in collaboration with:

Ramucirumab for treating advanced gastric cancer or gastro-oesophageal junction adenocarcinoma previously treated with chemotherapy

Produced by	Kleijnen Systematic Reviews Ltd in collaboration with Erasmus University Rotterdam (EUR) and Maastricht University
Authors	Rob Riemsma, Reviews Manager, KSR Ltd, UK Maiwenn Al, Health Economics Researcher, EUR, The Netherlands Nasuh Büyükkaramikli, Health Economics Researcher, EUR, NL Nigel Armstrong, Health Economist, KSR Ltd, UK Hedwig Blommestein, Health Economics Researcher, EUR, NL Fiona Clay, Systematic Reviewer, KSR Ltd, UK Janine Ross, Information Specialist, KSR Ltd, UK Gill Worthy, Statistician, KSR Ltd, UK Hans Severens, Professor of Evaluation in Healthcare, EUR, NL Jos Kleijnen, Director, KSR Ltd, Professor of Systematic Reviews in Health Care, Maastricht University, UK
Correspondence to	Rob Riemsma, Kleijnen Systematic Reviews Unit 6, Escrick Business Park Riccall Road, Escrick York, UK YO19 6FD

Date completed 21/07/2015

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 14/177/03 STA.

Declared competing interests of the authors
None.

Acknowledgements
None.

Commercial in confidence (CiC) data are highlighted in blue throughout the report.
Academic in confidence (AiC) data are highlighted in blue throughout the report.

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

This report should be referenced as follows:
Riemsma R, Al M, Büyükkaramikli N, Armstrong N, Blommestein H, Clay F, Ross J, Worthy G, Severens JL, Kleijnen J. Ramucirumab for treating advanced gastric cancer or gastro-oesophageal junction adenocarcinoma previously treated with chemotherapy. York: Kleijnen Systematic Reviews Ltd, 2015.

Contributions of authors
Rob Riemsma acted as project lead and systematic reviewer on this assessment, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Maiwenn Al acted as health economic project lead, critiqued the manufacturer's economic evaluation and contributed to the writing of the report. Nasuh Büyükkaramikli, Hedwig Blommestein and Nigel Armstrong acted as health economists on this assessment, critiqued the manufacturer's economic evaluation and contributed to the writing of the report. Fiona Clay acted as systematic reviewer, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Gill Worthy acted as statistician, critiqued the analyses in the manufacturer's submission and contributed to the writing of the report. Janine Ross critiqued the search methods in the submission and contributed to the writing of the report. Hans Severens critiqued the manufacturer's definition of the decision problem and their description of the underlying health problem and current service provision, contributed to the writing of the report. Jos Kleijnen critiqued the manufacturer's definition of the decision problem and their description of the underlying health problem and current service provision, contributed to the writing of the report and supervised the project.

Abbreviations

AE	Adverse Events
AF	Acceleration Factor
AIC	Akaike Information Criterion
ANC	Absolute Neutrophil Count
ASCO	American Society of Clinical Oncology
bd/b.i.d	Twice Daily
BI	Budget Impact
BIC	Bayesian Information Criterion
BORR	Best overall response rate
BSA	Body Surface Area
BSC	Best Supportive Care
CDF	Cancer Drugs Fund
CE	Cost Effectiveness
CEA	Cost-effectiveness Analysis
CEAC	Cost effectiveness Acceptability Curve
CHMP	Committee for Medicinal Products for Human Use
CI	Confidence Interval
CR	Complete response
CRD	Centre for Reviews and Dissemination
CrI	Credible interval
CSR	Clinical study report
CHMP	Committee for Medicinal Products for Human Use
CTCAE	Common Terminology Criteria for Adverse Events (National Cancer Institute)
DALY	Disability-adjusted life year
Den	Denominator
df	Degrees of freedom
DPG	Disease population group
ECOG	Eastern Cooperative Oncology Group
ECOG-PS	Eastern Cooperative Oncology Group performance status
EMA	European Medicines Agency
EORTC	European Organisation for Research and Treatment of Cancer
EORTC QLQ-C30	European Organisation for Research and Treatment of Cancer quality of life questionnaire
EPAR	European public assessment report
EQ-5D	European Quality of Life-5 Dimensions
EQ-5D-3L	European Quality of Life-5 Dimensions, three-level scale
ERG	Evidence Review Group
ESMO	European Society for Medical Oncology
EUR	Erasmus University Rotterdam
FDA	Food and Drug Administration
5-FU	5-Fluorouracil

GEJ/GOJ	Gastro-oesophageal junction
GIST	Gastrointestinal stromal tumour
HR	Hazard ratio
HRQL	Health-related Quality of Life
HTA	Health Technology Assessment
IC	Indirect Comparison
ICD	International Classification of Diseases
ICER	Incremental Cost-effectiveness Ratio
IDMC	Independent data monitoring committee
IRAC	Independent response adjudication committee
ITT	Intention to Treat
IV	Intravenous
KM	Kaplan–Meier
KSR	Kleijnen Systematic Reviews
LYS	Life Year Saved
MAH	Marketing authorisation holder
MedDRA	Medical Dictionary for Regulatory Activities
MeSH	Medical Subject Headings
MHRA	Medicines and Healthcare Products Regulatory Agency
mg	Milligram
MR	Minimal response
MRU	Medical resource utilisation
MS	Manufacturer’s Submission
MTC	Mixed Treatment Comparison
NA	Not applicable
NHS	National Health Services
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
NR	Not Reported
Num	Numerator
od	Once Daily
OR	Odds Ratio
ORR	Overall response rate
OS	Overall survival
PCT	Primary Care Trust
PD	Progressive disease
PFS	Progression-free survival
PK	Pharmacokinetic
PR	Partial response
PRESS	Peer Review of Electronic Search Strategies
PRO	Patient-reported outcome
PS	Performance status
PSA	Probabilistic Sensitivity Analyses
PSS	Personal Social Services

QALY(s)	Quality-adjusted Life Year(s)
QoL	Quality of life
RCT	Randomised Controlled Trial
RECIST	Response Evaluation Criteria In Solid Tumours
RPSFTM	Rank preserving structure failure time model
RR	Relative Risk; Risk Ratio
SAE	Serious Adverse Events
SC	Subcutaneous
ScHARR	School of Health and Related Research
sCR	Stringent complete response
SD	Standard deviation
SF-36	Short form 36
SHTAC	Southampton Health Technology Assessments Centre
SIGN	Scottish Intercollegiate Guidelines Network
SMC	Scottish Medicines Consortium
SPC	Summary of product characteristics
STA	Single Technology Appraisal
UMC	University Medical Centre
TCS	Treatment continuation scheme
TEAEs	Treatment-emergent adverse events
TESAEs	Treatment-emergent serious adverse events
TTF	Time to failure
TTP	Time to progression
UK	United Kingdom
VEGF	Vascular endothelial growth factor
VGPR	Very good partial response
WHO	World Health Organisation

Table of Contents

Abbreviations **3**

Table of Tables **8**

Table of Figures **9**

1. SUMMARY **11**

1.1 Critique of the decision problem in the manufacturer's submission 11

1.2 Summary of clinical effectiveness evidence submitted by the manufacturer 11

1.3 Summary of the ERG's critique of clinical effectiveness evidence submitted 12

1.4 Summary of cost-effectiveness submitted evidence by the manufacturer 13

1.5 Summary of the ERG's critique of cost-effectiveness evidence submitted 15

1.6 ERG commentary on the robustness of evidence submitted by the manufacturer 16

1.7 Summary of exploratory and sensitivity analyses undertaken by the ERG 17

2 BACKGROUND **18**

2.1 Critique of manufacturer's description of underlying health problem. 18

2.2 Critique of manufacturer's overview of current service provision 19

3 Critique of manufacturer's definition of decision problem **22**

3.1 Population 26

3.2 Intervention 26

3.3 Comparators 26

3.4 Outcomes 27

3.5 Other relevant factors 27

4 CLINICAL EFFECTIVENESS **28**

4.1 Critique of the methods of review(s) 28

4.2 Critique of trials of the technology of interest, their analysis and interpretation (and any standard meta-analyses of these) 37

4.3 Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison 61

4.4 Critique of the indirect comparison and/or multiple treatment comparison 66

4.5 Additional work on clinical effectiveness undertaken by the ERG 67

4.6 Conclusions of the clinical effectiveness section 68

5 COST-EFFECTIVENESS **71**

5.1 ERG comment on manufacturer's review of cost-effectiveness evidence 71

5.2 Summary and critique of manufacturer's submitted economic evaluation by the ERG 74

5.3 Exploratory and sensitivity analyses undertaken by the ERG 120

5.4 Conclusions of the cost-effectiveness section 130

6 IMPACT ON THE ICER OF ADDITIONAL CLINICAL AND ECONOMIC ANALYSES UNDERTAKEN BY THE ERG 132

7 END OF LIFE 134

8 OVERALL CONCLUSIONS 136

8.1 Implications for research 136

9 REFERENCES 137

Appendix 1: ERG Search Strategies 143

Appendix 2: Phillips et al checklist 144

Appendix 3: Details model changes implemented by the ERG 150

Appendix 4: PSA for monotherapy 151

Table of Tables

Table 3.1: Statement of the decision problem (as presented by the manufacturer) 22
Table 4.1: Eligibility criteria used in search strategy for systematic review 30
Table 4.2: Data sources for RCT's of ramucirumab: RAINBOW and REGARD 31
Table 4.3: Summary of the Risk of Bias assessment for the ramucirumab studies 33
Table 4.4: Summary of the Risk of Bias assessment for the docetaxel studies included in the evidence network 34
Table 4.5: Summary of the Risk of Bias assessment for the irinotecan studies included in the evidence network 36
Table 4.6: Overview of the RAINBOW and REGARD trials 37
Table 4.7: Summary of methodology of the RAINBOW and REGARD trials 39
Table 4.8: Major eligibility criteria of the RAINBOW and REGARD trials 41
Table 4.9: Reasons for treatment discontinuation 44
Table 4.10: Characteristics of participants in the trials by randomised group (ITT) 44
Table 4.11: Overall survival of patients in the RAINBOW trial (ITT) 47
Table 4.12: RAINBOW: Progression-free survival results (ITT population) 50
Table 4.13: RAINBOW: Response to treatment results – ITT population 51
Table 4.14: RAINBOW: EORTC QLQ-C30 - global health status results, ITT population 51
Table 4.15: RAINBOW: EQ-5D Results, ITT population 51
Table 4.16: RAINBOW: Overall safety results, safety population 52
Table 4.17: RAINBOW: Treatment-emergent adverse events occurring in at least 10% of patients on ramucirumab plus paclitaxel, irrespective of causality 53
Table 4.18: RAINBOW: Adverse events of special interest* 54
Table 4.19: Overall survival of patients in the REGARD trial (ITT) 56
Table 4.20: REGARD: Progression-free survival results (ITT population) 59
Table 4.21: REGARD: Response to treatment results – ITT population 60
Table 4.22: REGARD: Overall safety results, safety population 60
Table 4.23: REGARD: Adverse events of special interest 61
Table 4.24: Heterogeneity in network of studies for combination therapy 64
Table 4.25: ERG results of CS base-case for overall survival (HR, 95% Credible interval (CrI)) 67
Table 4.26: NMA results for OS (including Roy 2013) 67
Table 4.27: NMA results for PFS 68
Table 5.1: Summary of the manufacturer's economic evaluation 72
Table 5.2: Comparison of the CS model with the NICE reference case 74
Table 5.3: AIC and BIC statistics for PFS for combination therapy model. 82
Table 5.4: HRs (95% CrI) for PFS obtained from NMA 85
Table 5.5: HRs (95% CI) for OS obtained from ITC. 85
Table 5.6: AIC and BIC statistics for PFS for monotherapy model. 87
Table 5.7: AIC and BIC statistics for OS for the monotherapy model 89
Table 5.8: Events observed in the comparator treatments excluded from the analysis 92
Table 5.9: Utility values per health state for the combination therapy and monotherapy model 94
Table 5.10: Unit costs associated with the technology in the economic model 97
Table 5.11: Length of stay: combination and monotherapy model 99
Table 5.12: Hospitalisation probabilities: combination and monotherapy model 99
Table 5.13: Base case: Combination therapy results 103
Table 5.14: Summary of model results compared with clinical data: Combination therapy 104
Table 5.15: Results by technology and health state for the combination therapy 104
Table 5.16: Base case results for monotherapy 106

Table 5.17: Model results compared with clinical data for monotherapy 107
Table 5.18: Results by technology and health state of the monotherapy model.....107
Table 5.19: Summary of PSA distributions used for parameters.....108
Table 5.20: Cost-effectiveness results of probability sensitivity analysis for combination therapy.....109
Table 5.21: Results of deterministic sensitivity analysis (CS Table 141) 110
Table 5.22: Cost-effectiveness results of probability sensitivity analysis of monotherapy...113
Table 5.23: Results of the deterministic sensitivity analysis 115
Table 5.24: Summary of model results with different parametric curves compared with clinical data of the RAINBOW trial 117
Table 5.25: Summary of model results with different parametric curves compared with clinical data of the REGARD trial.....117
Table 5.26: Summary of model results compared with clinical data for docetaxel.....119
Table 5.27: Summary derived AE rates for IRI and FOLFIRI 121
Table 5.28: Base case model results with additional comparators 122
Table 5.29: Base case model results (intervention versus comparator only).....122
Table 5.30: Base case model results with additional comparators 123
Table 5.31: New ERG Base case for combination therapy 126
Table 5.32: New ERG Base case for monotherapy 126
Table 5.33: Combination therapy results with HRs calculated from the evidence network including Roy et al. 128
Table 5.34: Combination therapy RAM+PAC vs PAC results when only RAINBOW trial data is used.....129
Table 5.35: Combination therapy results when EQ-5D scores collected in different time points in the RAINBOW trial are implemented. 129
Table 5.36: Combination therapy RAM+PAC vs PAC results when only RAINBOW trial data is used and utility values from RAINBOW EQ-5D are directly used 130
Table 6.1: Revised base case cost-effectiveness analysis, incorporating corrections and amendments identified by the ERG: combination therapy RAM+PAC vs BSC.....132
Table 6.2: Revised base case cost-effectiveness analysis, incorporating corrections and amendments identified by the ERG for combination therapy RAM+PAC vs DOC 133
Table 6.3: Revised base case cost-effectiveness analysis, incorporating corrections and amendments identified by the ERG: monotherapy RAM vs BSC.....133

Table of Figures

Figure 2.1: Current treatment pathway based on current NICE recommendations for patients with GC/GOJ 20
Figure 2.2: Proposed treatment pathway for patients with GC/GOJ 21
Figure 4.1: Evidence network used in CS analyses 32
Figure 4.2: RAINBOW: Kaplan-Meier plot of overall survival (ITT population).....48
Figure 4.3: Forest plot for subgroup univariate analyses of overall survival 49
Figure 4.4: RAINBOW: Kaplan-Meier plot of Progression-Free Survival (ITT population).50
Figure 4.5: REGARD: Kaplan-Meier plot of overall survival (ITT population) 57
Figure 4.6: Forest plot for subgroup univariate analyses of overall survival 58
Figure 4.7: REGARD: Kaplan-Meier plot of Progression-Free Survival (ITT population)....59
Figure 4.8: Evidence network of relevant studies for the company analyses.....62
Figure 5.1: Economic model structure for monotherapy and combination therapy 75
Figure 5.2: PFS Kaplan-Meier plots from RAINBOW 81
Figure 5.3: Base case PFS estimated using Weibull distribution for combination therapy.....82
Figure 5.4: Overall survival Kaplan Meier plots for ITT population 83

Figure 5.5: Base case OS for RAM+PAC and PBO+PAC using the Kaplan-Meier data with exponential extrapolation....84
Figure 5.6: PFS Kaplan-Meier plots from REGARD....86
Figure 5.7: Base case PFS estimated using lognormal distribution for monotherapy....87
Figure 5.8: Overall survival Kaplan Meier plots for ITT population....88
Figure 5.9: Base case OS extrapolations for monotherapy model (gamma distribution)....89
Figure 5.10: Cost-effectiveness plane for combination therapy (QALYs)....109
Figure 5.11: Cost-effectiveness acceptability curve for combination therapy....110
Figure 5.12: Overall of KM and joint parametric estimates of OS using log-logistic....113
Figure 5.13: Cost-effectiveness plane for monotherapy (QALYs)....114
Figure 5.14: Cost-effectiveness acceptability curve for monotherapy....114
Figure 5.15: Overlay of KM and parametric estimates of OS for RAM....116
Figure 5.16: BSC OS parametric curve where the HR is derived from the ITC vs. BSC OS KM from REGARD....118
Figure 5.17: BSC PFS parametric curve where the HR is derived from the NMA vs. BSC PFS KM from REGARD....119
Figure 5.18: Cost-effectiveness plane for combination therapy (QALYs)....124
Figure 5.19: Cost-effectiveness acceptability curves for combination therapy....124
Figure 5.20: Cost-effectiveness plane for combination therapy (QALYs)....127
Figure 5.21: Cost-effectiveness acceptability curves for combination therapy....128

1. SUMMARY

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

There are two issues regarding the decision problem: the population and the comparators.

The patient population described in the final scope is as follows: "Adults with advanced gastric cancer or gastro-oesophageal junction adenocarcinoma previously treated with chemotherapy". This is in line with the patient population included in the company submission and in the main trials for this submission, the REGARD and RAINBOW studies.

However, according to the marketing authorisation ramucirumab was approved for the following indications:

- ramucirumab in combination with paclitaxel is indicated for the treatment of adult patients with advanced gastric cancer or gastro-oesophageal junction adenocarcinoma with disease progression after prior platinum and fluoropyrimidine chemotherapy.
- ramucirumab monotherapy is indicated for the treatment of adult patients with advanced gastric cancer or gastro-oesophageal junction adenocarcinoma with disease progression after prior platinum or fluoropyrimidine chemotherapy, for whom treatment in combination with paclitaxel is not appropriate.

The REGARD trial (ramucirumab monotherapy) did not specify whether patients were suitable for treatment in combination with paclitaxel. Therefore, patients for whom treatment in combination with paclitaxel is appropriate will have been included in the REGARD trial.

Regarding the comparators, the company used placebo as the comparator for ramucirumab monotherapy. This seems reasonable if the committee accepts the reasoning from the company that patients 'for whom treatment in combination with paclitaxel is not appropriate' are the same as patients 'for whom further cytotoxic chemotherapy is not appropriate'.

For ramucirumab in combination with paclitaxel, the comparator in the RAINBOW trial was paclitaxel monotherapy. In addition, the company presents results for comparisons with best supportive care and docetaxel using indirect comparisons (Bucher method) and a network meta-analysis. Results from the network meta-analysis are used in the economic analyses. Comparisons with irinotecan monotherapy and FOLFIRI are missing. It is particularly odd that the comparison with irinotecan monotherapy is missing, because a comparison of ramucirumab in combination with paclitaxel versus best supportive care or docetaxel is only possible in a network that includes irinotecan monotherapy. We asked the company to provide data for the comparisons with irinotecan monotherapy and FOLFIRI.

1.2 Summary of clinical effectiveness evidence submitted by the manufacturer

The majority of searches in the CS were well documented and easily reproducible; searches were carried out in line with NICE guidelines.

<u>Ramucirumab monotherapy</u>

Ramucirumab monotherapy compared to best supportive care (BSC) relies on one trial (REGARD). In the REGARD trial ramucirumab monotherapy was associated with a 1.4 month increase in median overall survival in comparison to placebo. Median overall survival was 5.2 months among patients treated with ramucirumab and 3.8 months among those

treated with placebo (HR=0.776, 95% CI: 0.603 to 0.998). The relative gain in PFS was 0.8 months for patients treated with ramucirumab with a median PFS of 2.1 months (95% CI: 1.5 to 2.7) in the ramucirumab group and 1.3 months (95% CI: 1.3 to 1.4) in the placebo group (HR=0.483, 95% CI: 0.376 to 0.620).

Ramucirumab plus paclitaxel

Ramucirumab+paclitaxel was used in one trial, the RAINBOW trial, which compares ramucirumab+paclitaxel with paclitaxel. The trial shows favourable results for ramucirumab+paclitaxel in terms of overall survival (HR=0.807 (95% CI: 0.678 to 0.962)) and PFS (HR=0.635 (95% CI: 0.536 to 0.752)). This represents a 31% (2.27 months) longer median overall survival in the ramucirumab+paclitaxel arm (9.63 months (95% CI 8.6-10.8) versus 7.36 (95% CI 6.3-8.4) months in the placebo+paclitaxel arm). The relative gain in PFS was 1.5 months for patients treated with ramucirumab+paclitaxel with a median PFS of 4.4 months (95% CI: 4.2 to 5.3) in the ramucirumab+paclitaxel group and 2.9 months (95% CI: 2.8 to 3.0) in the placebo+paclitaxel group.

Using indirect comparisons and network meta-analyses, the CS presents results comparing ramucirumab+paclitaxel with other relevant comparators such as docetaxel and best supportive care. In response to the clarification questions from the ERG, the company also presented results for ramucirumab+paclitaxel compared with irinotecan and (m)FOLFIRI.

In terms of overall survival, ramucirumab+paclitaxel was found significantly better than BSC (HR=0.41; 95% CrI: 0.24 to 0.70), paclitaxel (HR=0.81; 95% CrI: 0.68 to 0.96) and irinotecan (HR=0.71; 95% CrI: 0.52 to 0.99). In terms of progression free survival, ramucirumab+paclitaxel was found significantly better than paclitaxel (HR=0.64; 95% CrI: 0.54 to 0.75) and irinotecan (HR=0.56; 95% CrI: 0.41 to 0.76). A comparison with BSC was not possible for PFS.

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

Ramucirumab monotherapy

The main issue with the evidence for ramucirumab monotherapy (the REGARD trial) is that the REGARD trial (ramucirumab monotherapy) did not specify whether patients were suitable for treatment in combination with paclitaxel. Therefore, patients for whom treatment in combination with paclitaxel is appropriate will have been included in the REGARD trial. Given that eligibility criteria for RAINBOW and REGARD were almost the same and that all patients in the RAINBOW trial received paclitaxel, it is possible that all patients in the REGARD trial were eligible for paclitaxel. We did ask the company how many patients in each arm of the REGARD trial were not suitable for paclitaxel, but the company responded that it was not possible to estimate this.

The comparison of ramucirumab monotherapy versus BSC is sufficient and in line with the NICE final scope if it is accepted that 'not suitable for paclitaxel' is the same as 'not suitable for further cytotoxic chemotherapy'. If this is not accepted, comparisons with cytotoxic chemotherapy other than paclitaxel (docetaxel, irinotecan and FOLFIRI) are missing.

An indirect comparison with docetaxel, using the COUGAR-02 trial shows that the hazard ratio of overall survival of ramucirumab versus docetaxel is not significantly different, but actually favours docetaxel (HR=1.16 (95% CI: 0.77 to 1.73)).

Ramucirumab plus paclitaxel

The analyses for ramucirumab plus paclitaxel versus comparators other than paclitaxel rely on using data from a trial in a completely Japanese population (Hironaka 2013). As explained in the CS, "high rates of salvage therapy have been reported in previous Asian trials. A higher rate of post discontinuation (PDT) third-line therapy was expected in Region 3 (parts of Asia, including Japan), potentially confounding the OS treatment effect of the ramucirumab+paclitaxel regimen due to differing rates of PDT." In addition, the CS lists the following differences between Western and Asian countries:

Asian countries have a higher incidence and prevalence of Gastric cancer (GC) than Western nations. The adoption of national screening programs in Asian countries has resulted in diagnosis in the early stages of the disease in up to 50-60% of cases, while in Western nations patients are typically diagnosed at an advanced stage of GC and therefore have a poorer prognosis.

Differences also exist in GC histology (Western patients have a higher incidence of diffuse histology/proximal tumour types having a poorer prognosis than intestinal histology/distal tumour types seen in Asian patients). The surgical treatment of early GCs with extensive lymph node dissection (D2 resection) occurs more frequently in Asian countries. A higher proportion of patients receive second-line chemotherapy (and beyond) in Asia compared to US and Europe which extends survival in those patients.

As the network necessary to compare ramucirumab+paclitaxel with best supportive care and with other active treatments (docetaxel, irinotecan, and (m)FOLFIRI) relies on at least one trial in a multinational or Western population (RAINBOW, Thuss-Patience 2011, Roy 2013 and COUGAR-02) and one trial in an Asian population (Hironaka 2013 and Sym 2013), the network has a high level of heterogeneity and is therefore unreliable. Any results based on indirect comparisons or a network meta-analysis should be interpreted with great caution due to significant heterogeneity between studies.

1.4 Summary of cost-effectiveness submitted evidence by the manufacturer

The cost-effectiveness of ramucirumab+paclitaxel (RAM+PAC) and ramucirumab monotherapy (RAM) was evaluated using a partitioned survival model (which is essentially the same as a Markov model). The model included three different health states; pre-progression, post-progression and death. All patients enter the pre-progression health state and stay in this health state until they experience disease progression or death and receive treatment until disease progression or treatment cessation due to other reasons. Patients remain in the post-progression health state until they die. A weekly cycle length and a half-cycle correction were implemented. The model has an NHS perspective and a lifetime time horizon (which equates to 7.23 years). Costs and effects were discounted at an annual 3.5% rate. While the structure was similar for the combination therapy (RAM+PAC), and monotherapy (RAM) model, the input parameters were different.

For the combination therapy model transition probabilities between the health states were determined from the RAINBOW trial and for the non-trial comparators by the network meta-analysis and indirect comparison. Although best supportive care, docetaxel, irinotecan, and FOLFIRI were identified in the NICE scope, only best supportive care and docetaxel were included as comparators in the combination therapy model. Paclitaxel was included for validation of the economic model as it was the comparator treatment in the RAINBOW trial.

Data from the REGARD trial were used to model progression-free survival (PFS) and overall survival (OS) of best supportive care and ramucirumab in the monotherapy model.

To select the most appropriate distribution for PFS and OS, the functional form of the underlying hazard was assessed, goodness of fit tests were conducted and the suitability of parametric distribution was assessed.

In the combination therapy model, the Weibull distribution was chosen to parameterise the PFS of the RAM+PAC treatment. An alternative approach was used to model OS. OS KM data for the trial period were used with an exponential distribution to extrapolate from the end of the trial. The HR estimates from the network meta-analysis for best supportive care and docetaxel compared to RAM+PAC were applied to the baseline curves for RAM+PAC.

In the monotherapy model, the lognormal distribution was chosen in the base case to parameterise the PFS of the RAM and BSC treatment arms. The gamma distribution was chosen to estimate the OS of the RAM+BSC and PBO+BSC treatments

A baseline utility value for the pre-progression health state in the combination therapy model was obtained from the data collected during the RAINBOW trial. From the same data, a utility increment for responders to second line treatment was calculated. Utility decrements for adverse events were obtained from the literature and subtracted from the baseline utility estimate. The mean EQ-5D score of the RAINBOW trial for patients who discontinued treatment due to progression was used for the post-progression health state.

Data from the RAINBOW trial (baseline utility value pre-progression and utility value post-progression) were also used in the monotherapy model. Utility decrements for adverse events from the literature were applied to the baseline utility value.

Cost of the intervention and the comparators are comprised of the drug acquisition costs, drug administration costs and the cost of monitoring and the tests. In order to calculate the required drug doses for each regimen, an estimate of body weight and body surface area (BSA) was obtained from patients included in the RAINBOW and REGARD trial for the combination therapy and monotherapy model, respectively. A chart review of hospital medical records was used to identify the cost components of best supportive care. Costs further consist of follow-up costs, AE costs, hospitalisation costs, third-line therapy costs (including drug acquisition and administration cost and the cost of follow-up care) and terminal care costs.

The base case ICERs (cost per QALY gained) for the combination therapy model were £53,830 per QALY gained for DOC compared to BSC and £118,209 per QALY gained for RAM+PAC compared to BSC. The base case ICER for the monotherapy model was £188,640 per QALY gained for BSC compared to RAM (mono). Probabilistic and deterministic sensitivity analyses were conducted.

The deterministic sensitivity analysis of the combination therapy model showed that the ICER was most sensitive to the source of drug prices (eMIT vs BNF), length of hospitalisation stay, dose intensity and the BSA/body weight source data (all trial patients versus Region 1 trial patients). The deterministic sensitivity analysis of the monotherapy model revealed that the ICER was most sensitive to the hospitalisation rates, length of hospitalisation, assumptions on wastage (vial wastage versus vial sharing) and extrapolation of post-progression survival.

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

The model is clearly described and straightforward. However, a single post-progression health state is a simplification of the treatment pathway, as some patients may receive a third line treatment. The ITT population from the RAINBOW trial seems to reflect that of the licensed indication for combination therapy. However, a third of the ITT population from the trial were Asian patients, who are not reflective for the UK population. The ERG is highly uncertain to what extent the patient population in the monotherapy model aligns with the licensed population since the patient population in the economic model is based on the REGARD trial where eligibility for PAC combination therapy was not an exclusion criteria.

The ERG does not agree with excluding comparators that were included in the final scope and thinks that paclitaxel, docetaxel, irinotecan, and FOLFIRI should be included in the combination therapy model. Inclusion of docetaxel, irinotecan, and FOLFIRI in the monotherapy model depends on whether it is accepted that 'not suitable for paclitaxel' is the same as 'not suitable for further cytotoxic chemotherapy'. If this is not accepted, comparisons with cytotoxic chemotherapy other than paclitaxel (docetaxel, irinotecan and FOLFIRI) are missing.

In general, the process for the extrapolation of survival curves was clear; however, the choice of the survival modelling does not follow the same procedure for all PFS and OS curves in the combination therapy model and monotherapy model. According to the ERG, choosing the Weibull distribution for PFS over the log-logistic (with a better fit) for the sake of the PH assumption, was unnecessary as well as conflicting with the approach taken for the survival modelling of OS curves. Based on AIC/BIC test results, Cox-Snell residuals and visual check, the ERG finds the interval-censored log-normal distribution for PFS modelling and the Gamma distribution for OS modelling of BSC and RAM monotherapy arms plausible.

The ERG agrees with the company that AEs and tumour response are clinically meaningful events that may have an impact on HRQoL. Ideally, utility decrements should have been obtained from EQ-5D data of the RAINBOW or REGARD trial. Furthermore, the ERG could not verify all the AEs rates used in the combination and monotherapy models. Utility values for the monotherapy model were based on utility values from the combination therapy trial. This trial does not reflect the treatment options BSC and RAM nor the patient population for monotherapy (i.e. patients for whom treatment in combination with paclitaxel is not appropriate).

The ERG identified an error in the half cycle correction of the model submitted by the company. This error was confirmed in the clarification letter. The impact of this correction on

the ICER was negligible for RAM+PAC compared to BSC (£118,174) and RAM compared to BSC (£188,437).

The ERG also detected a technical error in the costs for docetaxel (both in the second and third line). Furthermore, according to the ERG, the drug acquisition costs for RAM+PAC are underestimated since these were derived based on the average weight of the patients in the RAINBOW trial (including a high proportion of Asian patients). The ERG identified double counting of hospitalisation costs because the hospitalisation obtained from the trial included hospitalisations due to AEs.

According to the ERG, a full incremental analysis of the comparators docetaxel, irinotecan and FOLFIRI should have been included in the base case results.

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

1.6.1 Strengths

Searches were carried out in line with the NICE guide to the methods of technology appraisal sections 5.2.2 and 5.2.4. The company's submission and response to clarification provided sufficient details for the ERG to appraise the searches. Additional searches were carried out for conference abstracts and other resources.

The two ramucirumab trials (REGARD and RAINBOW) are good quality RCTs including more than 300 patients per arm in the RAINBOW trial. The network meta-analyses methods were found to be appropriate and the results were reproducible.

OS and PFS from the RAINBOW and REGARD trial were mature and therefore the uncertainty regarding long-term follow-up is small. The impact of assumptions was extensively explored in the deterministic sensitivity analysis.

1.6.2 Weaknesses and areas of uncertainty

The reporting of database hosts, time spans and date of searches could have been more thorough in places. Searches in some sections of the submission were last run in 2013 and 2014; updates to theses searches could have provided more recent material. The ERG was concerned about the language bias of restricting some of the searches to English language only; this is not in line with current best practice.

There were problems with the population and with the comparators as described in section 1.1. In addition, it is not clear how representative the trials are for a UK population. The company presented separate data from region 1 patients in the REGARD and RAINBOW trials to make the results more applicable to the UK. However, region 1 patients only included very few UK patients and for ramucirumab plus paclitaxel, comparisons still need to be made with trials in non-region 1 patients. We found considerable heterogeneity in the indirect comparisons and network meta-analyses, making the results uncertain.

An indirect comparison was used to estimate the costs and effects for BSC and DOC in the combination therapy model. The assumptions used and the heterogeneity of the included trials in the indirect comparisons and network meta-analyses, make the estimates and thus the interpretation of the ICERs very uncertain. Furthermore, paclitaxel, irinotecan and FOLFIRI were not included as comparators in the combination therapy model.

Utility data for the monotherapy model was obtained from patients who may not reflect the patient population for monotherapy (i.e. patients for whom treatment in combination with paclitaxel is not appropriate).

1.7 Summary of exploratory and sensitivity analyses undertaken by the ERG

A new base case was identified by the ERG leading to ICERS of £129,431 and £188,055 compared to BSC for the combination therapy and monotherapy, respectively.

In the exploratory ERG analyses the comparators, as defined by the final scope, were included. Based on these analyses, it seems that the additional comparators irinotecan and FOLFIRI are on the efficient frontier for both combination therapy and monotherapy. However these analyses should be interpreted with caution since several assumptions were made.

In addition, the ERG explored three different scenarios. First, the study of Roy et al 2013 was included in the OS network meta-analysis. This scenario analysis revealed that the ICER is sensitive to the inclusion of Roy et al 2013 study. Second, the ICER was calculated if efficacy data is only based on direct evidence from the RAINBOW trial. This showed that the ICER of RAM+PAC compared to PAC would be almost £400,000 per QALY gained. Third, the utility values from the RAINBOW trial were directly implemented. The impact of this scenario on the ICER was limited.

2 BACKGROUND

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

Gastric cancer and gastro-oesophageal junction adenocarcinoma

The final scope stated that *"Gastric cancer (GC) is a malignant tumour arising from cells in the stomach. The most common type of gastric cancer is gastric or gastro-oesophageal junction adenocarcinoma, which affects about 95% of people with the disease."*[1]

According to page 36 of the company submission (CS), *"GC is a significant health issue globally. It is the third most common cause of cancer related deaths in men and women worldwide. GC/GOJs are rare and aggressive types of cancer that significantly shorten life expectancy".*

"In Western populations, GC/GOJs more commonly present in the proximal stomach or gastro-oesophageal junction and the risk factors for these more proximal tumours include gastro-oesophageal reflux disease (GORD), Barrett's oesophagus, a complication of gastro-oesophageal reflux disease (GORD), smoking, and obesity".

ERG comment: While the ERG note that there is lack of uniformity in the classification systems for both gastric cancer and oesophageal cancers, it is generally thought that gastric cancers originate in the lining of stomach, while a stomach cancer may refer to a malignancy found in any part of the stomach, Gastro-oesophageal junction (GOJ) usually originates within 5 cm proximal and distal of the anatomic cardia. According to the CS, GC/GOJs are rare types of cancer; however, gastric cancer is estimated to be the third most common cancer causing deaths worldwide. The high mortality rate associated with GC/GOJs stems from the aggressive nature of this cancer and that diagnosis often occurs in the later stages of the disease.

Diagnosis and prognosis

According to page 35 of the CS, *"Diagnosis often occurs in the later stages of the disease as the signs and symptoms may not appear until the disease is advanced"*

According to page 36 of the CS, *"Gastric and GOJ cancer have a very poor prognosis because of these vague symptoms and diagnosis late in the course of the disease"*

ERG comment: Patients diagnosed with GC/GOJ cancer stages 0-2 have a high chance of cure, for patients diagnosed with advanced GC/GOJ cancer (stages 3 and 4), the chance of cure diminishes drastically. The majority of patients from the Western countries are usually diagnosed in the advanced disease stage with unresectable tumours.

GC/GOJ cancers are usually asymptomatic at the early stages. Patients with GC/GOJ may have symptoms that include indigestion, abdominal discomfort, loss of appetite, nausea, vomiting, anaemia, early satiety, anorexia, difficulty or pain in swallowing, weight loss, bleeding etc. These symptoms are common to many gastrointestinal conditions. As a result, GC/GOJ is often diagnosed in the advanced stages. The overall prognosis of these patients remains poor with a five-year survival rate of approximately 5%.

Incidence of GC/GOJ

Page 15 of the CS states, *"The annual incidence of GC/GOJ is low in the UK at less than 5 per 100,000 persons"*

According to page 37 of the company submission (CS), *"In 2012, there were 5,637 new cases of GC and 3,085 cases of GOJ cancer in England with 4,118 deaths from GC in England in 2013. Of patients diagnosed with GC, approximately 80% are diagnosed with advanced, metastatic GC. For all patients diagnosed with GC between 2010-2011 in England and Wales, the predicted 5-year survival rate was 18.9%."*

"For the approximately 80% of patients with advanced, metastatic disease at diagnosis the 5-year survival rate is much lower at around 5%"

ERG comment: The estimated age standardised incidence rate for the UK of gastric cancer was 7.2%, which translates to 6,684 patients in 2013. Regarding deaths, the age standardised mortality rate is 4.7/100,000 (4,534 deaths). The prevalence of GOJ/GC as identified by the National Oesophago-Gastric Cancer Audit 2013 UK includes 4,835 patients (1,381 GOJ patients and 3,454 patients with stomach cancer).

