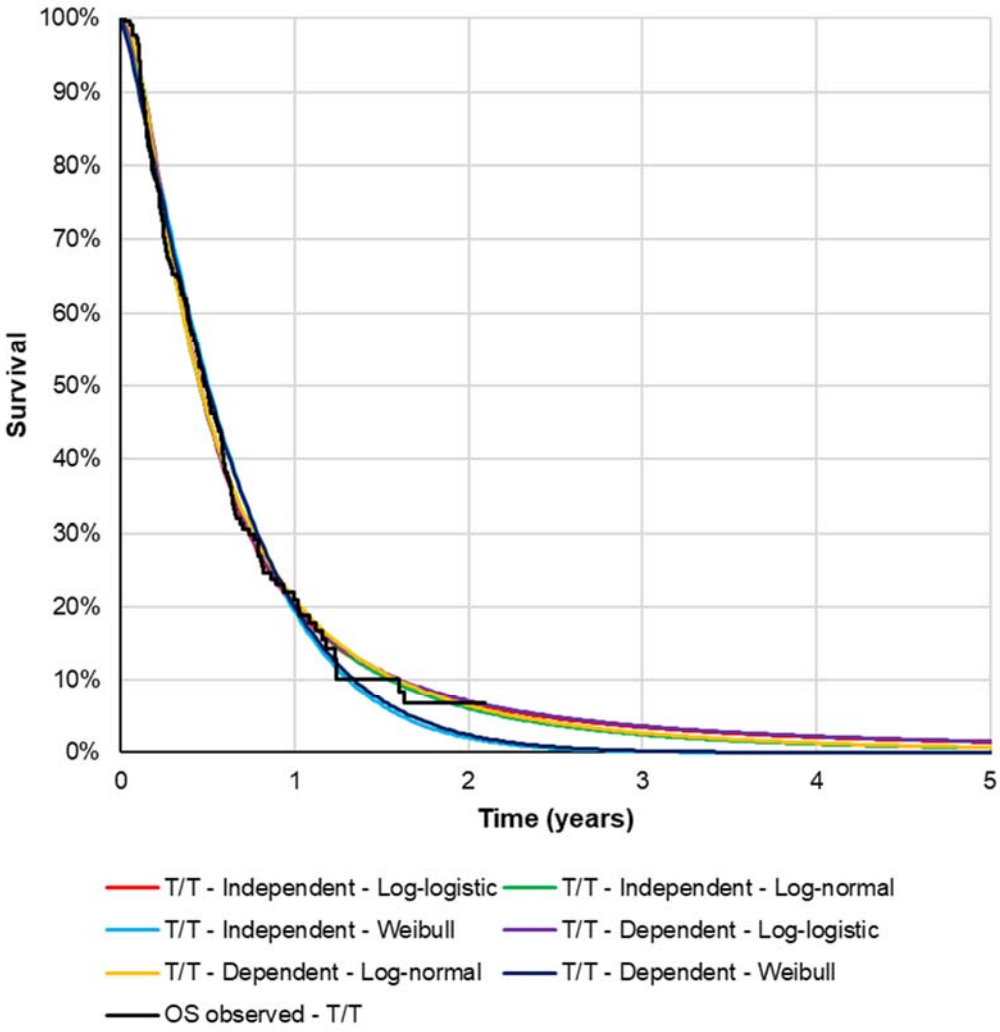
Table 21: Statistical goodness-of-fit scores (OS, TAGS, no prior ramucirumab)

Dependence	Parameterisation	Statistical goodness-of-fit					
		T/T + BSC		PBO + BSC		Combined	
		AIC	BIC	AIC	BIC	AIC	BIC
Independent	Exponential	2,124.11	2,127.52	1,138.16	1,140.91	3,262.27	3,268.42
	Generalised gamma	2,113.97	2,120.79	1,138.18	1,143.67	3,252.15	3,264.46
	Gompertz	2,124.47	2,131.28	1,139.93	1,145.42	3,264.40	3,276.70
	Log-logistic	2,102.91	2,109.72	1,125.44	1,130.93	3,228.35	3,240.66
	Log-normal	2,097.67	2,104.48	1,123.71	1,129.20	3,221.38	3,233.68
	Weibull	2,099.61	2,109.83	1,124.73	1,132.97	3,224.34	3,242.80
Dependent	Exponential					3,262.27	3,269.92
	Generalised gamma					3,251.27	3,262.74
	Gompertz					3,263.55	3,275.01
	Log-logistic					3,226.51	3,237.98
	Log-normal					3,219.65	3,231.11
	Weibull					3,220.91	3,236.20

Key: AIC, Akaike Information Criterion; BIC, Bayesian Information Criterion; BSC, best supportive care; OS, overall survival; PBO, placebo; T/T, trifluridine/tipiracil.

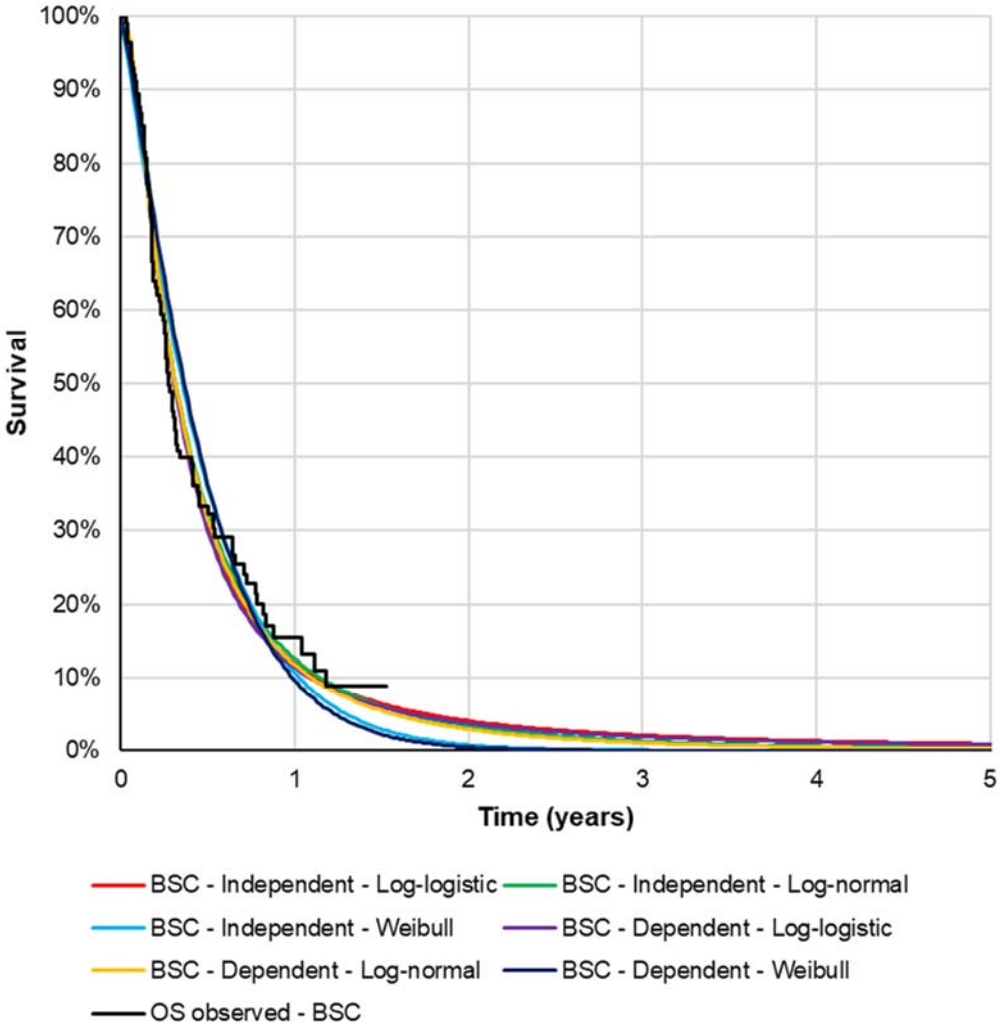
Note: AIC and BIC scores for the independent models were combined by simple addition – it is possible within the economic model to consider separate parameterisations for the independent models by treatment arm (for example, independent exponential for T/T + BSC and independent generalised gamma for PBO + BSC).

Figure 21: Top six curve fits: OS – T/T (TAGS, no prior ramucirumab)



Key: OS, overall survival; T/T, trifluridine/tipiracil.

Figure 22: Top six curve fits: OS – BSC (TAGS, no prior ramucirumab)



Key: BSC, best supportive care; OS, overall survival.

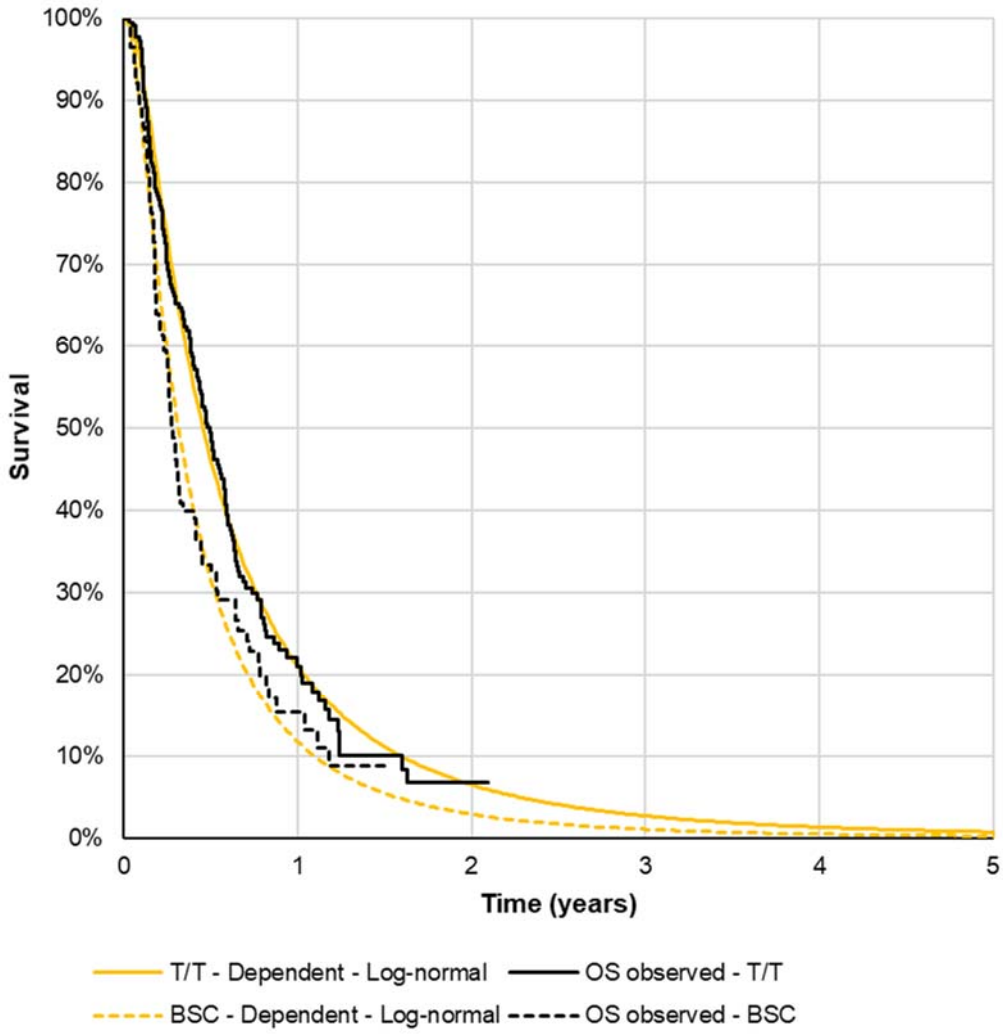
Based on the assessment of visual fit, statistical goodness-of-fit and long-term plausibility, the dependent log-normal model was chosen to inform the estimation of OS for both trifluridine/tipiracil + BSC and placebo + BSC groups. The base-case curve was selected due to the following:

- The dependent log-normal model provided a good visual fit to the Kaplan-Meier curves for OS across both treatment arms
- The dependent log-normal model had the lowest AIC and BIC of all models tested
- The long-term extrapolation of OS was aligned with clinical expectation:

- 5-year OS = 0.71% (T/T + BSC) and 0.23% (PBO + BSC)
- 10-year OS = 0.08% (T/T + BSC) and 0.02% (PBO + BSC)

Alternative survival extrapolations were explored within scenario analysis. Details of these scenarios may be found in Section B.3.8. The base-case curve fits are provided in Figure 23.

Figure 23: Base-case OS extrapolation (TAGS, no prior ramucirumab)



Key: BSC, best supportive care; OS, overall survival; PBO, placebo; T/T, trifluridine/tipiracil.

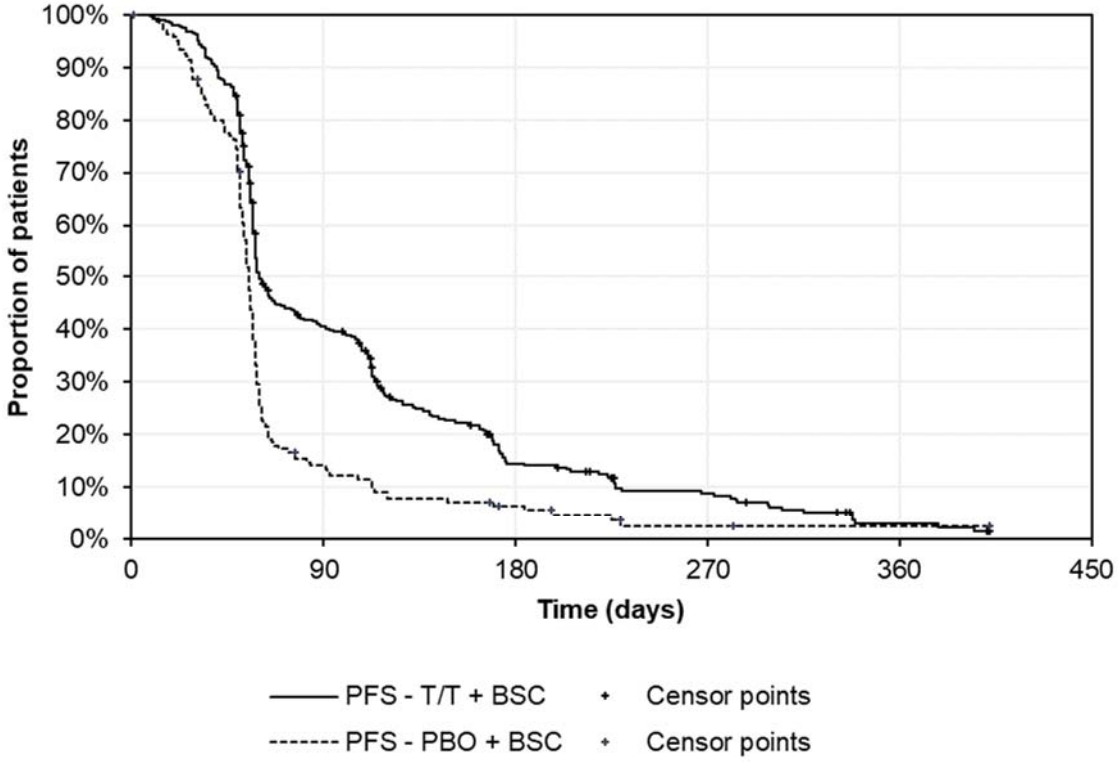
All extrapolations of OS are capped within the model according to the age- and sex-adjusted background mortality rates, taken from the Office for National Statistics Life Tables (released September 2018).[86] Should the predicted hazard of death for either treatment arm be less than the general population, the model will apply the hazard of

death for the general population instead. The impact of disabling adjustment of survival extrapolation in line with background mortality is explored within sensitivity analysis.

Progression-free survival

Progression-free survival (PFS) curves were used to inform the proportion of patients residing in the "progression-free" and "progressed disease" health states within the economic model. Data from the TAGS trial demonstrated that trifluridine/tipiracil significantly improved PFS versus placebo (HR = 0.57, 95% CI: 0.47-0.70). A summary of the available PFS data from the TAGS trial are provided in Figure 24.

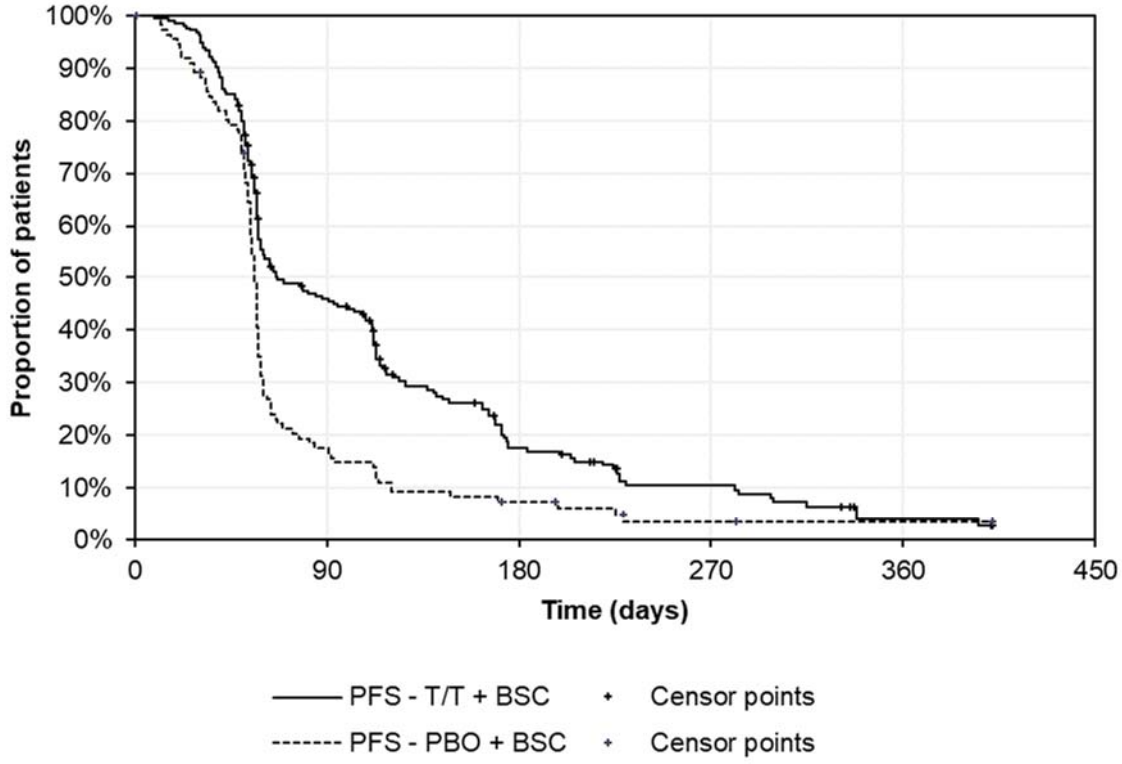
Figure 24: TAGS: Progression-free survival (ITT)



Key: BSC, best supportive care; PFS, progression-free survival; T/T, trifluridine/tipiracil.

For the population of patients whom have not previously received treatment with ramucirumab, the corresponding PFS Kaplan-Meier curve is presented in Figure 25.

Figure 25: TAGS: Progression-free survival (no prior ramucirumab)



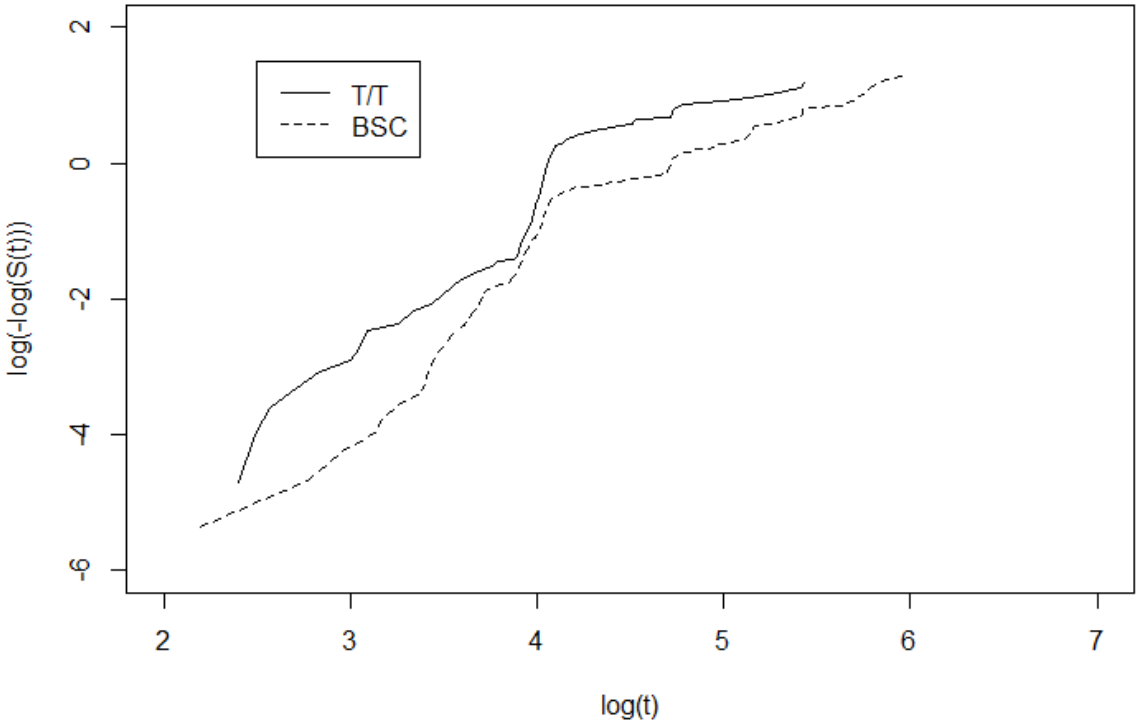
Key: BSC, best supportive care; PFS, progression-free survival; T/T, trifluridine/tipiracil.

The previous range of PSMs fitted to the OS data were also fitted to the PFS data from the TAGS trial. Per the assessment for OS, a series of hazard-based plots were produced to determine the most appropriate PSMs to fit to the PFS data, described in turn below.

A LCHP (to assess the appropriateness of fitting PSMs that assume PH, as well as the use of the exponential and Weibull PSMs specifically) is presented in Figure 26.

The curves in the LCHP are non-linear (as shown via the turning points at approximate log(t) = 4) and therefore PSMs that assume PH may be inappropriate for consideration. However, for completeness, these PSMs were not discounted from consideration, as the interpretation of the LCHP is subjective.

Figure 26: Log-cumulative hazard plot – Progression-free survival (TAGS, no prior ramucirumab)

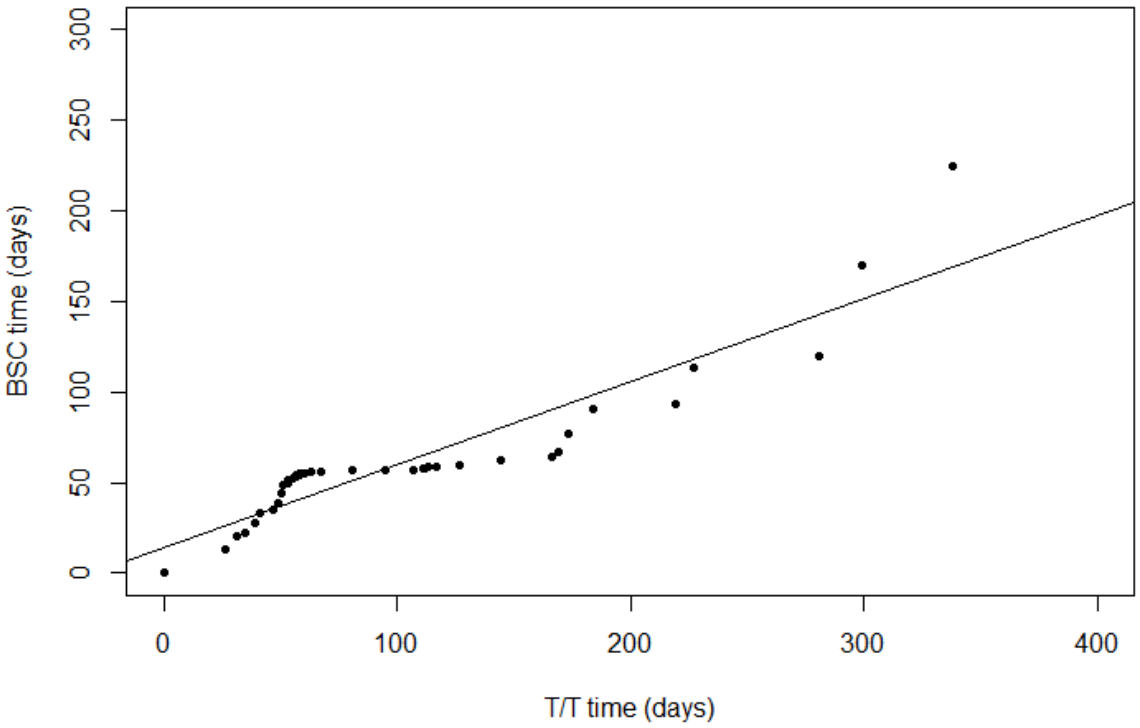


Key: BSC, best supportive care; S(t), survivor function; t, time; T/T, trifluridine/tipiracil.

Note: Approximately straight lines indicate that the survivor function is Weibull. If the gradient is approximately equal to 1, the survivor function is exponential.

A quantile-quantile plot was produced to assess the plausibility of AFT models, shown in Figure 27. The quantile-quantile plot demonstrates a non-linear pattern, which indicates a non-constant treatment effect over time and so AFT models fitted with a covariate for treatment assignment may be inappropriate.

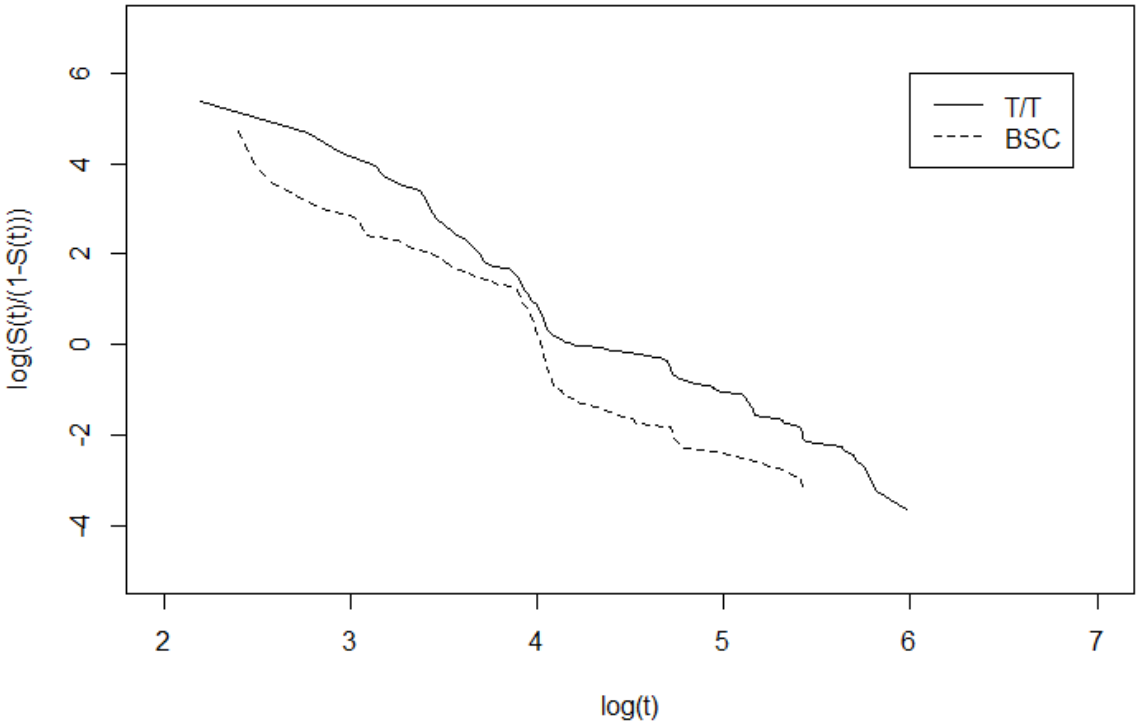
Figure 27: Quantile-quantile plot – Progression-free survival (TAGS, no prior ramucirumab)



Key: BSC, best supportive care; T/T, trifluridine/tipiracil.
Note: Straight line indicates non-violation of accelerated failure time (AFT) assumption

The logit function of survival versus the log of time is shown in Figure 28. The curve for T/T is approximately linear though the same may not be said for the curve for BSC. As such, the log-logistic PSMs may not provide a good fit to the data.

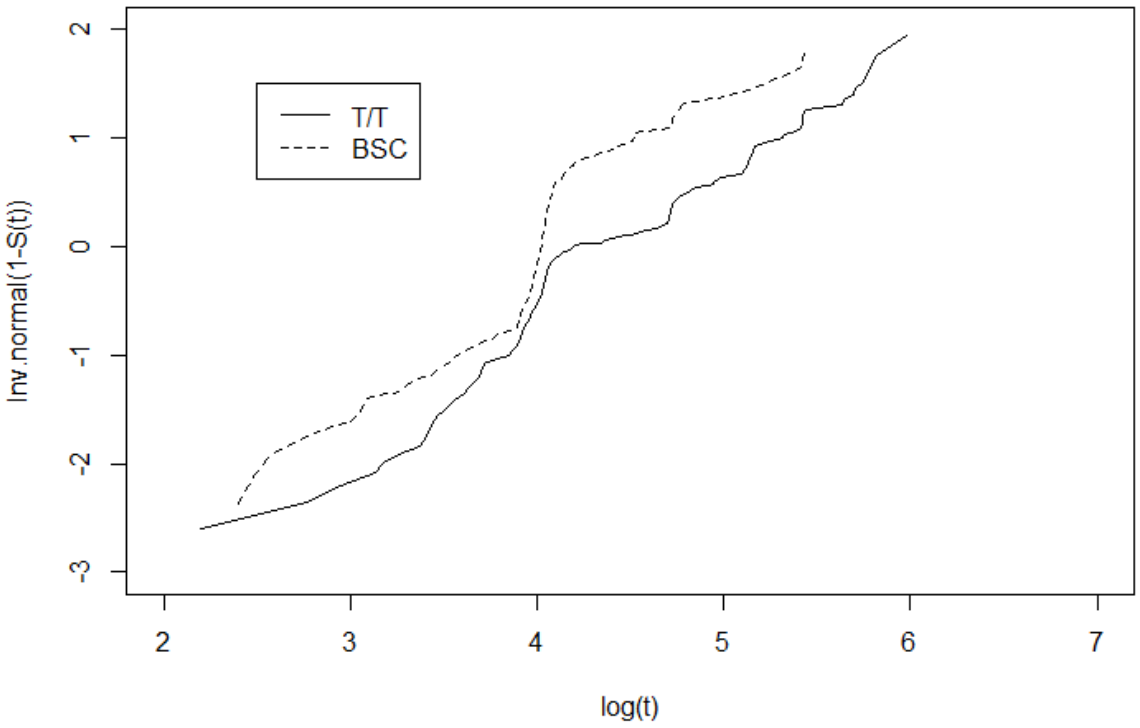
Figure 28: Logit survival plot – Progression-free survival (TAGS, no prior ramucirumab)



Key: BSC, best supportive care; S(t), survivor function; t, time; T/T, trifluridine/tipiracil.
Note: Approximately straight lines indicate that the survivor function is log-logistic.

The inverse Normal cumulative distribution function applied to the probability of a PFS event versus the log of time is shown in Figure 29. Neither curves for T/T or BSC are particularly linear, and so the log-normal PSMs may not provide a good fit to the data.

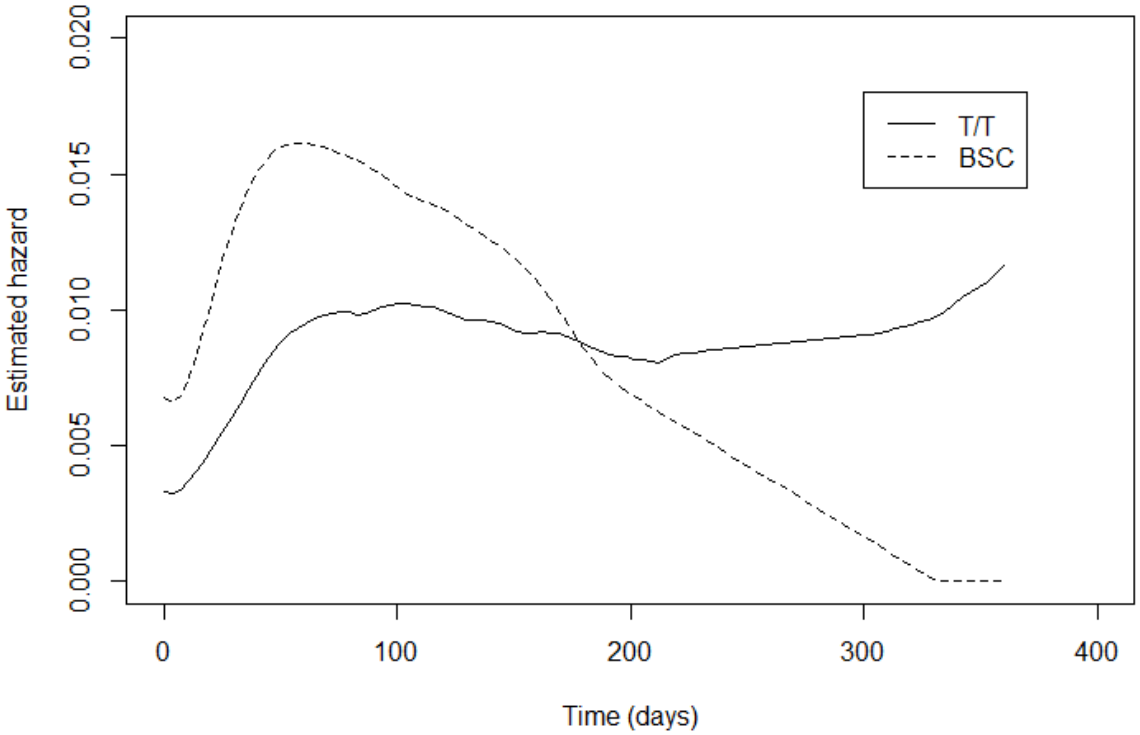
Figure 29: Inverse normal survival plot – Progression-free survival (TAGS, no prior ramucirumab)



Key: BSC, best supportive care; Inv.norm, Inverse normal S(t), survivor function; t, time; T/T, trifluridine/tipiracil.
Note: Approximately straight lines indicate that the survivor function is log-normal.

The final assessment of the survivor data undertaken comprises a smoothed hazard plot (Figure 30). The plots for both treatment arms demonstrate that the hazard of a PFS event is not constant over time, nor does it appear to be monotonic (that is, either consistently increasing or decreasing). This provides evidence to suggest that the exponential, Weibull, and Gompertz models may not provide a good fit to the data.

Figure 30: Smoothed hazard plots – Progression-free survival (TAGS, no prior ramucirumab)



Note: Turning points indicate the need for parametric survival models that are able to reflect non-monotonic hazard functions. A maximum time point of 360 days was selected to calculate the smoothed hazard estimation within the *muhaz* package.

Based on the diagnostic plots, it was determined that models with a covariate for treatment effect would be unlikely to provide a good fit to the data, as well as those that assume a constant or monotonic hazard function. Nevertheless, each of the 12 models were fitted for completeness.

To determine the best fitting PSM for PFS, the same criteria per the assessment for OS were used. Namely, the statistical goodness of fit, the visual fit within the observed period of data collection, and the plausibility of extrapolation beyond the observed period of data collection.

The statistical goodness-of-fit of all fitted PSMs is provided in Table 22. As the models fitted with a covariate for treatment would be unlikely to provide a better fit to those fitted independently by treatment arm, the “dependent” models were not considered further (but are available for testing within sensitivity analysis within the model). Based on the AIC and BIC scores, the 6 independent models were visually compared in order to select the base-case extrapolation (shown in Figure 31 and Figure 32 for T/T and

BSC, respectively). Visual fit within the observed period may be assessed on a per-model basis via the figures presented in Appendix O: Survival analysis.

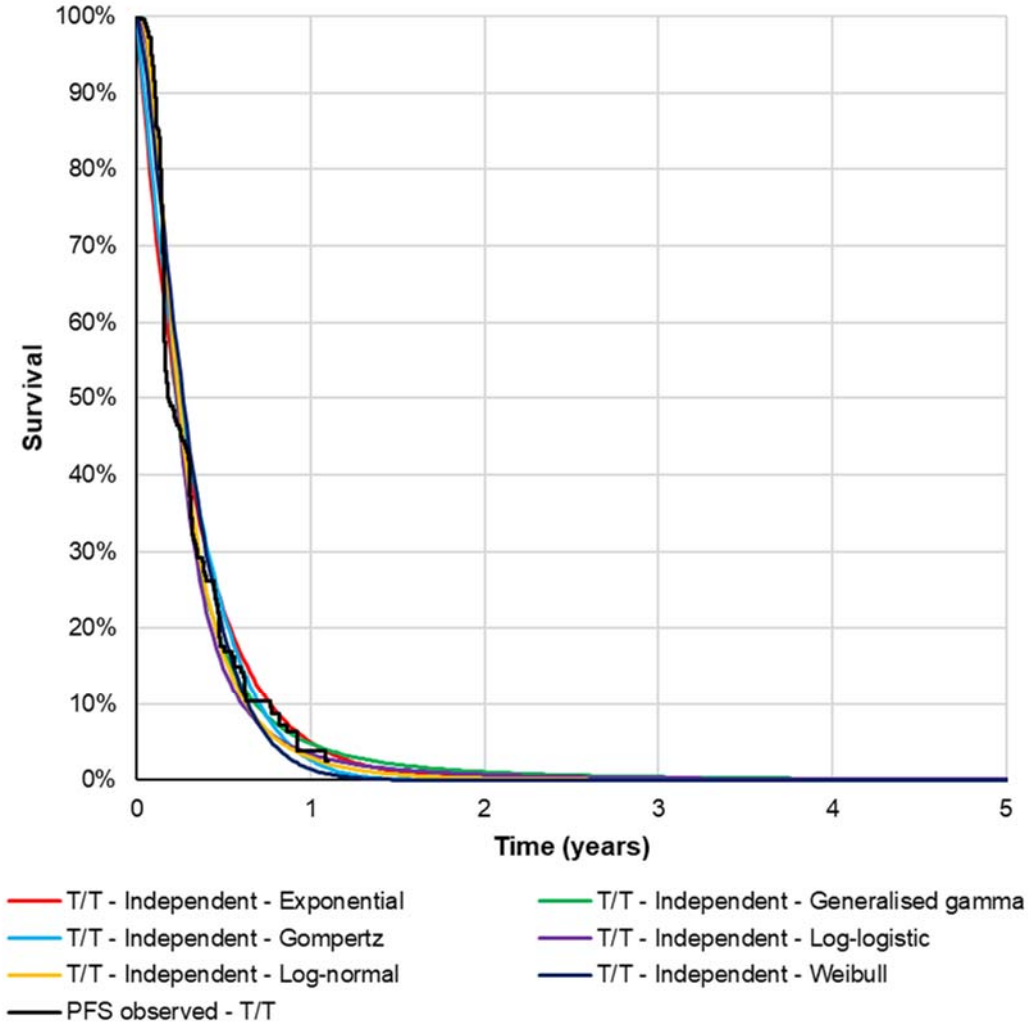
Table 22: Statistical goodness-of-fit scores (PFS, TAGS, no prior ramucirumab)

Dependence	Parameterisation	Statistical goodness-of-fit					
		T/T + BSC		PBO + BSC		Combined	
		AIC	BIC	AIC	BIC	AIC	BIC
Independent	Exponential	2,102.66	2,106.07	1,108.36	1,111.10	3,211.02	3,217.17
	Generalised gamma	2,077.30	2,084.12	1,090.75	1,096.24	3,168.06	3,180.36
	Gompertz	2,100.54	2,107.35	1,110.03	1,115.52	3,210.57	3,222.87
	Log-logistic	2,038.67	2,045.48	1,037.27	1,042.76	3,075.94	3,088.24
	Log-normal	2,033.75	2,040.57	1,052.00	1,057.49	3,085.75	3,098.06
	Weibull	2,028.75	2,038.97	1,051.25	1,059.48	3,080.00	3,098.46
Dependent	Exponential					3,211.02	3,218.67
	Generalised gamma					3,166.12	3,177.59
	Gompertz					3,208.93	3,220.40
	Log-logistic					3,082.12	3,093.59
	Log-normal					3,086.87	3,098.34
	Weibull					3,079.44	3,094.74

Note: AIC and BIC scores for the independent models were combined by simple addition – it is possible within the economic model to consider separate parameterisations for the independent models by treatment arm (for example, independent exponential for T/T + BSC and independent generalised gamma for PBO + BSC).

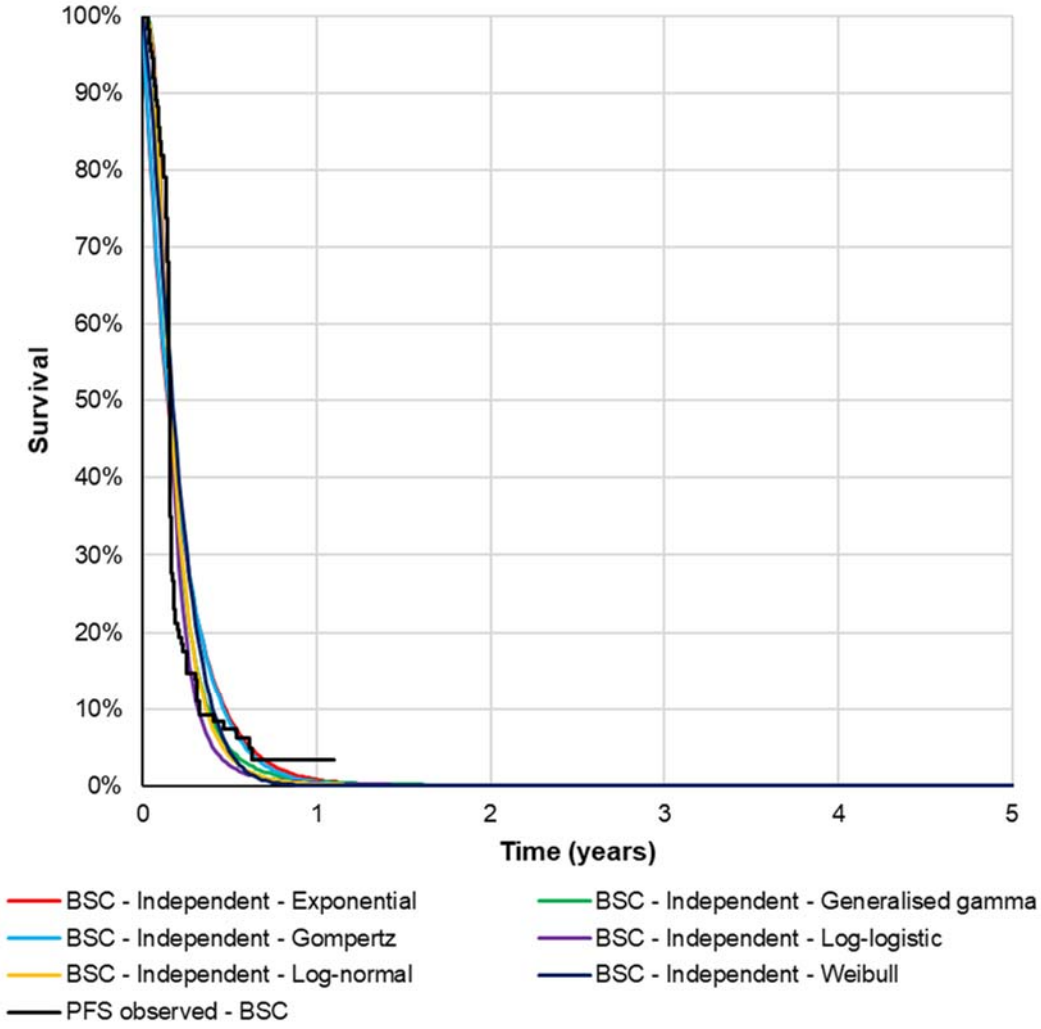
Key: AIC, Akaike Information Criterion; BIC, Bayesian Information Criterion; BSC, best supportive care; OS, overall survival; PBO, placebo; T/T, trifluridine/tipiracil.

Figure 31: Top six curve fits: PFS – T/T (TAGS, no prior ramucirumab)



Key: PFS, progression-free survival; T/T, trifluridine/tipiracil.

Figure 32: Top six curve fits: PFS – BSC (TAGS, no prior ramucirumab)



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

Owing to its flexibility, the independent generalised gamma was selected to inform the model base-case extrapolations for PFS. This choice of curve was based on the following:

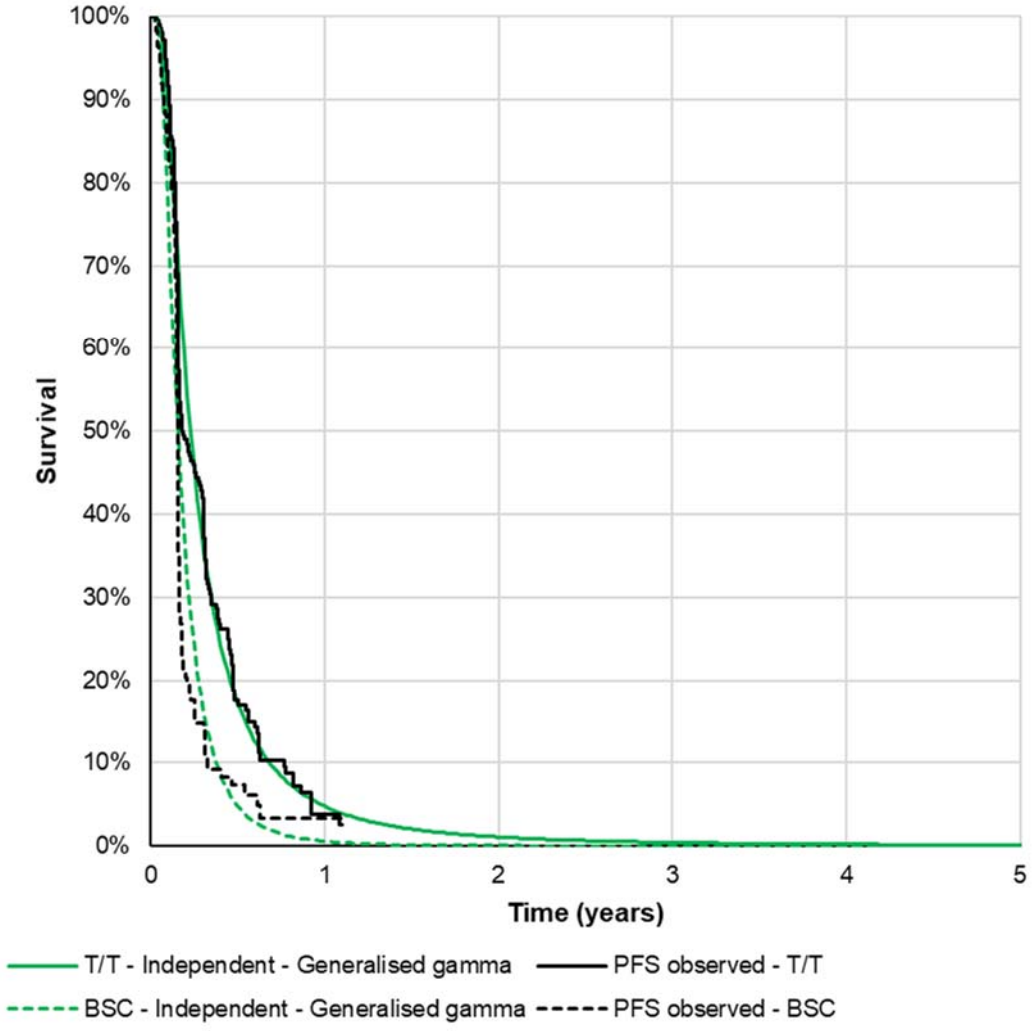
- The independent generalised gamma model provided a good visual fit to the Kaplan-Meier curves for PFS across both treatment arms
- While it does not yield the strongest statistical goodness-of-fit (measured via AIC and BIC), it was the only model fitted that could not be reasonably rejected based on an assessment of the underlying hazard function (for example, it may

be inferred from the logit survival plot [Figure 28] that a log-logistic model may not provide a good fit to the data)

- The long-term extrapolation of PFS was aligned with clinical expectation (0.00% for both arms by 5 years)

Alternative survival extrapolations were explored within scenario analysis. Details of these scenarios may be found in Section B.3.8. The base-case curve fits are provided in Figure 33.

Figure 33: Base-case PFS extrapolation (TAGS, no prior ramucirumab)



Key: BSC, best supportive care; OS, overall survival; PBO, placebo; T/T, trifluridine/tipiracil.

Scenario analysis: use of Kaplan-Meier curve

In the TAGS trial, the PFS curve is affected by the timing of assessments for progression (at approximately 2-monthly intervals). Assessments for progression were routinely performed every 8 weeks (in line with the protocol for the TAGS study. To address the potential influence of these “kinks” in the PFS Kaplan-Meier curves, a scenario was also undertaken using the Kaplan-Meier curve directly to inform the economic model. Due to the relative completeness of the curves (1.42% for trifluridine/tipiracil + BSC patients, and 2.46% for placebo + BSC patients [ITT population]; 2.55% for trifluridine/tipiracil + BSC patients, and 3.28% for placebo + BSC patients [no prior ramucirumab population]), the model allows for the option of assuming all patients immediately progressed at the end of the follow-up, or switching to a selected survival model to inform survival for the remainder of the model time horizon (at a given cut-point).

Selection of a relevant cut-point is somewhat arbitrary, and the choice of specific cut point is acknowledged to potentially have a large influence on cost-effectiveness results (which has been noted in several previous NICE appraisals of cancer treatments). [87, 88]

For this scenario, the cut-point was chosen based on the need to ensure a sufficient number of patients were still at risk, and after the first two “drops” in the PFS curve (as these constitute the key protocol-driven kinks in the curve). A cut-point of 84 days (12 weeks) was selected as the minimum plausible cut-point, as this is mid-way between the first and second assessments for progression (at 8 weeks and 16 weeks, respectively). An upper bound of 229 days (approximately 33 weeks) was selected as the maximum plausible cut-point, as this is approximately the last event time in the PFS curve for patients on the placebo arm of the TAGS trial.

In the scenario analysis, the PFS curve was assumed to utilise the KM curve until the cut-point, after which the base-case extrapolation was assumed to apply (that is, the estimated conditional survival estimates were lifted from the base-case extrapolation without re-basing the survival curve). The cut-point was varied from 12 to 33 weeks in weekly increments, and the impact on the ICER was recorded.

B.3.3.3 Safety

Safety data are available from the TAGS trial, allowing the occurrence of adverse events (AEs) to be captured by the economic model. Within the model, key AEs were associated with a cost of resolution and a HRQoL impact. The impact of AE occurrence on HRQoL and costs are discussed in Sections B.3.4 and B.3.5, respectively.

AEs of any grade were reported in 326 (97%) of 335 patients in the trifluridine/tipiracil group and 157 (93%) of 168 patients in the placebo group within the TAGS trial. Trifluridine/tipiracil was associated with a higher incidence of grade 3+ AEs than placebo (n=267, 80%; compared with n=97, 58%).

Treatment-emergent grade 3 or 4 AEs were included within the model, provided they occurred in at least 5% of patients in either treatment arm within the TAGS trial. This is consistent with the application of AEs in the previous NICE assessment of ramucirumab in gastric cancer (TA378), and similar to the approach used in the previous NICE assessment of trastuzumab for the treatment of HER2-positive mGC (TA208, though the impact on HRQL was excluded).

The only exceptions to the AE inclusion criteria were the addition of febrile neutropenia, which occurred in n=6 patients who received trifluridine/tipiracil; and nausea (n=14 trifluridine/tipiracil, n=5 BSC). Grade 3 or 4 febrile neutropenia was included within the cost-effectiveness model owing to its high impact on patient HRQL, as well as the cost of its treatment (which featured as the topic of a report by the NICE decision support unit).[89] Nausea was included based on clinical expert opinion that this was an important AE to consider.

Following the identification of relevant AEs, the total number of AE reports (grade 3 or 4) were used to inform the total costs incurred and QALYs lost as a result of the AE occurrence. These numbers are greater than or equal to the number of patients who experience each AE, as some patients may experience the same AE more than once.

A comprehensive presentation of safety data from the TAGS trial is provided within Section B.2.10. However, data used to inform the economic model are presented in Table 23.

Table 23: Adverse events included within the model

Adverse event	T/T + BSC (n= 335)		PBO + BSC (n=168)	
	Patients	Events	Patients	Events
Neutropenia	77 (23.0%)	155	-	-
Anaemia	37 (11.0%)	88	5 (3.0%)	16
Neutrophil count decreased	37 (11.0%)	86	-	-
Abdominal pain*	14 (4.2%)	16	15 (8.9%)	18
Decreased appetite	29 (8.7%)	33	11 (6.5%)	12
Fatigue	23 (6.9%)	24	10 (6.0%)	10
Leukopenia	23 (6.9%)	25	-	-
Asthenia	16 (4.8%)	18	11 (6.5%)	11
Ascites†	12 (3.6%)	21	10 (6.0%)	14
Febrile neutropenia	6 (1.8%)	6	-	-
Nausea	10 (3.0%)	14	5 (3.0%)	5

Key: BSC, best supportive care; PBO, placebo; T/T, trifluridine/tipiracil.

Note: Adverse events were included within the model subject to the following criteria: (1) treatment-emergent, (2) grade 3 or 4, (3) occurred in at least 5% of patients in either treatment arm, except for febrile neutropenia.
* Includes all adverse events labelled as 'abdominal pain', 'abdominal pain upper', and 'abdominal pain lower'.
† One patient on the PBO + BSC arm experienced grade 5 ascites (not included within the numbers presented).

B.3.3.4 Summary of clinical parameters used in the cost-effectiveness model

A summary of the clinical parameters used to inform the cost-effectiveness model is provided in Table 24.

Table 24: Summary of clinical parameters used in cost-effectiveness model

Outcome	Nature of inclusion within model	Rationale
Baseline patient characteristics	ITT population used to inform average age and proportion of female patients at baseline. T/T + BSC treated population used to inform estimation of patient BSA.	ITT population reflects both treatment arms at baseline, and constitutes the largest sample size to inform mean age and proportion of female patients. Treated T/T + BSC patients used to inform BSA given largest group exposed to T/T, and BSA is used only to inform cost of T/T in the model.
OS	Dependent log-normal PSM for both treatment arms.	Good fit to KM, lowest AIC and BIC scores, reasonable extrapolation of longer-term survival.
PFS	Independent generalised gamma PSM for both arms.	Good fit to KM, sufficiently flexible to reflect protocol-driven shape of curve, reasonable extrapolation of longer-term survival.
AEs	All treatment-emergent, grade 3 or 4 AEs occurring in (the equivalent of) at least 5% of patients in either treatment arm within the TAGS trial, plus all instances of grade 3 or 4 febrile neutropenia and nausea.	Most common AEs of high grade (per inclusion criteria of previous NICE assessments in advanced gastric cancer). Febrile neutropenia included due to high cost and utility impact. Nausea included following clinical expert feedback.

Key: AE, adverse event; AIC, Akaike Information Criterion; BIC, Bayesian Information Criterion; BSA, body surface area; BSC, best supportive care; KM, Kaplan-Meier; NICE, National Institute for Health and Care Excellence; OS, overall survival; PBO, placebo; PFS, progression-free survival; PSM, parametric survival model; T/T, trifluridine/tipiracil.

B.3.4 Measurement and valuation of health effects

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

In the TAGS clinical trial, the EQ-5D questionnaire was not administered to patients. However, data were collected using the EORTC-QLQ-C30 questionnaire – a cancer-specific preference-based measure of patient HRQoL.[90]

Patients completed the EORTC QLQ-C30 within 7 days before randomisation, before dose administration on Day 1 of treatment cycles ≥2, and at the safety follow-up 30 days after the last dose of treatment if not performed within the prior 4 weeks. The overall compliance rate was 84.0% for the QLQ-C30 questionnaire, which varied between treatment cycles from 72.5% to 100%.

To provide an indication of completion per cycle, Table 25 provides a summary of completion of three key aspects of the EORTC-QLQ-C30 relevant to the mGC population: global health status, emotional functioning, and physical functioning.

Table 25: Number of records for key aspects of EORTC-QLQ-C30 in TAGS

Cycle label	Number of recorded values		
	Global Health Status	Emotional Functioning	Physical Functioning
Baseline	493	493	496
Cycle 1	424	426	427
Cycle 2	267	267	269
Cycle 3	156	156	156
Cycle 4	107	107	107
Cycle 5	69	69	69
Cycle 6	47	47	47
Cycle 7	37	37	37
Cycle 8	29	29	29
Cycle 9	18	18	18
Cycle 10	13	13	13
Cycle 11	8	9	9
Cycle 12	3	4	4
Cycle 13	3	3	3
Cycle 14	1	1	1
Cycle 15	1	1	1
Safety Follow-Up	54	54	54

B.3.4.2 Mapping

As data from the EORTC-QLQ-C30 were collected within the TAGS trial, it was possible to apply a mapping algorithm to estimate EQ-5D values to inform the economic model.

A published mapping algorithm by Kontodimopoulos *et al.*, (2009) was selected to inform the model base-case.[91] This mapping algorithm was developed in a gastric cancer population, and so was considered the most appropriate to inform this appraisal. The Kontodimopoulos mapping algorithm is provided in Equation 1.

Equation 1: Mapping algorithm from Kontodimopoulos (2009)

$$EQ-5D = -0.18143 + GHS \times 0.00546 + EF \times 0.00313 + PF \times 0.00508$$

Key: GHS, Global health states; PF, physical functioning; EF, emotional functioning.

A recent conference abstract by Chau *et al.* found that the EORTC-QLQ-C30 was sensitive to clinical outcomes in advanced gastric cancer patients, particularly in global QoL, functional status and disease symptoms of fatigue, pain, and appetite loss.[92] This is aligned with the relatively simple algorithm developed by Kontodimopoulos *et al.* which found the most important aspects of the EORTC-QLQ-C30 to be global health status, physical functioning, and emotional functioning. When mapping the EORTC-QLQ-C30 to the SF-6D and 15D measures, Kontodimopoulos *et al.* found that only global health status and physical functioning were carried forward in each regression model (though emotional functioning also featured within the SF-6D model).

Based on the latest version of the University of Oxford Health Economics Research Centre (HERC) mapping database (24 April 2019), the algorithm by Kontodimopoulos *et al.* was the only published study that provides a mapping from the EORTC-QLQ-C30 for a gastric cancer population.[93] Other non-specific cancer mapping algorithms are available, though these were not considered as relevant to the patient population considered in this appraisal.

To estimate EQ-5D utilities, complete cases of the three relevant parameters from the EORTC QLQ-C30 were required (that is, global health score, physical functioning and emotional functioning). Where at least one of these responses was incomplete, the patient record was dropped. Duplicate entries were also removed – these were identified as observations recorded for a given patient on the same analysis day. In one instance, a patient reported four different sets of observations on the same analysis day. In this instance, all values were omitted from the analysis.

Patient records were separated by progression status such that utility values per model health state may be estimated. There were 74 distinct records where the EORTC-QLQ-C30 was administered after the patient was censored for PFS (that is, the patient was alive within an unknown progression status). These records were omitted from the analysis.

In total, 1,656 observations were used to calculate EQ- 5D utility values according to progression status using the Kontodimopoulos *et al.* mapping algorithm. 1,462 observations were collected for patient in the “progression-free” health state, and the remaining 194 observations were collected for progressed patients.

A tabulated summary of the mapped utility values by progression status is provided in Table 26. This table does not account for repeated measures for individual patients, and so should be interpreted with caution.

Table 26: Summary of mapped utility values by progression status

Health state	Mean	Median	Number of observations
Pre-progression	0.7644	0.7849	1,462
Post-progression	0.6522	0.6750	194

Generalised estimating equation (GEE) regressions were fitted to the utility data to account for repeated observations for individual patients. GEE regression methods have been used to estimate utility values to inform cost-effectiveness analysis in a number of previous NICE appraisals. Four regression models were considered:

1. Utility ~ Progression
2. Utility ~ Progression + treatment
3. Utility ~ Progression + no prior ramucirumab
4. Utility ~ Progression + treatment + no prior ramucirumab

An overview of the relevant statistical goodness-of-fit for each regression is provided in Table 27. The goodness of fit statistic QIC was used to assess the quality of the model fit, as AIC and BIC values cannot be calculated from a GEE model. The results of the GEE regressions are provided in Table 28.

Table 27: Statistical goodness-of-fit for GEE regressions

Model	QIC	Quasi Like	Trace	px
1 (progression)	99.83	-44.94	4.98	1,656
2 (progression + treatment)	104.73	-44.81	7.56	1,656
3 (progression + no prior ramucirumab)	106.42	-44.90	8.31	1,656
4 (progression + treatment + no prior ramucirumab)	111.40	-44.77	10.93	1,656

Table 28: GEE regressions output

Model 1 (progression)				
Coefficient	**Value**	**SE**	**Wald**	**p-value**
PF	0.7644	0.0105	5,262.3	<.0001
PP	0.6522	0.0236	765.8	<.0001
Estimated scale parameter: 0.0543 (SE 0.00388)				
Model 2 (progression + treatment)				
Coefficient	**Value**	**SE**	**Wald**	**p-value**
PF	0.7858	0.0164	2,290.38	<.0001
PP	0.6720	0.0281	572.75	<.0001
T/T	-0.0287	0.0210	1.87	0.17
Estimated scale parameter: 0.0541 (SE 0.00388)				
Model 3 (progression + no prior ramucirumab)				
Coefficient	**Value**	**SE**	**Wald**	**p-value**
PF	0.7602	0.0127	3,608.36	<.0001
PP	0.6472	0.0255	645.70	<.0001
Prior ramucirumab	0.0136	0.0224	0.37	0.54
Estimated scale parameter: 0.0542 (SE 0.00388)				
Model 4 (progression + treatment + no prior ramucirumab)				
Coefficient	**Value**	**SE**	**Wald**	**p-value**
PF	0.7816	0.0173	2,044.40	<.0001
PP	0.6670	0.0293	519.31	<.0001
Prior ramucirumab	0.0143	0.0224	0.41	0.52
T/T	-0.0290	0.0211	1.89	0.17
Estimated scale parameter: 0.0541 (SE 0.00386)				
Notes: Correlation: Structure = independence; Number of clusters: 492; Maximum cluster size: 15. **Key:** PF, progression-free; PP, post-progression; SE, standard error; T/T, trifluridine/tipiracil.				

The inclusion of treatment and prior ramucirumab experience as covariates did not improve model fit (as shown in Table 27) and were found to not be statistically significant predictors of utility (as indicated by the p values). Therefore, model 1 (progression) was selected to inform the cost-effectiveness analysis.

B.3.4.3 Health-related quality of life studies

Details of the systematic literature review undertaken to identify HRQoL studies relevant to the technology appraisal are provided within Appendix H: Health related quality of studies.

The review identified 1,938 records, and the relevant data were extracted from five unique studies identified from six publications. Of these, only one study was a primary utility elicitation study; the remaining four sourced utility data from other publications. The primary utility elicitation study was a population-based survey conducted in Japanese patients. This study reported that the mean health state utilities in advanced gastric cancer patients treated with third-line therapy were lower as compared to the general population.[94]

One Japanese study sourced the utility values for mGC patients treated in the third-line setting from an observational study and from the CheckMate 032 trial.[95] Two Chinese studies sourced utility values from the same study (Shiroiwa et al., 2011), which was conducted in treatment-naïve gastric cancer patients.[96] [97] In these Chinese studies, the utilities associated with the progression-free survival (PFS) health state were higher than those associated with the progressive disease (PD) health state. Another Chinese study sourced utility values from a study conducted by Carlson et al. in non-small cell lung cancer patients.[98]

None of the identified studies regarding the HRQoL of patients with mGC were considered directly relevant to the UK population considered within this appraisal.

B.3.4.4 Adverse reactions

Safety data from the TAGS trial were used within the economic model to explore the impact of AEs on patient utility. The safety data discussed within the context of the economic model in Section B.3.3, and more broadly in Section B.2.10.

To incorporate the impact of AEs on patient utility, the proportion of patients who experienced a given AE was taken from the TAGS trial and an associated loss in utility ('utility decrement' or 'disutility') was sourced from published sources. The loss in utility was multiplied by the duration over which the AE was expected to impact patient utility, and therefore a loss in QALYs due to each AE may be determined.

The disutility values and estimated durations of each AE impact are displayed in Table 29.

Table 29: Adverse event disutility values and duration of utility impact

Adverse event	Utility impact	Duration (days)	QALY loss	Source*
Neutropenia	-0.090	15.1	-0.004	Nafees (2008)
Anaemia	-0.119	16.1	-0.005	Swinburn (2010)†
Neutrophil count decreased	-0.090	15.1	-0.004	Assumed per neutropenia
Abdominal pain	-0.069	17.0	-0.003	Doyle (2008)‡
Decreased appetite	-0.048	3.0	0.000	Assumed per nausea
Fatigue	-0.073	21.0	-0.004	Nafees (2008), duration TA403
Leukopenia	-0.090	15.1	-0.004	Assumed per neutropenia
Asthenia	-0.073	21.0	-0.004	Assumed per fatigue
Ascites	-0.069	17.0	-0.003	Assumed per abdominal pain
Febrile neutropenia	-0.150	7.1	-0.003	Lloyd (2006)
Nausea	-0.048	3.0	0.000	Nafees (2008), duration TA403

Key: QALY, quality-adjusted life year.

Note: *All durations were taken from NICE TA378 (ramucirumab for treating advanced gastric cancer or gastro–oesophageal junction adenocarcinoma previously treated with chemotherapy) unless stated otherwise. †Calculated as the absolute difference between health state utility values for patients with and without anaemia, reported within this study. ‡Assumed to be equivalent to the disutility for 'pain' reported within this study.

The loss in utility due to AEs was accounted for within the economic model as a lump sum upon treatment initiation. This application allows for a simplistic calculation of the total QALYs expected to be lost, allowing also for repeated AEs to be captured. For the trifluridine/tipiracil + BSC group, this was a QALY loss of 0.00532, and for the placebo + BSC group this was a QALY loss of 0.00168.

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

A summary of the HRQoL data used to inform the cost-effectiveness analysis is provided in Table 30.

Table 30: Summary of utility values for cost-effectiveness analysis

State	Utility value: mean (SE)	95% CI	Reference in submission	Justification
PF	0.7644 (0.0105)	(0.7438, 0.7850)	Table 28	

PP	0.6522 (0.0236)	(0.6059, 0.6985)		GEE regression fitted to TAGS trial data.
AEs	See Table 29 for full explanation.			
Key: AEs, adverse events; CI, confidence interval; PF, progression-free; PP, post-progression; SE, standard error.				

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

B.3.5.1 Resource identification, measurement and valuation studies

A systematic literature review was undertaken to identify relevant cost and healthcare resource use data for patients in the third-line and beyond treatment setting (see Appendix I: Cost and healthcare resource identification, measurement and valuation, for further details).

Medical resource use estimates were taken from the most recent previous NICE appraisal conducted in mGC (TA378: ramucirumab for treating advanced gastric cancer or GEJ adenocarcinoma previously treated with chemotherapy). The estimates used to inform this submission formed the basis of medical resource utilisation within the current submission, which were validated with clinical experts and adjusted accordingly. A description of the changes made are provided within the dossier.

Other costs (such as treatment acquisition, administration, and end-of-life care) were informed through a combination of summaries of product characteristics, published literature, and clinical expert opinion.

B.3.5.2 Intervention and comparators' costs and resource use

Treatment acquisition

Trifluridine/tipiracil

Trifluridine/tipiracil is available in 15 mg/6.14 mg and 20 mg/8.19 mg tablets, in pack sizes of 20 and 60. For simplicity, the 15 mg/6.14mg and 20 mg/8.19mg tablets are termed "15 mg" and "20 mg", respectively (this labelling convention features within the Summary of Product Characteristics [SmPC] and other documentation regarding trifluridine/tipiracil). The unit costs of each pack size available for trifluridine/tipiracil

are provided in Table 31, at the list price and including the commercially-sensitive Patient Access Scheme (PAS) discount.

Table 31: Unit costs of treatment

Pack size	Cost (list)	Cost (including PAS)	Source
20 x 15 mg/6.14 mg ("15 mg")	£500.00	[redacted]	British National Formulary (BNF) online (Accessed 05-Feb-2019)
20 x 20 mg/8.19 mg ("20 mg")	£666.67	[redacted]	
60 x 15 mg/6.14 mg ("15 mg")	£1,500.00	[redacted]	
60 x 20 mg/8.19 mg ("20 mg")	£2,000.00	[redacted]	
Key: mg, milligram.			

BSC

No drug costs are associated with patients treated with BSC. As such, the drug cost per treatment cycle for patients treated with placebo + BSC was £0, and no additional drug cost for the BSC component for patients treated with trifluridine/tipiracil + BSC is included within the economic model.

Dosing

Trifluridine/tipiracil is administered orally at a dose of 35 mg/m² of body surface area (BSA) twice daily on days 1 to 5 and 8 to 12 of each 28-day treatment cycle. This dose was administered within the TAGS trial, and is representative of the anticipated licensed dose for mGC. Treatment with trifluridine/tipiracil is continued until the first occurrence of any of the following: disease progression (determined within TAGS per the RECIST criteria), unacceptable levels of toxicity, withdrawal of consent or death.

The licensed dose of trifluridine/tipiracil is based on patient body surface area (BSA), with pack sizes available to cater for all doses. The SmPC for trifluridine/tipiracil provides dosing bands based on BSA, which are presented in Table 32. Dosing bands in the SmPC are aligned with those considered within the TAGS trial.

Table 32: Dose bands for trifluridine/tipiracil based on body surface area

BSA band (m^2)	Dosage in mg (twice daily)	Tablet(s) per dose (twice daily)	
		15 mg	20 mg
< 1.07	35	1	1
1.07 - 1.22	40	0	2
1.23 - 1.37	45	3	0
1.38 - 1.52	50	2	1
1.53 - 1.68	55	1	2
1.69 - 1.83	60	0	3
1.84 - 1.98	65	3	1
1.99 - 2.14	70	2	2
2.15 - 2.29	75	1	3
≥ 2.30	80	0	4

Key: BSA, body surface area; mg, milligram.
Note: The figures provided in this table apply for the starting dose of trifluridine/tipiracil. Dosing adjustments may be required based on individual safety and tolerability.

BSA data are available from the TAGS trial. To determine the average cost of trifluridine/tipiracil per administration for inclusion within the economic model, the BSA of all patients who received at least one dose of trifluridine/tipiracil was considered (n=335 patients). The model allows the selection of patients irrespective of region (n=335), or only patients from the European Union (n=268). In addition, patients who were previously treated with ramucirumab may be excluded.

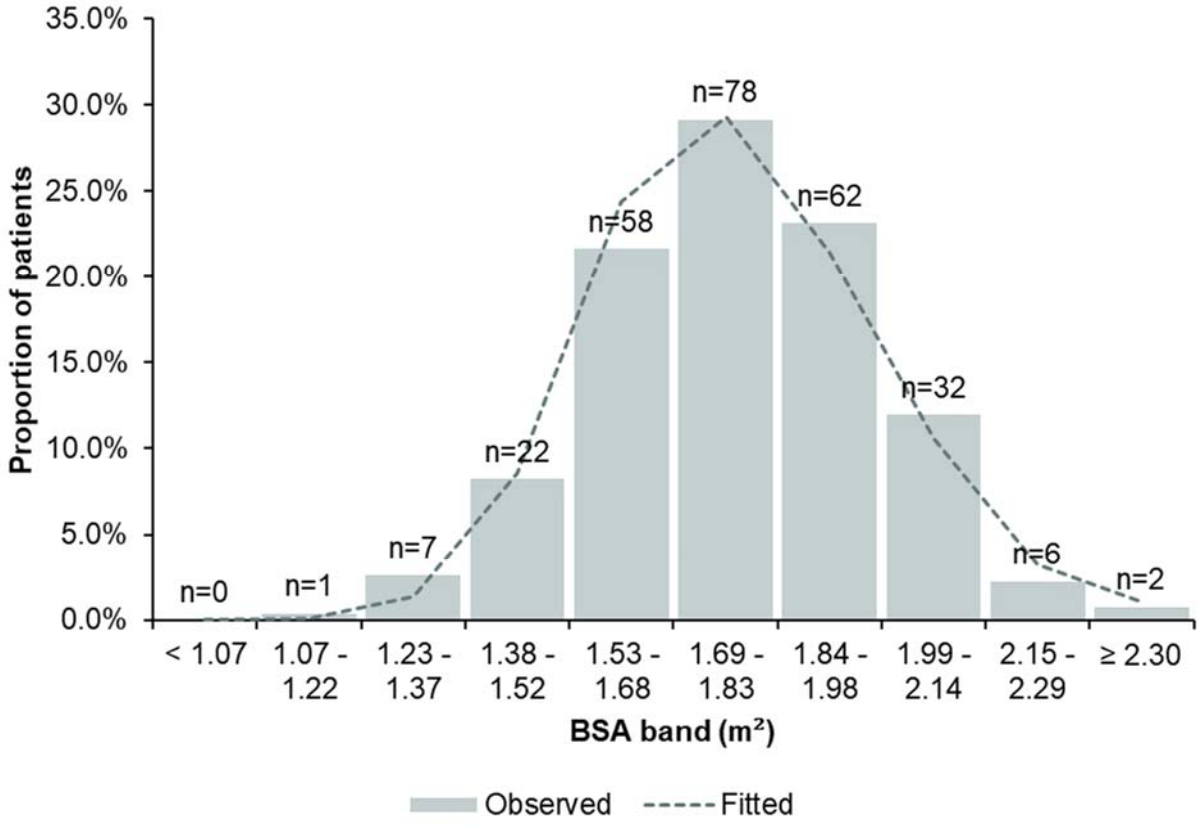
Clinical expert opinion noted that there was little evidence to suggest a difference in the BSA of patients with and without ramucirumab experience, and so the base-case analysis considers all patients regardless of treatment history. However, expert opinion noted that the use of EU patients only may be considered more reflective of the UK population.[3] Therefore, the base-case analysis considers all EU patients within the TAGS trial, with and without prior ramucirumab experience to inform the distribution of BSA.

A log-normal distribution was fitted to the distribution of BSA according to dosing band in acknowledgement of the limited sample size of the TAGS trial (which may influence the estimation of the proportion of patients in uncommon dosing bands [that is, at the limits]). The use of a log-normal distribution is advocated when considering the

"method of moments" approach to accurately costing the number of vials required for the administration of intravenous products, as *"fitting a distribution* [versus considering the observed data from the clinical trial] *would be less sensitive to the inclusion (or exclusion) of outliers in any given sample"*.[99]

The distribution of patients treated with trifluridine/tipiracil within the TAGS trial is presented in Figure 34 alongside the fitted distribution. The log-normal distribution provides a good fit to the observed proportions of patients in each dosing band.

Figure 34: TAGS: Distribution of body surface area (European patients who received at least one dose of trifluridine/tipiracil)



Key: BSA, body surface area; m, metre(s).

Note: The total number of European patients who received at least one dose of trifluridine/tipiracil within the TAGS trial was 268. 2 out of the 270 intention-to-treat (ITT) European patients randomised to trifluridine/tipiracil did not receive any doses.

Based on the dosing bands (Table 32) and fitted distribution for BSA (Figure 34), the weight average cost of trifluridine/tipiracil per 28-day treatment cycle was estimated at £2,184.01 (list price) [redacted].

Dose reductions

Three levels of dose reduction were reported within TAGS trial for patients treated with trifluridine/tipiracil. These dose reductions are described below:

- Level 1: reduction from 35 mg/m² to 30 mg/m² [redacted]
- Level 2: reduction from 30 mg/m² to 25 mg/m² [redacted]
- Level 3: reduction from 25 mg/m² to 20 mg/m² [redacted]

Each level of dose reduction is associated with specific BSA dosing bands, as displayed in Table 33.

Table 33: Dose bands associated with dose reduction levels for trifluridine/tipiracil based on body surface area

Level 1 Dose Reduction: 35 mg/m² to 30 mg/m²		**Level 2 Dose Reduction: 30 mg/m² to 25 mg/m²**		**Level 3 Dose Reduction: 25 mg/m² to 20 mg/m²**	
BSA (m²)	**Daily dose**	**BSA (m²)**	**Daily dose**	**BSA (m²)**	**Daily dose**
< 1.09	60 mg	< 1.10	50 mg*	< 1.14	40 mg
1.09 – 1.24	70 mg	1.10 – 1.29	60 mg	1.14 – 1.34	50 mg*
1.25 – 1.39	80 mg	1.30 – 1.49	70 mg	1.35 – 1.59	60 mg
1.40 – 1.54	90 mg	1.50 – 1.69	80 mg	1.60 – 1.94	70 mg
1.55 – 1.69	100 mg	1.70 – 1.89	90 mg	1.95 – 2.09	80 mg
1.70 – 1.94	110 mg	1.90 – 2.09	100 mg	2.10 – 2.34	90 mg
1.95 – 2.09	120 mg	2.10 – 2.29	110 mg	≥ 2.35	100 mg
2.10 – 2.28	130 mg	≥ 2.30	120 mg		
≥ 2.29	140 mg				

Key: BSA, body surface area; m, metre(s); mg, milligram.
Note: *At a total daily dose of 50 mg, patients should take 1 x 20 mg/8.19 mg tablet in the morning and 2 x 15 mg/6.14 mg tablets in the evening.