According to the CS, more than 50% patients can be cured if diagnosed early with a resectable localised distal GC/GOJ adenocarcinoma. Given that the majority of patients are not resectable due to late diagnosis they do not have any curative option and the only available option for them is palliative chemotherapy or best supportive care (BSC) to improve quality of life. Approximately, 657 patients diagnosed with advanced GC/GOJ living in the United Kingdom are estimated to be eligible for second-line treatment following the first-line chemotherapy (based on 2012 data).

Burden of disease

According to page 37 of the CS, *"It is clear however that GC/GOJ has a substantial economic burden as age-specific incidence rates rise significantly from the ages of 60-64 with a 1-year survival rate of 41.8%."*

"In addition to this, age-specific mortality rates rise gradually in GC from the ages of 40-44, with the sharpest increases taking place from ages 65-69. This relates to a loss of approximately 25-30 (working) years in terms of productivity."

ERG comment: Overall, the evidence presented in the CS on this section was in line with the background information given in the final scope and is also consistent with the ERG's understanding of the decision problem.

2.2 Critique of manufacturer's overview of current service provision

Current treatment pathway

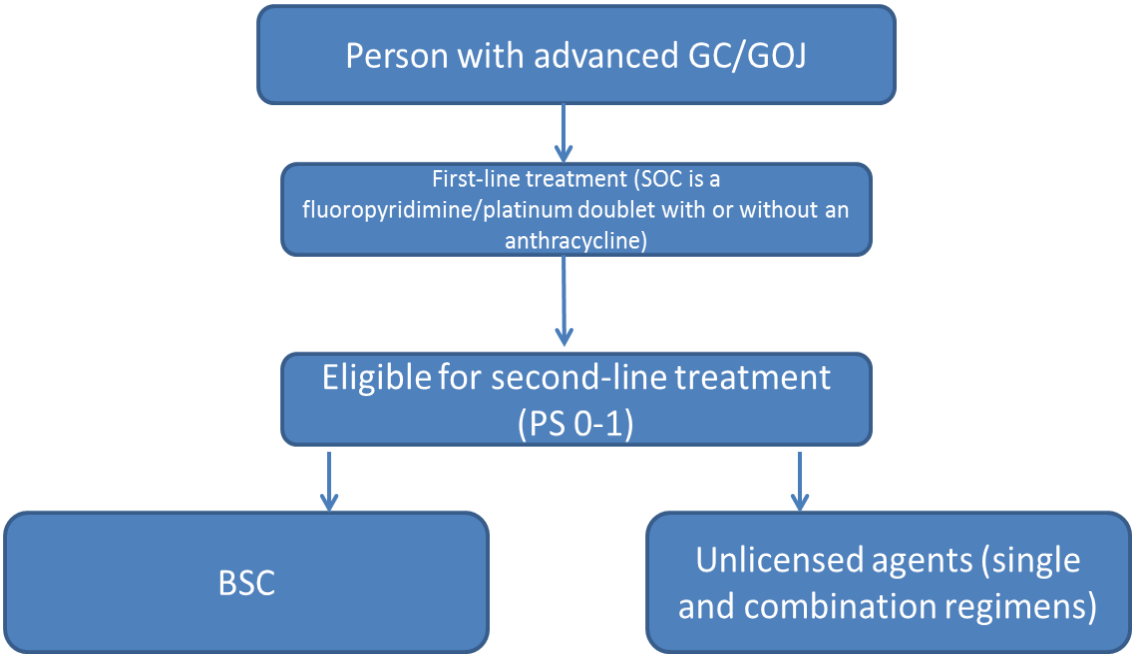
"For patients diagnosed with GC/GOJ in the early stages of the disease, or for those with localised disease, resection is the usual treatment with partial or total gastrectomy. However, only surgically fit patients are eligible for extensive resection."

"Most patients are treated with a fluoropyridimine/platinum doublet with or without an anthracycline as a first-line treatment. However, once patients progress, there is no universally accepted, standard treatment. The most widely reported treatment according to a recent chart review for patients who progress after first-line chemotherapy is BSC."

ERG comment: The current treatment pathway suggests that patients with early diagnosis and/or localised disease can be effectively cured following surgical resection. However, the

majority of undiagnosed patients have advanced GC/GOJ or are inoperable, hence there is a need for treatment to slow disease progression, extend survival and relieve symptoms with minimal adverse effects and without significant detriment to their quality of life. The first line treatment for patients with advanced GC/GOJ or unresectable disease consists of chemotherapy with a fluoropyridimine/platinum doublet with or without an anthracycline. For patients who progress after first-line chemotherapy, best supportive care is the treatment option. Currently, apart from ramucirumab there are no other licensed second-line treatments for patients with advanced GC/GOJ. The current UK clinical practice pathway for patients with GC/GOJ is presented in Figure 2.1.

Figure 2.1: Current treatment pathway based on current NICE recommendations for patients with GC/GOJ



(Source: Eli Lilly Submission, Section 3, P38)

Proposed treatment pathway

Ramucirumab (Cyramza, Eli Lilly and Company) is a fully human IgG1 monoclonal antibody, which acts as a vascular endothelial growth factor receptor-2 (VEGFR-2) angiogenic antagonist. It prevents activating ligands from interacting with VEGFR-2 receptors, which in turn may prevent the formation of new blood vessels, thereby limiting nutrient supply to the tumour causing death of tumour cells.

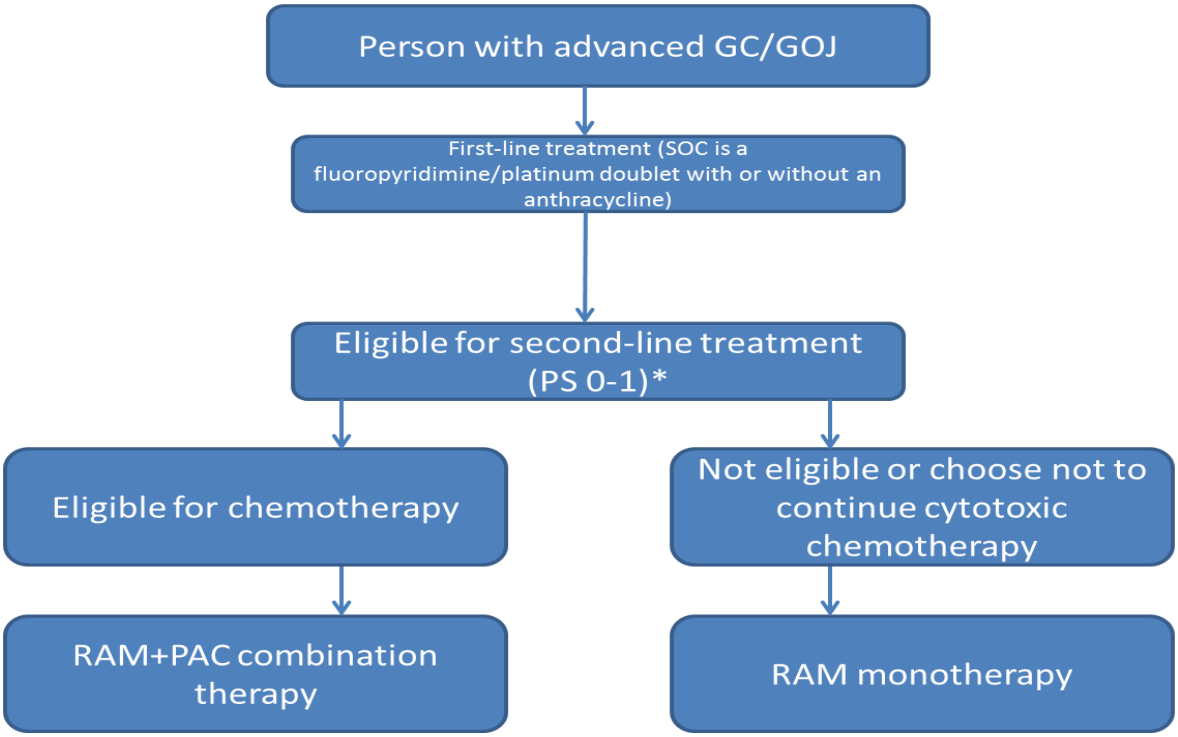
The company submission on page 16 states that, *"Ramucirumab is an orphan drug[2] approved for use for the treatment of adult patients with advanced GC/GOJ with disease progression:*

- *in combination with paclitaxel (RAM+PAC) after prior platinum and fluoropyrimidine chemotherapy and*
- *as a monotherapy (RAM) after prior platinum or fluoropyrimidine chemotherapy, for whom treatment in combination with paclitaxel is not appropriate."*

According to page 43 of the CS, *"The CHMP acknowledged the value of ramucirumab monotherapy even whilst recognising that the OS gain with monotherapy is less than the combination therapy. The CHMP outlined that this could still represent a useful therapeutic option in this second-line setting when chemotherapy in combination with ramucirumab is not the preferred option"*

ERG comment: Ramucirumab received European marketing authorisation on 19 December 2014. Ramucirumab has a marketing authorisation in the UK for use alone or in combination with paclitaxel. The CS claims that the introduction of ramucirumab monotherapy as a second-line therapy will provide clinicians with a treatment choice for a patient group with poor prognosis who are not suitable for cytotoxic chemotherapy and where currently no treatment option other than best supportive care exists. The proposed clinical practice pathway for patients with GC/GOJ is presented in Figure 2.2. Overall, the proposed clinical pathway outlined in the CS is in line with the ERG's understanding of the decision problem and final scope.

Figure 2.2: Proposed treatment pathway for patients with GC/GOJ



(Source: Eli Lilly Submission, Section 3, P42)

3 Critique of manufacturer's definition of decision problem

Table 3.1: Statement of the decision problem (as presented by the manufacturer)

	Final scope issued by NICE	Decision problem addressed in the submission	Rationale if different from the scope
Population	Adults with advanced gastric cancer or gastro-oesophageal junction adenocarcinoma previously treated with chemotherapy	*Ramucirumab in combination with paclitaxel* Adults with advanced gastric cancer or gastro-oesophageal junction adenocarcinoma with disease progression after prior platinum and fluoropyrimidine chemotherapy. *Ramucirumab monotherapy* Adults with advanced gastric cancer or gastro-oesophageal junction adenocarcinoma with disease progression after prior platinum or fluoropyrimidine chemotherapy for whom further cytotoxic chemotherapy is not appropriate.	*Ramucirumab monotherapy* The licence wording for ramucirumab as a monotherapy specifies its use in patients 'for whom treatment in combination with paclitaxel is not appropriate'. Clinically speaking, this patient population can be more broadly characterised as patients for whom further cytotoxic chemotherapy is not appropriate.
Intervention	Ramucirumab alone or in combination with paclitaxel	As per scope	n/a
Comparator	Chemotherapy including but	Ramucirumab in	*Ramucirumab in combination with paclitaxel:*

	Final scope issued by NICE	Decision problem addressed in the submission	Rationale if different from the scope
(s)	not limited to: • Docetaxel monotherapy • Irinotecan monotherapy • Irinotecan and fluorouracil-based therapy (FOLFIRI) • Paclitaxel monotherapy (only for people for whom treatment in combination with paclitaxel is appropriate) Best supportive care (including but not limited to antiemetics, blood transfusions, oesophageal stents, palliative radiotherapy and palliative surgery)	combination with paclitaxel: • Best supportive care • Docetaxel monotherapy • Paclitaxel monotherapy (only for model validation) Ramucirumab monotherapy: • Best supportive care	Prior to ramucirumab, there were no licensed treatments for advanced, previously treated GC/GOJ. The most commonly used treatment for patients who have progressed after first-line treatment with platinum and fluoropyrimidine containing treatment is BSC (71.5%).[3] The most commonly reported active treatment was docetaxel, but this was used in only 8% of patients who progressed after first-line therapy. The remaining patients actively treated received one of 20 different regimens. Only 3% of all patients received paclitaxel. This is only included in the economic model to validate the clinical outputs. Irinotecan and FOLFIRI are not included in the economic analysis as 2% and fewer of all patients were reported to receive these treatments. *Ramucirumab monotherapy:* The licence for RAM specifies its use for patients in whom treatment with PAC is not appropriate. In clinical practice patients who are 'not eligible for paclitaxel' can be more broadly characterised as 'not eligible for cytotoxic chemotherapy'. The standard of care for these patients is BSC only.
Outcomes	The outcome measures to be considered include: • overall survival • progression-free survival • response rate • adverse effects of treatment health-related quality of life	As per scope	n/a
Economic	The reference case stipulates	As per scope	n/a

	Final scope issued by NICE	Decision problem addressed in the submission	Rationale if different from the scope
analysis	that the cost-effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. The reference case stipulates that the time horizon for estimating clinical and cost-effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. Costs will be considered from an NHS and Personal Social Services perspective.		
Subgroups to be considered			A scenario analysis was conducted for Region 1 (North America, Europe, Australia and Israel) due to previously identified differences in treatment patterns and survival outcomes between Asian and non-Asian countries.[4-6] A subgroup analysis was not possible for the monotherapy model as the sample size in the clinical trial was too small.
Special considerations including issues related to equity or equality			Ramucirumab has been designated orphan status by the EMA. Orphan designation is reserved for treatments which provide an effective treatment option where currently none exist. Ramucirumab fulfils this criteria by being the only licensed treatment for previously treated advanced GC/GOJ, addressing a significant unmet need in this treatment setting. At present the only treatment options available to patients are off-licence

	Final scope issued by NICE	Decision problem addressed in the submission	Rationale if different from the scope
			chemotherapy regimens. These treatments are associated with notable toxicities and are therefore not suitable for all patients. For example, older patients are less likely to receive active treatment,[7] an inequality as there are currently no treatment options available to them. In the absence of any licensed treatment or any NICE clinical guidelines or national algorithms in this setting, trusts develop their own treatment protocols meaning access to treatment is contingent upon physician preference and upon any restrictions in place by the trust in which they practice. This leads to notable variation and inequity in provision of treatment for this disease across regions in England. This is clearly reflected in the wide variations in current clinical practice in GC/GOJ.[3] As ramucirumab is the first licensed second-line therapy, it provides the opportunity for equal access to a second-line treatment and reduces the continued use of off-licence drugs, providing a standard treatment pathway where none presently exists.

3.1 Population

The patient population described in the final scope is as follows: "Adults with advanced gastric cancer or gastro-oesophageal junction adenocarcinoma previously treated with chemotherapy".[1] This is in line with the patient population included in the company submission[8] and in the main trials for this submission, the REGARD and RAINBOW studies.[9, 10]

However, according to the marketing authorisation ramucirumab was approved for the following indications:

- ramucirumab in combination with paclitaxel is indicated for the treatment of adult patients with advanced gastric cancer or gastro-oesophageal junction adenocarcinoma with disease progression after prior platinum and fluoropyrimidine chemotherapy.
- ramucirumab monotherapy is indicated for the treatment of adult patients with advanced gastric cancer or gastro-oesophageal junction adenocarcinoma with disease progression after prior platinum or fluoropyrimidine chemotherapy, for whom treatment in combination with paclitaxel is not appropriate.

The REGARD trial (ramucirumab monotherapy) did not specify whether patients were suitable for treatment in combination with paclitaxel. Therefore, patients for whom treatment in combination with paclitaxel is appropriate will have been included in the REGARD trial. In order to assess the relevance of this discrepancy, we asked the company to provide numbers of patients in each arm and separate data of overall survival and progression-free survival from the REGARD trial for patients for whom treatment in combination with paclitaxel is not appropriate. However, the company replied that "It is not possible to estimate the number of patients in each arm of REGARD for whom treatment in combination with paclitaxel is not appropriate."(Response to Clarification Letter, Question A7).[11]

3.2 Intervention

There are two interventions in this STA: ramucirumab alone and ramucirumab in combination with paclitaxel. This is in line with the scope.

Patients in the REGARD trial received ramucirumab 8 mg/kg, intravenously once every two weeks, and patients in the RAINBOW trial received ramucirumab 8 mg/kg intravenously on days 1 and 15, plus paclitaxel 80 mg/m2 intravenously on days 1, 8, and 15 of a 28-day cycle.

3.3 Comparators

The comparators in the scope are defined as follows:

- Chemotherapy including but not limited to:
  - Docetaxel monotherapy
  - Irinotecan monotherapy
  - Irinotecan and fluorouracil-based therapy (FOLFIRI)
  - Paclitaxel monotherapy (only for people for whom treatment in combination with paclitaxel is appropriate)
- Best supportive care (including but not limited to antiemetics, blood transfusions, oesophageal stents, palliative radiotherapy and palliative surgery)

For ramucirumab monotherapy the company used placebo, the comparator in the REGARD trial as the comparator. In the REGARD trial patients in both arms received best supportive care, excluding other investigational antitumour drugs or antineoplastic chemotherapy, hormonal treatment, or immunotherapy. This seems reasonable if the committee accepts the reasoning from the company that although the "license wording for ramucirumab as a monotherapy specifies its use in patients 'for whom treatment in combination with paclitaxel is not appropriate'. Clinically speaking, this patient population can be more broadly characterised as patients for whom further cytotoxic chemotherapy is not appropriate."

For ramucirumab in combination with paclitaxel, the comparator in the RAINBOW trial was paclitaxel monotherapy. In addition, the company presents results for comparisons with best supportive care and docetaxel using indirect comparisons (Bucher method) and a network meta-analysis. Results from the network meta-analysis are used in the economic analyses.

Comparisons with irinotecan monotherapy and FOLFIRI are missing. It should be noted that the comparison with irinotecan monotherapy is missing, which is odd because a comparison of ramucirumab in combination with paclitaxel versus best supportive care or docetaxel is only possible in a network that includes irinotecan monotherapy (see CS: Figure 26, page 113). We asked the company to provide data for the comparisons with irinotecan monotherapy and FOLFIRI.

3.4 Outcomes

In the NICE final scope, outcomes are defined as follows:

- overall survival
- progression-free survival
- response rate
- adverse effects of treatment
- health-related quality of life.

These outcomes are all discussed in the company submission for the REGARD and RAINBOW trials. For indirect comparisons, only overall survival, progression free survival and withdrawals due to adverse events are reported. In the network meta-analysis results for the following adverse events are also reported: grade 3/4 anaemia, grade 3 or 4 neutropenia, and grade 3 or 4 febrile neutropenia.

3.5 Other relevant factors

Given the high ICERs for monotherapy and combination therapy we checked with NICE, and NICE confirmed that there is no Patient Access Scheme.

Regarding equity considerations, the company states that ramucirumab has been designated orphan status by the EMA.

According to the company, ramucirumab combination therapy fulfils the end-of-life criteria as specified by NICE. The company does not make this claim for ramucirumab monotherapy.

4 CLINICAL EFFECTIVENESS

4.1 Critique of the methods of review(s)

4.1.1 Searches

The Canadian Agency for Drugs and Technologies in Health (CADTH) evidence based checklist for the Peer Review of Electronic Search Strategies, was used to inform this critique.[12] The submission was checked against the Single Technology Appraisal (STA) specification for company/sponsor submission of evidence.[13] The ERG has presented only the major limitations of each search strategy in the main report. Further criticisms of each search strategy can be found in Appendix 1.

Clinical effectiveness

On Page 47 of the CS it states that a comprehensive review of the published literature was performed. Searches were reported for MEDLINE, MEDLINE In-Process, Embase and the Cochrane Library. These meet the requirements specified in current best practice guidance as detailed in the NICE guide to the methods of technology appraisal [14]. The company additionally searched the American Society of Clinical Oncology (ASCO) website for relevant abstracts, however no details of the conference search terms or date the searches took place were provided.

The database hosts for each database, date span and search date were listed. The searches were on the whole well reported and reproducible. The database searches were clearly structured and used combinations of index terms and free text. The search date provided in the submission text (CS, p35) was inconsistent with the date of search provided in Appendix 2. Following clarification, the company confirmed that the correct date was that provided in Appendix 2 (given as February 2015 in company response to clarification). The specific date in February 2015 that the searches were run was not provided, but from the database spans provided it can be assumed the searches were run in the second week of February.

The ERG has some concerns that the searches were limited to English language only. Current best practice states that 'Whenever possible review authors should attempt to identify and assess for eligibility all possibly relevant reports of trials irrespective of language of publication'.[15] In their response to clarification the company stated 'The restriction to English language studies only was for practical reasons and is commonly applied in searches', however the ERG is concerned that restricting to English language has the potential for relevant material to be missed.

The ERG queried the number of results retrieved by the searches reported in the PRISMA flowchart as this did not match the results given in the search strategy. In their response the company provided an updated PRISMA flowchart as the submission contained a typing error.

Indirect and network meta-analysis

A systematic literature review was conducted to identify RCT evidence for the efficacy and safety of RAM+PAC and other treatments in the second line treatment of adult patients with GC/GOJ adenocarcinoma. Searches were reported for MEDLINE, MEDLINE In-Process, Embase and the Cochrane Library. These meet the requirements specified in current best practice guidance as specified in NICE guide to the methods of technology appraisal.[14] The company additionally searched specific conference proceedings for eligible abstracts/posters,

however, no specific details of the search terms or date the conference searches were conducted were provided.

The database hosts for each database, date span and search date were listed. The database searches were clearly structured used a combination of index terms and free text and a study design filter was used to restrict results to RCTs although this was not referenced as an objectively derived or published filter.

The ERG queried the search date of December 2013 reported for the searches in the submission, more recent papers published in 2014 had been included in the submission analysis and these would not have been picked up in a search conducted in December 2013. Following clarification the company responded that the literature search was updated in May 2014. The updated search strategies and associated reporting information were not provided at clarification, therefore the ERG is unable to comment on these update searches, only on those reported for searches conducted on December 2013.

The searches were again limited to English language, so the same concerns regarding the use of language restriction apply here.

Non-randomised and non-controlled evidence

No non-RCT evidence was presented in the submission.

Adverse events

No specific AE searches were performed; the company states that AE data was taken from the RAINBOW and REGARD trials identified in the clinical effectiveness searches and from the NMA.

Cost-effectiveness

A literature review was conducted to identify any existing economic analyses in this area in order to address the decision problem and inform the economic model. Searches were reported for MEDLINE, MEDLINE In-Process, Embase and NHS EED. Additional searches of the NICE website, and the Centre for Reviews and Dissemination (CRD) database were also undertaken at an earlier date of September 2013. The database searches were limited from 2009-2014 (the earlier CRD search was limited to 2006 to 2013).

The reporting of hosts, date spans and specific dates searches were run was not as clear as for previous reported searches in the submission. More details are provided in Appendix 1 of this report.

The database searches were clearly structured and used a combination of index terms and free text and a good range of economics terms were included. A facet of economics terms were included in the NHS EED search, as this is an economics database the ERG believes it is not necessary to include this facet in the strategy for this database.

Measurement and value of health effects

A systematic review of HRQL studies was conducted for GC, GOJ or oesophageal cancer. Searches were reported for MEDLINE, MEDLINE In-Process, Embase, EconLit and NHS EED. Relevant submission/appraisal data from NICE, Scottish Medicines Consortium (SMC) and Pharmaceuticals Benefit Advisory Committee (PBAC) were also reviewed. Additional

conference searches, searches of the CEA Registry, Euro-Qol 3 dimensions (EQ-5D) and Research Papers in Economics (RePEc) websites were also conducted. The host, date span and search dates were reported for all resources, the data span for Embase was clarified with the company as there was a typographical error. The searches were well reported and reproducible.

The ERG feels that the search date of September 2013 is outdated; an update search could have potentially retrieved additional relevant results.

Cost and healthcare resource use identification, measurement and valuation

The same search methods were employed as in the cost-effectiveness searches and therefore the same limitations apply here.

4.1.2 Inclusion criteria

According to the response to the clarification letter, the inclusion criteria for the systematic review were as follows (see Table 4.1)

Table 4.1: Eligibility criteria used in search strategy for systematic review

	Description
Inclusion criteria	
Population	Adult patients (≥18 years) who received prior chemotherapy (no restriction on line of therapy) for the following: • Gastric cancer • Gastro-esophageal junction (GEJ) cancer
Interventions	Eligible studies included at least one treatment arm examining one of the following therapies: **Chemotherapy:** Capecitabine, Carboplatin, Cisplatin, Docetaxel, Epirubicin, Etoposide, Irinotecan, Mitomycin, Oxaliplatin, Paclitaxel, CAPOX, EOX, FLOT, FOLFIRI, mFOLFIRI, 5-FU, S-1, Tegafur **Targeted therapy:** Apatinib**,** Bevacizumab**,** Everolimus**,** Lapatinib, Ramucirumab, Trastuzumab
Comparators	Comparators to include any of the above interventions, placebo, BSC, or active symptom control (ASC).
Outcomes	Eligible studies reported at least one of the following outcomes listed below. The primary focus of the review was a measure of survival. However studies which report response rates only were flagged. • Overall Survival (OS) • Progression-free survival (PFS) with definition (N.B. RECIST 2009 guidelines change the definition of disease progression and therefore of PFS). • 12-week PFS rate • Time to progression (TTP) with definition • Objective response rate (ORR) (complete response (CR), partial response (PR), overall response (=CR+PR)) as measured by WHO or RECIST or revised v1.1 RECIST guidelines • Stable disease (SD) • Clinical benefit or disease control rate (DCR) (=CR+PR+SD) • Any of the above with respect to prognostic factor subgroups
Study design	Included: • Clinical (human) Phase II or III RCTs • Superiority or non-inferiority RCTs

	Description
	• Blinded or open label RCTs
Language restrictions	English language only

ERG comment: The systematic review was done to allow an indirect comparison of ramucirumab with or without paclitaxel versus the comparators mentioned in the NICE scope: docetaxel monotherapy, irinotecan monotherapy, irinotecan and fluorouracil-based therapy (FOLFIRI), and paclitaxel monotherapy. The inclusion criteria are appropriate for a systematic review in line with the scope of this appraisal. The review was limited to English language studies only.

4.1.3 Critique of data extraction

For trials looking directly at ramucirumab, two RCT's were included: the RAINBOW trial and the REGARD trial. Data related to these trials was extracted from the clinical study report, sponsor submission and journal publications (see Table 4.2).

Table 4.2: Data sources for RCT's of ramucirumab: RAINBOW and REGARD

Trial	Trial start and finish dates	Cut-off date for the analysis	Data availability	Data source
RAINBOW	Dec 2010 - July 2013	12 July 2013	Unpublished Published	Clinical study report Sponsor submission Journal articles: Wilke H, et al. 2014[9] Conference abstracts (ASCO, ESMO)[16-22]
REGARD	6 Oct 2009 - 26 Jan 2012	25 July 2012	Unpublished Published	Clinical study report Sponsor submission Journal article: Fuchs CS, et al 2014[10] Conference abstracts (AACR, ASCO, ESMO)[23-25]

Abbreviations: ECC, European Cancer Congress; AACR, American Association for Cancer Research Annual Meeting; ASCO, American Society of Clinical Oncology; ESMO, European Society for Medical Oncology

For studies retrieved through the systematic review, data extraction of 23 trials (30 publications) was based on primary study publications. These were full publications for 15 trials, posters for six trials, and abstracts only for two trials. Full data extraction of these studies is reported in Appendix 1 of the Response to the Clarification Letter.

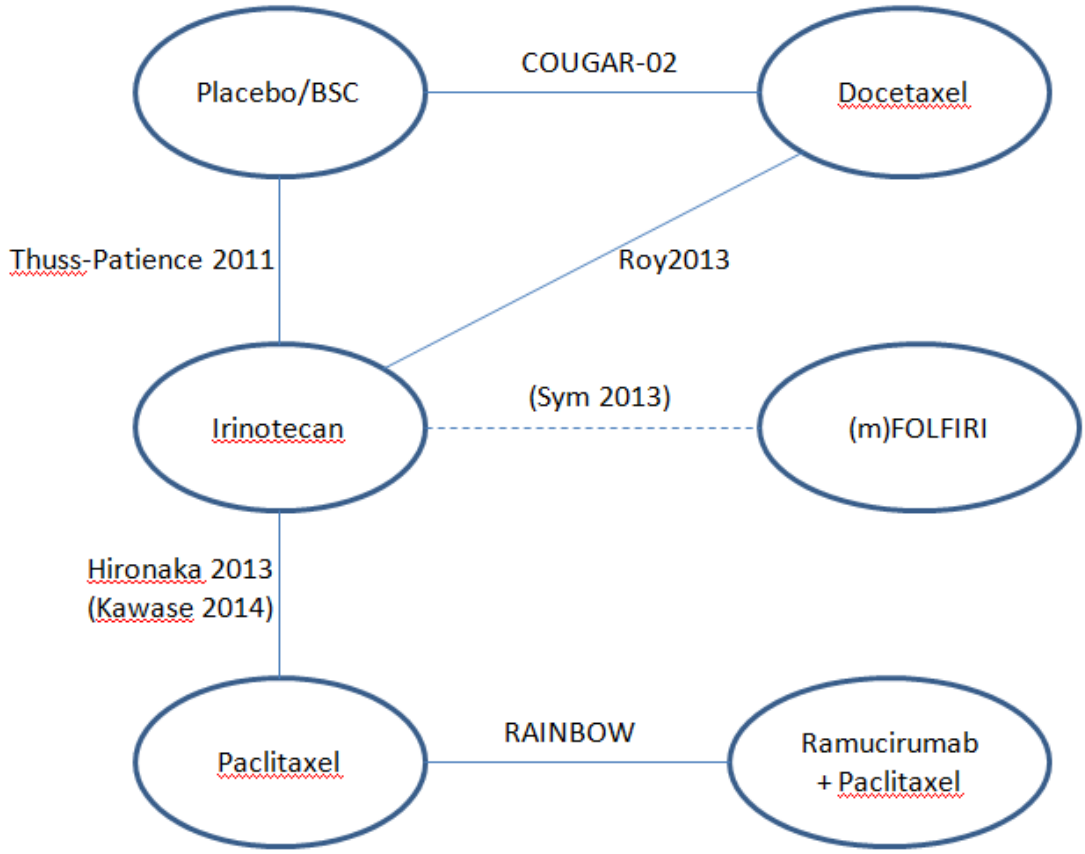
ERG comment: The full network included 23 trials. However, because the company considered only a limited number of comparators relevant for this appraisal, most trials were not used in the analyses.

For monotherapy, only the REGARD trial was considered relevant, as this compares ramucirumab with best supportive care.

For combination therapy, the company only considered paclitaxel, docetaxel and best supportive care to be relevant comparators. Therefore, only five trials were used in the analyses (see Figure 4.1). Although it was necessary to include irinotecan in the evidence network, results were not reported because the company considered irinotecan not relevant; and, although the comparison with (m)FOLFIRI was possible using Sym 2013,[26] this was also not reported because the company considered (m)FOLFIRI also not relevant. The reason why Kawase 2014 was not included in the network meta-analysis is not reported in the CS. We assume it is excluded because it is only reported as an abstract or poster and no relevant data have been reported. However, we were not able to check this as the publication was not included in the submission, nor was a reference provided.

The REGARD and RAINBOW trials are discussed in great detail, with extensive data extraction reported in the CS. Data extraction for all comparator trials (except Kawase 2014) are based on published journal articles and reported in Appendix 6 of the CS and in the Response to the Clarification Letter.

Figure 4.1: Evidence network used in CS analyses



4.1.4 Quality assessment

The risk of bias assessments for the two ramucirumab trials (REGARD and RAINBOW), as well as the trials included in the indirect comparisons and network meta-analyses are described in the tables below.

Table 4.3: Summary of the Risk of Bias assessment for the ramucirumab studies

	RAINBOW	ERG comment	REGARD	ERG comment
Was randomisation carried out appropriately?	Yes	Low risk of bias	Yes	Low risk of bias
Was the concealment of treatment allocation adequate?	Yes	Low risk of bias	Yes	Low risk of bias
Were the groups similar at the outset of the study in terms of prognostic factors?	No	Moderate risk of bias	No	Moderate risk of bias
	Imbalance with respect to ECOG performance status, number of metastatic sites and presence of ascites.		Imbalance with respect to histological subtype, % peritoneal metastases, number of metastatic sites and previous anticancer treatment	
Were the care providers, participants and outcome assessors blind to treatment allocation?	Yes	Low risk of bias	Yes	Low risk of bias
Were there any unexpected imbalances in drop-outs between groups?	No	Low risk of bias	No	Low risk of bias
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No	Low risk of bias	No	Low risk of bias
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	Low risk of bias	Yes, Yes	Low risk of bias

Both the REGARD trial and the RAINBOW trial were considered at low risk of bias (See Table 4.3). The only item which was rated as moderate risk of bias was the similarity between groups in both the RAINBOW and REGARD trials. In the RAINBOW trial there was an imbalance with respect to ECOG performance status, number of metastatic sites and presence of ascites. Regarding ECOG performance status, there was a greater proportion of patients with ECOG-1 performance status in the ramucirumab plus paclitaxel arm (65%) compared with the placebo plus paclitaxel arm (57%). Regarding the number of metastatic sites >=3, there was a greater proportion of patients in the ramucirumab plus paclitaxel arm (37%) compared with the placebo plus paclitaxel arm (31%). Regarding the presence of ascites, there were more patients with ascites in the ramucirumab plus paclitaxel arm (39%) compared with the placebo plus paclitaxel arm (32%). The direction of the imbalances in baseline characteristics in the RAINBOW trial was in favour of a worse outcome for patients receiving ramucirumab plus paclitaxel. ECOG performance status, region, and presence of ascites were the main contributors to the difference between adjusted and unadjusted efficacy estimates (Lancet supplementary files).[9]

In the REGARD trial there was an imbalance with respect to histological subtype, percentage peritoneal metastases, number of metastatic sites and previous anticancer treatment. Regarding histological subtype, the ramucirumab arm included a lower proportion of patients with intestinal type than the placebo arm. Intestinal histology/tumour types is associated with

better prognosis.[4] Regarding peritoneal metastases, the ramucirumab arm included a lower proportion of patients with peritoneal metastases (27%) compared with the placebo arm (38%). Regarding the number of metastatic sites, the ramucirumab arm included a lower proportion of patients with >=3 metastatic sites (32%) compared with the placebo arm (39%) Regarding previous anticancer treatment, the ramucirumab arm included a greater proportion of participants whose previous anticancer treatment had been fluropyrimidine plus platinum (84%) compared with the placebo arm (75%). However, the placebo arm included a greater proportion of patients whose previous anticancer treatment had been fluoropyrimidine plus other systemic drug compared with the ramucirumab arm. The majority of the imbalances in baseline characteristics in the REGARD trial favoured a worse outcome for patients in the placebo arm.

Table 4.4: Summary of the Risk of Bias assessment for the docetaxel studies included in the evidence network

	COUGAR-02[27]	**ERG comment**	**Roy 2013[28]**	**ERG comment**
Was randomisation carried out appropriately?	Yes	Low risk of bias	Yes	Low risk of bias
Was the concealment of treatment allocation adequate?	No, Open-label study	Moderate risk of bias	No, Open-label study	Moderate risk of bias
Were the groups similar at the outset of the study in terms of prognostic factors?	No	Moderate risk of bias	No	Moderate risk of bias
	Slight imbalance site of disease and older than 70 years		Imbalance re previous radiotherapy/surgery and tumour site	
Were the care providers, participants and outcome assessors blind to treatment allocation?	No, Open-label study	Moderate risk of bias	No, Open-label study	Moderate risk of bias
Were there any unexpected imbalances in drop-outs between groups?	No	Low risk of bias	unclear	Unclear risk of bias
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No	Low risk of bias	No	Low risk of bias
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	Low risk of bias	Yes, unclear	Unclear risk of bias

Both docetaxel studies were considered at moderate risk of bias, mainly because both trials were open-label studies (See Table 4.4). In addition, the numbers of drop-outs in the study by Roy et al (2013)[28] was unclear.

The three irinotecan trials were considered at high risk of bias (See Table 4.5). Two of these trials were open label trials and the randomisation process was unclear in the third. In addition, all three trials suffered from imbalances in the groups at the outset of the study in terms of prognostic factors.

Table 4.5: Summary of the Risk of Bias assessment for the irinotecan studies included in the evidence network

	Hironaka 2013[29]	**ERG comment**	**Thuss-patience 2011[30]**	**ERG comment**	**Sym 2013[26]**	**ERG comment**
Was randomisation carried out appropriately?	Yes	Low risk of bias	No, Centrally in blocks using a coin	High risk of bias	Unclear	Unclear risk of bias
Was the concealment of treatment allocation adequate?	No, Open-label study	Moderate risk of bias	No, Open-label study	Moderate risk of bias	Unclear	Unclear risk of bias
Were the groups similar at the outset of the study in terms of prognostic factors?	No	High risk of bias	No	High risk of bias	No	High risk of bias
	Imbalance re prior treatments		Strong Imbalance with respect to gender		Imbalance re gender and prior anticancer therapy	
Were the care providers, participants and outcome assessors blind to treatment allocation?	No, Open-label study	Moderate risk of bias	No, Open-label study	Moderate risk of bias	Unclear	Unclear risk of bias
Were there any unexpected imbalances in drop-outs between groups?	No	Low risk of bias	No	High risk of bias	No	Low risk of bias
			Study was stopped prematurely			
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No	Low risk of bias	No	Low risk of bias	No	Low risk of bias
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, unclear	Unclear risk of bias	Yes, unclear	Unclear risk of bias	Yes, Unclear	Unclear risk of bias

4.1.5 Evidence synthesis

A systematic review was done to allow an indirect comparison of ramucirumab with or without paclitaxel versus the comparators mentioned in the NICE scope: docetaxel monotherapy, irinotecan monotherapy, irinotecan and fluorouracil-based therapy (FOLFIRI), and paclitaxel monotherapy. Two types of syntheses are reported: in Chapter 4.10 of the company submission a series of indirect comparisons apparently using the Bucher method[31] are described; however, the methods for this are not reported. In Appendix 6 of the CS, the results of a full network meta-analysis (NMA), using WinBUGS, are reported.

The study selection and flowchart for the evidence synthesis is described in Chapter 4.10 of the company submission. However, it was clear that there were discrepancies between the flowchart of included studies (Figure 24, page 109 CS), the network of included studies (Figure 25, page 111 CS) and the data extraction tables for the included studies (Appendix 6 CS). In response to the clarification letter the company provided a corrected flowchart and a list of 23 trials (30 publications) included in the indirect comparisons and network meta-analysis.

The results for the indirect comparisons (Chapter 4.10 CS) and network meta-analysis (Appendix 6, CS) for OS and PFS are almost identical, and results of the NMA are used in the economic model. Therefore, we will discuss results of the NMA only.

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

For trials looking directly at ramucirumab, two RCT's were included: the RAINBOW trial and the REGARD trial. In this section, we will present the results from the RAINBOW and REGARD trials. The study characteristics are presented in Table 4.6, and a summary of the methodology in Table 4.7.

Table 4.6: Overview of the RAINBOW and REGARD trials

Trial (acronym)	Intervention	Comparator	Population
RAINBOW[9]	Ramucirumab 8 mg/kg + paclitaxel 80 mg/m^2 administered IV. RAM+PAC	Placebo + paclitaxel 80 mg/m^2 administered IV. PBO+PAC	Patients aged 18 years and older; having metastatic or nonresectable, locally advanced GC/GOJ; documented objective radiological or clinical disease progression during or within 4 months of the last dose of first-line platinum and fluoropyridimine doublet with or without anthracycline; an ECOG PS of 0 or 1; and measurable or non-measurable evaluable disease.
REGARD[10]	Ramucirumab administered IV at a dose of 8 mg/kg and best supportive care RAM+BSC	Equivalent volume of placebo administered IV and best supportive care PBO+BSC	Patients aged 18 years and older, had metastatic or unresectable, locally recurrent GC/GOJ; had experienced disease progression during or within 4 months after the last dose of first-line therapy for metastatic disease, or during or within 6 months after the last dose of adjuvant therapy; and had a ECOG PS of 0 or 1; and

Trial (acronym)	Intervention	Comparator	Population
			measurable or evaluable disease.

Source: CS, Table 12, Page 49
Abbreviations: RAM+ PAC, Ramucirumab + paclitaxel; GC, Gastric cancer; GOJ, gastroesophageal junction cancer; ECOG PS, Eastern Cooperative Oncology Group Performance Status.

ERG comment: In addition to the study drugs, both groups also received standard care. The company acknowledged that "Due to differences in standard care, geographic region was considered an important prognostic factor".

Trial methodology

The methodology of the RAINBOW and REGARD trials is outlined in Table 4.7.

ERG comment: In the clarification letter, the ERG sought clarification on the proportion of UK patients in the RAINBOW and REGARD trial. The sponsor advised that the RAINBOW trial (ITT population) included 15 UK patients (six intervention, nine comparator = 2.25%) and the REGARD trial included 17 UK patients (13 intervention, four comparator = 4.78%).

The trials used different stratification factors and the treatment cycles were different (RAINBOW 28 days, REGARD two weeks).

Eligibility criteria

Important inclusion/exclusion criteria for the RAINBOW and REGARD trials are presented in Table 4.8.

ERG comment: Eligibility criteria are broadly similar for the two trials.

Superseded – see Erratum

Table 4.7: Summary of methodology of the RAINBOW and REGARD trials

Trial title	RAINBOW	REGARD
Location	170 centres across 27 countries in North and South America, Europe, Asia, and Australia. (2.25% of patients were from the UK)	119 centres across 29 countries in North America, Central and South America, Europe, Asia, Australia, and Africa. (4.78% of patients were from the UK)
Design	Phase III, randomised, double-blind, placebo-controlled study of RAM +PAC versus PAC (1:1)	Phase III, randomised, double-blind, placebo-controlled study of RAM +PAC versus PAC (2:1)
Patient population	Patients with histologically or cytologically confirmed unresectable or metastatic gastric or GEJ carcinoma who had received at least 1 cycle of combination therapy with any platinum and any fluoropyrimidine as first-line treatment with/without anthracyclines	Patients with metastatic and locally advanced gastric cancer (including adenocarcinomas of the GEJ) and radiographic evidence of disease progression on prior first-line chemotherapeutic regimens.
Duration of study	2 years	2.3 years
Method of randomisation	Eligible patients were randomised (1:1) using a centralised IVRS/IWRS system and were stratified by time to progression from the start of first-line chemotherapy (<6 months vs. ≥6 months), disease measurability (measurable vs. nonmeasurable disease) and geographic region. • Region 1 (Europe, Israel, US, and Australia) • Region 3 (Hong Kong, Japan, South Korea, Singapore, Taiwan) • Region 2: (Argentina, Brazil, Chile, and Mexico).	Eligible patients were randomised (2:1) using a centralised IVRS/IWRS system to ramucirumab + BSC or placebo + BSC treatment, respectively. Randomisation was stratified by weight loss (≥10% over the prior 3 months vs. <10% over the prior 3 months), geographic region • (Region 1, 2 and 3) and location of the primary tumor (Gastric [including tumors of the gastric cardia that extend into the GEJ] vs. GEJ [including tumors of the distal esophagus that extend into the GEJ, and tumors involving the GEJ when precise identification of the organ of origin was not possible]).
Method of blinding (care provider, patient and outcome assessor)	Double-blinded: patients, investigators, and all other personnel involved in the conduct of the study were blinded to individual treatment assignments for the duration of the study. Ramucirumab and placebo for infusion were identical in appearance and there were no anticipated or identified toxicity of ramucirumab that would potentially unblind investigators to treatment assignment	Double-blinded: patients, investigators, sponsor and all other personnel involved in the conduct of the study were blinded to individual treatment assignments for the duration of the study. Ramucirumab and placebo for injection were identical in appearance. The study drug (ramucirumab or placebo) was uniquely labelled, and assigned to a patient by using IVRS/IWRS. Unblinding of the study team did not occur until the reporting database was validated and locked for final statistical analysis on 26 September 2012.
Intervention(s) (n =) and comparator(s) (n =)	1) Ramucirumab 8 mg/kg plus paclitaxel 80 mg/m2 administered intravenously (IV). Paclitaxel was given on Days 1, 8, and 15 of a 28-day cycle, in combination with ramucirumab given on Days	1) BSC plus ramucirumab administered intravenously (IV) every 2 weeks at a dose of 8 mg/kg (n=238) 2) BSC plus an equivalent volume of placebo administered IV

Trial title	RAINBOW	REGARD
	1 and 15. (N=330) 2) Placebo plus paclitaxel 80 mg/m2 administered IV. Paclitaxel was given on Days 1, 8, and 15 of a 28-day cycle, in combination with an equivalent volume of ramucirumab placebo (placebo) given on Days 1 and 15. (N=335) Each treatment cycle was 28 days in length.	every 2 weeks (n=117) Each treatment cycle was two weeks in length
Primary outcomes	The primary outcome was overall survival (OS), defined as the time from the date of randomisation to the date of death from any cause. OS was censored on the last date the patient was known to be alive (on or before data cut-off date or lost to follow-up). Patient survival status was collected every 8 weeks after treatment discontinuation, until the data cut-off date.	The primary efficacy variable was overall survival (OS), defined as the time from randomisation to the date of death from any cause.
Secondary outcomes	Progression-Free Survival, Time to Progression, Objective response rate (ORR) defined as the number of randomised patients who achieved complete response (CR) or partial response (PR) divided by the ITT population. Quality of Life (QoL): Assessed using EORTC QLQ-C30 (version 3.0), and the European Quality of Life-5 Dimensions (EQ-5D). Time to deterioration (TTD) in EORTC QLQ-C30. Time to deterioration (TTD) in ECOG PS and Safety (NCI CTCAE, version 4.02)	Progression-Free Survival, Investigator-assessed Objective response rate (ORR), Duration of Response (DCR), QoL (EORTC QLQ-C30 (version 3.0)), Time to deterioration of ECOG PS and Safety (NCI CTCAE, version 4.02)
Other endpoints	Pharmacodynamics, immunogenicity and pharmacokinetic parameters	

Abbreviations: IDMC, Independent Data Monitoring Committee; ITT, intent-to-treat; IVRS, Interactive Voice Response System; IWRS, Interactive Web Response System; N/A, not applicable; PS, performance status. AE, Adverse Events; BSC, Best supportive care; CI, Confidence Interval; CR, Complete response; TTP, Time to disease progression; RECIST, Response Evaluation Criteria in Solid Tumours; RCT, Randomised Controlled Trial; RR, Relative Risk; PR, Partial response; OS, Overall survival; OR, Odds Ratio; ORR, Objective response rate; ECOG PS, Eastern Cooperative Oncology Group Performance Status; EQ-5D, European Quality of Life-5 Dimensions; EORTC QLQ-C3,0 European Organization for Research and Treatment of Cancer QLQ-C30; TTD, Time to deterioration; HRQL, Health-related Quality of Life; NCI CTCAE, National Cancer Institute Common Terminology Criteria for Adverse Events; QOL, Quality of Life; DCR, Duration of Response.

Table 4.8: Major eligibility criteria of the RAINBOW and REGARD trials

RAINBOW		**REGARD**	
Inclusion	**Exclusion**	**Inclusion**	**Exclusion**
• Histologically or cytologically confirmed gastric carcinoma, including gastric or GOJ adenocarcinoma • Metastatic disease or locally recurrent, unresectable disease • Disease progression during or ≤4 months after the last dose of first-line platinum/fluoropyrimidine doublet with or without anthracycline (epirubicin, doxorubicin) for unresectable or metastatic disease • Age ≥18 years and the ECOG PS score of 0 or 1 • Adequate recovery from toxicities or effects of prior therapy • Adequate haematologic, hepatic, coagulation, and renal function	• Patients with squamous cell or undifferentiated gastric cancer. • Patients had undergone major surgery within 28-days prior to randomisation, or central venous access device placement within 7 days prior to randomisation • Patients received any chemotherapy other than platinum and fluoropyrimidine with or without anthracycline for advanced gastric or GOJ adenocarcinoma • History of GI perforation and/or fistulae within 6 months prior to randomisation. • Significant bleeding disorders, vasculitis, or had a significant bleeding episode from the GI tract within 3m prior to study entry • History of deep vein thrombosis, pulmonary embolism, or any other significant thromboembolism during the 3 months prior to randomisation. Symptomatic congestive heart failure • Uncontrolled arterial hypertension • Arterial thrombotic event (MI, transient ischemic attack, cerebrovascular accident, or unstable angina) within 6 months prior to randomisation • Serious or non-healing wound or peptic ulcer or bone fracture within 28-days prior to randomisation. • Current chronic platelet therapy, NSAIDs.	• Histologically or cytologically confirmed gastric carcinoma, including gastric or GOJ adenocarcinoma • Metastatic disease or locally recurrent, unresectable disease • Measurable disease (defined as at least one unidimensionally-measurable target lesion according to RECIST) and/or evaluable disease. • The patient has experienced disease progression during or within 4 months after the last dose of first-line therapy for metastatic disease, or during or within 6 months after the last dose of adjuvant therapy • Age ≥18 years, life expectancy of ≥12 weeks, and ECOG PS score of 0 or 1 • Adequate recovery from toxicities or effects of prior therapy	• Documented or symptomatic brain or leptomeningeal metastases • Any grade 3/4 gastrointestinal bleeding within 3 months prior to randomisation • Any arterial thromboembolic event (myocardial infarction, transient ischemic attack, cerebrovascular accident, unstable angina, or other within 6 months prior to randomisation) • Ongoing or active infection, symptomatic congestive heart failure, unstable angina pectoris, symptomatic or poorly controlled cardiac arrhythmia, uncontrolled throm-botic or haemorrhagic disorder, any other serious uncontrolled medical disorder • Uncontrolled or poorly controlled hypertension • Serious or unhealing wound, ulcer, or bone fracture within 28-days prior to randomisation • Current chronic antiplatelet therapy, including aspirin, NSAIDs, dipyridamole or clopidogrel, or similar agents. Once daily use of aspirin was permitted. • Known allergy to any treatment component • Pregnancy/breastfeeding • Concurrent active malignancy other than adequately-treated non-melanomatous skin cancer, other non-invasive carcinoma, or *in situ* neoplasm

	• Pregnancy/breastfeeding	• Adequate haematologic, hepatic, coagulation, and renal function	

Source: CS, Table 13, Page 52 & Table 30, Page 79

Treatment exposure and discontinuation

Patients continued treatment with ramucirumab plus paclitaxel or placebo plus paclitaxel (RAINBOW) and ramucirumab or placebo (REGARD) until disease progression (determined according to radiographic and clinical measures). The median duration of therapy in the RAINBOW trial was 18.9 weeks for the ramucirumab plus paclitaxel arm (with a median of five cycles received) compared with 12.1 weeks for the placebo plus paclitaxel arm (median of three cycles received). In the REGARD trial, median duration of therapy was eight weeks for the ramucirumab arm (with a median of four cycles received) and six weeks for the placebo arm (with a median of three cycles received).

In the RAINBOW trial, dose reductions of ramucirumab occurred in 16 (5%) of 327 patients in the ramucirumab plus paclitaxel group, and of placebo in three (<1%) of 329 patients in the placebo plus paclitaxel group. Paclitaxel dose reductions occurred in 78 (24%) patients in the ramucirumab plus paclitaxel group, and in 24 (7%) patients in the placebo plus paclitaxel group. The most common reason for treatment discontinuation was disease progression: 71.5% for the ramucirumab plus paclitaxel arm, 76.1% for the placebo plus paclitaxel arm. The proportion of patients who discontinued treatment due to an AE was similar between the two trial arms (11.8% for ramucirumab plus paclitaxel, 11.3% for placebo plus paclitaxel).

In the REGARD trial, progressive disease was the primary reason for treatment discontinuation (52.9% for ramucirumab, 62.4% for placebo), followed by symptomatic deterioration (17.2% for ramucirumab, 13.7% for placebo). A greater proportion of patients receiving the placebo discontinued treatment. Discontinuation following an adverse event: In the ramucirumab arm, 11.3% (25) of patients discontinued treatment due to AEs, 13 out of 25 discontinuations were considered by the investigator to be at least possibly related to study therapy.

Reasons for treatment discontinuation of the ITT populations in the RAINBOW trial (data cut-off date 12/07/ 2013) and in the REGARD trial (Data cut-off: 25 July 2012) are presented in Table 4.9.

Table 4.9: Reasons for treatment discontinuation

	RAINBOW		REGARD	
	Ramucirumab + Paclitaxel (N=330)	**Placebo + Paclitaxel (N=335)**	**Ramucirumab (N=238)**	**Placebo (N=117)**
Patients treated, n (%)	326 (98.8)	330 (98.5)	14 (5.9)	1 (0.9)
Never Treated	4 (1.2)	5 (1.5)	2 (0.8)	2 (1.7)
Treatment discontinued	313 (94.8)	323 (96.4)	222 (93.3)	114(97.4)
Treatment ongoing			16	3
Reasons for discontinuation, n (%)				
Progressive Disease	236 (71.5)	255 (76.1)	126(52.9)	73 (62.4)
Symptomatic deterioration			41 (17.2)	16 (13.7)
Death	12 (3.6)	13 (3.9)	20 (8.4)	13 (11.1)
AE	39 (11.8)	39 (11.8)	25 (10.5)	7 (6.0)
Withdrawal of consent	23 (7.0)	13 (3.9)	7 (2.9)	3 (3.6)
Other	3 (0.9)	3 (0.9)	3 (1.3)	2 (1.7)
Lost to follow-up	0	1 (0.3)	0	0

Source: CS, Figure 7, page 58 & Figure 17, page 91
Abbreviations: AE, adverse event; ITT, intent-to-treat; N, number of randomised patients; n, number of patients in category; SAE, serious adverse event; TEAE, treatment-emergent adverse event.

ERG comment: Reasons for treatment discontinuation are well balanced between the two trial arms of the RAINBOW trial. A greater proportion of patients in the ramucirumab+ paclitaxel group withdrew consent to treatment than in the placebo+paclitaxel group. According to the investigators, the majority of the patients who withdrew consent experienced a Grade ≥ 3 adverse event or serious adverse event within 2 weeks before or after discontinuation of therapy.

In the REGARD trial, a greater proportion of patients in the ramucirumab group discontinued due to an adverse event.

Patient characteristics in two trials

The demographics, baseline disease characteristics and medical history of patients in both trials by treatment arm are presented in Table 4.10.

Table 4.10: Characteristics of participants in the trials by randomised group (ITT)

	RAINBOW		REGARD	
Baseline characteristic	**RAM+PAC (n=330)**	**PBO+PAC (n=335)**	**RAM (n=238)**	**PBO (n=117)**
Age (years) <65 ≥65 Median (range) Median (IQR)	 204 (62) 126 (38) 61 (25 – 83)	 212 (63) 123 (37) 61 (24 – 84)	 60 (52-67)	 60 (51-71)
Gender Male Female	229 (69)	243 (73)	169 (71) 69 (29)	79 (68) 38 (32)
Geographic Regions* 1. 2.	 198 (60) 23 (7)	 200 (60) 21 (6)	 165 (69) 55 (23)	 80 (68) 29 (25)

	RAINBOW		REGARD	
Baseline characteristic	**RAM+PAC (n=330)**	**PBO+PAC (n=335)**	**RAM (n=238)**	**PBO (n=117)**
3.	109 (33)	114 (34)	18 (8)	8 (7)
Race				
White	208 (63)	199 (59)	181 (76)	91 (78)
Asian	110 (33)	121 (36)	39 (16)	17 (15)
Black or other	12 (4)	15 (4)	18 (8)	9 (8)
ECOG performance status				
0 (fully active)	117 (35)	144 (43)	67 (28)	31 (26)
1 (restricted in strenuous activity)	213 (65)	191 (57)	171 (72)	85 (73)
2 (ambulatory but unable to work)			0	1 (1)
Disease measurability				
Measurable	267 (81)	273 (81)	218 (92)	106 (91)
Non measurable	63 (19)	62 (19)	20 (8)	11 (9)
Time to progression from start of first-line therapy				
<6 months	250 (76)	256 (76)	NR	NR
≥6 months	80 (24)	79 (24)		
Progression-free interval on prior therapy, n (%)				
<6 months	NR	NR	154 (65)	83 (71)
≥6 months			81 (34)	34 (29)
Missing			3 (1)	0
Weight loss in the previous 3 months				
≥10%	53 (16)	47 (14)		
<10%	277 (84)	286 (85)	41 (17)	20 (17)
Missing	0	2 (1)	197 (83)	97 (83)
Primary tumour present				
Yes	209 (63)	209 (62)	174 (73)	86 (74)
No	121 (37)	126 (38)	64 (27)	31 (26)
Location of primary tumour				
Gastric	264 (80)	264 (79)	178 (75)	87 (74)
GOJ	66 (20)	71 (21)	60 (25)	30 (26)
Histological subtype				
Intestinal	145 (44)	135 (40)	52 (22)	35 (30)
Diffuse type	115 (35)	133 (40)	96 (40)	44 (38)
Mixed	21 (6)	14 (4)		
Unknown or not available	49 (15)	53 (16)	90 (38)	38 (32)
Tumour grade				
Well differentiated	28 (8)	22 (7)	NR	NR
Moderately differentiated	96 (29)	106 (32)		
Poorly differentiated	186 (56)	186 (56)		
Unknown/missing	20 (6)	21 (6)		
Peritoneal metastasis	163 (49)	152 (45)	64 (27)	45 (38)
Number of metastatic site				

	RAINBOW		REGARD	
Baseline characteristic	**RAM+PAC (n=330)**	**PBO+PAC (n=335)**	**RAM (n=238)**	**PBO (n=117)**
0 - 2	209 (63)	232 (69)	163 (68)	71 (61)
≥3	121 (37)	103 (31)	75 (32)	46 (39)
Presence of ascites				
Yes				NR
No				
Prior surgery (Yes)	52 (16)	65 (19)	NR	NR
Total gastrectomy	80 (24)	59 (18)		
Partial gastrectomy	1 (<1)	2 (<1)		
Other				
Prior treatment lines received				
Neoadjuvant therapy	24 (7)	15 (4)	2 (1)	0
Adjuvant therapy	31 (9)	32 (10)	37 (15)	14 (12)
First-line therapy	329 (100)	335 (100)	199 (84)	103 (88)
First-line platinum/fluoropyrimidine	76 (23)	87 (26)	NR	NR
Triplet: platinum/fluoropyrimidine with anthracycline	253 (77)	246 (73)		
Doublet: platinum/fluoropyrimi-dine without anthracycline				
Previous anti-cancer treatment (by type of drug), n (%)	NR	NR		
Fluoropyrimidine plus platinum			200 (84)	88 (75)
Fluoropyrimidine + other sys drug			13 (5)	17 (15)
Fluoropyrimidine alone			16 (7)	7 (6)
Platinum plus other systemic drug			9 (4)	5 (4)
Prior treatment with a regimen containing targeted agent (any targeted agent)	31 (9)	26 (8)	NR	NR

Superseded – see Erratum

Source: CS, Table 16, Page 61
Data are number (%) unless otherwise indicated.
Abbreviations: ECOG, Eastern Cooperative Oncology Group; PLAT, platinum; FLUO, fluoropyrimidine; GOJ, Gastro-oesophageal junction adenocarcinoma; IQR, interquartile range; RECIST, Response Evaluation Criteria in Solid Tumours. Data are number (%) unless otherwise indicated.
* REGARD: Region 1= North America, Europe, Australia, New Zealand, Israel, Region 2= South and Central America, India, South Africa, Middle East, Region 3=Asia.
RAINBOW: Region 1= Europe, Israel, US and Australia, Region 2=Argentina, Brazil, Chile and Mexico, Region 3= Hong Kong, Japan, South Korea, Singapore and Taiwan.

ERG comment: Overall, treatment arms were well balanced in the two trials. In the RAINBOW trial there was an imbalance with respect to the presence of ascites and ECOG Performance Status. In the REGARD trial there was an imbalance between the treatment

arms with respect to histological subtype, percentage of peritoneal metastases and previous anticancer treatment.

Stratification by geographic region occurred at the time of randomization as Geographic region was considered to be an important potential confounder. The investigators state that region 1 was most similar to the UK, (although the criterion by which this was determined was not stated). Overall the treatment arms for Region 1 participants are reasonably balanced.

As most patients in both trials had gastric cancer (75 to 80%), the evidence from this study is more limited with respect to gastroesophageal cancer and as a greater proportion of the recruited patients were male (70%), the evidence is more limited for female participants.

4.2.1 Results of the RAINBOW trial

The final scope lists the following outcome measures: overall survival, progression-free survival, response rate, health-related quality of life and adverse events. These results will now be discussed. Results presented in the CS are based on the data cut-off point of 22 May 2012. Efficacy analyses were performed using the ITT population, which included all randomised patients who received at least 1 dose of the allocated intervention.

Overall Survival

The primary endpoint of the study was overall survival. At the analysis cut-off date, 256 death events had been observed (256 [77.6%] in the ramucirumab+paclitaxel group and 260 [77.6%] in the placebo+paclitaxel group). According to the investigators ramucirumab+paclitaxel reduced the relative risk of death from any cause in this population by 19% (HR = 0.807; 95% CI: 0.678, 0.962; p=0.0169) compared with placebo+paclitaxel. This represents a 31% (2. 27 months) longer median overall survival in the ramucirumab+paclitaxel arm (9.63 months (95% CI 8.6-10.8) versus 7.36 (95% CI 6.3-8.4) months in the placebo+paclitaxel arm). The 6- and 12-month survival rates were (ramucirumab+paclitaxel vs placebo+paclitaxel) 71.5% versus 56.9% and 40.1% versus 30.2%, respectively. The Kaplan Meier survival curves overlapped during the first month, but separated within two months of treatment commencement and remained separate beyond one year of treatment.

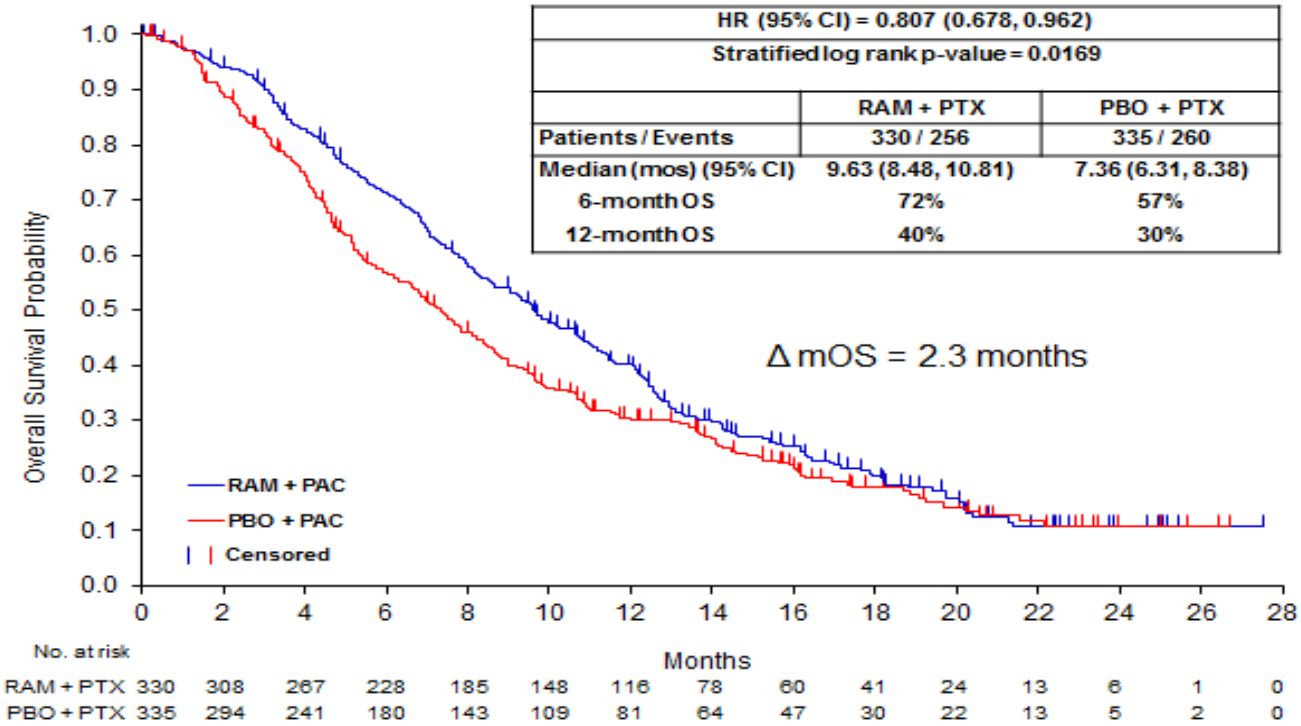
Table 4.11: Overall survival of patients in the RAINBOW trial (ITT)

	Median (95% CI) months to outcome	
Outcome	**RAM + PAC (N=330)**	**PBO + PAC (N=335)**
Number of deaths n%	256 (77.6)	260 (77.6)
Number censored	74 (22.4)	75 (22.4)
Median survival-months (95% CI)	9.63 (8.5, 10.8)	7.36 (6.3, 8.4)
HR (95% CI)	0.807 (0.678-0.962)	
P value	0.0169	

Source: CS, Table 17, Page 63
Abbreviations: CI, confidence interval; HR, hazard ratio; ITT, intent-to-treat; N, number of randomized patients; n, number of patients in category; RAM, ramucirumab; PAC, paclitaxel; PBO, placebo.
Note: Median and survival rates, along with 95% CIs, were estimated using the Kaplan-Meier method.

The Kaplan Meier plot of overall survival (ITT population) is presented in Figure 4.2.

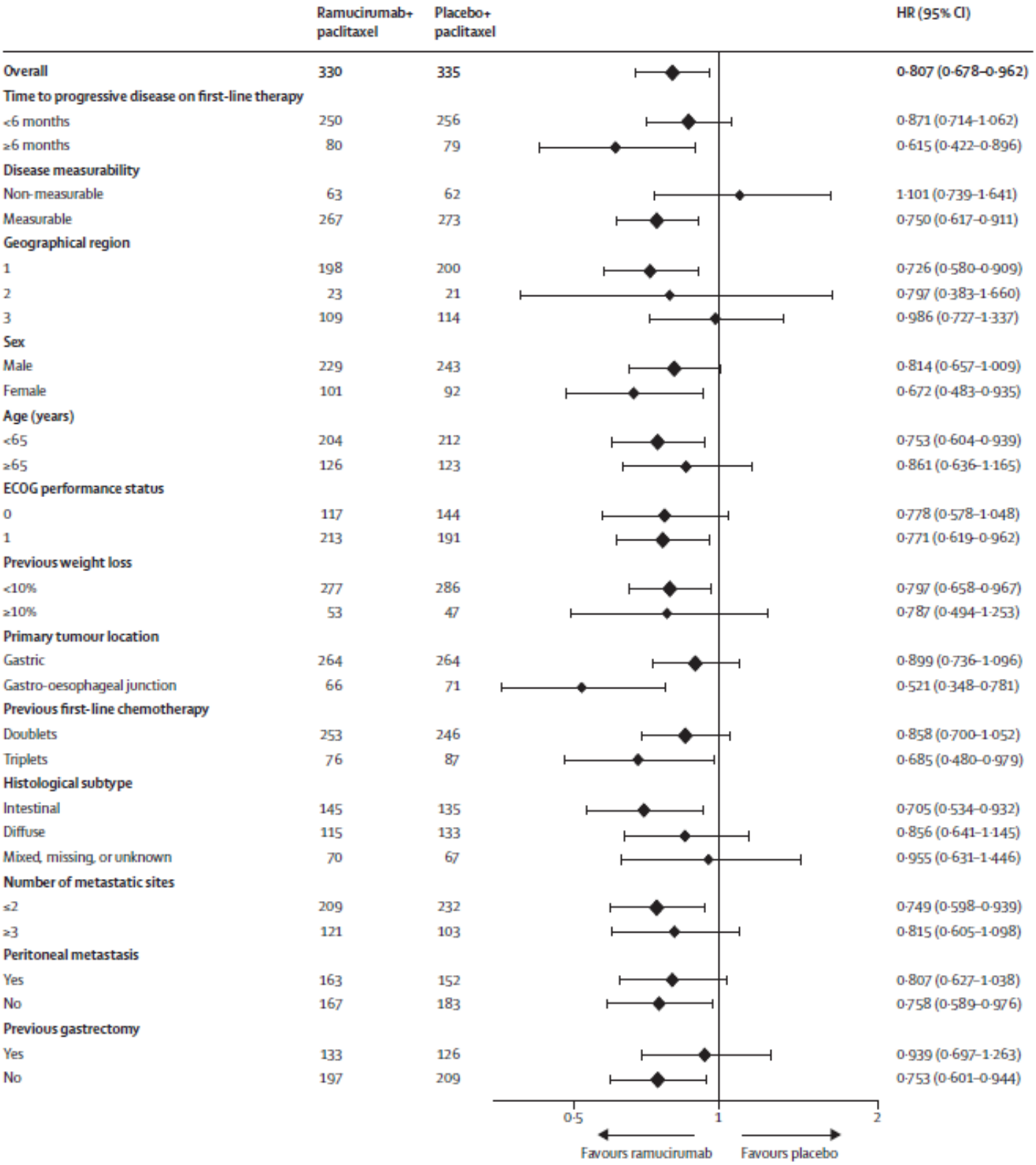
Figure 4.2: RAINBOW: Kaplan-Meier plot of overall survival (ITT population)



Source: CS, Figure 8, page 63
Abbreviations: CI, confidence interval; HR, Hazard ratio; ITT, intent to treat, PBO, placebo, PTX, paclitaxel, RAM, ramucirumab.

Subgroup analyses were performed showing forest plots for univariate analyses of overall survival (see Figure 4.3). The subgroup analyses show that ramucirumab+paclitaxel when compared with paclitaxel+placebo is less effective in non-measurable disease, region 3 (Asia), males, elderly, gastric cancer, diffuse and mixed, missing or unknown histological subtypes and in patients who had a previous gastrectomy.

Figure 4.3: Forest plot for subgroup univariate analyses of overall survival



Source: Wilke et al. 2014[9]
Abbreviations: ECOG, Eastern Cooperative Oncology Group; HR, Hazard ratio.

Progression-free survival

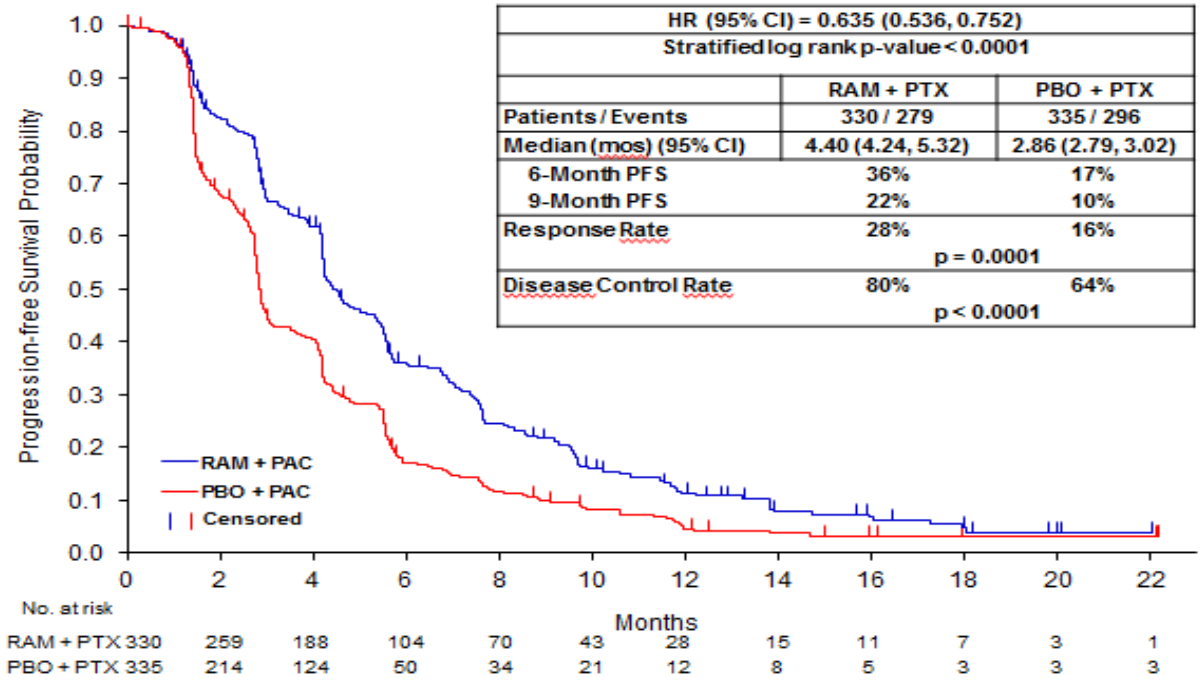
Treatment with ramucirumab+paclitaxel resulted in a 37% relative reduction in the risk of disease progression or death compared with placebo+paclitaxel and increased PFS by 1.5 months compared with the placebo plus paclitaxel arm (4.4 months versus 2.9 months. (See Table 4.12). The Kaplan–Meier plot for PFS (ITT) is presented in Figure 4.4.

Table 4.12: RAINBOW: Progression-free survival results (ITT population)

Outcome	Median (95% CI) months to outcome	
	RAM + PAC (N=330)	PBO +PAC (N=335)
Number of deaths n%	279 (84.5)	296 (88.4)
Number censored	51(15.5)	39 (11.6)
Median PFS survival-months (95% CI)	4.40 (4.24, 5.32)	2.86 (2.79, 3.02)
HR (95% CI)	0.635 (0.536, 0.752)	
P value	< 0.0001	

Source: CS, Page 64, Table 18
Abbreviations: CI = confidence interval; ITT = intent-to-treat; N = number of randomized patients; n = number of patients in category; PFS = progression-free survival; PAC, paclitaxel; PBO, placebo; RAM, ramucirumab.
Note: Median and survival rates, along with 95% CIs, were estimated using the Kaplan-Meier method.

Figure 4.4: RAINBOW: Kaplan-Meier plot of Progression-Free Survival (ITT population)



Source: Page 94, Table 11.5.2 RAINBOW clinical study report
Abbreviations: CI, confidence interval; HR, Hazard ratio; ITT, intent to treat; PBO, placebo; PFS, progression free survival; PTX, paclitaxel; RAM, ramucirumab.