The same log-normal distribution for the distribution of BSA was used to inform the proportion of patients falling within each dose band. The number of patients being treated at each dose level (that is, either at the target dose of 35 mg/m^2 or one of the three dose reduction levels) was calculated for each treatment cycle using patient-level data from the TAGS trial. These data are presented in Table 34. The maximum number of treatment cycles for which treatment duration data are available for patients

receiving trifluridine/tipiracil is 14 cycles. Therefore, the distribution of patients by dosing level beyond 14 cycles was assumed to be fixed.

Table 34: Number of patients by dosing level (TAGS, trifluridine/tipiracil)

Cycle	N	35 mg/m²		30 mg/m²		25 mg/m²		20 mg/m²	
		N	%	N	%	N	%	N	%
1	[redacted]	[redacted]	[redacted]			[redacted]	[redacted]	[redacted]	[redacted]
2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
3	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
4	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
5	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
6	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
7	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
8	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
9	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
10	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
11	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
12	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
13	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
14	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Key: m, metre(s); mg, milligram.

Due to the relatively small number of patients that fall within each dosing level, the distribution of patients falling within each dosing level for the no prior ramucirumab and European populations were assumed to be equivalent to the safety population. This assumption is also used for the application of AE rates (that is, also based on the safety population) for the same reason (sample size).

Dose delays

Some patients treated with trifluridine/tipiracil may experience a delay in treatment. [redacted].[44] Within the model, [redacted] of patients are assumed to be delayed by 7 days (cycle length of the model) at cycle 2, and are subsequently costed a week later that the non-delayed patients for the remainder of the model time horizon. The option to exclude dose delays is explored within sensitivity analysis for completeness.

Duration of treatment

In practice, patients are expected to be treated with trifluridine/tipiracil in line with the protocol from TAGS – that is, until confirmed disease progression, death, intolerant toxicity or withdrawal of consent. Data from TAGS were used to inform the estimation of treatment duration with trifluridine/tipiracil within the model.

In order to accurately capture the duration of treatment, the treatment start and end dates were extracted from patient level data. Equation 2 was then used in order to calculate the length of treatment for patients within the TAGS trial. For example, a patient starting treatment on January 1st and discontinuing on January 4th (in the same year) would have a treatment duration of 4 days, and be recorded as having experienced the event of treatment discontinuation.

Equation 2: Treatment duration formula

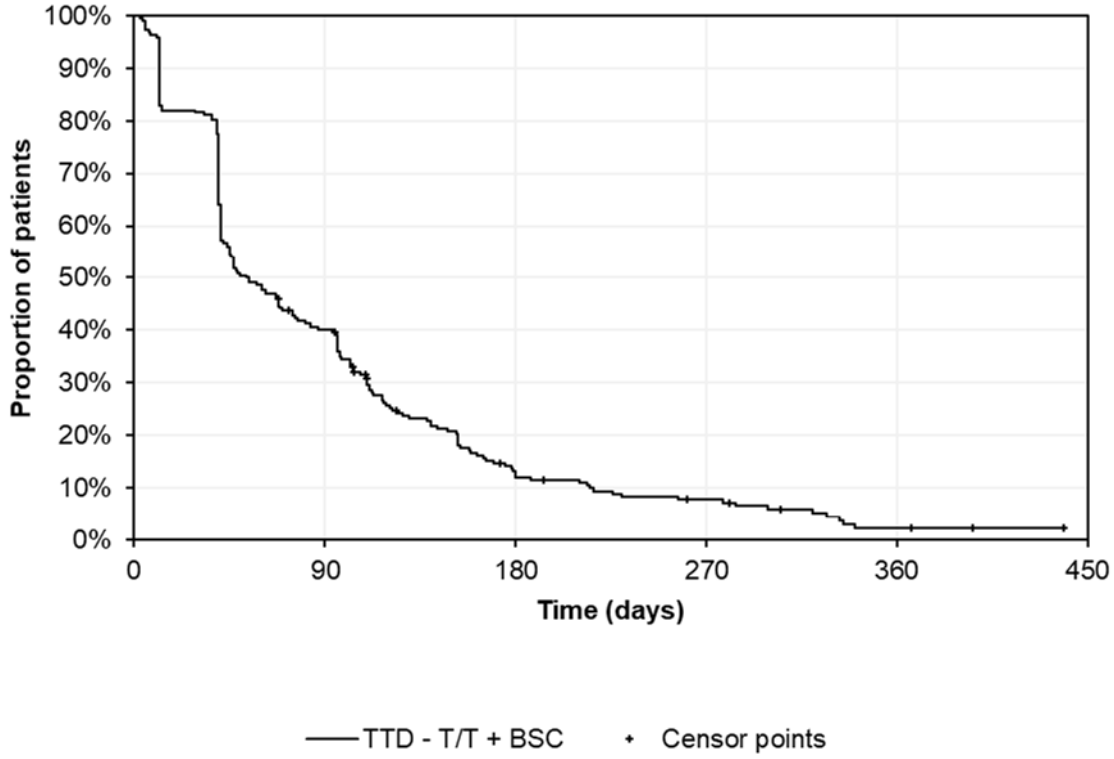
$$Treatment\ duration\ (days) = End\ date - Start\ date + 1$$

Censor events were assigned to patients who were not assigned a reason for treatment discontinuation within the patient-level data from the TAGS trial (that is, treatment was ongoing). All other events were defined as events for the purpose of informing the estimation of treatment duration. A summary of the reasons for treatment discontinuation are provided in Table 35, and the resultant time on treatment curve is presented in Figure 35.

Table 35: Reasons for treatment discontinuation (TAGS, trifluridine/tipiracil)

Reason for treatment discontinuation	Number of patients	
	All (n=335)	No prior ramucirumab (n=222)
Radiological progression	[redacted]	[redacted]
Clinical progression	[redacted]	[redacted]
Withdrawal by subject	[redacted]	[redacted]
Adverse event	[redacted]	[redacted]
Physician decision	[redacted]	[redacted]
Other (including deaths)	[redacted]	[redacted]
Total patients	**335 (100.0%)**	**222 (100.0%)**

Figure 35: TAGS: Time to treatment discontinuation (T/T, no prior ramucirumab)



Key: BSC, best supportive care; TTD, time to treatment discontinuation; T/T, trifluridine/tipiracil.

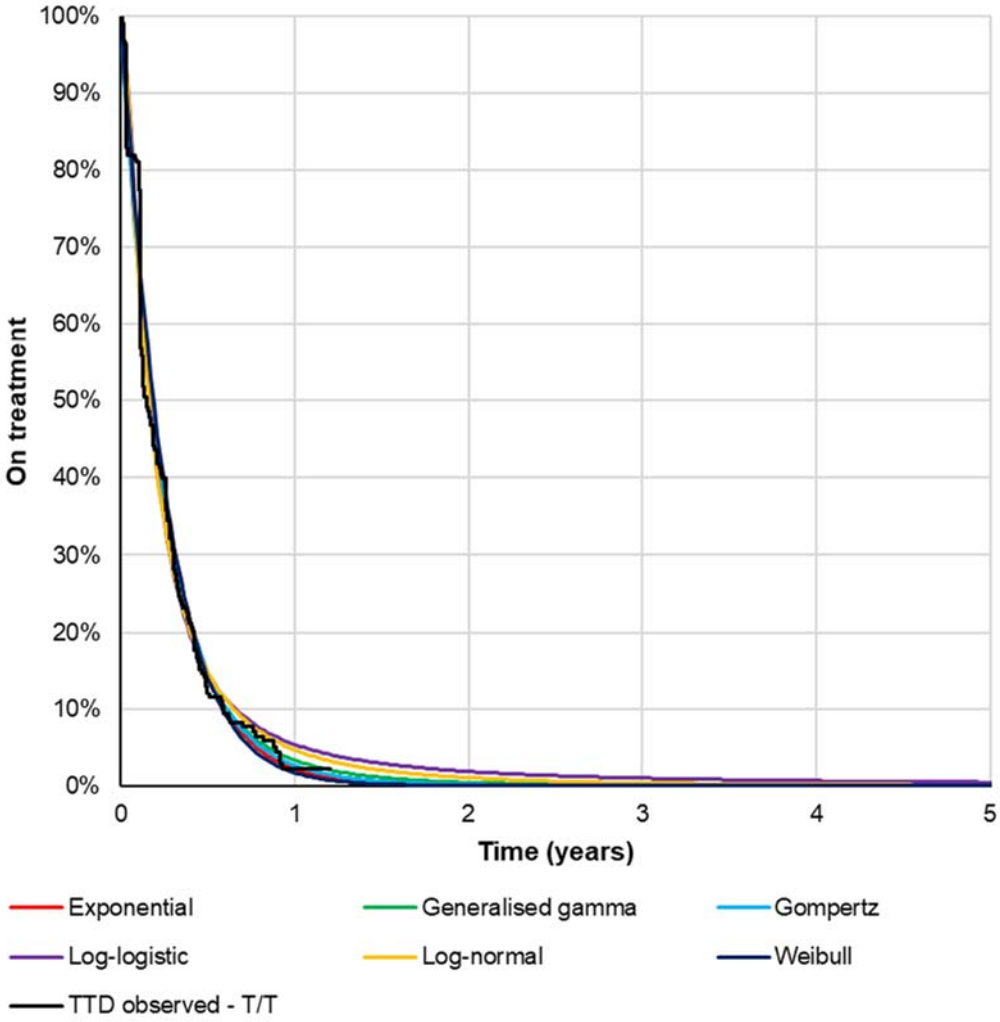
Independent PSMs were fitted to the TTD data from TAGS. For brevity, a full series of hazard plots is not presented within the dossier as each of the six candidate PSMs were considered based on an assessment of statistical goodness-of-fit, visual fit, and long-term plausibility. The statistical goodness-of-fit for each model fitted is presented in Table 36, and a plot of each curve is presented in Figure 36.

Table 36: Statistical goodness-of-fit scores (TTD, TAGS, no prior ramucirumab)

Model	AIC	BIC
Exponential	2,290.95	2,294.36
Weibull	2,291.80	2,298.61
Gompertz	2,292.57	2,299.37
Log-logistic	2,286.15	2,292.96
Log-normal	2,283.24	2,290.05
Generalised gamma	2,282.73	2,292.93

Key: AIC, Akaike Information Criterion; BIC, Bayesian Information Criterion; BSC, best supportive care; TTD, time to treatment discontinuation; T/T, trifluridine/tipiracil.

Figure 36: Curve fits: TTD – T/T (TAGS, no prior ramucirumab)



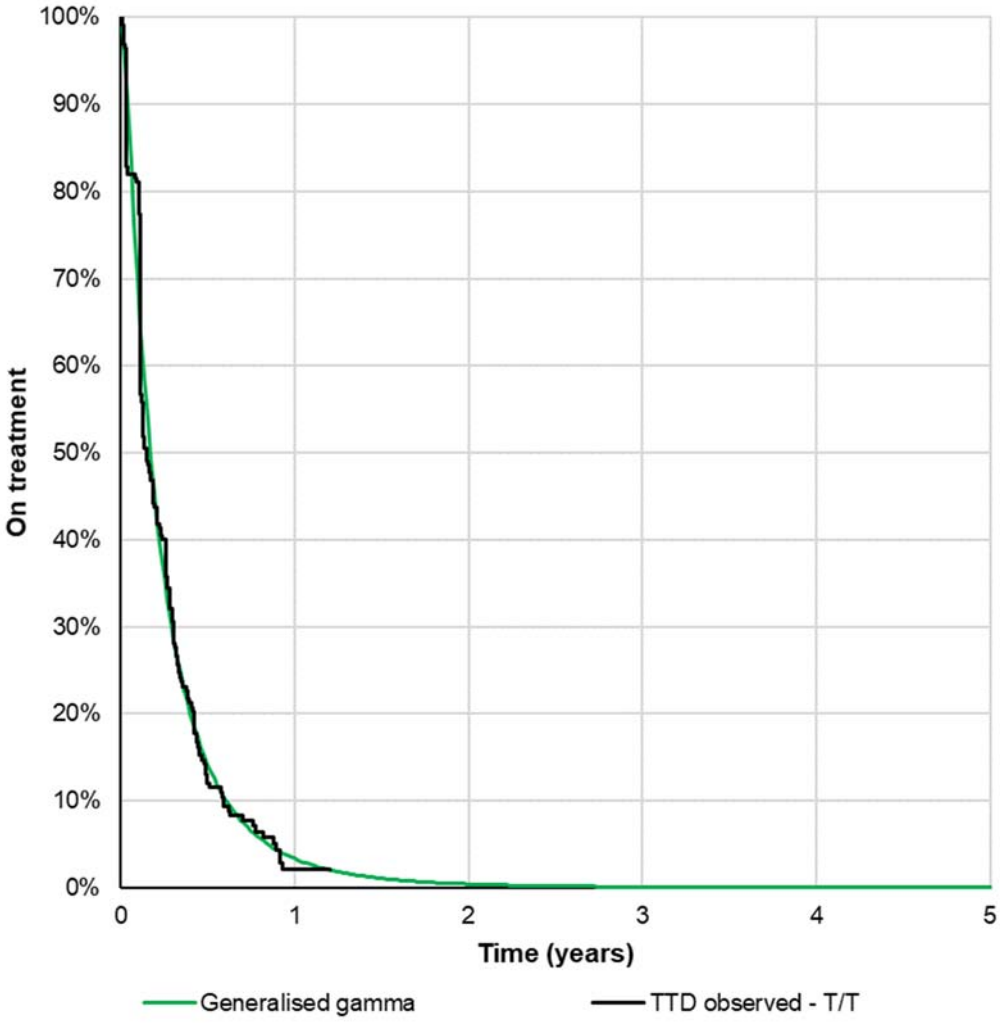
Key: TTD, time to treatment discontinuation; T/T, trifluridine/tipiracil.

Based on the assessment of visual fit, statistical goodness-of-fit and long-term plausibility, the generalised gamma model was chosen to inform the estimation of TTD. The base-case curve was selected due to the following:

- The independent generalised gamma model provided a good visual fit to the Kaplan-Meier curve for TTD
- The independent generalised gamma model had the lowest AIC and second-lowest BIC of all models tested
- The choice of TTD curve aligned with the choice of PFS curve (which are expected to follow a similar shape)

Alternative survival extrapolations were explored within scenario analysis. Details of these scenarios may be found in Section B.3.8. The base-case curve fits are provided in Figure 37.

Figure 37: Base-case TTD extrapolation (TAGS, no prior ramucirumab)



Key: TTD, time to treatment discontinuation; T/T, trifluridine/tipiracil.

To ensure model projections exhibit face validity, the TTD curve is capped according to the selected PFS curve (that is, should the TTD and PFS curves cross, TTD is set to the minimum of the two extrapolations). This may occur in probabilistic sensitivity analysis, or if two substantially different projections were selected to inform the base case. The independent of extrapolated survival curves is a known limitation of the partitioned-survival model structure, however the impact of this limitation is relatively small owing to the maturity of the data from the TAGS trial.

Administration

Trifluridine/tipiracil is an orally administered chemotherapy, and is expected to be taken by patients in the community setting (that is, not in a clinic). When discussed at the clinical advisory board meeting held by Servier, it was noted that some clinicians may choose to send patients to be seen by a chemotherapy nurse while they take their first treatment (though not all clinicians referred patients to see a nurse in practice).

Within the model base-case, it is assumed that all patients initiating trifluridine/tipiracil are seen by a chemotherapy nurse (assuming 30 minutes of Band 6 nurse time) to reflect the expected cost of initiation with treatment (£45 per working hour, based on the PSSRU Unit Costs of Health & Social Care 2018, Section 10.1).[100] After the first cycle, no further costs relating to administration are assumed to apply. The removal of this cost is explored within sensitivity analysis.

B.3.5.3 Health-state unit costs and resource use

The costs associated with each health state in the model are displayed in Table 19. The derivation of these costs (based on the associated unit costs and frequencies) are described within the remainder of this section of the dossier.

Table 37: Summary of health-state costs

Health states	Items	Cost		Reference in submission
		T/T + BSC	PBO + BSC	
Costs applied per 28-day cycle				
Pre-progression	Primary care	£0.00	£0.00	Table 38 and Table 39
	Secondary care	£162.05	£54.02	
	Scans	£44.10	£0.00	
	Tests	£4.72	£0.00	
	Total	**£210.88**	**£54.02**	Calculated
Post-progression	Primary care	£0.00	£0.00	Table 38 and Table 39
	Secondary care	£54.02	£54.02	
	Scans	£0.00	£0.00	
	Tests	£0.00	£0.00	
	Total	**£54.02**	**£54.02**	Calculated
One-off costs applied				
AEs	**Total**	**£306.26**	**£86.86**	Table 23 and Table 41
Progression	Surgery	£913.94	£1,079.34	Table 42
	Radiotherapy	£53.01	£65.68	
	SACT	£359.60	£386.48	
	Total	**£1,326.55**	**£1,531.51**	Calculated
Death	**Total**	**£306.26**	**£86.86**	Table 43

Key: BSC, best supportive care; PBO, placebo; T/T, trifluridine/tipiracil.

Note: Health-state costs for T/T + BSC apply only for patients currently receiving treatment with T/T. Upon discontinuation, the health-state costs for PBO + BSC patients are assumed to apply.

Medical resource use unit costs

Medical resource use (MRU) unit costs were taken from published sources, primarily NHS Reference costs (2016/17) and are displayed in Table 38.

Table 38: Medical resource use unit costs

Unit description	Cost	Source(s)
Consultation	£162.05	NHS Reference costs (2017/18). 370: Outpatient attendance - Medical Oncology
CT scan	£88.21	NHS Reference costs (2017/18). RD20A: Computerised Tomography Scan of One Area, without Contrast, 19 years and over
FBC	£2.51	NHS Reference costs (2017/18). DAPS05: Haematology
LFT	£1.11	NHS Reference costs (2017/18). DAPS04: Clinical Biochemistry
RFT	£1.11	

Key: CT, computerised tomography; FBC, Full blood count; LFT, Liver function tests; RFT, renal function test.

Medical resource use frequencies

MRU frequencies separated by progression status and treatment assignment were estimated based on previous NICE single technology appraisal submissions in mGC, and consultation with clinical experts. Consultations with an oncologist are expected to be required each 28-day treatment cycle with trifluridine/tipiracil, alongside a routine LFT, RFT, and FBC test. CT scans are expected to be every other treatment cycle. Following progression, patients may be expected to be seen by an oncologist once every 3 months.

Table 39: Medical resource use frequencies by treatment status

Medical resource use item	Receiving T/T	Not receiving T/T
Consultant	1.000	0.333
CT scan	0.500	0.000
FBC	1.000	0.000
LFT	1.000	0.000
RFT	1.000	0.000

Key: CT, computed tomography; FBC, full blood count; LFT, liver function test; RFT, renal function test; T/T, trifluridine/tipiracil.

In the previous NICE assessment of ramucirumab, a number of differences in medical resource use estimates were considered. CT scans were assumed to be required only once every 3 months. Costs relating to BSC were also included based on a retrospective chart review of medical resource utilisation for patients who received a platinum plus fluoropyrimidine in the first-line setting in the UK.[101]

In the model base-case, these costs are not included for the following reasons:

- CT scans are expected to be required once every other treatment cycle (that is, every 8 weeks instead of every 12 weeks) per the protocol of the TAGS trial, and the anticipated use of trifluridine/tipiracil in UK practice
- Several of the BSC costs identified are indicative of medical resource that would be utilised by patients towards the end of life, which are captured as an end-of-life care cost (for example, morphine for pain control)

- Based on the information available, it is not possible to establish how often these medical resources were used (for example, attendance of multiple counselling sessions), and thus accurately cost these within an economic model

- The inclusion of some costs may double count subsequent treatment costs (for example radiotherapy) or the costs associated with the resolution of adverse events (for example, antiemetics for nausea)

Notwithstanding the issues highlighted above, the model incorporates a scenario to explore the sensitivity of the cost-effectiveness results should the same medical resource use costs be applied per the previous ramucirumab appraisal in second-line mGC. Caution is emphasized when interpreting the results of this scenario, as some costs are expected to be double counted (though the extent to which they are double counted is not possible to establish). The costs used to inform this scenario are summarised in Table 40.

Table 40: Medical resource use frequencies by treatment status (ramucirumab appraisal scenario)

Medical resource use item	Receiving T/T	Not receiving T/T
Consultant	1.000	0.333
CT scan	0.333	0.000
FBC	1.000	0.000
LFT	1.000	0.000
RFT	1.000	0.000
Pain control (1)	42.10% of patients require 40 mg of morphine per day	62.90% of patients require 40 mg of morphine per day
Distress management (2)	10.50% of patients undergo 6 x CBT sessions per week	16.90% of patients undergo 6 x CBT sessions per week
Blood transfusion	8.80% of patients require 1 x RBC transfusion per month	23.80% of patients require 1 x RBC transfusion per month
RT	14.00% of patients require 1 x fraction of RT per month	11.90% of patients require 1 x fraction of RT per month

Key: CBT, cognitive behavioral therapy; CT, computed tomography; FBC, full blood count; LFT, liver function test; RBC, red blood cell; RFT, renal function test; RT, radiotherapy; T/T, trifluridine/tipiracil.

Notes: (1) Cost of morphine taken from eMIT (£10.04 for a pack of 56 x 20 mg tablets); (2) Cost per session of CBT estimated from PSSRU (2017) - £280 for 6 sessions inflated from 2016/17 to 2017/18; (3) Cost of one unit of RBC at £49 taken from Stokes et al., (2017) inflated from 2015/2016 to 2017/2018; (4) RT costed per post-progression costs (see Table 42).

B.3.5.4 Adverse reaction unit costs and resource use

The costs associated with adverse events were primarily taken from the NHS reference costs (2017/18), however values from other published sources were also used where applicable. These costs were used in combination with the probabilities of AE occurrence (identified from the TAGS trial). Adverse event costs are displayed in Table 41. The total costs of resolving AEs for trifluridine/tipiracil + BSC were £306.26 and for placebo + BSC were £86.86, applied within the model as a lump sum in the first cycle.

Table 41: Adverse event costs

Adverse event	Cost	Source(s)
Neutropenia, anaemia, neutrophil count decreased, leukopenia	£164.55	Assumed cost of FBC + outpatient medical oncology, based on clinical expert opinion
Abdominal pain. ascites	£319.68	NHS Reference costs (2017/18). Weighted average of day case abdominal pain with and without interventions (FD05A and FD05B)
Decreased appetite	£75.98	Unit cost of a dietician appointment, PSSRU (2018)
Fatigue, asthenia	£0.00	Assumed zero cost per clinical expert opinion
Febrile neutropenia	£4,619.81*	NICE decision support unit (2007) Risks and Costs of Febrile Neutropenia
Note: *£4,444.00 (2014 cost year) inflated using PSSSRU HCHS indices to 2016/17, then PSSRU HS index used (as the HCHS indices were discontinued).		

B.3.5.5 Miscellaneous unit costs and resource use

Post-progression costs

Following progression, patients may undergo surgery, radiotherapy or further rounds of systematic anti-cancer treatment (SACT). The proportion of patients in the TAGS trial undergoing these subsequent procedures and treatments following progression was recorded. The associated costs were extracted from NHS Reference costs (2016/17) and are displayed in Table 42.

Table 42: Post-progression costs and occurrence

Unit description	Proportion		Cost	Source
	T/T	BSC		
Surgery	13.9%	16.5%	£2,406.09	NHS Reference costs (2017/18). Weighted average of Malignant Gastrointestinal Tract Disorders, Elective inpatient (FD11A to FD11K) + 12 bed days (same code, duration based on CRUK [2019]).
RT	2.4%	2.9%	£2,233.23	NHS Reference costs (2017/18). SC47Z Preparation for Simple Radiotherapy with Imaging and Simple Calculation + SC31Z Deliver a Fraction of Adaptive Radiotherapy on a Megavoltage Machine. Assume 4 fractions in total (based on NICE TA378 assumption).
SACT	24.6%	26.5%	£1,460.05	Assumed 3-cycle course of docetaxel with same costs for resolving AEs as T/T.
Key: RT, radiotherapy; SACT, Systemic Anti-Cancer Therapy				

The expected cost is applied for all patients leaving the progression-free health state. For T/T patients, the total cost applied is £1,327 versus £1,532 for BSC patients (due to the increased use of post-progression treatment recorded in the TAGS trial). Post-progression costs are disabled within sensitivity analysis.

End of life care

In the final phases of life, patients with advanced cancer require a range of health and social care; as well as informal and charity care. End of life costs associated with the aforementioned categories for cancer patients with lung, breast, colorectal and prostate cancer are reported in a published study by Round *et al.*, (2015).[102] Costs for patients with gastric cancer were not reported, and so the costs incurred by colorectal cancer patients were applied as a proxy for gastric cancer (owing to the similar histological properties of colorectal and gastric cancers). This assumption was considered appropriate when discussed at the clinical advisory board meeting held by Servier.[3]

Health, social, charity and informal care costs were reported by Round *et al.*, (2015). While the costs associated with charity and informal care are quantified within the study, these costs were not included within the cost-effectiveness analysis owing to the NHS and PSS perspective adopted. This is consistent with the application of the

end-of-life care costs in previous appraisals conducted by NICE that have used this same reference.[103, 104]

The end of life care costs for colorectal cancer patients were reported separately for health (£4,854) and social care (£1,489), which were added for a total of £6,343 (cost year 2014). This cost was inflated via the Hospital and Community Health Services (HCHS) inflation indices from the PSSRU to determine the end of life care cost in the most recent cost year (2017), for a total of £6,594. This cost is applied within the cost-effectiveness analysis upon entry to the "Dead" health state.

Table 43: Cost of end-of-life care

Unit description	Proportion		Cost	Source
	T/T	BSC		
End-of-life care	100%	100%	£6,593.94	Round *et al.,* (2015) inflated using PSSRU indices
Key: BSC, best supportive care; T/T, trifluridine/tipiracil.				

B.3.6 Summary of base-case analysis inputs and assumptions

B.3.6.1 Summary of base-case analysis inputs

A summary of the base-case analysis inputs for the economic model is provided in Table 44.

Table 44: Summary of variables applied in the economic model

Variable	Value	Measurement of uncertainty	Reference to section in submission
Time horizon	**10**	**Fixed**	**B.3.2**
Cycle length	**7**	**Fixed**	
ADR - costs	**0.035**	**Fixed**	
ADR - QALYs	**0.035**	**Fixed**	
ADR - LYs	**0**	**Fixed**	
Age	**62.5**	**Normal (61.58, 63.42)**	
Female	**0.272**	**Beta (0.23, 0.31)**	
BSA	**1.77**	**Normal (1.77, 1.77)**	
AE prop: T/T - Neutropenia	**0.463**	**Beta (0.41, 0.52)**	**B.3.3**

AE prop: T/T - Anaemia	**0.263**	**Beta (0.22, 0.31)**	
AE prop: T/T - Neutrophil count decreased	**0.257**	**Beta (0.21, 0.31)**	
AE prop: T/T - Abdominal pain	**0.048**	**Beta (0.03, 0.07)**	
AE prop: T/T - Decreased appetite	**0.099**	**Beta (0.07, 0.13)**	
AE prop: T/T - Fatigue	**0.072**	**Beta (0.05, 0.1)**	
AE prop: T/T - Leukopenia	**0.075**	**Beta (0.05, 0.1)**	
AE prop: T/T - Asthenia	**0.054**	**Beta (0.03, 0.08)**	
AE prop: T/T - Ascites	**0.063**	**Beta (0.04, 0.09)**	
AE prop: T/T - Febrile neutropenia	**0.018**	**Beta (0.01, 0.03)**	
AE prop: T/T - Nausea	**0.042**	**Beta (0.02, 0.07)**	
AE prop: BSC - Neutropenia	**0**	**Beta (0, 0)**	
AE prop: BSC - Anaemia	**0.095**	**Beta (0.06, 0.14)**	
AE prop: BSC - Neutrophil count decreased	**0**	**Beta (0, 0)**	
AE prop: BSC - Abdominal pain	**0.107**	**Beta (0.06, 0.16)**	
AE prop: BSC - Decreased appetite	**0.071**	**Beta (0.04, 0.11)**	
AE prop: BSC - Fatigue	**0.06**	**Beta (0.03, 0.1)**	
AE prop: BSC - Leukopenia	**0**	**Beta (0, 0)**	
AE prop: BSC - Asthenia	**0.065**	**Beta (0.03, 0.11)**	
AE prop: BSC - Ascites	**0.083**	**Beta (0.05, 0.13)**	
AE prop: BSC - Febrile neutropenia	**0**	**Beta (0, 0)**	
AE prop: BSC - Nausea	**0.03**	**Beta (0.01, 0.06)**	
Utility: PF	**0.764**	**Multivariate normal**	**B.3.4**
Utility: PP	**0.652**	**Multivariate normal**	
AE Disutility: Neutropenia	**-0.09**	**Normal (-0.09, -0.09)**	

AE Disutility: Anaemia	**-0.119**	**Normal (-0.12, -0.12)**	
AE Disutility: Neutrophil count decreased	**-0.09**	**Normal (-0.09, -0.09)**	
AE Disutility: Abdominal pain	**-0.069**	**Normal (-0.07, -0.07)**	
AE Disutility: Decreased appetite	**-0.048**	**Normal (-0.05, -0.05)**	
AE Disutility: Fatigue	**-0.073**	**Normal (-0.07, -0.07)**	
AE Disutility: Leukopenia	**-0.09**	**Normal (-0.09, -0.09)**	
AE Disutility: Asthenia	**-0.073**	**Normal (-0.07, -0.07)**	
AE Disutility: Ascites	**-0.069**	**Normal (-0.07, -0.07)**	
AE Disutility: Febrile neutropenia	**-0.15**	**Normal (-0.15, -0.15)**	
AE Disutility: Nausea	**-0.048**	**Normal (-0.05, -0.05)**	
AE Duration: Neutropenia	**15.1**	**Normal (12.14, 18.06)**	
AE Duration: Anaemia	**16.1**	**Normal (12.94, 19.26)**	
AE Duration: Neutrophil count decreased	**15.1**	**Normal (12.14, 18.06)**	
AE Duration: Abdominal pain	**17**	**Normal (13.67, 20.33)**	
AE Duration: Decreased appetite	**3**	**Normal (2.41, 3.59)**	
AE Duration: Fatigue	**21**	**Normal (16.88, 25.12)**	
AE Duration: Leukopenia	**15.1**	**Normal (12.14, 18.06)**	
AE Duration: Asthenia	**21**	**Normal (16.88, 25.12)**	
AE Duration: Ascites	**17**	**Normal (13.67, 20.33)**	
AE Duration: Febrile neutropenia	**7.1**	**Normal (5.71, 8.49)**	
AE Duration: Nausea	**3**	**Normal (2.41, 3.59)**	
Cost: T/T - 20 x 15 mg	**500**	**Fixed**	**B.3.5**
Cost: T/T - 20 x 20 mg	**666.67**	**Fixed**	
Cost: T/T - 60 x 15 mg	**1500**	**Fixed**	
Cost: T/T - 60 x 20 mg	**2000**	**Fixed**	
Cost: T/T proportion delayed >7d	**0.136**	**Beta (0.11, 0.16)**	

Cost: T/T - admin	**22.5**	**Normal (18.09, 26.91)**	
Cost: N supervised doses /28d (first cycle)	**1**	**Fixed**	
Cost: MRU - Consultant	**162.049**	**Normal (161.99, 162.11)**	
Cost: MRU - CT scan	**88.207**	**Normal (88.17, 88.24)**	
Cost: MRU - FBC	**2.506**	**Normal (2.51, 2.51)**	
Cost: MRU - LFT	**1.109**	**Normal (1.11, 1.11)**	
Cost: MRU - RFT	**1.109**	**Normal (1.11, 1.11)**	
Cost: MRU - PP surgery cost	**6553.15**	**Normal (5268.76, 7837.54)**	
Cost: MRU - PP RT prep cost	**374.573**	**Normal (301.16, 447.99)**	
Cost: MRU - PP RT cost	**183.734**	**Normal (147.72, 219.74)**	
Cost: MRU - PP SACT cost	**1460.047**	**Normal (1173.88, 1746.21)**	
Cost: MRU - PP surgery prop (T/T)	**0.139**	**Beta (0.11, 0.17)**	
Cost: MRU - PP RT prop (T/T)	**0.024**	**Beta (0.02, 0.03)**	
Cost: MRU - PP SACT prop (T/T)	**0.246**	**Beta (0.2, 0.3)**	
Cost: MRU - PP surgery prop (BSC)	**0.165**	**Beta (0.13, 0.2)**	
Cost: MRU - PP RT prop (BSC)	**0.029**	**Beta (0.02, 0.04)**	
Cost: MRU - PP SACT prop (BSC)	**0.265**	**Beta (0.21, 0.32)**	
Cost: MRU - PP RT fractions	**4**	**Normal (3.22, 4.78)**	
Cost: AE - Neutropenia	**164.554**	**Normal (132.3, 196.81)**	
Cost: AE - Anaemia	**164.554**	**Normal (132.3, 196.81)**	
Cost: AE - Neutrophil count decreased	**164.554**	**Normal (132.3, 196.81)**	
Cost: AE - Abdominal pain	**319.677**	**Normal (257.02, 382.33)**	
Cost: AE - Decreased appetite	**75.98**	**Normal (61.09, 90.87)**	
Cost: AE - Fatigue	**0**	**Normal (0, 0)**	
Cost: AE - Leukopenia	**164.554**	**Normal (132.3, 196.81)**	

Cost: AE - Asthenia	**0**	**Normal (0, 0)**	
Cost: AE - Ascites	**319.677**	**Normal (257.02, 382.33)**	
Cost: AE - Febrile neutropenia	**4619.811**	**Normal (3714.34, 5525.28)**	
Cost: AE - Nausea	**163.583**	**Normal (131.52, 195.65)**	
Cost: EoL (health) - Lung	**3157**	**Normal (2538.24, 3775.76)**	
Cost: EoL (health) - Breast	**4346**	**Normal (3494.2, 5197.8)**	
Cost: EoL (health) - Colorectal	**4854**	**Normal (3902.63, 5805.37)**	
Cost: EoL (health) - Prostate	**6687**	**Normal (5376.37, 7997.63)**	
Cost: EoL (social) - Lung	**1358**	**Normal (1091.84, 1624.16)**	
Cost: EoL (social) - Breast	**2843**	**Normal (2285.78, 3400.22)**	
Cost: EoL (social) - Colorectal	**1489**	**Normal (1197.16, 1780.84)**	
Cost: EoL (social) - Prostate	**2728**	**Normal (2193.32, 3262.68)**	
MRU freq: PF (T/T) On - Consultant	**1**	**Normal (0.8, 1.2)**	
MRU freq: PF (T/T) On - CT scan	**0.5**	**Normal (0.4, 0.6)**	
MRU freq: PF (T/T) On - FBC	**1**	**Normal (0.8, 1.2)**	
MRU freq: PF (T/T) On - LFT	**1**	**Normal (0.8, 1.2)**	
MRU freq: PF (T/T) On - RFT	**1**	**Normal (0.8, 1.2)**	
MRU freq: PF (T/T) Off - Consultant	**0.333**	**Normal (0.27, 0.4)**	
MRU freq: PF (T/T) Off - CT scan	**0**	**Normal (0, 0)**	
MRU freq: PF (T/T) Off - FBC	**0**	**Normal (0, 0)**	
MRU freq: PF (T/T) Off - LFT	**0**	**Normal (0, 0)**	
MRU freq: PF (T/T) Off - RFT	**0**	**Normal (0, 0)**	
MRU freq: PP (T/T) - Consultant	**0.333**	**Normal (0.27, 0.4)**	

MRU freq: PP (T/T) - CT scan	**0**	**Normal (0, 0)**	
MRU freq: PP (T/T) - FBC	**0**	**Normal (0, 0)**	
MRU freq: PP (T/T) - LFT	**0**	**Normal (0, 0)**	
MRU freq: PP (T/T) - RFT	**0**	**Normal (0, 0)**	
MRU freq: PF (BSC) - Consultant	**0.333**	**Normal (0.27, 0.4)**	
MRU freq: PF (BSC) - CT scan	**0**	**Normal (0, 0)**	
MRU freq: PF (BSC) - FBC	**0**	**Normal (0, 0)**	
MRU freq: PF (BSC) - LFT	**0**	**Normal (0, 0)**	
MRU freq: PF (BSC) - RFT	**0**	**Normal (0, 0)**	
MRU freq: PP (BSC) - Consultant	**0.333**	**Normal (0.27, 0.4)**	
MRU freq: PP (BSC) - CT scan	**0**	**Normal (0, 0)**	
MRU freq: PP (BSC) - FBC	**0**	**Normal (0, 0)**	
MRU freq: PP (BSC) - LFT	**0**	**Normal (0, 0)**	
MRU freq: PP (BSC) - RFT	**0**	**Normal (0, 0)**	

Key: admin, administration; ADR, annual discount rate; AE, adverse event; BSA, body surface area; BSC, best supportive care; CT, computed tomography; d, day(s); EoL, end-of-life; FBC, full blood count; freq, frequency; LFT, liver function test; LY, life year; MRU, medical resource use; N, number; PF, progression-free; PP, post-progression; prop, proportion; QALY, quality-adjusted life year; RFT, renal function test; RT, radiotherapy; SACT, systemic anti-cancer therapy; T/T, trifluridine/tipiracil.

B.3.6.2 Assumptions

A number of assumptions have been made within the economic model used to inform this submission. The key assumptions made within the model are summarised in Table 45. The rationale for each assumption is provided, alongside the corresponding section(s) of the dossier wherein further information may be found.

Table 45: Summary of key assumptions applied in the economic model

Assumption	Rationale	Section(s)
The distribution of BSA for European patients in the TAGS	These patients are the most comparable patients to those likely to be treated in UK settings. Option to use whole population considered in sensitivity analysis.	B.3.3, B.3.5

Assumption	Rationale	Section(s)
trial is generalisable to UK clinical practice.		
The placebo arm of the TAGS trial is representative of the likely outcomes for patients receiving BSC in UK practice.	Use of placebo as a proxy for BSC has been used to inform a number of previous NICE appraisals, and was considered appropriate by UK clinical experts.	B.3.3
Model cycle length of 7 days is appropriate for informing the economic model.	Cycle length sufficiently short to reflect the expected frequency of/ time between clinical events (for example, progression, death), while also ensuring the model is not overly burdensome to use efficiently.	B.3.2
Use of the no prior ramucirumab population to inform the model base-case is most reflective of the UK patient population.	This population reflects the UK treatment pathway, and is therefore more representative of patients who would be eligible for trifluridine/tipiracil in UK NHS practice. ITT population presented as a subgroup analysis.	B.3.2
Unused capsules of T/T are costed within the model as wastage.	T/T is expected to be administered to patients on day 1 of each 28-day treatment cycle (per current use of T/T for mCRC), and as such any unused treatment would be wasted.	B.3.5
Grade III/IV AEs occurring in >5% of patients in either arm of TAGS, plus febrile neutropenia and nausea, are the key AEs associated with detrimental HRQoL.	Inclusion criteria (Grade III/IV) aligned with both previous mGC appraisals conducted by NICE (TA208 and TA378), with additional AEs included based on clinical expert feedback. Option to exclude AE-related disutility considered in sensitivity analysis.	B.3.4, B.3.5
Dose delays were assumed to apply for the second cycle of treatment with T/T only.	Summary information regarding dose delays suggested a small proportion of patients experienced a delay of at least 8 days. This was assumed to apply on the second cycle (as no patients were delayed for the first cycle, per the definition of a delay in the TAGS trial). Further dose delays were not incorporated for simplicity. Option to exclude dose delays considered in sensitivity analysis.	B.3.5
Distribution of patients by dose level beyond 14 cycles of treatment with T/T was assumed to be fixed.	1 patient in the TAGS trial was still on treatment at cycle 14, and so the impact of varying the dose level beyond this point in time was expected to have a negligible impact on results. Option to exclude dose reductions considered in sensitivity analysis.	B.3.5
Colorectal cancer assumed to serve as a proxy for end-of-life care costs for mGC patients.	As mGC is relatively uncommon, no disease-specific end-of-life care costs were identified to inform the submission. Clinical experts advised end-of-life care for colorectal and gastric cancer patients would be expected to be similar owing to similarities in the underlying disease pathology of the cancers, hence colorectal was deemed a suitable proxy in the absence of gastric cancer specific data.	B.3.5

Key: AE, adverse event; BSA, body surface area; BSC, best supportive care; mGC, metastatic gastric cancer; NICE, National Institute for Health and Care Excellence.

B.3.7 Base-case results

Base-case incremental cost-effectiveness analysis results

The discounted base-case cost-effectiveness results for trifluridine/tipiracil versus BSC are provided in Table 46 (with PAS) and Table 47 (without PAS).

Trifluridine/tipiracil provides an additional 0.153 QALYs and 0.226 LYs (a 43.97% increase on baseline survival for patients receiving BSC), with incremental costs of £7,327 (with PAS) [redacted]. The incremental cost-effectiveness ratio (ICER) is £47,933 (with PAS) [redacted] per QALY gained.

Table 46: Base-case results (with PAS)

Technologies	Total			Incremental			ICER (£/QALY)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	
PBO + BSC	[redacted]	0.514	0.349				
T/T + BSC	[redacted]	0.740	0.502	7,327	0.226	0.153	47,933

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

Table 47: Base-case results (without PAS)

Technologies	Total			Incremental			ICER (£/QALY)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	
PBO + BSC	[redacted]	0.514	0.349				
T/T + BSC	[redacted]	0.740	0.502	[redacted]	0.226	0.153	[redacted]

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

A comparison of the clinical outcomes of the model versus the results from the TAGS clinical trial is provided in Appendix J. Disaggregated results of the base-case incremental cost effectiveness analysis are provided in Appendix J.

B.3.8 Sensitivity analyses

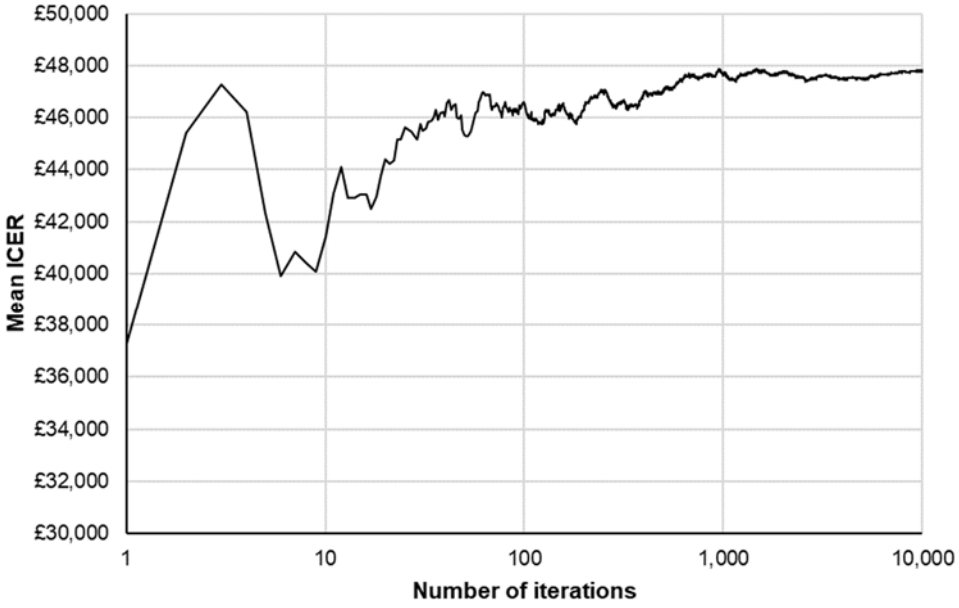
B.3.8.1 Probabilistic sensitivity analysis

Probabilistic sensitivity analysis (PSA) was conducted to explore the impact of parameter uncertainty on the results of the cost-effectiveness analysis. Within the PSA, all model parameters associated with parameter uncertainty were randomly sampled from their respective distributions, and the results of the cost-effectiveness analysis were recorded. This process was repeatedly performed until the mean results across the number of simulations were considered sufficiently stable.

To determine stability, a plot of the number of PSA iterations versus the mean ICER was produced, as shown in Figure 38 (ICER including PAS). From this plot, it may be

inferred that there is variation in the PSA results up until approximately 1,000 iterations, after which the ICER is relatively stable (i.e. the mean ICER is consistently between £47,000 and £48,000 per QALY gained). Consequently, 10,000 PSA iterations were used to inform the determination of average PSA results.

Figure 38: Stability of PSA



Key: ICER, incremental cost-effectiveness ratio.

The mean total costs and QALYs for each treatment arm were used to derive the mean incremental net monetary benefit (INMB) and ICER. The results of the PSA were also used to inform a PSA scatterplot and a cost-effectiveness acceptability curve (CEAC). The mean PSA results (including PAS) are presented in Table 48, with the corresponding results excluding the PAS discount provided in Table 49.

The results of the PSA are broadly aligned with the deterministic analysis results, however the results of the PSA highlight some of the limitations of the partitioned-survival analysis. For example, the parametric survival models for OS and PFS are sampled independently, and so this leads to a large volume of uncertainty in the estimation of the incremental QALY gain (which is not necessarily representative of the 'true' uncertainty).

Table 48: PSA results (with PAS)

Technologies	Total			Incremental			ICER (£/QALY)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	
PBO + BSC	[redacted]	0.518	0.351				
T/T + BSC	[redacted]	0.745	0.505	7,349	0.227	0.154	47,811

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

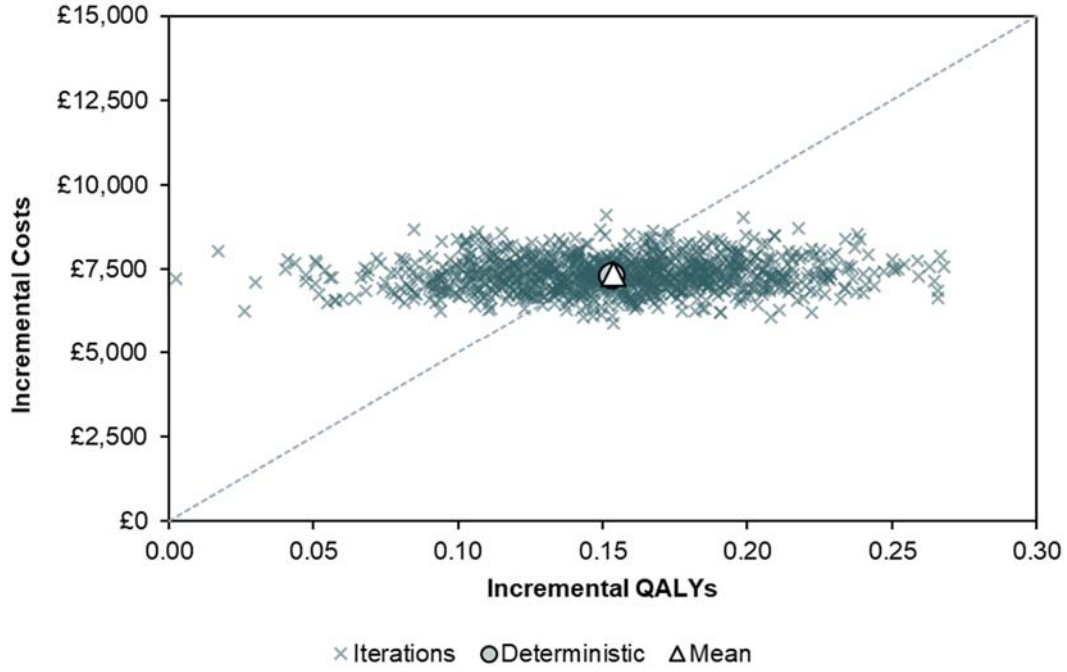
Table 49: PSA results (without PAS)

Technologies	Total			Incremental			ICER (£/QALY)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	
PBO + BSC	[redacted]	0.518	0.351				
T/T + BSC	[redacted]	0.745	0.505	[redacted]	0.227	0.154	[redacted]

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

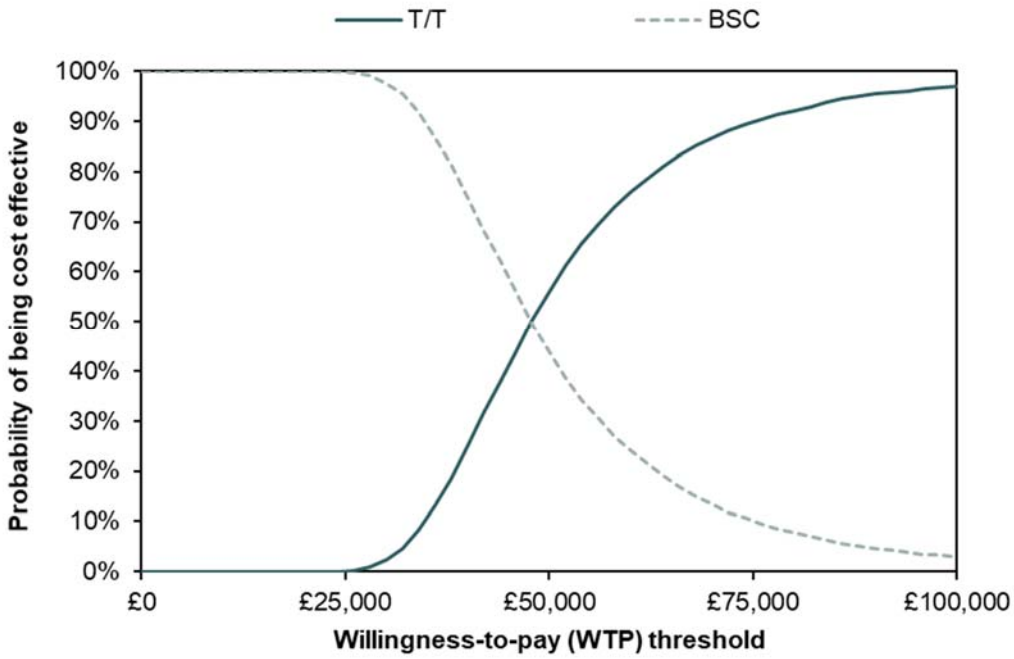
The PSA scatterplot and CEAC (including PAS) are presented in Figure 39 and Figure 40, respectively. The corresponding plots excluding the PAS discount provided in Figure 41 and Figure 42, respectively. These plots demonstrate that the majority of the uncertainty associated in the cost-effectiveness analysis is attributable to the estimation of the incremental QALY gain. The CEAC illustrates that at a willingness-to-pay threshold of £50,000 per QALY gained, trifluridine/tipiracil is associated with a probability of being a cost-effective treatment option of 56.2% (with PAS) [redacted].

Figure 39: PSA scatterplot (with PAS)



Key: QALY, quality-adjusted life year.

Figure 40: CEAC (with PAS)



Key: BSC, best supportive care; T/T, trifluridine/tipiracil.

Figure 41: PSA scatterplot (without PAS)

[redacted]

Key: QALY, quality-adjusted life year.

Figure 42: CEAC (without PAS)

Key: BSC, best supportive care; T/T, trifluridine/tipiracil.

B.3.8.2 Deterministic sensitivity analysis

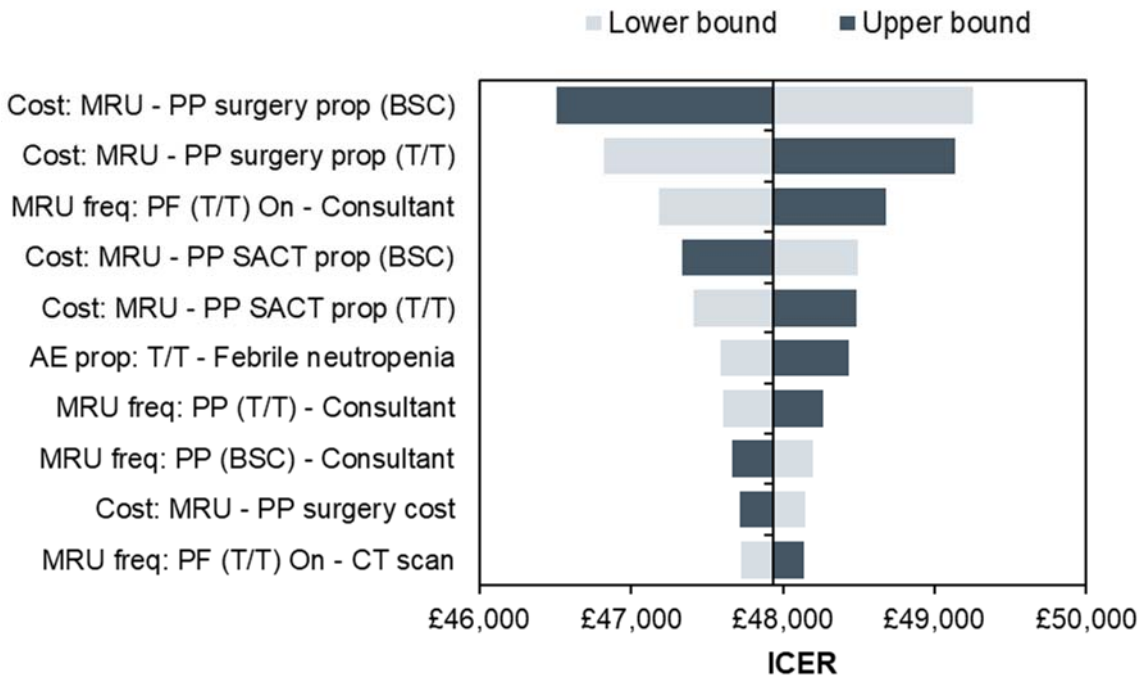
A deterministic one-way sensitivity analysis (OWSA) was implemented to evaluate which parameters were most influential on the outputs of the model. This analysis was performed by varying parameter inputs independently, at their upper and lower bounds. By recording the consequent impact on model results, the identification of key parameters of influence on the cost-effectiveness of trifluridine/tipiracil + BSC may be identified.

Model parameters with known covariance (for example, parametric survival model curve fit parameters) were not included in the deterministic sensitivity analysis, as the inclusion of these parameters may lead to counter-intuitive cost-effectiveness results. The exclusion of correlated parameters from the OWSA has also featured in recent previous NICE appraisals.[104, 105] The uncertainty associated with these parameters is instead explored in the probabilistic sensitivity analysis (PSA), and scenario analysis.

The outputs of the OWSA are comprised of a tornado diagram displaying the top ten most influential parameters. These results are also tabulated (should precise ICERs be required). The top ten most influential parameters are presented as a tornado diagram in Figure 43, with tabulated results in Table 50 (both with PAS). The corresponding results without the PAS discount for trifluridine/tipiracil are provided in Figure 44 and Table 51.

The most influential parameters on the ICER (within the context of the OWSA) were related to medical resource use, given that the key drivers of cost-effectiveness results are survival and HRQoL. Parameters relating to survival and utility were not included in this analysis as they are correlated (and so varying these in isolation would not be appropriate). Instead, the uncertainty associated with estimates of survival and utilities is discussed within the context of scenario analysis.

Figure 43: Tornado diagram (with PAS)



Key: AE, adverse event; BSC, best supportive care; CT, computed tomography; freq., frequency; MRU, medical resource use; PF, progression-free; PP, post-progression; prop, proportion; SACT, systemic anticancer therapy; T/T, trifluridine/tipiracil.

Table 50: Tabulated OWSA results (with PAS)

Rank	Parameter	LB	UB	Δ
1	Cost: MRU - PP surgery prop (BSC)	£49,252	£46,507	£2,745
2	Cost: MRU - PP surgery prop (T/T)	£46,824	£49,136	£2,313
3	MRU freq: PF (T/T) On - Consultant	£47,185	£48,681	£1,496
4	Cost: MRU - PP SACT prop (BSC)	£48,493	£47,339	£1,154
5	Cost: MRU - PP SACT prop (T/T)	£47,417	£48,486	£1,069
6	AE prop: T/T - Febrile neutropenia	£47,587	£48,436	£849
7	MRU freq: PP (T/T) - Consultant	£47,603	£48,263	£660
8	MRU freq: PP (BSC) - Consultant	£48,198	£47,668	£531
9	Cost: MRU - PP surgery cost	£48,150	£47,717	£433
10	MRU freq: PF (T/T) On - CT scan	£47,729	£48,137	£407

Key: AE, adverse event; BSC, best supportive care; CT, computed tomography; freq., frequency; MRU, medical resource use; PF, progression-free; PP, post-progression; prop, proportion; SACT, systemic anticancer therapy; T/T, trifluridine/tipiracil.

Figure 44: Tornado diagram (without PAS)

[redacted]

Key: AE, adverse event; BSC, best supportive care; CT, computed tomography; freq., frequency; MRU, medical resource use; PF, progression-free; PP, post-progression; prop, proportion; SACT, systemic anticancer therapy; T/T, trifluridine/tipiracil.

Table 51: Tabulated OWSA results (without PAS)

Rank	Parameter	LB	UB	Δ
1	Cost: MRU - PP surgery prop (BSC)	[redacted]	[redacted]	[redacted]
2	Cost: MRU - PP surgery prop (T/T)	[redacted]	[redacted]	[redacted]
3	MRU freq: PF (T/T) On - Consultant	[redacted]	[redacted]	[redacted]
4	Cost: MRU - PP SACT prop (BSC)	[redacted]	[redacted]	[redacted]
5	Cost: MRU - PP SACT prop (T/T)	[redacted]	[redacted]	[redacted]
6	AE prop: T/T - Febrile neutropenia	[redacted]	[redacted]	[redacted]
7	MRU freq: PP (T/T) - Consultant	[redacted]	[redacted]	[redacted]
8	MRU freq: PP (BSC) - Consultant	[redacted]	[redacted]	[redacted]
9	Cost: MRU - PP surgery cost	[redacted]	[redacted]	[redacted]
10	MRU freq: PF (T/T) On - CT scan	[redacted]	[redacted]	[redacted]

Key: AE, adverse event; BSC, best supportive care; CT, computed tomography; freq., frequency; MRU, medical resource use; PF, progression-free; PP, post-progression; prop, proportion; SACT, systemic anticancer therapy; T/T, trifluridine/tipiracil.

B.3.8.3 Scenario analyses

Scenario analyses were performed to explore the structural uncertainty within the model, including assumptions around the extrapolation of survival outcomes, specification of utility analyses, and other combinations of model input parameters. A summary of the scenarios considered is provided in Table 52.

Table 52: Scenario analyses performed

Scenario description	Rationale
Vary the time horizon from 1 to 10 years	Establish how influential the model time horizon is on the ICER.
Set annual discount rates to 0%	Understand how ICER is affected by discount rates.
Explore the use of all patients (including non-European) and restricting to only those with no prior ramucirumab treatment for BSA	Explore the impact on results were non-European patients accounted for (more patients, but potentially less representative for the UK), and the role of prior ramucirumab (no clear clinical basis for a difference, but provided for completeness).
Remove adjustments according to dose reductions and dose delays	Understand how influential these adjustments are on the costs associated with T/T.
Utilise same medical resource estimates per previous ramucirumab	Identify difference in total costs where the previous submission in second-line mGC estimates utilised, noting the risk of double counting.

Scenario description	Rationale
NICE technology appraisal in second-line mGC	
Exclude the cost of an initial nurse contact for patients receiving T/T	Not all patients are expected to require this appointment. Excluding the cost provides a plausible lower bound (and the base-case analysis constitutes a plausible upper bound, within the context of this model parameter).
Remove post-progression costs	Establish impact of post-progression costs on the ICER.
Vary the proxy cancer for the end-of-life care costs	Colorectal cancer chosen as a proxy, though the Round *et al.* study reports three other cancer types (lung, breast, prostate). Can also consider an average of the four.
Exclude the adjustment of survival curves for background mortality	Establish how the base-case survival estimates are affected by adjustment according to background mortality rates.
Vary OS curves*	Explore structural uncertainty in curve choice.
Vary PFS curves*	Explore structural uncertainty in curve choice.
Vary TTD curve	Explore structural uncertainty in curve choice.
Alternative regression models and sources for utility values	Choice of other parameters that may influence the utility values assigned (for example, treatment arm dependent, and experience of prior ramucirumab).
Exclude AE-related disutilities	Utility analyses did not adjust for the occurrence of AEs (such that the impact of these may be double counted). Scenario aims to show impact of double counting on results.
Exclude QALYs lost due to subsequent treatment	Illustrate how influential post-progression treatment QALY loss is on results.

Key: AE, adverse event; BSA, body surface area; ICER, incremental cost-effectiveness ratio; NICE, National Institute for Health and Care Excellence; OS, overall survival; PFS, progression-free survival; QALY, quality-adjusted life year; T/T, trifluridine/tipiracil; TA, technology appraisal; TTD, time to treatment discontinuation.

Note: *If the "KM" option is selected, by default this assumes the Kaplan-Meier curve is followed until 23 weeks, after which the base-case extrapolation is followed. 23 weeks was chosen as mid-way between 12 and 33 weeks (per the scenario analysis described in Section B.3.3).

Table 53: Scenario analysis results

Scenario label	ICER	
	With PAS	Without PAS
Base case	**£47,933**	[redacted]
Time horizon of 1 year	£76,552	[redacted]
Time horizon of 2 years	£57,809	[redacted]
Time horizon of 3 years	£52,627	[redacted]
Time horizon of 4 years	£50,455	[redacted]
Time horizon of 5 years	£49,369	[redacted]
Time horizon of 6 years	£48,757	[redacted]
Time horizon of 7 years	£48,401	[redacted]
Time horizon of 8 years	£48,178	[redacted]
Time horizon of 9 years	£48,032	[redacted]
Time horizon of 10 years	**£47,933**	[redacted]

Scenario label	ICER	
	With PAS	Without PAS
No discounting of costs, QALYs, or LYs	£46,438	[redacted]
Include non-European patients in BSA distribution	£47,455	[redacted]
Limit BSA distribution to only no prior ramucirumab patients	£48,125	[redacted]
Disable adjustment for background mortality	£47,932	[redacted]
Use mean BSA	£47,822	[redacted]
Use cost per mg assuming mean BSA	£45,862	[redacted]
Remove initial nurse cost	£47,786	[redacted]
Remove impact of down dosing	£49,727	[redacted]
Remove impact of dose delays	£48,255	[redacted]
Assume same medical resource use per TA378 (ramucirumab)	£50,880	[redacted]
Remove post-progression costs	£49,308	[redacted]
Assume lung cancer as a proxy for end-of-life care	£48,030	[redacted]
Assume breast cancer as a proxy for end-of-life care	£47,888	[redacted]
Assume prostate cancer as a proxy for end-of-life care	£47,770	[redacted]
Assume average of four cancers as a proxy for end-of-life care	£47,905	[redacted]
OS: Independent - Exponential, both arms	£54,520	[redacted]
OS: Independent - Generalised gamma, both arms	£47,339	[redacted]
OS: Independent - Gompertz, both arms	£48,506	[redacted]
OS: Independent - Log-logistic, both arms	£45,515	[redacted]
OS: Independent - Log-normal, both arms	£48,419	[redacted]
OS: Independent - Weibull, both arms	£47,535	[redacted]
OS: Dependent - Exponential, both arms	£54,520	[redacted]
OS: Dependent - Generalised gamma, both arms	£48,065	[redacted]
OS: Dependent - Gompertz, both arms	£53,665	[redacted]
OS: Dependent - Log-logistic, both arms	£45,248	[redacted]
OS: Dependent - Log-normal, both arms	**£47,933**	[redacted]
OS: Dependent - Weibull, both arms	£52,000	[redacted]
OS: KM, both arms	£41,130	[redacted]
PFS: Independent - Exponential, both arms	£47,636	[redacted]
PFS: Independent - Generalised gamma, both arms	**£47,933**	[redacted]
PFS: Independent - Gompertz, both arms	£47,287	[redacted]
PFS: Independent - Log-logistic, both arms	£47,865	[redacted]
PFS: Independent - Log-normal, both arms	£47,704	[redacted]
PFS: Independent - Weibull, both arms	£47,180	[redacted]
PFS: Dependent - Exponential, both arms	£48.301	[redacted]
PFS: Dependent - Generalised gamma, both arms	£48,749	[redacted]
PFS: Dependent - Gompertz, both arms	£47,400	[redacted]
PFS: Dependent - Log-logistic, both arms	£48,518	[redacted]

Scenario label	ICER	
	With PAS	Without PAS
PFS: Dependent - Log-normal, both arms	£48,227	[redacted]
PFS: Dependent - Weibull, both arms	£47,138	[redacted]
PFS: KM, both arms	£47,202	[redacted]
TTD: Exponential, T/T	£49,420	[redacted]
TTD: Generalised gamma, T/T	**£47,933**	[redacted]
TTD: Gompertz, T/T	£48,971	[redacted]
TTD: Log-logistic, T/T	£47,716	[redacted]
TTD: Log-normal, T/T	£47,526	[redacted]
TTD: Weibull, T/T	£49,847	[redacted]
TTD: KM, T/T	£48,042	[redacted]
Utility Model 1	**£47,933**	[redacted]
Utility Model 2	£49,774	[redacted]
Utility Model 3	£48,236	[redacted]
Utility Model 4	£50,137	[redacted]
Utility NICE TA378	£50,791	[redacted]
Utility NICE TA208	£48,791	[redacted]
Exclude AE-related disutilities	£46,848	[redacted]
Exclude QALYs lost due to AEs in subsequent treatment	£47,965	[redacted]

Key: AE, adverse event; BSC, best supportive care; BSA, body surface area; ICER, incremental cost-effectiveness ratio; KM, Kaplan-Meier; LY, life year; NICE, National Institute for Health and Care Excellence; OS, overall survival; PAS, patient access scheme; PFS, progression-free survival; QALY, quality-adjusted life year; T/T, trifluridine/tipiracil; TA, technology appraisal; TTD, time to treatment discontinuation.

Note: Base-case analyses highlighted in bold print throughout the table for context.

The majority of scenarios yielded an ICER of less than £50,000 (with PAS) [redacted]. The scenarios associated with the largest increase in the ICER were those relating to the restriction of the time horizon, the chosen survival curve(s), and alternative health state utility values.

Restricting the time horizon to 1 year is clearly inappropriate, as at this time approximately 21% of T/T patients are still alive. The base-case survival projections show a small proportion of patients may live for several years longer, necessitating a longer time horizon (10 years in the model base case) to reflect the lifetime costs and outcomes.

The survival curves shown to provide the largest estimates of the ICER are those that provide a relatively poor fit to the Kaplan-Meier curves (for example, exponential

models). Some curves caused a reduction in the ICER, in particular if the KM curve is applied for the first 23 weeks (see footnote of Table 52) followed by the base-case extrapolation, the ICER decreases markedly.

The base-case utility model was selected owing to it providing the best statistical fit to the available data from the TAGS trial. The ICER is shown to increase should covariates for T/T and no prior ramucirumab use be included. While provided for completeness, these models may double count the detrimental effects of T/T use (that is, disutility due to adverse events), and there is no clear rationale as to why utility for patients with prior ramucirumab experience should have a lower within-state utility score (confirmed by clinicians at the advisory board). As such, these scenario analyses should be interpreted with caution.

The utility values sourced from previous NICE appraisals are subject to a number of caveats.

- **TA378:** The post-progression utility value in TA378 was estimated using the mean utility score at the end of treatment for all patients who discontinued due to progressive disease (measured at the 30-day post-discontinuation visit), as opposed to being derived using a regression model that makes use of repeated observations at the individual level. Furthermore, the pre-progression utility value was assumed equal to baseline utility – this does not account for observations collected after treatment initiation and before disease progression

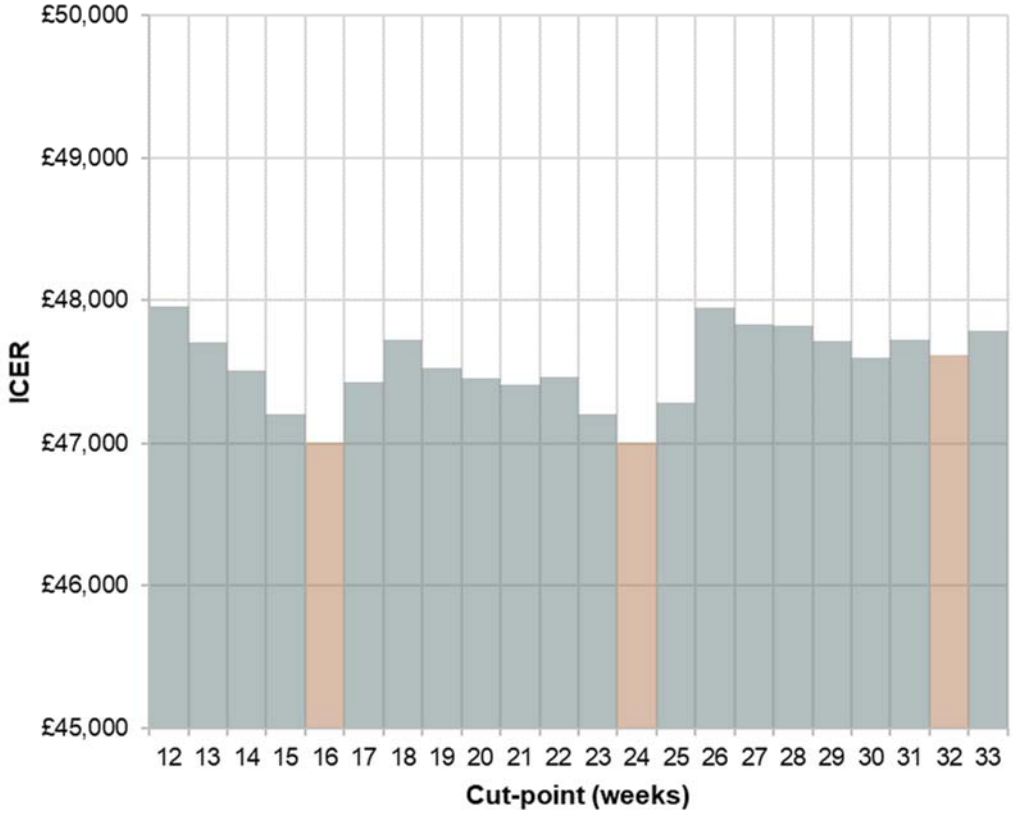
- **TA208:** The post-progression utility value in TA208 was taken from a previous NICE assessment of sunitinib for gastrointestinal stromal tumours (NICE TA179). In this appraisal, post-progression utility was calculated from the average EQ-5D score across both arms measured at the termination of the double-blind phase of the trial A6181004. Like TA378, repeated measures were not accounted for. For pre-progression, a daily utility increment of 0.000142 was assumed to apply on top of the baseline utility of 0.7292, which is not aligned with utility analyses conducted to inform contemporary cost-effectiveness analyses

Use of Kaplan-Meier followed by extrapolation for PFS

An additional scenario analysis was conducted to explore the effect of using the Kaplan-Meier curve for PFS (both arms) in combination with the base-case extrapolation. As discussed in Section B.3.3, the selection of a suitable cut-point may be considered arbitrary without a clear clinical rationale as to why the cut-point may be relevant. Instead, the cut-point was varied between 12 and 33 weeks and the ICER was recorded.

The results of this analysis are provided in Figure 45 (with PAS) and Figure 46 (without PAS). It was found that if the cut-point was set to the same time as an assessment point, the ICER reduced. This may be expected, as the majority of PFS events are expected to occur around the time of progression assessments. The ICER consistently fell between £47,000 and £48,000 (including the PAS, [redacted]).

Figure 45: Analysis of cut point for PFS curves (with PAS)



Key: ICER, incremental cost-effectiveness ratio.

Note: Protocol-specified assessment points (every 8 weeks) are highlighted in orange in the plot.

Figure 46: Analysis of cut point for PFS curves (without PAS)

Key: ICER, incremental cost-effectiveness ratio.
Note: Protocol-specified assessment points (every 8 weeks) are highlighted in orange in the plot.

B.3.8.4 Summary of sensitivity analyses results

A range of sensitivity analyses were undertaken to explore key areas of uncertainty in the submitted cost-effectiveness analysis. PSA demonstrated similar results to the deterministic base-case analysis, and highlighted the impact of small variation in the total QALYs on the ICER. Furthermore, the independence of modelled survival curves may (in part) explain the differences in probabilistic and deterministic results.

Deterministic one-way sensitivity analysis did not identify key model parameters that contribute to the overall uncertainty in the cost-effectiveness results. Instead, scenario analyses provided a comprehensive series of settings and assumptions that may impact results. Key drivers of results (excluding the specification of the model time horizon) were identified as the choice of survival curve for OS and PFS, as well as health state utility values used to inform the estimation of QALYs.

B.3.9 Subgroup analysis

The economic analysis was also performed using data for the ITT population, in line with the marketing authorisation (which does not exclude patients based on previous treatment with ramucirumab). However, as previously discussed, this population was not considered within the base-case analysis as ramucirumab is not recommended by NICE for use in the second line. The ITT population provides a larger sample size, at the cost of this sample being less relevant to the UK population (due to lack of recommendation of ramucirumab by NICE).

The headline results (with PAS) for the ITT population are presented in Table 54. The equivalent results excluding the PAS for trifluridine/tipiracil are provided in Table 55. Compared with the no prior ramucirumab results, the total incremental costs and QALYs are reduced, leading to a higher ICER.

Table 54: Headline economic analysis results for ITT population (with PAS)

Technologies	Total			Incremental			ICER (£/QALY)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	
PBO + BSC	[redacted]	0.513	0.346				
T/T + BSC	[redacted]	0.705	0.476	£6,928	0.192	0.130	£53,316
Key: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life years gained; PBO, placebo; QALYs, quality-adjusted life years; T/T, trifluridine/tipiracil.							

Table 55: Headline economic analysis results for ITT population (without PAS)

Technologies	Total			Incremental			ICER (£/QALY)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	
PBO + BSC	[redacted]	0.513	0.346				
T/T + BSC	[redacted]	0.705	0.476	[redacted]	0.192	0.130	[redacted]
Key: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life years gained; PBO, placebo; QALYs, quality-adjusted life years; T/T, trifluridine/tipiracil.							

B.3.10 Validation

Validation of cost-effectiveness analysis

The economic model was validated per the approaches detailed in Table 56.

Table 56: Validation of the cost-effectiveness analysis

Validation performed by	Nature of validation	Date	Aspects covered
Clinical advisory board attended by 12 UK practicing oncologists (in gastric cancer)	Review of key aspects of submission, including treatment pathway and model assumptions	April 2019	Overall submission and economic model
Delta Hat	Internal quality control check	May 2019	Economic model

B.3.11 Interpretation and conclusions of economic evidence

This submission presents the findings of an economic evaluation of trifluridine/tipiracil for the treatment of adult patients with mGC including adenocarcinoma of the gastroesophageal junction, who have been previously treated with at least two prior

systemic treatment regimens for advanced disease. A *de novo* cost-utility model was constructed to align with the NICE reference case, and estimate the likely cost-effectiveness of trifluridine/tipiracil in its anticipated positioning within UK NHS practice.

In the base-case analysis, trifluridine/tipiracil was associated with an ICER of £47,933 (including the approved, confidential PAS discount), and therefore represents a highly valuable end-of-life option for a population of patients with otherwise no recommended, effective treatment options. A series of sensitivity analyses were undertaken to demonstrate the role of uncertainty on the estimation of cost effectiveness – the results of these analyses support the conclusion that trifluridine/tipiracil offers a cost-effective treatment option.

The baseline survival for mGC patients receiving BSC may be one of the poorest assessed by NICE in recent history – lower than both baseline survival in the previous assessment of trifluridine/tipiracil in mCRC (TA405: 0.66 LYs, or 7.92 months) and nab-paclitaxel for (untreated) metastatic pancreatic cancer (TA476: 0.73 LYs, or 8.70 months). In these two previous appraisals, the incremental survival benefit was approximately 0.27 LYs (3.2 months, TA405) and 0.20 LYs (2.4 months, TA476) – both products met the end-of-life criteria specified by NICE. An extension in overall survival of 0.23 LYs (2.71 months) for trifluridine/tipiracil in mGC represents a mean 43.97% increase on baseline survival for patients receiving BSC alone (0.51 LYs, or 6.17 months). Thus, it may be inferred that the relative survival improvement provided by trifluridine/tipiracil in mGC is greater than both of these appraisals (40.90% and 27.40%, respectively); further illustrating the role of trifluridine/tipiracil for mGC as an end-of-life treatment option to address a high unmet medical need.

The economic model constructed to inform this appraisal was developed using similar methodology to other models constructed to inform NICE appraisals of treatments for mGC, as well as other late-stage cancers. The model is primarily informed by data collected as part of the pivotal, international, randomised, double-blind, placebo-controlled, phase III TAGS trial. The majority of patients recruited into the TAGS trial were European (n=408 of 507, 80.47%), and the majority received prior treatment in accordance with current UK guidelines. The use of prior ramucirumab was noted as a key limitation of the TAGS trial, and so the base-case analysis considers the sub-

population of patients with no experience of ramucirumab (aligned with current UK practice guidelines). The use of prior ramucirumab was a stratification factor within the TAGS trial.

The data available from the TAGS trial are mature, and as such the estimation of long-term survival outcomes within the economic model is not subject to a great deal of uncertainty. Quantification of medical resource utilisation is challenging with respect to a currently non-existent treatment line, and so clinical expert input was sought to validate the assumptions made within the model. HRQoL data were collected in TAGS in the form of the EORTC-QLQ-C30, which allows for mapping to NICE's preferred measure of utility to inform cost-effectiveness analysis (the EQ-5D-3L). While reliant upon mapping, the availability of utility data from the pivotal study allows for the calculation of health-state utility values to directly inform the economic model.

In conclusion, the economic evaluation presented in this submission demonstrates that trifluridine/tipiracil offers a cost-effective treatment option for patients with third-line and beyond mGC, for whom effective, recommended treatment options are currently unavailable. With a limited budgetary impact, minimally invasive route of administration, and a history of successful use for UK patients with mCRC; trifluridine/tipiracil provides invaluable extension in survival for patients at the end of life.

References

1. Ajani JA, D'Amico TA, Almhanna K, Bentrem DJ, Chao J, Das P, et al. Gastric Cancer, Version 3.2016, NCCN Clinical Practice Guidelines in Oncology. Journal of the National Comprehensive Cancer Network : JNCCN. 2016;14(10):1286-312.
2. Smyth EC, Committee obotEG, Verheij M, Committee obotEG, Allum W, Committee obotEG, et al. Gastric cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Annals of Oncology. 2016;27(suppl_5):v38-v49.
3. Servier. Data on file: UK Advisory board Meeting Report 22 March 2019. 2019.
4. Shitara K, Doi T, Dvorkin M, Mansoor W, Arkenau H-T, Prokharau A, et al. Trifluridine/tipiracil versus placebo in patients with heavily pretreated metastatic gastric cancer (TAGS): a randomised, double-blind, placebo-controlled, phase 3 trial. The Lancet Oncology. 2018;19(11):1437-48.
5. NICE. Ramucirumab for treating advanced gastric cancer or gastro–oesophageal junction adenocarcinoma previously treated with chemotherapy [TA378] 2016 [Available from: https://www.nice.org.uk/guidance/ta378/resources/ramucirumab-for-treating-advanced-gastric-cancer-or-gastrooesophageal-junction-adenocarcinoma-previously-treated-with-chemotherapy-82602795958981.
6. NICE. Palliative management for people with oesophageal and gastric cancer [Available from: https://pathways.nice.org.uk/pathways/oesophageal-and-gastric-cancer/palliative-management-for-people-with-oesophageal-and-gastric-cancer.
7. NICE. Oesophageal and gastric cancer overview [Available from: https://pathways.nice.org.uk/pathways/oesophageal-and-gastric-cancer.
8. Trifluridine/tipiracil. Summary of Product Characteristics 2017.
9. British Medical Association. British National Formulary: March -September 2019. 2019.
10. Nagini S. Carcinoma of the stomach: A review of epidemiology, pathogenesis, molecular genetics and chemoprevention. World J Gastrointest Oncol. 2012;4(7):156-69.
11. National Cancer Institute. Stages of Gastric Cancer." Gastric Cancer Treatment (PDQ®)–Patient Version. Retrieved August 14, 2018, from 2018 [Available from: https://www.cancer.gov/types/stomach/patient/stomach-treatment-pdq#section/_26.
12. Van Cutsem E, Sagaert X, Topal B, Haustermans K, Prenen H. Gastric cancer. The Lancet. 2016;388(10060):2654-64.
13. American Cancer Society. 2017 [Available from: https://www.cancer.org/cancer/stomach-cancer/about/what-is-stomach-cancer.html.
14. Riihimäki M, Hemminki A, Sundquist K, Sundquist J, Hemminki K. Metastatic spread in patients with gastric cancer. Oncotarget. 2016;7(32):52307-16.
15. National Cancer Institute. Gastric Cancer Treatment (PDQ®)–Patient Version 2018 [Available from: https://www.cancer.gov/types/stomach/patient/stomach-treatment-pdq.
16. Ramos M, Ribeiro Junior U, Viscondi JKY, Zilberstein B, Cecconello I, Eluf-Neto J. Risk factors associated with the development of gastric cancer - case-control study. Rev Assoc Med Bras (1992). 2018;64(7):611-9.
17. World Health Organisation. Cancer 2018 [Available from: https://www.who.int/news-room/fact-sheets/detail/cancer.
18. NICE. SINGLE TECHNOLOGY APPRAISAL: Ramucirumab for treating advanced gastric cancer or gastro-oesophageal junction adenocarcinoma after

chemotherapy [ID741]; Response to consultee, commentator and public comments on the Appraisal Consultation Document (ACD) 2015 [Available from: https://www.nice.org.uk/guidance/ta378/documents/committee-papers.
19. Maconi G, Manes G, Porro G-B. Role of symptoms in diagnosis and outcome of gastric cancer. World J Gastroenterol. 2008;14(8):1149-55.
20. Parkin DM, Bray F, Ferlay J, Pisani P. Global Cancer Statistics, 2002. CA: A Cancer Journal for Clinicians. 2005;55(2):74-108.
21. Edge SB, Compton CC. The American Joint Committee on Cancer: the 7th Edition of the AJCC Cancer Staging Manual and the Future of TNM. Annals of Surgical Oncology. 2010;17(6):1471-4.
22. De Mestier L. Recent insights in the therapeutic management of patients with gastric cancer. Dig Liver Dis 2016;48(9):984-94.
23. Wagner AD, Syn NL, Moehler M, Grothe W, Yong WP, Tai B-C, et al. Chemotherapy for advanced gastric cancer. Cochrane Database Syst Rev. 2017;8(8):CD004064-CD.
24. Kantar Health. Treatment architecture: Western Europe Gastric cancer. 2018.
25. Kim SM, Park SH. Chemotherapy beyond second-line in advanced gastric cancer. World J Gastroenterol. 2015;21(29):8811-6.
26. Zheng Y, Zhu X-Q, Ren X-G. Third-line chemotherapy in advanced gastric cancer: A systematic review and meta-analysis. Medicine (Baltimore). 2017;96(24):e6884-e.
27. Davidson M, Cafferkey C, Goode EF, Kouvelakis K, Hughes D, Reguera P, et al. Survival in Advanced Esophagogastric Adenocarcinoma Improves With Use of Multiple Lines of Therapy: Results From an Analysis of More Than 500 Patients. Clinical Colorectal Cancer. 2018;17(3):223-30.
28. Kang Y-K, Boku N, Satoh T, Ryu M-H, Chao Y, Kato K, et al. Nivolumab in patients with advanced gastric or gastro-oesophageal junction cancer refractory to, or intolerant of, at least two previous chemotherapy regimens (ONO-4538-12, ATTRACTION-2): a randomised, double-blind, placebo-controlled, phase 3 trial. The Lancet. 2017;390(10111):2461-71.
29. NICE. Developing NICE guidelines: the manual 2014 (updated October 2018) [Available from: https://www.nice.org.uk/process/pmg20/resources/developing-nice-guidelines-the-manual-2014-edition-pdf-6596134525.
30. Guo Y, Tang J, Huang XE, Cao J. Efficacy and toxicity of apatinib combined with or without chemotherapy for patients with advanced or metastatic chemotherapy-refractory gastric adenocarcinoma: A prospective clinical study. Medicine (Baltimore). 2019;98(6):e13908.
31. LI Jing, JIA Yongxu, Yanru Q. Clinical effect of apatinib combined with chemotherapy for advanced gastric cancer as second-line and above treatment. Chin J Cancer Biother. 2018;25(11):1135-9.
32. G Roviello, R Petroli, P Rosellini, A Giovanni Multari, R Conca, G Paganini. The influence of prior ramucirumab treatment on the clinical activity of FOLFIRI as third-line therapy in patients with metastatic gastric cancer. Invest New Drugs. 2019;37:524-30.
33. Bang YJ, Ruiz EY, Van Cutsem E, Lee KW, Wyrwicz L, Schenker M, et al. Phase III, randomised trial of avelumab versus physician's choice of chemotherapy as third-line treatment of patients with advanced gastric or gastro-oesophageal junction cancer: primary analysis of JAVELIN Gastric 300. Ann Oncol. 2018;29(10):2052-60.

34. Janjigian YY, Bendell J, Calvo E, Kim JW, Ascierto PA, Sharma P, et al. CheckMate-032 Study: Efficacy and Safety of Nivolumab and Nivolumab Plus Ipilimumab in Patients With Metastatic Esophagogastric Cancer. J Clin Oncol. 2018;36(28):2836-44.
35. Satoh T, Chen LT, Kang YK, Chao Y, Kato K, Chung HC, et al. 617PDA phase III study of nivolumab (nivo) in previously treated advanced gastric or gastric esophageal junction (G/GEJ) cancer (ATTRACTION-2): Two-years update data. Annals of Oncology. 2018;29(suppl_8).
36. Kang JH, Lee SI, Lim DH, Park KW, Oh SY, Kwon HC, et al. Salvage chemotherapy for pretreated gastric cancer: a randomized phase III trial comparing chemotherapy plus best supportive care with best supportive care alone. J Clin Oncol. 2012;30(13):1513-8.
37. Ohtsu A, Ajani JA, Bai Y-X, Bang Y-J, Chung H-C, Pan H-M, et al. Everolimus for previously treated advanced gastric cancer: results of the randomized, double-blind, phase III GRANITE-1 study. J Clin Oncol. 2013;31(31):3935-43.
38. Pauligk C, Lorenzen S, Goetze T, Riera Knorrenschild J, Hegewisch Becker S, Seraphin J, et al. 670PA randomized, double-blind, multi-center phase III study evaluating paclitaxel with and without RAD001 in patients with gastric or esophagogastric junction carcinoma who have progressed after therapy with a fluoropyrimidine/platinum-containing regimen (RADPAC). Annals of Oncology. 2017;28(suppl_5).
39. Pavlakis N, Sjoquist KM, Martin AJ, Tsobanis E, Yip S, Kang YK, et al. Regorafenib for the Treatment of Advanced Gastric Cancer (INTEGRATE): A Multinational Placebo-Controlled Phase II Trial. J Clin Oncol. 2016;34(23):2728-35.
40. Shitara K, Yuki S, Tahahari D, Nakamura M, Kondo C, Tsuda T, et al. Randomised phase II study comparing dose-escalated weekly paclitaxel vs standard-dose weekly paclitaxel for patients with previously treated advanced gastric cancer. Br J Cancer. 2014;110(2):271-7.
41. Fushida S, Kinoshita J, Kaji M, Oyama K, Hirono Y, Tsukada T, et al. Paclitaxel plus valproic acid versus paclitaxel alone as second- or third-line therapy for advanced gastric cancer: a randomized Phase II trial. Drug Des Devel Ther. 2016;10:2353-8.
42. Li J, Qin S, Xu J, Guo W, Xiong J, Bai Y, et al. Apatinib for chemotherapy-refractory advanced metastatic gastric cancer: results from a randomized, placebo-controlled, parallel-arm, phase II trial. J Clin Oncol. 2013;31(26):3219-25.
43. Li J, Qin S, Xu J, Xiong J, Wu C, Bai Y, et al. Randomized, Double-Blind, Placebo-Controlled Phase III Trial of Apatinib in Patients With Chemotherapy-Refractory Advanced or Metastatic Adenocarcinoma of the Stomach or Gastroesophageal Junction. Journal of Clinical Oncology. 2016;34(13):1448-54.
44. Data on file. Clinical Study Report TAGS. 2018.
45. Eric Van Cutsem, Lucjan Wyrwicz, Keun-Wook Lee, Fortunato Ciardiello, Rosine Guimbaud, Julien Taieb, et al. P-282JAVELIN Gastric 300: Phase 3 trial of avelumab (anti-PD-L1 antibody) + best supportive care (BSC) vs BSC ± chemotherapy as third-line treatment for advanced gastric or gastroesophageal junction cancer. Annals of Oncology. 2016;27(suppl 2):ii82.1-ii.
46. Bang Y-J, Wyrwicz L, Park YI, Ryu M-H, Muntean A, Gomez-Martin C, et al. Avelumab (MSB0010718C; anti-PD-L1) + best supportive care (BSC) vs BSC ± chemotherapy as third-line treatment for patients with unresectable, recurrent, or metastatic gastric cancer: The phase 3 JAVELIN Gastric 300 trial. Journal of Clinical Oncology. 2016;34(15_suppl):TPS4135-TPS.

47. Chau I, Bendell JC, Calvo E, Santana-Davila R, Ahnert JR, Penel N, et al. Interim safety and clinical activity in patients (pts) with advanced gastric or gastroesophageal junction (G/GEJ) adenocarcinoma from a multicohort phase 1 study of ramucirumab (R) plus pembrolizumab (P). Journal of Clinical Oncology. 2017;35(4_suppl):102-.
48. Liu J, Huang XE, Feng JF. Further study on pemetrexed based chemotherapy in treating patients with advanced gastric cancer (AGC). Asian Pac J Cancer Prev. 2014;15(16):6587-90.
49. Bando H, Doi T, Muro K, Yasui H, Nishina T, Yamaguchi K, et al. A multicenter phase II study of TAS-102 monotherapy in patients with pre-treated advanced gastric cancer (EPOC1201). European Journal of Cancer. 2016;62:46-53.
50. Ott P, Calvo E, Sharma P, Jaeger D, Ascierto P, Chau I, et al. Nivolumab monotherapy in patients with advanced gastric or gastroesophageal junction (GEJ) cancer and 2 or more prior treatment regimens: Sub-analysis of the CheckMate 032 study. . Ann Oncol. 2017;28(iii):139.
51. Fuchs CS, Denker AE, Tabernero J, Cutsem EV, Ohtsu A, Lam B, et al. Multicohort phase II KEYNOTE-059 study of pembrolizumab (MK-3475) for recurrent or metastatic gastric or gastroesophageal junction (GEJ) adenocarcinoma. Journal of Clinical Oncology. 2015;33(15_suppl):TPS4135-TPS.
52. Doi T, Muro K, Boku N, Yamada Y, Nishina T, Takiuchi H, et al. Multicenter phase II study of everolimus in patients with previously treated metastatic gastric cancer. J Clin Oncol. 2010;28(11):1904-10.
53. Ruan H, Dong J, Zhou X, Xiong J, Wang H, Zhong X, et al. Multicenter phase II study of apatinib treatment for metastatic gastric cancer after failure of second-line chemotherapy. Oncotarget. 2017;8(61):104552-9.
54. Kelly RJ, Lee J, Bang Y-J, Almhanna K, Murphy MAB, Catenacci DVT, et al. Safety and efficacy of durvalumab in combination with tremelimumab, durvalumab monotherapy, and tremelimumab monotherapy in patients with advanced gastric cancer. Journal of Clinical Oncology. 2018;36(15_suppl):4031-.
55. Li J, Wen L, Wang J, Deng W, Zhang G, Jia T, et al. Safety of apatinib as third-line or beyond treatment in advanced or metastatic gastric cancer: Results from a multicenter phase IV study (Ahead-G201). Journal of Clinical Oncology. 2018;36(15_suppl):e16019-e.
56. Schmalenberg H, Al-Batran S-E, Pauligk C, Zander T, Reichart A, Lindig U, et al. CabaGast: multicentre, Phase II study with cabazitaxel in previously treated patients with advanced or metastatic adenocarcinoma of the esophagogastric junction and stomach2017.
57. Nishimura T, Iwasa S, Nagashima K, Okita N, Takashima A, Honma Y, et al. Irinotecan monotherapy as third line treatment for advanced gastric cancer patients refractory fluoropyrimidine, platinum and taxanes harboring UGT1A1*1/*1. *1/*6 or *1/*28. Gastric Cancer. 2016;20:655-62.
58. Zhang Z, Chen Z, Zhu X, Zhang X, Zhao X, Zhang W, et al. A phase II study of irinotecan as single agent in the third-line treatment of unresectable or metastatic gastric cancer. Journal of Clinical Oncology. 2018;36(15_suppl):e16084-e.
59. Michael M, Bang YJ, Park YS, Kang YK, Kim TM, Hamid O, et al. A Phase 1 Study of LY2874455, an Oral Selective pan-FGFR Inhibitor, in Patients with Advanced Cancer. Targeted oncology. 2017;12(4):463-74.

60. Im CK, Rha SY, Jeung HC, Jeong J, Lee SH, Noh SH, et al. A Phase II Feasibility Study of Weekly Paclitaxel in Heavily Pretreated Advanced Gastric Cancer Patients with Poor Performance Status. Oncology. 2009;77(6):349-57.
61. Iwasa S, Shitara K, Takahashi S, Park H, Kadowaki S, Modi S, et al. Updated results of phase 1 study of DS-8201a in subjects with HER2-expressing gastric cancer. Journal of Clinical Oncology. 2018;36(4_suppl):118-.
62. Fuchs CS, Doi T, Jang RW, Muro K, Satoh T, Machado M, et al. Safety and Efficacy of Pembrolizumab Monotherapy in Patients With Previously Treated Advanced Gastric and Gastroesophageal Junction Cancer: Phase 2 Clinical KEYNOTE-059 Trial. JAMA Oncol. 2018;4(5):e180013.
63. Sugimoto N, Yoshinami T, Yagi T, Hasegawa A, Fujisawa F, Imamura F. 267P Irinotecan monotherapy as third line treatment for advanced gastric cancer patients refractory fluoropyrimidine, platinum and taxanes harboring UGT1A1*1/*1. *1/*6 or *1/*28. Annals of Oncology. 2016;27(suppl_9).
64. Jeong J, Jeung H-C, Rha SY, Im CK, Shin SJ, Ahn JB, et al. Phase II study of combination chemotherapy of 5-fluorouracil, low-dose leucovorin, and oxaliplatin (FLOX regimen) in pretreated advanced gastric cancer. Annals of Oncology. 2008;19(6):1135-40.
65. Sakai D, Boku N, Kodera Y, Komatsu Y, Fujii M, Iwasa S, et al. An intergroup phase III trial of ramucirumab plus irinotecan in third or more line beyond progression after ramucirumab for advanced gastric cancer (RINDBeRG trial). Journal of Clinical Oncology. 2018;36(15_suppl):TPS4138-TPS.
66. European Medicines Agency. Nivolumab 2018.
67. Ilson DH, Tabernero J, Doi T. Protocol No. TO-TAS-102-302 Metastatic Gastric Cancer. 2015.
68. Shitara K, Doi T, Dvorkin M, Mansoor W, Arkenau H-T, Prokharau A, et al. Supplementary appendix. The Lancet Oncology. 2018;19(11):1437-48.
69. Alsina M, Tabernero J, Shitara K, Doi T, Dvorkin M, Mansoor W, et al. Analysis of symptoms and functional HRQoL scales in TAGS, a phase III trial of trifluridine/tipiracil (FTD/TPI) in metastatic gastric cancer (mGC). Journal of Clinical Oncology. 2019;37(15_suppl):4043-.
70. Higgins JPT SJ, Savović J, Page MJ, Hróbjartsson A, Boutron I, Reeves B, Eldridge S. . A revised tool for assessing risk of bias in randomized trials In: Chandler J, McKenzie J, Boutron I, Welch V (editors). Cochrane Methods. Cochrane Database of Systematic Reviews 2016, Issue 10 (Suppl 1). dx.doi.org/10.1002/14651858.CD201601.
71. Arkenau T. ESMO TAGS Presentation. 2018.
72. M V, MH P, S S, D C, N M, T C, et al. National Oesophago-Gastric Cancer Audit [Internet]. 2019 [cited 2019 Jun 17]. Available from: https://www.nogca.org.uk/content/uploads/2019/04/NOGCA-2018-Annual-Report-V2.pdf.
73. Zaanan A, Bouché O, Benhaim L, Buecher B, Chapelle N, Dubreuil O, et al. Gastric cancer: French intergroup clinical practice guidelines for diagnosis, treatments and follow-up (SNFGE, FFCD, GERCOR, UNICANCER, SFCD, SFED, SFRO). Digestive and Liver Disease. 2018;50(8):768-79.
74. Muro K, Doi T, Bando H, Yasui H, Nishina T, Yamaguchi K, et al. 2621: Multicenter phase 2 study of TAS-102 monotherapy in patients with pretreated advanced gastric cancer. 2014.

75. Japanese Gastric Cancer A. Japanese gastric cancer treatment guidelines 2014 (ver. 4). Gastric Cancer. 2017;20(1):1-19.
76. Food and Drug Administration. FDA grants accelerated approval to pembrolizumab for advanced gastric cancer 2017 [Available from: https://www.fda.gov/drugs/resources-information-approved-drugs/fda-grants-accelerated-approval-pembrolizumab-advanced-gastric-cancer.
77. Jou E, Rajdev L. Current and emerging therapies in unresectable and recurrent gastric cancer. World J Gastroenterol. 2016;22(20):4812-23.
78. Al-Batran S-E, Riera-Knorrenschild J, Pauligk C, Goetze TO, Hegewisch-Becker S, Seraphin J, et al. A randomized, double-blind, multicenter phase III study evaluating paclitaxel with and without RAD001 in patients with gastric cancer who have progressed after therapy with a fluoropyrimidine/platinum-containing regimen (RADPAC). Journal of Clinical Oncology. 2017;35(4_suppl):4-.
79. Bang Y-J, Xu R-H, Chin K, Lee K-W, Park SH, Rha SY, et al. Olaparib in combination with paclitaxel in patients with advanced gastric cancer who have progressed following first-line therapy (GOLD): a double-blind, randomised, placebo-controlled, phase 3 trial. The Lancet Oncology. 2017;18(12):1637-51.
80. Kang Y-K, Boku N, Kang WK, Yoon HH, Cascinu S, Al-Batran S-E, et al. A prospective, randomized, double-blinded, placebo-controlled, phase III study to evaluate the efficacy and safety of apatinib plus best supportive care (BSC) compared to placebo plus BSC in patients with advanced or metastatic gastric cancer: The ANGEL study. Journal of Clinical Oncology. 2017;35(15_suppl):TPS4138-TPS.
81. L B. End of life decisions at NICE: Taking stock A report for Servier. 2019.
82. NICE. Trifluridine with tipiracil hydrochloride for treating metastatic colorectal cancer after standard therapy [ID876]
2016 [Available from: https://www.nice.org.uk/guidance/ta405/documents/committee-papers.
83. NICE. Guidance: Trastuzumab for the treatment of HER2-positive metastatic gastric cancer 2010 [Available from: https://www.nice.org.uk/guidance/ta208.
84. Drzal R, Szmurlo D, Plisko R. Can we determine the optimal cycle length for which half-cycle correction should always be applied? Value in Health. 2013;16(3).
85. N L. NICE DSU technical support document 14: Survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data.
86. Office of National Statistics. Dataset: National life tables: UK [Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/nationallifetablesunitedkingdomreferencetables.
87. Bullement A, Latimer NR, Bell Gorrod H. Survival Extrapolation in Cancer Immunotherapy: A Validation-Based Case Study. Value in Health. 2019;22(3):276-83.
88. Andrew Davies, Andrew Brigg, John Schneider, Levy A, Omar Ebeid, Samuel Wagner, et al. The Ends Justify the Mean: Outcome Measures for Estimating the Value of New Cancer Therapies. Health Outcomes Research in Medicine. 2012;3(1):e25-e36.
89. Morgan A, Sutton A, Wailoo A. Report by the Nice Decision Support Unit: The risk and costs of febrile neutropenia in patients with non small cell lung cancer treated with docetaxel 2007 [Available from: http://nicedsu.org.uk/wp-content/uploads/2016/03/Erlotinib-DSU-final-report1.pdf.
90. Life EQo. Quality of life of cancer patients [Available from: https://qol.eortc.org/questionnaire/eortc-qlq-c30/

91. Kontodimopoulos N, Aletras VH, Paliouras D, Niakas D. Mapping the cancer-specific EORTC QLQ-C30 to the preference-based EQ-5D, SF-6D, and 15D instruments. Value Health. 2009;12(8):1151-7.
92. Chau I, Fuchs CS, Ohtsu A, Liepa AM, Cui ZL, Hsu Y, et al. Meaningful changes in quality of life (QoL) in patients with gastric cancer: Exploratory analyses from RAINBOW and REGARD. Journal of Clinical Oncology. 2017;35(15_suppl):4047-.
93. Health Economics Research Centre. HERC Database of mapping studies [Available from: http://www.herc.ox.ac.uk/downloads/herc-database-of-mapping-studies.
94. K Muro, E Calvo, G Maglinte, A Rider, M Contente, C Middleton, et al. Real-world productivity, healthcare resource utilization (HRU), and quality of life (QOL) in patients with advanced gastric cancer (GC) in Korea and Japan. Annals of Oncology. 2017;28(suppl_10).
95. James Pike GAM, Alex Rider, Kim Cocks, Fiona Taylor, Marta Contente, Emiliano Calvo. Predicting EQ-5D Utility Index Scores for Gastric Cancer Patients in Japan: A Preliminary Model; Presented at ISPOR 20 th Annual European Congress; 4-8 November, 2017; Glasgow, Scotland.
96. Shukui Qin JL, Li Yang. Poster (PCN129) presented at: ISPOR 22nd Annual International Congress; Cost-effectiveness Analysis of Apatinib Compared with Best Supportive Care in Chemotherapy-refractory Advanced or Metastatic Gastric Cancer in China; May 22, 2017; Boston, MA, USA.
97. Hong-Dou Chen, Jing Zhou, Feng Wen, Peng-Fei Zhang, Ke-Xun Zhou, Han-Rui Zheng, et al. Cost-effectiveness analysis of apatinib treatment for chemotherapy-refractory advanced gastric cancer. Journal of Cancer Research and Clinical Oncology. 2017;143(2):361-8.
98. Bai Y, Xu Y, Wu B. Cost-Effectiveness and Budget Impact Analysis of Apatinib for Advanced Metastatic Gastric Cancer from the Perspective of Health Insurance System. Gastroenterol Res Pract. 2017;2017:2816737.
99. Hatswell AJ, Porter J, Lee D, Hertel N, Latimer NR. The Cost of Costing Treatments Incorrectly: Errors in the Application of Drug Prices in Economic Evaluation Due to Failing to Account for the Distribution of Patient Weight. Value in Health. 2016;19(8):1055-8.
100. Curtis L aB, A. Unit Costs of Health and Social Care 2018. Project report. University of Kent https://doi.org/10.22024/UniKent/01.02.70995 <https://doi.org/10.22024/UniKe 2018 [Available from: https://kar.kent.ac.uk/70995/1/Unit%20Costs%202018%20-%20FINAL%20with%20bookmarks%20and%20covers%20%282%29.pdf.
101. Liepa A BJ, Bapat B, Kaye J. ASCO Poster 184: Real-world treatment patterns of previously treated advanced gastric and gastroesophageal junction adenocarcinoma (GC) in the United Kingdom (UK). Gastrointenstinal Cancers Symposium, San Francisco, CA Jan 15-17, 2015.
102. 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.
103. NICE. Technology appraisal guidance [TA405]: Trifluridine–tipiracil for previously treated metastatic colorectal cancer [updated August 2016. Available from: https://www.nice.org.uk/guidance/ta405.
104. NICE. Technology appraisal guidance [TA517]: Avelumab for treating metastatic Merkel cell carcinoma [updated 11 April 2018. Available from: https://www.nice.org.uk/guidance/ta517.

105. NICE. Technology appraisal guidance [TA541]: Inotuzumab ozogamicin for treating relapsed or refractory B-cell acute lymphoblastic leukaemia [Available from: https://www.nice.org.uk/guidance/ta541.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Trifluridine–tipiracil for treating metastatic gastric or gastro-oesophageal junction cancer after 2 or more therapies [ID1507]

Clarification questions

July 2019

File name	Version	Contains confidential information	Date
	v1.0	Yes	25 July 2019

Section A: Clarification on effectiveness data

A1. Priority. In CSR Table 14, it is reported that 11.6% of the trifluridine–tipiracil (TFT) treatment group received the wrong dose or wrong treatment, but only 2.4% of the placebo group did. Please clarify how many patients received a higher dose (and how much higher), a lower dose (and how much lower) or an incorrect treatment (including the name of the treatment they received), for each treatment group, and how this has been accounted for in the analyses and cost effectiveness inputs? It appears that only major protocol deviations (which did not include dosing issues according to section CSR 10.3) were adjusted for in section 11.4.1.2 of the CSR.

There were 43 patients in total (39 (11.6%)[*] in the trifluridine/tipiracil treatment group and 4 (2.4%) in the placebo group) who received the wrong treatment or incorrect dose during the course of the TAGS trial.

The reasons for this may be categorised as the following:

1. The dose of study medication was not withheld even though the absolute neutrophil count (ANC) was below 1.5 x 10^9/L prior to dosing. (Per the protocol for the TAGS trial, the dose should have been withheld until neutrophil levels are above 1.5 x 10^9/L.)
    - This occurred for 29 patients in the trifluridine/tipiracil group and 2 patients in the placebo group
2. The patient took the incorrect dose
    - The incorrect dose was given to 6 patients in the trifluridine/tipiracil group and 2 patients in the placebo group, details of which can be seen in Table 1.
3. The kit (pack[s]) dispensed were incorrect
    - The incorrect kit being dispensed was given to 2 patients in the trifluridine/tipiracil group

[*] Please note that 2 patients in the trifluridine/tipiracil group were counted twice as they were both assigned to 2 separate categories (hence sum of trifluridine/tipiracil patients that received the wrong treatment equals 41).

4. There was an error in the calculation and/or recording of body surface area (or a contributing metric – i.e. height or weight)

    - 4 patients in the r trifluridine/tipiracil group

Table 1: Details of patients given incorrect dose during the TAGS trial

Incorrect dose of trifluridine/tipiracil given (6 patients)	Incorrect dose of placebo given (2 patients)
• Patient accidently given 145mg instead of 150mg total daily dose on cycle 1, day 1 (this was due to the wrong kit being assigned) • Patient decided to reduce dose from 110mg to 90mg on cycle 2, day 29 and to reduce dose from 130mg to 100mg on cycle 1, days 9-12 (self-decision due to nausea) • Patient accidentally only took 1 tablet (15mg) on cycle 1, day 1 • Patient took half required dose (35mg BD PO instead of 70mg BD PO) on cycle 2, days 1-5 • Patient took 2 instead of 3 pills per dose for cycle 1 • Patient took 5 more days of drug administration in cycle 2 (this was due to the wrong kit being assigned)	• Patient took 2 out of the 3 tablets in the morning dose in error (40mg instead of 60mg) • Patient lost a tablet of 15mg and took only 2 tablets of 20mg for one morning

Within the economic model, no explicit adjustments to the underlying efficacy data were made in light of dosing adjustments. However, as detailed in Document B, adjustments to the dose of trifluridine/tipiracil were made according to dose delays and dose reductions. For an explanation of these adjustments, please refer to Section B.3.5 of Document B.

A2. Priority. Please clarify what other cancer treatments were administered in each treatment group and to how many patients, and how this has been accounted for in the analyses.

It is assumed that this question is in reference to non-study anti-tumour treatment after the treatment period in the TAGS trial. Please find attached to this response the corresponding table from the CSR detailing the cancer treatments administered (stratified by treatment group), and the corresponding number of patients. To open the table, please double click the icon below:

No specific adjustments were made to the clinical efficacy data regarding post-progression treatment. Other than the regimen(s) given, limited data are available regarding the treatments administered (e.g. dose or duration). In addition, some of the treatments administered were experimental/unlicensed (e.g. nivolumab and apatinib). Post-progression treatment is similar across both arms, although a larger proportion of placebo patients received later lines of systemic anticancer therapy (26.5% versus 24.6%).

A3. Priority. Please provide the efficacy of TFT for all clinical outcomes, for patients who were from the EU and had no prior ramucirumab. Please clarify the baseline characteristics of these patients by treatment arm, and account for any imbalances between treatment arms in your analyses.

Provided below is a summary of clinical outcomes for trifluridine/tipiracil for European patients with no prior ramucirumab. We appreciate why the ERG have requested this post-hoc analysis, however, Servier urges extreme caution when interpreting the results for this subgroup, which is not based on stratification within the TAGS trial (patients were stratified based on region [Japan versus rest-of-the-world], ECOG performance status [0 or 1], and prior ramucirumab [yes or no]). Servier has provided these results for completeness, but does not consider this group appropriate to inform decision making. The removal of all non-European (i.e. Asian and American) patients is inconsistent with previous gastric cancer appraisals conducted by NICE in which these patients were not excluded within the estimation of efficacy outcomes – for example, in the ToGA study of capecitabine/5-FU versus capecitabine/5-FU+traztuzumab, ~55% of patients were Asian and ~9% were Central/South American.[1]

Patient numbers

Within the TAGS trial, n=169 patients were previously treated with ramucirumab, and n=338 were not. Of the n=338 patients with no previous ramucirumab treatment, [redacted] were European. Of these [redacted] patients, [redacted] were randomised to receive trifluridine/tipiracil, with the remaining [redacted] assigned to placebo. By excluding non-European patients from the rest-of-the-world, no prior ramucirumab population, [redacted] United States patients are excluded from consideration. For a full breakdown of patient numbers by region, prior ramucirumab treatment, and treatment assignment within the TAGS trial, please see the response to clarification question A22.

Baseline characteristics

Baseline patient characteristics for the European, no prior ramucirumab population are provided in Table 2, compared with those for the intention-to-treat (ITT) population. Other than the inherent differences in patients by region, there is one difference introduced by excluding all patients with previous ramucirumab treatment and non-European patients. There is a small shift in the distribution of the number of prior regimens towards the lower end – that is, a larger proportion with 2 prior regimens, and a smaller proportion with ≥4. This may be explained by the fact that by definition, patients with a history of treatment with ramucirumab are expected to have received a greater number of regimens.

Table 2: Baseline characteristics for the European, no prior ramucirumab population

	ITT		European, no prior RAM	
	T/T	PBO	T/T	PBO
Number of patients	337	170	203	109
Age (years)				
Median (IQR)	64.0 (24–89)	62.5 (32–82)	[redacted]	
<65	183 (54%)	96 (56%)		
≥65	154 (46%)	74 (44%)		
Sex				
Male	252 (75%)	117 (69%)		
Female	85 (25%)	53 (31%)		
Ethnicity				
White	244 (72%)	113 (66%)		
Asian	51 (15%)	29 (17%)		
Other or not available	4 (1%)	4 (2%)		
Not available	38 (11%)	24 (14%)		
Region				
USA	21 (6%)	5 (3%)	-	-
Europe*	270 (80%)	138 (81%)	203 (100%)	109 (100%)
Japan	46 (14%)	27 (16%)	-	-
ECOG PS				
0	123 (36%)	68 (40%)	[redacted]	
1	214 (64%)	102 (60%)		
Primary site				

Gastric	239 (71%)	121 (71%)	[redacted]	
GEJ	98 (29%)	47 (28%)		
Both	0	2 (1)		
Measurable disease	306 (91%)	150 (88%)		
Histology				
Diffused	53 (16%)	21 (12%)		
Intestinal	103 (31%)	52 (31%)		
Mixed	14 (4%)	8 (5%)		
Unknown	132 (39%)	69 (41%)		
Not available	35 (10%)	20 (12%)		
HER2 status				
Positive	67 (20%)	27 (16%)		
Negative	207 (61%)	106 (62%)		
Not assessed	62 (18%)	37 (22%)		
No. of metastatic sites				
1–2	155 (46%)	72 (42%)		
≥3	182 (54%)	98 (58%)		
Peritoneal metastases	87 (26%)	53 (31%)		
Previous gastrectomy	147 (44%)	74 (44%)		
No. of prior regimens				
2	126 (37%)	64 (38%)		
3	134 (40%)	60 (35%)		
≥4	77 (23%)	46 (27%)		
Prior systemic cancer therapeutic agents				
Platinum	337 (100%)	170 (100%)	203 (100%)	109 (100%)
Fluoropyrimidine	336 (>99%)	170 (100%)	203 (100%)	109 (100%)
Taxane	311 (92%)	148 (87%)	[redacted]	
Irinotecan	183 (54%)	98 (58%)		
Ramucirumab	114 (34%)	55 (32%)		
Anti-HER2 therapy	60 (18%)	24 (14%)	*	*
Immunotherapy (anti–PD-1/PD-L1)	25 (7%)	7 (4%)	9 (4%)	2 (2%)
Other	77 (23%)	41 (24%)	*	*

[redacted]

The request from the ERG suggested to account for any imbalances between treatment arms in the analyses produced. The largest imbalances (based on within-category difference in percentages between arms) are sex ([redacted] peritoneal metastases [redacted] HER-2 positivity ([redacted] %), and the proportion of patients with two prior regimens ([redacted] The majority of these imbalances are similar to that of the ITT population (e.g. male 75% versus 69%; δ=6%). Based on a review of the differences in patient characteristics between the treatment arms, Servier does not believe there to be any major imbalances in prognostic or predictive factors between treatment arms and so no adjustments have been made.

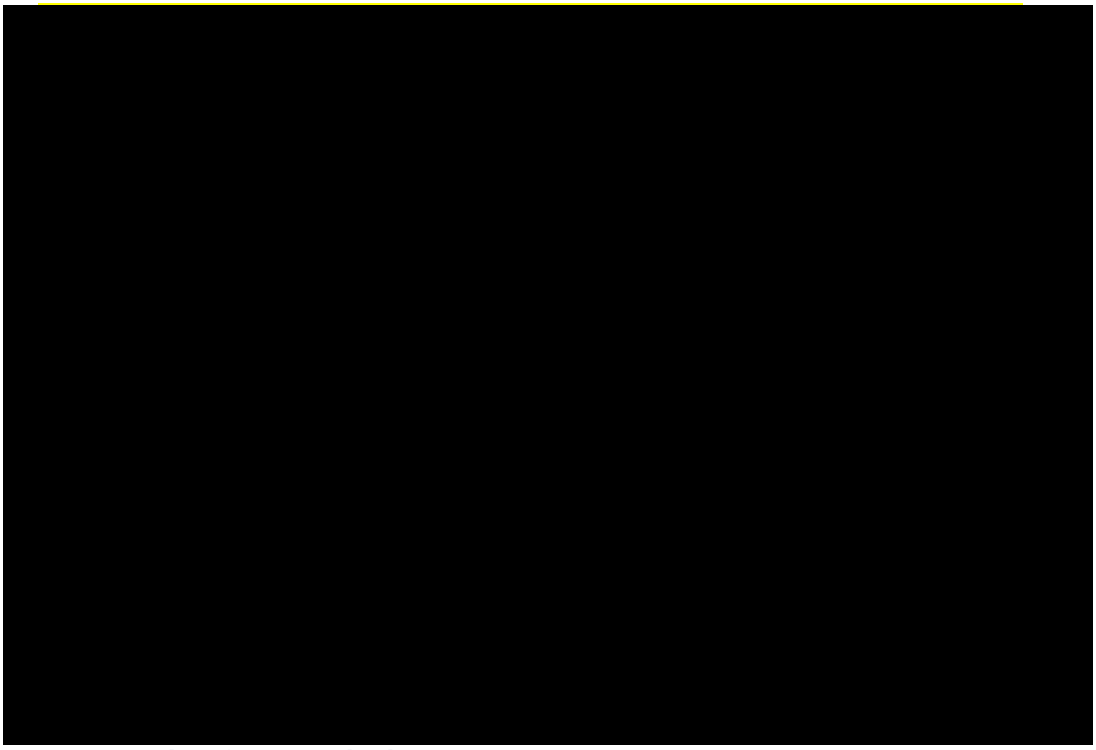
Clinical efficacy outcomes

The following clinical efficacy outcomes have been provided for the European, no prior ramucirumab population specifically: (1) overall survival (OS); (2) progression-free survival (PFS); (3) best overall response (BoR); and (4) duration of response (DoR).

Overall survival

The OS Kaplan-Meier curves for the European no prior ramucirumab population are provided in Figure 1. [redacted]

Figure 1: Overall survival for the European, no prior ramucirumab population



Progression-free survival

The PFS Kaplan-Meier curves for the European no prior ramucirumab population are provided in Figure 2. As per the OS analysis, the curves are very similar to the no prior ramucirumab population. However, the small number of patients at risk at 9 months (n=10 versus n=2) should be noted.

Figure 2: Progression-free survival for the European, no prior ramucirumab population

Best overall response

In the ITT population, the disease control rate (DCR, i.e. proportion of patients with complete response [CR], partial response [PR], or stable disease [SD]) was 44.1% in the trifluridine/tipiracil group versus 14.5% in the placebo group (Section B.2.6.3). For the European, no prior ramucirumab population, the DCR was [redacted] in the trifluridine/tipiracil group versus [redacted] in the placebo group.

The best overall response for European, no prior ramucirumab patients is provided in Table 3.

Table 3: Best overall response for the European, no prior ramucirumab population

	TAS-102	Placebo	Not treated	Total
CR				
PR				
SD				
PD				
NE				
Total				

Key: CR, complete response; NE, not evaluable; PD, progressive disease; PR, partial response; SD, stable disease.

Duration of response (DoR)

The DoR was recorded for patients achieving either CR or PR. As shown in Table 3, [redacted] trifluridine/tipiracil patients and [redacted] placebo patients achieved CR or PR. A Kaplan-Meier plot of DoR is presented in Figure 3. From this figure, relatively little may be inferred concerning the difference in the DoR between treatment arms, as only [redacted] placebo patients achieved a PR.

Figure 3: Duration of response for the European, no prior ramucirumab population

A4. Priority. Please clarify what software and package were used to perform generalised estimating equation (GEE) regression analysis. Please also provide the code use for GEE regression analysis.

The analysis was performed in the statistical software *R*, using the package *geepack*.[2] The code used for the analysis is provided below for the selected model. Other models follow the same format (and settings), but use different explanatory variables. The "-1" term is used to suppress the intercept term.

Table 4: Code used to run regression analyses

```
Utility modelled by progression:
Utility_prog <- geeglm(Utility ~ prog_state -1, data = qol_data, id =
USUBJID, corstr = "independence", waves = date)
```

A5. (Appendices – all reviews). Please clarify whether the MEDLINE and EMBASE searches (via Embase.com) were conducted simultaneously using a single strategy. If so clarify whether subject headings included are based on the controlled vocabulary of the former (MeSH) or the latter source (Emtree).

EMBASE and MEDLINE were searched using a single search strategy run via the Embase.com interface. Based on publications such as the white paper entitled *"A Comparison of Emtree and MeSH".*[3] Emtree was considered a more comprehensive and up-to-date thesaurus for biomedical research. As a result, subject headings in EMBASE were chosen by browsing the Emtree terms.

As the MEDLINE In-process search was run via the PubMed.com interface, MeSH controlled vocabulary was utilised.

A6. The ERG notes some errors in Servier's update searches covering the period 27th June 2018 – 28th February 2019. The Cochrane Library search (Appendix D, p15, unnumbered table) does not include a line 10; meanwhile the Medline In-process (PubMed) search (Appendix D, p17, unnumbered table) features two. Given that these errors could only have been caused by manual editing, clarify to what extent can the ERG be assured that searches were conducted as they are reported.

Servier acknowledge the errors highlighted by the ERG, on two of the tables in Appendix D (p15 and p17), and are appreciative of the concern raised. Having checked the original source documents used to compile Appendix D, the "*search number*" is free of such errors. Most importantly, the "*search query*" content for each respective table remains consistent and as intended. Please be assured the errors observed were caused by an attempt to aesthetically format the tables, and the searches executed were as they have been reported.

A7. Appendix D1.1.2 states that Servier's update searches in 2019 included DARE and HTAD, however the ERG is aware that these sources were removed from the Cochrane Library in August 2018 (https://www.cochranelibrary.com/about/CRD-database-info). Clarify the method used to search these databases.

The original systematic literature review (SLR) was conducted in June 2018. At this time, DARE and HTAD databases were available via Cochrane interface. At the time the SLR was updated in 2019, these two databases were searched via the Centre for Reviews and Dissemination (CRD) website. Although the databases are no longer updated, while they remain accessible, they were included in the supplementary search strategy for consistency.

A8. Please clarify the reason behind the cut-off date of 2008 for the cost effectiveness review.

The "Developing NICE Guideline manual" states that "*Inclusion criteria for sifting and selecting papers for each review should specify populations and interventions relevant to the review question. They should also specify: An appropriate date range, because older studies may reflect outdated practices".*

The cost-effectiveness literature search for this appraisal was restricted to the last 10 years because of the considerable change observed over this period in the following:

- costs and resource use
- advancement of technology (drug therapy, diagnostics etc.)
- quality/standards of care
- overall standards of living, and inflation

Based on this changing landscape, cost-effectiveness studies published after 2008 were considered most relevant to the decision problem, particularly as the earliest-published study identified within the clinical efficacy search was also published in 2008.[4]

Finally, while it is acknowledged that other potentially-relevant studies may have been published prior to 2008, there is no established standard of care for patients in the third-line and beyond setting, and so any identified publications were expected to have limited applicability to the decision problem. This is evident within the search conducted, as all four studies identified were based on the Chinese setting, and evaluated interventions that are not licensed within Europe.

A9. Please indicate the source of the search filter used to identify economic evaluations, providing a citation to any relevant validation studies.

The study design filter for the economic evaluation was developed based on search terms on the Scottish Intercollegiate Guidelines Network (SIGN) website. The searches were run using the EBSCO platform.

A10. Appendices G.1.4.5, Table 8, p12: EconLit search. The "search options" column contains the repeated statement "Find all my search terms". This does not seem to align with the "query" column which uses the OR operator (which logically means find ANY of these terms). Please explain this apparent contradiction.

As noted, the use of the boolean operator "OR" will retrieve all the studies that were retrieved via earlier search terms. When the electronic search is run and saved using the EBSCO

platform, the database reflects the searches wherein "*Search modes - Find all my search terms*" is mentioned against the rows that are using boolean operators as well (for instance S4, S8, S9, S11 and S12). With an objective to report searches in a transparent manner, we adopted the reporting style as mentioned on the EBSCO platform. To open the screenshot of the exemplary electronic search for your referral (run on 17 July 2019), please double click the icon below:

Screenshot of exemplary search

A11. p17. For administration of the drug, please clarify how "benefit" is defined. For example, should this be clinically or radiologically assessed? Which clinical staff judge benefit? How often is this assessed? Please indicate how benefit was defined in the trial and how generalisable this is to clinical practice in England.

During the TAGS trial, the risk vs. benefit of trifluridine/tipiracil was evaluated both clinically and radiologically. It was deemed not favourable for the following (protocol defined discontinuation) reasons:

- Patient request at any time irrespective of the reason
- Disease progression - defined by Response Evaluation Criteria in Solid Tumours (RECIST- (Version 1.1, 2009).
- Clinical progression
- Patient experiences an irreversible, treatment-related, Grade 4, clinically relevant, non-haematological event
- Unacceptable adverse events (AEs), or change in underlying condition such that the patient can no longer tolerate therapy, including:
    - A maximum dose delay >28 days from the scheduled start date of the next cycle of trifluridine/tipiracil
    - Need for more than 3 dose reductions of trifluridine/tipiracil
- Physician's decision including need for other anticancer therapy not specified in the protocol or surgery or radiotherapy to the only site(s) of disease being evaluated in protocol
- Pregnancy.

Ultimately, the treating physician, guided by the study protocol, cross-functional medical team and patient's wishes was responsible for treatment decisions.

Patients were radiologically assessed with computed tomography (CT) scans performed at baseline and then every 8 weeks thereafter until disease progression. On-site tumour assessments were performed by the Investigator/local radiologist. At a UK advisory board, it was discussed whether a CT scan every 8 weeks would align with UK practice. The general consensus was that this was consistent with what the 12 clinicians attending the advisory board would expect. However 1 clinician stated that they would perform a CT scan 4 weeks after initially commencing trifluridine/tipiracil to ensure they could stop treatment as early as possible if the patient was progressing, but that they would then have a CT scan every 8 weeks. Servier would therefore consider the assessments in the TAGs trial to be generalisable to the UK population.

A12. p34. Please clarify why duration of response not been reported? Please provide the duration of response data, some of which is available in the CSR (p95), but we do not have access to Table 14.2.9.

We note that duration of response was in the scope, and this was excluded as an oversight. Due to the small patient numbers, especially the placebo group (n=3), this wasn't considered meaningful. Please see Figure 4 for further information.

[redacted]

A13. p50. The answer to the risk of blinding (RoB) question about allocation concealment is answered with respect to blinding, not to concealment of the allocation before enrolment and randomisation. Please clarify if the personnel who were enrolling patients were aware of the group to which patients were assigned.

At enrolment and randomisation, a central interactive voice / web response system was used to assign kit-numbers for study medication. The study medications and all packaging were identical in appearance save for the kit numbers.

Personnel were not aware of the group to which study patients were assigned. As per the study protocol, treatment assignment was concealed, unknown to all patients, investigators, ancillary study personnel at the site and to employees of Taiho Oncology Inc. or Taiho Pharmaceutical Co Ltd.

A14. p51. Please clarify why the NICE quality assessment tool was not used, as per the submission template? Please clarify your scores by providing supporting evidence for the items without supporting evidence in Table 11 (i.e. were there any unexpected imbalances in

dropouts between groups?; Is there any evidence to suggest that the authors measured more outcomes than they reported?; Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?)

Table 11 (Quality assessment) in document B (page 51) is completed as per the NICE submission template, although some answers were not expanded upon. In addition, please note that it is not comparing multiple parallel group randomised controlled trials.

For further detail requested on answers on Table 11 (page 51, document B):

1. Were there any unexpected imbalances in drop-outs between groups?

    - Servier answered 'no' to this. To expand on this assertion please refer to the consort diagram in Figure 5, that demonstrates discontinuation rates between the two groups. The majority of patients who discontinued, did so due to progressive disease. In the trifluridine/tipiracil group, nine of the 11 deaths resulting in treatment discontinuation were attributed to disease progression (the cause of the other two deaths was septic shock that was judged to be unrelated to treatment).[5]

2. Is there any evidence to suggest that the authors measured more outcomes than they reported?

    - Servier answered 'no' to this. We would support this assertion by referring to the schedule of assessment provided in the protocol and table 2 of the clinical study report (CSR). This lists out all assessments and their timepoints that are expected to take place. There are no outcomes that are not accounted for in the clinical study report.

3. Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?

    - Servier answered 'yes' to this. The analysis populations were as follows:

	Number (%) of patients		
	Trifluridine/ tipiracil n (%)	**Placebo n (%)**	**Total n (%)**

Intention-to-treat (ITT)	337 (100)	170 (100)	507 (100)
As-treated (AT)	335 (99.4)	168 (98.8)	503 (99.2)
Tumour response (TR)	290 (86.1)	145 (85.3)	435 (85.8)

No missing data were estimated for efficacy variables with the exception of imputation of dates for partial dates of death or clinical progression in cases where only the day was missing; dates with missing month or year were not imputed

A15. Within Appendix D1.3 of the company submission, the Cochrane RoB2 tool has been used, but some items have not been scored. These relate to risk of bias due to deviations from the intended intervention. We have asked for clarification on some of these points in clarification question A1, but it would be helpful if you could clarify how these RoB items would be scored.

In the Cochrane RoB2 tool, where Domain 2 questions 2.1/2.2 are "No" (or even "Probably No"), subsequent questions 2.3, 2.4 and 2.5 are not applicable or accessible [greyed out in the tool], so they display unscored in the file exported from the tool.

These questions are as follows:

- **2.1** Were participants aware of their assigned intervention during the trial?
    - Servier answered 'No'
- **2.2** Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?
    - Servier answered 'No'
- **2.3** If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the experimental context?

- This question was not answered as it was inapplicable/not accessible (greyed out) due to previous answers given

- [redacted]

- **2.4** If Y/PY to 2.3: Were these deviations from intended intervention balanced between groups?

  - This question was not answered as it was inapplicable/not accessible (greyed out) due to previous answers given

- [redacted]

- **2.5** If N/PN/NI to 2.4: Were these deviations likely to have affected the outcome?

  - This question was not answered as it was inapplicable/not accessible (greyed out) due to previous answers given

- [redacted]

In terms of the overall result of the Cochran RoB tool, as the blinding and concealment were adequate, 2.3-2.5 do not change the outcome, which remains a “Low” risk of bias

A16. p51. Please clarify the criteria used to consider patient characteristics to be balanced. How may the following differences in characteristics have affected study results? Sex (75% vs 69%); Ethnicity (white 72% vs 66%); ECOG performance status 0 (36% vs 40%); Histology

(diffused 16% vs 12%); HER2 status 20% vs 16%; Number of metastatic sites (1-2 sites, 46% vs 42%); Number of prior chemotherapy regimens (⩾4, 23% vs 27%).

Despite these minor differences which are evident within the baseline characteristics in the TAGS trial, the investigators stated that the *"baseline demographic and disease characteristics were generally balanced between the two treatment groups"*.[5] This was validated at a UK advisory board by 12 UK clinicians.

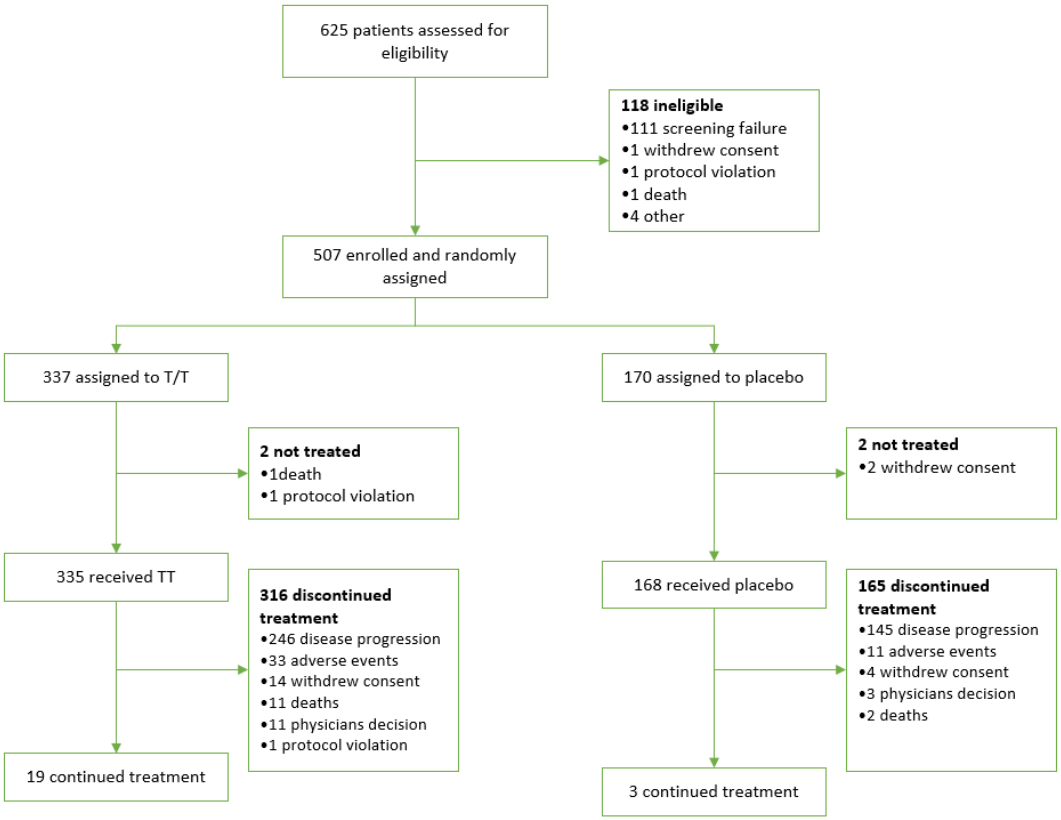
Regarding the effect these differences might have, it is worth noting how they might counteract each other, with both neutral, positive and negative effects. The placebo group was favoured by an increased number of patients with an ECOG PS 0 (36% vs 40%); number of metastatic sites 1-2 (rather than ≥3) (46% vs 42%); HER 2 positive status (20% vs 16%)[6] and also potentially histology (diffused 16% vs 12%).[7] The trifluridine/tipiracil group was favoured by the proportion of patients with ⩾4 prior chemotherapy regimens (23% vs 27%) and potentially gender (75% vs 69% male). The impact that the differences seen in ethnicity (72% vs 66% white) are difficult to estimate as there is a slightly higher Asian population in the placebo group.

Overall, we would consider that these variances in the baseline characteristics are minor and taken together, do not favour one group over the other. Therefore, they are unlikely to produce any significant differences in outcome.

A17. Please clarify the flow of patients through the trial with a CONSORT flow diagram. Please clarify reasons for withdrawals from the study for each arm of the trial.

The consort flow diagram for the TAGS trial can be found in Appendix D1.2 and is presented here:

Figure 5: Patient flow (CONSORT) diagram for TAGS trial



We have corrected some errors seen in this figure - please advise if the ERG would like Servier to edit parts of the submitted dossier to correct this in Appendix D.

A18. To provide a clinical rationale to support the stratification by prior ramucirumab and the subgroup analysis of these patients, please clarify why prior treatment with ramucirumab might affect the subsequent relative efficacy for TFT?

Stratification factors were selected *at the time of trial design,* shortly after ramucirumab became available in the EU. It was an important stratification factor to select to account for differences in availability (due to reimbursement) in numerous clinical care pathways.

Within the TAGS trial, patients who had previously been treated with ramucirumab were more heavily pre-treated and would have an expected poorer prognosis as a result. Based on this, we would speculate that patients who had not received prior ramucirumab are less heavily pre-treated and therefore could be expected to have better outcomes. However, in the TAGS trial, both patients who had received prior ramucirumab and those that had not had a significant improvement in overall survival. However, the study was not powered to detect a difference in relative improvement between these groups.

A19. Please clarify how many patients entered the trial due to intolerance, and how many entered the trial due to refractory disease, for each treatment arm. Please clarify how this matched the expected numbers of patients who would enter the trial in the UK context, and if relevant, how this has been accounted for in the submission. Please clarify the efficacy of TFT for patients who entered the trial due to refractory disease and those who entered due to intolerance separately.

The reason for discontinuation of the last prior regimen prior to randomisation in the TAGSs trial can be found in Table 5:

Table 5: Discontinuation of last regimen prior to randomisation in TAGS trials

Last regimen prior to randomisation	Trifluridine/tipiracil (n=337)	Placebo (n=170)
Contained fluoropyrimidine	**105 (31%)**	**59 (35%)**
Reason for discontinuation: Refractory to treatment	92 (88%)	53 (90%)
Reason for discontinuation: Intolerant of treatment	11 (10%)	5 (8%)
Reason for discontinuation: Neither	2 (2%)	1 (2%)
Contained platinum	**64 (19%)**	**39 (23%)**
Reason for discontinuation: Refractory to treatment	55 (86%)	33 (85%)
Reason for discontinuation: Intolerant of treatment	7 (11%)	5 (13%)
Reason for discontinuation: Neither	2 (3%)	1 (3%)
Contained taxane	**174 (52%)**	**73 (43%)**

Reason for discontinuation: Refractory to treatment	157 (90%)	63 (86%)
Reason for discontinuation: Intolerant of treatment	17 (10%)	10 (14%)
Reason for discontinuation: Neither	0	0
Contained irinotecan	**105 (31%)**	**62 (36%)**
Reason for discontinuation: Refractory to treatment	103 (98%)	61 (98%)
Reason for discontinuation: Intolerant of treatment	2 (2%)	1 (2%)
Reason for discontinuation: Neither	0	0

It is evident that the majority of patients discontinued their previous regimen due to becoming refractory to treatment. We would expect this to be consistent with UK practice, however to our knowledge there is no data to support this assertion.

A20. p55. Please clarify the effects on HRQoL by providing complete summary statistics for HRQoL, including between-group differences (with p values) in change from baseline for global health status and individual items. Please clarify why you have chosen 10 points (mentioned on p55) as a clinically meaningful difference by providing a reference to support this. A HRQoL publication is mentioned in Shitara 2018. Please provide this if available.

Between-group comparisons and corresponding p-values were not included in the clinical study report, standard deviations are provided in lieu of these. The HRQoL paper mentioned by Shitara et al has not been published.

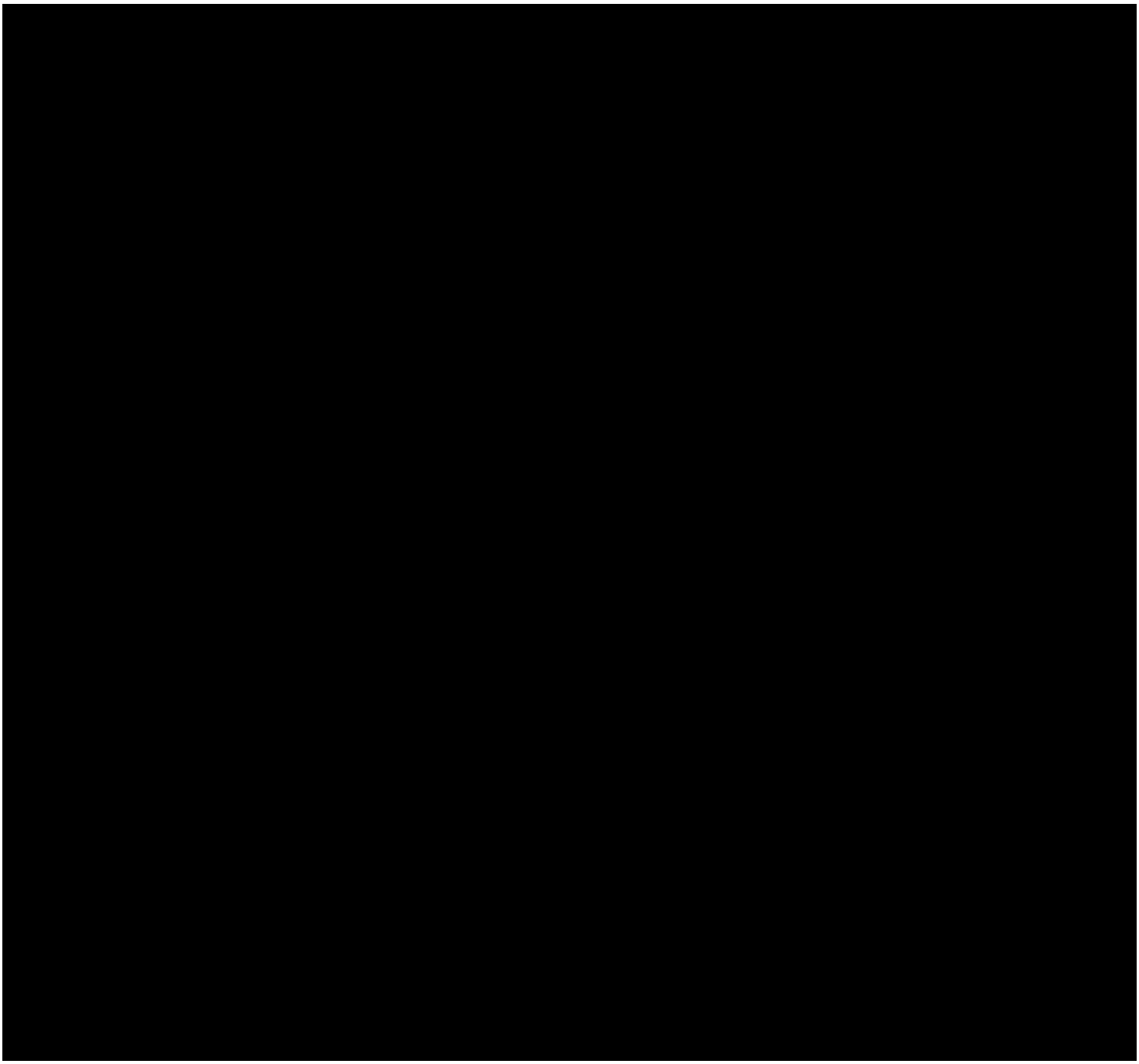
Complete summary statistics for HRQoL (mean and standard deviation) at baseline and changes vs baseline for Last Cycle completed / Safety Follow-up set (30 days after last cycle

if the data had not already been collected in the previous 4 weeks) tabulated below in Table 6 for:

- **EORTC QLQ-C30**
    - Global health status
    - 5 functional scales (physical, role, cognitive, emotional and social)
    - 3 symptom scales (fatigue, pain and nausea/vomiting)
    - 6 additional items (dyspnoea, loss of appetite, insomnia,constipation, diarrhoea and financial impact)
- **STO-22 (Gastric cancer-specific supplementary module)**
    - 5 multi-item scales (dysphagia, pain, reflux, dietary restriction and anxiety)
    - 4 additional single items (dry mouth, taste, body image and hair loss)

Table 6: Complete summary statistics for HRQoL (mean and standard deviation) at baseline and changes vs baseline for Last Cycle completed / Safety Follow-up set

[redacted]



To support the interpretation of changes in EORTC QLQ-C30 scores over time, a 10 point difference was identified as being clinically relevant. This is not necessarily synonymous with the term “clinically important,” which itself can be interpreted in different ways depending on perspective. A 10 point change in score from baseline was defined *a priori* in the statistical analysis plan as being clinically relevant. This is likely perceptible at individual level, and potentially of importance.

Osoba et al.[8] developed the Subjective Significance Questionnaire (SSQ) (Osoba et al., 1998). The SSQ covers patients’ perceived changes in physical, emotional, and social functioning and in global quality of life, using a 7-point scale ranging from ‘much worse’ through ‘no change’ to ‘much better’. Patients completed QLQ-C30 on two occasions, on the second they

also completed the SSQ. Patients reporting 'a little' change for better or worse on a particular scale (function or symptom) had QLQ-C30 changes about 5 to 10. Those reporting 'moderate' change had changed about 10 to 20, and 'very much' change corresponded to a change greater than 20.

[redacted]

A22. Please clarify how the statement "Other available guidelines focus on Asian populations, however due to biological differences between gastric cancer in Asian versus non-Asian patients relevance of these guidelines to European patients is not relevant" is consistent with the data that has entered the model, which includes Japanese patients?

The statement referenced by the ERG features within Section B.2.13 of Document B. Here, we refer to the existence of guidelines for the management of mGC from various international and national bodies, including ESMO and NICE. We also note the existence of other guidelines focused on Asian populations.

Within the statement referenced by the ERG, the use of the phrases "not relevant" and "biological differences" may be somewhat misleading – the statement was intended to clarify that the relevance of the guidelines to European patients is unclear based on current findings

from clinical trials, as there are known differences between European and Japanese mGC patients, including (but not limited to) biological differences.

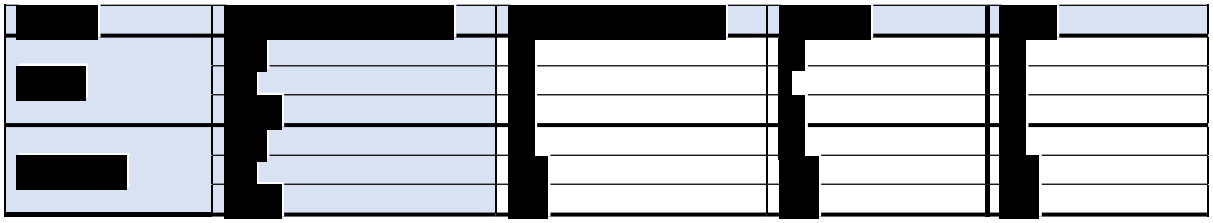
Later in Section B.2.13, the EMA's Committee for Medicinal Products for Human Use (CHMP) opinion concerning nivolumab (Opdivo®, Bristol-Myers Squibb) is discussed. In this report, the CHMP stated the following concerning the patient population studied:

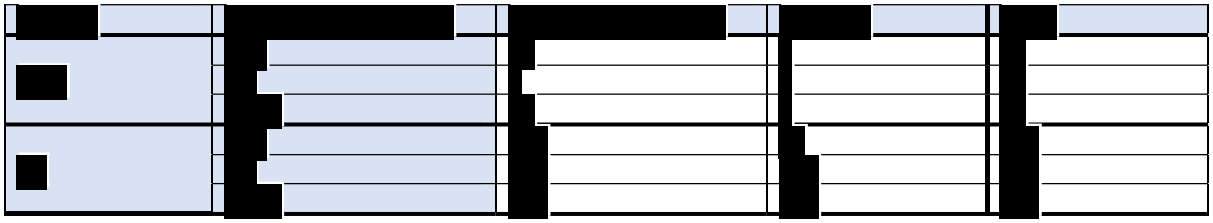
"… While [pharmacokinetic equivalence] has been demonstrated to be sufficiently comparable between Asian and non-Asian patients, the disease itself (gastric/gastroesophageal adenocarcinoma) differs in a number of relevant aspects between non-Asian and Asian patients – including differences in disease biology, patients' characteristics, and variability in treatment practices – which makes it is highly uncertain that non-Asian patients will derive a similar benefit from treatment with nivolumab..." Section 4. Benefit-risk balance, page 162 of 165.

Based on this report, it can be seen that nivolumab could not be recommended by the CHMP at least in part due to the fact that the trial was only conducted in an Asian population, and that the efficacy and safety of nivolumab in a non-Asian population had not been established. One of the key differences highlighted was the variability in treatment practices – for example, the use of paclitaxel + ramucirumab is noted as a new standard of care and was rated as recommendation category 1 *"treatment regimens that are recommended in clinical practice"* in the second-line setting by the Japanese Gastric Cancer Association.[9]

In the TAGS trial, n=46 out of n=337 (14%) trifluridine/tipiracil patients, and n=27 out of n=170 (16%) placebo patients were Japanese. Table 7 shows the number of Japanese patients stratified by prior treatment with ramucirumab. This table illustrates that the majority of Japanese patients were previously treated with ramucirumab, and are therefore excluded within the base-case analysis presented (i.e. the 'non-ramucirumab' population). The exclusion of all Japanese patients was considered inappropriate given that there is a non-zero number of Asian patients who may be treated in UK clinical practice – the 2011 census analysis estimated that amongst the 56 million residents in England and Wales, 86% were White, 8% were Asian/Asian British and 3% were Black/African/Caribbean/Black British.[10]

Table 7: Patient numbers by treatment arm and exposure to ramucirumab





The economic model considers the population considered most applicable to the UK setting, and aligned with treatment guidelines in the UK. Therefore, patients previously treated with ramucirumab (which is not recommended in the UK) were excluded which by extension excluded the majority of Japanese patients from the sample.

A23. Clinical advisors to the ERG suggested that patients in the UK do not generally go on to receive third or more line chemotherapy. Please clarify how inclusion of patients who had received three or more previous lines of chemotherapy impacts on the generalisability of the study results to England and how has this been accounted for in the submission.

Survival time is expected to decrease at each line of therapy, with a significant correlation found between the mortality rate and the number of treatment lines received ($P<0.001$).[11]Therefore the inclusion of patients who would be expected to a have a shorter survival time in the TAGS trial would be expected to negatively impact the outcomes. This was reflected in the advisory board, where the UK attending clinicians felt that the baseline characteristics were "*worse*" in the TAGS trial in comparison to what they observe in the UK population of patients potentially eligible for third-line treatment. This would suggest that patients in the UK would be expected to have better outcomes than would be anticipated based on the TAGs trial results.

A24. Please clarify the number of patients at risk in the Kaplan-Meier plots. For example, there are only 287/337 events in Figure 8. We can see from elsewhere in the submission that 19 remained on treatment and from the CSR Figure 2 that three withdrew, two refused follow-up, and 14 withdrew consent. However, this appears to leave 12 unaccounted for.

Figure 8 of Document B refers to the Kaplan-Meier curve of progression-free survival (PFS). This curve applies to the intention-to-treat (ITT) population. Referring to the numbers at risk in the trifluridine/tipiracil arm, it is noted that there are 337 patients at risk at t=0, and 287 events. For the purpose of the PFS curve, events are defined as either progression or death; and the following reasons for censoring were considered: discontinued follow-up, follow-up ongoing at the time of analysis, initiated antitumor therapy, and missed visit (>91 days since last response).

Figure 2 of the CSR refers to the patient disposition within the TAGS trial. The numbers referred to are related to those continuing to receive treatment (n=19), patients who were lost to follow-up (n=3), patients who discontinued the study and refused follow-up (n=2), and those who withdrew their consent (n=14). It is important to highlight that these numbers should not be taken as mutually exclude and exhaustive – it is entirely possible that one patient was considered to have withdrawn their consent (n=14, which is related to the use of treatment) and have refused follow-up (n=2, which is related to enrolment within the study, regardless of treatment use), for example. Furthermore, the discontinuation of treatment is not equivalent to the discontinuation of the study – this is evident within the difference in total numbers (n=316 versus n=257).

The numbers at risk in the PFS Kaplan-Meier curve are correct. The n=50 censored observations (337 – 287) were due to the following reasons: discontinued follow-up (n=12); follow-up ongoing at the time of analysis (n=20); initiated antitumor therapy (n=8); and missed visit (>91 days since last response) (n=10).

A25. p47 of the CS states that a one-sided stratified log-rank test was used for both progression-free survival (PFS) and overall survival (OS) and that the HR was based on a pre-specified stratified Cox model. Please clarify what covariates were adjusted in the stratified log-rank test and stratified Cox model.

Both the log-rank test the and Cox proportional-hazards model were assessed with and without the three stratification factors included within the TAGS trial. These were: region (Japan versus rest-of-the-world), ECOG status at baseline (0 or 1), prior treatment with ramucirumab (yes or no). For clarity, the Interactive Voice/Web Response System definition of each strata were used to inform the analysis (as opposed to the electronic case report form value[s]).

A26. p87 of the CSR states that the final model for the multivariate analysis included factors for treatment, region, ECOG status at baseline, prior treatment with ramucirumab, age group (< 65 vs ≥ 65 years), number of prior regimens, number of metastatic sites, histology subtype, and HER2 status at baseline. p56 of the CS states that after adjusting for ECOG status, age, number of previous chemotherapy regimens (two versus three), number of metastatic sites, and HER2 status, the adjusted HR was 0.69 with 95% CI 0.56-0.85. Please clarify why the covariates adjusted did not match the covariates listed for the final model in the CSR.

In the prespecified multivariate Cox regression analysis, no baseline patient characteristics or disease factors analysed were identified as being predictive of overall survival (all $p_{interaction}$ ≥0·24).[5]

The factors listed on page 56 of the company submission: ECOG performance status (p<0·0001), age (p=0·00041), number of previous chemotherapy regimens (two vs three or more; p=0·033), number of metastatic sites (p=0·0014), and HER2 status (p=0·016) were prognostic of improved overall survival. Nevertheless, the magnitude of the trifluridine/tipiracil treatment effect was maintained after adjustment for these factors (adjusted HR 0·69 [95% CI 0·56–0·85]).[5]

A27. Please clarify what software and package were used for extrapolation of survival curves.

The extrapolation of survival curves was performed using the statistical software *R*[12], using the package *flexsurv.*[13]

A28. None of the standard parametric distributions appears to fit the PFS data well. Please provide the extrapolation for PFS using more flexible models such as Royston and Parmar natural spline models and if these are more appropriate add as options within the model.

Servier acknowledges that due to the protocol-driven "kinks" in the Kaplan-Meier curve for PFS, the fitted parametric models provide a somewhat-limited fit to the data. Throughout development of the economic model, a number of alternative approaches were considered to address this feature of the data from the TAGS trial. These included:

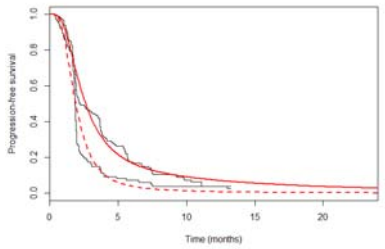
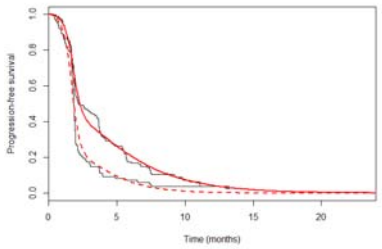
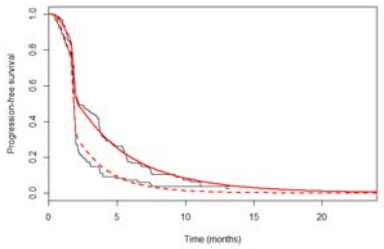
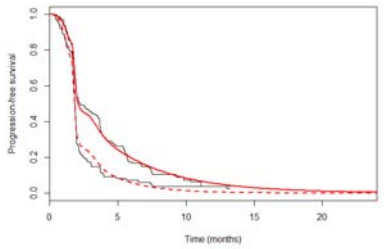
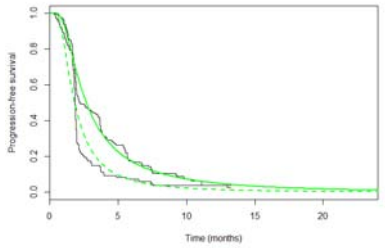
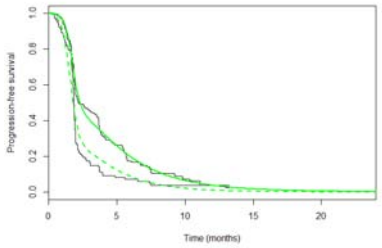
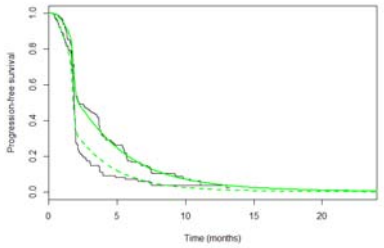
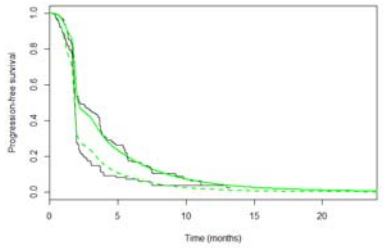
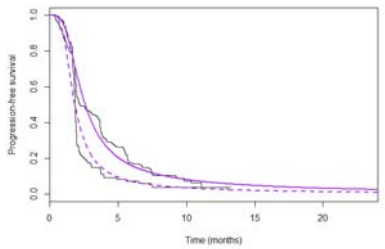
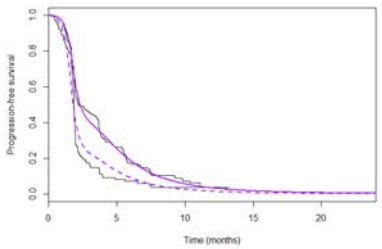
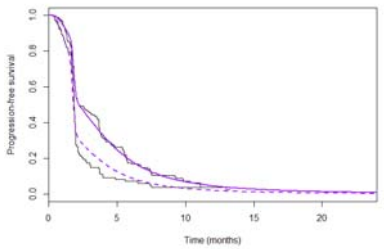
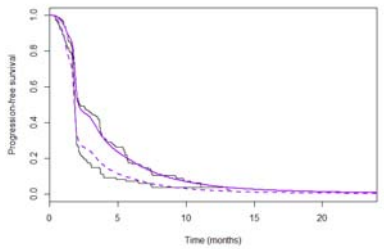
- Standard parametric models (accepting that these models would be unlikely to fully reflect the protocol-driven kinks in the PFS Kaplan-Meier curve)
- Piecewise approach using the Kaplan-Meier curve directly followed by a standard parametric model after a given time cut-off
- Other more flexible parametric modelling approaches, such as the use of spline-based models as highlighted by the ERG

The model allows the use of the first and second of these options for PFS, though a fully-parametric model is utilised to inform the model base-case analysis. A parametric approach was chosen in favour of a piecewise approach as literature notes the choice of cut-off is arbitrary, and dependent on the choice of cut-off the cost-effectiveness results could change markedly.[14] Nevertheless, both options were included within the economic model file for completeness.