Response rate

Response rates are reported below. Significant differences in favour of the ramucirumab+ paclitaxel group were observed for objective response rate (complete or partial response according to RECIST criteria), and duration of response.

Table 4.13: RAINBOW: Response to treatment results – ITT population

Best overall response	RAM + PAC N = 330	PBO + PAC N = 335
Patients with measurable disease at baseline	267 (81)	273 (81)
Patients with best overall response, n (%)	92 27.9%	54 16.1%
Complete response (CR)	2 (0.6)	1 (0.3)
Partial response (PR)	90 (27.3)	53 (15.8)
Stable disease (SD)	172 (52.1)	159 (47.5)
Progressive disease (PD)	43 (13.0)	83 (24.8)
Not evaluable (NE) /Not Done	23 (7.0)	39 (11.6)

Source: CS, Table 20, Page 65
Abbreviations: ITT, intent-to-treat; N, number of randomised patients; n, number of patients in category; PAC, paclitaxel; PBO, placebo; RAM, ramucirumab.

Health related quality of life

Health related quality of life in the RAINBOW trial was assessed using two validated instruments: the European Organisation for Research and Treatment of Cancer, Quality of Life Questionnaire, Core 30. Version 3.0 (EORTC-QLQ-C30) and EuroQol 5-dimension (EQ-5D).

Overall, the change in quality of life was similar in both treatments arms. There were no significant differences between treatment arms for both instruments, as can be seen in Tables 4.14 and 4.15.

Table 4.14: RAINBOW: EORTC QLQ-C30 - global health status results, ITT population

EORTC QLQ-C30, mean (SD)	Ramucirumab+paclitaxel N=330	Placebo+paclitaxel N=335
Baseline (N=322/326)	61.46 (21.952)	58.03 (22.031)
End of Treatment (N=211/204)	48.97 (22.979)	48.28 (23.897)
Change from Baseline (N=209/2042	-13.48 (23.238)	-12.13 (24.813)

Source: CSR RAINBOW, Table 14.2.36, page 516; Cut-off Date: 12 July 2013
Abbreviations: SD, standard deviation.
Based on a 100-point scale, with a higher score representing better quality of life

Table 4.15: RAINBOW: EQ-5D Results, ITT population

EQ-5D Index Score, mean (SD)	Ramucirumab+paclitaxel N=330	Placebo+paclitaxel N=335
Baseline (N=323/328)	0.741 (0.228)	0.732 (0.250)
End of Treatment (N=211/206)	0.581 (0.335)	0.570 (0.366)

Source: CS, Table 21, page
Abbreviations: SD, standard deviation.
Based on a –0.59 to 1 scale, with 1 representing perfect health. Calculated based on the UK population-based preference weights for EQ-5D. These are based on values elicited from a representative national sample using the time trade-off (TTO) method.

Adverse events

All adverse events data presented in the CS are from the RAINBOW and REGARD trials. Overall safety results for the RAINBOW trial are shown in Table 4.16. Similar numbers of patients had at least one serious adverse event (153 [47%] of 327 in the ramucirumab plus

paclitaxel group versus 139 [42%] of 329 in the placebo plus paclitaxel group), or treatment-emergent adverse event leading to death (39 [12%] versus 51 [16%], respectively).

Table 4.16: RAINBOW: Overall safety results, safety population

Safety outcome, n (%)	**RAM + PAC N=327**	**PBO + PAC N=329**
Number of patients with TEAEs	324 (99.1)	322 (97.9)
Number of patients with grade 3–4 TEAEs	267 (81.7)	206 (62.6)
Number of patients with treatment-emergent SAEs	153 (46.8)	139 (42.2)
Number of patients with TEAEs leading to treatment discontinuation	102 (31.2)	80 (24.3)
Number of patients with TEAE leading to death	39 (11.9)	51 (15.5)

Source: RAINBOW CSR, Table 12.3.1, Page 124
Abbreviations: TEAE, treatment-emergent adverse event; SAE serious adverse events
Patients may be counted in more than one category.

The incidence of grade 3 or 4 adverse events was higher in the ramucirumab plus paclitaxel group, including grade 3 or 4 neutropenia, leucopenia, and grade 3 hypertension, abdominal pain, and fatigue (See Table 4.17).

ERG comment: The ERG sought advice from the company on the proportion of neuropathy cases that were irreversible. The sponsor indicated that this level of information was not collected in the trial and that neuropathy was primarily a feature of paclitaxel therapy.

Table 4.17: RAINBOW: Treatment-emergent adverse events occurring in at least 10% of patients on ramucirumab plus paclitaxel, irrespective of causality

	Ramucirumab plus paclitaxel (n=327)				Placebo plus paclitaxel (n=329)			
	Grades 1-2	Grade 3	Grade 4	Grade 5	Grades 1-2	Grade 3	Grade 4	Grade 5
Any patients with a treatment-emergent adverse event	57 (17%)	155 (47%)	73 (22%)	39 (12%)	116 (35%)	128 (39%)	27 (8%)	51 (16%)
Non-haematological adverse events								
Fatigue*	147 (45%)	39 (12%)	0	0	126 (38%)	18 (5%)	0	0
Neuropathy*	123 (38%)	27 (8%)	0	0	104 (32%)	15 (5%)	0	0
Decreased appetite	121 (37%)	10 (3%)	0	0	92 (28%)	13 (4%)	0	0
Abdominal pain*	98 (30%)	20 (6%)	0	0	87 (26%)	10 (3%)	1 (<1%)	0
Nausea	109 (33%)	5 (2%)	1 (<1%)	0	100 (30%)	8 (2%)	0	0
Alopecia	107 (33%)	0	0	0	126 (38%)	1 (<1%)	0	0
Diarrhoea	94 (29%)	12 (4%)	0	0	71 (22%)	4 (1%)	1 (<1%)	0
Epistaxis	100 (31%)	0	0	0	23 (7%)	0	0	0
Vomiting	78 (24%)	9 (3%)	1 (<1%)	0	56 (17%)	12 (4%)	0	0
Peripheral oedema	77 (24%)	5 (2%)	0	0	43 (13%)	2 (<1%)	0	0
Hypertension	32 (10%)	46 (14%)	0	0	8 (2%)	8 (2%)	0	0
Constipation	70 (21%)	0	0	0	69 (21%)	2 (<1%)	0	0
Stomatitis	62 (19%)	2 (<1%)	0	0	22 (7%)	2 (<1%)	0	0
Pyrexia	56 (17%)	3 (<1%)	0	0	36 (11%)	1 (<1%)	0	0
Proteinuria	50 (15%)	4 (1%)	0	0	20 (6%)	0	0	0
Malignant neoplasm progression	5 (2%)	16 (5%)	4 (1%)	27 (8%)	1 (<1%)	24 (7%)	1 (<1%)	34 (10%)
Weight decreased	39 (12%)	6 (2%)	0	0	45 (14%)	4 (1%)	0	0
Dyspnoea	34 (10%)	8 (2%)	0	0	29 (9%)	2 (<1%)	0	0
Rash*	42 (13%)	0	0	0	31 (9%)	0	0	0
Cough	40 (12%)	0	0	0	25 (8%)	0	0	0
Back pain	35 (11%)	4 (1%)	0	0	35 (11%)	5 (2%)	0	0
Hypoalbuminaemia*	32 (10%)	4 (1%)	0	0	13 (4%)	2 (<1%)	0	1 (<1%)

	Ramucirumab plus paclitaxel (n=327)				Placebo plus paclitaxel (n=329)			
	Grades 1-2	Grade 3	Grade 4	Grade 5	Grades 1-2	Grade 3	Grade 4	Grade 5
Myalgia	34 (10%)	0	0	0	32 (10%)	1 (<1%)	0	0
Ascites	21 (6%)	11 (3%)	1 (<1%)	0	14 (4%)	13 (4%)	0	0
Headache	32 (10%)	0	0	0	21 (6%)	1 (<1%)	0	0
Haematological adverse events								
Neutropenia*	45 (14%)	71 (22%)	62 (19%)	0	40 (12%)	51 (16%)	11 (3%)	0
Anaemia*	84 (26%)	30 (9%)	0	0	85 (26%)	31 (9%)	3 (<1%)	0
Leucopenia*	54 (17%)	52 (16%)	5 (2%)	0	47 (14%)	19 (6%)	3 (<1%)	0
Thrombocytopenia*	38 (12%)	5 (2%)	0	0	14 (4%)	6 (2%)	0	0

Source: Wilke et al. 2014[9]

Table 4.18: RAINBOW: Adverse events of special interest*

	Ramucirumab plus paclitaxel (n=327)				Placebo plus paclitaxel (n=329)			
	Grades 1-2	Grade 3	Grade 4	Grade 5	Grades 1-2	Grade 3	Grade 4	Grade 5
Bleeding or haemorrhage	123 (38%)	12 (4%)	1 (<1%)	1 (<1%)	51 (16%)	4 (1%)	2 (<1%)	2 <1%)
Proteinuria	51 (16%)	4 (1%)	0	0	20 (6%)	0	0	0
Liver injury or failure	39 (12%)	12 (4%)	3 (<1%)	0	28 (9%)	11 (3%)	2 (<1%)	0
Hypertension	34 (10%)	48 (15%)	0	0	10 (3%)	9 (3%)	0	0
Gastrointestinal haemorrhage†	21 (6%)	10 (3%)	1 (<1%)	1 (<1%)	15 (5%)	3 (<1%)	1 (<1%)	1 (<1%)
Infusion-related reaction	17 (5%)	2 (<1%)	0	0	12 (4%)	0	0	0
Renal failure	16 (5%)	4 (1%)	2 (<1%)	0	11 (3%)	0	1 (<1%)	2 (<1%)
Congestive heart failure	6 (2%)	2 (<1%)	0	0	2 (<1%)	1 (<1%)	0	1 (<1%)
Venous thromboembolic events	5 (2%)	7 (2%)	0	1 (<1%)	7 (2%)	8 (2%)	1 (<1%)	2 (<1%)
Arterial thromboembolic events	3 (<1%)	1 (<1%)	2 (<1%)	0	2 (<1%)	2 (<1%)	0	1 (<1%)
Gastrointestinal perforation	0	1 (<1%)	2 (<1%)	1 (<1%)	1 (<1%)	0	0	0

Source: Wilke et al. 2014[9]
*Pooled adverse-event terms.
†Events pooled as gastrointestinal haemorrhage are also pooled as bleeding or haemorrhage.

Grade 3 adverse events that were potentially associated with the VEGF pathway—and thus were of special interest—that were more common in the ramucirumab plus paclitaxel group included hypertension, proteinuria, and bleeding or haemorrhage. The incidences of grade 4 and 5 adverse events of special interest were low in both groups, with no grade 4 or 5 hypertension, a similar incidence of gastrointestinal haemorrhage, and a higher incidence of gastrointestinal perforation in the ramucirumab plus paclitaxel group than the placebo plus paclitaxel group (See Table 4.18).

The most commonly reported subcategories of AEs of special interest were Bleeding (any grade: ramucirumab+paclitaxel 42%, placebo+paclitaxel 18%). Epistaxis (nose bleed) was the underlying cause of 75% of these events while GI bleeding accounted for the other 25%. The incidence of grade ≥3 bleeding events was low. Additionally, proteinuria (17% and 6%) and hypertension (25% and 6%) occurred more frequently in the ramucirumab+paclitaxel arm than in the placebo+paclitaxel arm.

4.2.2 Results of the REGARD trial

Overall survival

The primary endpoint of REGARD was overall survival. Primary and secondary endpoints were analysed using the ITT population. At the data cut-off date, 179 (75.2%) patients had died in the ramucirumab arm and 99 (84.6%) in the placebo arm (See Table 4.19). Median overall survival was 5.2 months among patients treated with ramucirumab and 3.8 months among those treated with placebo (HR=0.776; log-rank p=0.047 in stratified analysis), representing a 36.8% increase in median overall survival with ramucirumab.

Table 4.19: Overall survival of patients in the REGARD trial (ITT)

Outcome	Median (95% CI) months to outcome	
	Ramucirumab (N=238)	**Placebo (N=117)**
Number of deaths n%	179 (75.2)	99 (84.6)
Number censored	59 (24.8)	18 (15.4)
Median survival-months (95% CI)	5.2 (4.4, 5.7)	3.8 (2.8, 4.7)
HR (95% CI)	0.776 (0.603-0.998)	
P value	0.0473	

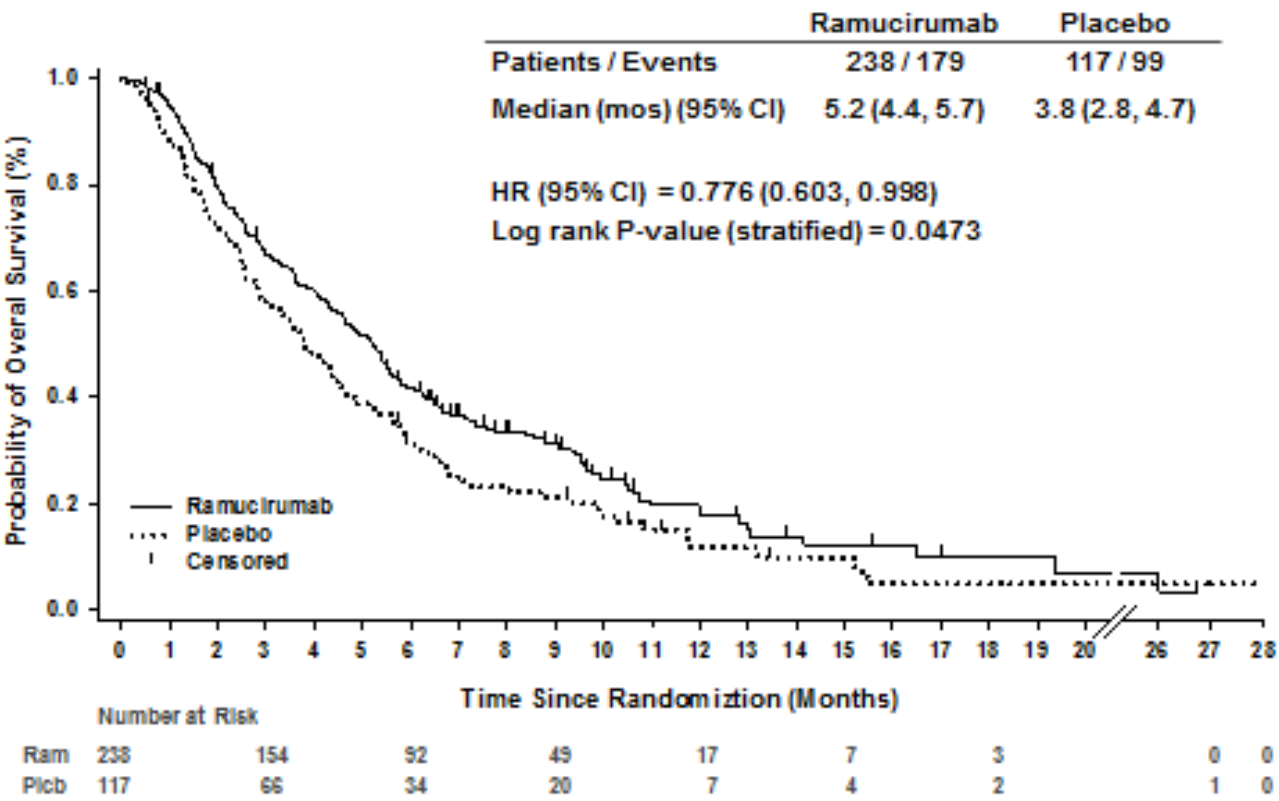
Source: CS, Table 35, Page 96

Abbreviations: CI, confidence interval; HR, hazard ratio; ITT, intent-to-treat; N, number of randomized patients; n, number of patients in category

Note: Median and survival rates, along with 95% CIs, were estimated using the Kaplan-Meier method.

The Kaplan Meier plot of overall survival (ITT population) is presented in Figure 4.5.

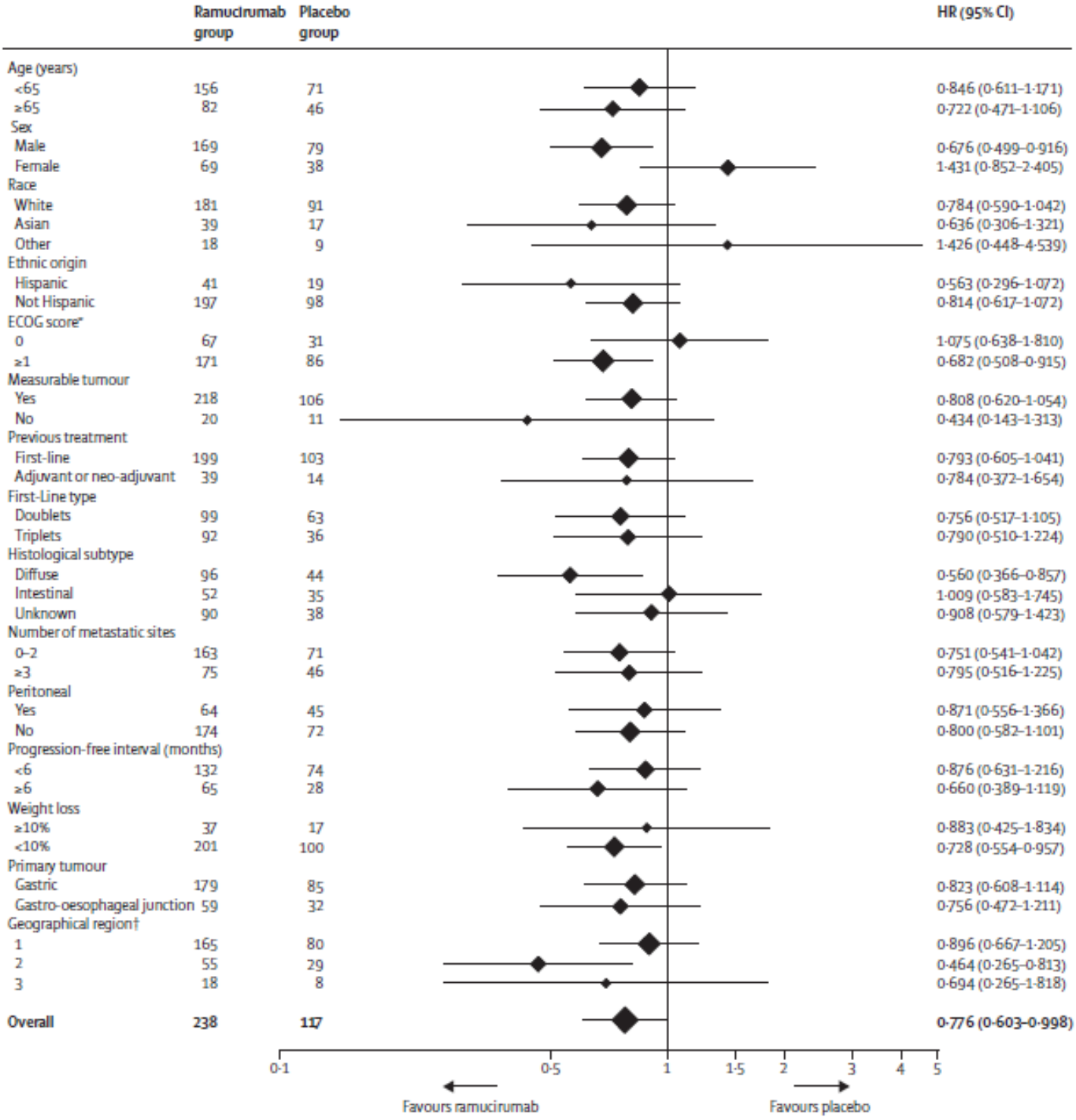
Figure 4.5: REGARD: Kaplan-Meier plot of overall survival (ITT population)



Source: CS, Figure 18, page 96
Abbreviations: CI, confidence interval; HR, Hazard ratio; ITT, intent to treat

Subgroup analyses were performed showing forest plots for univariate analyses of overall survival (see Figure 4.6). The subgroup analyses show that ramucirumab when compared with placebo is less effective in females, patients who are not white or Asian, and ECOG score 0.

Figure 4.6: Forest plot for subgroup univariate analyses of overall survival



Source: Fuchs et al. 2014[10]
Abbreviations: ECOG, Eastern Cooperative Oncology Group; HR, Hazard ratio.

Progression-free survival

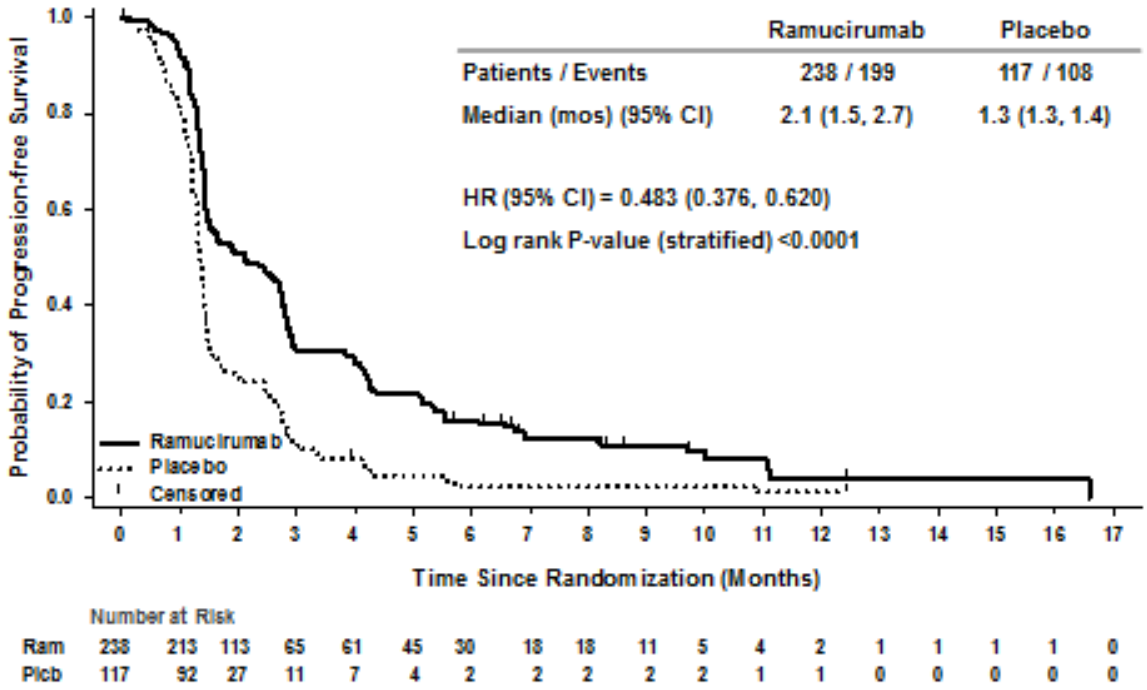
Treatment with RAM resulted in a 62% longer median time to disease progression in the ramucirumab arm (2.1 months versus 1.3 months) (See Table 4.20). The Kaplan–Meier plot for PFS is presented in Figure 4.7.

Table 4.20: REGARD: Progression-free survival results (ITT population)

Outcome	Median (95% CI) months to outcome	
	RAM (N=238)	PBO (N=117)
Number of deaths n%	199 (83.6)	108 (92.3)
Number censored	39 (16.4)	9 (7.7)
Median PFS survival-months (95% CI)	2.1 (1.5, 2.7)	1.3 (1.3, 1.4)
HR (95% CI)	0.483 (0.376, 0.620)	
P value	p<0.0001	

Source: CS, Page 97, Table 36
Abbreviations: CI = confidence interval; ITT = intent-to-treat; N = number of randomized patients; n = number of patients in category; PFS = progression-free survival; PBO, placebo; RAM, ramucirumab.
Note: Median and survival rates, along with 95% CIs, were estimated using the Kaplan-Meier method.

Figure 4.7: REGARD: Kaplan-Meier plot of Progression-Free Survival (ITT population)



Source: CS, Figure 19, Page 98
Abbreviations: mos = months.

Response rate

Response rates are reported in Table 4.21. A significant difference in favour of the ramucirumab group was observed for duration of response. There was no significant difference between groups for objective response rate (complete or partial response according to RECIST criteria).

Table 4.21: REGARD: Response to treatment results – ITT population

Best overall response	Ramucirumab (N=238)	Placebo (N=117)
Patients with measurable disease at baseline	218 (92)	106 (91)
Patients with best overall response, n (%)	8 (3.4)	3 (2.6)
Complete response (CR)	1 (0.4)	0 (0)
Partial response (PR)	7 (2.9)	3 (2.6)
Stable disease (SD)	108 (45.4)	24 (20.5)
Progressive disease (PD)	78 (32.8)	63 (53.8)
Not evaluable (NE) /Not Done	44 (18.5)	27 (23.1)

Source: CS, Table 37, Page 99
Abbreviations: ITT, intent-to-treat; N, number of randomized patients; n, number of patients in category; PBO, placebo; RAM, ramucirumab.

Health related quality of life

Health related quality of life in the REGARD trial was assessed using the European Organisation for Research and Treatment of Cancer, Quality of Life Questionnaire, Core 30. Version 3.0 (EORTC-QLQ-C30).

The number of completed QoL questionnaires decreased with time. At six weeks, only 114 (48%) patients in the ramucirumab arm versus 29 (25%) in the placebo arm provided QoL data, primarily due to disease progression and study discontinuation before the first scheduled post-baseline assessment, rather than non-compliance. At 18 weeks, only 38 (16%) patients in the ramucirumab arm versus five (4%) in the placebo arm provided QoL data

At six weeks, the proportion of patients with improved or stable QoL was higher for the ramucirumab arm (34.1%) than the placebo arm (13.7%); however, the difference was not statistically significant (p=0.23). Also at six weeks, there were higher proportions of patients in the RAM arm for the physical functioning, fatigue and pain subscales.

Adverse events

Overall safety results for the REGARD trial are shown in Table 4.22. Similar numbers of patients had at least one serious adverse event (106 [45%] of 236 in the ramucirumab group *vs* 51 [44%] of 115 in the placebo group), or adverse event leading to death (22 [9%] versus 15 [13%], respectively).

Table 4.22: REGARD: Overall safety results, safety population

Safety outcome, n (%)	RAM N=236	PBO N=115
Number of patients with TEAEs	223 (94.5)	101 (87.8)
Number of patients with grade 3–4 TEAEs	134 (56.8)	67 (58.3)
Number of patients with treatment-emergent SAEs	106 (44.9)	51 (44.3)
Number of patients with TEAEs leading to treatment discontinuation	25 (10.5)	7 (6.0)
Deaths due to an AE	22 (9.3)	15 (13.0)

Source: Table 12.3, REGARD clinical study report
Abbreviations: TEAE, treatment-emergent adverse event; SAE serious adverse events
Patients may be counted in more than one category.

The following adverse events were considered of special interest: infusion related reactions (IRR), hypertension, proteinuria, arterial and venous thromboembolic events, bleeding/ hemorrhagic events, gastrointestinal perforation, congestive heart failure, wound healing complications, fistula, liver failure / liver injury, and reversible posterior leukoence-phalopathy syndrome (RPLS).

The most commonly reported subcategories of AEs of special interest were any grade bleeding (13% ramucirumab versus 11% placebo), hypertension (16% ramucirumab versus 8% placebo), venous thromboembolism (4% ramucirumab versus 7% placebo), proteinuria (3% ramucirumab versus 3% placebo), IRRs (<1% ramucirumab and 2% placebo) and arterial thromboembolism (2% ramucirumab versus 0 placebo). Hypertension and arterial thromboembolism were more common in ramucirumab than placebo (See Table 4.23).

Table 4.23: REGARD: Adverse events of special interest

Adverse events of special interest	Ramucirumab (n=236)		Placebo (n=115)	
	Any event	Grade ≥3	Any event	Grade ≥3
Hypertension§	38 (16%)	18 (8%)	9 (8%)	3 (3%)
Bleeding or haemorrhage¶	30 (13%)	8 (3%)	13 (11%)	3 (3%)
Arterial thromboembolism‖	4 (2%)	3 (1%)	0	0
Venous thromboembolism**	9 (4%)	3 (1%)	8 (7%)	5 (4%)
Proteinuria	7 (3%)	1 (<1%)	3 (3%)	0
Gastrointestinal perforation	2 (<1%)	2 (<1%)	1 (<1%)	1 (<1%)
Fistula formation	1 (<1%)	1 (<1%)	1 (<1%)	1 (<1%)
Infusion-related reaction	1 (<1%)	0	2 (2%)	0
Cardiac failure	1 (<1%)	0	0	0

Source: Fuchs et al. 2014[10]
Data are n (%), unless otherwise indicated. §Includes increased blood pressure. ¶Includes epistaxis, gastric haemorrhage, gastrointestinal haemorrhage, gingival bleeding, haematemesis, haematoma, haematuria, haemoptysis, haemorrhage, haemorrhoidal haemorrhage, melaena, nail-bed bleeding, petechiae, rectal haemorrhage, and upper gastrointestinal haemorrhage. ‖Includes angina pectoris, cardiac arrest, cerebral ischaemia, cerebrovascular accident, myocardial infarction, and myocardial ischemia. **Includes pulmonary embolism, deep vein thrombosis, thrombosis, and venous thrombosis in a limb.

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

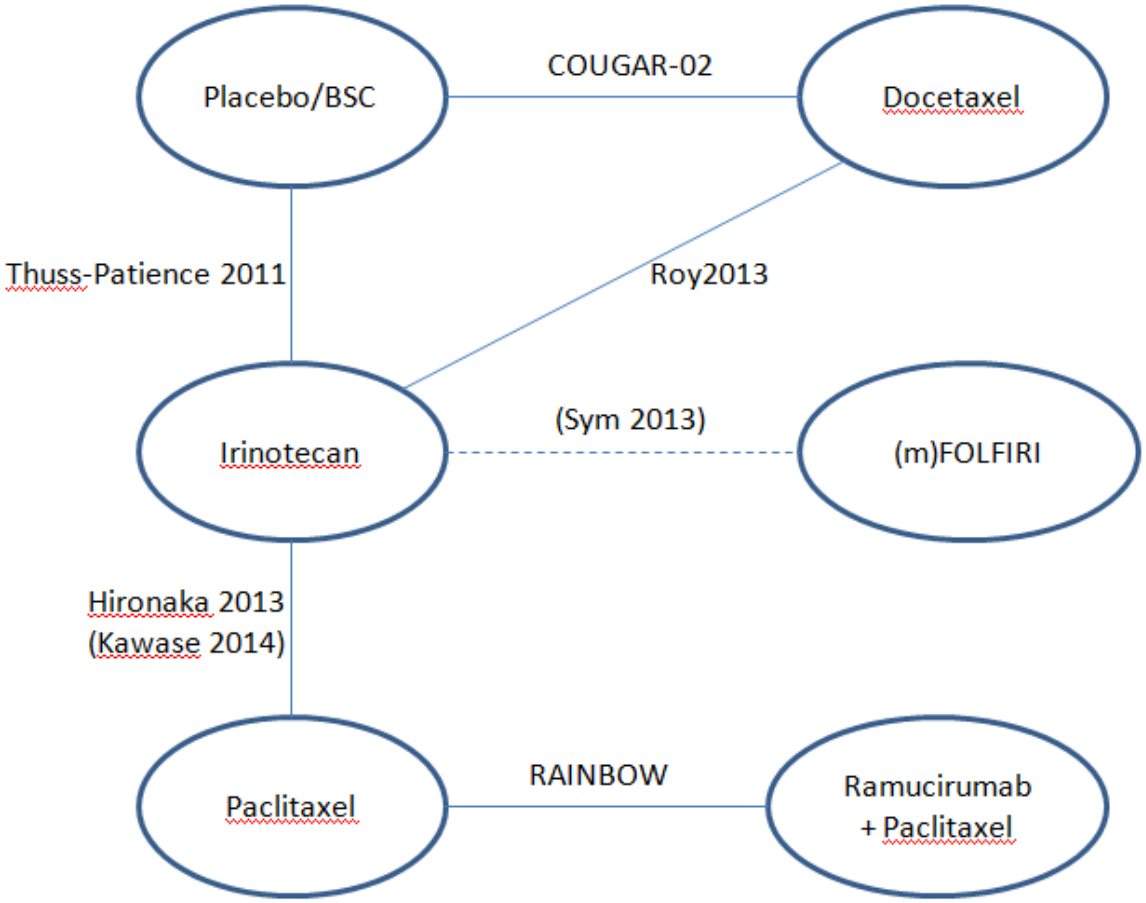
The company submission states (page 106, section 4.10) that a meta-analysis was performed to focus on the relative efficacy of RAM as a monotherapy or combination therapy in comparison with treatments of interest in adult patients who have received prior chemotherapy for GC/GOJ in terms of survival, tumour response and toxicity.

ERG Comment: The ERG sought clarification from the sponsor about the number of trials included in the network meta-analysis (NMA), reported in Appendix 6 of the company submission. In the response to the clarification letter, the company reports that 23 trials were included in the final potential NMA, which were presented in 30 publications.

However, for ramucirumab monotherapy, only one trial was used (Fuchs 2014 (REGARD)[10]), which compares ramucirumab with placebo. No other comparisons were reported for ramucirumab monotherapy.

For ramucirumab plus paclitaxel only six trials were included in the actual analyses. These trials are: Ford 2014 (COUGAR-02)[27]; Hironaka 2013[29]; Roy 2013[28]; Sym 2013[26]; Thuss-Patience 2011[30]; and Wilke 2014 (RAINBOW)[9]. The study by Kawase, 2014 was potentially of interest. Only an abstract for this study was available; but this abstract was not provided by the company, nor did the company explain why it was not included in the analyses. The evidence network of relevant studies for the analyses in the CS is presented in Figure 4.8. The study by Sym et al 2013[26] was only used in the additional analyses presented in the response to the clarification letter.

Figure 4.8: Evidence network of relevant studies for the company analyses



Heterogeneity

The company investigators noted several sources of heterogeneity which were identified across the studies included in the analyses. The evidence networks were too small (single studies for most treatment comparisons within the networks) to explore potential sources of heterogeneity as listed above or to employ meta-regression to explore the effects of various covariates. Sources of heterogeneity include:

- Site of primary disease (gastric or GOJ)
- Disease stage/PS
- Criteria/schedule used for tumour assessment
- Previous chemotherapy
- Treatment duration

- Patient cross-over
- Differential dosing of IRI across studies
- Lower dosing of DOC monotherapy in Asian studies compared with non-Asian studies
- Third-line chemotherapy administered post protocol
- Inconsistent reporting of safety outcomes across study publications
- The country where the study was conducted may be of importance as prognosis may vary across territories.

Table 4.24: Heterogeneity in network of studies for combination therapy

	RAINBOW	**Hironaka 2013**	**Thuss-Patience 2011**	**Roy 2013**	**Cougar-02**	**Sym 2013**
Study location	27 countries including Central/ North/ South America, Europe, Asia, Australia [34% recruited from Asia]	Japan	Germany	UK, Spain, Taiwan, Croatia, South Korea and Bosnia [46% recruited from Asia]	United Kingdom	South Korea
Nr of centres	Multinational	Multicentre	Multicentre	Multinational	Multicentre	Single centre
Site of primary disease	Gastric, 79% GEJ, 21%	Gastric, 100%	Gastric, 57% GEJ, 43%	Gastric, 77% GEJ, 23%	Gastric, 45% GEJ, 35% oesophagus, 20%	Gastric, 92% GEJ, 8%
ECOG status	0-1	0-2	0-2	0-2	0-2	0-2
Phase	PHASE III	PHASE III	PHASE III	PHASE III	PHASE III	PHASE II
Blind/Open Label	Double blind	Open label	Open label	NR	Open-label	Open label
Age range	24-84	37-75	35-73	33-81	28-84	30-76
Males%	70.9%	78%	72.5%	78.0%	80.1%	57.6%
Treatment ARMS	Ramucirumab 8 mg/kg, day 1 and 15 + paclitaxel, 80 mg/m^2, day 1, 8, and 15 Ramucirumab 8 mg/kg, + placebo, day 1 and 15	Paclitaxel 80 mg/m^2 on days 1, 8 and 15 Irinotecan 150 mg/m^2 on days 1 and 15	Irinotecan 250-350 mg/m^2 + BSC BSC	PEP02120 mg/m^2 Irinotecan 300 mg/m^2 Docetaxel 75 mg/m^2	Docetaxel 75 mg/m^2 iv + ASC ASC	Irinotecan 150 mg/m^2 mFOLFIRI (irinotecan 150 mg/m^2 plus LV 20 mg/m^2 on day 1 followed by 5-FU 2,000 mg/m^2)
Cycle length	day 1, (8) and 15 every 4 weeks	days 1, 8 and 15, every 4 weeks	3 weeks	day 1; every 3 weeks	3 weeks	2 weeks
Irinotecan dosing	Not applicable	150 mg/m^2	250 mg/m^2, day rising to 350 mg/m^2	300 mg/m^2	Not applicable	150 mg/m^2

	RAINBOW	Hironaka 2013	Thuss-Patience 2011	Roy 2013	Cougar-02	Sym 2013
			(37% of patients)			
Maximum chemotherapy cycles	Till progression	Till progression	Limit of 10 cycles	Limit of 6 cycles	Till progression	Limit of 12 cycles
Schedule used for tumour assessment	Every 6 weeks	Every 2 months	Every 3 weeks. Radiological assessments were conducted every 6 weeks, but only in the experimental arm	Every 6 weeks (every 2 treatment cycles)	Every 3 weeks (before each treatment cycle). Radiological assessments were conducted at 9 and 18 weeks, but only in the experimental arm	Every 6 weeks
Prior treatment	Platinum/ fluoro-pyrimidine + anthracycline, Platinum/ fluoro-pyrimidine only,	Fluoro-pyrimidine plus platinum therapy	First-line palliative chemotherapy; all patients were pre-treated with cisplatin, 39 patients with a cisplatin/5-FU based combination, 1 patient with cisplatin/ docetaxel	One prior systemic chemotherapy regimen	Platinum and fluoro-pyrimidine-based therapy	fluoropyrimidine and/or platinum chemotherapy taxanes and/or platinum chemotherapy
First/Second line prior treatment	Second line	Second line	Second line	Second line	Second line	First line

Table 4.24 addresses the heterogeneity of the particular populations and treatment characteristics for the included studies in the network meta-analyses. As can be seen, there is wide variety in the characteristics of the included studies. These include differences in ECOG performance status, chemotherapy cycle length, irinotecan dosing regimen, gender, proportion of GOJ patients recruited, time schedule used for tumour assessments, prior treatment, treatment duration and country where the study was conducted. Few studies included UK residents with only one study focusing solely on UK residents. Two studies focused only on participants recruited from Asia. This is important given the consideration of a more intensive prevention and treatment regimen in Asian countries.