Other more flexible approaches (i.e. the third option in the list above) were not considered at the time of model development. This is because these flexible models were expected to "over-fit" the Kaplan-Meier curve for PFS, particularly noting that to reflect each of the "kinks" in the curve a spline with several knots would be necessary to specify.

Table 8 provides a crude overview of spline models fitted to the PFS curves for the no prior ramucirumab population. Splines were fitted using 1, 3, 5, and 10 internal knots, using each of the functional forms permitted by the *flexsurv* package in the statistical software *R*. the selection of the number of knots was arbitrary, yet intended to provide a broad range of models to compare. From Table 8, it may be observed that as the number of knots is increased, the fit to the PFS curve improves (as an increasingly-flexible spline model is able to better reflect the “kinks” in the curve), and the Akaike Information Criterion (AIC) also improves. The functional form does not appear pivotal to the visual or statistical fit of each spline model, with the hazard- and odds-based splines providing generally better fit than the normal-based splines.

Table 8: Spline models fitted to progression-free survival for the no prior ramucirumab population

Function	1 knot	3 knots	5 knots	10 knots
Hazard	 **AIC:** 3,061.69	 **AIC:** 3,000.70	 **AIC:** 2,931.32	 **AIC:** 2,920.47
Normal	 **AIC:** 3,071.11	 **AIC:** 3,001.54	 **AIC:** 2,937.47	 **AIC:** 2,921.26
Odds	 **AIC:** 3,057.63	 **AIC:** 2,994.58	 **AIC:** 2,936.21	 **AIC:** 2,922.16

Key: AIC, Akaike information criterion.
Note: Solid line = trifluridine/tipiracil, dashed line = placebo. Models fitted in the *R* package *flexsurv* using the function *flexsurvreg*.

Servier does not consider the spline models to provide a substantially improved fit when compared within the generalised gamma model used in the base-case analysis, and urge caution be exercised when interpreting the statistical goodness-of-fit scores – the AIC is lowest for the 10-knot spline models, which are expected to "over-fit" the data and have been included only as a straw man. While a good visual fit to the Kaplan-Meier curve is important, a trade-off between the goodness-of-fit and the likely "true" survival curve (i.e. were progression captured within the TAGS trial in real time) is necessary to make when choosing appropriate curves to apply within the economic model.

Nevertheless, within the economic model the hazard- and odds-based spline models were added as exploratory options for informing PFS for completeness. The normal-based spline models were not implemented as for each number of knots tested, the AIC was consistently worse than other models. The spline models have been added for PFS in the no prior ramucirumab patient population only. The impact of each of these curves on the ICER is provided in response to clarification question B1.

Section B: Clarification on cost-effectiveness data

B1. Priority. Please provide a revised base case in the light of any changes made in response to the clarification questions.

In response to this clarification question and others, a revised economic model file has been developed and the base-case cost-effectiveness results have been updated. The following changes have been made to the base-case cost-effectiveness analysis in response to the clarification questions raised by the ERG:

1. **B7:** Re-estimation of the SE for medical resource use costs using previous versions of the NHS reference costs database
2. **B8:** Re-fitting of the log-normal distribution for BSA in the statistical software *R* (such that parameter uncertainty may be ascertained)
3. **B9:** Correction of the dosing of trifluridine/tipiracil (which was previously based on an incorrect hard-coding of values)
4. **B11:** Correction of the VBA code concerning the PFS KM scenario analysis (which was previously offsetting incorrectly)

Changes 1 and 4 do not affect the base-case deterministic ICER, whereas changes 2 and 3 affect the total costs associated with trifluridine/tipiracil and/or BSC (affecting the ICER). All changes are clearly documented in the revised company model.

Base-case analysis

The discounted base-case cost-effectiveness results for trifluridine/tipiracil versus BSC are provided in Table 9 (with PAS) and Table 10 (without PAS). Trifluridine/tipiracil provides an additional 0.153 QALYs and 0.226 LYs (a 43.97% increase on baseline survival for patients receiving BSC), with incremental costs of [redacted] (with PAS) and [redacted] (without PAS). The incremental cost-effectiveness ratio (ICER) is £45,164 (with PAS) and [redacted] (without PAS) per QALY gained.

Table 9: Base-case results (with PAS)

Technologies	Total			Incremental			ICER (£/QALY)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	
PBO + BSC	[redacted]	0.514	0.349				
T/T + BSC	[redacted]	0.740	0.502	[redacted]	0.226	0.153	45,164

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

Table 10: Base-case results (without PAS)

Technologies	Total			Incremental			ICER (£/QALY)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	
PBO + BSC	[redacted]	0.514	0.349				
T/T + BSC	[redacted]	0.740	0.502	[redacted]	0.226	0.153	[redacted]

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

The ICER (with PAS) decreased from £47,933 (Document B, Table 46) to £45,164 due to the error identified by the ERG in clarification question B9. As described in response to this question, the dosing for trifluridine/tipiracil was mistakenly hard-coded. This led to the values in the model including an additional 100 mg for each dosing band (or in other words, each total dose per cycle was offset incorrectly by one row). Consequently, in the submitted economic model, the average cost per administration for patients receiving the 35mg/m^2 dose was over-estimated.

Probabilistic sensitivity analysis (PSA)

The PSA was re-run with 10,000 iterations, per the submitted base-case analysis. The PSA results are very similar to the deterministic base-case results, as shown in Table 11 (with PAS) and Table 12 (without PAS). The corresponding PSA scatterplots and cost-effectiveness acceptability curves (CEACs) with and without the PAS are provided in Figure 6 and Figure 7.

Trifluridine/tipiracil is associated with a probability of being a cost-effective treatment option of 63.6% (with PAS) and [redacted] (without PAS).

Table 11: PSA results (with PAS)

Technologies	Total			Incremental			ICER (£/QALY)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	
PBO + BSC	[redacted]	0.518	0.351				
T/T + BSC	[redacted]	0.745	0.505	[redacted]	0.227	0.153	45,101

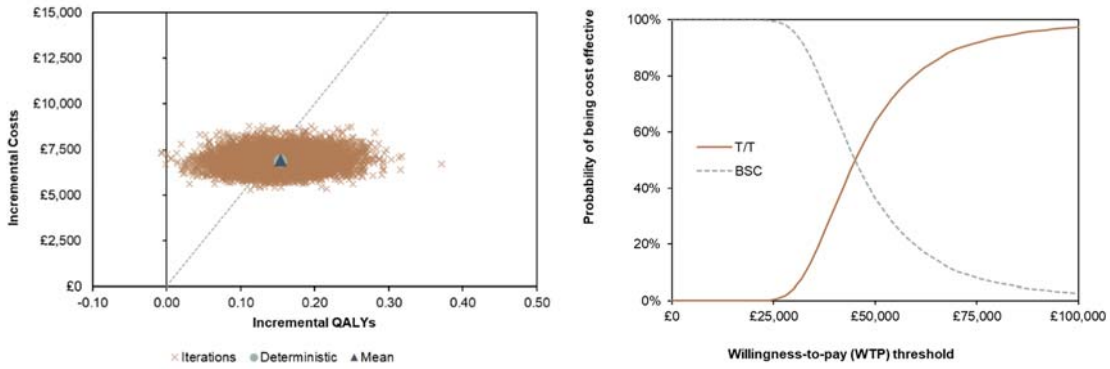
Key: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life years gained; PBO, placebo; QALYs, quality-adjusted life years; T/T, trifluridine/tipiracil.

Table 12: PSA results (without PAS)

Technologies	Total			Incremental			ICER (£/QALY)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	
PBO + BSC	[redacted]	0.518	0.351				
T/T + BSC	[redacted]	0.745	0.505	[redacted]	0.227	0.153	[redacted]

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

Figure 6: PSA scatterplot and CEAC (with PAS)



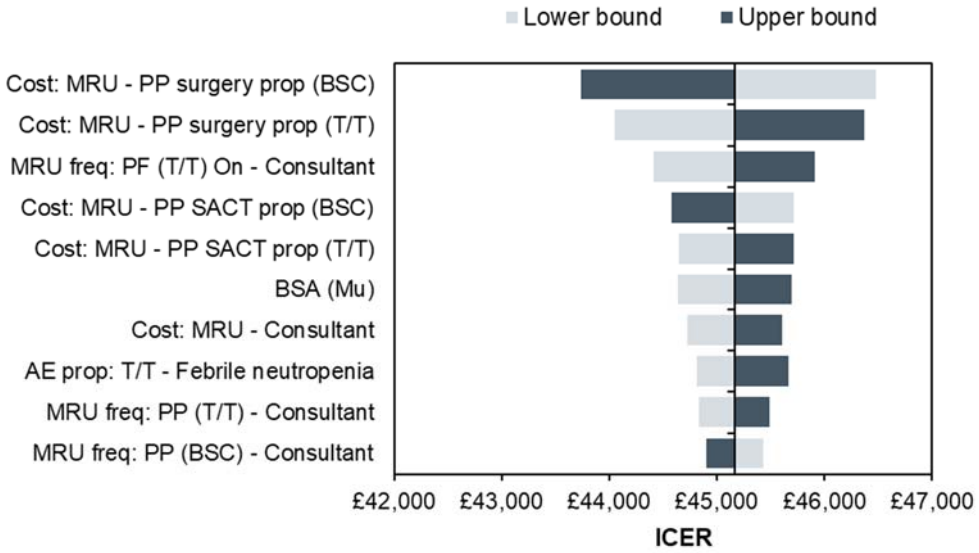
Key: BSC, best supportive care; QALY, quality-adjusted life year; T/T, trifluridine/tipiracil.

[redacted]

One-way sensitivity analysis (OWSA)

The OWSA was also repeated, with results presented as tornado diagrams with and without the PAS for trifluridine/tipiracil in Figure 8 and Figure 9. The key difference in the OWSA versus the results presented in Document B is the introduction of the mu (μ) parameter in the lognormal distribution for BSA ($BSA \sim Lognormal(\mu, \theta)$), which is rank #6 with PAS and rank #4 without the PAS.

Figure 8: Tornado diagram (with PAS)



Key: AE, adverse event; BSA, body surface area; BSC, best supportive care; freq., frequency; ICER, incremental cost-effectiveness ratio; MRU, medical resource use; PF, progression-free; PP, post-progression; prop, proportion; SACT, systemic anticancer therapy; T/T, trifluridine/tipiracil.

Figure 9: Tornado diagram (without PAS)

Key: AE, adverse event; BSA, body surface area; BSC, best supportive care; freq., frequency; ICER, incremental cost-effectiveness ratio; MRU, medical resource use; PF, progression-free; PP, post-progression; prop, proportion; SACT, systemic anticancer therapy; T/T, trifluridine/tipiracil.

Scenario analyses

The scenario analyses described within Section B.3.8.3 of Document B were also re-run with the updated base-case settings, with results presented in Table 13. As noted in response to clarification question B5, there were a number of errors in Table 53 of Document B.

The majority of scenarios yielded an ICER of less than £50,000 (with PAS) or [redacted] (without PAS). The scenarios associated with the largest increase in the ICER were those relating to the restriction of the time horizon, the chosen survival curve(s), and alternative health state utility values. A description of the rationale behind the choice of time horizon and utility model is provided within Document (Section B.3.8.3).

The choice of survival curves is shown to have a notable impact on the ICER. For PFS, the impact is limited (ICER [with PAS] ranging from £43,785 to £46,812). The impact of the time to treatment discontinuation (TTD) curve is slightly more impactful (ICER [with PAS] ranging from £43,916 to £48,707) yet still below £50,000. However, for OS the impact is much greater (ICER [with PAS] ranging from £42,208 to £68,950).

The majority of the OS curves associated with an ICER greater than £50,000 were fitted for each treatment arm independently which (based on an assessment of the underlying hazard function [see Section B.3.3.2 of Document B]) was not considered appropriate. The remaining options with an ICER exceeding £50,000 include the use of the restricted Kaplan-Meier curve (assuming no survival at the end of follow up) and parametric models with either a constant hazard rate over time (exponential) or monotonic hazard function (Gompertz or Weibull). These provide a relatively poor fit to the Kaplan-Meier curves for OS across both treatment arms, versus the more flexible parametric approaches presented (i.e. the lognormal, log-logistic, and generalised gamma); and are associated with substantially worse statistical goodness-of-fit scores (see Table 21 of Document B).

Table 13: Scenario analysis results

Scenario label	ICER	
	With PAS	Without PAS
Base case	**£45,164**	[redacted]
Time horizon of 1 year	£71,447	
Time horizon of 2 years	£54,327	
Time horizon of 3 years	£49,535	
Time horizon of 4 years	£47,517	
Time horizon of 5 years	£46,505	
Time horizon of 6 years	£45,934	
Time horizon of 7 years	£45,601	
Time horizon of 8 years	£45,392	
Time horizon of 9 years	£45,256	
Time horizon of 10 years	**£45,164**	
No discounting of costs, QALYs, or LYs	£43,785	
Include non-European patients in BSA distribution	£44,685	
Limit BSA distribution to only no prior ramucirumab patients	£45,356	
Disable adjustment for background mortality	£45,163	

Scenario label	ICER	
	With PAS	Without PAS
Use mean BSA	£47,822	[redacted]
Use cost per mg assuming mean BSA	£45,862	
Remove initial nurse cost	£45,016	
Remove impact of down dosing	£46,283	
Remove impact of dose delays	£45,465	
Assume same medical resource use per TA378 (ramucirumab)	£48,111	
Remove post-progression costs	£46,539	
Assume lung cancer as a proxy for end-of-life care	£45,261	
Assume breast cancer as a proxy for end-of-life care	£45,119	
Assume prostate cancer as a proxy for end-of-life care	£45,000	
Assume average of four cancers as a proxy for end-of-life care	£45,136	
OS: Independent - Exponential, both arms	£51,878	
OS: Independent - Generalised gamma, both arms	£65,496	
OS: Independent - Gompertz, both arms	£68,950	
OS: Independent - Log-logistic, both arms	£46,942	
OS: Independent - Log-normal, both arms	£51,642	
OS: Independent - Weibull, both arms	£61,310	
OS: Dependent - Exponential, both arms	£51,878	
OS: Dependent - Generalised gamma, both arms	£42,938	
OS: Dependent - Gompertz, both arms	£55,576	
OS: Dependent - Log-logistic, both arms	£42,208	
OS: Dependent - Log-normal, both arms	**£45,164**	
OS: Dependent - Weibull, both arms	£58,363	
OS: KM, both arms	£59,273	
PFS: Independent - Exponential, both arms	£45,495	
PFS: Independent - Generalised gamma, both arms	**£45,164**	
PFS: Independent - Gompertz, both arms	£44,639	
PFS: Independent - Log-logistic, both arms	£45,479	
PFS: Independent - Log-normal, both arms	£45,300	
PFS: Independent - Weibull, both arms	£43,859	
PFS: Dependent - Exponential, both arms	£45,495	
PFS: Dependent - Generalised gamma, both arms	£46,812	
PFS: Dependent - Gompertz, both arms	£44,766	
PFS: Dependent - Log-logistic, both arms	£45,405	
PFS: Dependent - Log-normal, both arms	£45,075	
PFS: Dependent - Weibull, both arms	£43,785	
PFS: KM, both arms	£44,686	
TTD: Exponential, T/T	£44,288	
TTD: Generalised gamma, T/T	**£45,164**	
TTD: Gompertz, T/T	£44,610	
TTD: Log-logistic, T/T	£48,707	
TTD: Log-normal, T/T	£47,803	
TTD: Weibull, T/T	£44,113	
TTD: KM, T/T	£43,916	
Utility Model 1	**£45,164**	
Utility Model 2	£46,898	
Utility Model 3	£45,449	
Utility Model 4	£47,240	
Utility NICE TA378	£47,857	
Utility NICE TA208	£48,473	
Exclude AE-related disutilities	£44,141	
Exclude QALYs lost due to AEs in subsequent treatment	£45,194	

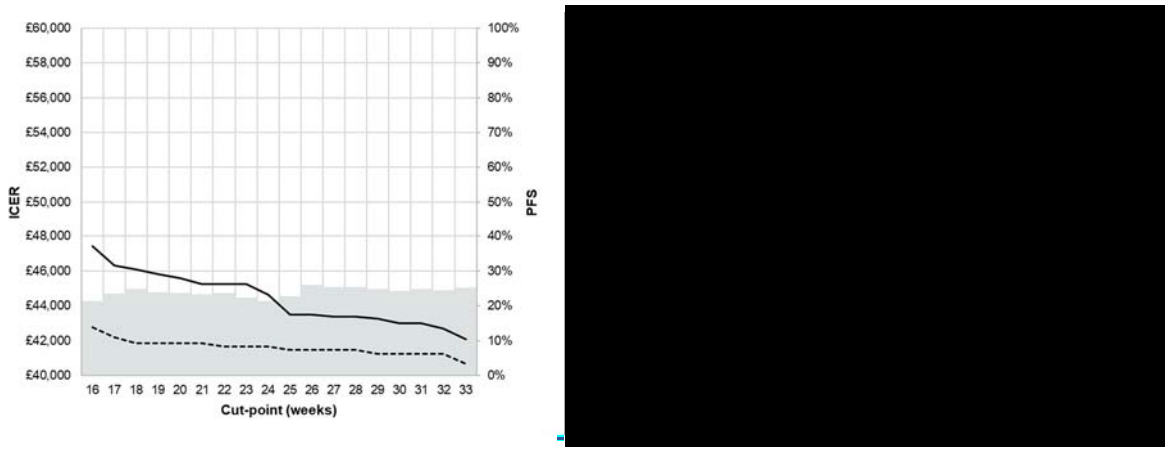
Key: AE, adverse event; BSC, best supportive care; BSA, body surface area; ICER, incremental cost-effectiveness ratio; KM, Kaplan-Meier; LY, life year; NICE, National Institute for Health and Care Excellence; OS, overall survival; PAS, patient access scheme; PFS, progression-free survival; QALY, quality-adjusted life year; T/T, trifluridine/tipiracil; TA, technology appraisal; TTD, time to treatment discontinuation.

Note: Base-case analyses highlighted in bold print throughout the table for context.

Scenarios concerning the estimation of progression-free survival (PFS)

In addition to the results of the scenario analyses provided in Table 13, additional scenario analyses were undertaken concerning PFS. The first of these analyses was provided in Document B, wherein PFS was derived directly from the KM curve until a specified cut-point, after which the base-case extrapolation was assumed to apply (that is, the estimated conditional survival estimates were lifted from the base-case extrapolation without re-basing the survival curve). The cut-point was varied from 16 to 33 weeks in weekly increments, and the impact on the ICER was recorded. The results of this analysis are provided in Figure 10, which show limited impact on the ICER of varying the cut-point in a piecewise modelling approach.

Figure 10: Analysis of cut point for PFS curves (with and without PAS)



Key: ICER, incremental cost-effectiveness ratio; PFS, progression-free survival.

In response to clarification question A28, spline-based models were also fitted to the PFS data from TAGS. The results of this analysis are provided in Table 14. The use of spline models causes the ICER to vary between £43,541 and £46,892 (with PAS), though the specification of a spline model with many knots is expected to "overfit" the data and is not something we would consider credible.

Table 14: Additional PFS scenario analysis results

Scenario label	ICER	
	With PAS	Without PAS
Base case (PFS: Independent - Generalised gamma, both arms)	**£45,164**	
PFS: 1-knot hazard-based spline	£43,541	
PFS: 3-knot hazard-based spline	£45,844	
PFS: 5-knot hazard-based spline	£45,628	
PFS: 10-knot hazard-based spline	£45,576	
PFS: 1-knot odds-based spline	£45,072	
PFS: 3-knot odds-based spline	£46,892	
PFS: 5-knot odds-based spline	£46,587	
PFS: 10-knot odds-based spline	£46,503	

Key: ICER, incremental cost-effectiveness ratio; PAS, patient access scheme; PFS, progression-free surviv

B2. Priority. If the analyses need to be amended for changes in treatment or dose (see clarification question A1), please provide an updated estimate of the ICER.

No changes to the model have been made in response to clarification question A1, hence no updated results are provided in response to this question. The points raised by the ERG concerning changes in treatment or dose are a feature of the TAGS clinical trial, and have been accounted for within the submitted economic model as described within Document B.

B3. Priority. If the analyses need to be amended for treatment after progression (see clarification question A2), please provide an updated estimate of the ICER. We note the model currently includes three cycles of docetaxel.

No changes to the model have been made in response to clarification question A2, hence no updated results are provided in response to this question. The ERG is correct that the model includes the cost of three cycles of docetaxel if post-progression costs are captured within the analysis. The use of docetaxel was assumed to serve as a proxy for further systemic anticancer therapy (SACT). Subsequent SACT was disabled within sensitivity analysis (Section B.3.8.3) as part of the "remove post-progression costs" scenario.

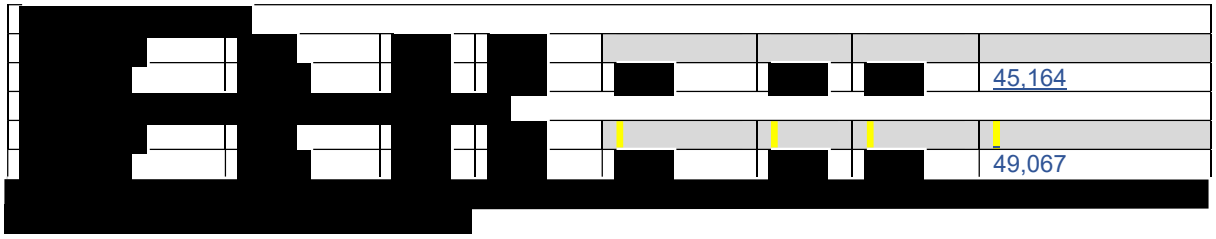
B4. Priority. Please provide a scenario analysis that estimates the ICER using the EU population who have not had ramucirumab treatment. As with the analyses originally submitted please provide the Kaplan-Meier data for the two arms.

To produce the response to this question, the following economic model input parameters were derived for the European, no prior ramucirumab population: (1) OS, (2) PFS, and (3) time on treatment (ToT). All other parameters (e.g. adverse event probabilities, body surface area, etc.) were assumed to be equal to the revised base-case settings (described in response to clarification question B1).

Compared with the base-case ICER for the no prior ramucirumab population, the ICER for European patients with no prior ramucirumab increases from [redacted] to [redacted]. This is primarily due to a reduction in the incremental survival benefit (the difference in life-years decreases from [redacted]). This reduction in calculated life-years is primarily due to the OS Kaplan-Meier curves being slightly closer together after approximately 8 months, which is reflected in the extrapolation of the curves. As discussed in response to clarification question A3, Servier urges caution when interpreting the cost-effectiveness results for this non-pre-specified subgroup.

[redacted]

[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	ICER (£/QALY)



B5. Scenario analyses figures reported in the CS appear different to those generated from the model (examples: fitting independent exponential curves for OS gives an ICER of [redacted] instead of [redacted], fitting a Weibull curve for time on treatment gives an ICER of [redacted] instead of [redacted]). This may only apply to the survival curve scenario. Please clarify the correct values. Additionally, clarify whether the estimated ICER for patients with prior ramucirumab use is £67,564 with PAS.

As highlighted by the ERG, there are several unintended errors within the scenario analysis results presented within Document B. The affected scenarios are related to alternative choices of survival curve and the use of utilities sourced from NICE TA208. The reason for the errors is as follows:

- **For the errors concerning the different choice of survival curves:** On the 'Controls' sheet, the settings relating to the piecewise modelling approach were accidentally enabled when the scenario analyses were run (i.e., ranges c_extrap_os, c_extrap_pfs, and c_extrap_tot were set to "Yes" from 25 weeks). Consequently, the scenarios run only varied the curve up until 25 weeks, after which the curve followed the base-case trajectory (hence the results for the base-case settings are correct). Given that the curve after 25 weeks followed the base-case, the impact of alternative survival curves on the ICER was limited, hence the relatively limited ICER range for these scenarios

- **For the error concerning the use of utility values from TA208:** This appears to be a text error, as the result for this scenario is correct in the submitted economic model.

Servier apologises for any confusion caused by these errors. The base-case results (including sensitivity analyses) have been updated in response to Question B1 (alongside a revised description of the findings of these scenario analyses).

Based on the use of survival curves for patients previously treated with ramucirumab (and all other base-case settings left as per the submitted base-case analysis), the ICER (using the settings per the submitted base-case analysis) is £67,564 (with PAS). However, it should be noted that the option to consider this population was provided within the economic model purely for context – this population is not relevant for decision making in the UK setting, and

is based on a relatively small sample versus the ITT or no prior ramucirumab populations within the TAGS trial.

B6. Age and gender are known confounders in estimating utilities. Please clarify whether including age and gender in the GEE regression analysis for EQ- 5D utility data would have an impact on the results.

While age and sex are known to affect utility in the general population, the role of these factors in the context of a randomised controlled trial (where patients are expected to be broadly balanced between arms) is not expected to be influential. However, for completeness, analysis is provided below for the progression-based utility regression model adding both age and sex as covariates.

The results of the analysis, in addition to the unadjusted regression previously submitted, are provided in Table 16. Two versions of the age- and sex-adjusted analysis are provided; firstly using age as a continuous variable, and secondly as a binary variable for age ≥65.

Table 16: Utility regressions including age and sex

B7. The uncertainty in NHS reference costs appears not to be captured correctly. Perform an analysis where the SE to mean ratio was estimated from the quartile data contained within previous NHS reference costs versions, rather than assuming that the SD was 10% of the mean. In the method used by the company the SE is underestimated as the number of observations was assumed to be the number of cases rather than number of data returns.

There are five costs implemented within the economic model that were identified from the NHS reference costs database. These are: consultant appointment, computed tomography (CT) scan, full blood count (FBC), liver function test (LFT), and renal function test (RFT). In the submitted economic model, the standard deviation (SD) around each of the unit costs was estimated to be 10% of the mean value, and the standard error (SE) was calculated using the quotient of the SD and the square-root of the activity.

As noted by the ERG, the latest version of the NHS reference costs do not provide quartiles for the unit costs, hence their omission within the submitted economic model. However, earlier versions of the NHS reference costs provide quartiles for some of these costs – to the best of Servier's knowledge, quartiles have never been provided for outpatient attendances.

Section 7.7.3.5 of the Cochrane handbook (Medians and interquartile ranges) suggests that the width of the interquartile range will be approximately 1.35 SDs.[15] Therefore, using data from the 2012-13 NHS reference costs database, the ratio of the mean to the standard deviation was estimated.[16] The values used to estimate these ratios are provided in Table 17.

Table 17: Derived ratio of standard error to average value

Item	Average	LQ	UQ	n	$\hat{\sigma}$	SE
RA08A[a]	94.79	59.44	116.68	80	42.40	4.74
DAPS04[b]	1.25	0.82	1.43	122	0.45	0.04
DAPS05[c]	3.01	1.68	3.97	128	1.70	0.15

Key: LQ, lower quartile; n, number of submissions; SE, standard error; UQ, upper quartile; $\hat{\sigma}$, sample standard deivation.
Notes: [a]Computerised Tomography Scan, one area, no contrast, 19 years and over; 370 Medical Oncology; [b]Clinical Biochemistry; [c]Haematology

Using the values reported in Table 17, the SE to mean ratio is estimated to be 3.28%, 4.98%, and 5.00%. Therefore, within the economic model the SE to mean ratio is set to 5% for each of these parameters. Updated one-way and probabilistic sensitivity analysis results are provided alongside the revised base-case deterministic analysis results in response to Question B1.

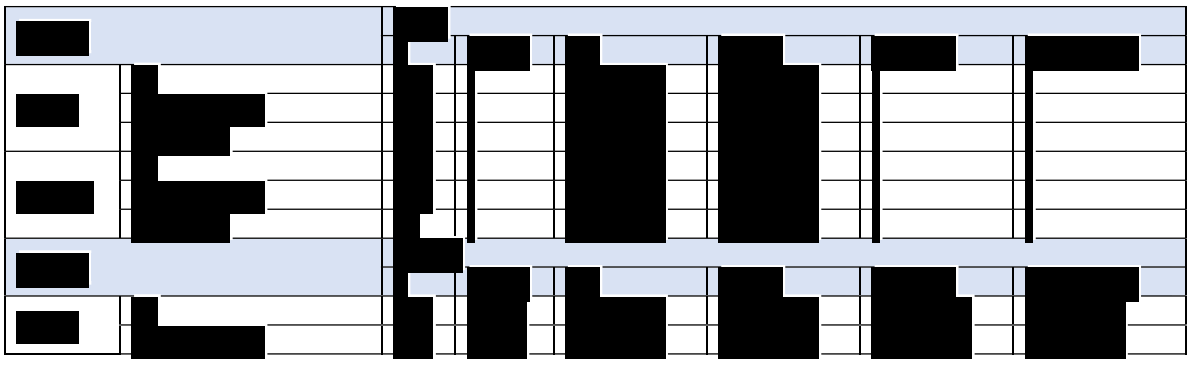
B8. No uncertainty was included in the distribution fitted to BSA. Please comment on the likely impact of omitting this uncertainty from the model.

No parameter uncertainty in the distribution of BSA was included within the originally-submitted economic model, as while there is uncertainty in the distribution of BSA, there are only a limited number of dosing bands for which patients may fall within. Consequently, the impact of the distribution shifting slightly in either direction was expected to only affect those patients on the cusp of a given dosing band. Instead of considering parameter uncertainty for this model input, a number of scenarios concerning the distribution of BSA (e.g. using all versus European patients) or the application of costing (e.g. mean dose) were performed.

To address the potential parameter uncertainty around the distribution of BSA, Servier is unaware of how uncertainty in the parameters derived using a method of moments (MoM) approach may be estimated within the model itself (without performing extensive bootstrapping). However, in lieu of estimating these parameters using the MoM approach, the distribution of BSA was estimating in the statistical software *R* using the *MASS* package. The *fitdistr* function was used, with the *densfun* argument set as "lognormal". This package allows for the estimation of the parameters themselves as well as the uncertainty around the parameter point estimates. This approach originally not used as the MoM has the advantage of increased transparency, given that the calculations may be directly integrated within the economic model.

Table 18 presents a comparison of the distribution parameters estimated using both approaches, including the estimated standard deviations for each of the parameters. It is noted that the parameters for Mu and Theta are correlated, and so while the corresponding variance-covariance matrices were also extracted, the off-diagonal elements were approximately zero (indicating very little correlation between the parameters). Therefore, it was considered appropriate to vary these parameters independently for the purpose of informing the economic model.

Table 18: Comparison of parameters estimated for the distribution of body surface area



Key: n, number of patients; loglik, log-likelihood; MoM, method of moments; MASS, Modern Applied Statistics with S; ram, ramucirumab; SD, standard deviation

Within the revised economic model supplied alongside this response, there is now an option to utilise these fitted BSA distribution parameters which in turn may be included within the one-way and probabilistic sensitivity analyses. In the deterministic base-case analysis, the ICER increases by £1 from £47,933 to £47,934. Therefore, for simplicity, the *R*-based analysis is assumed to constitute part of the revised base-case analysis, results for which are provided in response to clarification question B1.

B9. Model. In the "Data" sheet, cells DV12:21, the total mgs are values, whereas there are formulae for corresponding values in the tables below. Clarify why this is the case and what impact this has on the ICER if an amendment is needed.

The ERG is correct – the values were accidentally hard-coded and are incorrect. The affected table is intended to reflect Table 32 in Document B, in that the dosage per administration should range from 35 to 80mg, dependent on patient body surface area (BSA). The values in the model included an additional 100 mg for each band (or in other words, each total dose per cycle is offset incorrectly by one row). Consequently, in the submitted economic model, the average cost per administration for patients receiving the 35mg/m^2 dose was over-estimated.

This error has been amended within a revised economic model file, supplied alongside the response to this clarification letter. By amending this error, the average cost of administration for trifluridine/tipiracil (for patients receiving a dose of 35mg/m^2, including the PAS discount) is reduced from £2,184 to £2,017 (without PAS). This in turn affects the ICER, which reduces from £47,933 to £45,162 (with PAS). Please disregard the previous results which utilised the incorrectly-implemented dosing of trifluridine/tipiracil, and apologies for any confusion caused. Revised base-case cost-effectiveness results are provided in response to clarification question B1.

B10. Model. In the "BI" sheet, L51, there appears to be a calculation error in the total number of people with metastatic disease in the starting year. We believe it should be (**H47***I21) rather than (I47*I21) in this component of the formula. Clarify the impact of the adjustment, if required, on the ICER.

The ERG is correct – there is an unintended error within the BI sheet. Given that this sheet is concerned only with the budget impact estimate, amending this error does not affect the cost-

effectiveness results. A revised economic model file has been provided where this error has been amended.

B11. Model. In the PFS KM scenario, there appears to be a coding error in the relevant VBA code (line 185 of the DSA VBA module). We believe that the row offset should be i-14. Clarify the impact of the adjustment, if required, on the ICER.

The ERG is correct – there is an unintended error within the VBA code concerning the PFS Kaplan-Meier scenario analysis. The row offset should be i – 14, as suggested. In addition, some minor edits have been made to the code to improve its usability (for example, the code previously required the user to manually select the "KM" setting on the "Controls" sheet).

A revised economic model file has been provided where this error has been corrected. There is no impact on the ICER regarding this amend (the error was introduced following production of the results used to inform the dossier).

Section C: Textual clarification and additional points

C1. The SPC states "neutropenia (54% [35% ≥ Grade 3]), nausea (39% [1% ≥ Grade 3]), fatigue (35% [4% ≥ Grade 3]), anaemia (32% [13% ≥ Grade 3]) and leucopenia (31% [12% ≥ Grade 3])" whereas Table 12 reports Fatigue 27%, anaemia 45%, and leucopenia 23%. Please clarify which are the correct data.

The data referred to in the question are drawn from two distinct populations: colorectal cancer patients and gastric cancer patients. The SPC provided to the ERG is the current SPC (last updated in March 2017). It refers to previous data of trifluridine/tipiracil in colorectal cancer trials. The adverse events referred to in the submission relate to the safety outcomes in the TAGS trial (published in 2018). The updated SPC including the TAGS trial is currently under review by the European Medicines Agency (EMA) and will be submitted to NICE once it is available.

C2. Table 13, p61: some data is on a grey row with bold type, but some just has bold type. Should these also be grey background, to indicate System Organ Class?

Yes, please find the corrected table below. Please advise if the ERG would like Servier to edit parts of the submitted dossier with respect to this clarification question.

System Organ Class Preferred Term	Trifluridine/tipiracil (N=335) n (%)	Placebo (N=168) n (%)
Number of patients with at least 1 serious adverse event	143 (42.7)	70 (41.7)
Blood and lymphatic system disorders	**25 (7.5)**	**4 (2.4)**
Anaemia	13 (3.9)	4 (2.4)
Pancytopenia	7 (2.1)	0
Febrile neutropenia	4 (1.2)	0
Neutropenia	4 (1.2)	0
Gastrointestinal disorders	**55 (16.4)**	**31 (18.5)**
Vomiting	9 (2.7)	1 (0.6)
Abdominal pain	8 (2.4)	6 (3.6)
Diarrhoea	6 (1.8)	0
Dysphagia	6 (1.8)	2 (1.2)
Gastrointestinal haemorrhage	4 (1.2)	1 (0.6)
Intestinal obstruction	4 (1.2)	3 (1.8)
Ascites	3 (0.9)	7 (4.2)
Gastric haemorrhage	3 (0.9)	3 (1.8)

System Organ Class Preferred Term	Trifluridine/tipiracil (N=335) n (%)	Placebo (N=168) n (%)
Small intestinal obstruction	3 (0.9)	2 (1.2)
Upper gastrointestinal haemorrhage	2 (0.6)	2 (1.2)
General disorders and administration site conditions	**28 (8.4)**	**21 (12.5)**
General physical health deterioration	21 (6.3)	15 (8.9)
Asthenia	1 (0.3)	3 (1.8)
Infections and infestations	**20 (6.0)**	**9 (5.4)**
Neutropenic sepsis	4 (1.2)	0
Pneumonia	4 (1.2)	2 (1.2)
Metabolism and nutrition disorders	**18 (5.4)**	**7 (4.2)**
Decreased appetite	11 (3.3)	4 (2.4)
Musculoskeletal and connective tissue disorders	**1 (0.3)**	**3(1.8)**
Back pain	0	3 (1.8)
Neoplasms benign, malignant and unspecified (including cysts and polyps)	**8 (2.4)**	**4 (2.4)**
Malignant ascites	1 (0.3)	2 (1.2)
Respiratory, thoracic and mediastinal disorders	**15 (4.5)**	**4 (2.4)**
Pleural effusion	5 (1.5)	1 (0.6)
Pulmonary embolism	5 (1.5)	2 (1.2)
Dyspnoea	4 (1.2)	2 (1.2)
Note: At each level of summation (overall, system organ class, preferred term), patients were only counted once at the highest toxicity grade		

C3. p51. It is stated "Servier has submitted this as an additional population within its economic model. This is discussed further in section B.2.1.3.". However, section B.2.1.3 appears to be a section about the systematic review screening process, not about ramucirumab. Please correct this cross-reference.

The correct cross-reference should be to Section B.3.2.1 (not B.2.1.3). Please advise if the ERG would like Servier to edit parts of the submitted dossier with respect to this clarification question.

C4. Model. In the "Costs" sheet, we believe that J25 should be fixed and not change when J21 changes. We believe it should refer to 'Data!EA43' instead of 'Data!DU53'. This does not affect the ICER in the analyses presented.

The ERG is correct – this value should not change when the value of J21 changes. This error has been corrected in the revised economic model provided alongside these responses. As noted by the ERG, this does not affect the ICER.

References

1. Bang Y-J, Van Cutsem E, Feyereislova A, Chung HC, Shen L, Sawaki A, et al. Trastuzumab in combination with chemotherapy versus chemotherapy alone for treatment of HER2-positive advanced gastric or gastro-oesophageal junction cancer (ToGA): a phase 3, open-label, randomised controlled trial. The Lancet. 2010;376(9742):687-97.
2. Søren Højsgaard, Ulrich Halekoh, Yan. J. Package 'geepack': Generalized Estimating Equation Package Version 1.2-1 2016 [Available from: https://cran.r-project.org/web/packages/geepack/geepack.pdf.
3. Elselvier R&D Solutions. White Paper: A Comparison of Emtree and MeSH 2015 [Available from: http://supportcontent.elsevier.com/Support%20Hub/Embase/Files%20&%20Attachements/4685-Embase_White%20Paper_Comparison%20of%20Emtree%20and%20MeSH_July%202015.pdf.
4. Jeong J, Jeung HC, Rha SY, Im CK, Shin SJ, Ahn JB, et al. Phase II study of combination chemotherapy of 5-fluorouracil, low-dose leucovorin, and oxaliplatin (FLOX regimen) in pretreated advanced gastric cancer. Annals of Oncology. 2008;19(6):1135-40.
5. Shitara K, Doi T, Dvorkin M, Mansoor W, Arkenau H-T, Prokharau A, et al. Trifluridine/tipiracil versus placebo in patients with heavily pretreated metastatic gastric cancer (TAGS): a randomised, double-blind, placebo-controlled, phase 3 trial. The Lancet Oncology. 2018;19(11):1437-48.
6. Jørgensen JT HM. HER2 as a Prognostic Marker in Gastric Cancer - A Systematic Analysis of Data from the Literature. J Cancer. 2012;3:137-44.
7. Berlth F, Bollschweiler E, Drebber U, Hoelscher AH, Moenig S. Pathohistological classification systems in gastric cancer: diagnostic relevance and prognostic value. World J Gastroenterol. 2014;20(19):5679-84.
8. Osoba D, Rodrigues G, Myles J, Zee B, Pater J. Interpreting the significance of changes in health-related quality-of-life scores. Journal of Clinical Oncology. 1998;16(1):139-44.
9. Japanese Gastric Cancer A. Japanese gastric cancer treatment guidelines 2014 (ver. 4). Gastric Cancer. 2017;20(1):1-19.
10. Office for National Statistics. 2011 Census analysis: Ethnicity and religion of the non-UK born population in England and Wales: 2011 2011 [Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/culturalidentity/ethnicity/articles/2011censusanalysisethnicityandreligionofthenonukbornpopulationinenglandandwales/2015-06-18.
11. Davidson M, Cafferkey C, Goode EF, Kouvelakis K, Hughes D, Reguera P, et al. Survival in Advanced Esophagogastric Adenocarcinoma Improves With Use of Multiple Lines of Therapy: Results From an Analysis of More Than 500 Patients. Clinical Colorectal Cancer. 2018;17(3):223-30.
12. R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. [Available from: http://www.R-project.org/.

13. Jackson C, Metcalfe P, Amdahl Jf. Flexible Parametric Survival and Multi-State Models. R package version 1.1.1. 2019.
14. Andrew D, Andrew B, John S, Levy A, Omar E, Samuel W, et al. The Ends Justify the Mean: Outcome Measures for Estimating the Value of New Cancer Therapies. Health Outcomes Research in Medicine. 2012;3(1):e25-e36.
15. Higgins JPT GS. Cochrane Handbook for Systematic Reviews of Interventions: Wiley Online Library; 2008.
16. National Health Service England. NHS reference costs 2012 to 2013. 2013 [Available from: https://www.gov.uk/government/publications/nhs-reference-costs-2012-to-2013.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Trifluridine–tipiracil for treating metastatic gastric or gastro-oesophageal junction cancer after 2 or more therapies [ID1507]

Additional clarification questions (2)

August 2019

File name	Version	Contains confidential information	Date
2019-08-06_ID1507 Trifluridine ERG Additional clarif response (2)_v1-0	1-0	Yes	06/08/2019

Section A: Additional clarification on cost-effectiveness data

A2. In response to Question A2 the company has interpreted the question as relating to treatments after the treatment period, whereas the intended meaning was throughout the trial. Please could the company further clarify how many patients were censored in the PFS analysis due to commencing no-study anti-cancer treatments in each arm, and if different, how many patients stopped treatment to start a non-study anti-cancer treatment?

Within the TAGS clinical trial, the reasons for censoring within the progression-free survival (PFS) analysis were: (1) discontinued follow-up, (2) follow-up ongoing at the time of analysis, (3) missed visit (>91 days since last response), and (4) initiated anti-tumour therapy. [redacted]

Within the CSR, results of supportive analyses of PFS including clinical progression and initiation of anti-tumour therapy as PFS events were consistent with that of the primary analysis of PFS (please see CSR Table 26, replicated below in Table 1 for completeness). Please note however that hazard ratios (HR) should be interpreted with extreme caution within the context of the PFS outcome.

Table 1: Primary and Supportive Analyses of PFS (Intent-to-Treat Population)

	T/T (N=337)	PBO (N=170)	HR	95% CI
	PFS (months) Median (95% CI)	PFS (months) Median (95% CI)		
Radiologic progression only	[redacted]	[redacted]	[redacted]	[redacted]
Including clinical progression	[redacted]	[redacted]	[redacted]	[redacted]
Including clinical progression and initiation of anti-tumour therapy	[redacted]	[redacted]	[redacted]	[redacted]
Including clinical progression, initiation of anti-tumour therapy, or death without censoring missed visits	[redacted]	[redacted]	[redacted]	[redacted]

Key: CI, confidence interval; HR, hazard ratio; PBO, placebo; PFS, progression-free survival; T/T, trifluridine/tipiracil.
Note: Reproduced from Table 26 of the TAGS Clinical Study Report.

[redacted]

A3. In response to Question A3, the company has provided Kaplan-Meier plots, but no summary statistics for OS and PFS. Please could the company clarify what the HR was for each outcome, along with median OS/PFS for each arm? Please also clarify what the HR is for OS after adjusting for prognostic factors?

Median overall survival (OS) and PFS values are provided within Table 2. Please note that median estimates of PFS in particular should be interpreted with caution, owing to the protocol-driven 'kinks' in the PFS curve.

Table 2: Median OS and PFS (TAGS, no prior ramucirumab, European patients only)

Arm	OS	PFS
T/T	[redacted]	[redacted]
PBO	[redacted]	[redacted]

Key: OS, overall survival; PBO, placebo; PFS, progression-free survival; T/T, trifluridine/tipiracil.

As described within Document B, the proportional hazards (PH) assumption was not considered to hold for both OS and PFS in the ‘no prior ramucirumab’ population. Given the overlapping of this group and that European only no prior ramucirumab population, this assumption is also expected not to hold. For this reason, adjusted or unadjusted HRs were not provided in Servier’s response to clarification question A3.

Servier is unable to provide adjusted HRs accounting for prognostic factors for OS and PFS within the ‘no prior ramucirumab, European patients only’ population as this is not possible within the timeframe allocated to respond to this request. However, provision of any HR analyses (i.e. adjusted or unadjusted) is not considered appropriate by Servier due to the violation of the PH assumption. However, Servier hopes that the provision of the median estimates in this response and the Kaplan-Meier curves in the previous response is sufficient in order for the ERG to perform its critique of the submission.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Trifluridine–tipiracil for treating metastatic gastric or gastro-oesophageal junction cancer after 2 or more therapies [ID1507]

Additional clarification question

July 2019

File name	Version	Contains confidential information	Date
2019-08-01_ID1507 Trifluridine ERG Additional clarif response_v1-0	1-0	Yes	01/08/2019

Section A: Additional clarification on cost-effectiveness data

A1. Priority. The ERG has identified a recent review of algorithms mapping from EORTC-QLQ30 to EQ-5D-3L. The algorithm used in the company's submission is not among the 2 best mappings identified by the Woodcock and Doble (2018) review (Versteegh et al. 2012, Longworth et al. 2014). Please provide analyses exploring the effect of using one or both of these alternative mapping algorithms.

The ERG references a recent study by Woodcock and Doble[1] which reports the findings of an assessment of existing and newly-developed mapping algorithms from the EORTC-QLQ-C30 to the EQ-5D. In this study, the authors identified four previous mapping studies, from which seven distinct mapping algorithms are available:

- Versteegh et al. Mapping QLQC30, HAQ, and MSIS-29 on EQ-5D. Med Decis Making. 2012;32(4):554–68.[2]
- Longworth et al. Use of generic and condition-specific measures of health-related quality of life in NICE decision-making: a systematic review, statistical modelling and survey. Health Technol Assess. 2014;18(9): 1–224.[3]
- Khan et al. A non-linear beta-binomial regression model for mapping EORTC QLQ-C30 to the EQ-5D-3L in lung cancer patients: a comparison with existing approaches. Health Qual Life Outcomes. 2014;12:163.[4]
- Marriott et al. Mapping EORTC-QLQ-C30 to EQ-5D-3L in patients with colorectal cancer. Journal of medical economics. 2017;20(2):193–9.[5]

In assessing the appropriateness of each algorithm, Woodcock and Doble utilised data from the Cancer 2015 longitudinal data set. The sample consisted of patients with a ranger of tumour types and disease stages – approximately 3-6% of patients had oesophagogastric cancer (depending on the subset used), and 18-26% of patients had stage IV cancer (1-3% unknown, remainder stage I to III), based on Table 1 of the paper.

Of the other four studies referenced by Woodcock and Doble, only one considers a population of advanced gastrointestinal cancer patients (Marriott et al.). Each of the

other three studies considers populations from either a range of cancer types, or a single non- gastrointestinal cancer (i.e. Khan et al. consider a non-small-cell lung cancer population).

To inform the economic model, Servier selected the only mapping algorithm estimated within a gastric cancer population (Kontodimopoulos et al.[6]) included within the Health Economics Research Centre (HERC) mapping studies database. Servier acknowledges the existence of other mapping algorithms, but does not consider it appropriate to generalise the use of either of the Versteegh et al. or Longworth et al. algorithms to the population relevant to this appraisal.

In the study referenced by Marriott et al., the authors note that the Versteegh et al. study does not utilise a UK tariff (and so is not aligned with the NICE reference case). However, outside of the tariff used, the study was conducted in only haematological cancers (multiple myeloma and non-Hodgkin's lymphoma).

In NICE DSU TSD 10[7], with regards to the use of mapping algorithms from the published literature, it is stated that:

"… we recommend that careful consideration is given to the generalisability of the mapping function to the target population, including the range of disease severity over which the function was estimated and the potential for systematic differences in the populations that could impact on the health state utility values." (Section 3.2.5)

With this recommendation in mind, further interrogation of the dataset used by Longworth et al. was undertaken to assess its comparability to the TAGS trial. Longworth et al. primarily considers patients with multiple myeloma (n=572 of 771), as well as patients with breast or lung cancer (i.e. no patients with a gastrointestinal cancer). For the multiple myeloma cohort, patients were taken from the VISTA trial – a Phase III randomised open-label trial for newly-diagnosed patients. In the other two populations (breast and lung cancer), real-world data were collected from the Vancouver Cancer Clinic.

In the TAGS trial, all patients had heavily pre-treated (i.e. two or more previous lines of therapy) metastatic gastric cancer, and approximately 73% were male and mean age was approximately 64 years.[8] Across the three cohorts in the Longworth et al.

study, 44% were male and mean age was 68 years. Disease stage and treatment history was unclear though the multiple myeloma cohort were newly-diagnosed (hence receiving their first-line treatment). With these differences in populations in mind, the Longworth algorithm was also rejected to inform the economic model.

The other mapping algorithm referenced by Woodcock and Doble was by Khan et al. As well as being in only a non-small-cell lung cancer population, this study was shown to exhibit poor external validity even within a lung cancer population.[9]

However, Servier notes that similarities may be drawn between the population of the TAGS trial and the SIRFLOX trial used to inform the algorithm by Marriott et al. SIRFLOX was an RCT of patients with previously untreated metastatic colorectal cancer, with 68% of patients that were male, and a mean age of 62 years.[5]

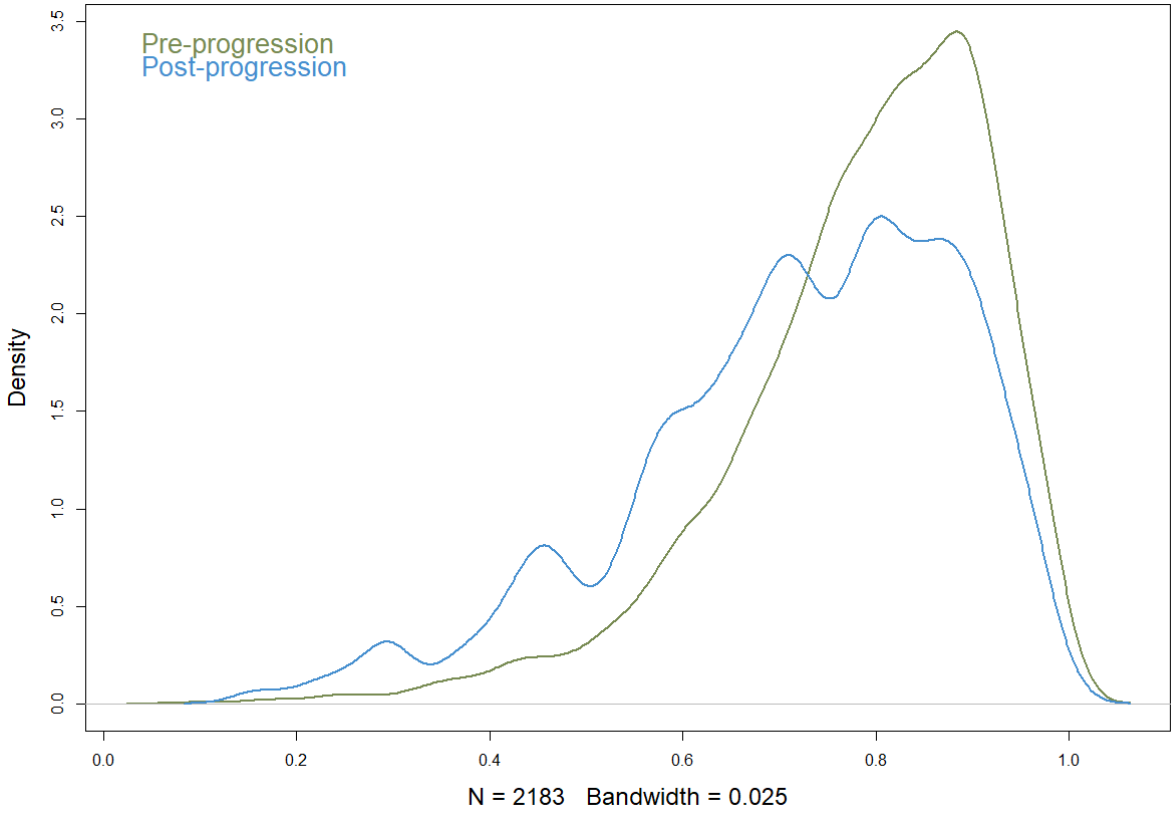
Consequently, Servier has re-run the base-case utility analysis using the Marriot mixed-effects algorithm (highlighted by the authors in the publication abstract as the better-performing algorithm of the two presented). The output of this analysis is provided in Table 1, with a corresponding density plot showing utility values by progression status presented in Figure 1. The pre-progression utility was 0.720 and the post-progression utility was 0.789.

Table 1: Statistical output

```
summary(Utilitygee1_m)

Call:
geeglm(formula = Utility2 ~ prog_state - 1, data = qol_data,
    id = USUBJID, waves = ADY, corstr = gee.costr)

 Coefficients:
               Estimate Std.err  Wald Pr(>|W|)
prog_statePost  0.71979 0.01350  2843   <2e-16 ***
prog_statePre   0.78944 0.00632 15604   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Estimated Scale Parameters:
            Estimate Std.err
(Intercept)   0.0201  0.0013

Correlation: Structure = independence Number of clusters:   489   Maximum
cluster size: 42
```

```
mae(Utilitygee1_m$residuals)
[1] 0.11
> rmse(Utilitygee1_m$residuals)
[1] 0.142
> QIC(Utilitygee1_m)
      QIC Quasi Lik     Trace        px
    84.90    -33.16      9.29   3306.00
> Utilitygee1_m$coefficients
prog_statePost  prog_statePre
         0.720          0.789
> Utilitygee1_m$geese$vbeta
          [,1]     [,2]
[1,] 1.82e-04 2.39e-05
[2,] 2.39e-05 3.99e-05
```

Figure 1: Density plot of utility values by progression status



References

1. Woodcock F, Doble B, CANCER 2015 Consortium. Mapping the EORTC-QLQ-C30 to the EQ-5D-3L: An Assessment of Existing and Newly Developed Algorithms. Med Decis Making. 2018;38(8):954–67.

2. Versteegh MM, Leunis A, Luime JJ, Boggild M, Uyl-de Groot CA, Stolk EA. Mapping QLQ-C30, HAQ, and MSIS-29 on EQ-5D. Med Decis Making. 2012 Aug;32(4):554–68.

3. Longworth L, Yang Y, Young T, Mulhern B, Hernández Alava M, Mukuria C, et al. Use of generic and condition-specific measures of health-related quality of life in NICE decision-making: a systematic review, statistical modelling and survey. Health Technol Assess. 2014 Feb;18(9):1–224.

4. Khan I, Morris S. A non-linear beta-binomial regression model for mapping EORTC QLQ- C30 to the EQ-5D-3L in lung cancer patients: a comparison with existing approaches. Health Qual Life Outcomes. 2014 Nov 12;12:163.

5. Marriott E-R, van Hazel G, Gibbs P, Hatswell AJ. Mapping EORTC-QLQ-C30 to EQ-5D-3L in patients with colorectal cancer. J Med Econ. 2017 Feb;20(2):193–9.

6. Kontodimopoulos N, Aletras VH, Paliouras D, Niakas D. Mapping the cancer-specific EORTC QLQ-C30 to the preference-based EQ-5D, SF-6D, and 15D instruments. Value Health. 2009 Dec;12(8):1151–7.

7. Longworth L, Rowen D. NICE DSU Technical Support Document 10: The Use of Mapping Methods to Estimate Health State Utility Values [Internet]. 2011 [cited 2019 Aug 1]. Available from: http://nicedsu.org.uk/wp-content/uploads/2016/03/TSD-10-mapping-FINAL.pdf

8. Shitara K, Doi T, Dvorkin M, Mansoor W, Arkenau H-T, Prokharau A, et al. Trifluridine/tipiracil versus placebo in patients with heavily pretreated metastatic gastric cancer (TAGS): a randomised, double-blind, placebo-controlled, phase 3 trial. Lancet Oncol. 2018;19(11):1437–48.

9. Gregory J, Dyer M, Hoyle C, Mann H, Hatswell AJ. How should we validate utility mapping algorithms before use? An example in non-small cell lung cancer. In Barcelona, Spain; 2018 [cited 2019 Aug 1]. Available from: https://www.ispor.org/docs/default-source/ispor-europe-2018/pcn356-gregory-j-sup-1-sup-dyer-m-sup-2-sup-hoyle-c-sup-2-sup-mann-h-sup-2-sup-strong-u-hatswell-a-u-sup-3-sup-strong-br-sup-1-sup-bresmed-health-solutions-sheffield-uk-sup-2-sup-astrazeneca-cambridge.pdf?sfvrsn=c5fd9a6a_0

NICE National Institute for Health and Care Excellence

Professional organisation submission

Trifluridine–tipiracil for treating metastatic gastric or gastro-oesophageal junction cancer after 2 or more therapies [ID1507]

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on the technology in the context of current clinical practice that is not typically available from the published literature.

To help you give your views, please use this questionnaire. You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 13 pages.

About you	
1. Your name	[redacted]
2. Name of organisation	**NCRI-ACP-RCP-RCR**

3. Job title or position	[redacted]
4. Are you (please tick all that apply):	☒ an employee or representative of a healthcare professional organisation that represents clinicians? ☐ a specialist in the treatment of people with this condition? ☐ a specialist in the clinical evidence base for this condition or technology? ☐ other (please specify):
5a. Brief description of the organisation (including who funds it).	
5b. Do you have any direct or indirect links with, or funding from, the tobacco industry?	**No**
The aim of treatment for this condition	
6. What is the main aim of treatment? (For example, to stop progression, to improve mobility, to cure the condition,	Improve overall survival (and progression free survival) for patients having received 2 or more lines of treatment for non resectable or advanced gastro-oesophageal or gastric adenocarcinoma.

or prevent progression or disability.)	
7. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by x cm, or a reduction in disease activity by a certain amount.)	For patients who have are being treated beyond Improvement in overall survival of greater than or equal to 2 months compared to best supportive care only (this would represent an 66% improvement in survival)
8. In your view, is there an unmet need for patients and healthcare professionals in this condition?	Yes, currently there is no active systemic treatment available for patients following 2 lines of therapy. The current SOC is BSC
What is the expected place of the technology in current practice?	
9. How is the condition currently treated in the NHS?	Following first and second line therapy, the only accepted treatment for patients is best supportive care only
• Are any clinical guidelines used in the treatment of the	Yes, NICE and ESMO guidelines

condition, and if so, which?	
• Is the pathway of care well defined? Does it vary or are there differences of opinion between professionals across the NHS? (Please state if your experience is from outside England.)	The pathway is well defined. It is nationally and internationally accepted that beyond 2nd line therapy, there is no further evidence based 3rd line therapy currently
• What impact would the technology have on the current pathway of care?	For 100 patients relapsing from first line palliative therapy, only 30 (ie 30%) will get second line treatment and following relapse of these patients – 10 (ie 10% of the initial relapses) will be eligible/ fit enough for 3rd line treatment (with trifluridine/ tipiracil) . For this 10% of the treated population, they will require less best supportive care intervention as their QOL will be better maintained on this active treatment.
10. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	
• How does healthcare resource use differ between the technology and current care?	The technology is an out-patient therapy which has been shown to maintain QOL in responsive patients for a longer duration of time than if no treatment was offered. It will most likely postpone and shorten the need to use best supportive care (eg in-patient stays for end of life care, requirement for supportive medicines such as analgesics.)
• In what clinical setting should the technology be	Specialist clinics (tertiary care)

used? (For example, primary or secondary care, specialist clinics.)	
• What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	This is a treatment delivered as an out-patient therapy. Would only need existing facilities
11. Do you expect the technology to provide clinically meaningful benefits compared with current care?	Yes – (explained in section 10)
• Do you expect the technology to increase length of life more than current care?	Yes (as per trial data)
• Do you expect the technology to increase health-related quality of life more than current care?	No deterioration (as per trial)

12. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	No
The use of the technology	
13. Will the technology be easier or more difficult to use for patients or healthcare professionals than current care? Are there any practical implications for its use (for example, any concomitant treatments needed, additional clinical requirements, factors affecting patient acceptability or ease of use or additional tests or monitoring needed.)	The technology is given as an oral tablet with no additional tests or scans above standard of care. No special equipment is required. The technology is similar to use as the current standard of care which is best supportive care

14. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	Stopping treatment will be dependent on radiological response and tolerability of medication.
15. Do you consider that the use of the technology will result in any substantial health-related benefits that are unlikely to be included in the quality-adjusted life year (QALY) calculation?	
16. Do you consider the technology to be innovative in its potential to make a significant and substantial impact on health-related benefits and how might it	Yes, currently there is no active treatment available for patients who have progressed through 2 lines of therapy in the advanced gastric setting. This technology significantly improves survival compare to the current standard of care (BSC)

improve the way that current need is met?	
• Is the technology a 'step-change' in the management of the condition?	Yes, the trial is the first global study to show efficacy (survival benefit) in this indication
• Does the use of the technology address any particular unmet need of the patient population?	Yes, currently there is no active treatment available for patients who have progressed through 2 lines of therapy in the advanced gastric setting. This technology is also given orally which is an option not available in either first or second line treatment.
17. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	Toxicities have not been shown to affect the management of this condition/ OOL
Sources of evidence	
18. Do the clinical trials on the technology reflect current UK clinical practice?	Yes, trial was a comparison between the test drug and BSC (which is the current SOC)

• If not, how could the results be extrapolated to the UK setting?	
• What, in your view, are the most important outcomes, and were they measured in the trials?	Overall survival and QOL Both were measured in the trial
• If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes?	
• Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	No
19. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	no

20. How do data on real-world experience compare with the trial data?	No experience as yet
Equality	
21a. Are there any potential equality issues that should be taken into account when considering this treatment?	no
21b. Consider whether these issues are different from issues with current care and why.	
Key messages	

24. In up to 5 bullet points, please summarise the key messages of your submission.

- Unmet need
- First global RCT to demonstrate survival benefit (66% improvement in overall survival compared to current SOC)
- Ease of administration (oral formulation)
- Safe and tolerable
- Requires no extra resource requirements in the real world management of this condition

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

……

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

x☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

……

NICE National Institute for Health and Care Excellence

Clinical expert statement

Trifluridine–tipiracil for treating metastatic gastric or gastro-oesophageal junction cancer after 2 or more therapies [ID1507]

Thank you for agreeing to give us your views on this technology and its possible use in the NHS.

You can provide a unique perspective on the technology in the context of current clinical practice that is not typically available from the published literature.

To help you give your views, please use this questionnaire. You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this expert statement

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 13 pages.

About you	
1. Your name	**Elizabeth Smyth**
2. Name of organisation	**Cambridge University Hospitals NHS Foundation Trust**

3. Job title or position	**Consultant Medical Oncologist**
4. Are you (please tick all that apply):	☐ an employee or representative of a healthcare professional organisation that represents clinicians? ☒ a specialist in the treatment of people with this condition? ☒ a specialist in the clinical evidence base for this condition or technology? ☐ other (please specify):
5. Do you wish to agree with your nominating organisation's submission? (We would encourage you to complete this form even if you agree with your nominating organisation's submission)	☒ yes, I agree with it ☐ no, I disagree with it ☐ I agree with some of it, but disagree with some of it ☐ other (they didn't submit one, I don't know if they submitted one etc.)
6. If you wrote the organisation submission and/ or do not have anything to add, tick here. (If you tick this box, the rest of this form will be deleted after submission.)	☐ yes

The aim of treatment for this condition	
7. What is the main aim of treatment? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability.)	The main aim is to stop progression and to improve overall survival. Maintaining quality of life is also important.
8. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by x cm, or a reduction in disease activity by a certain amount.)	When we treat patients with gastroesophageal cancer with second line chemotherapy (paclitaxel or irinotecan), we do so based on clinical trials which show a median survival benefit of 6 weeks (Kang et al, Ford et al). The TAGS trial which assessed trifluridine tipuracil compared to placebo in the third line setting improved survival by a median of >2 months. Therefore this is equivalent to or better than second line chemotherapy and is meaningful. The radiological response rate data is not so important, in second line treatment we expect response rates of ~7% for taxanes (Ford et al).
9. In your view, is there an unmet need for patients and healthcare professionals in this condition?	Yes, there are no other treatments which have level 1 evidence (randomised control trial) in the third line setting in a global population. This fulfils the need for an evidence based treatment.
What is the expected place of the technology in current practice?	

10. How is the condition currently treated in the NHS?	In third line most patients are not treated. For example approximately 30-40% are treated second line, and <15% third line. This could be with a taxane or irinotecan (whatever was not used in second line, but this is not very evidence based, see above) Also clinical trials could be considered.
• Are any clinical guidelines used in the treatment of the condition, and if so, which?	Yes, European Society of Medical Oncology Guidelines, or NICE guidelines. NICE guidelines do not discuss third line chemotherapy due to prior lack of evidence base.
• Is the pathway of care well defined? Does it vary or are there differences of opinion between professionals across the NHS? (Please state if your experience is from outside England.)	The pathway is well defined, but uptake of second line and subsequent chemotherapy may vary.
• What impact would the technology have on the current pathway of care?	For patients who are fit and currently have no treatment options after second line chemotherapy this will offer another choice.
11. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	Yes

• How does healthcare resource use differ between the technology and current care?	It depends on whether one compares to no treatment or treatment with non-evidence based chemotherapy. If compare to no treatment there will be one extra visit per month. If compared to taxane or irinotecan chemotherapy there will be 1-3 less visits per month.
• In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.)	Any hospital with oncologists
• What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	None, oncologists are already familiar with this drug through using it in colon cancer.
12. Do you expect the technology to provide clinically meaningful benefits compared with current care?	Yes, absolutely, a two month survival benefit is very meaningful in this context.
• Do you expect the technology to increase length of life more than current care?	Yes

• Do you expect the technology to increase health-related quality of life more than current care?	The QoL data from the TAGS study suggests that treatment will prevent a decline in QoL
13. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	This is not a biomarker selected drug and the treatment seemed equally effective in all subgroups based on the TAGS trial results.
The use of the technology	
14. Will the technology be easier or more difficult to use for patients or healthcare professionals than current care? Are there any practical implications for its use (for example, any concomitant treatments needed, additional clinical requirements, factors	This is a very easy to use treatment. Only one visit per month is required to outpatient clinic. There are no infusions and the patient takes the tablets at home. Blood tests are only needed once per month. As above if compared to no treatment there is a slight increase in health service use (for a defined benefit), however if compared to other commonly used chemotherapy regimens there will be a decrease of health service use.

affecting patient acceptability or ease of use or additional tests or monitoring needed.)	
15. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	Patients will start treatment if their cancer has grown after second line chemotherapy, and stop if their cancer grows on this treatment. This will require a CT scan.
16. Do you consider that the use of the technology will result in any substantial health-related benefits that are unlikely to be included in the quality-adjusted life year (QALY) calculation?	No
17. Do you consider the technology to be innovative in its potential to make a significant and substantial	Yes, there has not been any high quality evidence of a treatment which improved survival for patients in this setting before.

impact on health-related benefits and how might it improve the way that current need is met?	
• Is the technology a 'step-change' in the management of the condition?	Yes
• Does the use of the technology address any particular unmet need of the patient population?	Yes, need for effective treatment which can improve survival.
18. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	The side effects associated with the treatment are usually quite mild. There are some cases of low white cells but this does not usually result in infection. The side effects are usually managed by the oncologist who are familiar with this drug. The quality of life data does not suggest that there is an overall negative effect of side effects on patient quality of life.
Sources of evidence	

19. Do the clinical trials on the technology reflect current UK clinical practice?	Yes, the patients in the trial were typical of a similar gastric and gastroesophageal population in terms of demographics and previous treatments.
• If not, how could the results be extrapolated to the UK setting?	N/A
• What, in your view, are the most important outcomes, and were they measured in the trials?	Overall survival is the most important outcome. This was measured in the standard way.
• If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes?	N/A
• Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	No
20. Are you aware of any relevant evidence that might	No

not be found by a systematic review of the trial evidence?	
21. How do data on real-world experience compare with the trial data?	None available on real world experience as this has not been licensed in this indication before. But real world data in colon cancer reflect the trial data well.
Equality	
22a. Are there any potential equality issues that should be taken into account when considering this treatment?	No
22b. Consider whether these issues are different from issues with current care and why.	No
Topic-specific questions	
23. Currently what treatment would be used for advanced or metastatic oesophago-gastric	If chemotherapy is used it is either a taxane (docetaxel or paclitaxel) or irinotecan or irinotecan and 5FU (FOLFIRI). The clinical trial data to support this is not high quality. However, best supportive care is the treatment for the majority of patients. Best supportive care could include stents, nutritional advice,

cancer after 2 prior lines of treatment (i.e. for third line treatment)? a. Would a chemotherapy regimen be used as a third line treatment in combination with best supportive care? If so, what chemotherapy regimens are commonly used? b. What treatments would be provided as part of best supportive care?	analgesia, radiotherapy and other treatments as needed. Most patients would have visits from palliative care (hospice specialists).
24. Based on your clinical experience, would treatment outcomes for trifluridine-tipiracil be expected to differ for people who have not previously had ramucirumab?	No, there is no reason why prior ramucirumab would alter the outcome for trifluridine-tipiracil. They work on completely different pathways and cross resistance would not be expected.

25. Based on your clinical experience, what is the current average survival time for people with advanced or metastatic oesophago-gastric cancer who have had 2 prior lines of treatment?	Without treatment usually 3-4 months which is consistent with what was seen in the TAGS trial.

Key messages

26. In up to 5 bullet points, please summarise the key messages of your statement.

- Before now, there was no high quality evidence to support use of third line chemotherapy in GC patients
- Trifluridine-tipiracil improves survival by a median of >2 months based a global, phase III randomised trial
- Trifluridine-tipiracil is well tolerated and only requires once hospital visit per month
- Oncologists are very familiar with trifluridine-tipiracil through use in colorectal cancer and can manage side effects well.
- Trifluridine-tipiracil appears to prevent deterioration in quality of life for treated patients.

Thank you for your time.

Please log in to your NICE Docs account to upload your completed statement, declaration of interest form and consent form.

……

Your privacy

NICE National Institute for Health and Care Excellence

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

……

NICE National Institute for Health and Care Excellence

Clinical expert statement

Trifluridine–tipiracil for treating metastatic gastric or gastro-oesophageal junction cancer after 2 or more therapies [ID1507]

Thank you for agreeing to give us your views on this technology and its possible use in the NHS.

You can provide a unique perspective on the technology in the context of current clinical practice that is not typically available from the published literature.

To help you give your views, please use this questionnaire. You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this expert statement

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 13 pages.

About you	
1. Your name	**Wasat Mansoor**
2. Name of organisation	**Christie Hospital NHS Foundation Trust**

3. Job title or position	**Professor of Medical Oncology, Clinical director of medical oncology, Upper GI lead and GM NETWORK research lead for UGI Cancer research**
4. Are you (please tick all that apply):	☐ an employee or representative of a healthcare professional organisation that represents clinicians? x☐ a specialist in the treatment of people with this condition? x☐ a specialist in the clinical evidence base for this condition or technology? ☐ other (please specify):
5. Do you wish to agree with your nominating organisation's submission? (We would encourage you to complete this form even if you agree with your nominating organisation's submission)	x☐ yes, I agree with it ☐ no, I disagree with it ☐ I agree with some of it, but disagree with some of it ☐ other (they didn't submit one, I don't know if they submitted one etc.)
6. If you wrote the organisation submission and/ or do not have anything to add, tick here. (If you tick this box, the rest of this form will be deleted after submission.)	☐ yes

The aim of treatment for this condition	
7. What is the main aim of treatment? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability.)	The main aim of the treatment Tas102 (Lonsurf) in patients with stage IV adenocarcinoma of the gastric/ gastro-oesophageal junction (GEJ) is to prolong survival and to delay a patients deterioration of quality of life (QOL) during this time.
8. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by x cm, or a reduction in disease activity by a certain amount.)	For this condition, beyond second line treatment, a median survival advantage of greater than 2 months when compared to active symptom support is significant.
9. In your view, is there an unmet need for patients and healthcare professionals in this condition?	Yes, there are currently no licensed or commissioned active anti-cancer therapies for this condition beyond second line treatment. Furthermore, there are very few therapies that don't require intravenous administration for these patients.
What is the expected place of the technology in current practice?	

10. How is the condition currently treated in the NHS?	Beyond 2nd line , this condition is only treated with active symptom support (ie best supportive care)
• Are any clinical guidelines used in the treatment of the condition, and if so, which?	Yes, NICE guidelines, ESMO Guidelines
• Is the pathway of care well defined? Does it vary or are there differences of opinion between professionals across the NHS? (Please state if your experience is from outside England.)	The pathway is well defined. The consensus of opinion is that beyond 2nd line treatment there are currently no options available. At this stage it is largely accepted globally that active symptom support is the only standard of care. I sit in key meetings in the UK where a wide ranging spectrum of gastric oncologists meet. Similarly, I also sit in key international meetings where I have gauged opinion. My opinion is based on these interactions and general literature.
• What impact would the technology have on the current pathway of care?	If a 3rd line treatment was available, the impact would not be large based on numbers of patients eligible for this treatment as many patients are not fit at this stage in their pathway (approximately10 to 15% of patients commencing first line therapy would be fit enough. However, it may well have a significant change on the strategies employed by patients and physicians regarding how we try to preserve a patients fitness while on 1st and 2nd line therapies.
11. Will the technology be used (or is it already used) in	Yes – out patients.

the same way as current care in NHS clinical practice?	
• How does healthcare resource use differ between the technology and current care?	It does not
• In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.)	Specialist clinics
• What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	Nil
12. Do you expect the technology to provide clinically meaningful benefits compared with current care?	yes
• Do you expect the technology to increase	Yes (Median Overall Survival 5.7 months vs 3.6 months for best supportive care)

length of life more than current care?	
• Do you expect the technology to increase health-related quality of life more than current care?	Yes, the TAGS trial demonstrated that lonsurf prevents deterioration of QOL for a longer period than best supportive care
13. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	Not defined
The use of the technology	
14. Will the technology be easier or more difficult to use for patients or healthcare professionals than current care? Are there any practical implications for its use (for example, any concomitant	This technology is easier to apply and use than best supportive care (current technology being used). In general, no concomitant medications are required with lonsurf and this can be delivered in a standard oncology clinic. Administration of this drug may reduce the amount of supportive treatments required and the expensive expertise this requires to be administered (eg symptoms support teams, in patient admissions, disruption to patient and carers lives etc).

treatments needed, additional clinical requirements, factors affecting patient acceptability or ease of use or additional tests or monitoring needed.)	Lonsurf does not require any specialised tests or extra hospital visits.
15. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	Main rules applied to use of Lonsurf in this setting are the requirement to do CT scans to assess for progression (1 extra CT scan for the median patient)
16. Do you consider that the use of the technology will result in any substantial health-related benefits that are unlikely to be included in the quality-adjusted life year (QALY) calculation?	Yes, it is likely that responding patients will require LESS symptom support than patients only receiving best supportive care.
17. Do you consider the technology to be innovative in	As described above, this technology addresses an unmet need where there is currently no anti-cancer therapy available beyond two lines of treatment for the palliative patient. As an unintended benefit of this, It

its potential to make a significant and substantial impact on health-related benefits and how might it improve the way that current need is met?	is now being recognized that as more lines of treatment and more effective options become available, preserving our patients QOL is becoming more important to ensure they are well enough to benefit from the survival. This will have substantial health-related benefits in the patients pathway.
• Is the technology a 'step-change' in the management of the condition?	Yes, previously unaddressed unmet need
• Does the use of the technology address any particular unmet need of the patient population?	As above
18. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	It has been demonstrated within the TAGS trial that QOL does not deteriorate and toxicity profile of Lonsurf are managed without ant alterations in the management of the condition.
Sources of evidence	

19. Do the clinical trials on the technology reflect current UK clinical practice?	Yes
• If not, how could the results be extrapolated to the UK setting?	NA
• What, in your view, are the most important outcomes, and were they measured in the trials?	Overall survival and QOL measure – both were collected.
• If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes?	NA
• Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	No
20. Are you aware of any relevant evidence that might	No

not be found by a systematic review of the trial evidence?	
21. How do data on real-world experience compare with the trial data?	Toxicity data is similar in the real world as per trial. Real world efficacy data is not available as yet.
Equality	
22a. Are there any potential equality issues that should be taken into account when considering this treatment?	No
22b. Consider whether these issues are different from issues with current care and why.	NA
Topic-specific questions	
23. Currently what treatment would be used for advanced or metastatic oesophago-gastric	Best supportive care (symptom support). There is no evidence to support use of any anti-cancer therapy in 3rd line apart from the TAGS data.

cancer after 2 prior lines of treatment (i.e. for third line treatment)? a. Would a chemotherapy regimen be used as a third line treatment in combination with best supportive care? If so, what chemotherapy regimens are commonly used? b. What treatments would be provided as part of best supportive care?	As best supportive care, treatments such as analgesics, haematological support (blood transfusions), radiation therapy, anti-coagulation therapies, bone modifying therapies, anti-depressant therapies, anti-anorexia medications are all used to varying degrees either in sequence or concurrently and often as in patient care.
24. Based on your clinical experience, would treatment outcomes for trifluridine-tipiracil be expected to differ for people who have not previously had ramucirumab?	Possibly – ramucirumab in combination or without paclitaxel results in better preservation of patients reserve to tolerate Lonsurf than if conventional therapies like docetaxel, irinotecan etc are used at 2nd line. Therefore, the patient with commence third line treatment with Lonsurf in a better state, so, may do better (however this observation is only anecdotal!)