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

Two types of syntheses are reported: In Chapter 4.10 of the company submission a series of indirect comparisons apparently using the Bucher method[31] are described; however, the methods for this are not reported. In Appendix 6 of the CS, the results of a full network meta-analysis (NMA), using WinBUGS, are reported.

The results for the indirect comparisons (Chapter 4.10 CS) and network meta-analysis (Appendix 6, CS) for OS and PFS are almost identical, and results of the NMA are used in the economic model. Therefore, we will discuss the NMA only.

Analysis methods

The NMA used a Bayesian model and was performed in WinBUGS. The models used the standard code as recommended by the NICE Technical Support Documents and the analysis of hazard ratios (HR) for overall and progression-free survival used the natural logarithm of the HR with its standard error (SE) for each study. The analysis used non-informative priors (as recommended by NICE) and ran 100,000 iterations as a burn-in period followed by 50,000 iterations to estimate the results. Only the code for fixed effects models was provided. However, given that all the trials were two arm trials (or if more than two arms, only one two arm comparison was used in the analysis) and the numbers were low (six trials for PFS, eight for OS) the use of a fixed effect model is appropriate.

The code used for the binary outcomes (withdrawals and adverse events) was not provided.

Overall survival from the REGARD trial was included in the WinBUGS code (Response to the Clarification Letter, Appendices 1.1.2.1 for OS) but not in the data tables, it seems to have been included in the models but not in the data table (CS, Appendix Table 2) or results. Similarly Higuchi is in the WinBUGS code but not in the data table. The data used in the WinBUGS models should match the data reported in the tables for each outcome. If it did not inform the network then it should not have been included in the WinBUGS code.

The ERG ran the code provided for the treatment differences (survival outcomes, section 1.1.1) but was unable to run it as it contained terms for the mean effect (meanA) and precision (precA) for the ‘standard’ treatment A, which were not provided. However, on altering the code to remove that section, we ran the models using the data provided for the fixed effect models. The results are shown below in Table 4.25.

Table 4.25: ERG results of CS base-case for overall survival (HR, 95% Credible interval (CrI))

Treatment A	Treatment B				
	Placebo/BSC	**Docetaxel**	**Paclitaxel**	**Irinotecan**	**mFOLFIRI**
Docetaxel	0.67 (0.49, 0.91)				
Paclitaxel	0.43 (0.21, 0.86)	0.64 (0.29, 1.37)			
Irinotecan	0.48 (0.25, 0.91)	0.72 (0.35, 1.47)	1.13 (0.86, 1.49)		
mFOLFIRI	0.40 (0.17, 0.93)	0.59 (0.24, 1.47)	0.93 (0.50, 1.74)	0.83 (0.47, 1.45)	
Ramucirumab + paclitaxel	0.34 (0.17, 0.71)	0.51 (0.23, 1.13)	0.81 (0.68, 0.96)	0.71 (0.52, 0.99)	0.86 (0.45, 1.65)

Results are the HR with 95% CrI for the treatment A vs. treatment B

These results agree very closely with those presented in the Response to the Clarification Letter (Appendix 2, Table 3). We also checked the PFS base-case analysis and again matched the results reported in the submission.

In conclusion, apart from discrepancies between the studies in the data tables and WinBUGS code, the NMA analysis methods seem to be appropriate and the results were reproducible.

4.5 Additional work on clinical effectiveness undertaken by the ERG

We did not perform any additional work on clinical effectiveness ourselves. However, we will summarise the results from the NMA as reported in the Response to the Clarification Letter by the company, because we think these are most reliable.

Overall Survival

For overall survival there are two results from the NMA: with and without Roy 2013. In general it is better to include as much information as possible in a NMA, but in this case there is a second reason to include Roy 2013[28]. That is because the study by Thuss-Patience (2011)[30] was closed prematurely due to poor accrual, as a result only 40 patients were randomised instead of 120 patients. The results of the NMA for OS are reported in Table 4.26.

Table 4.26: NMA results for OS (including Roy 2013)

Treatment A	Treatment B				
	Placebo/BSC	**Docetaxel**	**Paclitaxel**	**Irinotecan**	**mFOLFIRI**
Docetaxel	***0.65 (0.48, 0.86)***				
Paclitaxel	***0.50 (0.31, 0.84)***	0.78 (0.49, 1.26)			
Irinotecan	***0.57 (0.37, 0.87)***	0.88 (0.59, 1.31)	1.13 (0.86, 1.48)		
mFOLFIRI	***0.47 (0.23, 0.95)***	0.73 (0.37, 1.45)	0.93 (0.50, 1.74)	0.83 (0.47, 1.45)	
Ramucirumab +paclitaxel	***0.41 (0.24, 0.70)***	0.63 (0.38, 1.05)	***0.81 (0.68, 0.96)***	***0.71 (0.52, 0.99)***	0.87 (0.45, 1.66)

BSC, best supportive care; Crl, credible interval; HR, hazard ratio; NMA, network meta-analysis. Results are the HR with 95% CrI for the treatment A vs. treatment B, Results where the CrI does not cross 1 are bold and italicised

Although we think the results for OS reported in Table 4.26 are the most reliable results given the available evidence, we need to stress that there is considerable uncertainty associated with these analyses. As reported above, the study by Thuss-Patience (2011)[30] is small and underpowered, which means the comparison of ramucirumab+paclitaxel versus BSC relies on four different studies, contributing to serious heterogeneity between studies. Especially the study by Hironaka (2013)[29] which was performed in a Japanese population causes significantly to the heterogeneity between studies. Therefore, all these results should be interpreted with great caution.

Progression Free Survival

For PFS it was not possible to calculate a result for ramucirumab+paclitaxel versus BSC because PFS was not reported in the studies by Thuss-Patience (2011)[30] and Roy (2013)[28]. The results of the NMA for PFS for the other comparisons are reported in Table 4.27.

Table 4.27: NMA results for PFS

	Treatment B			
Treatment A	**Docetaxel**	**Paclitaxel**	**Irinotecan**	**mFOLFIRI**
Paclitaxel	1.11 (0.76, 1.64)			
Irinotecan	1.26 (0.95, 1.69)	1.14 (0.88, 1.48)		
mFOLFIRI	1.05 (0.58, 1.89)	0.95 (0.53, 1.69)	0.83 (0.50, 1.39)	
Ramucirumab + paclitaxel	0.71 (0.46, 1.08)	***0.64 (0.54, 0.75)***	***0.56 (0.41, 0.76)***	0.67 (0.37, 1.22)

CrI, credible interval; HR, hazard ratio; NMA, network meta-analysis; PFS, progression-free survival
Results are the HR with 95% CrI for the treatment A vs. treatment B
Results where the CrI does not cross 1 are bold and italicised

Again, these results should be interpreted with great caution due to significant heterogeneity between studies.

4.6 Conclusions of the clinical effectiveness section

Ramucirumab monotherapy

Ramucirumab monotherapy compared to BSC relies on one trial (REGARD). This comparison is sufficient and in line with the NICE final scope if it is accepted that 'not suitable for paclitaxel' is the same as 'not suitable for further cytotoxic chemotherapy'. If this is not accepted, comparisons with cytotoxic chemotherapy other than paclitaxel (docetaxel, irinotecan and FOLFIRI) are missing.

In the REGARD trial ramucirumab monotherapy was associated with a 1.4 month increase in median overall survival in comparison to placebo. Median overall survival was 5.2 months among patients treated with ramucirumab and 3.8 months among those treated with placebo (HR=0.776, 95% CI: 0.603 to 0.998). The relative gain in PFS was 0.8 months for patients treated with ramucirumab with a median PFS of 2.1 months (95% CI: 1.5 to 2.7) in the

ramucirumab group and 1.3 months (95% CI: 1.3 to 1.4) in the placebo group (HR=0.483, 95% CI: 0.376 to 0.620).

An indirect comparison with docetaxel, using the COUGAR-02 trial shows that the hazard ratio of overall survival of ramucirumab versus docetaxel is not significantly different, but actually favours docetaxel (HR=1.16 (95% CI: 0.77 to 1.73)).

Time to progression was not measured in the active symptom control group of the COUGAR-02 trial because the authors decided that the value of measuring time to progression in a population not receiving cancer treatment but with known progressive disease at study entry was questionable. Therefore, the authors felt that it was not appropriate to subject these patients to additional unnecessary investigations. This means PFS could not be assessed in the COUGAR-02 trial.

Ramucirumab plus paclitaxel

Ramucirumab+paclitaxel was used in one trial, the RAINBOW trial, which compares ramucirumab+paclitaxel with paclitaxel. The trial shows favourable results for ramucirumab+paclitaxel in terms of overall survival (HR=0.807 (95% CI: 0.678 to 0.962)) and PFS (HR=0.635 (95% CI: 0.536 to 0.752)). This represents a 31% (2. 27 months) longer median overall survival in the ramucirumab+paclitaxel arm (9.63 months (95% CI 8.6-10.8) versus 7.36 (95% CI 6.3-8.4) months in the placebo+paclitaxel arm). The relative gain in PFS was 1.5 months for patients treated with ramucirumab+paclitaxel with a median PFS of 4.4 months (95% CI: 4.2 to 5.3) in the ramucirumab+paclitaxel group and 2.9 months (95% CI: 2.8 to 3.0) in the placebo+paclitaxel group.

Using indirect comparisons and network meta-analyses, the CS presents results comparing ramucirumab+paclitaxel with other relevant comparators such as docetaxel and best supportive care. In response to the clarification questions from the ERG, the company also presented results for ramucirumab+paclitaxel compared with irinotecan and (m)FOLFIRI. However, all these analyses rely on using data from a trial in a completely Japanese population (Hironaka 2013).[29] As explained in the CS, "high rates of salvage therapy have been reported in previous Asian trials.[32] A higher rate of post discontinuation (PDT) third-line therapy was expected in Region 3 (parts of Asia, including Japan), potentially confounding the OS treatment effect of the ramucirumab+paclitaxel regimen due to differing rates of PDT." In addition, the CS lists the following differences between Western and Asian countries:

Asian countries have a higher incidence and prevalence of Gastric cancer (GC) than Western nations.[33, 34] The adoption of national screening programs in Asian countries has resulted in diagnosis in the early stages of the disease in up to 50-60% of cases, while in Western nations patients are typically diagnosed at an advanced stage of GC and therefore have a poorer prognosis.[35]

Differences also exist in GC histology (Western patients have a higher incidence of diffuse histology/proximal tumour types having a poorer prognosis than intestinal histology/distal tumour types seen in Asian patients).[4] The surgical treatment of early GCs with extensive lymph node dissection (D2 resection) occurs more frequently in Asian countries.[4, 35] A higher

proportion of patients receive second-line chemotherapy (and beyond) in Asia compared to US and Europe which extends survival in those patients.[4]

As the network necessary to compare ramucirumab+paclitaxel with best supportive care and with other active treatments (docetaxel, irinotecan, and (m)FOLFIRI) relies on at least one trial in a multinational or Western population (RAINBOW,[9] Thuss-Patience 2011,[30] Roy 2013[28] and COUGAR-02[27]) and one trial in an Asian population (Hironaka 2013[29] and Sym 2013[26]), the network has a high level of heterogeneity and is therefore unreliable. Any results based on indirect comparisons or a network meta-analysis should be interpreted with great caution due to significant heterogeneity between studies.

5 COST-EFFECTIVENESS

5.1 ERG comment on manufacturer's review of cost-effectiveness evidence

5.1.1 Objective of cost-effectiveness review

A literature review was conducted to identify all economic studies relevant to the decision problem. The search strategies and ERG comments for the cost-effectiveness review are discussed in detail in section 4.1.1.

5.1.2 Inclusion/exclusion criteria used in the study selection

It was not clear to the ERG why cost studies, which may have provided input for the cost estimates in the model, were not included. The concern was raised in the clarification letter. The company explained in the clarification letter that the results of the literature search were reviewed for sources of costs and resource use but since all cost analysis were conducted in different countries, it was decided that the they were unlikely to contain information which would be applicable to English clinical practice.

5.1.3 Included/excluded studies in the cost-effectiveness review

The company identified 167 potentially relevant studies of which 11 remained after removing duplicates and reviewing titles and abstracts. Of these 11, no relevant economic analyses specific to the second-line setting were included.

Two papers,[36, 37] both adaptations of the analysis presented for TA 208 (trastuzumab for HER2-positive metastatic gastric cancer[38]), were identified from the additional targeted searches and included to inform the model structure.

5.1.4 Conclusions of the cost-effectiveness review

The ERG agrees with the conclusions from the manufacturer that none of the selected studies were relevant for the decision problem and that only the adaptations of the analysis presented for TA208 could be used to inform the models.

Table 5.1: Summary of the manufacturer's economic evaluation

	Approach	Source/Justification	Signpost (location in CS)
Model	Partitioned survival markov model with transition probabilities based on PFS and OS curves.		Section 5.2 (p.139)
States and events	Three health states are distinguished: progression-free, progressed disease or death.	The model structure and the health states are typical in the modelling of advanced/metastatic oncology and have been previously utilised in numerous NICE STAs and MTAs.	Section 5.2 (p. 140)
Comparators	For combination therapy: best supportive care (BSC) and docetaxel For monotherapy: BSC	No licensed therapies for second line besides ramucirumab. A chart review was conducted to explore real-world treatment patterns in previously treated GC/GOJ. Following inclusion criteria were applied: for additional comparators 1. Outlined in final scope 2. Routinely used in the NHS England (>10%) Even though docetaxel is used less than 10% threshold (8%) it is included due to the UK based COUGAR II study and clinical expert opinion	Section 5.2 (p. 144-148)
Natural history	Best supportive care arm represents the natural history of the disease. In the monotherapy population the BSC arm (i.e. natural history) is derived from the REGARD trial. In the combination therapy population, the treatment arm (RAM+PAC) is first estimated, and then the treatment effectiveness from indirect treatment comparison is used to estimate the BSC arm (i.e. natural history)	In combination therapy RAM+PAC PFS and OS were based on RAINBOW trial In monotherapy BSC and RAM PFS and OS were based on the REGARD trial.	Section 5.2 (p. 146)
Treatment effectiveness	Treatment influences PFS and OS compared to BSC	In combination therapy model, BSC and DOC PFS/OS were based on indirect treatment comparison applied on the (semi)parametric curves fitted to KM RAM+PAC data from RAINBOW trial. In the monotherapy model, both BSC and RAM PFS/OS were based on parametric curves fitted to KM data from REGARD trial.	Section 5.3 (p. 149-168)

	Approach	Source/Justification	Signpost (location in CS)
Model	Partitioned survival markov model with transition probabilities based on PFS and OS curves.		Section 5.2 (p.139)
Adverse events	Included in model as 1 time event, they have impact both on costs and QALYs	Based on observed grade 3/4 AEs with occurrence in more than 5% of the reference trial population of each comparator arm as well as on impact on costs and QALYs. REGARD is the reference trial for BSC and RAM, RAINBOW is the reference trial for RAM+PAC and COUGAR II is the reference trial of DOC.	Section 5.4 (p. 181) Section 5.5 (p. 210)
Health related QoL	Utility scores were assigned to the pre-progression and post-progression health states. Utility decrements were applied for adverse events and utility increments were applied for clinical response (ORR=PR+CR).	Utilities based on EQ5-D as administered in RAINBOW trial. UK weights are applied. Values for AE utility decrements and durations were taken from the literature. Values for utility increments due to response were also derived from RAINBOW trial,	Section 5.4 (p. 169)
Resource utilisation and costs	Treatment cost (adjusted for relative dosing intensity observed and including pre-medication and treatment administration), health state costs (Progression free and Progressed disease, accounting for follow-up care, consultation, hospitalization), terminal care, and costs of adverse events.	Based on UK reference costs, literature and expert opinion	Section 7.4.20 (p154) to 7.4.23 (p158).
Discount rates	A 3.5% discount rate was applied for both costs and effects	According to NICE reference case	Section 5.2 (p. 143)
Sub groups	An analysis using geographic subgroups was conducted.	PFS, OS and ToT were adjusted for Region1 (including North America, Europe, Australia and Israel) by including as a covariate.	Section 5.9 (p.240)
Sensitivity analysis	One-way deterministic sensitivity analysis, scenario analyses and probabilistic sensitivity analysis	Ranges based on observed confidence intervals and assumptions	Section 5.8 (p. 228)

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

The ERG has assessed the manufacturer's economic evaluation using the Philips et al checklist for quality assessing decision analytic models.[39] This is shown in Appendix 2 of this report and is used to assist the narrative critique in the following sections.

5.2.1 NICE reference case checklist (TABLE ONLY)

Table 5.2: Comparison of the CS model with the NICE reference case

Elements of the economic evaluation	Reference Case	Included in submission	Comment on whether de novo evaluation meets requirements of NICE reference case
Comparator(s)	Therapies routinely used in the NHS, including technologies regarded as current best practice	No	Except for DOC, many second-line treatment practices (e.g. PAC) were not included because they were used in <10% of all patients.
Type of economic evaluation	Cost-effectiveness analysis	Yes	
Perspective on costs	NHS and PSS	Yes	
Perspective on outcomes	All health effects on individuals	Yes	
Time horizon	Sufficient to capture differences in costs and outcomes	Yes	Time horizon is 7.23 years (lifetime), median OS ranges from 3.8[10] - 9.63[9] months.
Synthesis of evidence in outcomes	Systematic review	Yes	
Measure of health effects	QALYs Life-years	Yes	
Source of data for measurement HRQOL	Reported directly by patients and/or carers.	Yes, for Combo No, for mono	Both models obtained utility values from the RAINBOW trial with utility increments (response based) and decrements (AE based) from the literature.
Source of preference data for valuation of changes in HRQOL	Sample of public	Yes	
Discount rate	Annual rate of 3.5% on costs and health effects	Yes	
Equity weighting	No special weighting	Yes	
Sensitivity analysis	Probabilistic sensitivity analysis	Yes	

5.2.2 Model structure

The cost-effectiveness of ramucirumab+paclitaxel (RAM+PAC) and ramucirumab monotherapy (RAM) is evaluated using a partitioned survival model (which is essentially the same as a Markov model). The model structure for the combination therapy and monotherapy

economic models were the same. As shown in Figure 5.1 below, the model included three different health states; pre-progression, post-progression and death.

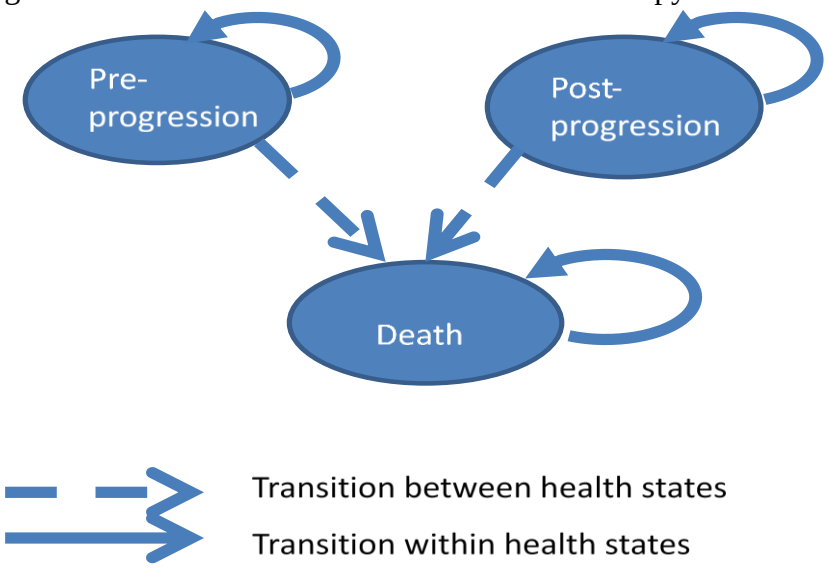
All patients enter the economic model in the pre-progression health state, which represents the initiation of second-line treatment or BSC only. Patients remain in this health state until they experience progression or death and receive treatment until disease progression or treatment cessation due to other reasons.

After progression, patients enter the post-progression health state. In this health state, the majority of the patients receive BSC while a small proportion of the patients receive a third line treatment. Patients remain in the post-progression health state until they die.

Death is an absorbing health state that patients enter from the pre- or post-progression health states.

Transition probabilities between the health states are determined from the RAINBOW and REGARD trial data for combination therapy and monotherapy, respectively and by the indirect comparison for the non-trial comparators. The model uses weekly cycle lengths and a half-cycle correction was applied to all calculations.

Figure 5.1: Economic model structure for monotherapy and combination therapy



ERG comment

The model structure is a common model structure for patients with cancer and, as stated in the CS, was used for first-line gastric cancer (TA208). The model is clearly described and straightforward. However, a single post-progression health state is a simplification of the treatment pathway that does not enable to distinguish (costs, effects and quality of life) between patients receiving third-line treatment and patients receiving BSC only.

Also, the way that the number of patients progressing was estimated (by calculating the number of patients leaving the pre-progression health state in a cycle) implicitly assumes that all patients leaving pre-progression move to post-progression, hence do not move to death. This is not made explicit in the above figure.

5.2.3 Population

The patient population as issued by the NICE scope included adults with advanced gastric cancer or gastro-oesophageal junction adenocarcinoma previously treated with chemotherapy. Ramucirumab is licensed for two indications, i.e. in combination therapy with PAC and as a monotherapy. The indications represent two different patient populations and therefore, the patient population will be described separately.

Combination therapy

Combination therapy is indicated for adult patients with advanced GC/GOJ with disease progression after prior platinum and fluoropyrimidine chemotherapy. The (primary) economic evaluation is based on the ITT population from the RAINBOW trial.[9] This trial included adult patients with confirmed metastatic or non-resectable, locally advanced gastric or oesophageal junction adenocarcinoma who had disease progression within 4 months of the last dose of first-line therapy with platinum and fluoropyrimidine doublet with or without anthracycline and who had an ECOG performance status score of 0 or 1 at study entry. Patients were enrolled from different countries. Sub analyses were performed on the patients enrolled in region 1 (i.e. North America, Europe, Australia, and Israel).

Monotherapy

Monotherapy is indicated for adult patients with advanced GC/GOJ with disease progression after prior platinum and fluoropyrimidine chemotherapy for whom treatment in combination with paclitaxel is not appropriate. According to the CS inappropriateness for treatment in combination with paclitaxel can be more broadly characterised as inappropriate for cytotoxic chemotherapy. The economic evaluation is based on the ITT population of the REGARD trial.[10] This trial included adult patients with confirmed metastatic or unresectable, locally recurrent gastric or oesophageal junction adenocarcinoma who had disease progression within 4 months or during first-line therapy with platinum- or fluoropyrimidine containing chemotherapy for metastatic disease, or within 6 months of the last dose of platinum-containing or flueropyrimidine containing adjuvant treatment and who had an ECOG performance status score of 0 or 1 at study entry.

The CS describes that the ITT population deviates slightly from the final licence wording as patients in the trial did not have to be unsuitable for treatment in combination with PAC upon entering the trial.

ERG comment

Combination therapy: 60% of the ITT population from the RAINBOW trial was from region 1 and therefore more similar to the UK population than the 7% from region 2 (Central and South America) and 33% from region 3 (Asian countries).

Although the trial specified additional inclusion criteria (e.g. resolution to adverse events of prior therapy, adequate organ and kidney function) and exclusion criteria (e.g. previous systemic chemotherapy with a cumulative dose of >900 mg/m2 of epirubicin or >400 mg/m2 of doxorubicin), the ITT population seems to reflect that of the licensed indication.

Monotherapy: The CS describes that the patient population in the economic model is aligned with the licensed population. The patient population in the economic model is based on the

REGARD trial. However, ineligibility for PAC combination therapy was not an inclusion criterion. The ERG asked to the company to provide additional information on the patients in the REGARD trial that were ineligible for combination therapy with PAC (Clarification letter question A7[11]). However, this could not be provided since the trial did not collect the data to identify this sub population and ineligibility for combination treatment with PAC is according to the company a matter of clinical judgement. Therefore, the ERG would cast serious doubt on whether the patient population used to inform the model aligns with the licensed population.

5.2.4 Interventions and comparators

Intervention: combination therapy

According to the license, ramucirumab is combined with paclitaxel (RAM+PAC) in 28 day cycles. RAM is administered on day 1 and 15 at 8mg/kg, PAC is administered on day 1, 8 and 15. Treatment cycles are recommended to be continued until disease progression or unacceptable toxicity.

Comparator: BSC

According to the CS, BSC is the most relevant comparator in the UK. Based on a review of treatment approaches in advanced gastric cancer[4] it can be concluded that second-line chemotherapy should be considered standard of care for patients as it provides a modest but significant improvement in survival in patients with good performance status. However, this review also states that second-line therapy is not standard care in the UK. Treatment in daily practice in the UK was studied in further detail in a commissioned chart review by the company.[3] Based on 200 patients with ECOG performance status 0, 1 or 2 at advanced stage diagnosis it was concluded that 28.5% (N=57) of the patients currently received an active second-line therapy, 71.5% received BSC.

BSC is not initiated to alter the natural history of the disease but provided to alleviate symptoms and maximise quality of life. The efficacy of BSC was obtained from the NMA and included three randomised Phase III trials.[4, 27, 30] BSC components and frequencies were obtained from the chart review commission by the company to ensure that BSC in the economic model reflects BSC in the UK. BSC consisted of morphine (62.9%), cognitive behavioural therapy (16.9%), blood transfusions (23.8%) and radiotherapy (11.9%).

Additional comparators

Based on the commissioned chart review by the company,[3] there is a big variation in second line treatment in clinical practice. DOC is the most commonly used active treatment as 8% of the patients receives this therapy. The NICE methods Guide[14] states specifically that non-licensed treatments can be considered by the appraisal committee when they are established clinical practice. To determine established clinical practice the company applied the following inclusion criteria for additional comparators, based on a previous technology appraisal (CS Ref 15) and verbally endorsement in the NICE/ABPI masterclass “How to do a good NICE submission”:

- Treatment is outlined in the final scope
- Treatment is routinely used in the NHS in England (>10% usage)

DOC

Although DOC did not meet the company's inclusion criteria it was included in the economic model by the company since the phase III study that investigated the efficacy of DOC was a UK study[27] and advice from their clinical experts confirmed that it is important in shaping clinical practice moving forward. Based on clinical trials included in the NMA, DOC was administered at a dose of 75 mg/m^2 every three weeks.

PAC

PAC was included for validation of the economic model as it was the comparator treatment in the RAINBOW trial and did not meet the standard inclusion criteria. Since PAC was only included in the model for validation of the clinical evidence, dosing was not reported in the CS.

IRI and FOLFIRI

IRI and FOLFIRI were listed in the final scope as potential appropriate comparators for RAM+PAC but were not included in the economic model by the company since their use was not sufficient according to the chart review.[3]

ERG comment combination therapy

The ERG agrees with the company that BSC and DOC are relevant comparators for RAM+PAC given the results from the literature and the commissioned chart review. However, the ERG does not agree with excluding comparators that were included in the final scope (i.e. DOC, IRI and FOLFIRI) based on the established use criteria. First, established NHS practice is already incorporated as criteria for defining the most appropriate scope. Second, the 10% is not a formal rule, although it was used previously, and also not implemented consequently in the CS. Third, the proportion of treated patients is very low and therefore the proportion of patients receiving certain comparators will always be low when it is calculated over all patients that progressed after first-line therapy. According to the CS the chart review illustrated that PAC was used for 3% of the patients, but this proportion is 10.5% if the number of patients that receives PAC as second-line therapy is divided by the total number of patients that receives second-line therapy (instead of by all the patients that progressed after first-line therapy). One might argue that it is more likely that the group of patients currently receiving a second-line treatment (who are in general in better health) will start treatment with RAM+PAC compared to the patients who currently receive BSC.

But also, since the company claims that RAM+PAC is a suitable treatment option for patients who currently receive BSC, the ERG considers it plausible that also the proportion of patients receiving PAC or other comparator treatments may increase once RAM+PAC receives positive guidance.

Finally, the chart review was based on data provided from June to July 2013. Favourable results for DOC from the COUGAR II study were published in December 2013. Although PAC is still not a licenced treatment for the current indication, increasing utilisation may be expected since results were available from December 2013 onwards.

In the wake of this, it may also be anticipated that off-label use of IRI and FOLFIRI could increase, and hence the inclusion of these treatments in the comparison could also be considered relevant.

Intervention: monotherapy

Ramucirumab is administered every two weeks at 8 mg/kg until disease progression or unacceptable toxicity. The licence specifies monotherapy for patients in whom combination with PAC is not appropriate. The final scope included BSC, DOC, IRI and FOLFIRI as comparator. However, according to the company BSC is the only comparator that should be included since in clinical practice, the licence for RAM is such that it is for use in a patient population who are currently receiving BSC.

ERG comment

The ERG agrees with the company that BSC is a relevant comparator. For the same reasons as mentioned for the combination therapy, the ERG is not convinced that other therapies are not appropriate comparators, unless 'inappropriate for paclitaxel' can be interpreted as 'inappropriate for further cytotoxic chemotherapy'. This is notwithstanding the view of the ERG that there are probably no data to inform any of these comparisons in the population eligible for monotherapy, in particular because the REGARD trial population might not be appropriate, since the eligibility criteria for RAINBOW and REGARD trial were more or less the same.

5.2.5 Perspective, time horizon and discounting

The model has a NHS perspective and a lifetime time horizon (which equates to 7.23 years) for both combination therapy and monotherapy. Costs and effects were discounted at an annual 3.5% discount rate

ERG comment

The ERG concludes that the discount rate and perspective are in line with the NICE reference case.

5.2.6 Treatment effectiveness and extrapolation

Progression-free survival (PFS) and overall survival (OS) probabilities were incorporated into the model using survival curves. As in the CS, treatment effectiveness and extrapolation are discussed first for the combination therapy model, and afterwards for the monotherapy model (for patients not eligible for a second-line combination therapy with PAC).

In the combination therapy model, survival analysis methods were carried out to incorporate the effects of the RAM+PAC treatment arm from the RAINBOW trial. For other comparators (BSC or DOC), results from the indirect treatment comparison/NMA were applied. PAC is included in the combination therapy model only for the validation purposes. In the monotherapy model, survival analysis methods were carried out to incorporate the effects of the BSC and RAM treatment arms from the REGARD trial. Since, according to the CS, there were no other relevant comparators and/or studies for patients not eligible for PAC, indirect comparisons or network meta-analysis were not conducted for the monotherapy model.

For both the combination therapy and monotherapy models, the following five parametric models were initially considered for goodness of fit to PFS and OS data from the RAINBOW

and REGARD trials: exponential, Weibull, lognormal, log-logistic and Gamma. However, the Gompertz distribution was excluded as it is not supported by the software used (i.e. SAS) for all the survival regression analyses described in the CS (p150). The following steps were followed for the survival modelling of PFS and OS:

1) Assessing for the functional form of the underlying hazard (e.g. checking the validity of the proportional hazards (PH) assumption)
2) Conducting goodness of fit tests and assessing the suitability of each parametric distribution
3) Selecting the most appropriate distribution

Akaike/Bayesian Information Criteria (AIC/BIC) statistics, Cox-Snell residuals, log-log hazard plots and visual inspection methods were applied for assessing the suitability of each parametric survival model during the observed trial period. During the extrapolation period, visual inspection method was used to assess how well the tail of the parametric function agreed with clinically expected outcomes or long term data if external data was available.

ERG comment

In general, the steps outlined in the DSU report on the extrapolation of survival curves[40] were followed in the CS, however the choice of the survival modelling does not follow the same procedure for all PFS and OS curves in combination therapy and monotherapy models. For instance in the combination therapy for PFS, the Weibull distribution was chosen for the sake of the PH assumption although other distributions had demonstrated a better fit in terms of AIC/BIC tests and Cox-Snell residuals. OS and PFS survival modelling will be thoroughly discussed in the next subsections below.

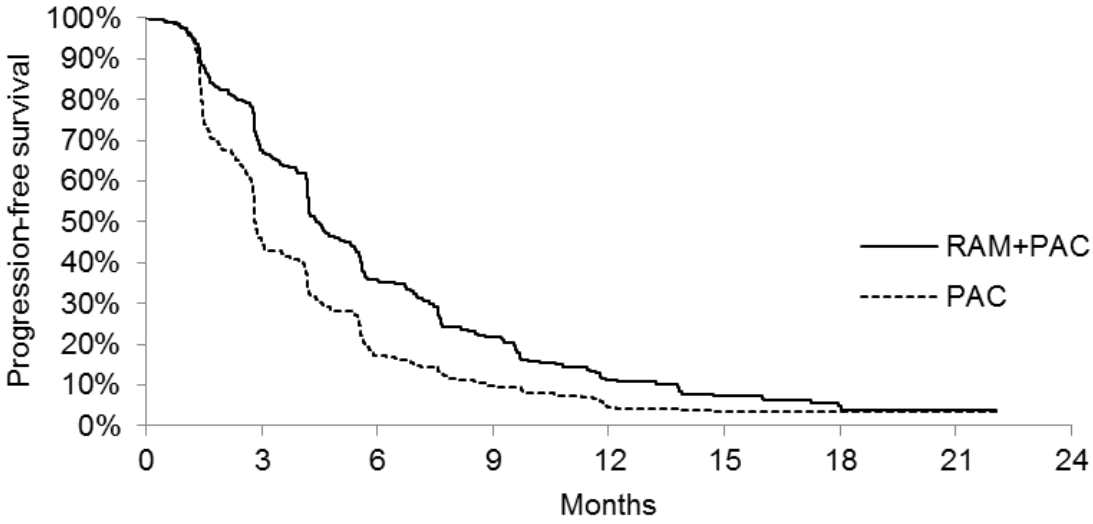
The ERG team understands the software incompatibility issues and recognises the attempts made to overcome the software ineligibility issues in the CS. However, the ERG would also like to point out that for example STATA can fit Gompertz distributions to survival data.

5.2.6.1 Combination Therapy Model (RAM+PAC)

Progression Free Survival (PFS)

The model used investigator assessed PFS data from the RAINBOW trial whereby progression was confirmed by radiological assessment. Collected PFS data were mature (less than 4% non-progressed at the end of the trial in both arms). The PFS Kaplan-Meier curves are presented in Figure 5.2

Figure 5.2: PFS Kaplan-Meier plots from RAINBOW



A stepped shape is observed in the PFS KM curves, which is caused by interval censoring (patients were assessed every six weeks). Therefore, parametric curves which can incorporate potential interval censoring were considered by the company rather than using the KM data.

The underlying hazards for PFS were assessed using log-log plots (CS Figure 34) and cumulative hazard plots (CS Figure 35). Some evidence for partial violation of PH assumption was detected in the log-log plots, but not in the cumulative hazard plots. In the CS, possible explanations were given such as censoring in the tails and overlapping of the KM curves in the first month. Given the presence of interval censoring, the PH assumption was considered plausible even though "an assessment of PH using the observational data should be taken with caution".

In the CS, it was mentioned that the parametric curves which took interval censoring into account had provided a more plausible fit to the trial data. In Appendix 9 of the CS (Figure 6 and Figure 7), overlays of the KM curve with interval censoring adjusted and non-adjusted parametric curves were presented for the PBO+PAC arm. From these figures, it could be seen that interval censoring adjusted parametric curves intersect the KM curve at the lower points of the steps. This was deemed to be clinically more plausible and avoids underestimating the hazard associated with progression/death, therefore, only interval censoring adjusted parametric curves were considered further by the company. This observation was reaffirmed by the same set of figures for RAM+PAC arm, which were asked for (and received) from the company in the clarification letter.

Goodness of fit tests showed that among the parametric models, lognormal, log-logistic, and gamma distributions provided the best fit to the PFS data from the RAINBOW trial as can be seen in Table 5.3.