25. Based on your clinical experience, what is the current average survival time for people with advanced or metastatic oesophago-gastric cancer who have had 2 prior lines of treatment?	Approximately 2-3 months

Key messages

26. In up to 5 bullet points, please summarise the key messages of your statement.

- Lonsurf significantly improves overall survival
- Lonsurf does not result in a deterioration in compared to best supportive care and may result in less intensive best supportive care requirments
- Lonsurf is easily administered as an oral reagent with minimal disruption to the patient or any special requirements for the NHS
- Lack of treatment beyond two lines of treatments for these patients is an important unmet need which Lonsurf meets effectively
- For patients with a median survival of approximately 3 on best supportive care, an extension of life by a further 40% is significant to them. This important psychologically for them.

Thank you for your time.

Please log in to your NICE Docs account to upload your completed statement, declaration of interest form and consent form.

NICE National Institute for Health and Care Excellence

……

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

……

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Questions for clinical expert

Trifluridine–tipiracil for treating metastatic gastric or gastro-oesophageal junction cancer after 2 or more therapies [ID1507]

1. Treatment pathway and current treatment options

1.1 What treatment is currently used as third-line treatment for metastatic gastric cancer? Would any chemotherapy regimens be used?

Clinical expert 1: a very small proportion of people are likely to get third-line therapy (approximately <15%) compared with around 30-40% who will receive second-line treatment. Current third-line treatment is in line with the ESMO Guideline for gastric cancer which recommends irinotecan or a taxane (docetaxel, paclitaxel). However, because paclitaxel and ramucirumab are generally used for second-line treatment, irinotecan is most commonly used as a third-line treatment.

Clinical expert 2: currently best supportive care (BSC) is standard of care for third-line treatment. There may be some people using a chemotherapy but almost everyone would be using BSC.

1.2 Is best supportive care the most appropriate comparator? Is this given with palliative intent? What treatments would be used as part of best supportive care?

Clinical expert 1: gastric cancer causes a quick decline in strength therefore many patients, even at second-line, are often not fit for active treatment. In this case BSC is offered to control symptoms (these include difficulty swallowing, pain, obstructions and most have nutritional requirements). However, BSC is also commonly used alongside active treatment with chemotherapy and includes several interventions such as treatment to strengthen bones, radiotherapy, opioids and regular nutritional

advice. Palliative care is usually offered alongside BSC. People who progress to third-line therapy are often more likely to have disease that is responsive to chemotherapy and may be younger and fitter, however these patients will still experience several symptoms and require BSC.

Clinical expert 2: yes, agree that BSC is the most appropriate comparator and this is given with palliative intent to improve health-related quality of life. This typically consists of treatments for pain relief, appetite support etc. in the early stages of disease but also includes other treatments for nausea, vomiting and inpatient stay for advanced disease towards the end of life.

2. Generalisability of trial population

2.1 **PRIORITY**: The full trial population in TAGS included people from the EU (80%) and Japan (14%). The most recent census data for England & Wales suggests 7.5% were Asian (mostly Indian or Pakistani). In your clinical opinion, is the full trial population or the subgroup from Europe most relevant to people in England?

Clinical expert 1: in the TAGS trial, recruitment in Asia was capped at 15% therefore this should not have an impact on the trial results. The reasons for different outcomes in people from East Asia are not fully known but compared with the UK where around 70% of gastric cancers tend to be at the top of the stomach and around the oesophagus; gastric cancers in East Asia tend to be located at the lower portion of the stomach. Furthermore, cancers are often detected earlier in Asia so patients tend to be fitter and have had more lines of treatment. The proportion of people from Japan in the TAGS trial is small therefore it is unlikely to have any impact on outcomes and the results from the full trial population should be generalisable to the NHS in England.

Clinical expert 2: subgroup analyses from TAGS showed that the Western population benefitted just as well as the population from Japan, therefore the full trial population is generalisable to the NHS in England.

2.2 In the trial, 63% of people had previously had 3 or more previous lines of chemotherapy. In your clinical experience does this reflect the population in England with metastatic gastric cancer?

Clinical expert 1: clinical trials do not always reflect clinical practice and this is the case here because we would expect a much smaller proportion to have had 3 previous lines of treatment in England. Although this may change over time as more treatments become available. However, based on the forest plots from the TAGS trial there is no evidence that the number of previous therapies has an impact on overall survival therefore it would be acceptable to generalise the full trial population to the NHS in England.

Clinical expert 2: no, the proportion having 3 or more previous lines of chemotherapy in the trial population is higher than what would be expected in the UK (less than 5%). This is because there are accepted first and second-line chemotherapy treatment options but third-line treatment is rare (although sometimes chemotherapy may be used to re-challenge disease depending on previous response). Overall, the trial population is more heavily treated compared with the NHS in England; however, this is not likely to make the results less valid.

3. Previous treatment with ramucirumab (not used in NHS)

3.1 Do you have any clinical experience with ramucirumab? Please describe.

Clinical expert 1: yes, I have clinical experience with ramucirumab as this is standard treatment in the ESMO guideline and although it's not used in the NHS, I have experience using it in private healthcare and in clinical trial settings.

Clinical expert 2: yes, I have clinical experience with ramucirumab within the phase 3 RAINFALL trial which examined first-line treatment (with ramucirumab) and in private practice.

3.2 **PRIORITY** In your clinical opinion, how might people be selected to receive ramucirumab? In particular:

a) Would you expect these people to be healthier, on average, than people not selected to have ramucirumab?

b) Would you expect these people to be more responsive to previous lines of chemotherapy compared with people not selected to have ramucirumab?

Clinical expert 1: if ramucirumab was available in the NHS, it would certainly be used as a second-line treatment because it's associated with very little toxicity and is considered a clinically effective treatment. The only selection criteria would be the use of well-defined contraindications which include recent myocardial infarction or stroke, but otherwise if it were available, it would be used. There would be no difference in the characteristics of people who have or have not had ramucirumab. The only difference is whether ramucirumab is publicly funded in the regions included in the trial but this would not have an impact on generalisability.

Clinical expert 2: no, I wouldn't expect prior ramucirumab to have an impact on response to treatment or expect those having ramucirumab to be healthier. It's generally given in combination with paclitaxel, but this is not much more effective compared with paclitaxel alone (which is what people in the UK would be offered).

3.3 **PRIORITY** The trial included subgroup analyses of people who had not previously had ramucirumab. In your clinical opinion, how might selection of patients having ramucirumab impact the results of this subgroup analysis? In your clinical opinion does prior ramucirumab affect overall survival?

Clinical expert 1: no, previous ramucirumab would not have an impact on treatment outcomes such as overall survival.

Clinical expert 2: no, previous ramucirumab should not have an impact on treatment outcomes.

3.4 **PRIORITY** Would you expect that patients who have had prior ramucirumab would respond to a) trifluridine–tipiracil and/or b) any treatment differently than those who have not? Why might this be the case?

Clinical expert 1: no, previous ramucirumab would not have an impact on response to treatment. People with disease that is responsive to chemotherapy would

generally have improved outcomes at all lines of therapy. Additionally, there is no evidence that use of ramucirumab changes the biology of the cancer to make it more or less sensitive to chemotherapy.

Clinical expert 2: there is no biological reason to expect a difference in the estimated treatment effect based on previous ramucirumab.

3.5 **PRIORITY** Do patients who have had prior ramucirumab have a worse prognosis than those who don't, once they stop treatment? How quickly would you expect disease to progress compared with people who have not had ramucirumab?

Clinical expert 1: no, there is no evidence of quicker progression after ramucirumab is stopped and ramucirumab is unlikely to influence post progression treatment. Overall there is no data to support different overall survival based on previous ramucirumab therefore selection bias is unlikely (this has been the case in colorectal cancer for which trifluridine–tipiracil is also indicated).

Clinical expert 2: yes, progression may be quicker once treatment stops but this wouldn't be due to prior ramucirumab, it's more likely to be due to the stage of disease and because life expectancy is less than 3 months. Prior ramucirumab doesn't cause rapid disease progression, but it is used at a later line of treatment when disease is more likely to be advanced.

4. Experience with trifluridine–tipiracil

4.1 Do you have any clinical experience with trifluridine–tipiracil? Please describe.

Clinical expert 1: yes, I've used trifluridine–tipiracil to treat colorectal cancer in a similar setting (after 2 prior lines of therapy) and also used it in trial settings to treat gastric cancer. Overall it's well tolerated and improves overall survival.

Clinical expert 2: yes, I've used trifluridine–tipiracil trial settings.

5. Survival and subsequent treatment

5.1 **PRIORITY:** In your clinical experience, what is the current average expected survival time for people with metastatic gastric cancer after 2 prior lines of treatment without trifluridine–tipiracil?

Clinical expert 1: in current practice the average expected survival would be less than 6 months. Expected survival is often associated with sensitivity to chemotherapy (a survival benefit is more likely with disease that has responded to previous chemotherapy). In this group of people with disease that has responded well to previous chemotherapy, survival may be over 6 months. Without treatment, survival is around 2 to 3 months.

Clinical expert 2: the current average expected survival time would be around 2 to 3 months with BSC alone.

5.2 **PRIORITY:** On average, how many people would you expect to be alive after 6 months, 1 year and 2 years? Would you expect anyone to be alive after 2 years?

Clinical expert 1: on average, survival at third-line would be less than 6 months. In current practice at third-line treatment I would expect around 20 to 25% to survive 6 months and around 10 to 15% to survive for 1 year. It may be useful to validate the data using the COUGAR-02 trial which examined second-line treatment – these results suggest around 15% survive at 1 year and 10% at 2 years.

Clinical expert 2: This would be approximately 10% at 6 months, 7% at 1 year and around 5% at 2 years, but this is uncertain. There may be some outliers, for example in the trial 1 person was having trifluridine–tipiracil for 16 months.

5.3 In clinical practice, what treatment is likely to be used after trifluridine–tipiracil? Would any chemotherapies be used or would best supportive care be given?

Clinical expert 1: if trifluridine–tipiracil was available it's likely that current treatment options would be delayed and used after trifluridine–tipiracil. Therefore, it's likely that irinotecan and FOLFIRI would be used as a fourth-line treatment option if patients

were fit, however you might expect that the majority of patients would not be fit. BSC would always be offered, independent of chemotherapy.

Clinical expert 2: if trifluridine–tipiracil was available current treatment options (that is, BSC) would be used as a fourth-line treatment option.

6. End of life

6.1 In your clinical opinion, what extension to life would you consider to be clinically meaningful for the population expected to be treated with trifluridine–tipiracil?

Clinical expert 1: a survival benefit of around 2 months would be good but this depends on the how toxic the treatment is. Generally, 2 months may be considered clinically meaningful particularly if it could be achieved with a good quality of life. What's interesting with the trifluridine–tipiracil data is that it gives the same or better benefit than second-line chemotherapy or ramucirumab monotherapy in the second-line, compared to BSC (these all give 6 weeks median overall survival benefit).

Clinical expert 2: a survival benefit of around 2 months would be about 40% life improvement and this is clinically meaningful. If it were 4 to 6 weeks, this may be considered be less meaningful. Health-related quality of life is fundamental and chemotherapy can help to preserve this.

Confidential until published

Trifluridine-tipiracil for treating metastatic gastric or gastro-oesophageal junction cancer after 2 or more therapies. A Single Technology Appraisal.

Produced by	School of Health and Related Research (ScHARR), The University of Sheffield
Authors	Matt Stevenson, Professor of Health Technology Assessment, ScHARR, University of Sheffield, Sheffield, UK
	Sue Harnan, Senior Research Fellow, ScHARR, University of Sheffield, Sheffield, UK
	Shijie Ren, Senior Research Fellow, ScHARR, University of Sheffield, Sheffield, UK
	Andrew Metry, Research Associate, ScHARR, University of Sheffield, Sheffield, UK
	Martin Orr, Research Associate, ScHARR, University of Sheffield, Sheffield, UK
	Fiona Campbell, Research Fellow, ScHARR, University of Sheffield, Sheffield, UK
	Mark Clowes, Information Specialist, ScHARR, University of Sheffield, Sheffield, UK
	Raj Sripadam, Consultant Clinical Oncologist, Clatterbridge Cancer Centre, NHS Foundation Trust, Bebington, Wirral CH63 4JY
Correspondence Author	Matt Stevenson, Professor of Health Technology Assessment, ScHARR, University of Sheffield, Sheffield, UK
Date completed	Date completed (21/08/2019)

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 12/93/72.

Declared competing interests of the authors

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

Acknowledgements

We would like to thank Dr Jamal Hayat for clinical advice relating to this project. We would also like to thank Paul Tappenden, ScHARR, for providing comments on the draft report and Andrea Shippam, Programme Manager, ScHARR, for providing administrative support and in preparing and formatting the report.

Rider on responsibility for report

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

This report should be referenced as follows:

Stevenson M, Harnan S, Ren S, Metry A, Orr M, Campbell F, Clowes M, Sripadam R. Trifluridine-tipiracil for treating metastatic gastric or gastro-oesophageal junction cancer after 2 or more therapies.: A Single Technology Appraisal. School of Health and Related Research (ScHARR), 2019.

Contributions of authors

Matt Stevenson and Andrew Metry critiqued the health economic analysis submitted by the company. Sue Harnan and Fiona Campbell summarised and critiqued the clinical effectiveness data reported within the company's submission. Shijie Ren and Martin Orr critiqued the statistical aspects of the submission. Mark Clowes critiqued the company's search strategy. All authors were involved in drafting and commenting on the final report.

CONTENTS

List of tables

List of figures

List of Boxes

Abbreviations

Adverse events	AEs
AF	Acceleration Factor
AFT	Accelerated Failure Time
AIC	Akaike Information Criterion
ASCO	American Society of Clinical Oncology
ASCO GI	The American Society of Clinical Oncology Gastrointestinal Cancers Symposium
AWMSG	All Wales Medicines Strategy Group
BIC	Bayesian Information Criterion
BSA	Body Surface Area
BSC	Best Supportive Care
CEAC	Cost-Effectiveness Acceptability Curve
CS	Company Submission
CT	Computed Tomography
DSU	Decision Support Unit
eMIT	electronic Market Information Tool
EORTC QLQ-C30	European Organisation for Research and Treatment of Cancer Quality-of-Life Questionnaire Core 30
EQ-5D-3L	EuroQol 5 dimensions 3 level
ERG	Evidence Review Group
ESMO	European Society of Medical Oncology
GEE	Generalised Estimating Equation
GEJ	Gastro-Oesophageal Junction Cancer
HAS	Haute Autorité de Santé
HER2	Human Epidermal Growth Factor Receptor 2
HERC	Health Economics Research Centre
HR	Hazard Ratio
HRQoL	Health-Related Quality of Life
HTAD	Health Technology Assessment Database
ICER	Incremental Cost Effectiveness Ratio
ISPOR	International Society for Pharmacoeconomics and Outcomes Research
KM	Kaplan-Meier
mGC	metastatic Gastric Cancer
MRU	Medical Resource Use
NHS EED	National Health Service Economic Evaluation Database

NICE	National Institute for Health and Care Excellence
NR	Not Reported
ORR	Overall Response Rate
OS	Overall Survival
PAS	Patient Access Scheme
PD	Progressed Disease
PF	Progression-Free
PFS	Progression-Free Survival
PH	Proportional Hazards
PSM	Partitioned Survival Model
PSSRU	Personal Social Services Research Unit
QALY	Quality-Adjusted Life Year
QIC	Quasi-likelihood under Independence Model Criterion
RCT	Randomised Controlled Trial
ROW	Rest Of World
SACT	Systematic Anti-Cancer Treatment
SLR	Systematic Literature Review
SLV	Statens Legemiddelverk
SMC	Scottish Medicines Consortium
STA	Single Technology Appraisal
TAGS	Trifluridine/tipiracil versus placebo in patients with heavily pre-treated mGC trial
TFT	Trifluridine-Tipiracil
TSD	Technical Support Document
TTD	Time To Treatment Discontinuation

1 SUMMARY

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

The company provided an appropriate description of metastatic gastric and gastro-oesophageal junction cancers, the current practice guidelines regarding lines of treatment and the potential positioning of trifluridine/tipiracil (TFT) (Lonsurf®) in the treatment pathway.

1.2 Summary of clinical effectiveness evidence submitted by the company

The company submitted a systematic literature review which the ERG believes identified all important studies.

The pivotal trial was TAGS, a Phase III, randomised, double-blind, placebo-controlled study in patients with heavily pre-treated metastatic gastric cancer, conducted at 110 sites in 18 countries, comparing 35mg/m2 TFT and BSC with placebo and BSC. The trial reported all key efficacy outcomes listed in the NICE scope. Overall survival (OS) was positively affected by TFT treatment with a hazard ratio (HR) of 0.69 (95% CI: 0.56–0.85) and a difference in median survival of 2.1 months between arms. Analyses adjusted for relevant prognostic factors gave a similar HRs. Progression free survival (PFS) was also positively affected by TFT treatment, with a HR 0.57 (95% CI: 0.47–0.70) and a 0.2 month difference between arms. Small benefits were reported for response rates and duration of response as may be expected given the stage of disease, however, the disease control rate was significantly improved. Health related quality of life was shown to be largely maintained with TFT treatment.

In subgroup analyses, for OS, patients with prior ramucirumab treatment had HR 0.76 (95% CI 0.53 – 1.09) and those without an HR 0.66 (95% CI 0.51 – 0.85). The HR in Japanese patients was 0.77 (95% CI 0.46 to 1.30) compared with 0.68; (95% CI: 0.54 to 0.85) in the rest of the world (ROW). The median survival in the placebo group was 5.9 in the Japanese population and 3.3 months in the ROW. The HR for European patients was 0.67 (95% CI 0.53-0.86), but no Kaplan-Meier plots were provided by the company.

The ERG requested an analysis of European patients without prior exposure to ramucirumab. The company urged caution in the interpretation of this analysis as the TAGS study was stratified on Japan versus the ROW, although the ERG comments that approximately 95% of the ROW group were European.

[redacted]

Key adverse events included nausea, anaemia, decreased appetite, vomiting, diarrhoea, fatigue, neutropenia and asthenia thrombocytopenia. Anaemia and neutropenia were two outcomes where the incidence appeared to be markedly greater in the TFT group compared with the placebo group (anaemia 45% vs 19%; neutropenia 53% vs 4%).

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

The ERG agreed that due to the low quality of the evidence available in the third-line setting, a network meta-analysis would have relied on strong assumptions and was not appropriate for this appraisal. Clinical advice to the ERG also indicated that chemotherapy was infrequently administered as a third-line treatment (approximately 10-15% of patients). The company did not synthesise a single-arm phase II study (EPOC1201) conducted in Japan with the pivotal TAGS study, and the ERG agreed it had low relevance to the decision problem.

Ramucirumab does not have a positive NICE recommendation, which means patients in England will largely be ramucirumab-naïve. There were mixed views from clinical advisors to the ERG and NICE about whether prior ramucirumab treatment would alter prognosis, but agreement that as ramucirumab and TFT work differently there should be no impact on treatment efficacy. In the absence of a strong indication that prior ramucirumab treatment alters prognosis, the ERG assumes there is no impact, though this is uncertain. The ERG therefore prefers an estimate of a HR or Acceleration Factor (AF) from the entire population rather than the non- ramucirumab patients only. However, the ERG notes that the non-ramucirumab population are less heavily pre-treated and their disease duration is shorter than the prior ramucirumab group.

Clinical advice to the ERG and NICE indicated that European patients have the highest generalisability to the decision problem, due to biological and/or treatment pathway differences between Europe and the USA and in particular, Japan. A subgroup analysis of European patients was reported, but no Kaplan-Meier curves were available. Exclusion of the Japanese patients from the full TAGS population leads to an under-representation of Asians compared with the English population (ERG-calculated 1% compared with 7.5% respectively), whilst their inclusion leads to over-representation (14.4%). The generalisability of Japanese patients to the more diverse Asian population in England is also unclear. The ERG concludes that analyses excluding Japanese and USA patients is preferred, although accepts that this breaks the stratification of the TAGS study.

In the requested analysis of European patients with no prior ramucirumab treatment, baseline imbalances in some prognostic characteristics were larger than in the full study population for which efficacy estimates were unadjusted.

Other issues were also noted. There were more gastric patients than would be usual in England, though this was thought unlikely to affect estimates of efficacy. Incorrect dosing within the trial [redacted]. It was not clear whether the discontinuation rules applied in the TAGS study were mandatory, and whether these will be applied in clinical practice in the UK. Monitoring in clinical practice may occur more frequently than in the TAGS study (every 4-6 weeks rather than every 2 months) which may lead to earlier discontinuations and may impact on efficacy, adverse event rates and costs.

1.4 Summary of cost effectiveness submitted evidence by the company

The submitted economic model was clear, relatively simple and generally well programmed, with minor errors amended in the clarification process. The company submitted a partitioned survival model comprising three health states (progression free, post progression, and death). The weekly transitions between health states were inferred via extrapolated PFS and OS curves fitted to data from TAGS. Health-related quality of life (HRQoL) data were collected using the European Organisation for Research and Treatment of Cancer Quality-of-Life Questionnaire Core 30 (EORTC QLQ-C30) within the TAGS trial and then mapped to EuroQol five dimensions three-level (EQ-5D-3L) values using a published algorithm. The time horizon in the base case was 10 years, with discounting of both benefits and costs at 3.5% per annum. The company's base case results suggested that TFT was cost-effective compared with BSC at an incremental cost-effectiveness ratio (ICER) threshold of £50,000 per quality-adjusted life year (QALY) gained. The probabilistic ICER for TFT compared with BSC was £45,314 per QALY gained when treating ramucirumab-naïve patients. The ICER was sensitive to the selected parametric survival models fitted to data from combinations of geographical region and prior ramucirumab use.

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

A key difference between the approach undertaken by the company and that preferred by the ERG is related to selection of the patient population relevant to England. Based on clinical advice the ERG believes that: prior ramucirumab treatment is unlikely to affect the relative effectiveness of TFT; that there is no strong signal that prior ramucirumab treatment affects prognosis; and that the European subgroup is the most relevant to the decision problem, albeit noting the limitation in breaking stratification.

The ERG also prefers the use of independent curve fits rather than the use of dependent curve fits using a HR or an AF to account for the efficacy of TFT.

The ERG noted the limitations in the study developing a mapping algorithm between the EORTC QLQ-C30 and the EQ5D-3L selected by the company to derive the model's base case utility values as the study involved 48 Greek patients with non-metastatic gastric cancer. The ERG requested the company to apply two alternative mapping algorithms (Versteegh *et al*. and Longworth *et al*.) to assess the impact on the ICER. The company did not consider these algorithms appropriate to inform the model citing differences in the patient populations used to derive the mapping algorithms, and instead provided an analysis calculating utility values from a mapping algorithm by Marriott *et al*.

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

1.6.1 Strengths

The clinical evidence was from a good quality phase III, randomised, double-blind, placebo-controlled study. All key efficacy and safety outcomes were reported. A good quality systematic literature review supports the submission.

The submitted mathematical model was of good quality. The company responded well to the clarification questions raised and provided a revised model and undertook the analyses requested.

1.6.2 Weaknesses and areas of uncertainty

A network meta-analysis was not possible, so the efficacy of TFT compared to other chemotherapy regimens used at third-line is unclear. Clinicians estimate between 10-15% of patients may receive chemotherapy at third-line in England.

It is unknown which subgroup's results, are of most relevance for the purpose of decision making. Subgroups of interest include those based on the prior use of ramucirumab treatment and geographical region. In addition, the TAGS study did not collect EQ-5D data and all mappings between the EORTC QLQ-C30 and the EQ5D-3L have limitations.

1.7 Summary of exploratory and sensitivity analyses undertaken by the ERG

As stated in Section 1.5, the ERG preferred alternative assumptions in the base case on multiple occasions to the company. The ERG explored relationships between both prior ramucirumab treatment and disease prognosis and prior ramucirumab treatment and the relative efficacy of TFT. The relationships were explored in both the whole population and the European cohort of the TAGS trial. The possible permutations resulted in eight scenarios, each of which was explored by fitting alternative parametric survival distributions. The ERG proffered a tentative base case but this could not be evaluated as the data were not available to the ERG.

Based on the analyses provided by the company and the ERG's exploratory analyses the ERG believes that the cost per QALY gained of TFT compared with BSC is likely to be in excess of £50,000. Whilst, the ERG's tentatively preferred scenario could not be evaluated, many component factors such as: using independent curves; assuming that prior ramucirumab use does not affect prognosis; assuming that prior ramucirumab use does not affect the relative treatment effect of TFT; using a European population; and reducing utility values, all increase the ICER. The ERG notes that some of these factors, in isolation, increase the ICER to greater than £50,000 per QALY gained.

2 BACKGROUND

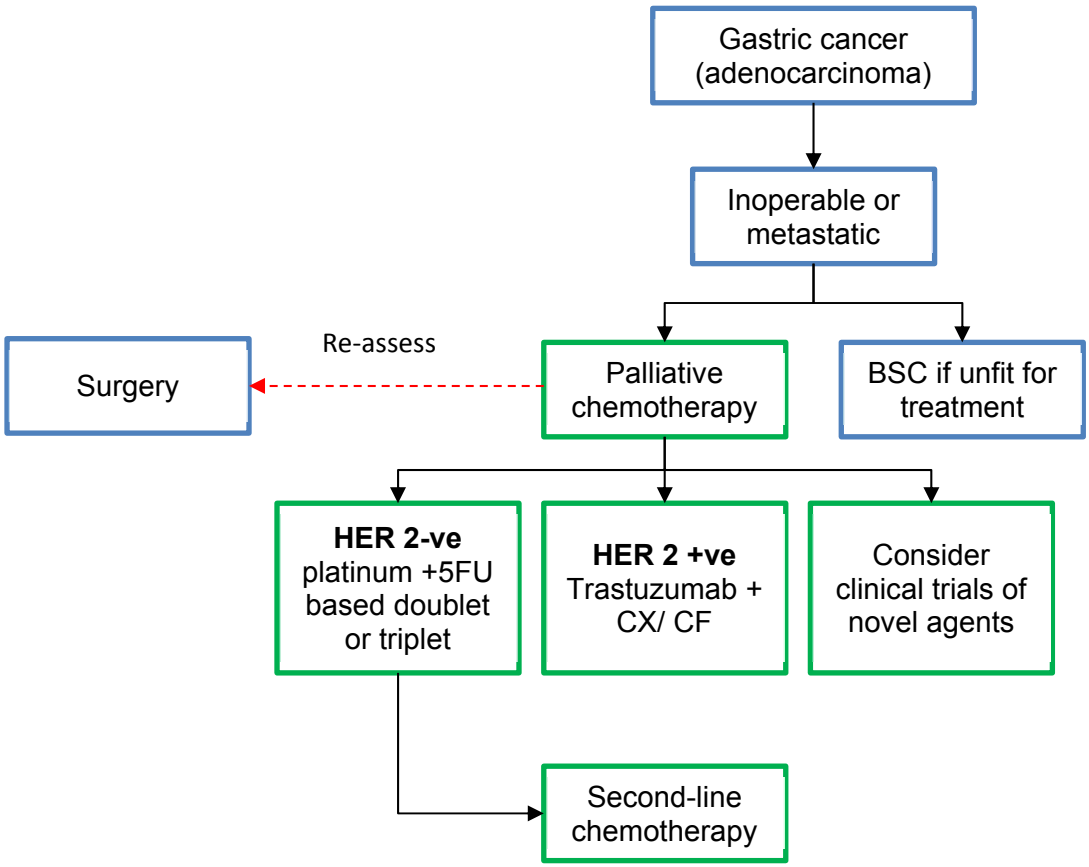
2.1 Critique of company's description of underlying health problem

Gastro-oesophageal cancers are malignant tumours characterised by uncontrolled cell growth in the tissues of the stomach. Such cancers rapidly progress and have significant impacts on patients. A detailed description of the epidemiology, risk factors, prognoses, diagnosis and methods to stage the severity of gastro-oesophageal cancers has been provided within the company submission (CS).[1]

2.2 Critique of company's overview of current service provision

The CS detailed the typical treatment pathway for patients with metastatic gastric cancer, and provided a schematic of current European Society of Medical Oncology (ESMO) guidelines. This has been reproduced in Figure 1, although this diagram appears to indicate that patients who are human epidermal growth factor receptor 2 (HER2) positive would not receive a second line of treatment, which is incorrect. The company stated that the majority of patients (95%) within the UK would receive a doublet regimen rather than a triplet regimen. Second-line treatment chemotherapy is provided after disease progression or recurrence. Few patients are deemed fit enough to receive third-line treatment and few treatments provide a meaningful benefit over best supportive care (BSC). Prognosis is very poor with the average survival time for patients in the third-line setting being less than six months.[2-4] NICE does not currently recommend any third-line treatment for metastatic gastric or gastro-oesophageal junction cancer.

It is worth noting that ramucirumab received marketing approval in 2015 for patients who had received previous chemotherapy although this was not recommended by NICE. However, within the multi-national pivotal study for trifluridine-tipiracil (TFT) (TAGS)[2] a proportion of patients had received ramucirumab, and the level to which these data are generalisable to England is unclear.



Key: BSC, Best supportive care; CF: cisplatin and 5-fluorouracil; CX: cisplatin and capecitabine; ECF: epirubicin, cisplatin and 5-fluorouracil; ECX: epirubicin, cisplatin and capecitabine; EOF: epirubicin, oxaliplatin and 5-fluorouracil; EOX: epirubicin, oxaliplatin and capecitabine; DCF: docetaxel, cisplatin and 5-fluorouracil; ESMO: European Society for Medical Oncology; FOLFIRI: folinic acid, fluorouracil and irinotecan; HER2 +ve/-ve, Human epidermal growth factor receptor 2 negative/positive.

Note: Doublet combinations of platinum and fluoropyrimidines are generally used, but triplet regimen options also include: ECF, ECX, EOF, EOX, DCF or FOLFIRI. Please note this is also the treatment pathway for inoperable advanced disease

Figure 1: The treatment pathway for metastatic gastric cancer provided by the company based on ESMO guidelines

2.3 Critique of company's definition of the decision problem

2.3.1 Population

The population within the decision problem matches that within the NICE scope[5] in considering patients with metastatic gastric cancer (mGC) or gastro-oesophageal junction cancer (GEJ) who have received two previous regimens of treatment. The TAGS study was multi-national which may mean that some sub-group data from the study may be more generalisable to England than the data from the entire study population. The company provided evidence on both a subgroup of the population that had no prior ramucirumab and on the full TAGS study population.

2.3.2 Intervention

The intervention matches that of the final NICE scope,[5] which is the use of TFT, a novel oral cytotoxic chemotherapy, in combination with BSC. TFT currently has a marketing authorisation for use in metastatic colorectal cancer, with an extension for mGC and GEJ expected by November 2019. The dose of TFT is dependent on body surface area (BSA) with a recommended starting dose of 35mg/m^2 administered orally twice daily on days 1-5 and 8-12 of each 28-day cycle. The minimum dose is 20 mg/m^2, with a maximum of 80mg.

2.3.3 Comparators

The comparators listed in the final NICE scope[5] are chemotherapy and BSC. The company have not included chemotherapy in the decision problem, stating that this deviation was based on "*current guidelines, a systematic literature review and expert opinion validating the lack of an evidence-based active chemotherapy option in the third-line setting.*" Clinical advice provided to the ERG supported the company's view that there is no established treatment for mGC or GEJ following two previous treatment regimens.

2.3.4 Outcomes

The outcomes in the CS are in line with those in the final scope issued by NICE.[5]

2.3.5 Other relevant factors

TFT has a patient access scheme (PAS) in place related to the treatment of metastatic colorectal cancer, which is a simple discount of [redacted]. This discount is also applicable to TFT for the use in mGC and GEJ. TFT is linearly priced with a pack of 20 15mg/6.14mg tablets costing £500 at list price and a pack of 60 20mg/8.19mg tablets costing £2000 at list price.[6]

3 CLINICAL EFFECTIVENESS

3.1 Critique of the methods of review(s)

The company presented a systematic review aiming to "*identify the available clinical efficacy, safety and tolerability evidence related to the third- or later-line treatment of advanced and/or metastatic GC (including gastroesophageal junction cancer [GOJ/GEJ]*)." (p10 of systematic literature review (SLR) report). An initial search was conducted until 27th June 2018 with an update searching to 28th February 2019 (CS Appendix D).

The CS suggests there are no standard comparators in the third-line setting, and that very few patients would receive anything other than palliative care at this point. As such, whilst the NICE scope includes chemotherapies as comparators, these are not contained in the company's economic analysis. The clinical advisors to the ERG agreed that they are not relevant comparators for the majority of patients. For the sake of transparency and completeness, the review of chemotherapy agents in a third-line setting is of some (but not crucial) relevance, and is discussed briefly in Section 3.1.53.1.5.

3.1.1 Searches

The evidence searches for the SLR report and subsequent update conducted by Servier Laboratories are reported in Appendices D1.1 and D1.2 of the CS, respectively. Both reviews used similar methods of identifying evidence, with only minor differences between the search strategies.

Searches covered all of the key databases recommended by NICE (MEDLINE including Medline-in-Process, Embase and Cochrane) plus two years' worth of relevant conference proceedings and HTA websites. For both the SLR report and the update, MEDLINE and Embase were searched simultaneously in a "multi-file" search (on Embase.com). This technique is not usually recommended since it limits the ability to optimise the strategy for each source.

Searches are generally well-designed and executed and include both Emtree subject headings and free text terms, with appropriate use of truncation and proximity search strings to increase the sensitivity.

The ERG noted some logical errors in Servier's update searches (a missing line in one search strategy, and incorrectly numbered lines in another) and queried these with the company (Clarification question, A6[7]). The company acknowledged the errors, blaming them on an attempt to re-format the search strategies for aesthetic reasons, but gave assurance that the actual searches were conducted correctly. (Clarification response, A6). On this basis, the ERG does not believe any relevant studies are likely to have been missed.

3.1.2 Inclusion criteria

The inclusion criteria are reported in Table 6 of the CS. These appear appropriate to the ERG, although one of NICE's decision problem-defined outcomes (duration of response) was omitted. However, this does appear to have been data extracted and reported in the SLR report.

For the original review, study selection was conducted by two independent reviewers, which is a high quality methodology.[8] For the update, only one reviewer conducted study selection, leaving the update at some risk of bias and error, though the extent to which this operated is unclear.

3.1.3 Critique of data extraction

For the original review, the company used a pre-agreed data extraction form, although it was not clear if this was piloted and tailored to the specifics of the review. Data were extracted by one reviewer and checked by a second, which is likely to result in reliable data extraction. For the update, only one reviewer extracted data, although it is not clear if this was into a data extraction form, or directly into data tables for presentation in the report. Data extraction in the update is therefore at some risk of bias and error; however, the extent to which this operated is unclear. This may be important as the results from the pivotal trial (Shitara *et al.* 2018[2]) were obtained from the updated search. The ERG checked key data and these were generally correct, though a few minor inconsistencies between the CSR, Shitara *et al.* 2018[2] and the CS were noted.

The information provided by the company in Appendix D and the SLR report appears to be complete, relevant, and to provide an appropriate level of detail. However, the ERG notes some possible mistakes in the original review. For example, in Table 5 of the SLR report, column 2 states that studies in all patients were at third or later lines, but column 5 shows one study (Li et al. 2016[9]) included patients at second- and third- lines of therapy. This is further at odds with a statement that "*The review identified only four RCTs that were solely conducted in GC patients receiving treatment in third- or further-line therapy.*" (p19, SLR report).

3.1.4 Quality assessment

Quality assessment for all included studies was provided in the SLR report. It used the items listed in the NICE user guide for evidence submissions[10] and was conducted by one reviewer and checked by a second. The CS and Appendix D of the CS focussed only on the TAGS study. The quality assessment in the CS was similar to the SLR assessment, but added reasons for scores for some items. It was unclear if the reasons for scores were checked by a second reviewer, as they were not presented in the SLR report. Reasons for scores were missing for some items, but were provided in the clarification response.[7]

Study quality was also assessed using the Cochrane Risk of Bias 2 (RoB2) tool[11] (Appendix D of the CS), but items relating to risk of bias due to deviations from the intended interventions were not scored clearly, as all options remained in column three. Clarification of the scores were requested by the ERG. This assessment appears to have been conducted by one reviewer and not checked. The company provided scores and reasons in their clarification response, with the explanation that RoB2 automatically greyed out some items due to answers to earlier questions. The answers relating to protocol deviations have been incorporated into Section 4.2.2.3.

The quality assessment of the key TAGS trial, comparing the CS scores with the ERG's own scores, is provided in Table 1. The company scored the trial at low risk of bias for all items. The ERG had concerns about imbalances between treatment arms in potentially prognostic baseline characteristics, although these are addressed for overall survival (OS) by an analysis adjusting for key factors. The ERG also noted some small imbalances in treatment discontinuation, withdrawals and loss to follow-up, but that these are unlikely to affect results greatly.

Table 1: Summary of risk of bias using the items listed in the NICE user guide for evidence submissions[10], as judged in the CS and by the ERG

Question	Company's score (Yes/No/Unclear) with reason (From CS[1])	ERG's score (Yes/No/Unclear) with reason (based on Shitara 2018[2])
Was randomisation carried out appropriately?	Yes Patients were enrolled by study investigators. Eligible patients were randomised (2:1) to trifluridine/tipiracil plus BSC or placebo plus BSC via a dynamic allocation method (biased coin) with an interactive-voice web-response system (IXRS). Almac (Craigavon, UK) operated the IXRS and created the algorithm that generated the individual patient allocation when the study site accessed the system. The company had no other role in the trial. Once a patient's eligibility was confirmed and the criteria for randomisation were met, study-site personnel logged on to the IXRS to allocate patients to treatment. The IXRS randomly assigned study medication (trifluridine/tipiracil or placebo) by	Yes As per the company's response in column 2

	assigning a kit number to that patient. Randomisation was stratified by region (Japan vs rest of world), ECOG performance status (0 vs 1), and previous treatment with ramucirumab (yes vs no).	
Was the concealment of treatment allocation adequate?	Yes Patients, investigators and study-site personnel, those assessing outcomes, and those analysing the data were masked to treatment assignment. Tablets of identical appearance were used to maintain masking. Only personnel from the contract research organisations involved in drug labelling and distribution (Fisher Clinical Services [Allentown, PA, USA] and Bell Medical Solutions [Tokyo, Japan]) and IXRS activities (Almac) were aware of treatment assignment.	Yes Almac (Craigavon, UK had no other role in the trial other than randomisation. Once a patient's eligibility was confirmed and the criteria for randomisation were met, study-site personnel logged on to the IXRS to allocate patients to treatment. The IXRS randomly assigned study medication (trifluridine/tipiracil or placebo) by assigning a kit number to that patient. (CS, p40) As per column 2.
Were the groups similar at the outset of the study in	Yes	Unclear - It is unclear if the small imbalances affected results in unadjusted analyses. Clinical advisors to the ERG stated that ECOG status, number of metastatic sites, HER2 status, and previous chemotherapy regimens (number and type) are prognostic of survival

terms of prognostic factors?	Baseline demographic and disease characteristics were generally balanced between the two treatment arms.	in the third-line setting, and that the impact of sex and ethnicity is uncertain because there is little data. Of these, there was some evidence of imbalance in some factors (TFT compared with placebo) (see list in Section 4.2.2.1) Clinical advisors to the ERG were not too concerned at these imbalances. The ERG notes that an analysis for OS, which adjusted for these factors, was presented, but an equivalent analysis for progression-free survival (PFS) and other outcomes was not provided.
Were the care providers, participants and outcome assessors blind to treatment allocation?	Yes Please see above regarding concealment of treatment allocation.	Yes As per the company's response in column 2
Were there any unexpected imbalances in drop-outs between groups?	No Reason from clarification response A14: *"The majority of patients who discontinued, did so due to progressive disease. In the trifluridine/tipiracil group, nine of the 11 deaths resulting in treatment discontinuation were attributed to disease progression (the cause of the*	Unclear - It is unclear whether small imbalances in drop outs are due to patients being "at risk" for longer. Patients in the TFT arm were slightly more likely to stop treatment due to withdrawal of consent (4.2% vs 3.6%); adverse events (9.8% vs 6.5%); physician decision (3.3% vs 1.8%). [redacted] but not reported for OS.

	other two deaths was septic shock that was judged to be unrelated to treatment).”	
Is there any evidence to suggest that the authors measured more outcomes than they reported?	*No* Reason from clarification response A14: *“There are no outcomes that are not accounted for in the clinical study report.”*	No All outcomes are reported in the CSR.
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes	Yes, an ITT analysis was performed. For OS, in the absence of death confirmation or for patients alive as of the OS cut-off date (30th April, 2018), the survival time was censored at the date of last study follow-up or the cut-off date, whichever was earlier. For PFS, patients who were alive with no disease progression as of the analysis cut-off date (31st March, 2018) were censored at the date of the last tumour assessment. Patients who received non-study cancer treatment before disease progression were censored at the date of the last evaluable tumour assessment before the non-study cancer treatment was initiated.

Adapted from Systematic reviews: CRD’s guidance for undertaking reviews in health care (University of York Centre for Reviews and Dissemination)

3.1.5 Evidence synthesis

There is some lack of clarity around the total number of included studies across the original and update reviews, but only four RCTs[2, 4, 9, 12] were in solely third-line (or later) patients, or reported the results for this subgroup separately (CS p29). The four RCTs related to TFT, apatinib and nivolumab. As apatinib and nivolumab are not licensed in England and cannot therefore be considered as comparators, only the study relating to TFT is relevant to the decision problem.

Additionally, the SLR report found two non-randomised trials and fourteen single arm studies, but the majority of these were not of treatments licensed in the England. Any attempt to compare these results to the TFT results would involve strong assumptions and are likely to produce highly uncertain results, with most studies having fewer than fifty participants. Due to the lack of data and the rarity of use of chemotherapy in the third-line setting, the ERG agrees that performing a network meta-analysis would not be useful for this appraisal.

A single arm study of 35mg/m^2 and 40mg/m^2 TFT from Japan (EPOC1201)[13] was not synthesised with the TAGS study.[2] The CS states that this is because some patients in EPOC1201 had only one prior round of chemotherapy, and some had the wrong dose of TFT (40mg/m^2 TFT).The company argues elsewhere in the CS that gastric cancer operates quite differently in Japanese patients "*Due to biological differences between gastric cancer in Asian versus non-Asian patients efficacy of these treatments* [apatinib and nivolumab] *in European patients is uncertain, as recognised by ESMO and JSMO, who developed a Pan-Asian adapted ESMO clinical practice guideline for the management of patients with mGC."* (p66 of the CS). Evidence from the TAGS trial may support this in that Japanese patients in the placebo group had a median OS of 5.9 months compared with 3.3 months in patients from the EU and US. Whilst these latter two issues could potentially have been overcome through subgroup analyses, the ERG agrees that due to biological and clinical practice differences, the study has low relevance to the decision problem and meta-analysis was not necessary.

As such, the only study presented in detail in the CS is the TAGS study of TFT and no synthesis was performed.

3.2 Critique of the TAGS study, its analysis and interpretation

The pivotal trial for TFT is the TAGS study (Shitara *et al.*[14]). TAGS was a Phase III, randomised, double-blind, placebo-controlled study in patients with heavily pre-treated mGC, conducted at 110 sites in 18 countries, comparing 35mg/m^2 TFT and BSC with placebo and BSC.

As discussed in Section 3.1.5, a single-armed study of TFT (EPOC1201)[13] was not reported in detail in the CS which the ERG deems appropriate. However, adverse event data was of interest to the appraisal and is included in Section 3.2.2.8.

The ERG verified that no other important studies were missed with a focussed search in Pubmed and citation searching in Google Scholar of TAGS and EPOC1201 key publications.[2, 13]

3.2.1 Study design: The TAGS study[2]

A table detailing the design of the TAGS study was provided in the CS (Table 7, p34)[1] and an adapted version is reproduced here for reference (

Table 2).

The TAGS study mostly matched the decision problem specified by NICE (see

Table 2). Clinical advice to the ERG indicated that the study was broadly in accordance with English populations and practice as detailed below, in sections *Population, Intervention, Comparator, Outcomes.*

A critical appraisal of the TAGS study is provided in Section 3.1.4.

Table 2: Key design features of TAGS trial,[2] adapted from Table 7 of the CS[1]

Study	NCT02500043 TFT versus placebo in patients with heavily pre-treated mGC (TAGS): a randomised, double-blind, placebo-controlled, phase III trial Shitara K *et al*. Lancet Oncol 2018.[2]				
Study design	Phase III, randomised, double-blind, placebo-controlled study at 110 sites in 17 countries* to evaluate the efficacy and safety of TFT versus placebo in patients with previously treated mGC. Countries included in the European dataset: Belarus, Belgium, Czech Republic, France, Germany, Ireland, Israel, Italy, Poland, Portugal, Romania, Russia, Spain, Turkey, the UK Countries included in the USA dataset: the USA Country included in the Japanese dataset: Japan				
Population	507 patients were enrolled and randomly assigned, 337 to the TFT group and 170 to the placebo group. All patients were aged 18 or older with histologically confirmed, non-resectable, metastatic gastric adenocarcinoma (including adenocarcinoma of the GEJ) who had undergone two* previous chemotherapy regimens (and had experienced radiological disease progression) that contained fluoropyrimidine, platinum agents, and taxanes or irinotecan.				
Intervention(s)	Oral TFT (35 mg/m² twice daily on days 1–5 and days 8–12 every 28 days) plus BSC				
Comparator(s)	Placebo plus BSC				
Indicate if trial supports application for marketing authorisation	Yes	✓	Indicate if trial used in the economic model	Yes	✓
	No			No	
Rationale for use/non-use in the model	Pivotal phase III RCT.				
Reported outcomes specified in the decision problem	OS PFS Disease control rate Objective response rate HRQoL Safety				
Other reported outcome	Time to ECOG PS ≥ 2				

* There was some confusion over whether 18 (stated in Table 7 of the CS)[1] or 17 (stated on p41 of the CS[1] [redacted])[14] countries were included in the study, and whether patients had to have undergone one (stated in Table 7 of the CS, p34) or two (stated in journal article[2] [redacted])[14] previous chemotherapy regimens. The ERG has assumed that [redacted] on both matters.

Population

The inclusion criteria for the TAGS study were very detailed (see Table 8 of the CS, p36), including previous types and number of rounds of chemotherapy, time since progression, reason for cessation of previous therapy (toxicity or refractory disease), and various other criteria. The exclusion criteria were also very detailed, and excluded patients on the basis of co-morbidities, previous treatment with TFT, pregnancy and other criteria. Recruitment was stratified by region of the world (Japan versus the rest of the world); ECOG performance status (0 versus 1); and prior treatment with ramucirumab (yes versus no). The clinical advice provided to the ERG suggested that the criteria should result in a population broadly in line with patients in England and that the stratification factors were appropriate, though there are other prognostic factors that could have been considered (see Section 3.2.2.1).

Intervention

The intervention matches the proposed license for oral TFT (35 mg/m² twice daily on days 1–5 and days 8–12 every 28 days) plus best supportive care. It is to be given for as long as "benefit" is being gained. In their clarification response (A11)[7], the company defines discontinuation reasons as patient request, disease progression, clinical progression, adverse events, physician's decision or pregnancy. It was not entirely clear whether the discontinuation rules described in A11[7] were mandatory or optional, and their application in a clinical setting may differ from that in a trial setting, as the drug has not been licensed yet, and a draft SPC was not included in the CS, nor requested by the ERG in the clarification process.

The CSR details specific conditions for dose reductions and resumption of treatment in Section 9.4.6, and summaries are given in the CS on p41 and p116. Patients who did not achieve the minimal criteria for resumption had treatment discontinued. Patients with haematological toxicity could have doses withheld until neutrophil and platelet levels returned to an acceptable level.

Comparator

The comparator was placebo twice daily plus BSC on days 1–5 and days 8–12 of each 28-day treatment cycle.

Clinical advice provided to the ERG indicated that patients under NHS care in the UK are likely to have better community care than some in other European countries, where the standard of care is more variable. This was thought unlikely to affect survival, but might mean patients receiving best supportive care in the UK have a higher HRQoL than patients in other European countries in the trial.

Outcomes

A summary of the outcomes included in the TAGS trial, their definition and statistical analysis is provided in

Table 3. The protocol-defined outcomes for TAGS that were also listed in the NICE scope were: OS; PFS; (objective) response rate; HRQoL; and safety. The TAGS study also measured disease control rate (DCR), a composite of complete response (CR), partial response (PR) and stable disease (SD), which could be classed as "response rate"; and time to ECOG PS ≥ 2, which was not listed in the NICE scope. Patients had a computed tomography (CT) scan at baseline and then every 8 weeks until disease progression.

The NICE scope listed duration of response, which was not provided in the CS, but is [redacted]. Data were provided in the company's clarification response, though the ERG was not able to identify a definition of duration of response in the CSR, the CS or the clarification response.

Adverse event data was not included in the CS for (EPOC1201)[13], though it was reported in the SLR report and the EPOC1201 study publication.[13]

Table 3: Summary of outcomes measured in TAGS and their relevance to the NICE scope. Collated from text in the CS, the CSR and the clarification responses

Outcome	Definition & Statistical analysis
Outcomes listed in NICE scope	
Primary outcome	
Overall survival (OS)	Time from date of randomisation to death. In the absence of death confirmation or where patients still alive, data censored at date of last study follow-up or the cut-off date (30th April, 2018), whichever was earlier. **Statistical analysis:** OS and radiologically confirmed PFS were analysed in the ITT population with a one-sided stratified log-rank test, with the Hazard Ratio (HR) and two-sided 95% CIs based on a prespecified stratified (region, ECOG performance status, ramucirumab exposure) Cox model and associated Kaplan-Meier survival estimates.
Key secondary outcome	
Progression free survival (PFS)	Time from date of randomisation until investigator-assessed radiological progression or death. Patients alive with no progression at analysis cut-off date (31st March, 2018), patients who received non-study cancer drug before disease progression, and those who had clinical but not radiological progression were censored at last tumour assessment. **Statistical analysis:** As for OS. patients were censored for (1) discontinued follow-up, (2) follow-up ongoing at the time of analysis, (3) missed visit (>91 days since last response), and (4) initiated anti-tumour therapy (Additional clarification response A2[7]).
Other secondary outcomes	
Objective response rate (ORR)	Objective CR or PR based on investigator review of radiological images and following RECIST criteria (version 1.1, 2009). Best overall response was best recorded response after randomisation but before disease progression or initiation of non-study cancer treatment. Best response of SD needed to be maintained for 6 weeks after randomisation. Best response of CR or PR did not have a minimum time limit, as per RECIST 1.1.
Disease control rate (DCR)	Proportion of patients with objective evidence of CR, PR or SD. **Statistical analysis:** NR

Duration of response	Not defined in the CS or the CSR, or the clarification response. **Statistical analysis:** Not defined.
HRQoL	EORTC QLQ-C30 – measures 5 functional scales (physical, role, cognitive, emotional, and social), 3 symptom scales (fatigue, pain, and nausea and vomiting), a global health status scale, and a number of single items assessing additional symptoms commonly reported by cancer patients (dyspnoea, loss of appetite, insomnia, constipation and diarrhoea) and perceived financial impact of the disease. Gastric Specific Module (QLQ-ST022) - 22-item instrument used alongside the 30-item QLQ-C30 core questionnaire, resulting in a total of 52 items. **Statistical analysis:** Based on Osoba *et al*. 1998,[15] for both questionnaires a mean change from baseline of at least 10 points was considered to be clinically relevant for the patients.
Adverse events	Graded according to the US National Cancer Institute's Common Terminology Criteria for Adverse Events (version 4.03) and recorded from the first dose of study drug (that is, day 1, cycle 1) until 30 days after the last dose of study drug. Includes haematology, serum chemistry and urinalysis. **Statistical analysis:** NR
Outcomes not listed in the NICE scope	
Time to Deterioration of ECOG Performance Status	Time from randomisation to ECOG performance status score of two or higher. **Statistical analysis:** NR

OS, overall survival; PFS, progression-free survival; CR, complete response; PR, partial response; SD, stable disease; NR, not reported.

Planned analyses

There were three analysis sets:

- **Intention to treat (ITT) population** – the primary population for all efficacy outcomes, included all randomised patients, according to treatment assigned at randomisation.
- **As-treated (AT) population**– patients who took any dose of study treatment, analysed according to the treatment they received. Used for the safety analyses.
- **Tumour-response (TR) evaluable population** – an ITT analysis only including patients with measurable lesions.

There were a number of subgroup and "supportive" analyses planned (CSR p56-58). Of most relevance were:

OS

- [redacted] The company did not provide the SAP, which was not identified by the ERG until after the clarification round.
- Multivariate analyses including the stratification factors and potential prognostic/predictive factors (age group (< 65, ≥ 65 years); race (White, Asian, other); gender; number of prior regimens (≤ 2, ≥ 3); prior therapy (taxane, irinotecan); previous gastrectomy; gastroesophageal junction involvement; presence of peritoneal metastases; presence of liver metastases; presence of lung metastases; number of metastatic sites (1-2, ≥ 3); [redacted] histology subtype (diffuse, intestinal), and human epidermal growth factor receptor (HER)2 status at baseline.)

PFS

- Various analyses including clinical progression, radiological progression, initiation of non-study drug as a PFS events and [redacted]

There was a lack of detail in the CS relating to the analysis plans for some outcomes (see

Table 3), especially relating to how missing data was handled. Stratification factors were adjusted in the analysis of OS and PFS, but whilst other potential prognostic or predictive factors were adjusted in an additional analysis for OS, a similar analysis was not presented for PFS. For PFS patients were censored for (1) discontinued follow-up, (2) follow-up ongoing at the time of analysis, (3) missed visit (>91 days since last response), and (4) initiated anti-tumour therapy (Additional clarification response A2[7]). [redacted]

[redacted]

3.2.2 Study results: The TAGS study[2]

3.2.2.1 Baseline characteristics

The baseline characteristics of patients in the study are provided in Table 4 (reproduction of Table 9 from the CS). Clinical advice provided to the ERG suggests that most proportions were in line with the patient population in England. One exception was that clinicians would expect the ratio of patients with gastric compared with gastroesophageal cancer to be around 40:60 in the third-line setting, as gastroesophageal cancer is generally more common. However, in the study, the ratio was roughly 71:29, indicating more gastric cancer patients than might be expect in an English population. Clinical advice provided to the ERG indicated that the two cancers behave similarly; however, the ERG notes that the graph of OS by subgroup presented in Shitara *et al.*[2] shows that the point estimate of the hazard ratio for those in the gastric cancer subgroup is more favourable (0.67 (95% CI 0.52 to 0.87) than for those with gastroesophageal cancer (0.75 (95% CI 0.50 to 1.11), although the confidence intervals overlap considerably. The impact of including more patients with gastric cancer on estimates of efficacy is unclear.

The trial includes around 62% of patients who have had three or more regimens of chemotherapy. This appears to be at odds with clinical advice provided to the ERG, which indicated that most patients in England do not get third-line therapy as by this stage, they are too ill and the burden of treatment outweighs the benefits. This view is echoed in the clinical statements submitted for this appraisal,[16, 17] where clinicians estimate that only 10-15% of patients have third-line therapy. However, the subgroup analysis in Figure 3 of Shitara *et al.*[2] indicates that previous lines of therapy (2, 3 or ≥4) did not impact much on the point estimates of the hazard ratios for each subgroup so this may not impact much on estimates of efficacy. It may, however, impact negatively on median survival compared to an English population at third-line therapy, as the prior ramucirumab population had more prior lines of treatment than the no prior ramucirumab population.

Table 4 Patient baseline characteristics. Reproduction, with correction of errors according to CSR, of Table 9 from the CS

	All regions		[redacted]	
	TFT (n=337)	**Placebo (n=170)**	[redacted]	[redacted]
Age (years)			[redacted]	[redacted]
Median (range*)	64.0 (24–89)	62.5 (32–82)		
<65	183 (54%)	96 (56%)		
≥65	154 (46%)	74 (44%)		
Sex			[redacted]	[redacted]
Male	252 (75%)	117 (69%)		
Female	85 (25%)	53 (31%)		
Ethnicity			[redacted]	[redacted]
White	244 (72%)	113 (66%)		
Asian	51 (15%)	29 (17%)		
Other	4 (1%)	4 (2%)		
Not available	38 (11%)	24 (14%)		
Region			[redacted]	[redacted]
USA	21 (6%)	5 (3%)		
Europe**	270 (80%)	138 (81%)		
Japan	46 (14%)	27 (16%)		
ECOG performance status			[redacted]	[redacted]
0	123 (36%)	68 (40%)		
1	214 (64%)	102 (60%)		
Primary site			[redacted]	[redacted]
Gastric	239 (71%)	121 (71%)		
GEJ	98 (29%)	47 (28%)		
Both	0	2 (1)		
Measurable disease	306 (91%)	150 (88%)	[redacted]	[redacted]
Histology			[redacted]	[redacted]
Diffused	53 (16%)	21 (12%)		
Intestinal	103 (31%)	52 (31%)		
Mixed	14 (4%)	8 (5%)		
Unknown	132 (39%)	69 (41%)		
Not available	35 (10%)	20 (12%)		
HER2 status			[redacted]	[redacted]
Positive	67 (20%)	27 (16%)		
Negative	207 (61%)	106 (62%)		
Not assessed	62 (18%)***	37 (22%)		
No. of metastatic sites			[redacted]	[redacted]
1–2	155 (46%)	72 (42%)		
≥3	182 (54%)	98 (58%)		
Peritoneal metastases	87 (26%)	53 (31%)	[redacted]	[redacted]
Previous gastrectomy	147 (44%)	74 (44%)	[redacted]	[redacted]
No. of prior regimens			[redacted]	[redacted]
2	126 (37%)	64 (38%)		
3	134 (40%)	60 (35%)		
≥4	77 (23%)	46 (27%)		

	All regions		[redacted]	
	TFT (n=337)	**Placebo (n=170)**	[redacted]	[redacted]
Prior systemic cancer therapeutic agents			[redacted]	[redacted]
Platinum	337 (100%)	170 (100%)		
Fluoropyrimidine	336 (>99%[a])	170 (100%)		
Taxane[†]	311 (92%)	148 (87%)		
Irinotecan[‡]	183 (54%)	98 (58%)		
Ramucirumab	114 (34%)	55 (32%)		
Anti-HER2 therapy	60 (18%)	24 (14%)		
Immunotherapy (anti–PD-1/PD-L1)	25 (7%)	7 (4%)		
Other	77 (23%)	41 (24%)		

ECOG PS: Eastern Cooperative Oncology Group performance status; HER2: human epidermal growth factor receptor 2; PD-1: programmed death-1; PD-L1: programmed death-ligand 1

Note: Data are n (%) unless noted otherwise. *Please note that Europe refers to Belarus, Belgium, Czech Republic, France, Germany, Ireland, Israel, Italy, Poland, Portugal, Romania, Russia, Spain, Turkey, and the UK; †One patient did not receive a fluoropyrimidine; ‡All patients received irinotecan or taxane or both.

* this was given as IQR in the CS, but range in the CSR. The IQR in Shitara[2] suggests the CSR is correct.

** Servier could not identify these values at the time of clarification as derivation requires further interrogation of patient level data from the TAGS trial.

*** This was given as 62 (18%) in the CS, but 63 (19%) in Shitara.[2] Shitara has been preferred so the total is 337 patients.

3.2.2.2 The potential impact of prior ramucirumab

Because ramucirumab does not have a positive NICE recommendation, the proportion of patients with prior exposure to ramucirumab is unlikely to reflect that of patients in England. This adds uncertainty to whether the full results from the TAGS study are generalisable to England, as firstly, the prognosis of patients who have received ramucirumab may be different from those that did not. Secondly, the relative efficacy of TFT may differ in patients who received ramucirumab and those that did not. These are discussed in turn.

The relative prognosis in patients with and without prior ramucirumab treatment is unknown. Clinical advice provided to the ERG was that prior ramucirumab use was unlikely to influence the natural history of the cancer. However, a view expressed in a clinical expert statement provided to NICE was that it may possibly mean that patients are in a "*better state*" although this observation was acknowledged to be anecdotal. Without a strong indication that prior ramucirumab treatment alters prognosis, the ERG prefers to assume that there is no impact associated with prior ramucirumab treatment but notes differences in the prior ramucirumab group and the no ramucirumab group in terms of prior lines of treatment and disease duration. Therefore, this assumption is uncertain.

In a further clinical expert statement provided to NICE it was commented, "*there is no reason why prior ramucirumab would alter the outcome for trifluridine-tipiracil. They work on completely different*

pathways and cross resistance would not be expected." As such, whilst the relative efficacy of TFT in patients with and without prior ramucirumab treatment is unknown, the ERG prefers an estimate of a HR or AF from the entire population rather than from only patients who had not received ramucirumab.

3.2.2.3 The potential impact of geographical region

Some recruitment took place in Japan, and the proportion of Japanese patients in TAGS was 14% in the TFT arm and 16% in the placebo arm respectively. Clinical advice provided to the ERG indicated that EU patients are likely to have the greatest generalisability to England, as disease prognosis and treatment practices are more similar within the EU than in Japan. This view is echoed in the CS (see Section 3.1.5). Being recruited in Japan, compared with recruitment in Europe or the USA was stratified for at baseline in the TAGS study, suggesting that the study investigators believed that being Japanese could affect the efficacy of TFT compared to that for EU or USA patients.

In its clarification response, the company asserted that Japanese patients should be included because England has an 8% Asian population (clarification response A22[7]), and that for their base-case (the no prior ramucirumab population) there were [redacted] Japanese patients. The most recently available census data for England and Wales (2011[18]) indicates that 7.5% of the population was Asian, and the majority were of Indian (approximately 3%) or Pakistani (approximately 2.5%) ethnicity. People of Japanese ethnicity were not reported separately, but probably included in the category of "other Asian" (approximately 1.5%). The ERG notes that this makes the Asian population in TAGS around double that of the English population and in the company's base-case (no prior ramucirumab) around half that of the English population. In both the whole trial and the no prior ramucirumab group, there are likely to be a higher proportion of Japanese patients than is found in England and the generalisability of Japanese patients to the broader category of "Asians" is unclear. The exclusion of the patients recruited in Japan and the USA leaves a [redacted] (clarification response A3[7]), which is also an under-representation of Asians within the trial results compared to the English population. The ERG concludes that whilst exclusion of the Japanese patients from the whole trial, or use of the no-prior ramucirumab population leads to an under-representation of Asians compared with the English population, their inclusion leads to over-representation, and the generalisability of Japanese patients to the more diverse Asian population in England is unclear. The ERG concludes that analyses of European patients, or where not available, the ROW have highest relevance to the decision problem.

3.2.2.4 The balance of prognostic factors between arms

Clinical advice provided to the ERG indicated that ECOG status, number of metastatic sites, HER2 status and previous chemotherapy regimens (number and type) are prognostic of survival in the third-line setting. The impact of sex and ethnicity was thought to be uncertain due to a lack of data. The CS

broadly agrees, stating that "*ECOG PS, age, number of previous chemotherapy regimens (two versus three), number of metastatic sites, and HER2 status were prognostic of improved OS*" (p56 of the CS). Of these, there was some evidence of imbalance in some factors (TFT compared with placebo): sex (75% and 69%); ethnicity (72% and 66% white); ECOG status 0 (36% and 40%); HER2 status (20% and 16% positive); number of metastatic sites (≥3 54% and 58%); number of previous regimens (≥4 23% and 27%).

The clinical advisors to the ERG noted the slight imbalances in potential prognostic factors but were not concerned. However, the ERG asked the company for clarification on how the results might affect efficacy estimates. The company responded (clarification response A16[7]) that their clinical advisors had not been concerned, and consequently no adjustments had been made in analyses. They added that the direction of effect of imbalances in prognostic factors was mixed and likely to counteract each other. In their OS analysis, however, the company has adjusted for [redacted]) and the adjusted result was similar to the ITT population primary analysis results (see Section 3.2.2.7).

[redacted]

[redacted]

3.2.2.5 Flow of patients through the trial

A flow diagram of patients through the trial was presented in the CS in the appendix relating to the systematic review (Appendix D) and in Shitara *et al*.[2] A version correcting identified errors was provided as Figure 5 in the clarification response and is reproduced here as Figure 2.

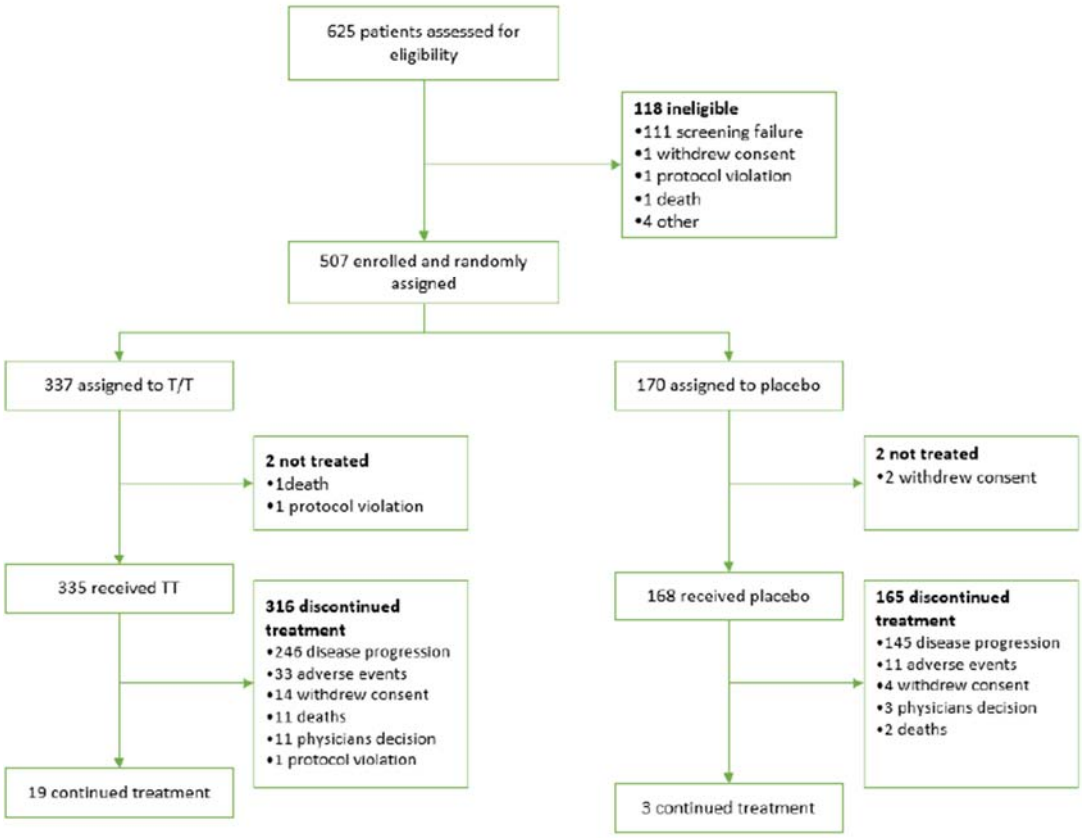


Figure 2 CONSORT flow diagram of patients in the TAGS study. Reproduction of Figure 5 of the clarification response[7]

The ERG calculated treatment discontinuation rates in

Table 5. In total, the number of treatment discontinuations not due to disease progression was 21.4% in the TFT arm, and 12.9% in the placebo arm. The proportion of deaths were 3.6% in the TFT arm, but 1.2% in the placebo arm. The difference in deaths appears large, but for 9/12 patients in the TFT arm the cause of death was disease progression, meaning patients died before progression was recorded as an outcome. All other reasons are generally higher in the TFT arm, which may be due to patients being on treatment for longer. The only discontinuation reason that might affect OS is withdrawal of consent as all other patients would be followed up for survival, and this is largely balanced between arms (4.2% in TFT arm and 3.6% in placebo arm). For PFS, it is not clear to the ERG how withdrawal of consent, physician decision, protocol deviation and adverse events (AEs) were handled in analyses.

Table 5 Reasons for treatment discontinuations in the TAGS study. Number of patients taken from Shitara *et al*.[2]

Reason for treatment discontinuation	TFT		Placebo	
	Number of patients (N=337)	% of patients in arm	Number of patients (N=170)	% of patients in arm
Death	12*	3.6	2	1.2
Protocol deviation	2 †	0.6	0	0
Disease progression	246	73.0	145	85.3
Adverse events	33	9.8	11	6.5
Withdrew consent	14	4.2	6‡	3.6
Physician decision	11	3.3	3	1.8
Still on treatment	19	5.6	3	1.8
Total	337	100	170	100

*One death occurred in the TFT arm before treatment started

†One protocol violation occurred in the TFT arm before treatment started.

‡Two withdrew consent from the placebo arm before treatment started.

3.2.2.6 Dosing delays, protocol deviations and non-study drugs

Dosing delays: 58% of patients receiving TFT and 22% of patients receiving placebo had dosing delays and reductions (Table 14 of CS); 11% and 1% had dose reductions; 13% and 17% had treatment discontinuation.

Protocol deviations: [redacted]

Non-study drugs: The ERG asked for clarification around how many patients received non-study drugs as this was not clear from the CS. [redacted] The ERG asked for clarification on how many patients discontinued treatment in order to receive a non-study drug; [redacted]

3.2.2.7 Efficacy of TFT

Key efficacy results for the TAGS study are presented in Table 6.

Table 6: The efficacy data from TAGS, with reference to NICE scope. Data taken from the CS[1], the CSR[7] and Shitara *et al*.[2]

	Whole trial population			**Ramucirumab–naïve subgroup (pre-specified analysis)**		
Outcome	**TFT N=337**	**Placebo N=170**	**Between group comparisons**	**TFT N=223**	**Placebo N=115**	**Between group comparisons**
Outcomes listed in NICE scope						
Primary outcome						
OS Median (months):	5.7, 95% CI: 4.8-6.2	3.6, 95% CI: 3.1-4.1	HR: 0.69 (95% CI: 0.56–0.85, one-sided p=0.0003, two-sided p=0.0006) Difference between medians: 2.1 months	6.0 (95% CI: 5.1-6.9)	3.3 (95% CI: 2.8-3.9)	HR: 0.66 (95% CI: 0.51–0.85)
Key secondary outcome						
PFS Median (months):	2.0 (95% CI1.9-2.3)	1.8 (95% CI 1.7-1.9)	HR: 0.57 (95% CI: 0.47–0.70, two-sided p<0.0001)	2.2 (95% CI:1.9-3.5)	1.8 (95% CI:1.8-1.9)	HR 0.5832 (95% CI: 0.4550- 0.7475)
Other secondary outcomes						
ORR* Rate	4.5%	2.1%	NR	NA	NA	NA
Disease control rate (DCR, composite of CR, PR and SD)	44.1%	14.5%	**p<0.0001**	NA	NA	NA
SD	39.7%	12.4%		NA	NA	NA
Duration of response	No summary statistics presented. See "**Response rates and duration of response"** below.			NA	NA	NA
HRQoL EORTC QLQ-c30 and QLQ-STO22	No comparative summary statistics presented in the clinical section. See clarification response A20 for full HRQoL summary statistics.			NA	NA	NA
Outcomes not listed in the NICE scope						
Time to Deterioration of ECOG Performance Status Median (months)	4.3 (95% CI: 3.7–4.7)	2.3 (95% CI: 2.0–2.8)	HR 0.69, 95% CI: 0.56–0.85, two-sided p=0.0005	NA	NA	NA

* Restricted to patients with measurable disease, i.e. 290/337 in the TFT group, 145/170 in the placebo group

Overall survival

In the ITT population, the hazard ratio for OS was 0.69; 95% CI: 0.56–0.85, p=0.0006, indicating patients lived statistically significantly longer in the TFT arm than in the placebo arm. Median OS was 5.7 months in the TFT arm and 3.6 months in the placebo arm; the difference in median survival was 2.1 months. At six months, 47% of TFT patients and 33% of placebo patients were alive. At one year, 21% and 13% respectively were alive. The Kaplan-Meier plot is provided as Figure 3a.

[redacted]

OS subgroup and supportive analyses

Prognostic factor subgroups: Subgroup analyses (including by stratification factor) were reported more fully in the CSR and are reproduced here as Figure 4. The study was not powered for these subgroup analyses, and no statistical comparisons were presented. Patients with measurable disease appeared to respond statistically significantly less well than those without measurable disease, on the basis of their confidence intervals not overlapping (0.74 (95% CI 0.59 to 0.93) compared with 0.21 (95% CI 0.09 to 0.52) respectively[2]). Where confidence intervals overlapped, the biggest differences between point estimates were seen for prior treatment with irinotecan (point estimate favours those without prior treatment) and prior treatment with taxane (point estimate favours TFT for those with prior treatment and favours placebo for those without prior treatment). Others with notable differences included age (<65 compared with ≥75), prior ramucirumab, "other" ethnicity, gastrectomy, tumour grade, peritoneal metastases, histology and HER2 status.

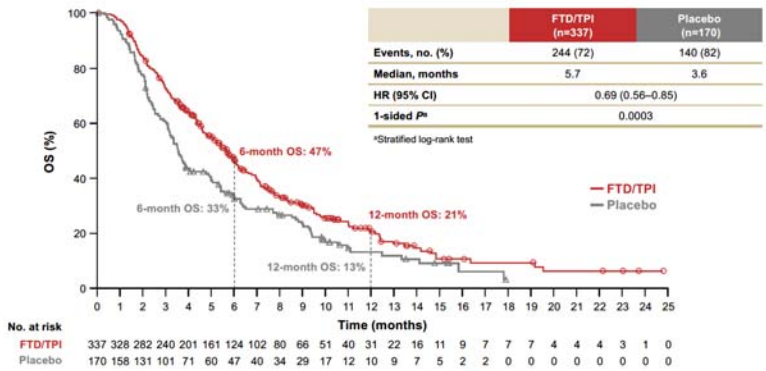
No prior ramucirumab: The pre-specified analysis of patients with no prior ramucirumab treatment indicated that the treatment effect was consistent with the main analysis with an HR 0.66 (95% CI: 0.51–0.85), and similar median survival (see Table 6).

Japanese patients compared with rest of the world: Patients in Japan had a median OS in the TFT (n=46) and placebo (n=27) groups of 6.3 months and 5.9 months, respectively, and a HR of 0.77 (95% CI 0.46–1.30). This compared with patients in rest of the world (ROW, comprising EU/US patients) who had a median OS of 5.4 months in the TFT (n=291) and 3.3 months in the placebo group (n=143), and a HR of 0.68 (95% CI 0.54–0.85).

EU patients with no prior ramucirumab: [redacted]

However, the ERG did not agree that subgrouping patients by prior ramucirumab treatment is necessarily appropriate (see Section 3.2.2.2). Subgroup data for patients in Europe regardless of prior ramucirumab treatment were presented in Shitara *et al.*[2] for OS, with an HR of 0.67; 95% CI: 0.53-0.86. Median survival was presented in the CSR (CSR[14], Figure 7) as [redacted]. No Kaplan-Meier plots were available to the ERG.

a)



b)

Figure 3: **Overall survival Kaplan-Meier curves for a) the whole population and for b) patients in the EU with no prior ramucirumab treatment. Reproduction of Figure 7 of the CS and Figure 1 of the clarification response, respectively**

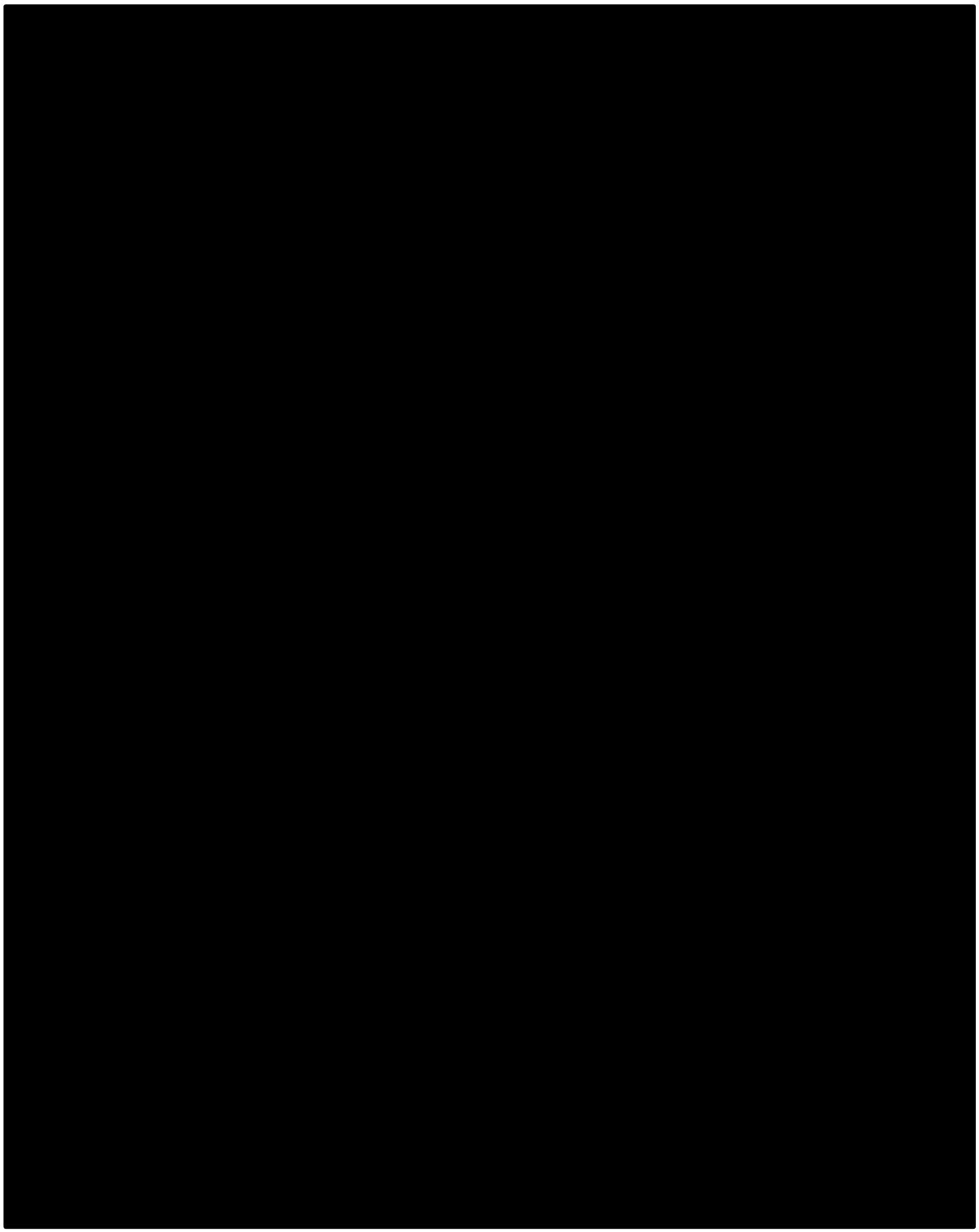


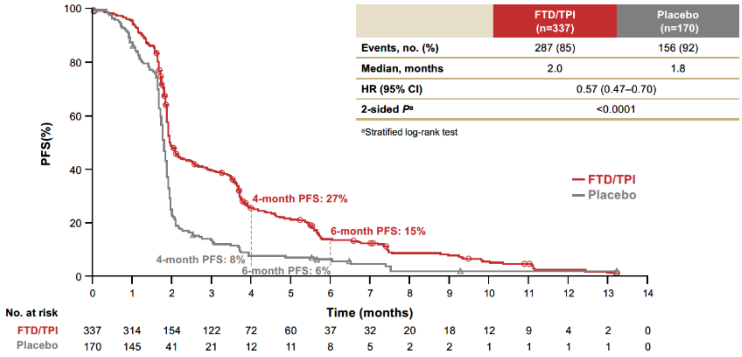
Figure 4: Reproduction of Figure 7 from the CSR, Hazard Ratio for Treatment Effect on Overall Survival by Selected Subgroups (Intent-to-Treat Population)

Progression-free survival

In the ITT population, the HR for PFS was 0.57, 95% CI: 0.47–0.70, p<0.0001, indicating patients progressed statistically significantly later in the TFT arm compared with the placebo arm. Median PFS

was 2.0 months in the TFT arm and 1.8 months in the placebo arm; the difference in median PFS was 0.2 months. At six months, 27% of TFT patients and 8% of placebo patients were progression free and alive. At one year, 15% and 6% respectively were progression free and alive. The Kaplan-Meier plot for the whole TAGS population is provided in Figure 5a with the value for European patients without prior ramucirumab use in shown in Figure 5b. Both plots have "steps" at two monthly intervals, presumably caused by radiological progression being observed at scheduled study assessment points. Clinical advice to the ERG suggested 4-6 weekly monitoring was usual in clinical practice in England. This may not always be a radiological assessment, but clinicians would make treatment decisions based on clinical or radiological progression, in accordance with the stopping rules described in Section 3.2.1. As such, patients may discontinue treatment earlier, which may affect efficacy, adverse events, and drug costs.

a)



b)

Figure 5: **Progression free survival Kaplan-Meier curves for a) the whole population in the TAGS study and for b) patients in the EU with no prior ramucirumab treatment. Reproduction of Figure 8 in the CS and Figure 2 in the clarification response, respectively**

PFS subgroup and supportive analyses

Prognostic factor subgroups: patients subgrouped according to potential prognostic factors were presented in the online supplement of Shitara *et al.*[2] Of most relevance to the appraisal was an analysis of European patients, regardless of ramucirumab treatment. The HR was 0.60; 95%CI: 0.48-0.75. The CS, Shitara *et al.* and the CSR report neither the median survival nor the Kaplan-Meier plots.

EU patients with no prior ramucirumab: [redacted]

Subgroups: Shitara *et al*[2] reports pre-specified subgroup analyses for PFS, but these were not reported in the CS, and are not reported here due to their low relevance to the health economic model (Chapter 4).

Response rates and duration of response

Response rate outcomes only included patients with measurable disease and ≥1 post-baseline assessment (the TR population; 287/337 (85%) patients receiving TFT, and 156/170 (92%) patients receiving placebo). Objective response rate (ORR, a composite of CR and PR) was 4.5% and 2.1% (*p*-value not reported) and were low in both arms, as would be expected in patients at third and later lines

of chemotherapy. DCR (a composite of CR, PR and SD) was 44.1% and 14.5% (p<0.0001) in the TFT and placebo arms respectively. DCR rates were largely due to SD (39.7% versus 12.4%, respectively). For the European, no prior ramucirumab population the DCR was [redacted] in the TFT group versus [redacted] in the placebo group.

Duration of response was only presented as a figure in the clarification response[7] and is presented here as Figure 6 a and b. [redacted]

No other subgroup or supporting analyses were presented in the CS. It is unclear what the results would be in patients without measurable disease.

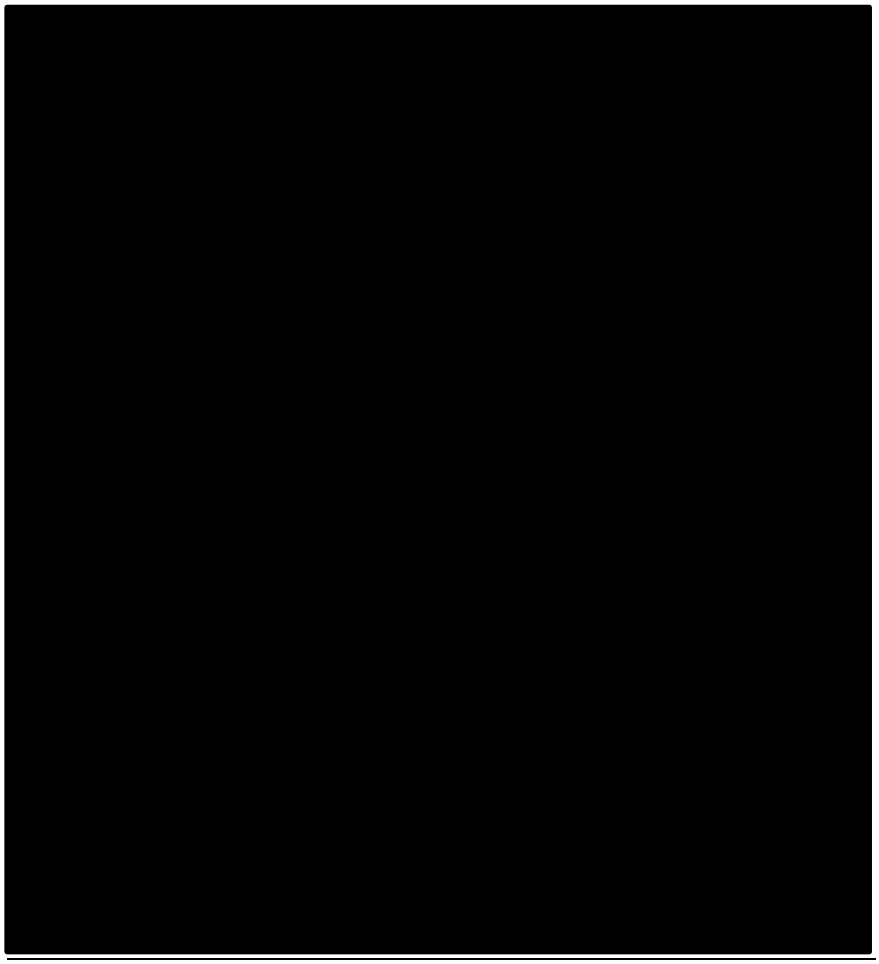


Figure 6: Duration of response in a) the whole population in the TAGS study and for [redacted] Reproduction of Figure 3 & 4 in the clarification response

Time to ECOG Performance Status ≥ 2

The HR for Time to ECOG performance status ≥ 2 was 0.69 (95% CI: 0.56–0.85, p=0.0005) with a median of 4.3 months (95% CI: 3.7–4.7) and 2.3 months (95% CI: 2.0–2.8) in the TFT and placebo groups, respectively. This was not an outcome listed in the NICE scope. However, it indicates that patients maintained ECOG performance status for longer in the TFT arm compared with the placebo group.

Health-related quality of life

HRQoL was measured with the EORTC QLQ-C30 and the EORTC QLQ-STO22. The company did not report HRQoL using comparative summary statistics with *p*-values. Instead, a simple table of mean change from baseline was presented (with neither confidence intervals nor *p*-values reported, p56 of the CS). Additionally, a large table of summary values with standard deviations for each subscale was provided in the clarification response (question A20).[7] Using the pre-specified cut-off of a mean change from baseline of ≥10 points, the company stated that there were no clinically relevant differences (between groups) in the mean change from baseline. For some items there was a difference of ≥10 points in the mean change from baseline (within group), and these included less hair loss, role functioning, fatigue, pain and appetite loss. The company also highlighted a between-group clinically relevant difference in pain: 11.3 (Cycle 2) in favour of TFT; and role functioning: 10.0 (Cycle 3) in favour of placebo. The ERG comments that the cut point of 10 points was derived for the EORTC QLQ-C30, and of unknown significance to the EORTC QLQ-STO22[15] but concludes that the results suggest that HRQoL is not affected in a clinically relevant way, either positively or negatively, by treatment with TFT.

3.2.2.8 Safety of TFT: Adverse events

Only one study (TAGS[2]) reporting AEs was included in the CS review. These AEs are summarised in Table 12 (page 59) of the CS.[1] The safety analysis included all patients who took any dose of study treatment. All analyses using this population were based on the treatment received. Safety was assessed in the TAGS trial by investigators throughout the study and AEs were graded according to the US National Cancer Institutes' common Terminology Criteria for Adverse Events (version 4.03[19]) and recorded from the first dose of study drug until 30 days after the last dose of study drug. The overall incidence of AE events was 97.3% for the TFT group and 93.5% for the placebo treatment group. However, grade III or worse AEs occurred in 267 (80%) patients in the TFT group, but only in 97 (58%) patients in the placebo group.

The most common AE reported included: nausea; anaemia; decreased appetite; vomiting; diarrhoea; fatigue; neutropenia; asthenia thrombocytopenia. Anaemia and neutropenia were two outcomes where

the incidence appeared to be markedly greater in the TFT group compared with the placebo group (anaemia 45% vs 19%; neutropenia 53% vs 4%).

AEs resulting in death occurred in 13.4% (n=45 patients) in the TFT group and in 11.3% (n=19 patients) in the placebo group. The most frequently reported AE resulting in death in both treatment groups was general physical health deterioration.

In the economic model, treatment-emergent grade 3 or 4 AEs were included provided they occurred in at least 5% of patients in either treatment arm. In addition, febrile neutropenia (occurring in n=6 of TFT patients in the TAGs trial) and nausea (n=14 TFT, n=5 control) were included within the cost-effectiveness model owing to their high impact on patient HRQL and the cost of its treatment. The ERG notes potential discrepancies between the data used in the model, the data reported in the clinical section of the CS, and data reported in the CSR. The biggest discrepancy is for the category anaemia, which is reported as [redacted]) in the TFT and placebo arms respectively in the CSR Table 35, as 64 (ERG-calculated 19.1%) and 13 (ERG-calculated 7.7%) patients respectively in Shitara *et al.*[2] and the CS Table 12, but as 37 (11%) and 5 (3%) patients respectively in the modelling section (Table 23 of the CS). The ERG were not able to determine why there was an apparent discrepancy, but speculate this may be due to composite outcomes ("anaemia" in the modelling compared with "anaemia or decreased haemoglobin concentration" in Shitara *et al.*), and/or different categories entering the analysis ("treatment emergent" versus "any cause", or events in 2% versus 5% of patients).

A phase II trial (Bando *et al.* 2016[13]) in Japanese patients was excluded from the review. The ERG has however, included the AEs observed in this study for comparison (Table 7). The most commonly occurring more serious AEs (grade III/IV) in patients receiving the 35 mg/m^2 were neutropenia (69.0%), leukopenia (41.4% and anaemia (20.7%). One case of febrile neutropenia occurred, although no treatment related deaths were reported. For those patients in the 40 mg/m^2 group, neutropenia (83.3%) and leukopenia (66.7%) were slightly more frequent. The findings of this study are tabulated below alongside the incidence of grade III-V AE in the TAGS study.[2] Neutropenia, anaemia and leukopenia were the most common serious AEs in both studies. The incidence of leukopenia was greater in the study by Bando *et al.* (41.4% vs 9.3%).[13] More AE-related deaths occurred in the TAGS study when compared with the Bando *et al.*[13] study (13.4% vs 0%); however, AE-related deaths in the intervention and control groups were similar in the TAGS study.

Table 7 Serious adverse events grade III or higher (Bando *et al.*[13], Shitara *et al.*[2] and CSR)

	EPOC1201		TAGS					
	Bando *et al.*[13]		Shitara *et al.*[2]		[redacted]		Model, CS Table 23	
	35 mg/m^2 (n=29) (%)	40 mg/m^2 (n=29) (%)	TFT (n=335) (%)	Placebo (n=168) (%)	[redacted] (%)	[redacted] (%)	TFT (n=335) (%)	BSC (n=168) (%)
Haematological								
Neutropenia	20 (69)	5 (83.3)	114 (34.0)*	0 (0)	[redacted]	[redacted]	77 (23)	√
Febrile neutropenia	1 (3.4)	0 (0)	4 (1.2)	0 (0)	[redacted]	[redacted]	6 (1.8)	√
Leukopenia	12 (41.4)	4 (66.7)	31 (9.3)	0 (0)	[redacted]	[redacted]	23 (6.9)	√
Anaemia	6 (20.7)	1 (16.7)	64 (19.1)	13 (7.7)	[redacted]	[redacted]	37 (11)	5 (3)
Asthenia			16 (4.8)	11 (6.5)	[redacted]	[redacted]	√	√
Thrombocytopenia			11 (3.3)	0 (0)	[redacted]	[redacted]	NR	NR
Hyponatraemia			4 (1.2)	7 (4.2)	[redacted]	[redacted]	NR	NR
Neutrophil count decreased			NR	NR	[redacted]	[redacted]	37 (11.0)	0 (0)
Non-haematological								
Anorexia/decreased appetite	3 (10.3)	1 (16.7)	29 (6.7)	11 (6.5)	[redacted]	[redacted]	√	√
Nausea	1 (3.4)	1 (16.7)	10 (3.0)	5 (3.0)	[redacted]	[redacted]	√	√
Vomiting	1(3.4)	1 (16.7)	12 (3.6)	3 (1.8)	[redacted]	[redacted]	NR	NR
Diarrhoea	0 (0)	0 (0)	9 (2.7)	3 (1.8)	[redacted]	[redacted]	NR	NR
Abdominal pain	0 (0)	0 (0)	14 (4.2)	15 (8.9)	[redacted]	[redacted]	√	√
Constipation	0 (0)	0 (0)	4 (1.2)	4 (2.4)	[redacted]	[redacted]	NR	NR
Deaths	0 (0)	0 (0)	45 (13.4)	19 (11.3)	[redacted]	[redacted]	NR	NR
At least one serious AE	NR	NR	143 (42.7)	70 (41.7)	[redacted]	[redacted]	NR	NR
Dysphagia	NR	NR	7 (2.1)	4 (2.4)	[redacted]	[redacted]	NR	NR
Gastrointestinal haemorrhage	NR	NR	4 (1.2)	1 (0.6)	[redacted]	[redacted]	NR	NR
Fatigue	NR	NR	23 (6.9)	10 (6.0)	[redacted]	[redacted]	NR	NR
Back pain	NR	NR	2 (0.6)	4 (2.4)	[redacted]	[redacted]	NR	NR
Blood (alkaline phosphatase	NR	NR	9 (2.7)	5 (3.0)				
Dyspnoea	NR	NR	6 (1.8)	6 (3.6)	[redacted]	[redacted]	NR	NR
Ascites	NR	NR	12 (3.6)	11 (6.5)	[redacted]	[redacted]	NR	NR
General health deterioration	NR	NR	22 (6.6)	15 (8.9)	[redacted]	[redacted]	NR	NR
y-[glutamylatransferase]	NR	NR	3 (0.9)	4 (3.0)	[redacted]	[redacted]	NR	NR

*This result includes both neutropenia and those with neutrophil count decreased. In the CSR these are reported separately.
√ represents where data is the same in all TAGs sources
AE, adverse event; NR, not reported; BSC, best supportive care.

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

No meta-analysis or indirect comparison was reported. The company states this is due to a paucity of relevant evidence in the third-line setting. The ERG agrees that an indirect comparison would not have been useful due to the quality and quantity of data available, and the infrequent use of chemotherapy in the third-line setting.

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

Not applicable.

3.5 Additional work on clinical effectiveness undertaken by the ERG

None, other than reported above.

3.6 Conclusions of the clinical effectiveness section

The ERG agrees that the CS has included all relevant trials, and that a network meta-analysis of the evidence relating to other third-line treatments was not feasible. The key trial (TAGS)[2] was a phase III double-blind, placebo-controlled, multicentre randomised controlled trial. It included all key outcomes identified in the NICE scope.

An OS advantage was observed (HR 0.69 (95% CI: 0.56–0.85, p <0.001)), with median survival of 5.7 months (95% CI: 4.8-6.2) in the TFT arm and 3.6 (95% CI: 3.1-4.1) in the placebo arm - a difference of 2.1 months. In subgroup analyses, for OS, patients with prior ramucirumab treatment had a HR of 0.76 (95% CI 0.53-1.09) and those without an HR 0.66 (95% CI 0.51-0.85). Patients from Japan had a HR of 0.77 (95% CI 0.46-1.30), and those from ROW had a HR of 0.68; (95% CI: 0.54-0.85). Patients from Europe had a HR of 0.67 (95% CI 0.53-0.86). The ERG requested an analysis of patients from Europe without prior ramucirumab treatment, for which the HR was not reported, [redacted]

A PFS advantage was observed (HR 0.57 (95% CI: 0.47–0.70, p<0.0001)), although the absolute benefit (difference in median PFS between arms) was only 0.2 months (2.0 months (95% 1.9-2.3) and 1.8 months (95% CI 1.7-1.9) in the TFT and placebo arms, respectively). The analysis was not adjusted for all prognostic factors.

Objective response rates were low (4.5% and 2.1% respectively) as would be expected in patients at third-line of treatment. Disease control rate was higher in the TFT arm than in the placebo arm (44.1%

and 14.5% respectively, p<0.0001), and was mostly driven by SD rather than CR or PR. HRQoL appeared largely maintained with TFT treatment.

Key adverse events were nausea, anaemia, decreased appetite, vomiting, diarrhoea, fatigue, neutropenia and asthenia thrombocytopenia. Anaemia and neutropenia were two outcomes where the incidence appeared to be markedly greater in the TFT group compared with the placebo group (anaemia 45% vs 19%, neutropenia 53% vs 4%).

The population recruited to TAGS was thought to be largely generalisable to the population in England, with some exceptions.

The study stratified patients according to prior exposure to ramucirumab and the company used the subgroup of patients without prior treatment with ramucirumab in the base case analyses within its model. The ERG has noted uncertainty in the clinical views about whether prior ramucirumab use affects the natural history of the disease, whilst two clinicians agreed that it is unlikely to affect the efficacy of TFT. Without a strong indication that prior ramucirumab treatment alters prognosis, the ERG prefers to assume that there is no impact but notes differences in the prior ramucirumab group and the no ramucirumab group in terms of prior lines of treatment and disease duration. Therefore, although this assumption is uncertain the ERG prefers an estimate of a HR or AF from the entire population rather than the non-the ramucirumab patients only

The inclusion of a larger proportion of Japanese patients in TAGS than are in the English population was potentially problematic as Japanese patients have a different natural history and treatment pathway than European patients. Whilst exclusion of the Japanese patients leads to an under-representation of Asians compared with the English population, their inclusion leads to over-representation, and the generalisability of Japanese patients to the more diverse Asian population in England is unclear. The ERG concludes that analyses of European patients, or where not available, the ROW have highest relevance to the decision problem.

There were also some minor imbalances in prognostic factors between arms at baseline. In the primary analysis of OS, adjustment for these factors did not affect the HR. [redacted]

There were more gastric patients (rather than gastroesophageal patients) than clinical advice to the ERG indicated would be expected, but clinical advice suggested that this is unlikely to affect estimates of efficacy.

In terms of the intervention, it was not entirely clear whether the discontinuation rules were mandatory or optional, and their application in a clinical setting may differ from that in a trial setting.

Within the trial, [redacted]

Patients may be assessed for treatment continuation more frequently (4- to 6-weekly rather than every two months) in England than in the TAGS study, and this could lead to earlier discontinuations, with an unknown impact on efficacy, adverse events and costs.

For objective response rate and disease control rate, no adjustments were made and it was not clear how missing data were handled. Because there was a statistically significant difference in efficacy for patients with measurable disease compared with those without measurable disease, and this analysis only included those with measurable disease, the data may not be generalisable to the whole population.

[redacted]

In conclusion, TFT appears to confer an overall survival advantage with a HR of 0.69 in the whole TAGS study population. The midpoint for the HR is lower in subgroup analyses that may be more applicable to England: 0.68 for the ROW group; 0.67 for the European group and 0.66 for those who had not received prior ramucirumab treatment. Other outcomes indicate that the key benefit of TFT is the gain in OS, as there is only a 0.2 month absolute difference in median PFS, although there was a clear improvement after 2 months int eh whole TAGS study population, and no improvement in HRQoL compared with baseline values.

4 COST EFFECTIVENESS

This section provides a structured critique of the economic evidence submitted by the company to support the cost effectiveness of TFT for mGC patients who have received two or more lines of treatment.

The two key components of the economic evidence presented in the CS are: (i) a systematic review of the relevant literature and (ii) a report of the company's *de novo* economic evaluation. The company also provided an electronic version of their economic model developed in Microsoft Excel. For brevity, in most cases treatment with TFT + BSC will be referred to as TFT, and placebo + BSC will be referred to as BSC.

4.1 ERG's comment on company's review of cost-effectiveness evidence

4.1.1 Objective of cost effectiveness review

The company undertook an SLR to identify published evidence to support the company's cost effectiveness model. Details of the search strategies employed by the company are provided in Appendix G of the CS.

Searches were conducted in June 2018 and covered an appropriate range of databases (MEDLINE & Medline-in-process; Embase; Econlit and NHS EED and HTAD); relevant conference series; and international HTA websites (see Table 1 below for more detail). The search strategies are generally well-designed, although as with the clinical SLR, the ERG notes the use of a multi-file search to interrogate Medline and Embase simultaneously, with some associated loss of functionality. However, it is unlikely that the SLR has missed any relevant studies.

No citation is provided for the search filters used to identify economic evaluations, although in their response to the ERG's clarification letter (question A9) the company explained that the terms used were based on filters developed by the Scottish Intercollegiate Guidelines Network (SIGN). While SIGN filters are not formally validated, the ERG recognises that they are expert-designed and likely to retrieve most of the studies eligible for inclusion.

A cut-off date of 10 years was applied; this was justified by the company on the grounds that considerable change had been observed in this period in terms of technology evolution and quality of care, and that 2008 was also the date of the earliest study identified in the clinical effectiveness review (Clarification response, question A8).

Additionally, the company undertook searches as needed to populate its economic model (including utility studies and cost/resource use studies (reported in Appendices H and I, respectively)). The same sources have been used as for the economic SLR, and again the ERG is broadly satisfied with the searches as reported in the relevant sections.

Table 8: Data sources for the economic systematic review

Search strategy component	Sources	Date limits
Electronic database searches Key biomedical electronic literature databases recommended by HTA agencies	-MEDLINE® -MEDLINE® In-process -Embase® -The Cochrane Library including National Health Service Economic Evaluation Database (NHS EED) -EconLit® -Health Technology Assessment Database (HTAD)	2008-2018
Conference proceedings	-International Society for Pharmacoeconomics and Outcomes Research (ISPOR) -American Society of Clinical Oncology (ASCO) -The American Society of Clinical Oncology Gastrointestinal Cancers Symposium (ASCO GI).	2016-2018
Key international HTA websites	-National Institute for Health and Care Excellence (NICE) -Scottish Medicines Consortium (SMC) -All Wales Medicines Strategy Group (AWMSG) -Haute Autorité de Santé (HAS) -Statens legemiddelverk (SLV)	Not specified

4.1.2 The inclusion and exclusion criteria used in the study selection

The inclusion/exclusion criteria used by the company to facilitate study selection are presented in

Table 9.

Table 9: Inclusion/exclusion criteria for the economic review

Category	Inclusion criteria	Exclusion criteria
Population (P)	-Age: adults aged ≥18 years -Gender: any -Race: any -Disease: patients with either unresectable advanced/metastatic GC who are at stage IIIb and IV according to the American Joint Committee on Cancer guidelines, or GEJ cancer	• Paediatric patients • Patients with early stage GC • Oesophageal cancer • Localised GC
Intervention (I)	All pharmacological interventions	Non-pharmacological interventions
Comparator (C)	• Any pharmacological intervention • Placebo • Best supportive care	None
Outcome (O)	Studies were not excluded based on the reported outcome	None
Study design	-All economic evaluation studies based on models -Cost-effectiveness analysis -Cost-utility analysis -Cost-minimisation analysis -Cost-benefit analysis -Budget impact models -Cost-consequence analysis	-Letters, comments and editorials -Studies reporting clinical data only -Simple costing analysis studies
Line of therapy	Third- or further-line of therapy	First- or second-line of therapy
Search timeframe	2008 to 2018	Studies published prior to 2008
Language	No restrictions	None

GC, gastric cancer; GEJ, gastroesophageal junction

4.1.3 Findings of the cost effectiveness review

Four studies were identified that were of relevance to the decision problem; however, none of these included TFT as an option. All four models used a three-state partitioned survival model within their analyses.

4.1.4 Conclusions of the cost effectiveness review

As the company's searches did not identify any relevant studies of TFT, they developed a *de novo* health economic model.

4.2 Summary of the company's submitted economic evaluation

4.2.1 Population

The population included in the company's health economic analysis reflects adult patients with mGC including adenocarcinoma of the GEJ who have received at least two prior lines of treatment. The modelled patient characteristics reflect those of the full patient population within the TAGS trial[2] with

an average age of 62.5 years, and 27% of the population are assumed to be female. However, the selected BSA distribution for the base case used the European patient cohort of the TAGS trial (with an average of 1.77 m^2) as this was deemed more clinically appropriate for patients in England.

4.2.2 Interventions and comparators

In the TAGS trial, TFT was administered in combination with BSC, where BSC is provided to alleviate cancer-related symptoms and maximise the patient's health-related quality of life. TFT was taken orally at a dose of 35 mg/m^2 twice daily for 10 days (1-5 and 8-12) per 28-day treatment cycle. The comparator was placebo in combination with BSC.

The company stated that no treatments have been recommended by NICE for mGC patients who had received two or more prior lines of treatment, and that chemotherapy (as included in the final NICE scope[5]) is rarely used for such population. This was confirmed by a clinician advisory board held by the company, and therefore TFT + BSC was only compared with BSC within the company's economic analyses.

4.2.3 Perspective, time horizon and discounting

The base case model adopts an NHS and Personal Social Services perspective. The base case model uses a 10-year time horizon; shorter values were included in the company's scenario analyses. Both costs and QALYs were discounted at 3.5% per annum as recommended by NICE.[20]

4.2.4 Model structure

As part of its submission to NICE, the company developed a fully executable partitioned survival model (PSM) that comprised three mutually exclusive and exhaustive health states: (i) progression-free (PF); (ii) progressed disease (PD); and (iii) death. The model is similar to that of other treatments for advanced/metastatic cancer previously submitted to NICE as part of the STA process. The health states and possible transitions between these are shown in Figure 7, with the arrows for remaining in the same state added by the ERG. A weekly cycle length was used; according to the CS, this obviated the need for half-cycle correction.

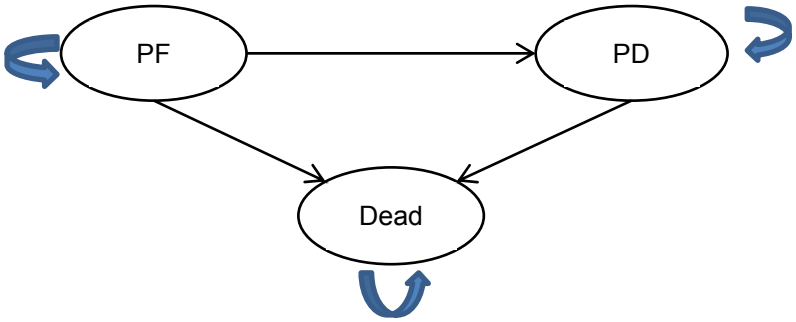


Figure 7: The company's model structure

All patients are assumed to enter the model in the PF health state and remain there until progression or death. As with a standard PSM, the transition probabilities between the health states are inferred via extrapolated PFS and OS curves fitted to the clinical trial data. For patients on TFT, parametric curves were fitted to time to treatment discontinuation data from the TAGS RCT in order to estimate the cost of TFT treatment. The company assumed that should the treatment discontinuation curve be higher than the progression-free survival (PFS) curve in the extrapolated period, then the discontinuation curve would be set equal to the PFS curve.

4.2.5 Evidence used to inform the company's model parameters

4.2.5.1 Treatment effectiveness and extrapolation in the base case

Data from the TAGS trial for patients with no prior ramucirumab treatment (n=222) were used for the extrapolation of PFS and OS in the TFT and BSC arms in the company's base case. At the time of data cut-off (30th April, 2018 for OS and 31st March, 2018 for all other clinical data), more than 90% patients in both arms had experienced the event of interest for both PFS and OS endpoints.

The company followed standard guidance for fitting and selecting survival models based on NICE Decision Support Unit (DSU) Technical Support Document (TSD) 14.[21] A full description of the survival extrapolation analyses undertaken by the company is presented in Section B.3.3.2 of the CS.[1]

The company investigated the use of a range of parametric survival models: exponential, Weibull, Gompertz, lognormal, log-logistic and generalised gamma. The company also explored two approaches to model the treatment effect: combined models (with a covariate for treatment assignment) and independent models (models were fitted independently to data for each treatment arm). Hence, 12 distinct extrapolations were available for use in each treatment arm for each of the PFS and OS dataset (6 from the combined modelling approach and 6 from the independent modelling approach). All survival analysis was performed using the "flexsurv" package in R.[22, 23]

4.2.5.1.1 Extrapolating OS

The company firstly assessed the appropriateness of using either a proportional hazards (PH) model or an accelerated failure time (AFT) model in the combined modelling approach for OS. The company concluded that exponential and Weibull PH models with a covariate for treatment assignment may be inappropriate and AFT models with a covariate for treatment assignment could be considered as appropriate. From assessment of the hazard plots, the company concluded that no specific models were ruled out, but the lognormal and log-logistic models may provide a better fit to the data than the exponential, Weibull and Gompertz models.

Based on the assessment of the visual fit (Figures 21 and 22 of the CS[1]), the statistical goodness-of-fit using Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) scores (Table 21 of the CS[1]), and long-term plausibility, the company selected the combined lognormal model as the base case for OS. Figure 8 presents the fitted lognormal model for both treatment arms. The company commented that the long-term extrapolation of OS using the combined lognormal was aligned with clinical expectation, with 5-year OS being 0.71% and 0.23% in the TFT and BSC arms, respectively, and 10-year OS being 0.08% and 0.02% in the TFT and BSC arms, respectively.

The company explored alternative survival models within its scenario analysis. The model assumes that the probabilities of death are always higher, or equal, to those in the general population at the corresponding age.[24]

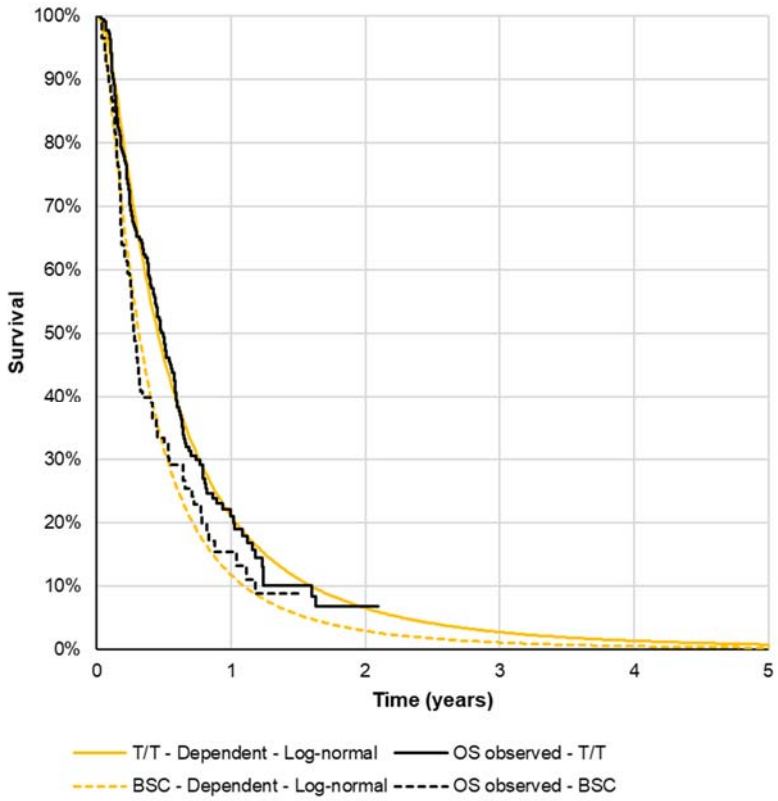


Figure 8: The company's base case OS extrapolation

4.2.5.1.2 Extrapolating PFS

From the assessment of the appropriateness of using either a PH model or AFT model in the combined modelling approach, it was concluded by the company that both PH models and AFT models with a covariate for treatment assignment may be inappropriate for PFS. From assessment of the hazard plots, the company concluded that the exponential, Weibull, Gompertz, lognormal and log-logistic models may be inappropriate as the hazard of a PFS event did not follow the trend associated with these parametric models.

The company selected the generalised gamma model fitted independently to both arms as the base case for PFS due to a good visual fit (Figures 31-32 of the CS[1]), and also that the long-term PFS extrapolation was aligned with clinical expectation (0.00% for both arms at 5-years). The company argued that although the generalised gamma model was not associated with the lowest AIC or BIC, and was considerably higher than some alternative models, it was the only model that allows for a more flexible hazard shape. Figure 9 presents the fitted generalised gamma curves for both treatment arms. The company explored alternative parametric survival models in scenario analyses.

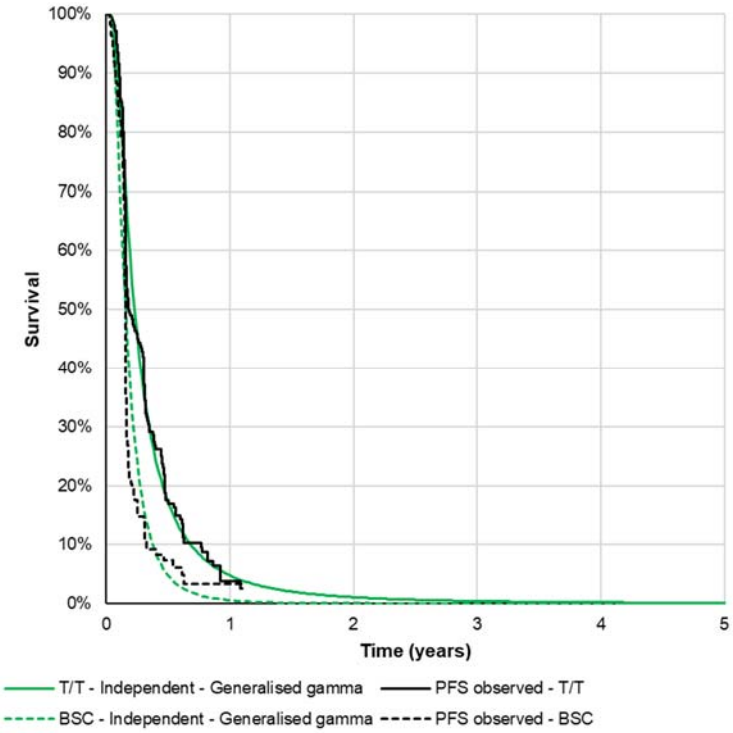


Figure 9: The company's base case PFS extrapolation

The company also explored the use of a hybrid Kaplan-Meier (KM) approach in scenario analyses because of the presence of "*kinks*" in the PFS curves due to the 8-weekly progression assessment visits. In these scenario analyses, PFS was modelled directly from the KM data up to a specified cut-point and following this time point, the failure rates associated with the parametric model within the base case was used. The company's scenario analyses tested different cut-points ranging from 12 weeks (selected as the minimum plausible cut-point) to 33 weeks (selected as the maximum plausible cut-point that represented nearly every observed PFS event).

During the clarification process, the ERG asked the company to provide the extrapolation for PFS using more flexible models such as Royston and Parmar[25] natural spline models because of the complex shape of the observed hazard (clarification response A28).[7] The company provided the extrapolation results using spline models with treatment as a covariate using 1, 3, 5 and 10 internal knots, and concluded that the hazard- and odds-based splines provided better fit than the normal-based splines and the fit improves with the increase in the number of knots (clarification response A28). The company argued that the spline models did not provide a substantial improved fit when compared with the generalised gamma model (the model chosen in the base case), and are expected to "over-fit" the data with 10 knots spline models.

The ERG also asked the company to estimate the ICER using the European population who have not had ramucirumab treatment (clarification question B4). The company did not describe the extrapolation analyses performed for this subpopulation group in the clarification response, but the extrapolation analyses results were provided within the submitted economic model. It was assumed that the types of statistical distribution chosen in the base case (i.e. dependent lognormal for OS and independent generalised Gamma for PFS) were appropriate in these analyses.

4.2.5.2 Treatment safety

AEs were included in the model to account for the potential cost and HRQoL burden of experiencing events whilst on treatment. Treatment-emergent grade III or IV AEs were included in the model where at least 5% of patients experienced them in one or more treatment arm within the TAGS trial. The only exceptions were the inclusion of febrile neutropenia and nausea. Febrile neutropenia was included due to its significant impact on HRQoL and costs; nausea was included based on expert opinion sought by the company. The incidence rates used to inform the economic model are presented in Table 23 of the CS. The company applied the impact of adverse events on costs and quality of life as one-off events for one cycle at the start of the model. The values are discussed in Section 4.2.5.4.

4.2.5.3 Duration of treatment

In the TAGS trial, the company reported that patients discontinued treatment on TFT mainly due to disease progression (73%) or suffering adverse events (10%). Treatment duration data were collected from individual patients and time to treatment discontinuation (TTD) KM curves were constructed for patients with no prior ramucirumab.

The six independent parametric survival models were fitted to the TTD data, and the generalised gamma model was selected for inclusion in the base case due to its good visual fit (Figure 36 of the CS[1]) and because it had lower AIC and BIC values than the other models (Table 36 of the CS[1]).

Figure 10 presents the TTD KM curve and the fitted generalised gamma model for patients on TFT with no prior ramucirumab experience used in the company's model base case. In order to preserve the structural correlation between progression and treatment discontinuation, the company capped the TTD curves according to the selected PFS curve.

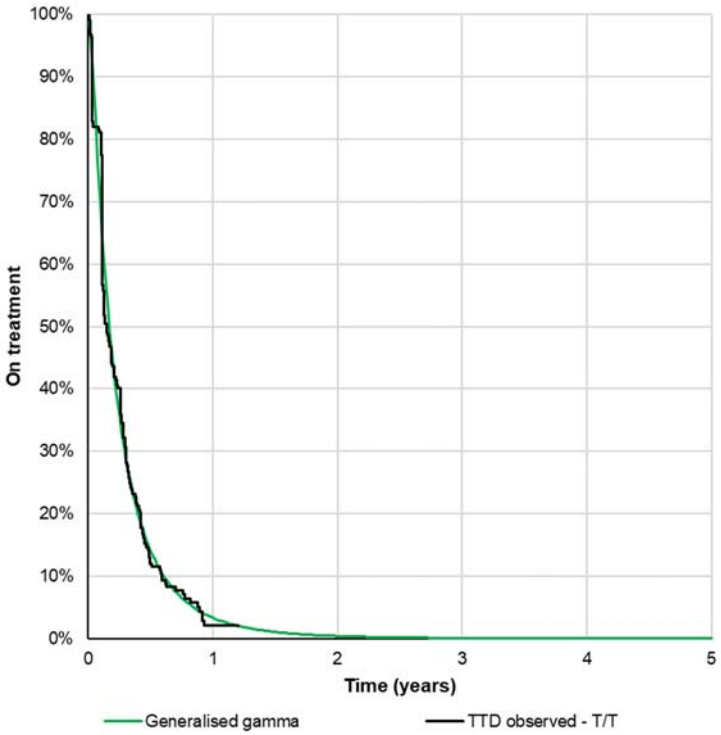


Figure 10: The company's base case TTD extrapolation

4.2.5.4 Health related quality of life

The SLR carried out by the company identified five unique HRQoL studies relevant to the technology appraisal; however, none of these considered relevant to the UK population as all were either Japanese or Chinese studies. In addition, four of the studies sourced the utility values from other studies or trials whose patient populations were considered irrelevant by the company to the decision problem.

HRQoL data were collected using the EORTC QLQ-C30 within the TAGS trial at three different time points, with compliance rates varying between 73% and 100%. A mapping algorithm was applied to these data to estimate the corresponding EQ-5D-3L values. The company used a published mapping algorithm (Kontodimopoulos *et al.*[26]) in its base case as it was the only published algorithm developed in a gastric cancer population according to the latest version of the University Of Oxford Health Economics Research (HERC) mapping database.[27] In an additional clarification question, the ERG questioned why this relatively small study (n=48) was used in preference to other mappings of EORTC-QLQ30 to EQ-5D-3L. This is discussed in detail within Section 5.3.4.2.

The company fitted generalised estimating equation (GEE) regressions using the "geepack" package[28] in R to the mapped utility values to account for the repeated observations for individual patients with four models considered: Model 1 with progression as a covariate; Model 2 with progression and treatment as covariates; Model 3 with progression and no prior ramucirumab as covariates; and Model

4 with progression, treatment and no prior ramucirumab as covariates. The goodness-of-fit statistic, Quasi-likelihood under Independence Model Criterion (QIC), was used to compare between the four models. The company stated that the inclusion of covariates for treatment and prior ramucirumab experience did not improve model fit and these were found to be statistically insignificant predictors of utility. Therefore, the company selected the simplest model (Model 1) whereby utility varies with progression status, and tested the other models in its scenario analyses.

During the clarification process, the ERG asked the company to clarify whether including known confounders, such as age and gender, in the GEE regression analysis would have an impact on the results (clarification question B6). In response, the company performed a regression analysis including age, gender and progression status as covariates.

[redacted]

In addition, utility values from previous relevant technology appraisals (TA208 and TA378) were considered for the scenario analysis. These were derived from EQ-5D data collected directly from the appraisals' main clinical trials. Table 10 summarises the six sets of utility values used in the company's economic model.

Table 10: The different sets of utility values included in the company's economic model

	Model 1 (*base case*)	Model 2	Model 3	Model 4	NICE TA378	NICE TA208
Line of treatment	3L+	3L+	3L+	3L+	2L+	1L
Covariates included	Progression	Progression and treatment	Progression and ramucirumab experience	Progression, treatment and ramucirumab experience		
PF utility	0.764	0.786	0.760	0.782	0.737	0.729
PP utility	0.652	0.672	0.647	0.667	0.587	0.577
TFT associated utility		-0.029		-0.029		

PF, progression-free; PP, post-progression

In addition to utility values associated with the two health states, the company's base case analysis applied utility decrements due to AEs. The proportion of patients experiencing a given AE was taken from the safety data of the TAGS trial, with the associated utility decrements being sourced from published literature. For any given AE, the associated QALY loss was calculated by multiplying the

utility decrement by the duration over which the AE impact was expected to last. Table 29 of the CS presents the disutility values and duration of AEs included in the company's base case analysis. The frequency of each event is provided in Section 5.2.5.5.5 in this report. AEs attributed to a QALY loss of 0.005 and 0.002 associated with TFT and BSC respectively in the company's base case, which were deducted in the first cycle of the model.

4.2.5.5 Resources and costs

The costs and resource use included in the base case model comprised: drug acquisition costs; drug administration costs; medical resource use (MRU) associated with TFT or BSC; off-treatment and post-progression related costs; AE costs; and end of life care costs. These are discussed in the following sections.

4.2.5.5.1 Drug acquisition costs

TFT tablets are available in two concentrations; 15mg trifluridine/6.14mg tipiracil and 20mg trifluridine/8.19mg tipiracil (referred to as "15mg" and "20mg" respectively) and in two package sizes of 20 and 60. The four formulations had the same cost of £33.33 ([redacted] including the Patient Access Scheme (PAS) discount) per 1mg trifluridine/0.41mg tipiracil.

TFT is administered at a dose of 35mg/m^2 of BSA twice daily on 10 days each 28-day treatment cycle. This represents the licensed dose as well as the dosing followed within the TAGS trial. Table 32 of the CS presents the BSA bands with the required tablets per dose. The company used the BSA distribution of the European population of the TAGS trial within its base case analysis with an average BSA of 1.77m^2. A lognormal distribution fitted to the BSA distribution was combined with a "method of moments" approach to give a weighted average cost of £2,184.01 for TFT per 28-day treatment cycle ([redacted] with the PAS discount). In answering clarification question B9, the company amended a calculation error resulting in an average cost of £2,017.47 ([redacted] with PAS) per 28-day treatment cycle,

Within the TAGS study, three levels of dose reduction were reported (from 35mg/m^2 to 30mg/m^2, 30mg/m^2 to 25mg/m^2, and 25mg/m^2 to 20mg/m^2). The dosing associated with each BSA band is detailed in Table 33 of the CS. Table 34 of the CS presents the number of patients by dosing level for 14 cycles of TFT treatment. Data from the TAGS trial regarding dose delays were also applied in the company's base case model where [redacted] of patients were assumed to start treatment in the second model cycle.

BSC was assumed to have no associated costs within the company's base case analysis; however, post-progression drug costs were considered in the model as detailed in Section 4.2.5.5.4.

4.2.5.5.2 Drug administration costs

Owing to its oral administration route, the company assumed that no medical resources are needed for TFT administration. Clinical advice provided to the company stated that some clinicians might send patients to a chemotherapy nurse before taking their first treatment cycle. Therefore, the company applied an administration cost of £22.5 (equivalent to 30 minutes of Band 6 nurse time) for the first treatment cycle within its model base case.

4.2.5.5.3 Medical resource use associated with treatment assignment

MRU data were estimated by the company based on consultation with clinical experts. These resources included oncologist consultations, computed tomography (CT) scans, and laboratory tests (full blood count, liver function test, and renal function test) as presented in Table 39 of the CS. The company stated that the MRU estimates in its base case analysis were different from those reported in the ramucirumab appraisal (TA378[29]) for reasons of following the TAGS trial protocol and avoiding potential double counting issues with AE costs or end of life care costs. For completeness, the company conducted a scenario where MRU resources were assumed the same as reported in TA378 as shown in Table 40 of the CS.

The company sourced MRU unit costs mainly from NHS Reference Costs 2017/18[30], the Commercial Medicines Unit (CMU) electronic Marketing Information Tool (eMIT), and inflated values using Personal Social Services Research Unit (PSSRU[31]) indices as appropriate.

4.2.5.5.4 Off-treatment and post progression related costs

In its base case, the company assumed that routine MRU costs for patients were based on treatment status (i.e. receiving TFT or not receiving TFT), as opposed to progression status. These included an oncologist consultation every 3 cycles of treatment. In the TA378 scenario, the company, as in the case with on-treatment MRU, also included costs of pain control, distress management, blood transfusion, and radiotherapy. These costs are detailed in Appendix 1; in the company's base case the costs were £211 per 28 days for patients receiving TFT and £54 per 28 days for people not receiving TFT.

Following progression in the TAGS trial, patients could undergo surgery, radiotherapy or continue onto further rounds of systematic anti-cancer treatment (SACT) which was assumed to involve a 3-cycle course of docetaxel. These costs were applied during the first model cycle following progression. Post-progression MRU estimates were extracted from the TAGS study with unit costs sourced from NHS Reference Costs 2017/18[30] and eMIT. In the company's base case, the average total costs incurred upon progression were £1,327 for patients who had received TFT and £1,532 for those who had received

BSC. The higher costs for BSC were due to the greater observed level of post-progression treatment in the BSC arm of the TAGS study.

4.2.5.5.5 AE costs

The rationale for the AEs included in the model is provided in Section 4.2.5.2. The costs associated with each were primarily sourced from NHS Reference Costs 2017/18.[30] Table 11 presents the frequency of AEs observed within the TAGS study and the costs associated with their management. This resulted in an average total cost of £306 and £87 to resolve AEs associated with TFT and BSC, respectively. These cost estimates were applied as a fixed sum within the model's first cycle.

Table 11: AE costs

Adverse event	Occurrence rates		Assumed cost to resolve	Source
	TFT	BSC		
Neutropenia	46%	0%	£164.55	Assumption of FBC cost + outpatient medical oncologist consultation (based on clinical experts' opinion)
Anaemia	26%	10%		
Decreased neutrophil count	26%	0%		
Leukopenia	7%	0%		
Abdominal pain	5%	11%	£319.68	NHS Reference Costs (2017/18[30]): Weighted average of day case abdominal pain with and without interventions (FD05A and FD05B)
Ascites	6%	8%		
Decreased appetite	10%	7%	£75.98	PSSRU 2018: Unit cost of a dietician appointment[31]
Fatigue	7%	6%	£0.00	Assumption of zero cost was based on clinical experts' opinion
Ascites	6%	8%		
Febrile neutropenia	2%	0%	£4,619.81	Wehler *et al.* (2017[32]) and inflated using PSSRU inflation indices[31]
Nausea	4%	3%	£163.58	NHS Reference Costs (2017/18[30]): Outpatient attendance – General Medicine

BSC, best supportive care; FBC, full blood count; TFT, trifluridine/tipiracil

For those patients in both the TFT and BSC arm who receive SACT upon progression, AE costs were set equal to those of people initially receiving TFT.

4.2.5.5.6 End of life care costs

In the company's base case, end of life care costs (health and social care costs) for colorectal cancer patients reported within Round *et al.* (2015[33]) were inflated. These were applied for all patients upon entry to the "Dead" health state. Alternative costs for cancer patients with lung, breast, and prostate cancers were also reported in the same publication. Based on clinical advice, the company decided that

the end of life care costs incurred by colorectal cancer patients were relevant to this appraisal; the costs of other cancer types were used in scenario analysis.

4.2.6 Model validation and face validity check

The company validated its economic model using two approaches. The first was holding a clinical advisory board attended by twelve UK practicing oncologists specialising in GC who validated the key aspects and assumptions of the model. The second approach was an internal quality control check of the company's model by a third party.

4.2.7 Cost effectiveness results

Following the clarification process the company submitted a revised version of the model that included updated estimates of the cost-effectiveness of TFT. All the results presented in this section and in Section 4.2.8 use the revised model and all results use the established price for TFT after consideration of the PAS. Table 12 shows the results of the company's base case analysis for both the deterministic and probabilistic versions of the model. The probabilistic sensitivity analyses (PSA) results are based on 10,000 iterations run by the ERG. Based on the probabilistic version of the model, TFT plus BSC is expected to generate 0.153 additional QALYs at an additional cost of £6,923, compared with placebo + BSC. The corresponding ICER is £45,314 per QALY gained. The deterministic version of the company's model produces a similar ICER of £45,164 per QALY gained.

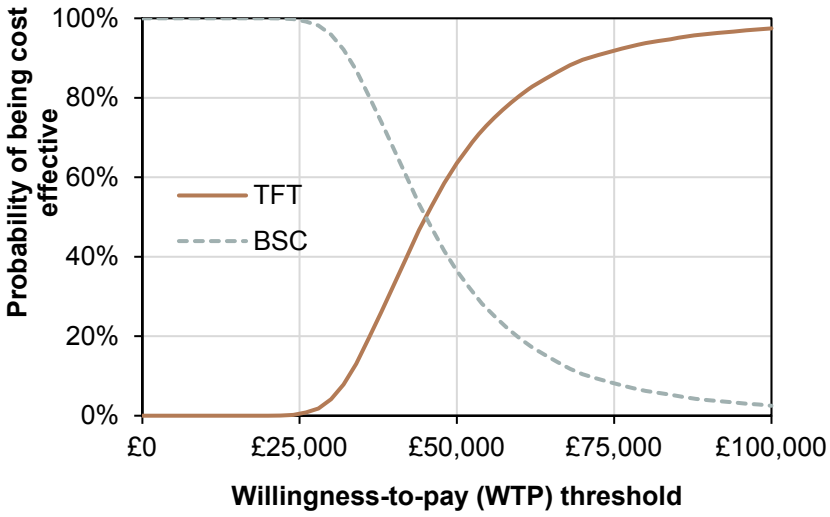


Figure 11 shows the cost-effectiveness acceptability curve (CEAC) produced by the ERG when running the company's base case. Figure 12 plots the PSA results on the cost-effectiveness plane. Figure 13 presents the resultant survival curves for the first five years of the company's model.

Table 12: The Company's base case results

Treatment	Total QALYs	Total Costs	ICER (£ per QALY gained)

Deterministic			
Placebo + BSC	0.349	[redacted]	
TFT + BSC	0.502	[redacted]	£45,164
PSA (run by the Evidence Review Group)			
Placebo + BSC	0.351	[redacted]	
TFT + BSC	0.504	[redacted]	£45,314

BSC, best supportive care; ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life year; TFT, trifluridine/tipiracil

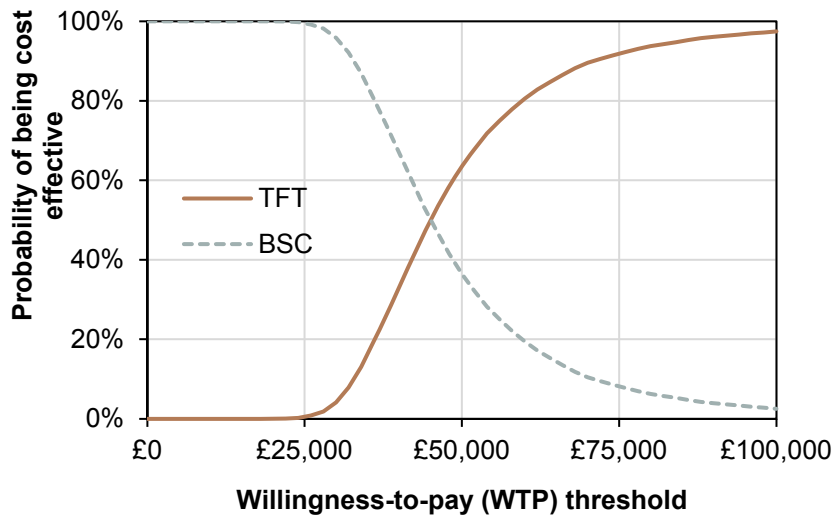


Figure 11: Company's base case cost-effectiveness acceptability curve

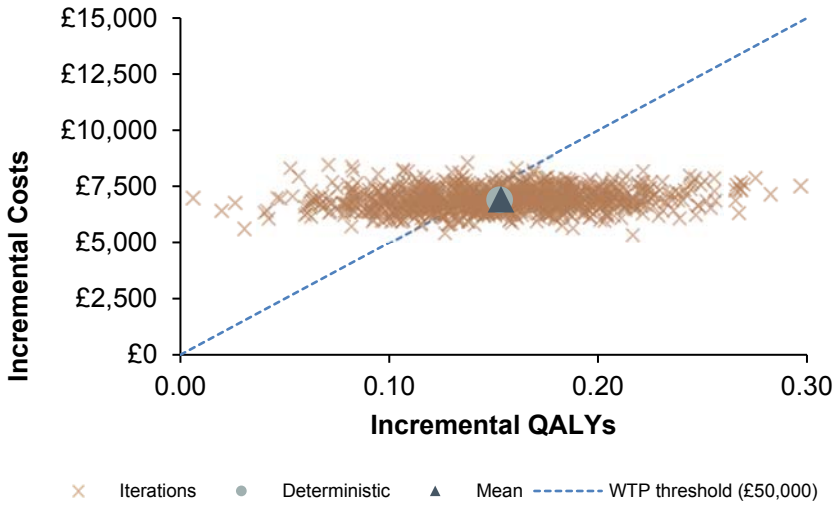


Figure 12: Company's base case cost-effectiveness plane

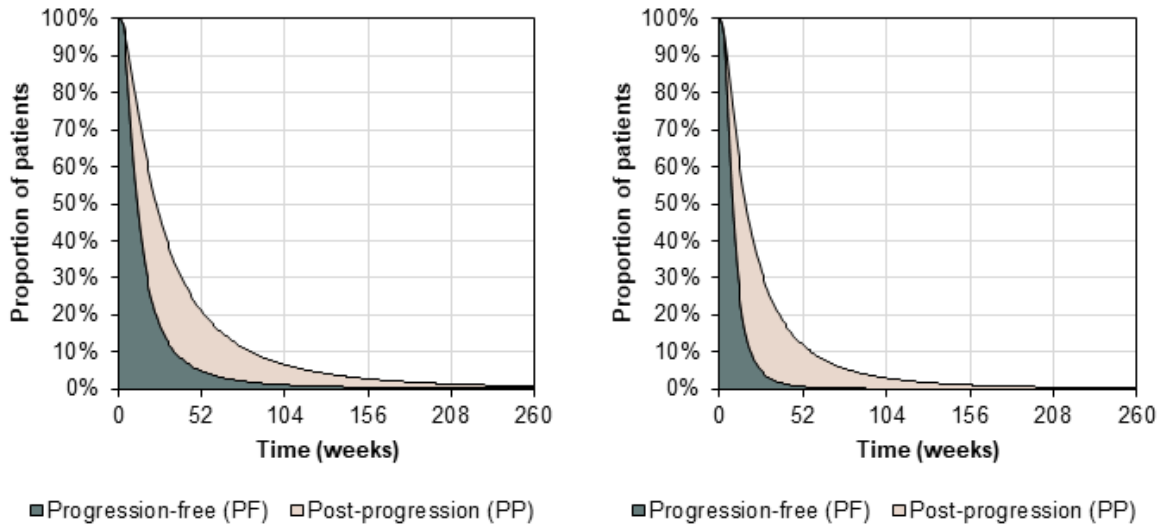


Figure 13: Company's base case survival curves (model traces) – BSC (left) and TFT (right)

4.2.7.1 Tornado diagrams

The company's tornado diagram, which shows the ten most influential parameters in terms of impact on ICER, is presented in Figure 43 of the CS. Within the tornado diagram, all parameters were varied between the upper and lower bounds of the 95% confidence intervals. The company stated that this analysis could not include parameters related to survival and utility as "*they are correlated (and so varying these in isolation would not be appropriate). Instead, the uncertainty associated with estimates of survival and utilities is discussed within the context of scenario analysis.*" The ERG noted that BSA was not incorporated in the company's original one-way sensitivity analysis. In its response to clarification question B8,[7] the company estimated the standard deviation around the two parameters (mu and theta) of the fitted lognormal distribution. The resultant uncertainty was included in both one-way and probabilistic sensitivity analyses. This had a small impact on the ICER value.

The most influential parameters in this analysis were related to MRU frequencies and costs. None of the ICERs on the tornado plot exceeded £50,000 per QALY gained.

4.2.8 Sensitivity analyses

The company conducted sensitivity analyses, which included: (1) a range of scenario analyses, which included the effects of alternative survival extrapolations and data on the results; and (2) exploring the use of KM curves at different cut-points followed by extrapolation for the PFS data.

4.2.8.1 Scenario and subgroup analyses

The company undertook several scenario analyses, which are presented in Tables 13 of the company's response to the clarification questions.[7] Generally, most scenarios produced ICERs that were similar to the company's base case ICER. Reducing the model time horizon resulted in non-linear increase in ICER values, and time horizons of two years or less resulted in ICERs, which were higher than £50,000 per QALY gained.

Eight out of the 12 tested survival parametric modelling scenarios of OS data presented in Table 13 of the company's clarification response[7] were associated with ICERs higher than £50,000 per QALY gained. These were: fitting independent exponential, independent generalised gamma, independent Gompertz, independent lognormal, independent Weibull, dependent exponential, dependent Gompertz, and the dependent Weibull. The company claimed that these curves provided a relatively poor fit to the KM curves.

All alternative PFS and TTD parametric models tested by the company resulted in ICERs, which were lower than £50,000 per QALY gained. The use of different spline-based models fitted to the PFS data produced ICERs, which were between £43,500 and £47,000 per QALY gained.

Within subgroup analyses, the company considered the entire TAGS population (with and without ramucirumab experience) as the data source for efficacy. Using the same survival models from the base case, this scenario produced an ICER, which was slightly higher than £50,000 per QALY gained. The ERG comments that the company did not undertake an assessment of which was the most appropriate curves to use in this population.

The ERG requested, in its clarification questions, a scenario analysis using only the European population with no prior treatment with ramucirumab. The company estimated new parameter values for the same curves fitted to the OS, PFS and TTD in the base case, which produced an ICER which was slightly above £49,000 per QALY gained. The company did not consider this scenario in its base case claiming that caution should be used when interpreting these results as this subgroup was not stratified for in the TAGS study.

4.2.8.2 Use of KM curves followed by parametric curve extrapolation for the PFS data

Due to the presence of "*kinks*" in the KM PFS data, the company conducted an additional scenario analysis to explore the impact of using the KM PFS curves for both treatment arms followed by the generalised gamma curves used within the base case analysis. The time cut-point between both sets of PFS curves were varied between 16 and 33 weeks in weekly increments. All cut-points produced ICERs that were between £44,000 and £45,200 per QALY gained.

4.3 Critique of company's submitted economic evaluation by the ERG

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

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

- Scrutiny of the company's model and discussion of issues identified amongst the members of the ERG.
- Examination of the correspondence between the description of the model reported within the CS and the company's executable model.
- Re-running the DSA and PSA presented within the CS.
- Where possible, checking the parameter values used in the company's model against their original data sources.
- The use of expert clinical input to judge the credibility of the company's economic evaluation and the assumptions underpinning the model.

4.3.2 Adherence of the company's model to the NICE reference case

As shown in Table 13, the company's economic evaluation is generally in line with the NICE reference case.[30]

Table 13: Adherence of the company's model to the NICE reference case[30]

Element	Reference case	ERG comments
Type of economic evaluation	Cost-utility analysis with fully incremental analysis	The CS met the NICE reference case.[30]
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	The CS met the NICE reference case.[30] A 10-year time horizon was adopted. By this point, almost 100% of simulated patients were dead.
Synthesis of evidence on health effects	Based on trial outcome data and systematic review	The CS met the NICE reference case.[30] Health outcomes are modelled using the data collected in the TAGS study. The base case used a subgroup with no prior ramucirumab experience.

Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of HRQoL in adults.	The CS met the NICE reference case.[30]
Source of data for measurement of health-related quality of life	Reported directly by patients and/or carers	The CS met the NICE reference case.[30]
Source of preference data for valuation of changes in HRQoL	Representative sample of the UK population	EORTC-QLQ-C30 data collected in the TAGS study were mapped to EQ-5D-3L values. The mapping algorithm used in the company's base case was developed in a gastric cancer population. However, none of the patients were suffering from metastatic cancer and the sample size was small (n=48).
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	The CS met the NICE reference case, although the company makes a case for the end of life criteria being met.
Evidence on resource use and costs	Costs should relate to NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS	The CS met the NICE reference case.[30]
Discount rate	The same annual rate for both costs and health effects (currently 3.5%)	The CS met the NICE reference case.[30]

4.3.3 ERG Critique of the modelling performed by the company

4.3.3.1 Model verification

The ERG checked and verified the implementation of the model and the methods for generating results. During this process, the ERG identified one minor implementation error, which was addressed by the

company in their clarification response to question B9. The implemented model appears to be generally in line with its description within the CS. KM curves were available for OS, PFS and TTD and provided in the model.

4.3.3.2 Correspondence of the model inputs and the original sources of parameter values

The ERG is satisfied that model parameters corresponded with their original source values. These were in line also with the parameter values reported in the CS. The only possible exception to this was potential discrepancies in the AE inputs, see Section 4.2.2.4. However, the ERG's exploratory analyses indicated that these issues would not affect the ICER significantly.

The ERG noted that the company's model uses arbitrary values to characterise the uncertainty in NHS Reference Costs by assuming 10% of the mean cost as its standard deviation and dividing it by the number of cases rather than number of data returns. In its response to clarification question B7,[7] the company reverted to 2012-13 NHS Reference Costs database to estimate the ratio of the standard error (SE) to the mean cost from the quartile data. The company subsequently concluded that a ratio of 5% could be used to account for the uncertainty of all NHS Reference Costs included in the model. The ERG highlights the relatively old NHS reference costs version used by the company and that the same ratio was used for all costs. However, the ERG expects these limitations to have minimal impact on the uncertainty in the probabilistic ICERs.

4.3.4 The main issues identified by the critical appraisal

Generally, the model was implemented well and the company provided reasonable responses to the ERG's clarification questions. However, the ERG identified four main issues within the model. These points are summarised in Box 1, with further details provided in the subsequent sections. The small number of issues are testament to the implementation of the decision problem by the company and the relative simple decision problem.

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

Summary of identified concerns within the company's health economic model:

1) Selection of the appropriate population for the base case analyses
2) Extrapolation of OS and PFS
3) The mapping of EORTC-QLQ30 to EQ-5D-3L
4) Exclusion of oral chemotherapy delivery fees

4.3.4.1 Selection of the appropriate population for the base case analyses

As indicated in Section 3.2.2.1, clinical advice sought by the ERG, and provided to NICE, suggested exposure to ramucirumab is not expected to influence the relative efficacy of TFT or prognosis. The company's base case uses the no prior ramucirumab population, which, therefore, may not be the most appropriate estimates for the purpose of decision making, although this population had fewer lines of previous treatment and a lower proportion of Japanese patients. The company's base case also included patients from Japan and the United States; clinical advice to the ERG suggested that a European subgroup may be more appropriate.

4.3.4.2 Extrapolation of OS and PFS

The ERG notes that the company considered a number of approaches in selecting its preferred base-case model, including the use of statistical goodness of fit, a quantile-quantile plot, a cumulative hazard plot, an empirical hazard plot, visual inspection and assessing the plausibility of longer-term projections. For extrapolating OS, the combined modelling approach provided lower AIC/BIC scores when compared with the independent modelling approach. However, the difference in scores were less than 3 points, hence it indicated that both models provided similar statistical goodness of fit to the data. By examining the plots for assessing the appropriateness of the combined modelling approach (with treatment as a covariate), the ERG believes that it was not clear that the combined modelling approach would be more appropriate for the OS. If the OS data were associated with a constant AF over time, the fitted survival curves would theoretically be the same using either the combined modelling or independent modelling approach (though this would be difficult to establish using "real" trial data, owing to limited sample sizes). The ERG notes that when using independent lognormal models increased ICER to above £50,000.

The company fitted spline models for PFS during the clarification process. The ERG notes that the combined modelling approach was used without justification, and the impact on the use of independent modelling approach is unclear.

The company, in its reply to clarification question B4,[7] chose the same parametric curves fitted for the whole TAGS trial population with no prior ramucirumab experience to be the selected curves for the European subpopulation (i.e. dependent lognormal distributions for the OS data, and independent generalised gamma distributions for the PFS data). The company did not mention the rationale their curve choice although AIC and BIC data were contained in the Excel model.

4.3.4.3 The mapping of EORTC-QLQ30 to EQ-5D-3L

The ERG noted that the mapping study used by the company, that of Kontodimopoulos *et al.*,[26] was derived from a small population (n=48) and whilst the patients all had gastric cancer, none had

metastatic cancer. The estimated utilities appeared to have a lack of face validity compared with those used in previous STAs where EQ-5D-3L data had been collected within the trial and where the patients were less heavily pre-treated (Table 10) as it would be expected that the PD state after third-line treatment would be lower than after first-, or second-line. Whilst the company identify limitations in the utilities collected in TA378 and TA208, (p146-147 of the CS[1]) this would not address the potential face validity concern.

A recent review of mapping algorithms from the EORTC-QLQ30 to the EQ-5D-3L[34] was identified by the ERG which stated that two algorithms were the best performing in external validation studies.[35, 36] These mapping algorithms use more of the EORTC-QLQ30 domains than Kontodimopoulos *et al.*[26] which only uses physical functioning, emotional functioning and global health status as predictors of EQ-5D-3L.

The ERG asked the company to provide ICERs when each of the two mappings were used, in order to inform the committee of the sensitivity of the results to the chosen mapping algorithm. However, this was not undertaken by the company for the following reasons. The company stated that the "*Versteegh et al. study does not utilise a UK tariff (and so is not aligned with the NICE reference case). However, outside of the tariff used, the study was conducted in only haematological cancers (multiple myeloma and non-Hodgkin's lymphoma)*" and that "*Longworth et al. primarily considers patients with multiple myeloma (n=572 of 771), as well as patients with breast or lung cancer (i.e. no patients with a gastrointestinal cancer). For the multiple myeloma cohort, patients were taken from the VISTA trial – a Phase III randomised open-label trial for newly-diagnosed patients. In the other two populations (breast and lung cancer), real-world data were collected from the Vancouver Cancer Clinic.*" The company quote guidance from NICE DSU TSD 10[37] stating that ""*… we recommend that careful consideration is given to the generalisability of the mapping function to the target population, including the range of disease severity over which the function was estimated and the potential for systematic differences in the populations that could impact on the health state utility values.*" (Section 3.2.5 of NICE DSU TSD 10[37]).

A similar conclusion is also made by Woodcock and Doble who state that "*The most appropriate mapping algorithm to apply in practice may depend on the disease severity of the patient sample whose utility values are being predicted.*" Both the NICE DSU TSD 10[37] and Woodcock and Doble[34] would lead the ERG to question whether the mapping algorithm from Kontodimopoulos *et al.*[26] which did not include patients with metastatic cancer would be appropriate in a population in which "*all patients had heavily pre-treated (i.e. two or more previous lines of therapy) metastatic gastric cancer*" and whereby the estimated life expectancy under current standard care was in the region of six months. The Kontodimopoulos *et al.* paper[26] states that "*No patients were suffering from metastases of the cancer to*

other organs, which could further affect their HRQoL negatively." The ERG notes, however, that an alternative mapping algorithm by Marriott *et al*.[38] which considers a metastatic colorectal cancer population was provided by the company. These values were higher for both PFS and PD than those generated using the mapping of Kontodimopoulos *et al.*[26]

The ERG is not contending that the mapping algorithms produced by Versteegh *et al.*[35] and Longworth *et al*.[36] are unquestionably better than that of Kontodimopoulos *et al.,*[26] and accept the criticisms of the alternative mappings put forward by the company. However, the ERG believes that the ICERs produced when these mappings are used would be informative to the committee and that the sensitivity analyses should have been performed. As the ERG does not have access to the data required to calculate utility estimates based on the alternative mapping algorithms, this remains an area of considerable uncertainty.

The ERG also notes that the compliance rate for filling in the EORTC QLQ-C30 was 84% and thus there may be the potential for responder bias within the study, however, if there was, the extent to which this would influence the results in unknown.

4.3.4.4 Exclusion of oral chemotherapy delivery fees

The company assumed in its base case analysis that there would be no administration costs regularly associated with TFT due to its oral route of administration. However, as NHS England noted in an recent STA[39] "*Trusts will regard [TFT] as chemotherapy and may charge the oral delivery tariff SB11Z (£120) each time [TFT] is given to patients*." This will be in addition to other consultation costs already included in the economic model. The ERG does not know NICE's position on this but contend that this may be seen as a transfer payment and has excluded this from the base case but has explored it within scenario analyses.

4.4 Exploratory analyses undertaken by the ERG

This section presents the methods of the ERG's exploratory analyses.

4.4.1 Exploratory analyses based on and whether it should be assumed that the prognoses of the patients and the efficacy of TFT are affected by prior ramucirumab use and amending the chosen geographical region for the analyses

The ERG has explored the impacts of alternative assumptions relating to prognoses of patients considered for TFT, whether prior ramucirumab use affects the HR for TFT compared to BSC, and which geographical region is most appropriate. Each of the three components had two choices, which leads to eight potential scenarios. These are summarised in Section 4.1.4.4, however beforehand each component will be detailed.

4.4.1.1 Exploring the relationship between prior ramucirumab treatment and prognoses

The company's base case assumes that prior ramucirumab treatment affects the survival of patients who would be considered for TFT. As such, the company use the OS data in its base case for patients who have not had prior ramucirumab treatment. Clinical advice provided to the ERG suggested that it was unclear whether prior ramucirumab leads to a different prognosis and that the OS related to all patients in the TAGS study[2] may be more appropriate. (See Section 3.2.2.2).

4.4.1.2 Exploring the relationship between prior ramucirumab treatment and the relative efficacy of TFT

The company's base case assumes that the most appropriate estimation of the relative efficacy of TFT is derived from the no prior ramucirumab group. As such, the company use the AF in its base case for patients who have not had prior ramucirumab treatment, which matches the prognosis group in Section 4.4.1.1. Clinical advice provided to the ERG and to NICE suggested that the HR or AF would be expected to be independent of prior ramucirumab use (See Section 3.2.2.2). The ERG believes that a more accurate estimate of the efficacy of TFT would therefore come from all patients irrespective of prior ramucirumab use.

4.4.1.3 Exploring the relationship between prior ramucirumab treatment and TFT relative efficacy on different populations

The company's base case assumes that the HR or AF would be independent of geographical region and combines patients from Japan and the ROW and uses this in its base case. However, it is argued in the CS that gastric cancer operates differently in Japanese patients and it is noted that results from Japanese patients cannot be assumed to be generalisable to non-Asian patients. (Section 3.1.5). Furthermore, the clinical advice provided to the ERG was that the EU subgroup would be more generalisable than a group including Japanese and American patients (see Section 3.2.2.3). Whilst an analysis of the Europe only geographical area breaks the stratification within the TAGS study,[2] the ERG notes that the study was stratified on Japanese vs the ROW and that the European component was approximately 95% of the ROW so it is anticipated that the inaccuracy caused by this limitation may not be large.

4.4.1.4 Summarising the eight potential scenarios

The eight scenarios are shown in

Table 14. The first scenario is the company's base case where prior ramucirumab is assumed to affect both prognosis and the relative efficacy of TFT and all geographical regions are used. The second scenario is a company scenario analysis, which uses all patients from the TAGS study. The third scenario assumes that prior ramucirumab treatment may affect disease prognosis, but does not affect the relative efficacy of TFT. In the fourth scenario, prior ramucirumab treatment does not affect disease prognosis but impacts on TFT relative efficacy. To run scenarios 3 and 4, an estimate of the treatment effect is taken from scenarios 2 and 1 respectively, meaning that only the use of dependent models could be explored. Scenarios 1 to 4 are replicated for the European population in Scenarios 5-8. However, the ERG did not have the data required to explore Scenarios 6 to 8. The company provided the data for Scenario 5 in its response to clarification question B4.[7]

Table 14: The eight scenarios defined by the ERG

Scenario	Is prognosis from non-ramucirumab patients most appropriate?	Is the HR or AF from non-ramucirumab patients most appropriate?	Is the entire TAGS study population more appropriate than the European geographical area?
1*	✓	✓	✓
2†	✗	✗	✓
3	✓	✗	✓
4	✗	✓	✓
5†	✓	✓	✗
6☯	✗	✗	✗
7	✓	✗	✗
8	✗	✓	✗

*The company base case; †A company scenario analysis; ☯Tentative ERG base case.