The authors stated that Cox-Snell residual plots given in Appendix 9 of the CS (Figure 8), confirmed the findings of the AIC and BIC tests, and had indicated that joint parametric models with log-logistics and lognormal distribution provided the best fit to the underlying

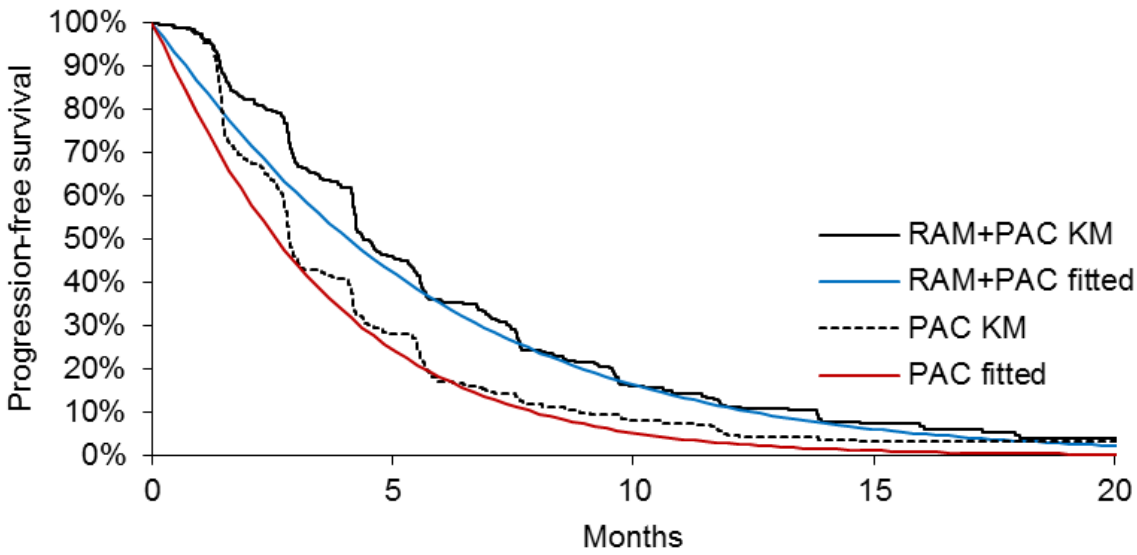
hazard, while proportional hazard models with Weibull and exponential distributions also provided a reasonable fit according to the CS.

Table 5.3: AIC and BIC statistics for PFS for combination therapy model.

Model	N	LogL	DF	AIC	BIC	AIC ranking	BIC ranking
Exponential	665	-1181.50	2	2367.00	2376.00	5	4
Weibull	665	-1179.76	3	2365.51	2379.01	4	5
Lognormal	665	-1165.47	3	2336.95	2350.45	3	2
Log-logistic	665	-1161.50	3	2329.01	2342.51	1	1
Gamma	665	-1164.27	4	2336.54	2354.54	2	3

In the end, the Weibull distribution was chosen in the base case to parameterize the PFS of the RAM+PAC treatment. The company motivated this choice by the fact that an indirect treatment comparison was used to estimate the treatment effects (a hazard ratio) of BSC and DOC, and Weibull is preferred to make the comparison. As a scenario analysis, log logistics was used. Base case PFS parametric curves from the trial for the combination therapy model can be seen in Figure 5.3.

Figure 5.3: Base case PFS estimated using Weibull distribution for combination therapy



ERG comment

The ERG can understand the reasons for interval censoring adjustments. However, from the CS it was not clear to the ERG how this interval censoring was adjusted. Also from Figure 5.3, it seems that the interval censoring adjusted curve approach slightly underestimates the PAC+PBO arm more compared to the RAM+PAC arm.

The PH assumption was justified in the CS, despite evidences suggesting violation, on several grounds such as censoring in the tails, overlapping of KM curves in the first month and interval censoring. Since KM plots were limited due to interval censoring, true observed

event times were not reflected, which prevents the ERG team from making claims on the validity of any assumption on the PH property of the KM curves.

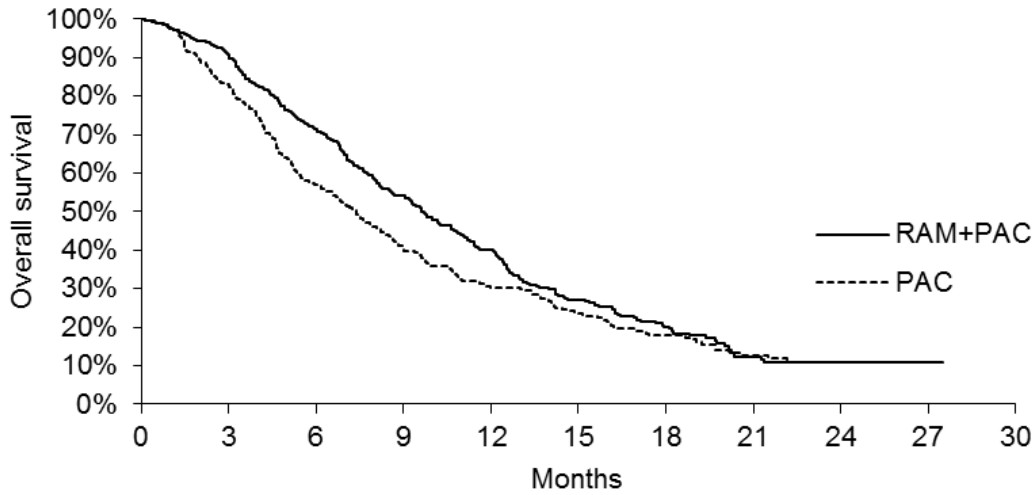
Even though PH is assessed only between PAC and PAC+RAM, it is assumed to hold for all other comparators' PFS curves (e.g. BSC or DOC).

Based on the AIC/BIC statistics and Cox-Snell residual plots, the log-logistic distribution was the best fit; however, the Weibull distribution was chosen to parameterise PFS in the base case as it was deemed to be more appropriate to the way the hazards estimated from ITC are applied (HRs were applied as in the PH model). According to the ERG, this choice is unnecessary, as well as conflicting with the approach taken for the survival modelling of OS curves, in which the best fitting approach was targeted without any consideration on whether the chosen distribution can be represented as a PH type model or not.

Overall Survival (OS)

The model used OS data from the RAINBOW trial. The data were quite mature as survival in both arms was about 10% by the end of the follow up. The OS Kaplan-Meier curves for the ITT population are presented in Figure 5.4.

Figure 5.4: Overall survival Kaplan Meier plots for ITT population



The underlying hazards for OS were assessed using log-log plots (CS Figure 38), which demonstrated a change in the hazard rate over time in both the RAM+PAC and PAC+PBO arms, indicating a potential violation of the PH assumption.

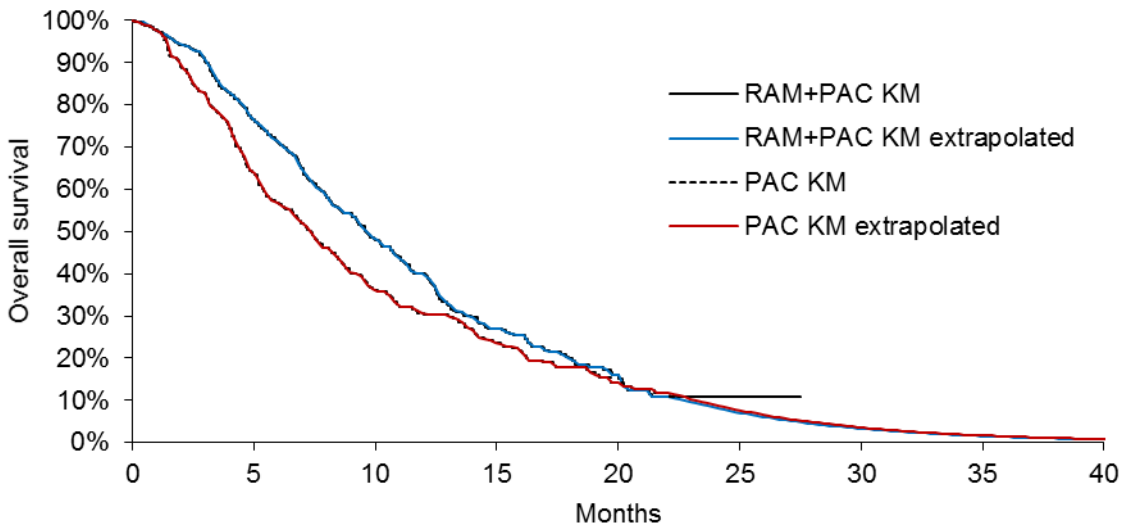
After the assessment of the underlying hazards, the standard approach of jointly fitting parametric distributions to OS KM curves was attempted and in the CS it was stated that visual inspection had shown that all the parametric models did not fit the KM data well and overestimated the survival benefit of RAM+PAC arm (see Appendix 9 in the CS, Figures 9 to 13). Therefore joint modelling of the parametric functions was not considered further. Next, parametric curves were fitted independently to each arm of the trial. Goodness of fit was first explored using AIC and BIC tests for both RAM+PAC (CS Table 63) and PBO+PAC (CS Table 64) arms. According to the AIC and BIC tests, log-logistic was the best fitting

distribution for the PBO+PAC arm, and Weibull was the best fitting distribution for the RAM+PAC arm. In order to prevent differing functional forms of hazard between two arms (clinically not plausible), fitted independent models were forced to have the same distribution for both arms. When Cox-Snell residuals of RAM+PAC (CS Figure 14 in Appendix 9) and PBO+PAC (CS Figure 15 in Appendix 9) arms were inspected visually, there was evidence that parametric models with Weibull and log-logistic distributions were deviating from the KM curves towards the end.

In order to address the concerns with the independent parametric modelling estimates, an alternative modelling approach was explored. In this approach, OS KM data for the trial period were used and then extrapolated from the end of the trial until the point at which all patients are believed to have died via an exponential distribution, which assumed no difference in the hazard between the RAM+PAC arm and the PBO+PAC arm. The hazard rate of the exponential distribution used in the extrapolation was estimated using two time points. The point at which the extrapolation was implemented was determined by the time at which the last event was observed in the PBO+PAC arm (22.14 months). The second point used was the time at which survival is 0.1% in the PBO+PAC arm when extrapolated using the Weibull distribution (53.5 months). Different extrapolation starting points were explored in the sensitivity analyses.

In the end, the alternative approach explained above (KM data+exponential extrapolation) was chosen in the base case to estimate the OS of the RAM+PAC treatment. In the CS this choice was motivated by the fact that this approach was the most conservative one. As scenario analyses, different cut points for extrapolation and independent modelling of log logistic and Weibull distribution were also implemented. Base case OS curves from the KM trial and exponential extrapolation can be seen in Figure 5.5.

Figure 5.5: Base case OS for RAM+PAC and PBO+PAC using the Kaplan-Meier data with exponential extrapolation



ERG comment

In the CS it is mentioned that parametric models fitted by the joint modelling approach had overestimated the survival benefits. According to the ERG Figures 9-13 in the Appendix 9 of the CS do not present a uniform overestimation of the survival benefit. However, the ERG agrees that none of the distributions provided a plausible fit to the OS KM curves.

Conclusions on the poor fit of parametric functions from independent modelling based on the AIC/BIC statistics and Cox-Snell residual plots seem reasonable. According to the ERG, the Kaplan-Meier curve with exponential extrapolation is more plausible than other parametric functions fitted. Nevertheless, there are minor concerns in this approach like selecting the time point which at which survival is 0.1% is based on the Weibull distribution. However, the sensitivity analyses conducted in CS showed that the impact of the chosen time for extrapolation was minimal on ICER.

Application of treatment effect alternative comparators

In order to conduct the comparison of RAM+PAC to BSC and DOC, the HR estimates from the NMA for BSC and DOC compared to RAM+PAC were applied to the baseline curves for RAM+PAC. An assumption of PH is made across the trials included in the evidence network to allow this methodology to be implemented.

For PFS, the evidence network presented in Figure 27 in the CS was disconnected. PFS HRs were not reported in some of the key studies for the comparators. In order to connect the network, an equal hazard rate assumption (HR=1) is made between RAM and irinotecan (IRI), between IRI and DOC and between DOC and RAM. This assumption is based on the naïve comparison of the PFS KM estimates from Roy et al[28] and the REGARD trial which yielded comparable median PFS estimates for BSC, RAM and DOC. A standard error of 0.01 was used in the base case for the assumed HRs equal to 1. Different standard errors were explored (0 and 0.278) in scenario analyses. The PFS HR estimates resulting from the NMA on the final evidence network (with assumptions) are given in Table 5.4 below:

Table 5.4: HRs (95% CrI) for PFS obtained from NMA

Treatment	NMA using SE = 0.01; HR (95% Crl)
RAM+PAC	-
BSC	3.7(1.89, 7.14)
DOC	1.79 (1.32, 2.44)

For OS, the evidence network presented in Figure 26 in the CS was used. As all relevant comparators were connected to the network no further assumptions were needed. The OS HR estimates resulting from the indirect treatment comparison are given in Table 5.5 below:

Table 5.5: HRs (95% CI) for OS obtained from ITC.

Treatment	NMA using SE = 0.01; HR (95% Cl)
RAM+PAC	-
BSC	2.94 (1.41, 5.88)
DOC	1.96 (0.88, 4.35)

ERG comment

In the indirect treatment comparison for PFS, several assumptions on the HRs were made. However a standard error of 0.01 is not based on any evidence. Therefore the ERG thinks a standard error of 0.287 based on Roy et al[28] may be a better approach.

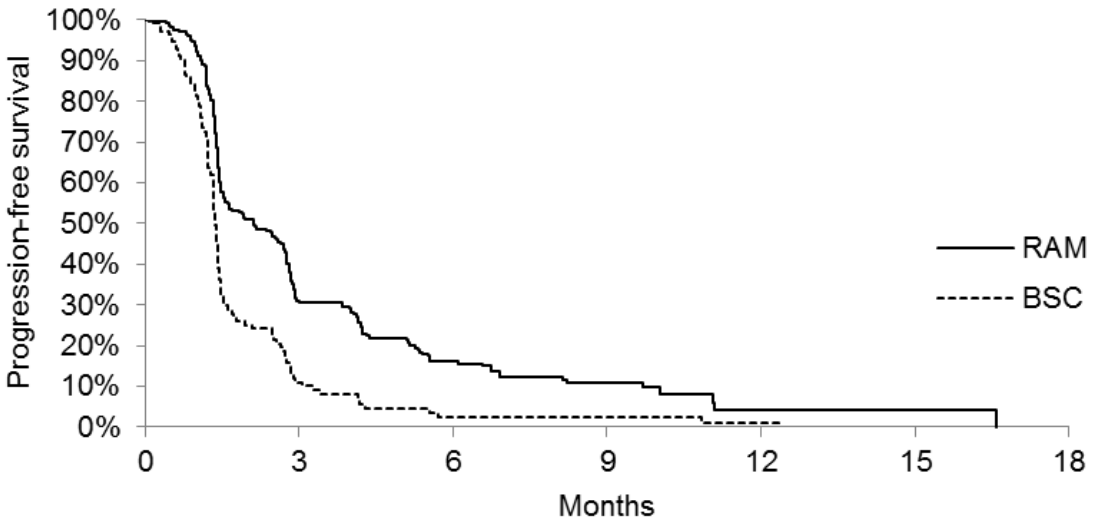
For the indirect treatment comparison for OS, in the base case, results from the evidence network excluding Roy et al[28] is used. Roy et al 2013 is a randomised phase II study including three treatment arms (i.e. DOC (N=44), IRI (N=44) and a highly stable liposomal nanocarrier formulation of irinotecan (N-44)). Therefore, the ERG thinks that NMA results including Roy et al 2013 may be more reliable. Including Roy et al 20132 in the OS NMA for combination therapy yields more favourable results in terms of HR for the comparators of RAM+PAC. Therefore, in one of the additional scenario analyses, results from the NMA including Roy et al 2013[28] will be used. Also in another additional scenario analysis, PFS HR from Roy et al[28] is used (1.26) instead of 1 between IRI and DOC.

5.2.6.2 Monotherapy model (RAM)

Progression free survival (PFS)

The model used investigator assessed PFS data from the REGARD trial. Collected PFS data were mature (less than 1.5% non-progressed at the end of the trial in both arms). The PFS Kaplan-Meier curves are presented in Figure 5.6

Figure 5.6: PFS Kaplan-Meier plots from REGARD



In the REGARD trial, similar to the RAINBOW trial, a stepped shape is observed in the PFS KM curves, which is caused by interval censoring (patients were assessed every six weeks). Therefore, parametric curves which can incorporate potential interval censoring were considered rather than KM data.

The underlying hazards for PFS are assessed using log-log plots (Figure 44 of the CS). No clear violation of PH assumption was detected in the log-log plots.

In Appendix 9 (Figure 17 and Figure 18) of the CS, overlays of KM curve with interval censoring adjusted and non-adjusted parametric estimates were presented for the PBO+BSC arm. From these figures, it could be seen that interval censoring adjusted parametric estimates

intersect KM curve at the lower points of the steps. Similar to the combination therapy, this was deemed to be clinically more plausible and less risky to underestimate the hazard associated with progression/death, therefore, only interval censoring adjusted parametric curves were considered further.

Goodness of fit tests showed that among the parametric models, lognormal, log-logistic, and gamma distributions provided the best fit to the PFS data from the REGARD trial as can be seen in Table 5.6.

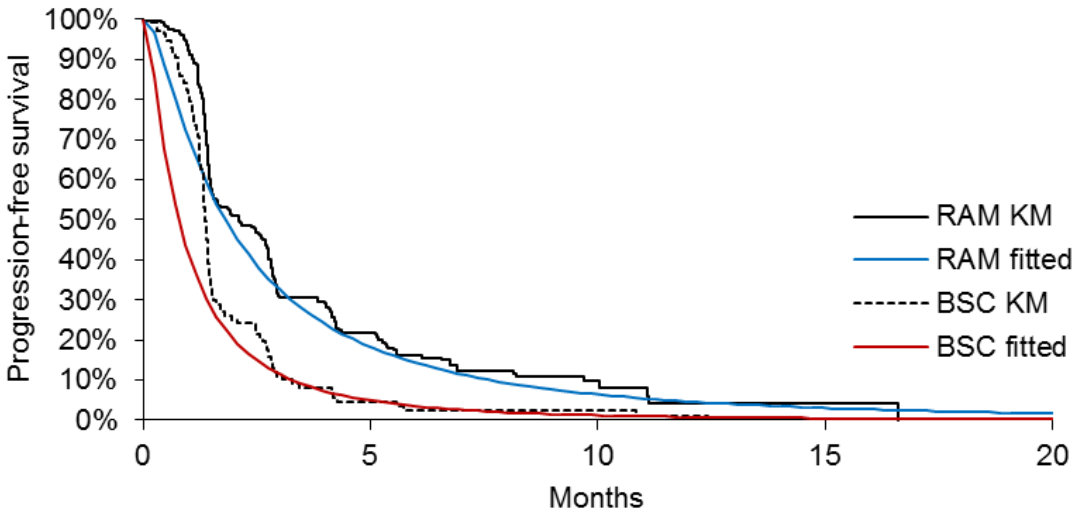
Table 5.6: AIC and BIC statistics for PFS for monotherapy model.

Model	**N**	**AIC**	**BIC**	**AIC ranking**	**BIC ranking**
Exponential	355	946.99	954.74	5	5
Weibull	355	941.91	953.53	4	4
Lognormal	355	914.62	926.23	2	2
Log-logistic	355	913.99	925.61	1	1
Gamma	335	915.24	930.73	3	3

The authors stated that Cox-Snell residual plots given in Appendix 9 of the CS (Figure 20), also had reaffirmed the findings of AIC and BIC tests, and had indicated that joint parametric models with log-logistics and lognormal distribution provided the best fit to the underlying hazard.

In the end, lognormal distribution was chosen in the base case to parameterise the PFS of the RAM and BSC treatment arms. Lognormal was superior to other distributions because of shorter tails of extrapolation, which is more conservative. Log-logistics and Weibull distributions were included as scenario analyses. Base case PFS parametric curves from the trial for the monotherapy model can be seen in Figure 5.7.

Figure 5.7: Base case PFS estimated using lognormal distribution for monotherapy



ERG comment

Similar to the combination therapy, the ERG can understand the reasons behind interval censoring adjustments, but which approach had been followed in interval censoring adjustments was not clear.

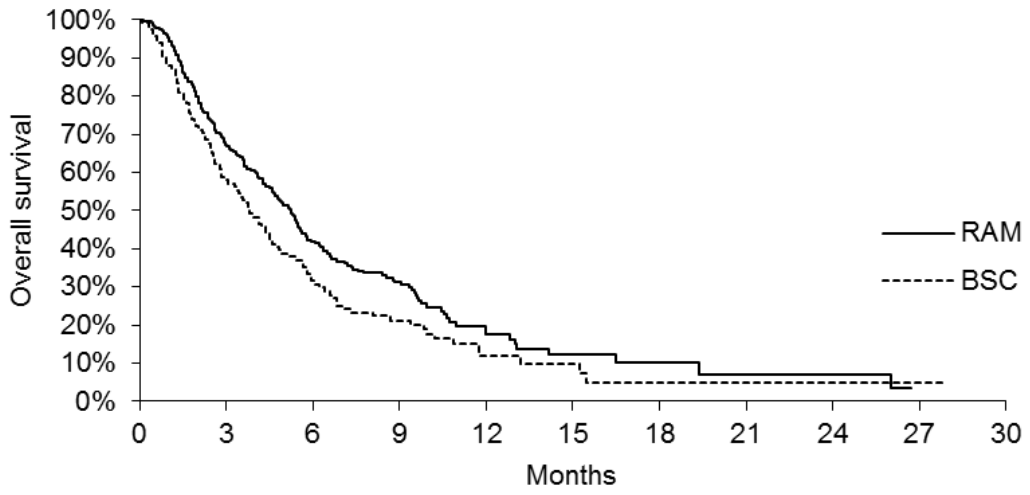
Although no clear violation is observed, commenting on the validity of any assumptions on the PH property of the KM curves is difficult, since KM plots were limited due to interval censoring.

The lognormal distribution was selected as the base-case for clinical plausibility. However, in the Cox-Snell residuals in Figure 20 in Appendix 9 of the CS, the log-logistic distribution demonstrated a smaller deviation than the lognormal. Considering AIC/BIC fit and Cox-Snell residuals, the log-logistic distribution might have been a more appropriate choice; however, from Figure 18 and 19 in Appendix 9 of the CS, it can be seen that log-normal and log-logistic parametric PFS estimates of RAM and BSC are almost the same.

Overall survival (OS)

The model used OS data from the REGARD trial. The data were quite mature as survival in both arms was less than 5% by the end of the follow up. The OS Kaplan-Meier curves for the ITT population are presented in Figure 5.8.

Figure 5.8: Overall survival Kaplan Meier plots for ITT population



The underlying hazards for OS were assessed using log-log plots (Figure 47 CS,), which demonstrated no clear evidence of PH violation.

Goodness of fit tests showed that among the parametric models, lognormal, log-logistic and gamma distributions provided the best fit to the OS data from the REGARD trial as can be seen in Table 5.7.

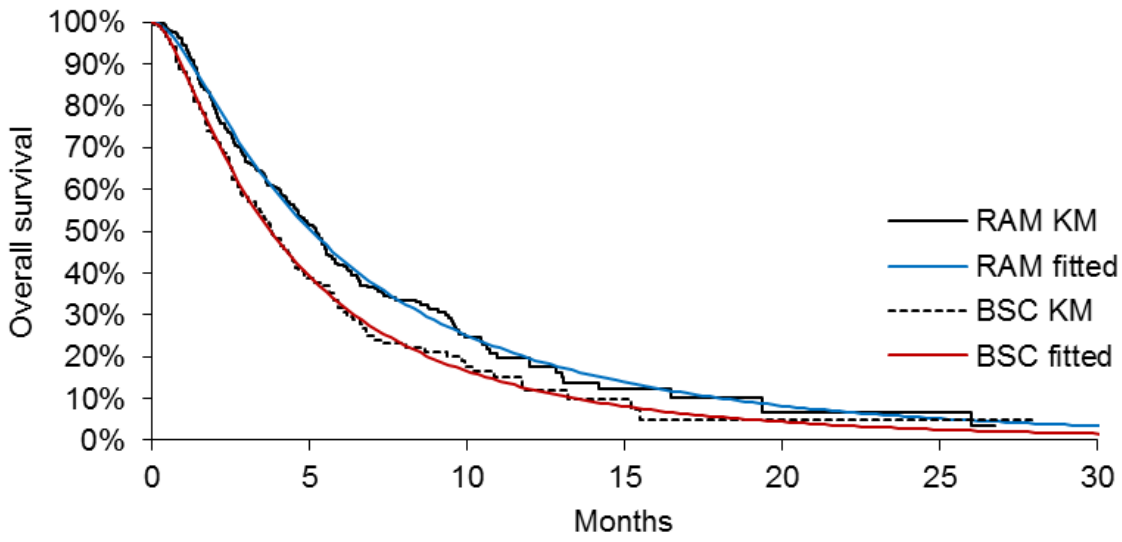
Table 5.7: AIC and BIC statistics for OS for the monotherapy model

Model	N	AIC	BIC	AIC ranking	BIC ranking
Exponential	355	987.05	994.79	5	5
Weibull	355	982.35	993.96	4	4
Lognormal	355	964.14	975.76	2	2
Log-logistic	355	963.82	975.44	1	1
Gamma	355	964.51	980.00	3	3

Cox-Snell residuals (CS, Figure 21 in Appendix 9) supported the findings from the AIC and BIC tests, with some minor divergence at the end of the trial period.

In the end, the Gamma distribution was chosen in the base case to estimate the OS of the RAM+BSC and PBO+BSC treatments. Authors motivated this choice due to the fact that Gamma provided a better fit and was the most conservative. As a scenario analysis, log-normal distribution was also included. Base case OS extrapolations for the monotherapy model can be seen in Figure 5.9.

Figure 5.9: Base case OS extrapolations for monotherapy model (gamma distribution)



As BSC was considered to be the only relevant comparator, (which was the trial comparator in REGARD), no indirect treatment comparison was required for the monotherapy model.

ERG comment

Similar procedures are followed in the monotherapy model, and the Gamma distribution is chosen as the base case. Based on AIC/BIC test results, Cox-Snell residuals and visual check, the ERG finds the Gamma distribution for OS modelling of BSC and RAM monotherapy plausible.

In the final scope, other comparators than BSC were mentioned as well (DOC, IRI, FOLFIRI), however in the CS, these comparators are excluded. For these comparators' PFS estimates, the ERG will use the estimates calculated for the combination therapy model. For OS estimates, HRs of DOC, IRI and FOLFIRI with respect to BSC will be applied to the OS estimated curve from REGARD.

5.2.6.3 Transition probabilities in combination and monotherapy models

Transitions between health states are estimated by parametric survival functions, which are used to estimate the proportion of the cohort remaining free of endpoints that defined the three health states - pre-progression, post-progression and death. Time in the 'pre-progression' state was estimated directly from the PFS curve. Time in the 'post-progression' state was estimated by the difference between the PFS and the OS curve at each time point. As the PFS and OS curves were modelled independently it was ensured that the PFS curve does not lie above the OS curve.

ERG Comment

In the clarification letter, question 21,[11] the ERG asked about the underlying assumption in the calculation of the newly progressed patients. The way it is calculated in the CS (newly progressed at time *t* = Patients at pre-progression state at time *t* – Patients at pre-progression state at time *t-1*) is an upper limit of the percentage of newly progressed patients, as it assumes that all patients leaving the pre-progression health state move to the progressed health state. In reality (and also according to the conceptual structure of the model) patients may also die whilst in the pre-progression state. Since this percentage is used only in a couple of cost calculations related to third line, the impact of this assumption on costs/ effects and ICER are minimal.

5.2.7 Health related quality of life

5.2.7.1 Literature review

Advanced GC/GOJ has a significant effect on patient's health related quality of life (HRQL). Inability to swallow, early satiety, weight loss, fatigue and breathlessness are some of the most commonly reported symptoms.[41]

A systematic review in patients with metastatic esophagogastric cancer (Jan 1988-Jun 2008)[42] was identified by the company. This review investigated the impact of chemotherapy on HRQL. According to this review, HRQL scores at baseline did not improve after chemotherapy. Nevertheless, chemotherapy was able to maintain HRQL. Improvements in HRQL were correlated with the efficacy of treatments but not to increased toxicity. This could be explained by the rapid progression of the disease; if chemotherapy fails, tumour related-symptoms and deterioration in ECOG will immediately impact HRQL.

The company conducted a systematic review on 25 September 2013 to identify a suitable source of utility data for the model. The review found nine studies which met the inclusion criteria. After review, three studies[43-45] were consistent with the NICE reference case i.e. health states were described using a validated generic questionnaire and valued using appropriate social preferences. The patient populations in these studies were patients with GC (without metastases), chemotherapy naïve advanced metastatic GC, and advanced or metastatic gastrointestinal disease but according to the company, not relevant for the decision problem.

ERG Comment

The ERG feels that the search date of September 2013 is outdated; an update search could have potentially retrieved additional relevant results.

5.2.7.2 Evidence available from RAINBOW and REGARD

The RAINBOW trial included both the EQ-5D and the EORTC QLQ C-30 (version 3). Data were collected at baseline and then every six weeks and at the end of treatment. 60% of the patients completed the 30-day post discontinuation EQ-5D questionnaire. Data collection of the EORTC QLQ C-30 was similar to the EQ-5D (i.e. at baseline and then every six weeks). Although the CS reports a trend towards improvement for RAM+PAC for all components except diarrhoea, significantly improvement was only observed for the two components emotional functioning and nausea and vomiting. Time to deterioration for diarrhoea was significantly better for PBO+PAC (HR 1.333, 95% CI 1.007-1.764).

In the REGARD trial, HRQL was only measured with the QLQ C-30, the EQ-5D was not included. Mapping was not performed by the company. First, this was considered second best to direct use of EQ-5D. Second, this would require data imputation since there were insufficient post-baseline data available due to the rapid disease progression in both arms.

Therefore the EQ-5D data from the RAINBOW trial was used in both the combination therapy and monotherapy model.

ERG comment

Utility values for the monotherapy model were based on utilities from the combination therapy trial. While the company claims that there was insufficient data in the REGARD trial due to rapid disease progression, the ERG would like to emphasise that the lack of data is caused by both the rapid disease progression and the administration of the questionnaires. Since these periods were rather long (i.e. six weeks), more people had already progressed before the first post-baseline questionnaire. This could have been resolved by administering the questionnaire every week or per two weeks.

Furthermore, the ERG would like to note that the EQ-5D data as used from the RAINBOW trial does not reflect the treatment options BSC and RAM nor the patient population for monotherapy (i.e. patients for whom treatment in combination with paclitaxel is not appropriate).

5.2.7.3 Pre-progression health state

HRQL is assumed to be constant in the pre-progression health state. The utility value for the pre-progression health state (0.737) was calculated by the baseline mean EQ-5D index score for the ITT population of the RAINBOW trial and applying UK weights.[46]

According to the CS, AEs and tumour response are clinically meaningful events that impact HRQL, but are not captured in the health state utility value used. Baseline utilities, obtained from the RAINBOW trial, were adjusted for treatment-related AEs in both the combination therapy model and monotherapy model. Adjustments for response rate (CR/PR) were made in the combination therapy model and applied from the first cycle onwards. No adjustment for response rate was made in the monotherapy model since the response rates to RAM and BSC were very small (i.e. 3.4% and 2.6%, respectively).

ERG comment

The ERG agrees with the company that AEs and tumour response are clinically meaningful events that may impact HRQL. Including the utility increments for responders from the first

cycle onwards (slightly) overestimates the QALYs obtained by RAM+PAC and DOC compared to BSC because usually it takes time before the response is established.

Utility increments and decrements combination therapy model

Response rates

The utility increment for responding patients was obtained from the change in utility at cycle 2 from baseline for all patients who experienced a response in the trial (+0.014). This increment was applied to the proportion of patients who responded (27.9% for RAM+PAC in the RAINBOW trial) from the first cycle onwards for simplicity. It was assumed that the response rate for DOC was similar to the response rate of PBO+PAC in the RAINBOW trial (16.1%). No response rate was applied to BSC since the response rate observed in trials was very low (2.6% in REGARD trial).

Adverse events

The type of AEs included in the model were based on the AEs (grade 3 and 4) that occurred in >5% of patients in either arm of the RAINBOW trial or in >5% of patients in the BSC and DOC arms of the relevant clinical trials.[10, 25, 27] According to the appendix, AEs occurring in >5% were reviewed in their expected costs and QoL impact. So although leukopenia (17.4% / 6.7%), asthenia (5.5% / 1.8%) and malignant neoplasm (14.4% / 17.9%) also occurred in more than 5% of the patients in either arm of the RAINBOW trial, these were not included in the model. From the REGARD trial dyspnoea (6.1%) was excluded and from the COUGAR trial, the constitutional symptoms (16.05%), gastrointestinal (25.93%), metabolic or laboratory (6.17%) and pulmonary or upper respiratory (6.17%)). Table 5.8 presents the reasons for exclusion.

Table 5.8: Events observed in the comparator treatments excluded from the analysis

Treatment	Adverse events	Description of treatment pathway/symptoms
Asthenia	Assumed to be capture in G3-4 fatigue	Asthenia
Malignant neoplasm progression	Assumed to be capture in disease progression within the model	Malignant neoplasm progression
Abdominal pain upper	Assumed to be capture in G3-4 abdominal pain upper, which is already included within the model	Abdominal pain upper
Leukopenia	Subtype of neutropenia	Leukopenia
DOC	All blood/BM AEs	Aggregation of blood-related AEs, individual blood-related AEs are also reported in study
	Constitutional symptoms	Cost/HRQL impact deemed insufficient as symptoms thought to be non-specific
	Metabolic or laboratory	Laboratory-related events; assume that no treatment is required, no QoL impact expected
	Dermatological	Cost/HRQL impact deemed insufficient as symptoms thought to be non-specific

	Pulmonary or upper respiratory	Cost/HRQL impact deemed insufficient as symptoms thought to be non-specific
BSC	Dyspnoea	Event is expected to be representative of the background toxicity profile of patients rather than being treatment related

ERG comment

The proportion of AEs for BSC in the combination therapy model could be verified with CS ref 7 from the REGARD trial, except for fatigue (fatigue 9.6% in CS[10] while 8.7% in the model and 3.5% in the clinical study report REGARD Table JVBD.12.4., p99). While proportions in the model for abdominal pain, anaemia, neutropenia, infection, febrile neutropenia, hypertension and pain could be verified, the proportion of AEs for anorexia (6.1%) and fatigue (25.8%) could not be retrieved with the COUGAR-02 reference.[27]

For RAM+PAC and PAC, the proportion of patients could be verified with the clinical study report of the RAINBOW trial (RAINBOW CSR Table JVBE.12.3.2, p126). According to the ERG, the reasons for excluding AEs based on effect on costs and QoL impact is arbitrary.

The proportions of AEs were obtained from different trials but incorporated without adjustment in the model. According to the ERG, this approach is questionable. For example, the proportion of patients experiencing pain in the DOC arm (i.e. 11%) is obtained from the COUGAR-02 trial and much higher than the proportion of patients experiencing pain in the RAM+PAC arm (i.e. 0.9%) of the RAINBOW trial and the BSC arm (0%) of the REGARD trial. However, the proportion of patients in the active symptom control arm of the COUGAR-02 trial was also high (20.3%). This raises questions to what extend the 11% is representative for DOC in a comparison with RAM+PAC and BSC.

<u>Utility decrements for the monotherapy model</u>

Utility decrements were included for AEs that occurred in more than 5% in either arm of the REGARD trial. Thus, the model includes abdominal pain, anaemia, and hypertension. Dyspnoea was excluded as it was expected to be representative of the background toxicity profile of patients rather than being treatment related.

As in the combination therapy model, utility decrements and duration were obtained from data in a single technology appraisal of pixantrone for the treatment of adults with relapsed or refractory aggressive B-cell non-Hodgkin lymphoma (NICE TA306[47])

ERG comment

It is not clear to the ERG why fatigue was not included in the monotherapy model especially since it was included in the combination therapy model for BSC (8.7%). However, there is confusion about the proportion of patients experiencing grade ≥3 fatigue in the REGARD trial. According to the REGARD trial,[10] fatigue occurred in 6.4 % of the RAM arm and 9.6% in the BSC arm. According to the clinical study report, fatigue occurred in 4.2% and 3.5% in RAM and BSC, respectively. (REGARD clinical study report Table JVBD.12.4., p99**)**.

Utility decrements

Utility decrements are applied for AEs based on the proportion of patients experiencing the AEs, the utility decrement associated with the AEs and the duration of AEs. The utility decrements for AEs were obtained from the references[48-51] as used in a single technology appraisal of pixantrone for the treatment of adults with relapsed or refractory aggressive B-cell non-Hodgkin lymphoma (NICE TA306[47]) and applied from the first cycle onwards for both models. Duration of AEs was estimated from the trial[52] used in the same STA.

ERG comment

Utility decrements are applied from the first cycle onwards. However, since AEs will mostly occur at the beginning phases of the treatment (and always within the first year so that discounting is not yet applied) this will not influence the results.

5.2.7.4 Post-progression health state

The post-progression health state utility was obtained from the mean EQ-5D index score of the REGARD trial at the end of treatment for all patients who discontinued due to progressive disease (0.587). This was measured at the 30-day post discontinuation visit.