4.4.2 Analyses exploring the uncertainty in survival curve fits

The ERG selected the three best fitting survival distributions to run the analysis for each of the eight scenarios defined in

Table 14. These were determined based on the AIC/BIC scores provided in Table 21 of the CS for the overall population with no prior ramucirumab, and in the revised model for the European subgroup. These were the lognormal, the Weibull and the log-logistic for the all geographical region analyses, and the lognormal, the log-logistic and the generalised gamma for the European population. Where possible, independent models were evaluated as well as dependent models.

4.4.3 Impact of alternative mapping studies

As indicated in Section 4.3.4.3, the impact of using different mapping algorithms was not explored by the company. The ERG noted that the company's preferred mapping study did not include metastatic patients; the ERG believes that it is likely that these patients would have a lower utility than patients without metastases. The ERG performed an analysis using the values from the most recent STA for advanced gastric cancer or gastro-oesophageal junction adenocarcinoma which were 0.729 for those in the PFS and 0.587 for those in the PD state. (Table 10)

4.4.4 Impact of including oral chemotherapy delivery fees

As indicated in Section 4.3.4.4, the ERG explored the inclusion of the oral delivery tariff for chemotherapy (SB11Z) for outpatient setting in its scenario analyses. This resulted in a cost of £131.61 applied every 28 days for patients receiving TFT.

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

This section presents the results of the ERG's exploratory analyses.

5.1 The impact of selecting different populations and using different curve fits

Table 15 shows the results for the scenarios 1-5 (as defined in

Table 14). The ICERs varied considerably among different scenarios and models. Generally, independent models gave higher ICERs, whereas log-logistic models produced lower ICERs followed by lognormal and Weibull models. The company selected a dependent lognormal to use within the base case as it had the lowest AIC, and the ERG agrees that the lognormal is more likely to be appropriate than the log-logistic or the Weibull although the other two distributions remain plausible. The ERG prefers the use of independent models because the use of independent modelling approaches avoids making any assumptions about constant HR or AF over time, and would provide the same fitted curves as to the combined modelling approach if either assumption holds.

The ERG could not produce ICERs for its tentative base case (Scenario 6) but notes that when moving from Scenario 1 to 2, the ICER increases by £4,000 to £5,000. In the absence of further evidence, it may be appropriate to assume that this level of increase would also apply when moving from Scenario 5 to Scenario 6. Such calculations would indicate ICERs of over £64,000 when using independent models and in excess of £50,000 per QALY gained when assuming dependent models. However, the analyses would need to be undertaken to provide an accurate estimation. As the model appeared linear, only deterministic analyses have been run by the ERG.

Table 15: ERG's exploratory analysis regarding the impact of prior ramucirumab treatment and geographical region

Scenario	Independent models			Dependent models		
	Lognormal	Log-logistic	Weibull	Lognormal	Log-logistic	Weibull
1	£51,642	£46,942	£61,310	£45,164*	£42,208	£58,363
2	£55,600	£52,655	£66,137	£50,191	£47,449	£64,318
3				£50,278	£47,750	£65,129
4				£45,076	£41,926	£57,652
	Lognormal	Log-logistic	Generalised gamma	Lognormal	Log-logistic	Generalised gamma
5	£68,061	£59,564	£169,370	£49,067	£45,068	£46,024
6	Not evaluable due to data unavailability					
7						
8						

*Company's base case

Appendix 2 presents the results of

Table 15 in terms of differential costs and QALYs. The different scenarios and model selection have little impact on cost differences but have a proportionately higher impact of the difference in QALYs. In Scenario 5 the ICER from the independent generalised gamma models was markedly larger than other fits

[redacted]

5.2 Impact of decrementing utility values due to patients having metastatic disease

The impact of using utility values from TA378 increased the ICER. The company's base case ICER increased from £45,164 to £47,857 and in Scenario 5, using independent lognormal models, from £68,061 to £70,905 per QALY gained.

5.3 Impact of including the oral administration delivery fees

In a scenario analysis, the ERG explored the impact of adding the delivery fees as detailed in Section 4.4.3. This increased the differential costs between the two compared interventions by approximately [redacted] and increased the company's base case ICER from £45,164 to £48,592.

6 END OF LIFE

The company puts forward the case, in Section B.2.13 of its CS, that TFT meets the NICE End of Life criteria. These criteria are:

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

The ERG believes that TFT meets the first criterion because the mean life years associated with BSC in the model was 0.514 years (6.2 months) for patients without prior ramucirumab use.

Whether TFT meets the second criterion is less straightforward as the life extension associated with TFT estimated in the model is 0.226 years (2.7 months) which is below the 3 months normally required. The company cite precedent in two prior NICE appraisals to support their case, although only one had a survival extension of less than 3 months. This is nab-paclitaxel for (untreated) metastatic pancreatic cancer (TA476[40]) which was estimated to have a mean life extension of 2.4 months, but was assumed to meet the end of life criteria due to the short median life expectancy without treatment of 6.6 months. The company states that the proportional life improvement associated with TFT in mGC and GEJ is superior to that estimated in TA476. For completeness the ERG has reproduced the text from TA476. "*The committee noted that the survival data were mature and therefore considered that the survival gain estimate was robust. It recognised that this survival gain should be considered in the context of the very poor prognosis for metastatic pancreatic cancer. The committee noted that the survival gain was below what is normally considered appropriate for the extension-to-life criterion to be met (that is, it was less than 3 months). However, it agreed that the survival gain was particularly important relative to the average survival of people with this condition, and therefore this criterion could be accepted as met in this circumstance. The committee concluded that, for the comparison with gemcitabine monotherapy, nab-paclitaxel plus gemcitabine met the criteria to be considered a life-extending end-of-life treatment.*" The ERG leaves the decision on whether TFT meets the second criterion to the NICE Appraisal Committee.

Table 16 presents the life expectancy gains associated with TFT and other relevant technologies appraised in different scenarios.

Table 16: Survival gain associated with TFT and with the precedent cited by the company

Scenario	OS associated with SoC (months)	OS associated with the appraised technology (months)	OS gained with the technology (months, % gained)
The TAGS trial (no prior ramucirumab, whole population)	Median: 3.3	Median: 6.0	2.7 (82%)
The TAGS trial (whole population regardless of ramucirumab use)	Median: 3.6	Median: 5.7	2.1 (58%)
The company's base case model (no prior ramucirumab, whole population)	Mean: 6.2	Mean: 8.9	2.7 (44%)
The company's model (whole population regardless of ramucirumab use)	Mean: 6.2	Mean: 8.5	2.3 (37%)
TA476 main trial results	Median: 6.6	Median: 8.7	2.1 (32%)
TA476 economic model results	Mean: 8.7	Mean: 11.1	2.4 (28%)

OS, overall survival; SoC, standard of care

7 OVERALL CONCLUSIONS

The TAGS study reported a HR 0.69 (95% CI 0.56–0.85) for OS, and a difference in median survival of 2.1 months between arms. Analyses adjusted for relevant prognostic factors gave a similar HR. PFS was also positively affected, with a HR 0.57 (95% CI 0.47–0.70) and a 0.2 month difference between arms. Small benefits were reported for response rates and duration of response as may be expected given the stage of disease, however, the disease control rate was significantly improved. Health related quality of life was shown to be largely maintained with TFT treatment.

In subgroup analyses, for OS, patients with prior ramucirumab treatment had a HR of 0.76 (95% CI 0.53-1.09) and those without a HR of 0.66 (95% CI 0.51-0.85). Patients from Japan had a HR of 0.77 (95% CI 0.46-1.30), and those from ROW had a HR of 0.68; (95% CI 0.54-0.85). Patients from Europe had a HR of 0.67 (95% CI 0.53-0.86). The ERG requested an analysis of patients from Europe without prior ramucirumab treatment, for which a HR was not reported, [redacted]. PFS subgroup analyses were largely similar to the main analysis of PFS.

Clinical advice to the ERG and NICE suggested that there is no strong indication that prior ramucirumab treatment affects prognosis, and was unlikely to affect the efficacy of TFT. Clinical advice also indicated that European patients would have the highest generalisability to the decision problem due to biological and/or treatment pathway differences between Europe, the USA and particularly, Japan.

As shown within the ERG exploratory analyses the company's base case ICER is one of the lower estimates amongst the analyses undertaken by the company and the ERG. Factors that increase the ICER include: the use of independent rather than dependent models; assuming that prior use of ramucirumab does not affect prognosis; assuming that prior use of ramucirumab does not affect the efficacy of TFT; and assuming a European geographical area rather than the full TAGS study. The clinical advice provided to the ERG and NICE resulted in a tentative base case being put forward by the ERG (Scenario 6). This scenario could not be evaluated, but it is expected to have ICERs which are higher than those for Scenario 5 whereby the ICER was approximately £68,000 when an independent lognormal model was used and £49,000 when a dependent lognormal model was used. The ERG prefers the use of independent curves rather than dependent ones.

The ERG notes that the company explored an alternative mapping in sensitivity analysis, but declined to explore the alternative mappings proposed by the ERG. The ERG believes that sensitivity analyses should have been performed that assessed the impact of these mapping algorithms on the ICER. The

ERG performed exploratory analyses that showed that reducing the assumed utility value in the PFS and PD state increased the ICER.

In summary, Based on the analyses provided by the company and the ERG's exploratory analyses the ERG believes that the cost per QALY gained of TFT compared with BSC is likely to be in excess of £50,000. Whilst, the ERG's tentatively preferred scenario could not be evaluated, many component factors such as: using independent curves; assuming that prior ramucirumab use does not affect prognosis; assuming that prior ramucirumab use does not affect the relative treatment effect of TFT; using a European population; and reducing utility values, all increase the ICER (£45,164). The ERG notes that some of these factors, in isolation, increase the ICER to greater than £50,000 per QALY gained.

8 REFERENCES

1. Servier Laboraties. Trifluridine–tipiracil for treating metastatic gastric or gastro-oesophageal junction cancer after 2 or more therapies [ID1507]. Company evidence submission. 2019.
2. Shitara K, Doi T, Dvorkin M, Mansoor W, Arkenau H-T, Prokharau A, *et al.* Trifluridine/tipiracil versus placebo in patients with heavily pretreated metastatic gastric cancer (TAGS): a randomised, double-blind, placebo-controlled, phase 3 trial. *The Lancet Oncology* 2018;19:1437-48.
3. Davidson M, Cafferkey C, Goode EF, Kouvelakis K, Hughes D, Reguera P, *et al.* Survival in Advanced Esophagogastric Adenocarcinoma Improves With Use of Multiple Lines of Therapy: Results From an Analysis of More Than 500 Patients. *Clinical Colorectal Cancer* 2018;17:223-30.
4. Kang Y-K, Boku N, Satoh T, Ryu M-H, Chao Y, Kato K, *et al.* Nivolumab in patients with advanced gastric or gastro-oesophageal junction cancer refractory to, or intolerant of, at least two previous chemotherapy regimens (ONO-4538-12, ATTRACTION-2): a randomised, double-blind, placebo-controlled, phase 3 trial. *The Lancet* 2017;390:2461-71.
5. National Institute for Health and Care Excellence. Trifluridine–tipiracil for treating metastatic gastric or gastro-oesophageal junction cancer after 2 or more therapies. Final scope. 2019.
6. British Medical Association. British National Formulary: March -September 2019. 2019.
7. Servier Laboraties. Trifluridine–tipiracil for treating metastatic gastric or gastro-oesophageal junction cancer after 2 or more therapies [ID1507]. Clarification Response. 2019.
8. Centre for Reviews and Dissemination. CRD's guidance for undertaking reviews in health care.: University of York.; 2008.
9. Li J, Qin S, Xu J, Xiong J, Wu C, Bai Y, *et al.* Randomized, Double-Blind, Placebo-Controlled Phase III Trial of Apatinib in Patients With Chemotherapy-Refractory Advanced or Metastatic Adenocarcinoma of the Stomach or Gastroesophageal Junction. *Journal of Clinical Oncology* 2016;34:1448-54.
10. NICE. Single technology appraisal: User guide for company evidence submission template. Instructions for companies.; 2017.
11. Higgins JPT, Savovic J, Page MJ, Sterne JAC. Revised Cochrane risk-of-bias tool for randomized trials (RoB 2) TEMPLATE FOR COMPLETION. *https://sitesgooglecom/site/riskofbiastool/welcome/rob-2-0-tool* 2019.
12. Li J, Qin S, Xu J, Guo W, Xiong J, Bai Y, *et al.* Apatinib for chemotherapy-refractory advanced metastatic gastric cancer: results from a randomized, placebo-controlled, parallel-arm, phase II trial. *J Clin Oncol* 2013;31:3219-25.
13. Bando H, Doi T, Muro K, Yasui H, Nishina T, Yamaguchi K, *et al.* A multicenter phase II study of TAS-102 monotherapy in patients with pre-treated advanced gastric cancer (EPOC1201). *European Journal of Cancer* 2016;62:46-53.
14. Laboratories S. Clinical Study Report TAGS. 2018.
15. Osoba D, Rodrigues G, Myles J, Zee B, Pater J. Interpreting the significance of changes in health-related quality-of-life scores. *Journal of clinical oncology* 1998;16:139-44.
16. Mansoor W. Trifluridine–tipiracil for treating metastatic gastric or gastro-oesophageal junction cancer after 2 or more therapies [ID1507]. Clinical expert statement. 2019.
17. Smyth E. Trifluridine–tipiracil for treating metastatic gastric or gastro-oesophageal junction cancer after 2 or more therapies [ID1507]. Clinical expert statement. 2019.
18. Office of National Statistics. Ethnicity and National Identity in England and Wales: 2011.; 2012. https://www.ons.gov.uk/peoplepopulationandcommunity/culturalidentity/ethnicity/articles/ethnicityandnationalidentityinenglandandwales/2012-12-11 (Accessed 06.08.19).
19. CTCAE Files.; 2019. https://evs.nci.nih.gov/ftp1/CTCAE/About.html (Accessed 07.08.19).
20. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal 2013. 2013.

21. Latimer NR. NICE DSU Technical Support Document 14: Survival Analysis For Economic Evaluations Alongside Clinical Trials - Extrapolation With Patient-Level Data. School of Health and Related Research, University of Sheffield, UK 2011.
22. Jackson C, Metcalfe P, Amdahl J. Flexible Parametric Survival and Multi-State Models. R package version 1.1.1. 2019.
23. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. . 2013.
24. Office of National Statistics. Dataset: National life tables: UK. https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/nationallifetablesunitedkingdomreferencetables (Accessed June).
25. Royston P, Parmar M. Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. . *Stat Med* 2002;21:2175-97.
26. Kontodimopoulos N, Aletras VH, Paliouras D, Niakas D. Mapping the cancer-specific EORTC QLQ-C30 to the preference-based EQ-5D, SF-6D, and 15D instruments. *Value Health* 2009;12:1151-7.
27. Health Economics Research Centre. HERC Database of mapping studies. http://www.herc.ox.ac.uk/downloads/herc-database-of-mapping-studies. (Accessed June).
28. Søren H, Ulrich H, Yan JP. Package 'geepack': Generalized Estimating Equation Package Version 1.2-1. 2016.
29. NICE. Ramucirumab for treating advanced gastric cancer or gastro–oesophageal junction adenocarcinoma previously treated with chemotherapy [TA378]. 2016. https://www.nice.org.uk/guidance/ta378/resources/ramucirumab-for-treating-advanced-gastric-cancer-or-gastrooesophageal-junction-adenocarcinoma-previously-treated-with-chemotherapy-82602795958981 (Accessed May).
30. NHS Improvement. Reference costs. 2018. https://improvement.nhs.uk/resources/reference-costs/ (Accessed 29 July 2019).
31. Personal Social Services Research Unit. Unit Costs of Health and Social Care. 2015.
32. Wehler E, Zhao Z, Pinar Bilir S, Munakata J, Barber B. Economic burden of toxicities associated with treating metastatic melanoma in eight countries. *The European journal of health economics : HEPAC : health economics in prevention and care* 2017;18:49-58.
33. Round J, Jones L, Morris S. Estimating the cost of caring for people with cancer at the end of life: A modelling study. *Palliative medicine* 2015;29:899-907.
34. Woodcock F, Doble B, Consortium C. Mapping the EORTC-QLQ-C30 to the EQ-5D-3L: An Assessment of Existing and Newly Developed Algorithms. *Med Decis Making* 2018;38:954-67.
35. Versteegh M, Leunis A, Luime J, Boggild M, Uyl-de Groot C, Stolk E. Mapping QLQ-C30, HAQ, and MSIS-29 on EQ-5D. . *Med Decis Making* 2012;32:554-68.
36. Longworth L, Yang Y, Young T, Mulhern B, Hernández Alava M, Mukuria C*, et al.* Use of generic and condition-specific measures of health-related quality of life in NICE decision-making: a systematic review, statistical modelling and survey. . *Health Technol Assess* 2014;18:1-224.
37. Longworth L, Rowen D. NICE DSU Technical Support Document 10: The Use of Mapping Methods to Estimate Health State Utility Values 2011.
38. Marriott E-R, van Hazel G, Gibbs P, Hatswell A. Mapping EORTC-QLQ-C30 to EQ-5D-3L in patients with colorectal cancer. . *J Med Econ* 2017;20:193-9.
39. NICE. Olaparib for maintenance treatment of recurrent, platinum-sensitive ovarian, fallopian tube and peritoneal cancer that has responded to platinum-based chemotherapy (including a review of technology appraisal no. 381) [ID1296]. 2019. https://www.nice.org.uk/guidance/indevelopment/gid-ta10303 (Accessed 13.08.19).
40. NICE. Paclitaxel as albumin-bound nanoparticles with gemcitabine for untreated metastatic pancreatic cancer. TA476.; 2017. https://www.nice.org.uk/guidance/TA476 (Accessed 29 July 2019).

9 APPENDICES

Appendix 1: Summary of Medical Resource Costs used within the model

Table 17 summarises the frequencies of MRU used in the company's base case and scenario analyses and the resultant costs applied per 28-day treatment cycle. Differences between the scenario analysis and the base case are underlined in the scenario analysis data.

Table 17: MRU frequencies used in the company's base case and scenario analyses per treatment and progression status

MRU item	Company's base case				Company's scenario analysis (TA378)			
	TFT + BSC		Placebo + BSC		TFT + BSC		Placebo + BSC	
	PF on treatment	PF off treatment or PP	PF on treatment	PF off treatment or PP	PF on treatment	PF off treatment or PP	PF on treatment	PF off treatment or PP
Oncologist consultations	1.00	0.33	0.33	0.33	1.00	0.33	0.33	0.33
CT scan	0.50	0.00	0.00	0.00	0.33	0.00	0.00	0.00
FBC	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00
LFT	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00
RFT	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00
Pain control*	0.00	0.00	0.00	0.00	471.5	704.5	704.5	704.5
Distress management†	0.00	0.00	0.00	0.00	2.52	4.06	4.06	4.06
Blood transfusion**	0.00	0.00	0.00	0.00	0.08	0.22	0.22	0.22
Radiotherapy††	0.00	0.00	0.00	0.00	0.13	0.11	0.11	0.11
Total MRU cost***	£210.88	£54.02	£54.02	£54.02	£395.65	£324.45	£324.45	£324.45

BSC, best supportive care; CT, computerised tomography; FBC, full blood count; LFT, liver function test; PF, progression-free; PP, post-progression; RFT, renal function test; TFT, trifluridine/tipiracil

* Average number of mg of morphine required per patient per 28-day treatment cycle. This is based on 42.1% of patients on TFT requiring 40 mg of morphine per day versus 62.9% of those who are not on TFT.

† Average number of cognitive behavioural therapy (CBT) sessions undergone per patient per 28-day treatment cycle. This is based on 10.5% of patients on TFT undergoing six CBT sessions per week versus 16.9% of those who are not on TFT.

** Average number of red blood cell (RBC) transfusions required per patient per 28-day treatment cycle. This is based on 8.8% of patients on TFT requiring 1 RBC transfusion per month versus 23.8% of those who are not on TFT.

†† Average number of fractions of radiotherapy required per patient per 28-day treatment cycle. This is based on 14.0% of patients on TFT requiring 1 radiotherapy fraction per month versus 11.9% of those who are not on TFT.

*** Total cost was calculated by multiplying the item cost by its frequency of usage.

Appendix 2: Detailed Results for the exploratory analyses undertaken by the ERG

Table 18 provides more detailed results for each of the scenarios and statistical fits explored by the ERG than in the main document.

Table 18: Granular results for each of the scenarios and statistical fits explored by the ERG

Treatment	Total QALYs	Total Costs	ICER (£ per QALY gained)
Scenario 1 (independent lognormal models)			
Placebo + BSC	0.360	[redacted]	
TFT + BSC	0.493	[redacted]	£51,642
Scenario 1 (independent log-logistic models)			
Placebo + BSC	0.377	[redacted]	
TFT + BSC	0.524	[redacted]	£46,942
Scenario 1 (independent Weibull models)			
Placebo + BSC	0.328	[redacted]	
TFT + BSC	0.440	[redacted]	£61,310
Scenario 1 (dependent lognormal models)			
Placebo + BSC	0.349	[redacted]	
TFT + BSC	0.502	[redacted]	£51,642
Scenario 1 (dependent log-logistic models)			
Placebo + BSC	0.367	[redacted]	
TFT + BSC	0.531	[redacted]	£42,208
Scenario 1 (dependent Weibull models)			
Placebo + BSC	0.326	[redacted]	
TFT + BSC	0.444	[redacted]	£58,363
Scenario 2 (independent lognormal models)			
Placebo + BSC	0.354	[redacted]	
TFT + BSC	0.471	[redacted]	£55,600
Scenario 2 (independent log-logistic models)			
Placebo + BSC	0.372	[redacted]	
TFT + BSC	0.496	[redacted]	£52,655
Scenario 2 (independent Weibull models)			
Placebo + BSC	0.329	[redacted]	
TFT + BSC	0.427	[redacted]	£66,137
Scenario 2 (dependent lognormal models)			
Placebo + BSC	0.346	[redacted]	

TFT + BSC	0.476	[redacted]	£50,191
Scenario 2 (dependent log-logistic models)			
Placebo + BSC	0.364	[redacted]	
TFT + BSC	0.501	[redacted]	£47,449
Scenario 2 (dependent Weibull models)			
Placebo + BSC	0.328	[redacted]	
TFT + BSC	0.429	[redacted]	£64,318
Scenario 3 (dependent lognormal models)			
Placebo + BSC	0.347	[redacted]	
TFT + BSC	0.477	[redacted]	£50,278
Scenario 3 (dependent log-logistic models)			
Placebo + BSC	0.365	[redacted]	
TFT + BSC	0.502	[redacted]	£47,750
Scenario 3 (dependent Weibull models)			
Placebo + BSC	0.324	[redacted]	
TFT + BSC	0.423	[redacted]	£65,129
Scenario 4 (dependent lognormal models)			
Placebo + BSC	0.348	[redacted]	
TFT + BSC	0.502	[redacted]	£45,076
Scenario 4 (dependent log-logistic models)			
Placebo + BSC	0.366	[redacted]	
TFT + BSC	0.531	[redacted]	£41,926
Scenario 4 (dependent Weibull models)			
Placebo + BSC	0.330	[redacted]	
TFT + BSC	0.449	[redacted]	£57,652
Scenario 5 (independent lognormal models)			
Placebo + BSC	0.363	[redacted]	
TFT + BSC	0.462	[redacted]	£68,061
Scenario 5 (independent log-logistic models)			
Placebo + BSC	0.379	[redacted]	
TFT + BSC	0.493	[redacted]	£59,564
Scenario 5 (independent generalised gamma models)			
Placebo + BSC	0.417	[redacted]	
TFT + BSC	0.457	[redacted]	£169,370
Scenario 5 (dependent lognormal models)			

Placebo + BSC	0.342	[redacted]	
TFT + BSC	0.480	[redacted]	£49,067
Scenario 5 (dependent log-logistic models)			
Placebo + BSC	0.358	[redacted]	
TFT + BSC	0.508	[redacted]	£45,068
Scenario 5 (dependent generalised gamma models)			
Placebo + BSC	0.356	[redacted]	
TFT + BSC	0.503	[redacted]	£46,024

BSC, best supportive care; ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life year; TFT, trifluridine/tipiracil

National Institute for Health and Care Excellence

Centre for Health Technology Evaluation

ERG report – factual accuracy check

Trifluridine–tipiracil for treating metastatic gastric or gastro-oesophageal junction cancer after 2 or more therapies [ID1507]

You are asked to check the ERG report to ensure there are no factual inaccuracies contained within it.

If you do identify any factual inaccuracies you must inform NICE by **5pm on Thursday 5 September 2019,** using the below comments table. All factual errors will be highlighted in a report and presented to the Appraisal Committee and will subsequently be published on the NICE website with the committee papers.

The factual accuracy check form should act as a method of detailing any inaccuracies found and how and why they should be corrected.

Issue 1 Population most relevant to the decision problem

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
On page 9, the ERG states: *"Ramucirumab does not have a positive NICE recommendation, which means patients in England will largely be ramucirumab-naïve. There were mixed views from clinical advisors to the ERG and NICE about whether prior ramucirumab treatment would alter prognosis, but agreement that as ramucirumab and TFT work differently there should be no impact on treatment efficacy. In the absence of a strong indication that prior ramucirumab treatment alters prognosis, the ERG assumes there is no impact, though this is uncertain. The ERG therefore expects that the best estimate of a HR or Acceleration Factor (AF) would be from the entire population rather than the non- ramucirumab patients only."* Servier agrees that there is uncertainty concerning whether or not prior exposure to ramucirumab is a treatment effect modifier for trifluridine/tipiracil. However, there are other differences in the populations with and without prior ramucirumab exposure that are expected to influence both prognosis and treatment efficacy. One of the most important of these factors is the number of prior treatment lines – patients with prior ramucirumab exposure on average have a greater number of prior treatment lines than those without prior exposure. For clarity, please see below a tabulated summary of the number of prior lines for each subgroup (based on the ITT population):	To ensure the ERG report appropriately reflects the differences in these populations (i.e. that the differences extend beyond simply whether or not patients were previously treated with ramucirumab), Servier requests that the ERG revise its text to the following: *"In the absence of a strong indication that prior ramucirumab treatment alters prognosis, the ERG assumes there is no impact, though this is uncertain. The ERG therefore expects that the best estimate of a HR or Acceleration Factor (AF) <u>may</u> be from the entire population rather than the non- ramucirumab patients only.* ***<u>However, the ERG also notes that the non-ramucirumab population has different characteristics versus the prior ramucirumab group.</u>***	Servier appreciates that there is limited evidence concerning the impact of prior ramucirumab treatment on the efficacy of trifluridine/tipiracil, and consequently that *ceteris paribus*, it would be preferable to inform the estimation of relative efficacy from the broader patient population (were there no clear rationale for the use of the 'no prior ramucirumab' subgroup). However, Servier contends that there is a clear justification for the use of this subgroup that extends beyond treatment history alone – the 'no prior ramucirumab' population is less heavily pre-treated versus the 'prior ramucirumab' population, which is expected to influence prognosis and treatment efficacy. The TAGS trial was not powered to detect a difference in efficacy by subgroup according to the number of prior treatment lines, yet it is expected that in general, patients with an improved prognosis are expected to have an increased capacity to benefit from treatment with trifluridine/tipiracil.	We have amended the text using some of the suggested text. However, we believe it is unknown whether treatment duration or depth of pre-treatment affects the HR of TFT

Number of prior lines	Prior ramucirumab?	
	Yes (n=169)	No (n=338)
2	[redacted]	[redacted]
3	[redacted]	[redacted]
≥4	[redacted]	[redacted]

If the population of patients with ≥4 prior lines are assumed to have only received 4 prior lines, the average number of prior lines for each subgroup may be estimated as [redacted] (no prior ramucirumab) versus [redacted] (prior ramucirumab).

The intended position of trifluridine/tipiracil is UK practice is for patients with at least two prior lines of therapy. However, as heard by clinical experts (consulted by both Servier and the ERG), a very small percentage of UK patients are expected to be treated with third-line (off-label) chemotherapy. As such, the population of primary relevance to the decision problem would ideally comprise of predominantly third-line (i.e. two previous lines) patients. The 'no prior ramucirumab' comprises a less heavily pre-treated population versus the 'prior ramucirumab' population, yet on average still considers a population with more than 2 prior lines – [redacted] of patients in the 'no prior ramucirumab' population received three or more prior lines, versus [redacted] of the prior ramucirumab population.

In addition to the above, the patient-level data from the TAGS trial demonstrates a shorter time between diagnosis/confirmation of metastases and randomisation for patients not receiving ramucirumab versus those that were previously treated (see tables below):

Treatment arm	Prior ramucirumab?		
	Yes (n=169)	No (n=338)	All (n=507)

They are less heavily pre-treated and their disease duration is shorter, factors which are known to alter treatment response. Therefore, the estimate of a HR or AF from the non-ramucirumab population (a factor stratified for at randomisation) may also be appropriate to consider."

Revision of the text to reflect this feature of the patient population (and consequently, interpretation of the most relevant group(s) to consider within analysis) is expected to improve the clarity of the ERG's report.

Time since diagnosis TFT [redacted] Placebo [redacted] All [redacted] *Time since confirmed metastases* TFT [redacted] Placebo [redacted] All [redacted] This finding is to be expected, given that within the context of fewer treatment options being available (in this case, ramucirumab not being available), patients would be expected to receive TFT sooner. Conversely, where ramucirumab is routinely prescribed, patients would likely receive treatment with TFT later and in potentially a more progressive disease state. The latter statement in the ERG's paragraph is misleading, as all other potentially-important variables that are affected through the removal/addition of patients according to prior ramucirumab exposure are not discussed.			
On page 11, the ERG states: *"It is unknown which subgroups based on prior ramucirumab treatment and geographical region are of most relevance to the decision problem."* This statement is misleading, and does not appropriately reflect the uncertainties surrounding the patient population. The subgroup of most relevance is known – i.e. this would be a subgroup of purely UK patients predominantly treated in the third-line setting who have not previously been treated with ramucirumab, with all relevant characteristics reflecting the UK population.	In order to more accurately reflect the uncertainties surrounding the patient population, Servier requests the text of page 11 to be changed accordingly: Page 11: *"**There is uncertainty surrounding which** subgroups are of most relevance to the decision problem.*	Servier notes that this statement is a simplification of the broader question concerning the population of most relevance to the decision problem, which includes several factors including (but not limited to) prior ramucirumab treatment and geographical region. The relevance of a given subgroup to the decision problem extends beyond purely whether or not a given patient has previously received	We have amended the text so that it is clear that it is which subgroup's results are of most relevance rather than the subgroup. We have also listed the two most relevant subgroups, but have reserved mention of disease duration and lines of treatment for the main text.

The ERG is correct to highlight that there is uncertainty concerning which subgroups within the TAGS trial best reflect this population for the purpose of decision making, but the ideal group to consider within the context of the decision problem is known. Further to this, the subgroup(s) to consider from the TAGS trial should be determined through a broad range of factors, including but not limited to those stated by the ERG.	***Subgroups of interest include those based on prior ramucirumab treatment and geographical region, as well as other related factors such as the number of prior treatment lines**."*	ramucirumab or is within a given geography. For example, as highlighted in response to clarification question A3, the 'no prior ramucirumab' European population exhibit a lower proportion with ≥4 prior lines (for the T/T arm, 23% versus 13% and for the PBO arm, 27% versus 11%). In addition, by removing patients with prior ramucirumab exposure, a large proportion of patients treated with non-recommended options (such as immunotherapies) are also omitted from the analysis (for the T/T arm, 7% versus 4% and for the PBO arm, 4% versus 2%). Correction of the wording surrounding this criticism is intended to clarify the potential uncertainties surrounding the patient population, and highlight that the two characteristics highlighted by the ERG are not an exhaustive list of potentially relevant parameters. No impact on cost-effectiveness results are noted based upon this change.	
On page 30, the ERG state: *"*[The increased number of prior treatment lines in the TAGS study] *may, however, impact negatively on median survival compared to an English population at third line therapy"* Servier considers it important to clarify that while there are several issues with comparing survival outcomes	This difference in patient populations is particularly important within the context of determining the most relevant population. As such, Servier requests that the text be revised in line	As discussed above, the relationship between the number of prior lines and exposure to ramucirumab is extremely important to consider when determining the population of most relevant to the decision problem. Servier therefore requests	We have changed the text to be similar to that suggested by the company

across patient populations according to line of therapy, in general it is accepted that average survival generally decreases as patients progress through sequential lines of therapy. Written in isolation of prior ramucirumab use, this statement is misleading.	with the following: *"It may, however, impact negatively on median survival compared to an English population at third line therapy**, and the ERG notes that the prior ramucirumab population considers a population with more prior lines of treatment versus the no prior ramucirumab population**."*	that this be addressed by the ERG in its report, so that the reader may understand that neither of these features of the patient population should be considered independently of the other (given that they are intrinsically linked). Amending the report will have no impact on the cost-effectiveness results, but will improve the transparency of the key issues affecting the TAGS trial population.	
On page 32, the ERG state: *"Without a strong indication that prior ramucirumab treatment alters prognosis, the ERG prefers to assume that there is no impact."* Similar text is also reported on page 50 of the ERG report: *"Without a strong indication that prior ramucirumab treatment alters prognosis, the ERG prefers to assume that there is no impact. However, this assumption is uncertain. The ERG expects that the best estimate of a HR or AF would be from the entire population rather than the non-the ramucirumab patients only"* Servier acknowledges that this statement concerns the ERG's preference. However, for the avoidance of doubt, Servier believes this statement should be amended to highlight that prior ramucirumab treatment alone may not influence prognosis, yet the characteristics of the "no prior ramucirumab" versus "prior ramucirumab" subgroups may (and indeed are expected to) affect prognosis – e.g. number of prior	Servier requests that the ERG includes a minor revision to the text such that the interpretation applies specifically to this patient characteristic, rather than the subgroup of the TAGS trial (for which other characteristics are not necessarily balanced). Please see suggested text below: *"Without a strong indication that prior ramucirumab treatment alters prognosis, the ERG prefers to assume that there is no impact **associated with this characteristic specifically (noting that there are***	As discussed previously, it is important to distinguish between the impact of ramucirumab exposure specifically (i.e. in isolation of all other differences) and the subgroup analyses of the TAGS trial based on ramucirumab exposure. By amending the report, the distinction between these related yet distinct topics may be clearly presented. Amending the report will have no impact on the cost-effectiveness results, but will improve the transparency of factors which may influence survival.	We have changed the text to be similar to that suggested by the company

lines, proportion of Japanese patients. Servier considers this an important distinction to make within the ERG's report.	***differences in the subgroups of the TAGS trial based on prior ramucirumab exposure)****."*		
On page 32, the ERG states: *"In a further clinical expert statement provided to NICE it was commented, "there is no reason why prior ramucirumab would alter the outcome for trifluridine-tipiracil. They work on completely different pathways and cross resistance would not be expected." However, the ERG notes that the TAGS study was stratified based on prior ramucirumab use which suggests that it was believed that prior ramucirumab use could affect the relative efficacy of TFT. As such, whilst the relative efficacy of TFT in patients with and without prior ramucirumab treatment is unknown, the ERG expects that the best estimate of a HR or AF would be from the entire population rather than from only patients who had not received ramucirumab"* As communicated in response to clarification question A18, stratification factors were selected at the time of trial design, shortly after ramucirumab became available in the EU. It was an important stratification factor to select to account for differences in availability (due to reimbursement) in numerous clinical care pathways. Servier speculates that patients who had not received prior ramucirumab are less heavily pre-treated and therefore could be expected to have better outcomes (noting that in the TAGS trial, both patients with and without prior ramucirumab had a significant improvement in overall survival). The statement included in the ERG's report implies that Servier believes previous ramucirumab treatment may	For transparency, Servier requests the text be amended to the following: *In a further clinical expert statement provided to NICE it was commented, "there is no reason why prior ramucirumab would alter the outcome for trifluridine-tipiracil. They work on completely different pathways and cross resistance would not be expected." The ERG notes that the TAGS study was stratified based on prior ramucirumab use,* ***yet the company notes that there is no evidence to suggest prior ramucirumab exposure is a treatment effect modifier. Communicated in response to clarification question (A18), the company highlighted that stratification factors were selected at the time of trial design, shortly after ramucirumab became***	Servier acknowledges that stratification factors are typically used to identify subgroups based on factors expected to be potential treatment effect modifiers. However, in the case of the TAGS trial, this is not the case for prior ramucirumab exposure. Amending the report will have no impact on the cost-effectiveness results, but will improve the explanation concerning stratification in the TAGS trial.	We felt it simpler to delete the sentence, and slightly modify the paragraph.

be a treatment effect modifier. This is factually inaccurate – Servier has previously stated that there is insufficient evidence to ascertain whether or not ramucirumab exposure is a treatment effect modifier. For clarity, Servier does not (and previously has not) claimed the prior ramucirumab is a treatment effect modifier for mGC patients treated with TFT.	***available in the EU, and that prior ramucirumab treatment was considered an important stratification factor to select to account for differences in availability (due to reimbursement) in numerous clinical care pathways.*** *As such, whilst the relative efficacy of TFT in patients with and without prior ramucirumab treatment is unknown, the ERG expects that the best estimate of a HR or AF would be from the entire population rather than from only patients who had not received ramucirumab"*		
On page 32, the ERG state: *"In its clarification response, the company asserted that Japanese patients should be included because England has an 8% Asian population (clarification response A22[7]).."* While this was one point raised with respect to why Japanese patients were suggested to be included within the analysis, several other points were also raised in response to clarification question A22. Servier noted that by excluding patients with prior treatment with ramucirumab, the majority of Japanese patients were also excluded (i.e. of the n=73 total Japanese patients, only [redacted] were not previously treated	To ensure the points made by Servier in response to this clarification questions are appropriately reflected in the ERG's report, the following revised text is proposed: *"In its clarification response, the company asserted that Japanese patients should be included because England has an*	Servier understands that the TAGS trial is not a perfect representation of the UK patient population. However, Servier considers it extremely important that the differences in populations are appropriately reflected within the ERG's report, and that the point raised by Servier is considered within the context in which it was raised – i.e., that the proportion of Japanese patients in Servier's preferred base-case analysis is relatively low (~[redacted]) and	The ERG has included reference to the [redacted] Japanese patients in the no prior ramucirumab population as a response to clarification question A22, and noted that this represents an under-representation of Asians compared with the UK population.

with ramucirumab). More specifically, the response stated: *"This table illustrates that the majority of Japanese patients were previously treated with ramucirumab, and are therefore excluded within the base-case analysis presented (i.e. the 'non-ramucirumab' population)."* For the 'no prior ramucirumab' population, only [redacted] of the population are Japanese, which is substantially lower than the proportion of the ITT population which is 14.4%. The statement included by the ERG misrepresents the points made by Servier in its response to this clarification question (which include a tabulated summary of patient numbers by region).	*8% Asian population **and that after removing the population of patients with prior ramucirumab exposure, only n=[redacted] Japanese patients remained within the sample (equivalent to [redacted] of the 'no prior ramucirumab' subgroup)** (clarification response A22[7]).*" The ERG may also wish to highlight the proportion of Japanese patients in the ITT population, but Servier suggests such a revision to the text be made at the discretion of the ERG.	that the UK general population includes a larger proportion of Asian patients (~7.5%), which although are not all Japanese, highlights that a purely European population would also not be perfectly representative of the UK patient population. Amending the report will have no impact on the cost-effectiveness results, but will ensure the ERG's report presents a true reflection of the points made by Servier in its response to the clarification question asked.	
On page 50, the ERG states: *"The inclusion of a larger proportion of Japanese patients in TAGS than are in the English population was potentially problematic as Japanese patients have a different natural history and treatment pathway than European patients."* This statement is only true with respect to the ITT population – the opposite is true for the subgroup analysis preferred by Servier. More specifically, in the 'no prior ramucirumab' population only [redacted] of the population are Japanese, which is substantially lower than the proportion of the ITT population which is 14.4%.	For the avoidance of doubt, the following revised text is proposed by Servier: *"The inclusion of a larger proportion of Japanese patients in TAGS **ITT population** than are in the English population was potentially problematic as Japanese patients have a different natural history and treatment pathway than European patients.* [redacted]	As above, Servier considers it important to clarify to which groups statements raised by the ERG apply to the TAGS population. Amending the report will have no impact on the cost-effectiveness results, but will ensure the ERG's report presents a true interpretation of differences in the TAGS and UK populations.	Text has been amended to note the proportions in the whole trial population and the no prior ramucirumab population, and how these relate to the proportion in England.

	[redacted]		
On page 73, the ERG states: *"As indicated in Section 3.2.2.1, the company's base case suggested that prior ramucirumab plays a role in both prognoses (in terms of overall survival) and the efficacy of TFT compared with BSC; however, clinical advice sought by the ERG, and provided to NICE, suggested that this may not be the most appropriate population. The company's base case also included patients from Japan and the United States; clinical advice to the ERG suggested that a European subgroup may be more appropriate."* As previously highlighted, Servier's position is that prior ramucirumab may or may not have an impact on the relative efficacy of TFT, though this is unknown. In addition, the base-case analysis presented implies the 'no prior ramucirumab' subgroup has a differential prognosis and exhibits different relative efficacy as this population includes key differences versus the 'prior ramucirumab' subgroup (i.e. less heavily pre-treated, lower proportion of Japanese patients). This paragraph in the ERG report conflates the patient characteristic with the TAGS subgroup, which in turns leads to misinterpretation of the evidence base.	To ensure that patient characteristics and subgroups analyses are considered appropriately, Servier requests the text be changed in accordance with the following: *"As indicated in Section 3.2.2.1, clinical advice sought by the ERG, and provided to NICE, suggested* ***exposure to ramucirumab is not expected to influence prognosis or treatment effect specifically. The company's base-case considers the 'no prior ramucirumab' population, which may not be the most relevant population to the decision problem, yet this population includes a less heavily-pre-treated and lower proportion of Japanese patients.*** *The company's base case also included*	As previously highlighted, Servier notes that the separation of issues relating to the subgroup analyses and potentially influential patient characteristics is paramount to understanding which subgroups and/or characteristics are important to acknowledge within the interpretation of the clinical- and cost-effectiveness of TFT. Revision of the text has no impact on cost-effectiveness results, yet aims to clarify these differences.	We have changed the text to be similar to that suggested by the company.

	patients from Japan and the United States; clinical advice to the ERG suggested that a European subgroup may be more appropriate."		
On page 76, the ERG states: *"The company's base case assumes that prior ramucirumab treatment is a treatment effect modifier on TFT and that patients with previous ramucirumab use would have a different HR or AF, compared with those who did not have prior ramucirumab use."* Though the AF is derived for the subgroup of patients without prior ramucirumab treatment, Servier does not claim the difference in treatment effect is entirely related to the use of prior ramucirumab. The statement currently suggests that this is the only reason for a difference (as described for a number of other previous statements). For brevity, please refer to the other similar clarifications raised earlier in this table for a more detailed description of why Servier considers it extremely important to distinguish between patient characteristics and the 'no prior ramucirumab' subgroup.	For transparency, Servier requests the text be amended to the following: *"The company's base case assumes that* ***the most appropriate estimation of the relative efficacy of TFT is derived from the 'no prior ramucirumab' group****."*	The revised description is aligned with the analysis presented, and (as described previously) distinguishes between patient characteristics and properties of a given patient subgroup. This revision does not impact cost-effectiveness results, but clarifies what is claimed within the submitted analysis.	We have changed the text to be similar to that suggested by the company.

Issue 2 Use of mapping algorithms to inform the estimation of utility values

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Within the submitted model, the mapping algorithm by Kontodimopoulos *et al.* was applied to inform the estimation of utility values. At the clarification stage, the ERG requested the use of alternative mapping functions which Servier does not consider appropriate to inform the model (with an explanation provided in response to the first additional clarification question A1). Instead, Servier provided a sensitivity analysis using an alternative published mapping algorithm by Marriott *et al.*, and provided an explanation as to why this mapping algorithm is considered more appropriate. In the ERG's report, it is stated and/or implied in several sections that an analysis using an alternative mapping algorithm was not provided by Servier – for example on page 10: *"The ERG preferred the company to use other mapping algorithms to judge the impact on the ICER but the company declined."*. In addition, on page 84 it is stated: *"The ERG notes that the company declined to explore other mappings than the one used in the base case, despite this mapping being in patients with gastric cancer but without metastatic disease."* This is factually inaccurate – utility values based upon an alternative mapping algorithm (Marriott *et al.*) were provided to the ERG in response to an additional clarification request. The ERG's critique of the submitted model contains no description of the analysis provided using the Marriott *et al.* mapping	Servier requests that the ERG amends discussion concerning the provision of alternative mapping algorithms to acknowledge that alternative values were provided using the mapping algorithm by Marriott *et al.* Please see proposed amendments below: Page 10: *"The ERG* ***requested the company to apply two alternative mapping algorithms (Versteegh et al. and Longworth et al.)*** *to judge the impact on the ICER.* ***The company did not consider these algorithms appropriate to inform the model owing to differences in the patient populations used to derive the mapping algorithms, and instead provided an analysis using utility values from a mapping algorithm by Marriott et al.****"* Page 84: *"The ERG notes that the company* ***explored an alternative mapping in sensitivity analysis (by Marriott et al.), yet declined to explore the alternative mappings proposed by the ERG****."*	By amending this error, the ERG report will exhibit a true reflection of the analyses provided to the ERG to inform its critique of Servier's submission, as well as those not provided. Servier understands that the ERG requested the use of specific alternative mapping algorithms, though the current text within the report implies no alternatives were provided. Furthermore, Servier understands that the ERG may agree or disagree with the explanation provided concerning why alternative mapping algorithms were not applied/ considered appropriate. However, it is important that the ERG report acknowledges that an explanation was provided as to why these algorithms were not considered appropriate by Servier. Without this explanation, it may be inferred that the algorithms were not	We have changed the text to be similar to that suggested by the company

algorithm (see Section 4.2.5.4). Furthermore, the publication by Marriott *et al.* provides important information concerning the external validity of the Kontodimopoulos *et al.* algorithm, wherein it is stated: *"The mapping that provided the closest fit to the observed data (Kontodimopoulos et al.[15]; 0.80 vs 0.79) was conducted in gastric cancer, which we would expect to have symptoms more similar to liver-only or liver-dominant mCRC than other conditions used in the mappings."*, supporting the use of this algorithm as a sensitivity analysis.	Should the ERG wish to do so, a critique of the analysis provided using the Marriott *et al.* algorithm may also be helpful to include within the ERG report for completeness.	provided without justification. Amending this error will have no impact on the cost-effectiveness results, but will improve the transparency of the ERG report.	
On page 74, the ERG states: *"Both the NICE DSU TSD 10[37] and Woodcock and Doble[34] would lead the ERG to question whether the mapping algorithm from Kontodimopoulos et al.[26] which did not include patients with metastatic cancer would be appropriate in a population in which "all patients had heavily pre-treated (i.e. two or more previous lines of therapy) metastatic gastric cancer" and whereby the estimated life expectancy under current standard care was in the region of six months."* For the Versteegh *et al.* publication, disease stage for the multiple myeloma cohort was unclear, and for the non-Hodgkin's lymphoma cohort patients had Ann Arbor stage II to IV, or intermediate or high-grade malignancy. For the Longworth *et al.* population, disease stage is discussed, but no clear figures are presented. However, the Marriott *et al.* algorithm considers a population of patients with previously untreated metastatic colorectal cancer. Servier acknowledges that the Kontodimopoulos *et al.* algorithm considers a non-metastatic population; however, it remains unclear how many patients had	Servier requests that the ERG amends discussion concerning the limitations of the base-case analysis mapping algorithm by noting that the alternative mapping algorithm by Marriott *et al.* addresses one of the limitations highlighted (concerning metastatic disease). Please see proposed amendments below: Page 74: *"Both the NICE DSU TSD 10[37] and Woodcock and Doble[34] would lead the ERG to question whether the mapping algorithm from Kontodimopoulos et al.[26] which did not include patients with metastatic cancer would be appropriate in a population in which "all patients had heavily pre-treated (i.e. two or more previous lines of therapy) metastatic gastric cancer" and whereby the estimated life expectancy under current standard care was in the*	Servier understands that a limitation of the Kontodimopoulos *et al.* algorithm is that it does not consider a metastatic population. However, this same criticism may also apply to the alternative algorithms presented (though this is unclear); and the criticism does not apply to the Marriott *et al.* algorithm which is not discussed within the ERG report. By amending the ERG's report to highlight that an analysis utilising a mapping algorithm wherein the population had metastatic disease, there is improved clarity concerning (a) the analyses provided, and (b) the differences in the	We have changed the text to be similar to that suggested by the company

metastatic disease in the other mapping studies proposed by the ERG (Versteegh *et al.* and Longworth *et al.*). The mapping algorithm by Marriott *et al.* was developed in a purely metastatic, gastrointestinal cancer population, yet this is not highlighted within the ERG's report. Exclusion of this important feature of the identified mapping studies should be noted within the ERG's report, as this omission implies (a) that the two studies highlighted by the ERG consider a metastatic cohort (which is unknown), and (b) that the Marriott algorithm does not consider a metastatic cohort (which is incorrect).	*region of six months.* ***The ERG notes however that the alternative mapping algorithm used in response to a clarification question by Marriott et al. considers a metastatic colorectal cancer population****."*	populations considered. Amending this error will have no impact on the cost-effectiveness results, but will improve the transparency of the analyses provided to the ERG, along with highlighting important differences in the studies considered.	

Issue 3 Survival outcomes

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
On page 51, the ERG states: *"Other outcomes indicate that the key efficacy gain is in OS, as there is only a 0.2 month absolute difference in median PFS and no clinically relevant differences in HRQoL."* While the difference in median PFS is relatively small, there is a clear PFS advantage associated with TFT versus placebo, as indicated via the Kaplan-Meier curves from the TAGS trial. For example, 4-month PFS is 27% versus 8%, and 6-month PFS is 15% versus 6%. Extensive interpretation of median PFS within the context of a trial affected by protocol-driven progression assessments is inappropriate (as this does not reflect the true "average" outcome for patients). In addition, the statement of clinically relevant differences in HRQoL relates to the difference in HRQoL from baseline, which is indicative of HRQoL stabilisation with TFT. This is discussed further in response to Issue 1. The comments raised by the ERG should ideally be considered within the context of this rapidly progressing, aggressive cancer – i.e., stabilisation in HRQoL and improvements in PFS outside of the median statistic are important considerations. The current text within the ERG's report is potentially misleading, and does not fairly represent the data from the TAGS trial.	Servier requests the following revisions be made to this statement for cohesion with the evidence presented within the CS: *"Other outcomes indicate that the key efficacy gain is in OS. **There is a 0.2 month absolute difference in median PFS, though a clear improvement in PFS after approximately 2 months (e.g. 4-month PFS was 27% versus 6%). In addition, there were no clinically relevant differences in HRQoL for TFT versus baseline (indicative of stabilisation)."***	This minor amendment to the text in the ERG's report provides a more accurate reflection of the evidence base presented within the CS. Amending the report will have no impact on the cost-effectiveness results, but will ensure the ERG's report reports these clinical outcomes appropriately – that is, without focusing on a single metric which is not representative of the totality of the evidence base, and acknowledges that the lack of a clinically relevant difference in HRQoL applies to comparisons made versus baseline measures.	We have changed the text to be similar to that suggested by the company
On page 65, the ERG states: *"The ERG notes that the rates of*	Servier requests this statement be	The ERG notes a	The ERG's

AEs were generally lower in Asian patients[14] and therefore that the average costs would increase if a ROW population was used. However, the ERG anticipated that the impact was low and has ignored this." Servier notes that the number of Japanese patients is low and specific events are relatively rare, and so there is a limited basis from which to make conclusions concerning the differences in AE rates across groups. However, the statement raised by the ERG is false – in the CSR (Section 12.5.1.1), it is shown that 80.4% of Japanese patients receiving TFT experienced at least one Grade 3 or higher AE, versus 79.6% of the ROW subgroup (CSR Table 48). In addition, the aforementioned table demonstrates that 89.1% Japanese patients experienced at least one treatment-related AE, versus 79.6% of the ROW subgroup. As such, the opposite of the statement made by the ERG may indeed be the case for the TAGS trial, though Servier notes that such differences are highly uncertain and the difference in AE rates by region is not always in favour of the ROW group (e.g. there are more serious AEs in the ROW subgroup versus the Japanese subgroup).	removed from the ERG report – differences in AE rates by region are highly uncertain, and so it would be inappropriate to suggest the impact would be in any particular direction (i.e. either in favour of or against TFT).	potential difference in AE rates by region, yet this is not substantiated with conclusive evidence from the TAGS trial. By removing this statement, the ERG report presents a more accurate reflection of the evidence base (i.e. that any differences by region are unclear). Cost-effectiveness results are unaffected by this omission.	statement was based on the following excerpt from the CSR, which is an analysis by race (section 12.5.2.2: "For the TAS-102 group, the incidences of nearly all adverse events (preferred terms) were reported at lower incidences in Asian patients than for other racial groups with few exceptions (Table 14.3.1.2.9)." Servier are referring to an analysis by region (section 12.5.1.1). As the data are uncertain the ERG have removed the sentence.
On page 73, the ERG states: *"For extrapolating OS, the combined modelling approach provided lower AIC/BIC scores*	To align with the approaches taken to select the most appropriate base-case	Servier acknowledges the ERG's opinion	We have changed the text to be

when compared with the independent modelling approach. However, the difference in scores were less than 3 points, hence it indicated that both models provided similar statistical goodness of fit to the data. By examining the plots for assessing the appropriateness of the combined modelling approach (with treatment as a covariate), the ERG believes that it was not clear that the combined modelling approach would be more appropriate for the OS. If the OS data were associated with a constant AF over time, the fitted survival curves would be the same using either the combined modelling or independent modelling approach. The ERG notes that when using independent lognormal models increased ICER to above £50,000." The ERG is correct to note that the difference in AIC/BIC scores is relatively small, indicating similar levels of statistical goodness-of-fit. However, further inspection of the AIC values leads to the following: • AIC, dependent log-normal model: 3,219.65 • AIC, independent log-normal model: 3,221.38 The quantity $\exp(AIC_{min} - AIC_i)/2$ is proportional to the probability that the i^{th} model minimises the estimated information loss. In this case, the independent model is 0.42 times as probable as the first model to minimise the information loss (i.e. less than half as probable). However, this was not the sole argument provided within the CS as to why these models were chosen. Firstly, the dependent log-normal model provides a good visual fit to the Kaplan-Meier curves for both treatment arms. In addition, the long-term extrapolations were aligned with clinical expectation. Furthermore, the AFT assumption was tested via inspection of a quantile-quantile plot which (given the illustrated linear pattern over time) indicated a constant treatment effect over time. When considering the range of survival models fitted, it was determined that the simplest model that provided a reasonable fit	model fits, Servier requests the ERG to amend its description in line with the following: *"For extrapolating OS, the combined modelling approach provided lower AIC/BIC scores when compared with the independent modelling approach. However, the difference in scores were less than 3 points, hence it indicated that both models provided similar statistical goodness of fit to the data.* ***The ERG notes that the company considered a number of other approaches in selecting its preferred base-case model, including the use of a quantile-quantile plot, visual inspection and assessing the plausibility of longer-term projections.*** *By examining the plots for assessing the appropriateness of the combined modelling approach (with treatment as a covariate), the ERG believes that it was not clear that the combined modelling approach would be more appropriate for the OS. If the OS data were associated with a constant AF over time, the fitted survival curves would* ***theoretically*** *be the same using either the combined modelling or independent modelling approach* ***(though this would be difficult to establish using "real" trial data, owing to limited sample sizes****). The ERG notes that when using independent lognormal models increased ICER to*	concerning appropriate model selection, yet notes that the current text within the report implies that only statistical goodness-of-fit was used when Servier determined the base-case analysis settings. In addition, the ERG's comment concerning evidence for a constant AF is based in theory, which in practice would be very difficult to conclusively demonstrate (i.e. the AF is an estimated parameter, which is affected by the estimation of all other model parameters). The proposed amendment aims to address each of these points, such that the report appropriately reflects the CS and the issues associated with extrapolating from "real" clinical trial data. No changes to the cost-effectiveness results are introduced as a	similar to that suggested by the company

to the data should be preferred over an alternative. The ERG's comment that should the AF be constant over time then the independent models would be identical to the dependent models is theoretically true, though fails to acknowledge several practical issues when fitting survival models to trial data. Most relevant to the TAGS trial is that patients were randomised 2:1 and so the base curve parameters (e.g. shape and scale) for the independent model fits to the placebo arm are based upon a third of the sample size of the entire population – as such, these parameters are fitted to a substantially smaller sample size, and as such are more likely to be influenced by individual observations. Servier appreciates that the ERG is entitled to its opinion concerning the choice of models. However, Servier equally considers it important that its justification of the preferred survival model is not implied to be based on a comparison of statistical goodness-of-fit scores alone – the current text within the ERG's report is a misrepresentation of the steps taken to select the preferred base-case model fits. In addition, the theoretical argument concerning the difference between dependent and independent models should be made in light of the data prevalent to this appraisal, as in particular the randomisation of the TAGS trial influences the practical ability to establish if the models would be identical.	*above £50,000."*	consequence of this amendment.	
[redacted]	[redacted]	[redacted]	We have changed the text to be similar to that suggested by the company

[redacted]	[redacted]	[redacted]	
On page 84, the ERG states: *"The ERG requested an analysis of patients from Europe without prior ramucirumab treatment, for which a HR was not reported, [redacted]."* Servier acknowledges that a HR was not provided for this subgroup, yet considers it inappropriate to speculate what the HR would be.	Servier requests that the text concerning the HR for this subgroup analysis be removed, given that this is unknown (and that speculation around this estimate is potentially misleading). The revised text would read as follows: *"The ERG requested an analysis of patients from Europe without prior ramucirumab treatment, for which a HR was not reported."*	Servier appreciates that the ERG would have liked to see the HR for this subgroup analysis, yet this does not warrant speculation of the output of the analysis. The estimation of HRs requires firstly the PH assumption to be tested, and then the value calculated using statistical software. Servier considers that the ERG report should not comment on this value given that it has not been estimated. This revision has no impact on cost-effectiveness results, and serves to ensure the ERG report remains evidence-based.	[redacted]

Issue 4 Other clinical outcomes (HRQoL, response)

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
On page 50, the ERG states: *"HRQoL appeared largely unaffected by TFT treatment."* In addition, on pages 8 and 84 it is stated that: *"Other outcomes (response rate, duration of response, health-related quality of life) reported small or no benefits."* A conference presentation by Alsina *et al.* concluded that: *"HRQoL remained stable for most functional and symptom scales in both arms, suggesting that HRQoL is largely maintained on treatment with* [TFT]*"* and *"There was a positive trend toward a lower risk of QoL deterioration with* [TFT] *versus placebo across most of the scores"* Furthermore, as presented within the CS, the lack of a clinically relevant change versus baseline is indicative of HRQoL maintenance with TFT, which is an important outcome for patients. In addition to the extension in survival offered by TFT versus placebo, the maintenance of HRQoL is noted to be an advantage of TFT versus other treatments that have been studied in the third-line and beyond setting. In regards to response, the ERG is correct to highlight that the response rate (measured as patient achieving complete or partial response [CR/PR]) is small, yet this is to be expected at this stage of disease. However, the disease control rate (DCR, defined as CR/PR or stable disease [SD]) was 44.1% of patients in the TFT group, compared to 14.5% of patients in the placebo group (p<0.0001), therefore the TFT group demonstrated a significant three-fold increase in the proportion of patients	To align with published analyses of the HRQoL data from the TAGS trial, Servier asks that the text in the ERG report be revised in accordance with the following: *"HRQoL appeared largely* ***maintained with*** *TFT treatment."* For the text on pages 8 and 84, Servier requests this be revised in line with the following: *"****Response rates were limited, though this is to be expected given the stage of disease. However, the disease control rate was markedly improved for TFT versus placebo (44.1% versus 14.5%, p<0.0001). Health-related quality of life was shown to be largely maintained with TFT treatment.****"*	These amendments to the text in the ERG's report are intended to clarify that the goal of treatment at this stage of disease is to maintain HRQoL and extend survival. As such, the fact that TFT has been shown to largely maintain HRQoL, and be associated with a significant improvement in DCR, is important to acknowledge within the ERG report. Amending the report will have no impact on the cost-effectiveness results, but will ensure the ERG's report discusses the impact of TFT on HRQoL within the context of third-line and beyond mGC.	The text on p50 has been changed as requested. The text on p8 and 84 have been changed to be similar to that requested.

with tumour shrinkage or SD when directly compared to placebo. This is not reflected in the “other outcomes” stated by the ERG. Servier considers the statements raised by the ERG to be inaccurate summaries of the available TAGS data, and potentially implies that TFT has zero (or potentially negative) impact on HRQoL and DCR.			

Issue 5 End-of-life criteria

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
On page 82, the ERG states: *"Whether TFT meets the second criterion is less straightforward as the life extension associated with TFT estimated in the model is 0.226 years (2.7 months) which is below the 3 months normally required. The company cite precedent in two prior NICE appraisals to support their case, although only one appears to be directly relevant by having a survival extension of less than 3 months."* Servier notes that of the two examples presented, only one (TA476, nab-paclitaxel for [untreated] pancreatic cancer) included a committee-estimated base-case life extension of less than three months. However, Servier contends that the other identified example (TFT for metastatic colorectal cancer) is equivalently "relevant", given that the estimated survival benefit was close to 3 months (3.2 months) and was based on a relatively better prognosis (i.e. mean survival with placebo was approximately 7.9 months versus 6.0 months based on the base-case analysis presented for the current	Servier requests that the text be revised to note that while the other cited case study was associated with a committee-accepted life extension of more than 3 months, it remains a relevant case study nevertheless: *"Whether TFT meets the second criterion is less straightforward as the life extension associated with TFT estimated in the model is 0.226 years (2.7 months) which is below the 3 months normally required. The company cite precedent in two prior NICE appraisals to support their case, although only one **had** a survival extension of less than 3 months."*	This minor amend is intended to clarify that of the stated case studies, only one included a life extension of less than 3 months. However, the other case study is still relevant for consideration, given that the survival extension was marginally greater than 3 months, and the evidence base is considerably similar to the current appraisal (which is to be expected given that the other example is the previous appraisal of TFT in metastatic colorectal cancer). No changes are introduced to the cost-effectiveness results as a consequence of this amendment.	The text has been changed as requested.

appraisal).			

Issue 6 Minor typographical and/or grammatical errors

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
On page 25, the ERG states: *"This was though unlikely to affect survival."* This appears to be a minor typographical error (and should read "thought" instead of "though")	Servier suggests the text be revised to: *"This was **thought** unlikely to affect survival."*	Minor amend to typographical error for clarity.	The text has been changed as proposed.
On page 28, the ERG states: *"Tumour-response (TR) evaluable population – an ITT analysis only including patients with measureable lesions"* Also, on page 44, the ERG states: *"Response rate outcomes only included patients with measureable disease and ≥1 post-baseline assessment (the TR population; 287/337 (85%) patients receiving TFT, and 156/170 (92%) patients receiving placebo)."* Finally, on page 51, the ERG states: *"Because there was a statistically significant difference in efficacy for patients with measureable disease compared with those without measureable disease, and this analysis only included those with measureable disease, the data may not be generalisable to the whole population."* There is a minor typographical error ("measureable" should be "measurable")	Servier suggests the text be revised to: *"Tumour-response (TR) evaluable population – an ITT analysis only including patients with **measurable** lesions"* *"Response rate outcomes only included patients with **measurable** disease and ≥1 post-baseline assessment (the TR population; 287/337 (85%) patients receiving TFT, and 156/170 (92%) patients receiving placebo)."* *"Because there was a statistically significant difference in efficacy for patients with **measurable** disease compared with those without **measurable** disease, and this analysis only included those with **measurable** disease, the data may not be generalisable to the whole population."*	Minor amend to typographical error for clarity.	The text has been changed as proposed.
On page 42-43, the ERG states: *"Median OS was 2.0 months in the TFT arm and 1.8 months in the placebo arm; the difference in median survival was 0.2*	Servier suggests the text be revised to: *"Median **PFS** was 2.0 months in the TFT*	Minor amend to typographical error for clarity.	The text has been changed as proposed.

months." This should consider the outcome of progression-free survival (PFS).	*arm and 1.8 months in the placebo arm; the difference in median **PFS** was 0.2 months."*		
On page 46, the ERG states: *"The overall incidence of AE evens was 97.3% for the TFT group and 93.5% for the placebo treatment group."* There is a minor typographical error ("evens" should be "events")	Servier suggests the text be revised to: *"The overall incidence of AE **events** was 97.3% for the TFT group and 93.5% for the placebo treatment group."*	Minor amend to typographical error for clarity.	The text has been changed as proposed.
On page 52, the ERG states: *"While SIGN filters are not formally validated, the ERG recognises that they expert-designed and likely to retrieve most of the studies eligible for inclusion."* There appears to be a missing word ("are")	Servier suggests the text be revised to: *"While SIGN filters are not formally validated, the ERG recognises that they **are** expert-designed and likely to retrieve most of the studies eligible for inclusion."*	Minor amend to typographical error for clarity.	The text has been changed as proposed.
On page 64, the ERG states: *"In its base case, the company assumed that MRU costs for patients who were not on treatment independent of health state were: progression-free or post-progression."* The phrasing within this sentence may cause confusion, and so Servier proposes alternative phrasing for the avoidance of doubt.	Servier suggests the text be revised to: *"In its base case, the company assumed that **routine** MRU costs for patients **were based on treatment status (i.e. receiving TFT or not receiving TFT), as opposed to progression status**."*	Minor amend to grammatical error for clarity. Addition of "routine" added to clarify that not all medical costs are purely based on treatment status.	The text has been changed as proposed.

NICE National Institute for Health and Care Excellence

Technical engagement response form

Trifluridine–tipiracil for treating metastatic gastric or gastro-oesophageal junction cancer after 2 or more therapies [ID1507]

As a stakeholder you have been invited to comment on the technical report for this appraisal. The technical report and stakeholders responses are used by the appraisal committee to help it make decisions at the appraisal committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

We need your comments and feedback on the questions below. You do not have to answer every question. The text boxes will expand as you type. Please read the notes about completing this form. We cannot accept forms that are not filled in correctly. Your comments will be summarised and used by the technical team to amend or update the scientific judgement and rationale in the technical report.

Deadline for comments **Monday 4 November 2019**

Thank you for your time.

Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Notes on completing this form

- Please see the technical report which summarises the background and submitted evidence. This will provide context and describe the questions below in greater detail.
- Please do not embed documents (such as PDFs or tables) because this may lead to the information being mislaid or make the response unreadable. Please type information directly into the form.
- Do not include medical information about yourself or another person that could identify you or the other person.
- Do not use abbreviations.
- Do not include attachments such as journal articles, letters or leaflets. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.
- If you provide journal articles to support your comments, you must have copyright clearance for these articles.
- Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.

- Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, and all information submitted under 'depersonalised data' in pink. If confidential information is submitted, please also send a second version of your comments with that information replaced with the following text: 'academic/commercial in confidence information removed'. See the Guide to the processes of technology appraisal (sections 3.1.23 to 3.1.29) for more information.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

About you

Your name	[redacted]
Organisation name – stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder please leave blank)	Servier Laboratories Ltd
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	None

Questions for engagement

Issue 1: Comparator	
1. What treatment is currently used as third-line treatment for metastatic gastric cancer? a. Is chemotherapy used? If so, please specify commonly used regimens and approximately what proportion of people at third line will be offered this treatment?	In NHS practice, the majority of patients would receive best supportive care (BSC) in the third-line setting. Servier highlights that the use of the term *"third-line"* to describe the use of BSC in current NHS practice may cause some confusion (i.e. BSC is not an active pharmacological intervention). For the purpose of this response, third-line refers to patients receiving a pharmacological intervention after two prior lines. The use of chemotherapy, as highlighted by clinical opinion provided to Servier, the ERG, and NICE; is considered for a *"very small proportion of people"*. Clinical opinion provided to Servier noted that the use of third-line chemotherapy is usually restricted to a clinical trial setting, due to the lack of recommended options that are available and the lack of evidence produced within a randomised controlled trial. At an advisory board held by Servier in March 2019, some clinicians said that they ***"****could not remember"* the last time they actively treated a patient in the third-line setting. In addition, the experts noted that there is a difference between the proportion of patients for whom third-line treatment may be considered, versus the proportion of patients who go on to receive a third-line treatment. Accounting for variability in practice across the UK, Servier expects the true proportion of patients who currently receive third-line chemotherapy within routine NHS practice to be *"fractions of a percent"*. There are no specific chemotherapy treatments recommended for use in the third-line and beyond (3L+) settings in published guidelines (ESMO and NICE pathways). ESMO guidelines (published in 2016) state: *"Treatment options may be used sequentially in second and third line, but there is no clear evidence for a benefit beyond second line treatment"*. Where chemotherapy is used in NHS practice, it is expected that this is either based on enrolment within a clinical trial, or use of a regimen not previously administered to patients (i.e. either irinotecan or a taxane, such as docetaxel or paclitaxel).

2. What treatments are currently used as part of best supportive care?	There are no specific treatment(s) used for all patients as part of BSC, as patients are discharged to the community setting and would be managed palliatively, with any treatments required tailored to individual patient needs. However, BSC is expected to encompass a range of interventions that may be used to manage cancer patients at the end of life (and are captured within the submitted economic model as 'end-of-life' care costs). These are expected to include analgesics, antiemetics, haematological support, palliative radiotherapy for symptoms, nutritional support, distress management, and admission to hospice care.
Issue 2a: Generalisability of the TAGS trial: geographical region and prior ramucirumab	
3. Can the full trial population in TAGS (includes 15% of people from Japan, 5% from USA and 80% from Europe) be generalised to the population expected to receive trifluridine–tipiracil in the NHS in England? a. If not, how would you expect any differences to affect trial outcomes?	In general, the characteristics of the TAGS trial population are considered reasonably similar to those who would be expected to be treated with trifluridine-tipiracil in NHS practice. However, there are some important differences expected between the trial and NHS patient populations which are important to consider when determining the population from the TAGS trial which is of the most relevance for decision making. The full trial population included a third of patients who received prior ramucirumab, not routinely recommended for use in UK NHS practice. The prior ramucirumab subgroup also differed from the population expected to receive trifluridine–tipiracil in the NHS in England in other ways, notably in their number of prior treatment lines. Patients in the TAGS trial were stratified according to prior use of ramucirumab. Consequently, Servier believes that the 'no prior ramucirumab' subgroup is the most suitable population for decision making. This is based on two key reasons: **<u>Patients with prior exposure to ramucirumab have received more prior lines</u>** Patients that are less heavily pre-treated are expected to have better outcomes. Improved prognosis is expected to be associated with an increased capacity to benefit from treatment with trifluridine-tipiracil. In the TAGS trial, patients had received at least two prior lines. Notably however, a large proportion of patients had received three or more (3+) prior lines (62.5% of the ITT population). The use of trifluridine-tipiracil in NHS practice is expected to be predominantly in the third-line setting given the lack of other treatment options, and so the TAGS trial includes a population of patients that has been more heavily treated. As would be expected, the proportion of patients with 3+ prior lines is higher in patients with prior ramucirumab exposure:

	• [redacted] of the 'no prior ramucirumab' population had received 3+ prior lines • [redacted] of the 'prior ramucirumab' population had received 3+ prior lines A study by Davidson et al., (2018) showed that as patients progress through multiple lines of therapy, the median survival of the cohort decreases. At first-line, median survival was approximately 11.48 months, decreasing to approximately 6.02 months at the second-line, and 4.61 months at the third-line. As referenced in Servier's submission (Section B.2.13), the analysis found a significant correlation between the survival rate and the number of treatment lines received (P<0.001). In the TAGS trial, median survival for the ITT population treated with placebo was lower still, at 3.6 months. This is unsurprising, given that the TAGS trial covers a third-line and beyond population. **The majority of patients from Japan have prior exposure to ramucirumab** Epidemiological studies in gastric cancer have found that in general, patients from Asian countries (including Japan) have better outcomes than those from European countries. The proportion of patients from Japan is heavily correlated with prior ramucirumab use, as ramucirumab (in combination with paclitaxel) has emerged as a new standard of care in Japan and was rated as recommendation category 1 in the second-line setting (based on the Japanese Gastric Cancer Association guidelines, 2017). The link between prior ramucirumab use as the number of patients from Japan is evident in the TAGS trial: • 14.4% (n=73 of 507) of the ITT population were from Japan • [redacted] of the 'no prior ramucirumab' population were from Japan • [redacted] of the 'prior ramucirumab' population were from Japan **Expected impact on trial outcomes** It is expected that outcomes for the 'no prior ramucirumab' subgroup are likely an under-estimate of the outcomes that would be expected in a predominantly third-line population of patients in routine NHS practice. This is because in practice, NHS patients would have a better baseline overall prognosis than the whole trial population, because of the difference in lines of therapy. A poorer prognosis is expected to be associated with a decreased capacity to benefit from treatment with trifluridine-tipiracil.