Although disease progression is associated with a decline in patient HRQL,[42] no attempts were made by the company to model HRQL across time or with respect to proximity to death due to the lack of data.

The following Table 5.9 shows the utility values for the pre- and post-progression health states for both the combination therapy and monotherapy model.

Table 5.9: Utility values per health state for the combination therapy and monotherapy model

Health state	N	Mean	s.e.
Pre-progression baseline value	646	0.737	0.009
Adjustment for combination therapy			
Adjustment response rate RAM+PAC each cycle		+0.004	
Decrement AEs (weighted average) RAM+PAC cycle 1 only		-0.004	
Adjustment response rate DOC each cycle		+0.002	
Decrement AEs (weighted average) DOC cycle 1 only		-0.005	
Adjustment response rate BSC each cycle		+0.000	
Decrement AEs (weighted average) BSC cycle 1 only		-0.002	
Adjustment for monotherapy			
Decrement AEs (weighted average) RAM cycle 1 only		-0.001	
Decrement AEs (weighted average) BSC cycle 1 only		-0.001	
Post-progression	357	0.587	0.018

Health state utility values were comparable to the utilities used in the NICE technology appraisal of trastuzumab (i.e. 0.7292 [increasing by 0.000142 daily during PFS) and 0.577 for progression-free and progressive disease, respectively).

ERG Comment

According to the ERG, response rate and AEs are clinically meaningful events that may have an impact on QoL. However, it is not clear for the ERG what the rational was for using adjusted baseline EQ-5D utility scores instead of using the treatment arm specific utility values derived at different time points (including response rate and AEs). This question was addressed in the clarification letter and according to the company this was not used in the model because these data were unavailable for BSC or DOC (and using BCO+PAC was considered inappropriate for these comparators). However, according to the ERG this data could have been used to validate the approach currently chosen in the model.

5.2.8 Resources and costs

5.2.8.1 Intervention and comparators' costs and resource use

Cost of the intervention and the comparators are comprised of the drug acquisition, drug administration costs and the cost of monitoring and the tests.

Drug acquisition costs depend on five main components:

- Cost of the drug(s)
- Average dose required
- Treatment duration
- Relative dose intensity
- Required pre-medication

Cost of the drug(s)

Cost of the generically available chemotherapies are taken from eMIT (CS Table 87), which includes the actual prices paid by hospitals over the last 12 months. The prices from eMIT are different from BNF prices (CS Table 88[53]), due to tendering of generic pharmaceuticals. eMIT prices were used in the base case analysis and BNF prices were used in the scenario analysis.

Drug dosing and regimen

Ramucirumab combination and monotherapy treatment regimens were from the SPC, which were the same as the relevant clinical trials.[54] For DOC, a dosage of 75 mg/m2 three-weekly was used from the clinical trials included in the NMA[27, 28] and deemed appropriate conforming to English clinical practice by the authors. Dosing/ regimen info of RAM, PAC in RAM+PAC combination therapy, of RAM as a monotherapy, and of DOC as a monotherapy applied as second line treatment were given in Table 89 in the CS.

Patient weight/BSA

In order to calculate the required drug doses for each regimen, an estimate of body weight and body surface area (BSA) was needed.

For combination therapy, patient characteristics from the RAINBOW study were used for all comparators in the combination therapy. The base case analysis used the average of all patients (weight/ BSA) in the trial (i.e. 63.33 kg), and as a scenario analysis weight/BSA from the patient population in region 1 (i.e. 64.83) (CS Table 90).

For monotherapy, patient characteristics from the REGARD study were used. The base case analysis used the average of all patients (weight=63.33 kg, BSA=1.71m^2) in the trial (i.e.

65.19), and as a scenario analysis weight/BSA from the patient population in region 1 (weight=68.15kg, BSA=1.78m^2, given in CS Table 91).

For each regimen, planned dosage was calculated by multiplying the dose by the mean weight or BSA per treatment cycle (CS Table 92).

Treatment duration

Treatment duration was estimated using parametric curves to determine the time on treatment from trial data. The methodology used to select the most appropriate distribution was the same as outlined for PFS and OS distribution selection. Five parametric distributions were estimated and their goodness of fit to the KM data is assessed.

For combination therapy, KM plots for time on treatment were given in CS Figure 53, and according to the AIC/BIC tests (CS Table 93), and Cox-Snell residuals (CS, Figure 24 in Appendix 11), authors state that the log-normal was the best fitting distribution, and therefore chosen as the base case and the log-logistic distribution was selected as a scenario analysis. Time on treatment for DOC was calculated from Ford et al,[27] based on the reported proportion of patients completing different number of cycles and the maximum threshold of six cycles. This resulted in a mean of 3.07 cycles.

For monotherapy, the KM plot (only for RAM) for time on treatment was given in Figure 55 of the CS. According to the AIC/BIC tests (CS Table 94), Cox-Snell residuals (CS, Figure 25 in Appendix 11) and visual inspection, it is stated that lognormal was the most-appropriate distribution, and therefore chosen as the base case. The log-logistic distribution was selected as a scenario analysis. Gamma distribution (with minimum AIC/BIC test score) was disregarded because of the overestimated extrapolation of time on treatment (1% on treatment at 113 weeks) compared to lognormal/log-logistics, which was deemed clinically infeasible when compared with PFS results (99% had progressed at week 68). A summary of the time on treatment for all comparators were provided in Table 95 of the CS.

Relative dose intensity

RDI adjustments were applied in the base case to correct for the differences seen in the real clinical practice from the planned treatment dosages. In one of the scenario analyses, RDI adjustment was switched off.

RDI data for RAM and RAM+PAC were taken from the RAINBOW and REGARD clinical trials and both RDIs were high (RAM mono: 97.52%, RAM combi: 96.7% PAC: 83.6%) and for DOC, the RDI derived from the Kang study[32] (95%) was applied. Authors substantiate the choice of the Kang study compared to the COUGAR-II study (which was in the NMA) due to the incompatibility of RDI calculations of COUGAR-II study to the RDI calculations for RAM and RAM+PAC. Summary of RDI data for second line treatments were given in Table 96 of the CS.

Drug wastage

In the base case vial wastage from opened vials were assumed, since the leftover would be discarded probably due to the low chance of having multiple GC/GJC patients in a hospital at the same time. As an alternative, no vial wastage assumption was explored in a scenario analysis.

Pre-medication

Pre-medication was included based on SPC for ramucirumab and paclitaxel.[54, 55] For DOC, premedication instructions for first line treatment were used (since DOC is not licensed for second-line). The costs for pre-medication were based on BNF prices and premedication costs per cycle were given in Table 100 in the CS.

Drug administration

Drug administration was costed according to the appropriate HRG codes.[28] Since RAM was not included in the chemotherapy regimen list, existing day case HRG codes were assumed as an estimate for RAM (SB12Z: £214)) and RAM+PAC (SB14Z: £330 and SB15Z: £302) in the base case. Outpatient costs were used as sensitivity analysis. The costs were given in Table 101 of the CS.

Follow-up care

Follow-up costs include tests and clinical monitoring both applied in pre- and post-progression health states dependent on whether the patient was on- or off-treatment. Rates of the tests/monitoring were based on clinical expert input.

In the model it was assumed that consultant visits take place every four weeks for all treatments. For RAM+PAC, treatment cycles and consultation visits were aligned, but for DOC, consultation visits are underestimated as the treatment cycle for DOC is every three weeks. Upon progression, consultation was assumed to be every 12 weeks. For BSC, consultation visits were also every 12 weeks.

CT scanning was assumed to take place every 12 weeks while patient is on active treatment. Patients on BSC do not receive any CT scans.

Full-blood count was assumed to be conducted prior to every cycle of treatment in order to ensure patients eligibility for treatment continuation. Similarly, hepatic and renal function tests were assumed to be assessed at the beginning of each treatment cycle. Upon treatment stopping, the tests are ceased, as well. Patients who receive 3rd line treatment are assumed to receive consultant visits, CT scanning and tests at the rates that were assumed for 2nd line treatment. Overview of follow-up costs (with rates and frequency values for on- and off-treatment) were provided in Tables 103, 104 and 105 of the CS.

Table 5.10 summarises the unit costs associated with the technology in the economic model.

Table 5.10: Unit costs associated with the technology in the economic model

	RAM+PAC	DOC	RAM	BSC
Drug cost	£1,257.64	£14.15	£1,500	N/A
Pre-med cost	£15.94	£4.02	£1.40	N/A

Administration cost	£233.50	£71.33	£107	N/A
Consultation	£34.81	£34.81	£34.81	£11.14
Follow-up (tests and CT scan)	£25.82	£30.74	£39.68	N/A
Total	£1,567.71	£155.05	£1,682.89	£11.14

ERG Comment

The ERG group could not replicate the same values in Table 97 in CS for RAM dose per treatment cycle in the RAM+PAC treatment. 979.84 mg was found instead of 972.84 mg reported in the table. In the economic model, 979.84 was used as well.

Although in the table £34.29 was reported as the drug acquisition cost per treatment cycle for docetaxel, in the economic model £42.45 was used due to a technical coding error. It was corrected in the ERG base case. The effect of the correction of this error was minimal for the comparison RAM+PAC and BSC.

The drug acquisition costs of RAM for the RAM+PAC treatment were derived based on RAINBOW trial. Average weight of the patients in the RAINBOW trial was 63.33 kg. If the average weight of the patients in the RAINBOW trial had been 64.83 kg, cost per treatment cycle would increase by £1,000. Expecting that the average weight of UK patients would be higher than the RAINBOW baseline patient population (which included a high percentage of Asian patients), in the ERG base case, the weight/BSA values from region 1 (weight = 68.15kg, BSA =1.78m^2) will be used. Therefore, in the ERG base case analysis, treatment cycle cost will be £6,030.56 for RAM+PAC instead of £5,030.56 in the CS.

5.2.8.2 Health state unit costs and resource use

A summary table of the unit costs associated with the health states is given in Table 106 of the CS.

Other non-intervention related health state costs include hospitalisation, BSC costs, adverse event costs and third-line costs

Hospitalisation costs

Since authors claim that hospitalisation may occur due to reasons other than treatment related AEs, hospitalisation costs apart from AEs were included in the model. Hospitalisation data from the trials was incorporated and costed using NHS reference costs.

For both combination and monotherapies, hospitalisation data was collected from RAINBOW and REGARD trials, respectively, which informed the model on:

- Length of stay
- Rate of hospitalisations

Length of stay by treatment arm for combination therapy and for monotherapy are presented in Table 107 and Table 109 in CS. Observed data was used to inform mean rate of

hospitalisations in the models (observed no of hospitalisations/follow up length). Authors mentioned that other model fits (Poisson/negative binomial) were not chosen due to poor fit to the data, but still provided as scenario analyses. Table 5.11 presents the length of stay as used in the models.

Table 5.11: Length of stay: combination and monotherapy model

Treatment	Mean days	s.d	Source
RAM+PAC	14.60	1.20	RAINBOW safety population
PBO+PAC	18.20	1.68	RAINBOW safety population
RAM	13.40	10.68	REGARD safety population
BSC	10.40	6.79	REGARD safety population

For both combination and monotherapy models, rates of hospitalisation based on pre-progression state were also applied in post-progression state. Alternative scenario analyses with augmented risks of hospitalisation (RR =2 and 3) were also included in the economic model. Table 5.12 presents the hospitalisation probabilities of the combination and monotherapy model.

Table 5.12: Hospitalisation probabilities: combination and monotherapy model

Cohort	Probabilities per week		Source
	PFS	PPS	
RAM+PAC	0.024	0.024	RAINBOW
DOC*	0.036	0.036	RAINBOW
BSC	0.055	0.055	REGARD
RAM	0.034	0.034	REGARD

* assumed to be the same as PAC due to lack of alternative data

Hospitalisation cost per day for UK was £526.05 (s.e. £242.68), which is the average weighted by activity of the non-elective inpatients stays for 'Malignant Gastrointestinal Tract Disorders, with multiple Interventions' with various complication and comorbidity scores (given in Table 111 of the CS).

ERG Comment

In the clarification letter, the ERG indicated a potential double counting of hospitalisation costs because HRGs referring to AEs also take hospitalisations into account. In the response to the clarification letter, a scenario had been added to the model, which reduces the rate of hospitalisations by an estimate of proportion of hospitalisations due to adverse events due to AEs. The ERG team will use these AE adjusted hospitalisation rates in the additional analyses that will be conducted.

Also in the CS, it was mentioned that hospitalisation cost per day was calculated based on the weighted average of different hospitalisation costs (with HRG codes) given in Table 84 of the CS. However, it was not very clear which weights were used and the original cost values of the HRG coded hospitalisations.

Terminal care costs

All patients are assumed to receive terminal care as a one-off cost prior to death. The cost for terminal care was taken from TA208 for trastuzumab[38] and were based on Coyle et al,[51] where an 'end of life' cost of £4,000 (2010 values) was applied. This cost was then inflated to 2014 values (£4,467.60).

Best supportive care

Since there is a lack of data in the literature concerning the treatment of GC/GOJ patients in the second-line setting, the company had conducted a treatment pattern study of patients who received a platinum plus fluoropyrimidine in the first-line setting in the UK. The study was a chart review of hospital medical records which assessed the supportive care received by all patients (n=200), disaggregated by whether they received a second-line therapy (n=57) or supportive care only (n=143).[3]

Specific BSC components were identified for inclusion if they were received by at least 10% of patients receiving supportive care alone or receiving active treatment. Afterwards, refinements were conducted based on expected resource impact and the possibility of double counting with the treatment of AEs. The BSC components received in >10% patients were outlined in Table 85 of the CS as well as the rationale for their inclusion or exclusion. The included BSC components (morphine, distress management, transfusions and radiation therapy) were costed using relevant HRG codes and other alternative sources, which were listed in Table 86 in the CS. The components and rates of supportive care as well as their frequencies, which were summarised in Table 113 of the CS. In the base case, supportive care components and rates differ between additionally treated (with an active treatment) and supportive care only patients. As a scenario analysis, the same rates for additionally treated and supportive care only patients were applied.

Based on BNF guidance,[56] NHS choices and other assumptions, BSC components' frequency and cost per administration were given in Table 114 of the CS. The same assumptions for patients who receive BSC post-progression are similar to BSC assumptions for patients in the pre-progression state.

ERG Comment

Components of the BSC were identified by the treatment pattern study and clinical expert judgement. In the original CS, it was not clear how the clinical expert opinions had been incorporated. More details were requested in the clarification letter by the ERG. The CS provided more details in response to the clarification letter. It was mentioned that three medical oncologists were approached to review the submission and validate the clinical assumptions and their answers were provided in Appendix 5 of the clarification letter.

5.2.8.3 Adverse reaction unit costs and resource use

Costs of AE treatment were included in the model. The inclusion criteria were earlier explained in section 5.2.7.3.

Combination therapy model

AE counts were used directly from the RAINBOW trial CSR for RAM+PAC, from REGARD trial CSR for BSC and COUGAR-II trial for DOC. Grade-3/4 AEs were included if they occurred in more than 5% of patients on any treatment and if they had a significant

impact on costs/QALYs. The exclusion of AEs based on their impact were discussed in the CS Appendix 11 in Tables 31 and 32. In Appendix 6 of the CS (subsections 1.5 to 1.7), the incidence of AEs were estimated via NMA, but authors claimed that this approach did not yield the rates for each relevant AEs for each comparator. AE rates used in the economic model for combination therapy were given in Table 117 of the CS.

Monotherapy model

AE counts from REGARD trial CSR were used both for RAM and BSC. Grade-3/4 AEs were included if they occurred in more than 5% of patients on the RAM treatment and if they had a significant impact on costs/QALYs. The exclusion of AEs based on their impact were not discussed in Appendix 11 of the CS. AE rates used in the economic model for monotherapy was given in Table 118 in the CS.

Adverse event HRGs

NHS reference costs were used in costing the AEs. Table 115 and 116 in the CS show the AE costs and their relevant HRG codes for combination and monotherapy.

ERG Comment

AE incidence based threshold criteria (5% in each relevant treatment arm) results in different selections of AEs for BSC in the combination therapy and in the monotherapy models. This will result in different AE costs for BSC in the combination therapy and monotherapy models.

A naïve indirect comparison method based on the AE counts were used in calculating the AE rates. The NMA for AEs give results only for grade 3 and 4 anaemia, neutropenia and febrile neutropenia. Since the number of AEs for which an NMA was available is limited and given the heterogeneity in the network, no alternative approach to the AE count approach in the CS will be presented.

5.2.8.4 Miscellaneous unit costs and resource use

Third-line treatment

All patients were assumed to continue to receive BSC upon disease progression, and a proportion of patients went on to receive third-line chemotherapy. Third-line therapies were included only with respect to the drug acquisition and administration cost, and the cost of follow-up care. Total cost of receiving third-line therapy is applied to the first cycle following progression. A rate of 12% (from Market research[57]) was used, which was lower than the post-discontinuation chemotherapy rates from the RAINBOW (47.9% RAM+PAC versus 46% PAC) and REGARD (30.3% RAM+BSC versus 37.6% BSC+PBO) trials. In the CS, it was mentioned that this reflects the clinical practice in UK better. Trial post-discontinuation rates and 0% as third-line chemotherapy probability were given as scenario analysis. Distribution of third-line regimens (second-line treatment specific) in the combination and in the monotherapy model (Tables 120 and 121 in the CS) were estimated using market shared data and clinical expert opinion. Treatment duration was assumed to be 50% of PPS duration. In the estimation of third-line treatment costs same assumptions and sources were used as second-line treatment cost estimation. The details of the calculations were given in Appendix 11 of the CS.

ERG Comment

Similar to the second line, there was an error in calculating the third-line docetaxel costs and weekly costs of docetaxel acquisition is changed to £11.43 (using 140 mg/unit) from £14.15 (using 80 mg/unit). The effect of the correction of this error was minimal for the comparison RAM+PAC and BSC.

In the CS, 12%, which had been derived from market research, was used for the probability that a patient who progressed after second-line treatment receives a third-line treatment. However the actual probability of third-line treatment was different in REGARD and RAINBOW trials. Since the observed OS was realised in part due to the third-line treatment, any change in the rate of third-line treatment could potentially lead to changes in the overall survival. In the CS, the OS estimates were not adjusted according to this post-discontinuation treatment probability from market research. Therefore, the ERG team considers it to be inconsistent to apply this adjustment for the costs. However from the deterministic sensitivity analysis results, it was seen that the impact of this probability is very limited.

5.2.9 Cost-effectiveness results

Combination therapy

This section describes the results of the base case analysis of the CS for combination therapy. The following Table 5.13 presents the total costs, LYs and QALYs for the combination therapy model including BSC, DOC, and RAM+PAC. The incremental values are shown versus BSC. While DOC is included in these base case results, the company decided not to discuss DOC further in the CS. The reasoning was as follow; it was not a licenced treatment option and does not appear to be cost-effective at a standard willingness to pay threshold.

Table 5.13: Base case: Combination therapy results

Technologies	Total costs (£)	Total LYs	Total QALYs	Incremental costs (£) vs BSC	Incremental LYs vs BSC	Incremental QALYs vs BSC	Incremental cost/LYs vs BSC	ICER (£) incremental (QALYs) vs BSC
BSC	£13,400	0.45	0.29	-	-	-	-	
DOC	£18,779	0.59	0.39	£5,378	0.14	0.10	£38,498	£53,830
RAM+PAC	£52,996	0.94	0.62	£39,595	0.48	0.33	£81,809	£118,209

The incremental effects for RAM+PAC compared to BSC were 0.48 LYs and 0.33 QALYs. Incremental costs were £81,809 corresponding to an ICER of £118,209.

Clinical validation

A comparison is made between the outputs of the clinical model and the output of the economic model (Table 5.14). According to the CS, the model estimated median and mean values for RAM+PAC represent the clinical data reasonable well.

Table 5.14: Summary of model results compared with clinical data: Combination therapy

Outcome	Clinical trial results (CI)			Model results		
	RAM+PAC	BSC	Incremental	RAM+PAC	BSC	Incremental
Median OS	9.63 months (8.5, 10.8)	N/A	N/A	9.43 months	4.60 months	4.83 months
Median PFS	4.40 months (4.24, 5.32)	N/A	N/A	4.14 Months	1.15 months	2.99 months
Mean OS	11.13 months (10.30, 11.96) (restricted mean)	N/A	N/A	11.63 Months	5.60 months	6.03 months
Mean PFS	6.20 months (5.62, 6.78) (restricted mean)	N/A	N/A	6.00 Months	1.85 months	4.15 months

Abbreviations: N/A = not applicable

The results by health state for BSC and RAM+PAC are presented in Table 5.15. This table shows that patients treated with RAM+PAC spend the majority of their remaining life-time in the pre-progression health state (57%) while patients receiving BSC spent most of their remaining life time in the post-progression health state (64%).

Table 5.15: Results by technology and health state for the combination therapy

Technology	Outcome	Lys	% of overall LYs	QALYs	% of overall QALYs
BSC	Progression-free survival	0.14	32%	0.10	36%
	Progressed disease	0.31	68%	0.18	64%
	Overall survival	0.45	100%	0.29	100%
RAM+PAC	Progression-free survival	0.48	52%	0.35	57%
	Progressed disease	0.45	48%	0.27	43%
	Overall survival	0.94	100%	0.62	100%

Note: Numbers may not compute due to rounding

As a consequence, the disaggregated results for effects show that incremental QALYs and LYs for RAM+PAC are mainly accrued in the pre-progression health state, 75% and 70%, respectively. Disaggregated costs revealed that incremental costs were mainly due to the therapy costs associated with RAM+PAC (76.2%). Costs for terminal care were however, slightly lower (-0.2%) because of discounting.

Monotherapy model

This section describes the results of the base case analysis for monotherapy of the CS. Table 5.16 describes the total costs, LYs and QALYS estimated from the model for RAM+BSC and Incremental effects for RAM+BSC are 0.16 LYs and 0.12 QALYs while incremental costs are £22,542. The ICER for RAM+BSC compared to PBO+BSC is £188,640 per QALY.

Table 5.16: Base case results for monotherapy

Technologies	Total costs (£)	Total LYs	Total QALYs	Incremental costs (£)	Incremental Lys	Incremental QALYs	Incremental cost/LYs	ICER (£) incremental (QALYs)
PBO+BSC	£14,137	0.49	0.31	-	-	-	-	-
RAM+BSC	£36,678	0.66	0.43	£22,542	0.16	0.12	£137,723	£188,640

Note: Numbers may not compute due to rounding

Clinical validation

Clinical results from the monotherapy model are compared to the results from the REGARD trial. Table 5.17 shows the model results and clinical data for the monotherapy model. The model underestimates PFS and OS for both arms. The incremental difference for OS is 0.02 and 0.12 for PFS.

Table 5.17: Model results compared with clinical data for monotherapy

Outcome	Clinical trial results (CI)			Model results		
	RAM+BSC	PBO+BSC	Inc	RAM+BSC	PBO+BSC	Inc
Median OS	5.2 months (2.3, 9.9)	3.8 months (1.7, 7.1)	1.4 months	5.06 months	3.68 months	1.38 months
Median PFS	2.1 months (1.5, 2.7)	1.3 months (1.34, 1.4)	0.8 months	1.61 months	0.69 months	0.92 months

Abbreviations: inc = incremental

Results per health state are presented in Table 5.18. This table reveals that patients receiving RAM+BSC spend a greater proportion of their remaining life expectancy in the pre-progression health state.

Table 5.18: Results by technology and health state of the monotherapy model

Technology	Outcome	LYs	% of overall LY	QALY	% of overall QALYs
PBO+BSC	Progression-free survival	0.12	24%	0.09	29%
	Progressed disease	0.37	76%	0.22	71%
	Overall survival	0.49	100%	0.31	100%
RAM+BSC	Progression-free survival	0.28	42%	0.20	48%
	Progressed disease	0.38	58%	0.22	52%
	Overall survival	0.66	100%	0.43	100%

Note: Numbers may not compute due to rounding

Disaggregated results show that almost all (LY: 96% QALY: 97%) of the survival benefit of RAM+BSC compared to PBO+BSC occurs in the pre-progression health state. The majority of the incremental costs of RAM+BSC were due to the administration and acquisition costs of ramucirumab.

ERG comment

Combination therapy

The ERG does not agree with the reasoning of the manufacturer not to discuss DOC in further detail. According to the ERG, the usual approach would be to compare RAM+PAC to the next best comparator, i.e. DOC. Also, the company aimed to compare RAM+PAC to DOC because it was included in the model. The difference in costs between DOC and RAM+PAC was £34,217, the difference in QALYs was 0.23. The ICER would be £148,770.

The ERG identified an error in the half cycle correction of the model submitted by the company. This error was confirmed in the clarification letter and the company edited the model so that the half-cycle correction was applied to all cost components. The impact of this correction on the ICER was negligible (£118,174).

The results for IRI and FOLFIRI are not included in the base case results, despite the fact that these comparators were part of the scope. Additional analyses were conducted by the ERG and are presented in section 5.3.

Monotherapy

The monotherapy model also included an error in the half cycle correction. This was confirmed and corrected in the clarification letter. The impact on the ICER was negligible (£188,437).

Both OS and PFS are underestimated in the model However, the underestimation for PBO+RAM is larger than the underestimation for RAM+BSC. Therefore, the incremental difference for the median PFS is overestimated (0.12 months) and this also influences the incremental QALYs.

5.2.10 Sensitivity analyses

Probabilistic sensitivity analysis

Probabilistic sensitivity analysis (PSA) was performed to capture the uncertainty in the estimation of input parameters. Distributions were assigned to the input parameters and randomly sampled using 10,000 Monte Carlo simulations.

In case standard errors were unavailable, these were assumed to be 10% of the mean. Table 5.19 shows the choice of distributions for each parameter included in the PSA.

Table 5.19: Summary of PSA distributions used for parameters

Parameter	Distribution	Rationale
Discount rates	N/A	Not subject to sampling uncertainty, explored in the deterministic sensitivity analysis
Regression parameters*	Multivariate normal	To capture correlation between normally distribution regression parameters
KM data for OS**	Normal	To capture uncertainty in the KM trial data
Drug costs	N/A	Not subject to sampling uncertainty
Other unit costs	Gamma	Constrained to interval of 0 and positive infinity
Resource use rates	Gamma	Constrained to interval of 0 and positive infinity
Resource use probabilities	Beta	Constrained to interval of 0 and 1
Health state utilities	Beta	Both the ‘pre-progression’ and ‘post-progression’ health states are sufficiently far from death hence constrained on interval 0 and 1
Adverse event utility decrement	Gamma	Constrained to interval of 0 and positive infinity
Response utility increment and	Beta	Constrained to interval of 0 and 1

Parameter	Distribution	Rationale
proportion of overall response rate**		

*PFS, OS with parametric distributions, time on treatment and hospitalisation rates
**For combination therapy only

The results of the PSA for the combination therapy model are presented in Table 5.20. Figure 5.10 presents the cost-effectiveness plane and Figure 5.11 shows the cost-effectiveness acceptability curve.

Table 5.20: Cost-effectiveness results of probability sensitivity analysis for combination therapy

Technologies	Total costs (£)	Total LY	Total QALYs	Incremental costs (£)	Incremental LY	Incremental QALYs	Incremental cost/LYG	ICER (£) incremental (QALYs)
BSC	£13,787	0.47	0.30	-	-	-	-	-
RAM+PAC	£53,444	0.96	0.64	£39,657	0.49	0.34	£80,858	£116,820

Figure 5.10: Cost-effectiveness plane for combination therapy (QALYs)

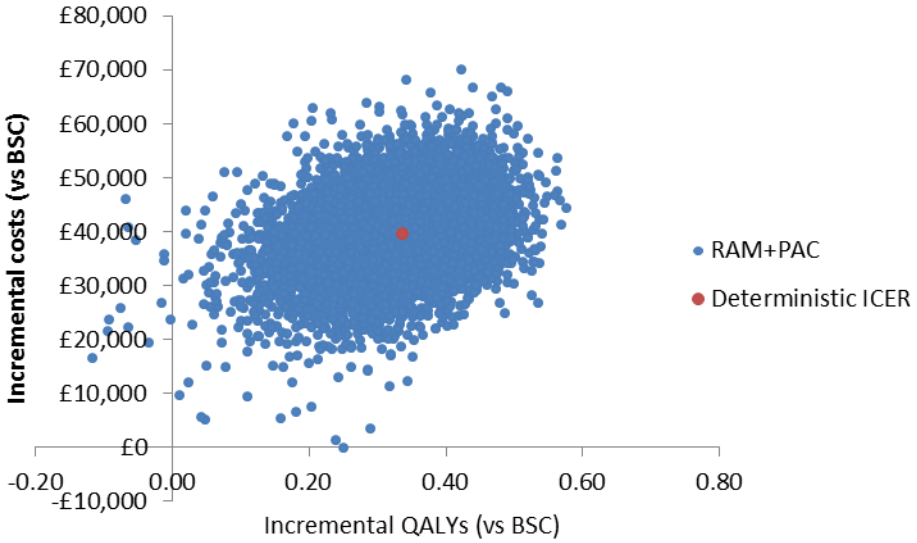
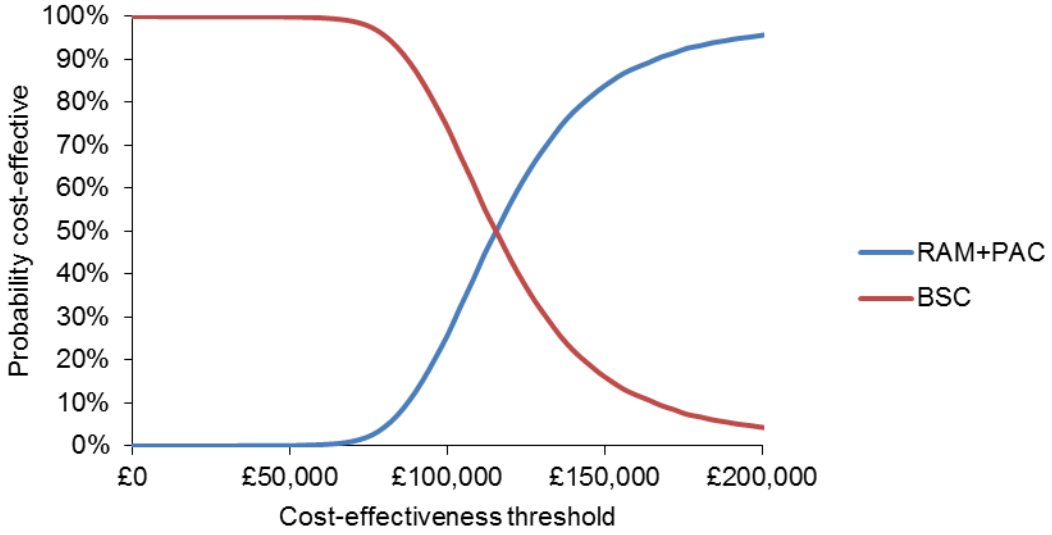


Figure 5.11: Cost-effectiveness acceptability curve for combination therapy



ERG comment

The ICER from the PSA (£116,820) is slightly lower than the ICER from the base case analysis £118,209. This is caused by slightly higher LYs and QALYs for both BSC and RAM+PAC. As a consequence, the incremental difference in LYs and QALYs increases with 0.01. While this difference is very small, it causes the ICER to decrease with £1,389 per QALY.

Deterministic sensitivity analysis

Deterministic sensitivity analyses were conducted by the manufacturer to test the assumptions in the economic model. Analyses were conducted for second-line drug costs, utilities, resource use (regarding hospitalisation, BSC and AEs), efficacy, discount rates and third-line treatment costs. Table 5.21 shows for each item the base case value, value in the deterministic analysis and the new ICER.

Table 5.21: Results of deterministic sensitivity analysis (CS Table 141)

	Base case value	**Deterministic sensitivity analysis value**	**ICER**
Base case			**£118,209**
Costs			
Drug cost	eMIT prices	BNF	**£131,537**
Drug administration	Day case setting	Outpatient setting	£115,781
Hospitalisation rate	Observed, adjusted for treatment	Negative binomial, adjusted for treatment	£119,127
	Observed, adjusted for treatment	Poisson, adjusted for treatment	£113,191
	Observed, adjusted for treatment	Observed, unadjusted	£123,858
Length of hospitalisation stay stratified by	Treatment	Region 1	**£109,434**
	Treatment	All patients	**£109,143**
	Treatment	Region 1 and treatment	£116,420
Hospitalisation rate	Relative risk of post-	Relative risk of post-	£116,944

	Base case value	Deterministic sensitivity analysis value	ICER
post-progression	progression hospitalisation = 1	progression hospitalisation = 2	
	Relative risk of post-progression survival = 1	Relative risk of post-progression survival = 3	£116,146
Wastage	Vial wastage	Vial sharing	£116,167
Relative dose intensity	On	Off	**£127,186**
Time on treatment	Lognormal distribution	Log-logistic distribution	£120,360
Third-line therapy	12% in each arm	0% in each arm	£118,097
	12% in each arm	Trial rate (47% in both arms)	£118,495
Proportion of time in post-progression spent on treatment	Proportion = 50%	Proportion = 25%	£118,035
	Proportion = 50%	Proportion = 75%	£118,383
AEs costs	100%	Scaled down by 50%	£117,581
	Infection length = 15.1 days	Infection length = 10 days	£118,209
BSC resource rates	Differential rates across active treatments and BSC	Equal rates	£119,455
BSA and body-weight data source	All patients	Region 1	**£127,128**
Efficacy			
Progression-free survival	Unadjusted, Weibull	Unadjusted, log-normal	£117,588
	Unadjusted, Weibull	Unadjusted, log-logistic	£117,040
Overall survival	Kaplan-Meier extrapolated from T1= last event in PAC arm	Kaplan-Meier extrapolated from t_1= last event in PAC arm + 2 months (t_1=24.12)	£115,886
	Kaplan-Meier extrapolated from T1= last event in PAC arm	Kaplan-Meier extrapolated from t_1= last event in PAC arm - 2 months ($t_{1=}$20.14)	£117,304
	Kaplan-Meier extrapolated from T1= last event in PAC arm	Kaplan-Meier extrapolated from t_1= last event in RAM+PAC arm (t_1=21.39)	£119,198
Other			
Discount rate (costs and health)	3.5%	1.5%	£117,253
Utility	Response utility applied	No response utility applied	£118,878

ERG comment

The ICER was mostly influenced (+/- £7,000) by the source for drug costs (BNF instead of eMIT prices +£13,328), length of hospitalisation stay (stratified by region 1 instead of treatment -£8,775 or one LOS for all patients (stratified by all patients) instead of treatment stratified -£9,066), relative dose intensity (off instead of on +£8,977) and source for BSA and body-weight data (region 1 instead of all patients +£8,919).

Regarding these parameters, the ERG agrees with the manufacturer that it is appropriate to use eMIT prices and apply relative dose intensity. However, the ERG considers it more appropriate to use region 1 data for BSA and body-weight since it is believed that region data better reflects the BSA and body-weight of the UK population. The ERG also considers it more appropriate to use region 1 data for length of hospitalisation stay stratification but in addition to treatment stratification. These assumptions will be used in the ERG base case which is presented in section 5.3

Scenario analysis

For OS, the Kaplan-Meier data and extrapolating with an exponential distribution from the end of the trial period was chosen as the base case as it represented the most conservative approach and utilises the trial data to the fullest extent. However, since this methodology has its limitations, alternative approaches were explored as scenario analysis.

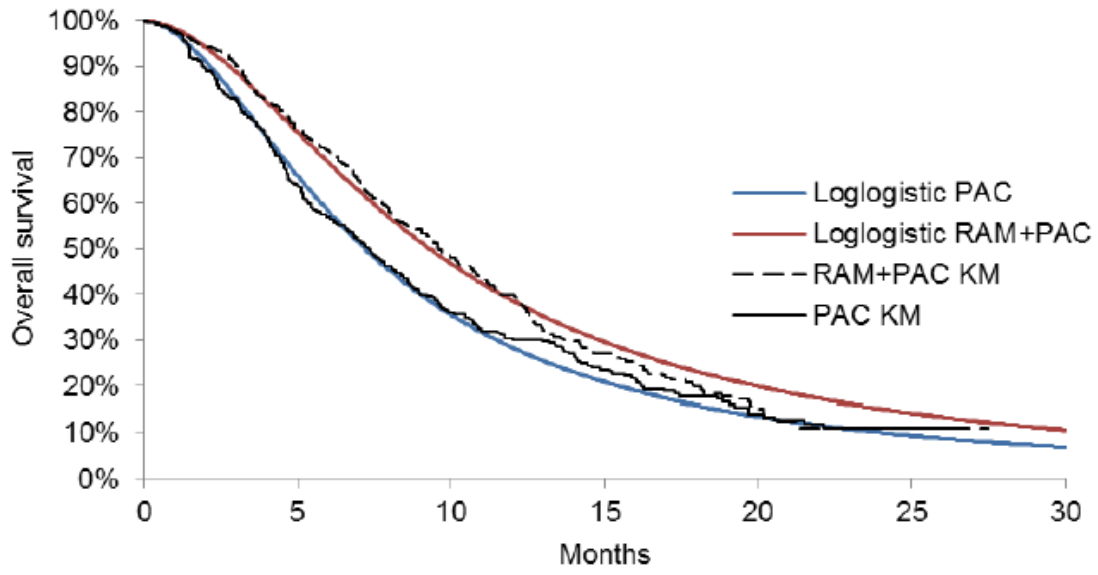
Independently fitted OS curves showed that the Weibull distribution followed by the log-logistic distribution had the best fit to the observed trial data for RAM+PAC. The log-logistic distribution was the best fit for the PBO+PAC and the Weibull the second worst fitting distribution. The Weibull and log-logistic distribution were used in two scenarios.