4. Is prior ramucirumab expected to influence the relative effectiveness of trifluridine–tipiracil compared with BSC? If so, how?	It is uncertain whether prior exposure to ramucirumab influences the relative effectiveness of trifluridine-tipiracil versus BSC (either positively, or negatively). There is some (primarily anecdotal) evidence to suggest that previous treatment with ramucirumab *may* result in better preservation of patients reserve to tolerate trifluridine-tipiracil in the third-line setting (i.e. patients may initiate third-line trifluridine-tipiracil in a better state, and so may have better outcomes than those who have not previously received ramucirumab). However, as described above, the population of patients who received ramucirumab before enrolment within the TAGS trial have several characteristics that influence patient prognosis (or lower baseline overall survival) and thus the relative effectiveness of trifluridine-tipiracil – most notably, the number of prior lines. In addition, the TAGS trial was not designed to detect a difference in outcomes between these groups. Servier considers it important to highlight that within this context, the absence of evidence to support a difference in relative effectiveness is not evidence of an absence of effect. However, for the purpose of NICE's decision making, Servier considers it appropriate to assume that exposure to ramucirumab is not linked to the relative effectiveness of trifluridine–tipiracil compared with BSC.
5. Would you expect people selected to have prior ramucirumab to have a different baseline overall survival than people who do not have prior ramucirumab? If so, how?	As highlighted previously, there is no evidence which demonstrates that patients with prior ramucirumab exposure have a different baseline overall survival compared to those who have not previously receive ramucirumab, all other things being equal. However, the subgroup of patients with prior ramucirumab exposure in the TAGS trial have an increased number of prior lines of therapy, which is associated with a different (lower) baseline overall survival and thus their inclusion means that the ITT is not the most appropriate group for decision making.
Issue 2b: Generalisability of the TAGS trial: number of prior therapies and ECOG	
6. Can the full trial population in TAGS (includes 63% of people with 3 or more prior lines of treatment) be generalised to the population expected to receive trifluridine–tipiracil in the NHS in England?	As discussed previously, an increased number of prior lines is associated with poorer baseline survival, and potentially a reduced capacity for trifluridine-tipiracil to provide benefit to patients. Were the TAGS trial conducted in a purely third-line population, outcomes would be expected to improve, though the extent to which outcomes may be improved is unclear (as the TAGS trial was not designed to detect a difference in outcomes according to the number of prior lines of therapy).

a. If not, how would you expect any differences to affect trial outcomes?	
7. The TAGS trial only included people with ECOG performance score of 0 or 1. Is this generalisable to the expected population eligible for third-line treatment in the NHS in England?	This is generalisable to the expected patient population for whom trifluridine-tipiracil may be considered in routine NHS practice.
8. Are there any other clinically relevant subgroups where trifluridine–tipiracil is expected to be more clinically effective and cost effective (for example HER-2 status)?	Any speculation regarding the clinical- and cost-effectiveness of trifluridine-tipiracil in specific subgroups is based on limited evidence. However, based on the forest plot from the TAGS trial, the following clinical features may be expected to improve estimates of clinical-effectiveness: patients with gastric (versus gastroesophageal) cancer, HER-2 positivity, fewer metastatic sites, without peritoneal metastases, with previous gastrectomy and fewer prior lines. However, these findings are purely based on an assessment of the forest plot from the TAGS trial, and should be interpreted with caution. The translation of any changes in clinical effectiveness to the outputs of the economic model are even more uncertain, and therefore Servier does not consider it appropriate to speculate the likely directional effect based on these subgroups, nor on the suitability of restricting access by subgroup on this limited evidence.
Issue 3: Overall survival extrapolation	
9. In current practice, for people with metastatic gastric cancer after 2 prior lines of treatment: a. On average, what is the expected survival time? b. How many people would you expect to be alive at 6 months, 1 & 2 years?	The TAGS population considers patients that were treated at the 3L+ setting, as opposed to those that have just progressed following second-line chemotherapy (i.e. third-line only). However, most patients are expected to have a survival time with current care that is less than one year if considered eligible for third-line chemotherapy. From the TAGS trial, median survival was 3.6 months for the BSC arm (based on the ITT population), and so most patients are expected to have a survival time of less than 6 months. In the TAGS trial, the longest surviving patient on the placebo arm had a survival time of approximately 18 months and was censored at this time. As such, it may be reasonable to expect patients to live beyond 18

c. After how long would you expect no people to remain alive?	months, but this is a small proportion. With trifluridine-tipiracil, there were n=7 patients at risk at 18 months, after which two patients died, but the remaining five patients were censored, with one patient censored after 2 years. While both Servier's and the ERG's preferred curves are similar (as shown by Table 4 in the technical report), Servier's curves are slightly closer to the opinions provided by the experts. For example, one clinical expert predicted that in current practice, approximately 20% to 25% of people would survive to 6 months and this would decline to approximately 10% to 15% at 1 year. The company's curves estimated 32% and 12%, whereas the ERG's preferred curves estimated 33% and 12% for 6 months and 1 year, respectively.
Issue 4: End-of-life	
10. What would you consider to be a clinically meaningful extension to life in the population with metastatic gastric cancer after 2 prior treatments?	Within the context of heavily pre-treated, metastatic, gastric/ gastro-oesophageal junction cancer, an improvement in survival of at least 2 months is considered to be very meaningful (Clinical Expert Statement by Elizabeth Smyth). Assuming a baseline survival of 6 months, an improvement of two months would represent a 33.3% increase, which is extremely important for this patient population, particularly as there are no other recommended treatment options available for those who wish to pursue further treatment. The model base-case estimates a 2.7-month extension in survival, equivalent to a 44.0% improvement on baseline survival.
Issue 5: Utility values	
11. Would you expect health-related quality of life to be lower in people with metastatic disease compared with a population with gastric cancer without metastatic disease?	Owing to the increased burden of disease for patients with metastatic gastric cancer compared with those without metastases, Servier considers it likely that metastatic patients would have poorer health-related quality of life compared with a non-metastatic population. The EORTC-QLQ-C30 data from the TAGS trial were of course collected within a metastatic gastric/ gastro-oesophageal junction cancer population, yet these were mapped to the EQ-5D using a published mapping algorithm (Kontodimopoulos *et al.*) in Servier's preferred base-case analysis. The mapping algorithm used considers a non-metastatic gastric cancer population, yet as referenced in Servier's submission, the mapping includes the key domains expected to be affected by gastric cancer. In a conference abstract by Chau *et al.* concerning an analysis of the RAINBOW and REGARD trials of ramucirumab, the authors highlighted that the EORTC-QLQ-C30 is *"sensitive to clinical outcomes in advanced gastric cancer*

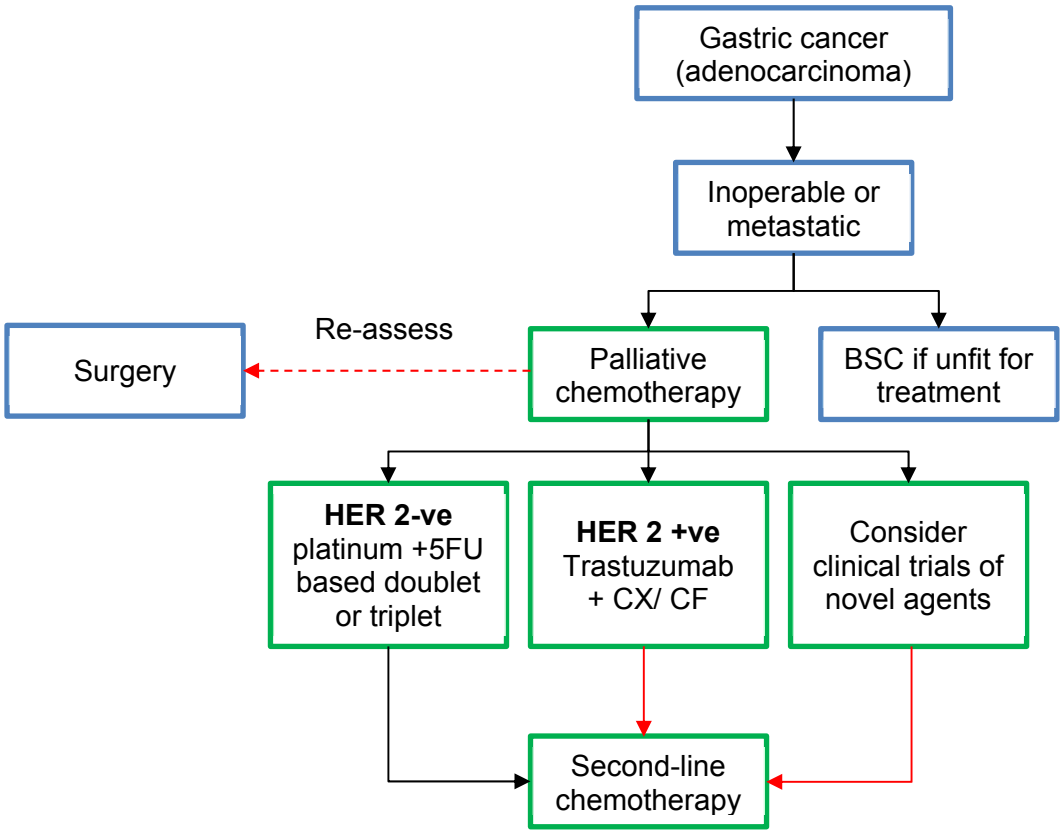
	patients, particularly in global QoL, functional status and disease symptoms of fatigue, pain, and appetite loss." The mapping by Kontodimopoulos includes global health status, as well as physical and emotional functioning (i.e. functional status).
12. Which of the following is a more clinically appropriate method of deriving health-related quality of life values? a. Using EQ-5D values from TA378, a previous appraisal in a similar disease area but with only 1 previous treatment. b. Mapping disease-specific values from the TAGS trial to obtain equivalent EQ-5D values, using an algorithm from a study that did not include people with metastatic disease.	Servier considers the latter of these options (based on mapping values recorded in the TAGS trial) to be the most clinically appropriate methods of deriving utility values. The utility values from TA378 (ramucirumab treating advanced gastric cancer or gastro–oesophageal junction adenocarcinoma previously treated with chemotherapy) have a number of limitations which we are unable to rectify: • Baseline mean EQ-5D-3L index score was applied as the utility value for the pre-progression health state. The utility value for the post-progression health state was estimated using the mean EQ-5D index score at the end of treatment for all patients who discontinued due to progressive disease (measured at the 30-day post-discontinuation visit). Therefore, the utility values are based on empirical mean values taken at a single point in time (and so all other measures of utility were not included). • No specific consideration was taken into account for the correlation between utility scores for the same patient (i.e. a model was not fitted to the utility data, and so only observed values at two fixed time points were considered). This means that it remains unclear how affected the post-progression utility value would be if intra-patient correlation was taken into account, which may radically change the findings of analysis. Servier provided the TA378 utility values as a sensitivity analysis within the submitted economic model for completeness, given that they were used to inform previous decision making. However, these values are associated with limitations which have been highlighted, and therefore the mapped utility values should be considered a more methodologically appropriate analysis to inform decision making.
13. Are there any health-related quality of life benefits that may not be captured in the model?	The impact of metastatic gastric or gastro-oesophageal junction cancer on family and carers was not included within the analysis. In addition, the oral mode of administration is associated with reduced patient burden compared to other treatments administered via intravenous infusion, which patients will have experience of given that trifluridine-tipiracil is administered in a 3L+ setting. For patients wishing to continue treatment in this setting, trifluridine-tipiracil addresses an unmet need for an easily-administered option with proven efficacy and an acceptable safety profile.

Additional information

Clinical pathway of care (ESMO guidelines)

In Servier's submission, a treatment pathway overview diagram for mGC was re-created from the published ESMO guidelines. As highlighted during the technical engagement teleconference on 21 October, 2019, the re-created diagram omitted two arrows concerning second-line chemotherapy. A revised diagram is provided in Figure 1 with this error resolved – the two solid arrows in red have been added, which were originally missing.

Figure 1: Treatment pathway overview for mGC (ESMO guidelines)



Key: BSC, Best supportive care; CF: cisplatin and 5-fluorouracil; CX: cisplatin and capecitabine; ECF: epirubicin, cisplatin and 5-fluorouracil; ECX: epirubicin, cisplatin and capecitabine; EOF: epirubicin, oxaliplatin and 5-fluorouracil; EOX: epirubicin, oxaliplatin and capecitabine; DCF: docetaxel, cisplatin and 5-fluorouracil; ESMO: European Society for Medical Oncology; FOLFIRI: folinic acid, fluorouracil and irinotecan; HER2 +ve/-ve, Human epidermal growth factor receptor 2 negative/positive.

Note: Doublet combinations of platinum and fluoropyrimidines are generally used, but triplet regimen options also include: ECF, ECX, EOF, EOX, DCF or FOLFIRI. Please note this is also the treatment pathway for inoperable advanced disease

Rationale for not providing subgroup analyses for all of the ERG's scenarios

A re-created version of "Table 2" from the technical report is provided in Table 1.

Table 1: ERG's 8 scenario analyses (taken from Technical Report, Table 2)

Prior ramucirumab affects OS	Prior ramucirumab affects relative effectiveness of trifluridine-tipiracil†	
	Yes	**No**
Yes	*Scenario 1: all regions included* Scenario 5: EU subgroup	Scenario 3: all regions included Scenario 7: EU subgroup*
No	Scenario 4: all regions included Scenario 8: EU subgroup*	Scenario 2: all regions included **Scenario 6: EU subgroup***

Notes: *The ERG was not able to calculate ICERS for scenarios 6, 7 and 8 because the data was not available. †The hazard ratio (or acceleration factor) comparing trifluridine–tipiracil + BSC vs. placebo + BSC for overall survival (see issue 3). Scenario 6 in **bold** is the ERG's preferred analysis and scenario 1 in italics is the company's base case.

Servier submitted the analysis referred to as Scenario 1 as its preferred base-case, with a sensitivity analysis concerning Scenario 2 (the ITT population). However, the description of these scenarios as being whether or not prior ramucirumab affecting relative effectiveness or OS is factually inaccurate and misleading. Servier's response within the main form provides further information regarding this, however in summary it is not just exposure to ramucirumab which differs for these subgroups, as there are other factors which are also affected (i.e. number of prior lines).

It is unclear to Servier how the Scenarios referred to as 3, 4, 7, and 8 may be programmed appropriately within standard statistical software. In other words, it is not clear how a relative effect based on a population different to the base curve parameters may be estimated, while also maintaining the correlation between parameters for informing the economic model and its sensitivity analyses. Consequently, these analyses have been attempted by Servier.

Servier provided data for the ERG to consider an analysis of European-only patients with no prior ramucirumab exposure (referred to as Scenario 5), but highlighted extreme caution when interpreting these findings as this subgroup was not pre-specified in the TAGS trial.

This leaves Scenario 6 - European patients regardless of prior ramucirumab exposure. However, as previously highlighted, this population was not pre-specified (as the stratification by region was binary: Japan or Rest of the World), and does not avoid the issue of prior ramucirumab exposure being highly correlated with the number of prior lines of therapy. More specifically, of the European population with prior ramucirumab exposure (n=96), only 24.0% (n=23) had two prior lines, with the remaining 76.0% (n= 73) having three or more prior lines. Therefore, a subgroup analysis of European patients regardless of prior ramucirumab exposure is not provided. However, Servier hopes the explanation provided above clarifies why such an analysis has not been performed.

Trifluridine-tipiracil for treating metastatic gastric or gastro-oesophageal junction cancer after 2 or more therapies. Addendum following Technical Engagement.

Produced by	School of Health and Related Research (ScHARR), The University of Sheffield
Correspondence Author	Matt Stevenson, Professor of Health Technology Assessment, ScHARR, University of Sheffield, Sheffield, UK Andrew Metry, Research Associate, ScHARR, University of Sheffield, Sheffield, UK
Date completed	Matt Stevenson, Professor of Health Technology Assessment, ScHARR, University of Sheffield, Sheffield, UK Date completed (08/11/2019)

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 12/93/72.

1 Background

In June 2019, the company submitted to the National Institute for Health and Care Excellence (NICE) the evidence for use of trifluridine/tipiracil (TFT) (Lonsurf®) in the treatment of metastatic gastric or gastro-oesophageal junction cancer beyond second-line therapy. The Evidence Review Group (ERG) submitted a report in August 2019. The ERG report explored the impacts of alternative assumptions relating to prognoses of patients considered for TFT, whether prior ramucirumab use affects the hazard ratio (HR) for TFT compared with BSC, and the most appropriate geographical region. This culminated in eight different scenarios. These scenarios are shown in Table 1

Table 1: The eight scenarios defined by the ERG

Scenario	Is prognosis from non-ramucirumab patients most appropriate?	Is the HR or AF from non-ramucirumab patients most appropriate?	Is the entire TAGS study population more appropriate than the European geographical area?
1*	✓	✓	✓
2†	✗	✗	✓
3	✓	✗	✓
4	✗	✓	✓
5†	✓	✓	✗
6☯	✗	✗	✗
7	✓	✗	✗
8	✗	✓	✗

*The company base case; †A company scenario analysis; ☯Tentative ERG base case.

The costs per QALY gained from these eight scenarios were reported in Table 15 of the ERG report assuming the use of parametric models for the entire modelling horizon (and are presented again in this addendum in Table 2). This approach was preferred by the ERG as it removed the need for arbitrary assumptions related to: when the parametric curve should replace the Kaplan Meier (KM) data; the duration for which data would be assumed to contribute to the extrapolated portion; and removing any steps in the KM that may be caused by pre-specified time intervals for follow-up.

Table 2: The ERG's base case within the ERG report

Scenario	Independent models			Dependent models		
	Lognormal	Log-logistic	Weibull	Lognormal	Log-logistic	Weibull
1	£51,642	£46,942	£61,310	£45,164*	£42,208	£58,363
2	£55,600	£52,655	£66,137	£50,191	£47,449	£64,318
3				£50,278	£47,750	£65,129
4				£45,076	£41,926	£57,652
	Lognormal	Log-logistic	Generalised gamma	Lognormal	Log-logistic	Generalised gamma
5	£68,061	£59,564	£169,370	£49,067	£45,068	£46,024
6	Not evaluable due to data unavailability					
7						
8						

*Company's base case

Following the Technical Engagement step, the NICE technical team requested cost-effectiveness estimates using the observed, mature TAGS data to model overall survival (OS), applying parametric curves only to extrapolate beyond the data. The company noted that their submitted model had the functionality to calculate these ICERs, but did not provide these results. As such, NICE has requested that the ERG perform further analyses that are reported in this addendum.

During this process, the ERG identified two coding errors with the use of KM data within the company's model to derive ICERs. The first error was the presence of "no value available errors" (#N/A errors) in the time on treatment (ToT) KM data in the excel model provided by the company following the clarification questions. The ERG removed the first error by deleting cells KQ418:KS457 and LX456:LZ457 in the 'ClarQ' sheet. The second error involved using only 524 data points of the ToT KM data in the 'Costs' sheet which resulted in the estimated acquisition costs of TFT being inappropriately reduced. To remove the second error, the ERG made two changes to the company model (i) extended the time horizon of the ToT KM data from R237:AA761 to R237:AA909 in the 'Costs' sheet and (ii) amending the formula in cell Y235 in the 'Costs' sheet to refer to the longer TOT KM time horizon.

2 ERG's exploratory analysis incorporating Kaplan Meier data

The eight analyses presented in Table 2 are reproduced assuming that the KM data are used until ERG defined cut-points. Due to the small number of patients at longer time points the ERG have assumed that 12 months would be appropriate for overall survival within the additional analyses. At this point, there were 31 patients-at-risk in the TFT arm and 10 in the best supportive care (BSC) arm in the full

TAGS study population, and fewer in subgroups. For information, a sensitivity analysis was undertaken increasing the cut-point to 18 months where there were seven patients-at-risk in the TFT arm and zero in the BSC arm in the full TAGS study population. The ERG used a 52-week cut-point for both progression free survival (PFS) and time on treatment (ToT) KM data, which was not altered.

The functionality of the model submitted by the company was such that the hazards assumed after the cut point were equal to the hazards after at the same time point as in the chosen parametric distribution fit to the full modelling horizon. This contrasts with a method that would use a shorter time period to extrapolate hazards from the KM data.

The costs per QALYs gained assuming a cut-point for overall survival of 52 weeks are shown in Table 3, with the results when a cut-point of 78 weeks is used are provided in Table 4. The preference of the ERG between these two scenarios is to use a 52-week cut-point, however, the ERG has not altered its view that fitting a parametric curve through the entire modelling period is a better approach, with these results provided in Table 2.

Comparing Table 2 and Table 3 it is seen that the cost per QALY gained are greater using the KM data in all of the five scenarios. The ERG's tentative base case could not be evaluated and making inferences is more difficult using the KM data due to the small number of patients-at-risk within subgroup analyses.

The results using an 18-month KM cut-off for overall survival (Table 4) increases the cost per QALY gained compared with the 12-month cut-point particularly in Scenario 5 [redacted].

Table 3: The ERG's exploratory analysis using KM data for the first 52 weeks for OS, PFS, and ToT

Scenario	Independent models			Dependent models		
	Lognormal	Log-logistic	Weibull	Lognormal	Log-logistic	Weibull
1	£54,786	£53,322	£65,795	£50,815*	£50,082	£60,676
2	£57,605	£54,740	£71,130	£54,876	£52,069	£67,054
3				£53,437	£50,554	£66,441
4				£51,963	£51,311	£61,139
	Lognormal	Log-logistic	Generalised gamma	Lognormal	Log-logistic	Generalised gamma
5	£81,798	£84,620	£243,801	£66,003	£68,404	£64,039
6	Not evaluable due to data unavailability					
7						
8						

*Company's base case

Table 4: The ERG's exploratory analysis using KM data for the first 78 weeks for OS and first 52 weeks for PFS, and ToT

Scenario	Independent models			Dependent models		
	Lognormal	Log-logistic	Weibull	Lognormal	Log-logistic	Weibull
1	£66,616	£67,941	£71,407	£63,024*	£64,889	£67,815
2	£72,700	£72,876	£78,931	£70,182	£70,251	£76,130
3				£69,480	£69,545	£75,848
4				£63,562	£65,399	£68,051
	Lognormal	Log-logistic	Generalised gamma	Lognormal	Log-logistic	Generalised gamma
5	£111,000	£131,559	£318,139	£93,319	£109,259	£95,940
6	Not evaluable due to data unavailability					
7						
8						

*Company's base case

3 ERG's exploratory analysis incorporating Kaplan Meier data and using a lower utility for the progression free survival and progressed disease health states

The ERG report commented that the mapping algorithm used by the company to transform EORTC-30 values into EQ5D-3L values was derived from a data set where no patient had metastatic cancer. Further, the estimated utilities appeared to lack face validity compared with those used in previous STAs, which had collected EQ-5D-3L within the pivotal studies. As such, NICE requested that the analyses shown in Tables 3 and 4 be rerun, using utility values from TA378 which had a utility of 0.737 for progression free survival and 0.587 for progressed disease which were lower than the 0.764 and the 0.652 respectively that were assumed by the company within its base case. The costs per QALYs gained assuming a cut-point for overall survival of 52 weeks are shown in Table 5, with the results when a cut-point of 78 weeks is used are provided in Table 6. For reference, the impact of using utility values from TA378 when the KM curves were not used increased the company's base case ICER from £45,164 to £47,857 and in Scenario 5, using independent lognormal models, from £68,061 to £70,905 per QALY gained.

In the current analysis, for the majority of results reducing the utility to that used in TA378 increased the ICER by less than £4000. The major exception is when the independent generalised gamma is selected for OS. This is caused by [redacted]. The ERG does not believe this to be likely and cautions against the use of the generalised gamma for OS in Scenario 5.

Table 5: The ERG's exploratory analysis using KM data for the first 52 weeks for OS, PFS, and ToT and using TA378 utility values

Scenario	Independent models			Dependent models		
	Lognormal	Log-logistic	Weibull	Lognormal	Log-logistic	Weibull
1	£57,864	£56,392	£68,950	£53,865*	£53,122	£63,846
2	£61,126	£58,216	£74,741	£58,356	£55,492	£70,670
3				£56,889	£53,941	£70,053
4				£55,025	£54,365	£64,315
	Lognormal	Log-logistic	Generalised gamma	Lognormal	Log-logistic	Generalised gamma
5	£85,034	£87,773	£224,668	£69,485	£71,871	£67,522
6	Not evaluable due to data unavailability					
7						
8						

*Company's base case

Table 6: The ERG's exploratory analysis using KM data for the first 78 weeks for OS and first 52 weeks for PFS, and ToT and using TA378 utility values

Scenario	Independent models			Dependent models		
	Lognormal	Log-logistic	Weibull	Lognormal	Log-logistic	Weibull
1	£69,606	£70,906	£74,453	£66,068*	£67,907	£70,906
2	£76,233	£76,407	£82,436	£73,738	£73,805	£79,672
3				£73,040	£73,103	£79,388
4				£66,599	£68,409	£71,148
	Lognormal	Log-logistic	Generalised gamma	Lognormal	Log-logistic	Generalised gamma
5	£112,778	£131,580	£278,632	£96,131	£111,160	£98,626
6	Not evaluable due to data unavailability					
7						
8						

*Company's base case

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Final technical report

Trifluridine–tipiracil for treating metastatic gastric or gastro-oesophageal junction cancer after 2 or more therapies [ID1507]

This document is the draft technical report for this appraisal. It has been prepared by the technical team with input from the lead team and chair of the appraisal committee.

The technical report and stakeholder's responses to it are used by the appraisal committee to help it make decisions at the appraisal committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the appraisal committee meeting.

The technical report includes:

- topic background based on the company's submission
- a commentary on the evidence received and written statements
- technical judgements on the evidence by the technical team
- reflections on NICE's structured decision-making framework.

This report is based on:

- the evidence and views submitted by the company, consultees and their nominated clinical experts and patient experts and
- the evidence review group (ERG) report.

The technical report should be read with the full supporting documents for this appraisal.

Issue date: December 2019

Contents

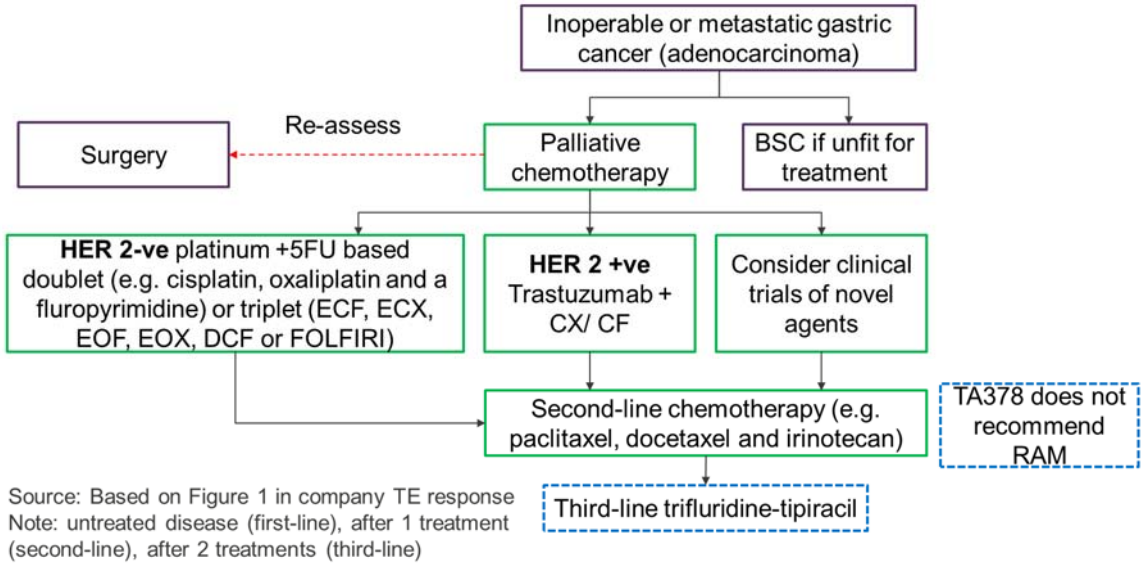
1. Topic background

1.1 Disease background: Gastric or gastro-oesophageal junction cancer

- Stomach cancer is a malignant tumour arising from cells in the stomach. The most common type of stomach cancer is gastric or gastro-oesophageal junction adenocarcinoma, which affects about 95% of people with the disease.
- The company submission and TAGS trial include people with gastric or gastro-oesophageal junction cancer. Unless otherwise specified, gastric cancer is used in this document to refer to both conditions.
- Initial symptoms of gastric cancer are vague and are similar to other stomach conditions, but symptoms of advanced stages may include a lack of appetite and subsequent weight loss, fluid in the abdomen and blood in the stool.
- In England in 2016, around 18% of gastric cancer was diagnosed at stage 3 (locally advanced) and 38% was diagnosed at stage 4 (metastatic).
- There is no standard treatment for treated advanced or metastatic disease. The ESMO clinical practice guideline for gastric cancer recommends:
  - for untreated disease (first line): chemotherapy (such as doublet or triplet platinum or fluoropyrimidine combinations)
  - after 1 or more treatments (second- and subsequent-lines): taxane (docetaxel, paclitaxel) or irinotecan (ramucirumab is not recommended in TA378)
  - trifluridine–tipiracil is recommended as third-line treatment in people with PS 0-1 (Nov 2019 update)
- In clinical practice, paclitaxel is generally used after 1 treatment (second line), and irinotecan is more likely to be used after 2 treatments (third line). Best supportive care is often used after 2 treatments (third line).

- For gastro-oesophageal junction cancer, the ESMO clinical practice guideline for oesophageal cancer notes that treatment largely follows the ESMO recommendations for gastric cancer.

1.2 Treatment pathway for metastatic gastric cancer based on ESMO guideline



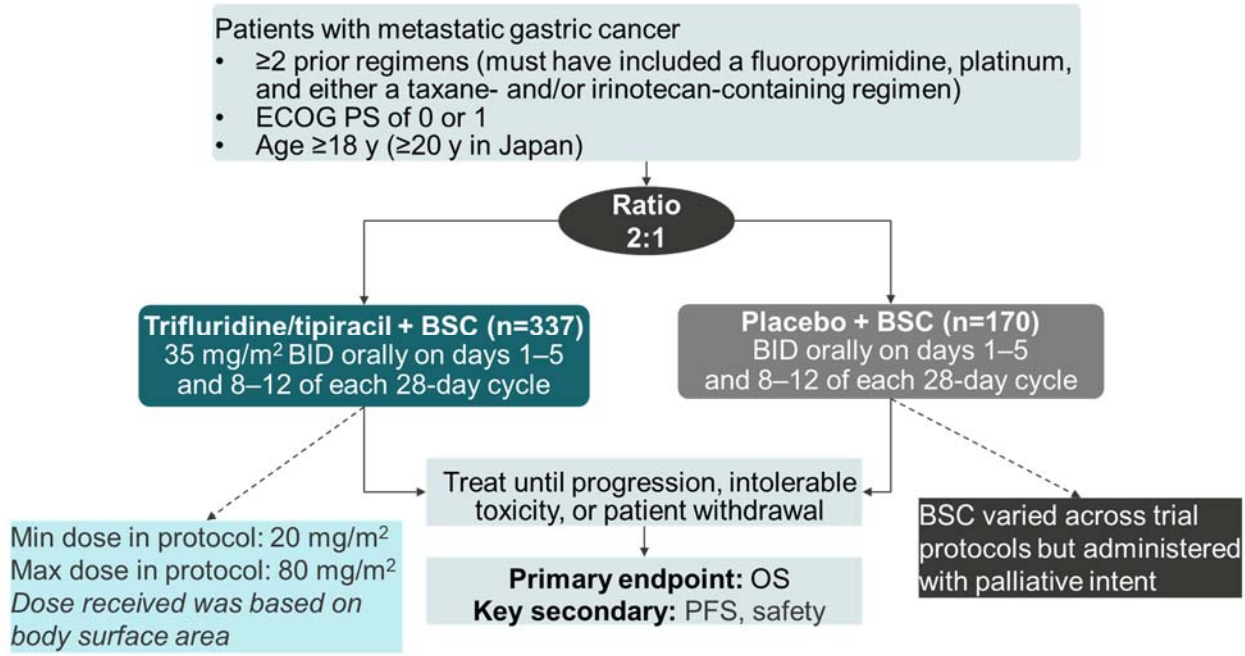
Source: Based on Figure 1 in company TE response
Note: untreated disease (first-line), after 1 treatment (second-line), after 2 treatments (third-line)

Abbreviations: BSC, Best supportive care; CF: cisplatin and 5-fluorouracil; CX: cisplatin and capecitabine; ECF: epirubicin, cisplatin and 5-fluorouracil; ECX: epirubicin, cisplatin and capecitabine; EOF: epirubicin, oxaliplatin and 5-fluorouracil; EOX: epirubicin, oxaliplatin and capecitabine; DCF: docetaxel, cisplatin and 5-fluorouracil; FOLFIRI: folinic acid, fluorouracil and irinotecan; HER2 +ve/-ve, Human epidermal growth factor receptor 2 negative/positive; RAM, ramucirumab

1.3 Trifluridine–tipiracil

Marketing authorisation (received September 2019)	License extension: Trifluridine-tipiracil is indicated as monotherapy for the treatment of adult patients with metastatic gastric cancer including adenocarcinoma of the gastroesophageal junction, who have been previously treated with at least two prior systemic treatment regimens for advanced disease.
Mechanism of action	Antineoplastic thymidine-based nucleoside analogue, trifluridine, and the thymidine phosphorylase (TPase) inhibitor, tipiracil hydrochloride.
Administration	The recommended starting dose of trifluridine-tipiracil in adults is 35 mg/m^2 administered orally twice daily on days 1 to 5 and days 8 to 12 of each 28-day cycle as long as benefit is observed or until unacceptable toxicity occurs. The dosage is calculated according to body surface area (BSA). The dosage must not exceed 80 mg/dose.
Price	List price: 20 x 15 mg tablets: £500; 60 x 15 mg tablets: £1,500; 20 x 20 mg tablets: £666.67; 60 x 20 mg tablets: £2,000 Commercial arrangement (simple discount patient access scheme [PAS]) approved for trifluridine-tipiracil as part of previous appraisal TA405. Average cost per 28-day cycle (excluding the commercial arrangement and using body surface area distribution) is £2,017.

1.4 TAGS trial for people with gastric cancer (including gastro-oesophageal junction)



1.5 Summary of clinical effectiveness results from TAGS

The summary table below presents the company's clinical effectiveness results from the full intention to treat (ITT) population and results for a pre-specified subgroup of people in TAGS who did not have prior ramucirumab. Prior ramucirumab, is not recommended in NICE TA378, but is used in TAGS, and some clinical advice to the company suggests that it's use may affect clinical outcomes.

Outcome	Full ITT population (N=507)†		Subgroup: no prior ramucirumab (N=338)	
	Trifluridine-tipiracil + BSC (n=337)	Placebo + BSC (n=170)	Trifluridine-tipiracil + BSC (n=223)	Placebo + BSC (n=115)
Median overall survival, months	5.7 (4.8 to 6.2)	3.6 (3.1 to 4.1)	6.0 (5.1 to 6.9)	3.3 (2.8 to 3.9)
Overall survival	HR 0.69 (0.56 to 0.85)		HR 0.66 (0.51 to 0.85)	
Median PFS, months	2.0 (1.9 to 2.3)	1.8 (1.7 to 1.9)	NR	NR
PFS	HR 0.57 (0.47 to 0.70)		HR 0.58 (0.46 to 0.75)	
Disease control rate*	44.1%	14.5%	NR	NR
ORR	4.5%	2.1%	NR	NR

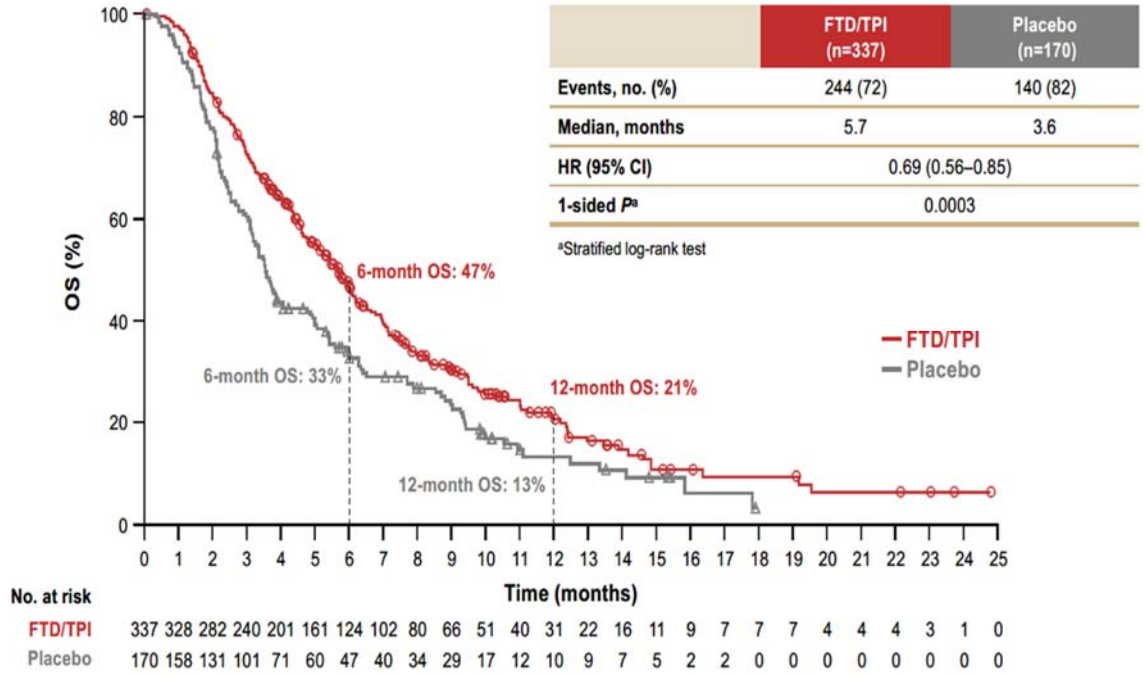
All data include (95% confidence intervals)

*Proportion with complete response, partial response or stable disease

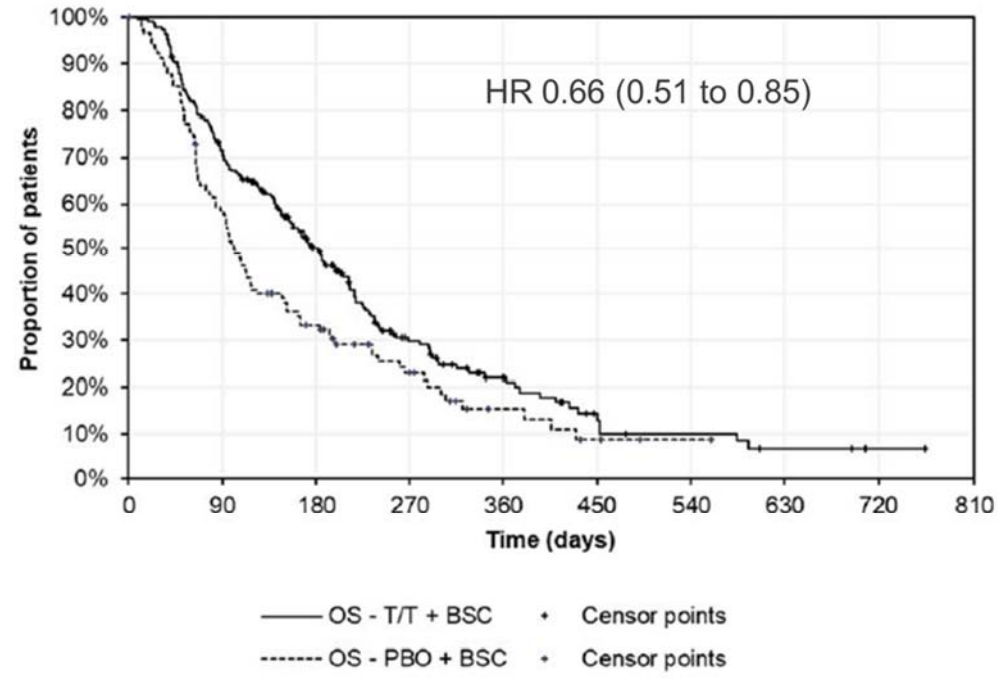
† The intention to treat population included all patients regardless of prior ramucirumab

Abbreviations: BSC, best supportive care; HR, hazard ratio; ITT, intention to treat; NR, not reported; ORR, objective response rate; PFS, progression-free survival

1.6 Overall survival Kaplan-Meier from full ITT population from TAGS



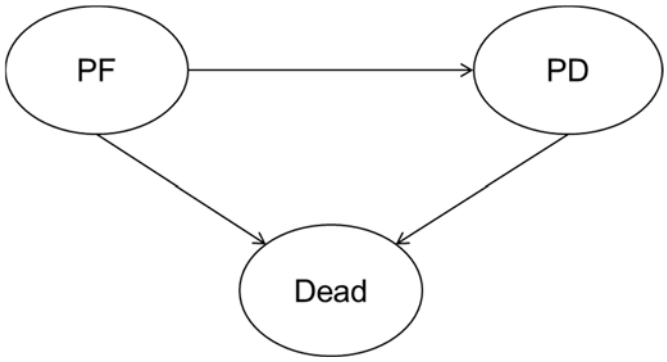
1.7 Overall survival from TAGS subgroup without prior ramucirumab



Key: BSC, best supportive care; OS, overall survival; T/T, trifluridine/tipiracil.
Source: Based on Figure 15 in company submission

1.8 Company's model structure and inputs

- 3-state partitioned survival model with 7-day cycle length
- 10-year time horizon & 3.5% discount rate
- 33% of people in TAGS had previous ramucirumab (not recommended in NICE TA378), therefore the company's base case uses the subgroup without prior ramucirumab from all geographical regions to represent the population in England
- Long-term extrapolations for both treatment arms (see issue 3):
  - Overall survival: lognormal curve with treatment variable (dependent model)
  - Progression-free survival: generalised gamma with separate (independent) curves for each treatment arm
  - Time to treatment discontinuation: generalised gamma with separate (independent) curves for each treatment arm
- Utility values from EORTC-QLQ-C30 collected in TAGS and mapped to EQ-5D-3L
- Adverse event utility decrements from a targeted literature review
- No drug costs for best supportive care or oral administration costs for trifluridine-tipiracil



1.9 Key model assumptions

	Company base case	Technical team
Population	• Data from a subgroup in TAGS (no prior ramucirumab treatment) is used to model clinical outcomes and the treatment effect of trifluridine-tipiracil compared with best supportive care • The model uses TAGS data that includes all geographical regions.	The technical team suggests that ERG scenarios 2 and 6, which both assume that prior ramucirumab does not impact overall survival or the estimated treatment effect, are likely to be the most clinically plausible. The technical team suggest both scenarios are plausible because there is mixed clinical advice, therefore both the full trial population (includes all geographical regions) and the EU subgroup should be considered.
Dose	• Trifluridine-tipiracil is administered orally at a dose of 35 mg/m² of body surface area (BSA) twice daily • Company base case considers all patients from EU with and without prior ramucirumab to inform the distribution of BSA • Dose delays ([redacted]) from TAGS were included in model	The technical team accepts the company's approach
Overall survival	An accelerated failure time model is used to extrapolate overall survival using a log-normal curve (fitted with treatment arm as a covariate)	To extrapolate overall survival the technical team: • prefers independent models (a separate curve is fitted independently to each treatment arm) because it does not assume the treatment effect is constant over time • considers the lognormal and log-logistic curves to be clinically plausible because survival predictions are in line with clinical advice
Utility	Utility values are estimated using a mapping algorithm (Kontodimopoulos et al 2009) to map EORTC QLQ-C30 from TAGS to estimate corresponding EQ-5D-3L values	The technical team considers that the company's utility values may be reasonable, but lower utility values may also be clinically plausible because the company's preferred mapping study didn't include people with metastatic disease
Costs	• No treatment costs included for best supportive care • No costs applied for oral administration of trifluridine-tipiracil • Includes costs of subsequent treatment from TAGS after disease progression	The technical team accepts the company's approach

2. Summary of the draft technical report

After technical engagement the technical team has collated the comments received and, if relevant, updated its judgement and rationale. Judgements that have been updated after engagement are highlighted in bold below.

2.1 In summary, the technical team considered the following:

Issue		Technical team's preliminary judgement
1	**Comparator**	The technical team accepts the company's **suggestion** that best supportive care (BSC) is the main comparator **but notes that a small proportion of people may have third-line chemotherapy alongside BSC in NHS practice, and the cost effectiveness of trifluridine-tipiracil compared with third-line chemotherapy is unknown.**
2a	**Generalisability of TAGS trial (1)**	The technical team **prefers** ERG scenarios 2 and 6 which **use the full trial population (regardless of prior ramucirumab) to estimate overall survival and the relative effectiveness of trifluridine-tipiracil. The technical team recognises that the use of ramucirumab and number of prior treatments in the TAGS trial may differ from clinical practice in the NHS in England but does not expect this to impact clinical outcomes.** Both the full trial population and the EU subgroup should be considered because there is mixed clinical advice on whether results from the full trial population are generalisable to the population in England.
2b	**Generalisability of TAGS trial (2)**	The technical team **recognises that ECOG performance status and the number of prior treatments in the TAGS trial may differ from clinical practice in England but does not expect** this to influence the relative treatment effect.
3	**Overall survival extrapolation**	The technical team prefers models to extrapolate overall survival that are fitted independently to each treatment and considers that the lognormal curve is the most appropriate function, because its survival predictions are consistent with clinical expert advice. However, the technical team recognises that the log-logistic function also provides clinically plausible survival estimates.
4	**End-of-life**	The mean and median extension to life with trifluridine-tipiracil is less than 3 months in the TAGS trial and in the company's model. However, given the poor prognosis, the extension to life with trifluridine-tipiracil may be clinically meaningful.
5	**Utility values**	The company's utility values may be reasonable, **but the company's post-progression values are considerably higher than other published technology appraisals in the same disease area. Therefore,** lower utility values may also be clinically plausible for this population.

2.2 The technical team recognised that the following uncertainties would remain in the analyses and cannot be resolved:

- The inclusion criteria in the TAGS trial were more restrictive compared with the full marketing authorisation for trifluridine-tipiracil because the trial only included people:
  - with ECOG performance score 0 or 1
  - who have had 2 prior regimens that must have included a fluoropyrimidine, platinum, and either a taxane and/or irinotecan-containing regimen

2.3 The cost-effectiveness results include a commercial arrangement (simple discount patient access scheme) for trifluridine-tipiracil.

2.4 Taking these issues into account, the technical team's preferred assumptions result in an incremental cost-effectiveness ratio (ICER) between £52,655 and £58,651 per quality-adjusted life year (QALY) gained for trifluridine-tipiracil compared with BSC. However, the upper limits of this could be higher because the cost-effectiveness estimates are not known for some clinically plausible scenarios (see table 7).

2.5 The company did not make a case for including trifluridine-tipiracil in the cancer drugs fund (CDF).

2.6 Based on the company's economic model, it is uncertain whether trifluridine-tipiracil meets the life extension end-of-life criterion specified in NICE's guide to the methods of technology appraisal (see issue 4).

2.7 Trifluridine-tipiracil is unlikely to be considered innovative. All relevant benefits associated with the drug are adequately captured in the model (see table 9).

2.8 No relevant equality issues were identified.

3. Key issues for consideration

Issue 1 – Comparator

Questions for engagement	1. What treatment is currently used as third-line treatment for metastatic gastric cancer? a. Is chemotherapy used? If so, please specify commonly used regimens and approximately what proportion of people at third line will be offered this treatment? 2. What treatments are currently used as part of best supportive care?
Background/description of issue	The TAGS trial compared trifluridine-tipiracil with placebo and best supportive care (the definition of BSC varied across trial protocols but it was administered with palliative intent), and the inclusion criteria meant only people with ECOG performance scores 0 or 1 were eligible. Approximately 63% of the trial population had 3 or more prior treatments at baseline and approximately 26% had subsequent anti-cancer treatment after disease progression in TAGS. **The company** only includes cost-effectiveness estimates comparing trifluridine-tipiracil with best supportive care because there is a lack of evidence to support the use of chemotherapy in a third-line setting therefore it was not considered to be a relevant comparator. **The ERG** agrees there are no established treatments after 2 prior treatments based on clinical advice. One **clinical expert** suggested that third-line anti-cancer treatment is consistent with the ESMO Guideline for gastric cancer which recommends irinotecan or a taxane (docetaxel, paclitaxel). However, because paclitaxel and ramucirumab are generally used as second-line treatments, irinotecan is the most commonly used third-line treatment. One clinical expert gave an alternative view, advising that chemotherapy is not routinely used in the NHS as a third-line treatment, and that all patients receive BSC with palliative intent (e.g. symptom relief). **The technical team** is concerned that everyone in the comparator arm is assumed to have best supportive care alone, but in clinical practice in England some people may be offered third-line chemotherapy alongside BSC. However, the technical team recognises that clinical expert advice is mixed on this issue.
Why this issue is important	The cost-effectiveness estimates may be biased because it excludes an active treatment comparator (such as irinotecan) which could improve survival and affect treatment costs compared with best supportive care alone.

Technical team preliminary judgement and rationale	The technical team accepts the company's approach that BSC is the main comparator. However, it would be useful to know the clinical and cost-effectiveness of trifluridine-tipiracil compared with a comparator that includes a proportion of people who have chemotherapy because in the NHS in England, some people may be offered third-line chemotherapy alongside BSC.
Summary of comments	**Company**: • The proportion of patients expected to have third-line chemotherapy in current practice is very small. This is supported by clinical opinion that third-line chemotherapy is usually restricted to a clinical trial setting, because there is a lack of treatments available and RCT evidence to support the use of chemotherapy. • BSC is expected to encompass a range of interventions that may be used to manage cancer patients at the end of life (and are captured within the submitted economic model as 'end-of-life' care costs). These are expected to include analgesics, antiemetics, haematological support, palliative radiotherapy for symptoms, nutritional support, distress management, and admission to hospice care.
Technical team judgement after engagement	The technical team accepts the company's suggestion that BSC is the main comparator, but notes that a small proportion of people may have third-line chemotherapy alongside BSC in NHS practice, and the cost effectiveness of trifluridine-tipiracil compared with third-line chemotherapy is unknown.

Issue 2a – Generalisability of the TAGS trial: geographical region and prior ramucirumab

Questions for engagement	3. Can the full trial population in TAGS (includes 15% of people from Japan, 5% from USA and 80% from Europe) be generalised to the population expected to receive trifluridine–tipiracil in the NHS in England? a. If not, how would you expect any differences to affect trial outcomes? 4. Is prior ramucirumab expected to influence the relative effectiveness of trifluridine–tipiracil compared with BSC? If so, how? 5. Would you expect people selected to have prior ramucirumab to have a different baseline overall survival than people who do not have prior ramucirumab? If so, how?
Background/description of issue	The full TAGS trial results may not be generalisable to people who will be eligible for trifluridine–tipiracil in the NHS in England, based on the regions in which the trial was conducted and the use of prior ramucirumab, which is not recommended in TA378. These factors may influence the cost effectiveness of trifluridine-tipiracil

compared with BSC. Table 1 provides a summary of these baseline characteristics from the TAGS trial, the company's base case assumptions, clinical expert advice and the ERG's comments.

Table 1. Overview of generalisability issues

Item	Region	Impact of prior ramucirumab	
		Overall survival (prognosis)	**Relative treatment effect**
TAGS population	The trial included people from Japan (15%), USA (5%) and Europe (80%). Randomisation was stratified to maintain balanced groups by region (Japan vs. rest of world).	33% (n=169) of people in the trial had received prior ramucirumab. Randomisation was stratified by prior ramucirumab (yes/no). Ramucirumab is currently not recommended as a treatment option in the NHS in England for advanced gastric cancer or gastro–oesophageal junction (see NICE TA387).	
TAGS results	OS ITT population (all regions): HR 0.69 (0.56 to 0.85) OS Japan: HR 0.77 (0.46 to 1.30) OS EU: HR 0.67 (0.53 to 0.86)	OS no prior ramucirumab: HR 0.66 (0.51 to 0.85) OS with prior ramucirumab: HR 0.76 (0.53 to 1.09)	
Company	• The company's base case uses the TAGS population from all regions (that includes [redacted] from Japan in the prior ramucirumab subgroup) because it considers the removal of non-European patients to be inconsistent with previous NICE gastric cancer appraisals and because England has an 8% Asian population. • The company also highlights that the EU subgroup without prior ramucirumab is a post-hoc	• The company's base case uses overall survival data and estimates of treatment efficacy from the subgroup of people who have not had prior ramucirumab because it considers this to be more reflective of the UK population who would be eligible for trifluridine-tipiracil. • The company also highlights that at baseline, the subgroup who have not had prior ramucirumab are less heavily pre-treated ([redacted] had 4 or more prior treatments) compared with the subgroup who had prior ramucirumab ([redacted] had 4 or more prior treatments) and this may affect response to treatment.	

		analysis and is not based on stratification in TAGS.		
	Clinical expert	Clinical experts advised that recruitment from Japan was capped at 15%, which is a low proportion, and it is unlikely to affect trial results. Both clinical experts agreed the full TAGS population was likely to be generalisable to the NHS in England.	• Clinical advice to the ERG suggests rapid disease progression after ramucirumab is stopped, suggesting a poor prognosis after treatment • Both clinical experts advised that previous ramucirumab would not affect prognosis and subsequent treatment outcomes such as survival. Furthermore, there is unlikely to be differences in characteristics based on ramucirumab use. • One clinical expert suggested rapid progression after treatment is likely to be due to the stage of disease rather than prior ramucirumab.	Both clinical experts advised that there would not be any expected differences in the relative treatment effect for trifluridine–tipiracil compared with best supportive care based on prior ramucirumab use.
	ERG	Based on clinical advice, the ERG suggests the EU subgroup may be appropriate because disease prognosis and treatment practices are more similar within the EU than in Japan. The ERG notes that in England, people from Japan are likely to make	The ERG prefers to use survival data and estimates of treatment effect from the full TAGS population (regardless of prior ramucirumab), because clinical advice was mixed and there is no strong evidence that prior ramucirumab affects prognosis. However, the ERG notes that the subgroup without prior ramucirumab is less heavily pre-treated and has a	

	up around 1.5% of the population, as the 'Other Asian' group suggests. Therefore, the company's base case (assuming [redacted] from Japan) is likely to be over-estimated for the population in England.	shorter disease duration than the prior ramucirumab group.

The ERG noted that clinical advice is mixed and so it reported cost-effectiveness estimates from 8 scenarios that vary the assumptions around prior ramucirumab and region (see table 2 below).

Table 2. ERG's 8 scenario analyses – sources of data used for OS and relative effectiveness

OS data	Relative effectiveness of trifluridine–tipiracil†	
	No prior RAM subgroup	**Full ITT population**
No prior RAM subgroup	*Scenario 1: all regions included* Scenario 5: EU subgroup	Scenario 3: all regions included Scenario 7: EU subgroup*
Full ITT population	Scenario 4: all regions included Scenario 8: EU subgroup*	Scenario 2: all regions included **Scenario 6: EU subgroup***

*The ERG was not able to calculate ICERS for scenarios 6, 7 and 8 because the data was not available

† the hazard ratio (or acceleration factor) comparing trifluridine–tipiracil + BSC vs. placebo + BSC for overall survival (see issue 3)

Note: Scenario 6 in **bold** is the ERG's preferred analysis and scenario 1 in *italics* is the company's base case.

The technical team is concerned that the company's base case analyses (using the subgroup without prior ramucirumab):

- excludes data from prior ramucirumab patients when it may not be appropriate to do so, and
- includes data with a relatively high proportion of east Asian patients (even if prior ramucirumab patients are excluded).

Why this issue is important	The long-term overall survival predictions and estimates of the relative effectiveness of trifluridine–tipiracil are more favourable in the company's base case (ERG scenario 1). **Company's base case ICER (scenario 1):** £45,164 per QALY gained **Region**: The company's cost-effectiveness estimates increase when using the EU subgroup from TAGS. In scenario 5 (uses the no prior ramucirumab subgroup for overall survival and estimated treatment effect) the ICER compared with BSC increases from £45,164 in scenario 1 to £49,067 per QALY gained. However, there is no data available to estimate cost-effectiveness results in any of the other scenarios that use the EU subgroup (that is, scenarios 6 to 8). **Prior ramucirumab**: Cost-effectiveness estimates using the full TAGS population (regardless of prior ramucirumab) for the estimated treatment effect of trifluridine–tipiracil compared with BSC increases the company's base case ICER (scenario 2: £50,191 and scenario 3: £50,278 per QALY gained)
Technical team preliminary judgement and rationale	The technical team suggest that the ERG scenarios 2 and 6, which assume that prior ramucirumab does not impact overall survival or the estimated treatment effect, are the most clinically plausible.
Summary of comments	**Company**: The full trial population included a third of patients who received prior ramucirumab, not routinely recommended for use in UK NHS practice. The prior ramucirumab subgroup also differed from the population expected to receive trifluridine–tipiracil in the NHS in England in other ways: 1) Patients with prior exposure to ramucirumab have received more prior lines of treatment a. patients that are less heavily pre-treated are expected to have better outcomes b. [redacted] of the 'no prior ramucirumab' population had received 3 or more prior lines [redacted] of the 'prior ramucirumab' population had received 3 or more prior lines. c. Davidson et al., (2018) showed that as patients progress through multiple lines of therapy, the median survival of the cohort decreases 2) The majority of patients from Japan have prior exposure to ramucirumab a. 14.4% (n=73 of 507) of the ITT population were from Japan b. [redacted] (n=[redacted] of 338) of the 'no prior ramucirumab' population were from Japan c. [redacted] (n=[redacted] of 169) of the 'prior ramucirumab' population were from Japan 3) Expected impact on trial outcomes → It is expected that outcomes for the 'no prior ramucirumab' subgroup are likely an under-estimate of the outcomes that would be expected in a predominantly third-

	line population of patients in routine NHS practice. This is because in practice, NHS patients would have a better baseline overall prognosis than the whole trial population, because of the difference in lines of therapy. A poorer prognosis is expected to be associated with a decreased capacity to benefit from treatment with trifluridine-tipiracil. The company agrees that there is a lack of evidence to suggest prior ramucirumab affects overall survival or the relative treatment effect. The company did not provide further evidence for scenarios 3, 4, 7 or 8, because it was unclear how to estimate a relative effect based on a population different to the base curve parameters while also maintaining the correlation between parameters for informing the model. For scenario 6, the company reiterated that patients who had prior ramucirumab were more likely to have had 3 or more treatment lines, and randomisation was not stratified by EU/non-EU status.
Technical team judgement after engagement	The technical team considers this to be a generalisability issue and prefers ERG scenarios 2 and 6 which use the full trial population (regardless of prior ramucirumab) to estimate overall survival and the relative effectiveness of trifluridine-tipiracil. The technical team recognises that the use of ramucirumab and number of prior treatments in the TAGS trial may differ from clinical practice in the NHS in England but does not expect this to impact clinical outcomes. The technical team also recognises that the EU subgroup is not a prespecified subgroup, therefore it is still important to consider the sensitivity of cost-effectiveness estimates to both scenarios 2 and 6.

Issue 2b - Generalisability of the TAGS trial: number of prior therapies and ECOG

Questions for engagement	6. Can the full trial population in TAGS (includes 63% of people with 3 or more prior lines of treatment) be generalised to the population expected to receive trifluridine–tipiracil in the NHS in England? a. If not, how would you expect any differences to affect trial outcomes? 7. The TAGS trial only included people with ECOG performance score of 0 or 1. Is this generalisable to the expected population eligible for third-line treatment in the NHS in England? 8. Are there any other clinically relevant subgroups where trifluridine–tipiracil is expected to be more clinically effective and cost effective (for example HER-2 status)?

<table>
<tr><td>Background/description of issue</td><td>The full TAGS trial results may not be generalisable to people who will be eligible for trifluridine–tipiracil in the NHS in England, in terms of number of prior treatments and ECOG performance status. Table 3 provides a summary of these baseline characteristics from the TAGS trial, the company's base case assumptions, clinical expert advice and the ERG's comments.
Table 3. Overview of generalisability issues
<table>
<tr><th></th><th>Previous therapy</th><th>ECOG performance score</th></tr>
<tr><td>TAGS</td><td>The trial includes around 63% of patients who have had three or more prior regimens of chemotherapy.</td><td>The trial included people with ECOG performance status 0 (38%) or 1 (62%).</td></tr>
<tr><td>Company</td><td>The company highlight that that the subgroup who have not had prior ramucirumab are less heavily pre-treated compared with people who had prior ramucirumab (see issue 2a).</td><td>No specific comments</td></tr>
<tr><td>Clinical expert</td><td>Both clinical experts agreed that the proportion of people having 3 prior treatments would be much lower in England (approximately < 5%).</td><td>No specific comments</td></tr>
<tr><td>ERG</td><td>Based on clinical advice received by the ERG, most people in England do not get a third-line chemotherapy because the burden of treatment outweighs the benefits. The full trial population might have a lower baseline survival prognosis compared with the third-line population in England, because the prior ramucirumab subgroup had more prior lines of treatment than the no prior ramucirumab subgroup.</td><td>Clinical advice received by the ERG indicated that ECOG status, number of metastatic sites, HER-2 status and previous chemotherapy regimens (number and type) are prognostic factors in the third-line setting. The clinical advisors to the ERG noted the slight imbalances in potential prognostic factors, but were not concerned about their potential impact on generalisability to the NHS in England.</td></tr>
</table>
The technical team is concerned that the TAGS trial may not be representative of the population eligible for third-line treatment in the NHS in England.</td></tr>
</table>

Why this issue is important	The TAGS trial included people who were heavily pre-treated and had good ECOG performance status, which may affect the generalisability of its results to the expected population in the NHS in England.
Technical team preliminary judgement and rationale	The technical team notes that the trial population was more heavily pre-treated compared with the population who would be eligible to have trifluridine–tipiracil in the NHS in England, but recognises that this is unlikely to influence the relative treatment effect.
Summary of comments	**Company**: • As discussed previously, an increased number of prior lines is associated with poorer baseline survival, and potentially a reduced capacity for trifluridine-tipiracil to provide benefit to patients. • The ECOG status in the trial is generalisable to the expected patient population for whom trifluridine-tipiracil may be considered in routine NHS practice. • Any speculation regarding the clinical- and cost-effectiveness of trifluridine-tipiracil in specific subgroups is based on limited evidence. Findings based on an assessment of the forest plot from the TAGS trial should be interpreted with caution.
Technical team judgement after engagement	The technical team recognises that ECOG performance status and the number of prior treatments in the TAGS trial may differ from clinical practice in England but does not expect this to impact the relative treatment effect of trifluridine-tipiracil.

Issue 3 – Overall survival extrapolation

Questions for engagement	9. In current practice, for people with metastatic gastric cancer after 2 prior lines of treatment: a. On average, what is the expected survival time? b. How many people would you expect to be alive at 6 months, 1 & 2 years? c. After how long would you expect no people to remain alive?
Background/description of issue	The TAGS trial collected overall survival data over 2 years and although the data is mature, a statistical extrapolation is needed to predict longer-term outcomes up to 10 years after starting treatment. Table 4

shows the proportion of people alive using the 3 survival curves that the ERG believe to be the best fitting and that are independently fitted to each treatment arm.

Table 4. Proportion of people alive at different time points with independent survival models

	6 months		1 year		2 years		5 years		10 years	
Parameter	**TFT**	**BSC**	**TFT**	**BSC**	**TFT**	**BSC**	**TFT**	**BSC**	**TFT**	**BSC**
Full TAGS population										
Company base case	46	32	20	12	6	3	<1	<1	<1	<1
Lognormal	46	33	19	12	5	3	<1	<1	<1	<1
Log-logistic	45	31	19	12	6	4	1	<1	<1	<1
Weibull	50	37	18	10	2	<1	0	0	0	0
Subgroup: no prior ramucirumab										
Company base case	46	32	21	12	6	3	<1	<1	<1	<1
Lognormal	46	32	21	13	6	3	<1	<1	<1	<1
Log-logistic	46	31	20	12	7	4	1	<1	<1	<1
Weibull	50	36	20	11	2	<1	0	0	0	0

Note: company base case uses a dependent model with a lognormal curve but all other data are for an independent model

The company modelled overall survival using an accelerated failure time model that included a dependent variable to capture the effect of treatment with trifluridine-tipiracil. A log-normal model was used because it had the best statistical fit (that is, had the lowest AIC/BIC statistics). The overall survival data in the subgroup of people without prior ramucirumab suggests the treatment effect is constant over time and the long-term predictions were in line with clinical expectations.

The ERG prefers independent models (that is, separate overall survival curves are fitted independently for each treatment arm) because doing so avoids assuming that the relative treatment effect remains constant over time. This is supported by subgroup analyses from the EU subgroup showing that survival curves become extremely close at around 8 months, and a similar convergence at around 15 months in the primary analysis. Furthermore, if the dependent model fitted the data appropriately, then the

	independent model approach would theoretically result in the same survival curves as the dependent approach (though this would be difficult to establish using "real" trial data, due to limited sample sizes). The ERG agrees that the lognormal curve is likely to be the most appropriate based on statistical fit to the data, but recognises that the log-logistic and Weibull curves remain plausible distributions. One **clinical expert** predicted that in current practice, approximately 20% to 25% of people would survive to 6 months and this would decline to approximately 10% to 15% at 1 year. One clinical expert predicted that around 10% would survive at 6 months, around 7% at 1 year and 5% at 2 years. This expert advised that a small proportion of patients does experience very long-term survival. **The technical team** is concerned that the company's overall survival extrapolation may overestimate long-term survival.
Why this issue is important	The overall survival predictions are affected by: • a larger QALY gain for trifluridine–tipiracil in the company's preferred dependently fitted models compared with the ERG's preferred independent models • the curve used to extrapolate overall survival, for example the Weibull curve has a shorter tail compared with the lognormal and log-logistic and predicts no survival after 5 years.
Technical team preliminary judgement and rationale	The technical team: • prefers independent models to extrapolate overall survival because it does not assume the treatment effect is constant over time and • considers that the lognormal curves are the most appropriate because its prediction for BSC is the most consistent with clinical expert advice. However, the technical team recognises that the log-logistic curves may also be clinically plausible. It also notes that it might be useful to see cost-effectiveness estimates using the observed, mature TAGS data to model overall survival, applying parametric curves only to extrapolate beyond the data.
Summary of comments	**Company**: • The TAGS population included people having treatment at third or subsequent line rather than third-line only. However, most patients are expected to have a survival time with current care that is less than one year if considered eligible for third-line chemotherapy. • In the TAGS trial, the longest surviving patient on the placebo arm had a survival time of approximately 18 months and was censored at this time. As such, it may be reasonable to expect

patients to live beyond 18 months, but this is a small proportion. With trifluridine-tipiracil, there were 7 patients at risk at 18 months, after which two patients died, but the remaining five patients were censored, with one patient censored after 2 years.

- While both the company and ERG's preferred OS curves are similar (see Table 4 above), the company's preferred curves are slightly closer to estimates from the clinical experts (approximately 20% to 25% of people would survive to 6 months and this would decline to approximately 10% to 15% at 1 year). The company's curves estimated 32% and 12%, whereas the ERG's preferred curves estimated 33% and 12% for 6 months and 1 year, respectively.

ERG:

Cost-effectiveness estimates for ERG exploratory scenarios using Kaplan-Meier (KM) data for the first 12 months or first 18 months for the technical team's preferred scenario 2 (independent model) are lower compared with the ERG's original and preferred scenario 2 (see table 5). However, the exploratory analyses are limited because:

a. The extrapolated portion of OS in these exploratory scenarios use the same parametric curves as before, which were fitted to the whole dataset. Ideally, it would be preferable to only fit the extrapolation curves to the last portion of the KM data.

b. This analysis required the ERG to select the 'cut points' at which the model stops using the KM data and switches to the parametric curves (this is potentially an arbitrary decision).

Table 5. ERG exploratory scenario analyses using Kaplan-Meier data and parametric curve to extrapolate beyond the data

	Lognormal	Log-logistic	Weibull
ERG original and preferred scenario 2 (parametric curve to model OS for the entire duration of the model)	**£55,600**	**£52,655**	**£66,137**
Kaplan-Meier data for 12 months (OS, PFS and time on treatment)	£41,832	£39,752	£51,610
Kaplan-Meier data for 18 months (OS) and 12 months for PFS and time on treatment	£52,742	£52,867	£57,248

Technical team judgement after engagement	The technical team prefers independent models to extrapolate overall survival using a lognormal curve for the entire duration of the model because its predictions for BSC is most consistent with clinical expert

	advice and it does not assume the treatment effect is constant over time. However, the technical team recognises that the log-logistic function also provides clinically plausible survival estimates. The technical team notes that cost-effectiveness estimates for scenario 2 are more favourable in exploratory analyses that use mature TAGS data to model overall survival and apply parametric curves only to extrapolate beyond the data. The company did not provide these, so the technical team have requested the ERG to provide exploratory scenarios for committee consideration.

Issue 4 – End-of-life

Questions for engagement	10. What would you consider to be a clinically meaningful extension to life in the population with metastatic gastric cancer after 2 prior treatments?
Background/description of issue	The survival benefit associated with trifluridine–tipiracil is less than 3 months in the TAGS trials and the company's model. **The company** suggests that the extension to life criterion (that is, the survival benefit with trifluridine–tipiracil is normally at least 3 months) is met because the company model predicts a median overall survival gain of 2.7 months in the subgroup of people without prior ramucirumab, which is an 82% extension in survival compared with best supportive care. The company states this should be considered in relation to the poor prognosis of this population and highlight that this was also the case in another technology appraisal for metastatic pancreatic cancer (see NICE TA476). Specifically, the company state the survival benefit for trifluridine–tipiracil is superior to that in TA476 (main trial results used in TA476 show a benefit of 2.1 months and the model predicted a mean benefit of 2.4 months) **The ERG** explained that all mean and median survival benefits from the TAGS trial and the company's model were less than 3 months. **The clinical experts** suggest a survival benefit of around 2 months would be considered clinically meaningful for this population, particularly if it can be achieved with a good quality of life. **The technical team** is concerned that an extension to life of less than 3 months may not be clinically meaningful.
Why this issue is important	The company's base case ICER is £45,164 per QALY gained which is above the threshold normally considered a cost-effective use of NHS resource (that is, £20,000 to £30,000 per QALY gained).

Technical team preliminary judgement and rationale	The extension to life with trifluridine-tipiracil is less than 3 months in the TAGS trial and in the company's model. However, given the poor prognosis, the extension to life with trifluridine-tipiracil may be clinically meaningful.
Summary of comments	**Company**: • Within the context of heavily pre-treated, metastatic, gastric or gastro-oesophageal junction cancer, an improvement in survival of at least 2 months is considered to be very meaningful (Clinical Expert Statement). • Assuming a baseline survival of 6 months, an improvement of two months would represent a 33.3% increase, which is extremely important for this patient population, particularly as there are no other recommended treatment options available. The model base-case estimates a 2.7 month extension in survival, equivalent to a 44.0% improvement on baseline survival.
Technical team judgement after engagement	The extension to life with trifluridine-tipiracil is less than 3 months in the TAGS trial and in the company's model. However, given the poor prognosis, the extension to life with trifluridine-tipiracil may be clinically meaningful.

Issue 5 – Utility values

Questions for engagement	11. Would you expect health-related quality of life to be lower in people with metastatic disease compared with a population with gastric cancer without metastatic disease? 12. Which of the following is a more clinically appropriate method of deriving health-related quality of life values? a. Using EQ-5D values from TA378, a previous appraisal in a similar disease area but with only 1 previous treatment. b. Mapping disease-specific values from the TAGS trial to obtain equivalent EQ-5D values, using an algorithm from a study that did not include people with metastatic disease. 13. Are there any health-related quality of life benefits that may not be captured in the model?

Background/description of issue	The TAGS trial collected health-related quality of life data using the EORTC QLQ-C30 at three different time points, with compliance rates varying between 73% and 100%. **The company's** base-case analysis used a mapping algorithm from Kontodimopoulos et al (2009), a Greek study of 48 people with gastric cancer, to map EORTC QLQ-C30 collected in TAGS to estimate corresponding EQ-5D-3L values. The company used this algorithm in its base case because it was the only published algorithm developed in a gastric cancer population according to the latest version of the University of Oxford Health Economics Research (HERC) mapping database. **The ERG** explained that the mapping algorithm used by the company was based on a small population (n=48). The ERG noted that the company's preferred mapping study only included people without metastatic disease, and it is likely that people with metastatic disease would have lower utility values. The ERG also received clinical expert advice that suggests community care in the UK may be higher quality than in many other European countries; therefore, BSC in the UK may be associated with better patient quality of life than elsewhere. At clarification, the ERG requested scenario analyses using 2 alternative mapping algorithms (Versteegh et al 2012 & Longworth et al 2014) based on a recent review (Woodcock et al 2018) that concluded these algorithms were the best performing in external validation studies. The company did not consider these algorithms appropriate to inform the model citing differences in the patient populations used to derive the mapping algorithms, and instead provided an analysis calculating utility values from a mapping algorithm by Marriott et al (2017), which includes a metastatic colorectal cancer population. These values were higher for both progression-free and progressed health states compared with the company's base case. The ERG did an exploratory analysis using the values from the most recent technology appraisal for advanced gastric cancer or gastro-oesophageal junction adenocarcinoma (TA378). Doing so increased the company's base case ICER (see table 6). **The technical team** is concerned that the company's utility values may overestimate quality of life for a population with metastatic disease eligible for third-line treatment.
Why this issue is important	**Table 6. Alternative utility values**

Data source	Population	Progression-free	Progressed disease	Company base case	
				ICER	QALY gain

	Company's base case: mapping algorithm from Kontodimopoulos et al (2009)	Gastric cancer (none had metastatic disease)	0.764	0.652	£45,164	0.153
	TA378*: EQ-5D from RAINBOW trial	Metastatic or non-resectable locally advanced gastric cancer after 1 previous therapy	0.737	0.587	£47,857	0.144
	Company scenario: mapping algorithm from Marriot et al (2017)	Previously untreated metastatic colorectal cancer	0.789	0.720	Not reported	Not reported
	TA208†: EQ-5D from ToGA trial	Previously untreated inoperable locally advanced or recurrent and/or metastatic adenocarcinoma of the stomach or gastro-oesophageal junction	0.729	0.577	£48,473¥	0.142¥
	Note: the company submission also included a scenario analysis using utility values from TA378 but all ICERs reported in this table are taken from the ERG report *In TA378, the committee's preferred progression-free utility value was based on EQ-5D data from the RAINBOW trial (ramucirumab plus paclitaxel vs. placebo plus paclitaxel in 665 adults who had disease progression on or within 4 months after treatment with platinum-containing and fluoropyrimidine-containing chemotherapeutic regimens with or without an anthracycline and ECOG status 0 or 1) from different timepoints during the pre-progression period. †In TA208, the committee's preferred progression-free utility value was based on EQ-5D data from the ToGA trial and included an increase in values over time that was capped so they did not go above those of the general population in the UK of a comparable age.					

	¥ ICER calculated by technical team using the ERG model.
Technical team preliminary judgement and rationale	It is uncertain which utility values are the most appropriate. The company's utility values may be reasonable, but lower utility values may also be clinically plausible for this population.
Summary of comments	**Company**: • The company considers that people with metastatic disease would have a poorer health-related quality of life than people without metastases, because there is an increased burden of disease. • The mapping algorithm used considers a non-metastatic gastric cancer population, but includes the key domains expected to be affected by gastric cancer. In a conference abstract by Chau *et al.* concerning an analysis of the RAINBOW and REGARD trials of ramucirumab, the authors highlighted that the EORTC-QLQ-C30 is *"sensitive to clinical outcomes in advanced gastric cancer patients, particularly in global QoL, functional status and disease symptoms of fatigue, pain, and appetite loss."* The mapping by Kontodimopoulos includes global health status, as well as physical and emotional functioning (i.e. functional status). • The company considers mapped utility values from TAGS to be the most clinically appropriate methods of deriving utility values. The utility values from TA378 (ramucirumab treating advanced gastric cancer or gastro–oesophageal junction adenocarcinoma previously treated with chemotherapy) have a number of limitations which we are unable to rectify: o Baseline mean EQ-5D-3L index score was applied as the utility value for the pre-progression health state. The utility value for the post-progression health state was estimated using the mean EQ-5D index score at the end of treatment for all patients who discontinued due to progressive disease (measured at the 30-day post-discontinuation visit). Therefore, the utility values are based on empirical mean values taken at a single point in time (and so all other measures of utility were not included). o No specific consideration was taken into account for the correlation between utility scores for the same patient (i.e. a model was not fitted to the utility data, and so only observed values at two fixed time points were considered). This means that it remains unclear how affected

	the post-progression utility value would be if intra-patient correlation was taken into account, which may radically change the findings of analysis • The impact of metastatic gastric or gastro-oesophageal junction cancer on family and carers was not included within the analysis. In addition, the oral mode of administration is associated with reduced patient burden compared to other treatments administered via intravenous infusion.
Technical team judgement after engagement	The company's utility values may be reasonable, but the company's post-progression values are considerably higher than those used in other published technology appraisals in the same disease area. Therefore, lower utility values may also be clinically plausible for this population, and sensitivity of cost-effectiveness estimates to alternative values should be considered. The ERG and technical team have reported scenario analyses with utility values from TA378 for committee consideration.

4. Issues for information

Tables 7 to 9 are provided to stakeholders for information only and are not included in the technical report comments table provided.

Table 7: Technical team preferred assumptions and impact on the cost-effectiveness estimate

Alteration	Technical team rationale	ICER	Change from base case
ERG scenario 1: No prior ramucirumab subgroup data used for both OS and treatment effect			
Company base case (scenario 1) using dependent OS extrapolation with lognormal curve	–	**£45,164†**	**-**
ERG scenarios 2 & 6: Full TAGS data (regardless of prior ramucirumab) for both OS and treatment effect			
A. Scenario 2 with independent OS extrapolation using lognormal curve	Alternative OS extrapolations using the lognormal or log-logistic curves applied independently to each arm are clinically plausible (see issue 3). Alternative utility values from TA378 are clinically plausible (see issue 5)	£55,600	+£10,436
B. A + utility values from TA378		£58,651	+13,487
C. Scenario 2 with independent OS extrapolation using log-logistic curve		£52,655	+£7,491
D. C + utility values from TA378		£55,691	+10,527
E. ERG preferred scenario 6 (EU subgroup) with alternative independent OS extrapolation using lognormal or log-logistic curves	It is plausible that the EU subgroup is more generalisable to the NHS in England (see issue 2)	No data (Kaplan-Meier plots for EU subgroup not available to ERG)	
F. E + utility values from TA378			
Cumulative impact of the technical team's preferred assumptions on the cost-effectiveness estimate (A or B, C or D, E or F)	–	£52,655 to £58,651*	+£7,491 to +£13,487*
† PSA result of ERG run company base case £45,314 per QALY gained			

Alteration	Technical team rationale	ICER	Change from base case
* There is uncertainty around the upper limit because scenarios E and F cannot be calculated. However, the ERG note that from scenario 1 to 2, the ICER increases by around £5,000 and in the absence of further evidence, it may be appropriate to assume that this level of increase would also apply when moving from Scenario 5 to Scenario 6. Such calculations would indicate ICERs in excess of £60,000 when using independent models and £50,000 per QALY gained when assuming dependent models.			

Table 8: Outstanding uncertainties in the evidence base

Area of uncertainty	Why this issue is important	Likely impact on the cost-effectiveness estimate
Trial population	The inclusion criteria of the TAGS trial were more restrictive compared with the full marketing authorisation for trifluridine-tipiracil because the trial only included people: • with ECOG performance score 0 or 1 • who have had 2 prior regimens that must have included a fluoropyrimidine, platinum, and either a taxane and/or irinotecan-containing regimen	Unknown
Potential clinical subgroups	Subgroup analyses from TAGS suggest there may be clinically relevant subgroups based on: • prior irinotecan • prior taxane • site of cancer (gastric or gastro-oesophageal junction) • HER-2 status	Unknown

Table 9: Other issues for information

Issue	Comments
Exclusion of oral chemotherapy administration costs	The ERG's preferred analyses exclude oral administration costs for trifluridine-tipiracil (included in the company base case), because: (1) this was accepted by the committee in TA405, and (2) they may be seen as transfer payments, with the unit cost reimbursed to the NHS Trust being much higher than the actual cost of administering an oral tablet, such that the money received is simply reallocated for other uses. This is explored in a scenario analysis which increased the company's base case ICER from £45,164 to £48,592.
Innovation	The company considers the drug to be innovative because it offers a third-line orally administered treatment option and allows treatment to be continued in community settings. However, the technical team considers that all relevant benefits associated with the drug are adequately captured in the model.
Equality considerations	No equalities issues were identified by the company, consultees and their nominated clinical experts and patient experts.

Authors

Peter Selby

Appraisal committee chair

Abi Senthinathan

Technical lead

Jamie Elvidge

Technical adviser

Frances Sutcliffe

Associate director

With input from the lead team:

Alex Cale

Lead team member

Mike Chambers

Lead team member

Stella O'Brien

Lead team member