Using the Weibull distribution yielded almost similar results as the base case analysis (ICER £117,236). Using the log-logistic distribution reduced the ICER to £96,103 per QALY.

ERG comment

The ERG does not consider the log-logistic distribution a realistic scenario. Figure 5.12 shows the extrapolation of the log-logistic distribution. The parameterised OS from 12 months onwards shows a moderate difference between RAM+PAC and PBO+PAC while the KM data shows more or less similar OS curves from 12 months onwards.

Figure 5.12: Overall of KM and joint parametric estimates of OS using log-logistic



Subgroup analysis

A scenario analysis using a geographical subgroup was conducted to validate the results of the ITT analysis in the combination therapy model. In this scenario, region 1 (including English patients, i.e. North America, Europe, Australia and Israel) adjusted OS, PFS and time on treatment were used. A log-logistic and Weibull distribution were used. Costs per QALY were £114,474 for the Weibull distribution and £95,618 for the log-logistic distribution.

ERG comment

While the ERG considers it plausible to adjust analysis for region 1, the ERG believes this to be more relevant for BSA and body-weight and hospitalisations. Therefore, the scenario that only adjusts for OS, PFS and time on treatment is not considered appropriate as new base case by the ERG.

Monotherapy

The results of the PSA are presented in Table 5.22. Figure 5.13 shows the cost-effectiveness plane and Figure 5.14 the CEAC.

Table 5.22: Cost-effectiveness results of probability sensitivity analysis of monotherapy

Technologies	Total costs (£)	Total LY	Total QALYs	Incremental costs (£)	Incremental LY	Incremental QALYs	Incremental cost/LYG	ICER (£) incremental (QALYs)
PBO+BSC	£14,149	0.50	0.31	-	-	-	-	-
RAM	£36,376	0.66	0.43	£22,227	0.16	0.12	£138,573	£189,232

Figure 5.13: Cost-effectiveness plane for monotherapy (QALYs)

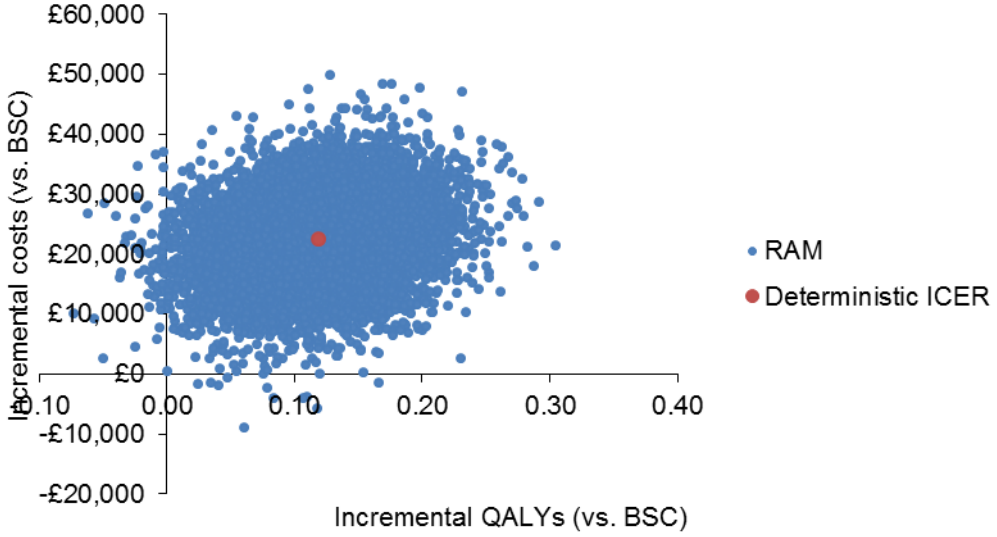
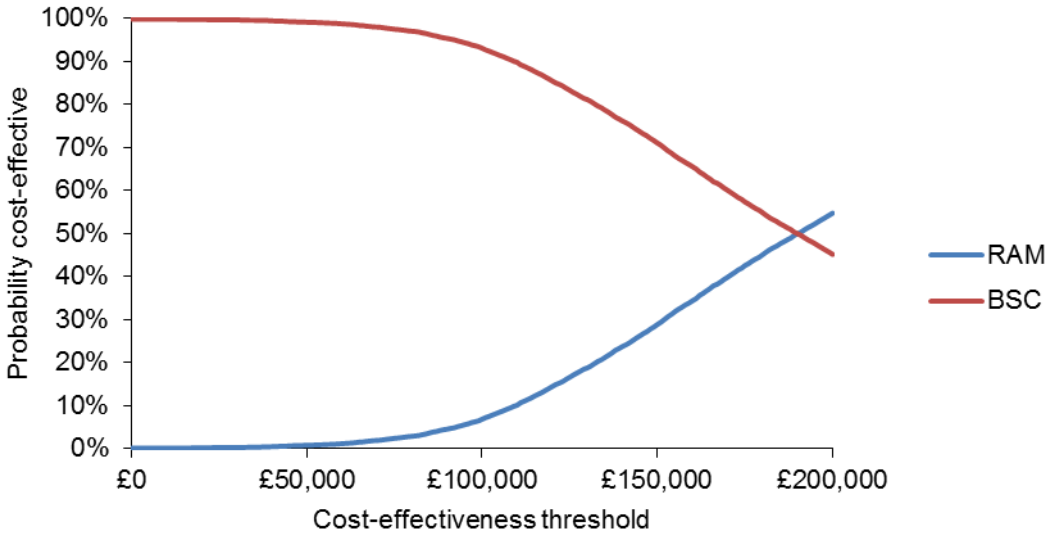


Figure 5.14: Cost-effectiveness acceptability curve for monotherapy



ERG comment

The ICER from the PSA (£189,232) is almost similar to the ICER from the base case analysis £188,640.

Deterministic sensitivity analysis

Deterministic sensitivity analysis to test the assumptions was also performed for the monotherapy model and also included second-line drug costs, utilities, resource use (regarding hospitalisation, BSC and AEs), efficacy, discount rates and third-line treatment costs. Table 5.23 shows the base case values and deterministic sensitivity analysis value and ICER for the monotherapy model.

Table 5.23: Results of the deterministic sensitivity analysis

	Base case value	Deterministic sensitivity analysis value	ICER
Base case			**£188,640**
Costs			
Drug administration	Day case setting	Outpatient setting	£185,379
Hospitalisation rate	Observed, adjusted	Negative binomial, adjusted	£192,735
	Observed, adjusted	Poisson, adjusted	£188,705
	Observed, adjusted	Observed, unadjusted	£217,891
Length of hospitalisation stay stratified by	Treatment	Region 1	£171,426
	Treatment	All patients	£170,974
	Treatment	Region 1 and treatment	£173,706
Hospitalisation rate post-progression	Relative risk of post-progression hospitalisation = 1	Relative risk of post-progression hospitalisation = 2	£180,476
	Relative risk of post-progression hospitalisation = 1	Relative risk of post-progression survival = 3	£173,582
Proportion of time in post-progression spent on treatment	Proportion = 50%	Proportion = 25%	£188,644
Proportion of time in post-progression spent on treatment	Proportion = 50%	Proportion = 75%	£188,637
Wastage	Vial wastage	Vial sharing	£163,319
Relative dose intensity	On	Off	£188,640
Time on treatment	Lognormal	Log-logistic	£189,358
Third-line therapy	5% in each arm	RAM = 30.3% BSC = 37.6%	£187,043
	5% in each arm	0% in each arm	£188,690
Drug cost source	eMIT	BNF	£188,600
AEs costs	100%	Scaled down by 50%	£188,369
BSC resource rates	Differential rates across arms	Equal rates	£191,347
BSA and body-weight data source	All patients	Region 1 patients	£188,640
Efficacy			
Progression-free survival	Lognormal	Log-logistic	£188,926
	Lognormal	Weibull	£197,904
Overall survival	Gamma	Lognormal	£174,485
Other			
Discount rate (costs and benefit)	3.5%	1.5%	£185,708

ERG comment

The ICER was influenced by the assumptions. An increase or decrease of the costs per QALY of more than £15,000 was observed for hospitalisation rate (observed, unadjusted instead of observed adjusted +£29,251), length of hospitalisation stay (stratified by region 1 instead of treatment -£17,214, stratified by all patients instead of treatment -£17,666), the relative risk of post progression survival is 3 instead of 1 -£15,058) and wastage (vial sharing

instead of vial wastage -£25,321). The ICER was also influenced by the method for estimating OS but to a lesser extent.

The ERG considers only the length of hospitalisation stratified by region 1 and treatment as a more plausible assumption.

Scenario analysis

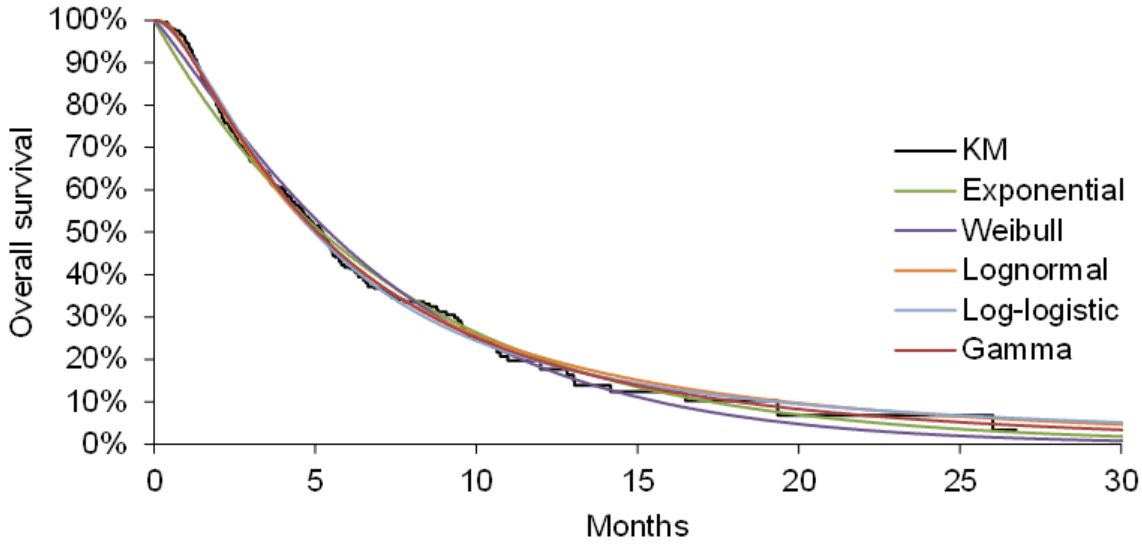
The Gamma distribution was chosen as the base case for extrapolating OS data since this was the most conservative approach because it had the shortest overall length of the extrapolation. The lack of long-term registry data or more mature trial data made it difficult to assess the appropriateness of long-term extrapolation.

An alternative analysis was conducted using the lognormal distribution (i.e. the distribution with a better fit using the goodness-of-fit diagnostic tests). This decreases the ICER to £174,485.

ERG comment

The ERG agrees with the company to use the more conservative extrapolation in the base case. However it was not clear to the ERG from Appendix 9 of the CS, Figure 23, p.69 (Figure 5.15 below) why the Gamma distribution had the shortest overall length of extrapolation.

Figure 5.15: Overlay of KM and parametric estimates of OS for RAM



5.2.11 Model validation and <u>face validity check</u>

In the CS, it was mentioned that for the technical validation of both combination therapy and monotherapy models, an external consultant did a comprehensive cell-by-cell review of the formulae and all subsequent adaptations for UK were cross-checked by two members of the external consultancy team.

For the internal validation, comparisons was made between the model results and results of the trials (i.e. RAINBOW for combination therapy and REGARD for the monotherapy model).

Table 5.24: Summary of model results with different parametric curves compared with clinical data of the RAINBOW trial

Outcome	RAINBOW Clinical trial results (CI)			Model results		
	RAM+PAC	PBO+PAC	Inc*	RAM+PAC	PBO+PAC	Inc*
Weibull distribution						
Median OS	9.63 months (8.5, 10.8)	7.36 months (6.3, 8.4)	2.27 months	9.66 months	7.82 months	1.84 months
Median PFS	4.40 months (4.24 , 5.32)	2.86 months (2.79, 3.02)	1.54 months	4.14 months	2.53 months	1.61 months
Log-logistic distribution						
Median OS	9.63 months (8.5, 10.8)	7.36 months (6.3, 8.4)	2.27 months	9.20 months	7.13 months	2.07 months
Median PFS	4.40 months (4.24 , 5.32)	2.86 months (2.79, 3.02)	1.54 months	4.14 months	2.53 months	1.61 months
Kaplan-Meier extrapolation						
Median OS	9.63 months (8.5, 10.8)	7.36 months (6.3, 8.4)	2.27 months	9.43 months	7.13 months	2.3 months

Table 5.25: Summary of model results with different parametric curves compared with clinical data of the REGARD trial

Outcome	REGARD Clinical trial results (CI)			Model results		
	RAM+BSC	PBO+BSC	Inc	RAM+BSC	PBO+BSC	Inc
Median OS	5.2 months (2.3, 9.9)	3.8 months (1.7, 7.1)	1.4 months	5.06 months	3.68 months	1.38 months
Median PFS	2.1 months (1.5, 2.7)	1.3 months (1.3, 1.4)	0.8 months	1.61 months	0.69 months	0.92 months

Authors concluded that the comparison of model results versus clinical trial data demonstrated that the incorporation of the PFS and OS data from the rainbow trial into the economic model was done in a way that is reflective of the RAINBOW and REGARD trials.

As an external validation exercise, authors plotted the OS survival curve of BSC, which had been derived from the survival curve from RAM+PAC and the hazard ratio from NMA, with the OS KM curve of the BSC arm from COUGAR-II trial in Figure 65 in CS. The OS curve from the combination therapy model slightly overestimates the OS curve from COUGAR-II trial.

ERG Comment

From the CS, the validation procedure was not clearly explained. Only a cell-by-cell review by one external reviewer is prone to human mistakes and several errors may be overlooked. A checklist form (i.e. a predefined list of tests to be performed with expected outcomes) would help the ERG team to assess the technical integrity of the model.

Further minor technical errors were found in the excel model, which were mentioned in the clarification letter and in previous sections. These errors were having small impacts on the ICER.

For PFS, it seems that the median PFS from the model is not within the 95% CI of the PFS estimate from the RAINBOW trial. Furthermore, according to the ERG, the fact that the OS median estimates predicted by the model are within the 95% CI of the clinical trial estimates does not guarantee that the internal validity for OS is ensured.

As an additional interval validation exercise, BSC OS and PFS survival curve from the combination therapy models are compared with the BSC OS and PFS curves from the REGARD trial. The HR derived from the ITC/network meta-analysis is applied on the RAM+PAC curves for OS and PFS. Figure 5.16 and Figure 5.17 show that the PFS and OS fitted parametric curves for BSC seem to be different from the KM PFS and OS curves from REGARD trial. Nevertheless, although some differences between the KM curves (REGARD) and estimated curves from HRs were observed, we do not observe an overall overestimation or underestimation of the PFS or OS KM curves.

Figure 5.16: BSC OS parametric curve where the HR is derived from the ITC vs. BSC OS KM from REGARD

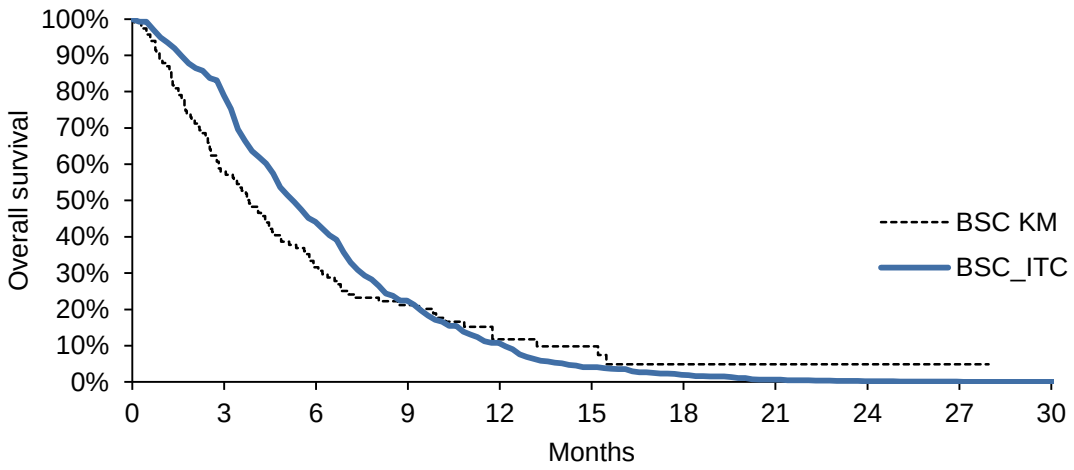
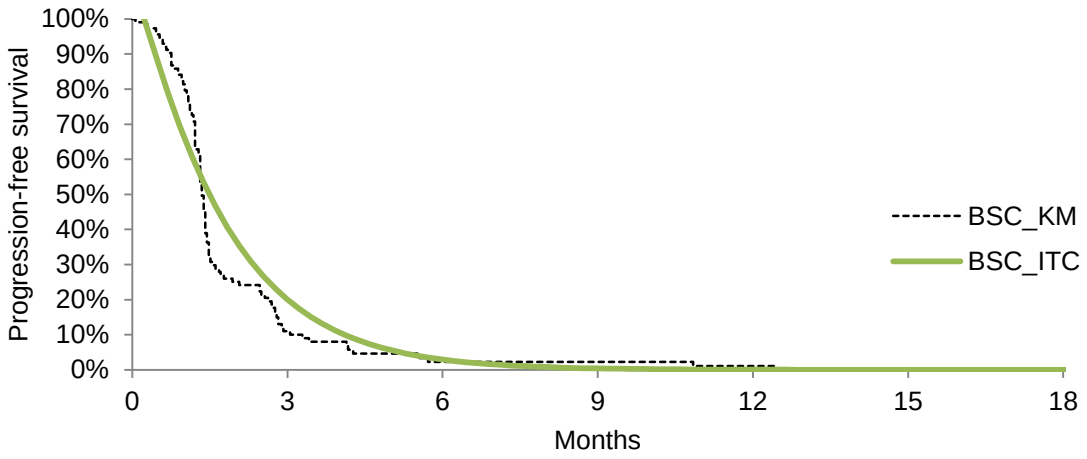


Figure 5.17: BSC PFS parametric curve where the HR is derived from the NMA vs. BSC PFS KM from REGARD



Regarding external validity, the modelled OS for BSC in the combination therapy model is higher than the OS for BSC from the COUGAR-II study (Figure 65 in CS). This may imply that the OS for BSC is overestimated.

Although DOC was one of the comparators, no validation exercise had been performed in the CS for DOC. Therefore the ERG compares the DOC model outcomes with the results from COUGAR-II trial below. It seems that the model overestimates the OS that is observed in the clinical trial.

Table 5.26: Summary of model results compared with clinical data for docetaxel.

Outcome	COUGAR II Clinical trial results (CI)	Model results
	DOC	DOC
Median OS	5.2 months (4.1, 5.9)	6.95 months
Median PFS	Not reported	2.52 months

The differences between the model OS and COUGAR-II OS results for DOC (Table 5.26 above) and BSC (Figure 65 in CS) indicate an overestimation. These differences may be explained by the differences of baseline characteristics between the COUGAR-II and RAINBOW trials. In the COUGAR-II trial, median age at the baseline was higher (65 compared to 61 in the RAINBOW) and ECOG performance status was worse. Furthermore, the COUGAR-II is a UK study, whereas in the RAINBOW, patients from Asian countries were also included. Asian patients are more aggressively treated and receive more lines of active treatments, as it was shown in Table 26 in the CS.[8]

5.3 Exploratory and sensitivity analyses undertaken by the ERG

5.3.1 Inclusion of the additional comparators

In line with the final scope, the following additional comparators were added to the combination therapy and monotherapy models respectively.

For combination therapy these are paclitaxel, irinotecan (IRI), and FOLFIRI (irinotecan, leucovorin and fluorouracil). For monotherapy the added comparator are: docetaxel, irinotecan and FOLFIRI.

Treatment effectiveness of the additional comparators

Overall survival

In the response to the clarification letter,[11] in Table 3 of Appendix 2, the company presented indirect treatment comparison results for base-case OS HRs for all relevant alternative treatments in the combination therapy model. From this table, HRs for IRI, PAC and FOLFIRI (versus RAM+PAC) were listed as follows: 1/0.72, 1/0.81 and 1/0.86.

Parallel to the approach described in Section 5.2.6, under the PH assumption, these HR values were applied to the fitted OS curve for the RAM+PAC arm from the RAINBOW trial to derive OS curves for the alternative treatments.

For the monotherapy model, HRs for DOC, IRI and FOLFIRI (versus BSC) could be observed, 0.67, 0.48 and 0.40, respectively (Table 3, in Appendix 2 of the response to the clarification letter[11]). Under the PH assumption, these HRs were applied to the fitted OS curve for the BSC arm from the REGARD trial, which generates the OS curve estimates to be used in the monotherapy model for DOC, IRI and FOLFIRI.

Progression free survival

For PFS, not all the necessary HR estimates from the NMA were available to the ERG. Therefore, an assumption was made similar to these in Figure 42 in the CS.[8] Since there the HR for PFS was assumed to be 1 for IRI versus DOC, the estimated PFS curve of IRI was assumed to be identical to the estimated PFS curve of DOC. Also, in the Sym et al 2012 paper,[26] no significant difference between the median PFS values of IRI and FOLFIRI was found. Therefore, the PFS HR for IRI versus FOLFIRI was assumed to be 1 as well. Under this assumption, the PFS estimated curves of IRI, FOLFIRI and DOC are identical. Since the PFS curve for PAC was already estimated in the model, no further assumption was required. The PFS curve estimates for IRI, FOLFIRI and DOC were used in both the combination and monotherapy model.

Input parameters

Drug acquisition costs

Drug acquisition costs for IRI and FOLFIRI treatments were obtained from the eMIT database. For irinotecan, the price of one package (1 pack size) of 300 mg/15ml solution for infusion vial was £20.11. For 5-FU (fluorouracil), the price of a package (1 pack size) of 2.5 gr/50 ml (5%) solution for infusion vial was £2.04. Finally for leucovorin (folic acid), the price of a package (28 pack size) of 5 mg tablets was £0.29.

Under IRI treatment, irinotecan is administered 150 mg/m2 every two weeks. Treatment duration for IRI was assumed to be eight weeks, which was the median treatment duration derived from Sym et al 2012.[26]

Under FOLFIRI treatment, irinotecan is administered 150 mg/m^2 every two weeks, leucovorin is administered 20 mg/m^2 and 5-FU is administered 1000 mg/m^2 per day for two days. Treatment duration for IRI was assumed to be 12 weeks, which was the median treatment duration derived from Sym et al 2012.[26]

For both IRI and FOLFIRI treatments, administration is preceded with atropine 0.25 gr (with a price of £1.26 for 10 pack of 0.6 mg/ml solutions for injection ampules), dexamethasone (dosage/cycle and price info is already available as it was used as a premedication for other comparators) and 5HT-3 receptor antagonist (e.g. Ondansetron), with a price of £1.82 for 10 pack of 8 mg tablets. Relative dose intensity for IRI and FOLFIRI were assumed to be the same as DOC.

Based in the information above, the drug acquisition cost for IRI is £16.47 per week and for FOLFIRI £19.91 per week.

Other costs

Drug administration, follow-up care costs, hospitalisation, terminal and miscellaneous costs were calculated in the same way as the DOC costs were calculated. Frequency based adjustments due to the treatment cycle differences were applied if necessary. AE rates were derived from the event counts reported in Sym et al 2012[26] and are tabulated in Table 5.27 as below:

Table 5.27: Summary derived AE rates for IRI and FOLFIRI

Event	IRI	FOLFIRI
Abdominal pain	0.0%	0.0%
Anaemia	0.0%	10.3%
Neutropenia	27.6%	37.9%
Infection	0.0%	0.0%
Febrile neutropenia	0.0%	3.4%
Pain	0.0%	0.0%
Anorexia	10.3%	13.8%
Hypertension	0.0%	0.0%
Fatigue	10.3%	3.4%
Diarrhoea	3%	7%

Since diarrhoea was not among the AEs from other interventions, cost (£475) and disutility data for diarrhoea (-0.103 for 1.44 months) are obtained from TA307 (technology appraisal for aflibercept in combination with irinotecan and fluoroucil-based therapy in metastatic colorectal cancer).[58]

Results of the cost effectiveness analyses - combination therapy model:

The base case results from the full incremental analysis are summarised in Table 5.28.

Table 5.28: Base case model results with additional comparators

	QALYs	Cost	Incr. QALY	Incr. Cost	ICER
BSC	0.29	£13,419			
DOC	0.39	£18,849			Extendedly dominated† by IRI
IRI	0.47	£21,765	0.18[a]	£8,346[a]	£44,319[a]
PAC	0.52	£26,213			Dominated by FOLFIRI
FOLFIRI	0.52	£24,837	0.05[b]	£3,072[b]	£60,278[b]
RAM+PAC	0.62	£53,003	0.10[c]	£28,166[c]	£294,362[c]

† An extendedly dominated strategy has an ICER higher than that of the next most effective strategy.
[a] IRI vs. BSC [b] FOLFIRI vs. IRI [C] RAM+PAC vs. FOLFIRI

BSC is the cheapest treatment but also the one providing the lowest amount of QALYs. DOC is extendedly dominated by IRI, whilst PAC is dominated by FOLFIRI, as it provides almost the same QALYs (0.5222 vs 0.5244, respectively) with higher costs. Thus, the efficiency frontier consists of BSC, IRI, FOLFIRI, and RAM+PAC. The ICER of RAM+PAC compared to FOLFIRI is £294,362.

Alternatively, we present in Table 5.29 the base case ICERs for the RAM+PAC against each comparator.

Table 5.29: Base case model results (intervention versus comparator only)

Intervention	Comparator	Incr. QALY	Incr. Cost	ICER
RAM+PAC	BSC	0.33	£39,584	£118,174
	DOC	0.24	£34,153	£145,302
	IRI	0.15	£31,238	£213,015
	PAC	0.1	£26,790	£273,657
	FOLFIRI	0.1	£28,166	£294,362

Note that the results in this scenario analysis are derived from the model which was received by the clarification letter answer after some corrections (e.g. half cycle correction etc.), therefore some of the results in the above table may deviate from the results that were provided in Table 5.13.

Results of the cost effectiveness analyses for the monotherapy model

The base case results from the full incremental analysis are summarised in Table 5.30. It should be noted that the results for some of the additional comparator results for the monotherapy model should be interpreted with care, since the OS HRs used in the model for DOC, IRI and FOLFIRI were derived from an evidence network that was prepared for the combination therapy model, and therefore did not include the REGARD trial. Also, the PFS curve estimates from the combination therapy are used in the monotherapy model for DOC, IRI and FOLFIRI, since no HR with respect to BSC was available. Therefore results should be interpreted with some caution.

Table 5.30: Base case model results with additional comparators

	QALYs	Cost	Incr. QALY	Incr. Cost	ICER
BSC	0.31	£14,168			
RAM	0.43	£36,685			Dominated by IRI
DOC	0.50	£22,856			Extendedly dominated[†] by IRI
IRI	0.72	£30,504	0.41[a]	£16,336[a]	£39,843[a]
FOLFIRI	0.87	£37,102	0.15[b]	£6,614[b]	£44,093[b]

[†] An extendedly dominated strategy has an ICER higher than that of the next most effective strategy.
[a] IRI vs BSC [b] FOLFIRI vs IRI

Similar to the combination therapy, BSC has the lowest costs but also the lowest amount of QALYs. RAM and DOC are dominated and extendedly dominated by IRI, respectively. Thus, the efficiency frontier consists of BSC, IRI, and FOLFIRI.

Probabilistic sensitivity analysis

Probabilistic sensitivity analysis (PSA) was performed for all comparators to capture the uncertainty in the estimation of input parameters.

Figure 5.18 presents the cost-effectiveness plane and Figure 5.19 shows the cost-effectiveness acceptability curves (CEACs) with additional comparators for the combination therapy model.

For combination therapy, the CEACs confirmed that BSC is the treatment with the highest probability of being cost-effective, for threshold ratios lower than £22,500. When that threshold is exceeded then IRI is the treatment with the highest probability of being cost-effective until £25,000. For thresholds above £25,000, FOLFIRI becomes the treatment with the highest probability of being cost-effective.

A similar pattern can be seen in the CEACs for monotherapy model. BSC is the treatment with the highest probability of being cost-effective, for threshold ratios lower than £32,500. When that threshold is exceeded then IRI is the treatment with the highest probability of being cost-effective until threshold is lower than £100,000. After £100,000, FOLFIRI is the treatment with the highest probability of being cost-effective. PSA results of the monotherapy model are similar to the combination therapy, and the scatterplot and CEAC figures are given in Appendix 4 of this report.

Figure 5.18: Cost-effectiveness plane for combination therapy (QALYs)

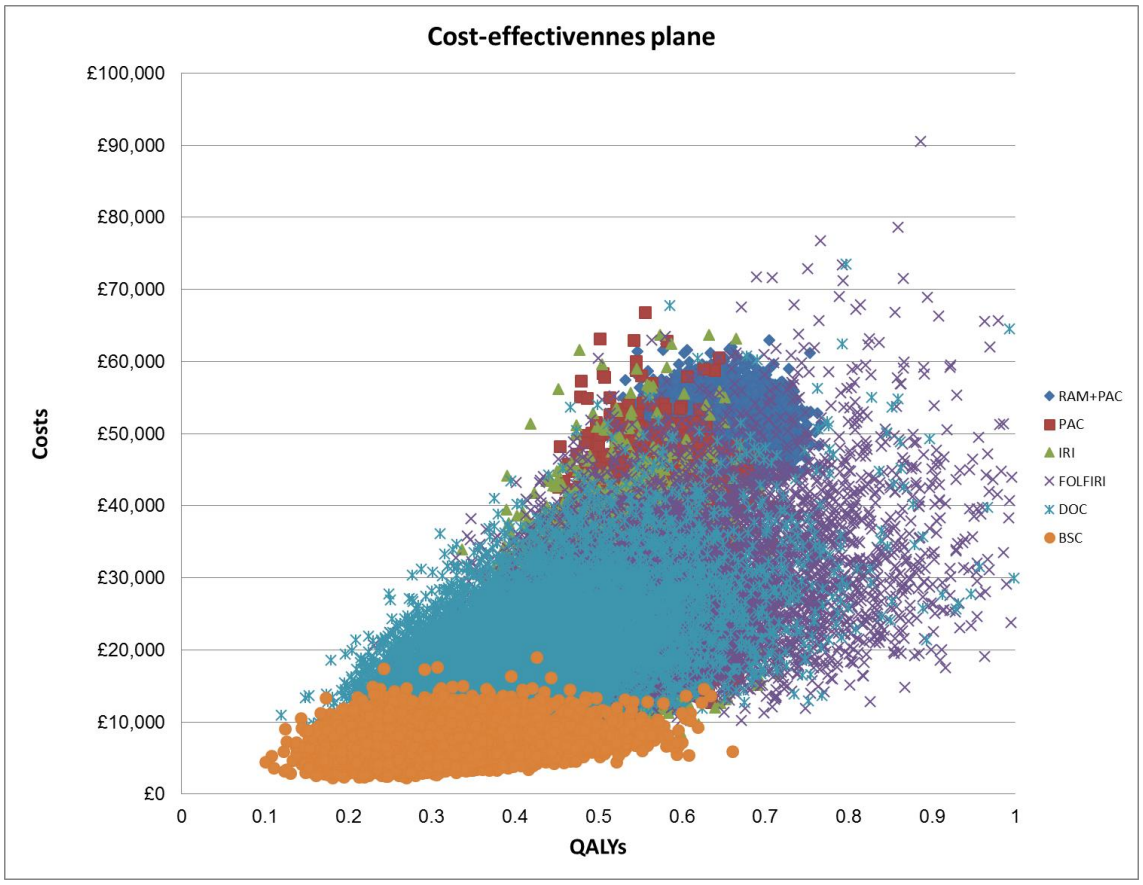
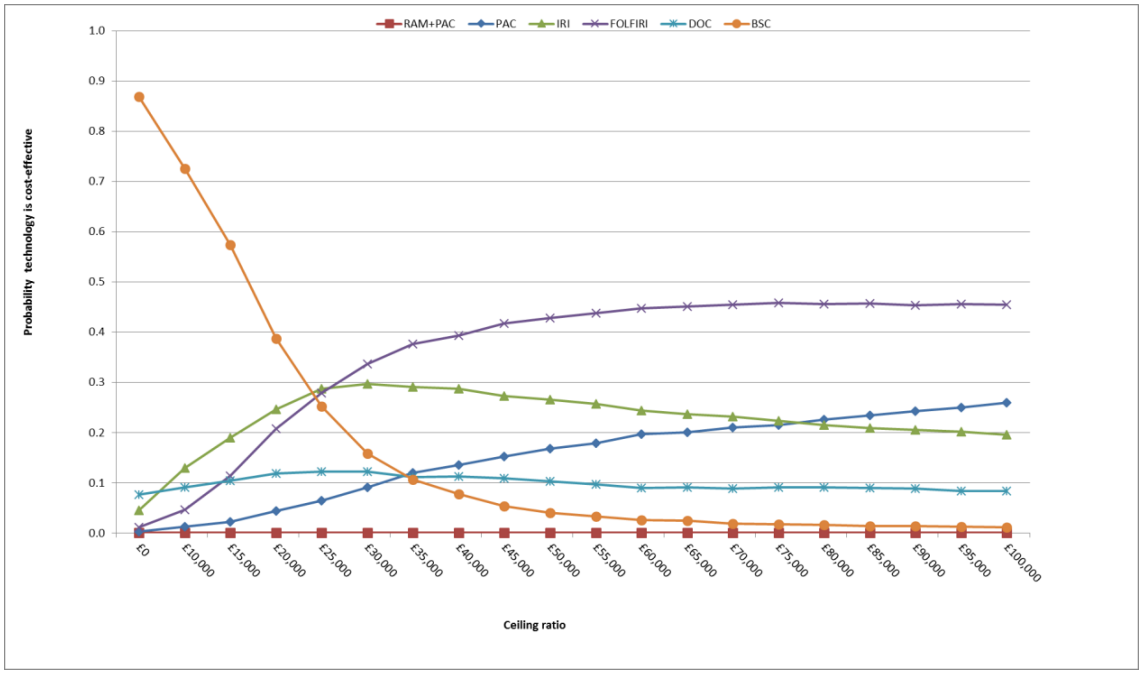


Figure 5.19: Cost-effectiveness acceptability curves for combination therapy



5.3.2 New ERG base case analysis

Based on several remarks in section 5.2 of this report the ERG defined a new base case analysis. This new ERG base case included the following adjustments:

- Programming errors confirmed by the manufacturer have been removed.
- Programming errors relating to DOC price (detected after the clarification letter)
- Hospitalization stratification is based on treatment and region
- BSA/weight based on region 1 are used instead of ITT population
- Double counting of hospitalizations due to AEs are corrected

The results of the adjusted ERG base case are presented in table 5.31 and 5.32 for both the combination therapy and monotherapy model, respectively. It can be seen that in the new base case, in the combination therapy model, the ICER of RAM+PAC vs. BSC has increase by £9,000 compared to the CS base case. The ICER of DOC vs. BSC changed from £53,830 to £38,280. This can be explained by the detected error in the docetaxel cost calculations. Further details are provided in section 6 for both the combination and monotherapy model.

Table 5.31: New ERG Base case for combination therapy

Technologies	Total costs (£)	Total LYs	Total QALYs	Incremental costs (£) vs BSC	Incremental LYs vs BSC	Incremental QALYs vs BSC	Incremental cost/LYs vs BSC	ICER (£) incremental (QALYs) vs BSC
BSC	£6,695	0.45	0.29	-	-	-	-	
DOC	£10,523	0.59	0.39	£3,828	0.14	0.10	£27,343	£38,280
RAM+PAC	£50,050	0.94	0.62	£43,354	0.48	0.33	£89,576	£129,431

Table 5.32: New ERG Base case for monotherapy

Technologies	Total costs (£)	Total LYs	Total QALYs	Incremental costs (£) vs BSC	Incremental LYs vs BSC	Incremental QALYs vs BSC	Incremental cost/LYs vs BSC	ICER (£) incremental (QALYs) vs BSC
BSC	£6,853	0.49	0.31	-	-	-	-	
RAM	£29,325	0.66	0.43	£22,472	0.16	0.12	£137,296	£188,055

Probabilistic sensitivity analysis

A PSA was performed for all comparators to capture the uncertainty in the estimation of input parameters in the new ERG base case. Figure 5.20 presents the cost-effectiveness plane and Figure 5.21 shows the CEACs with additional comparators for the combination therapy model.

Compared to the previous base case, the CEAC results are more or less the same in the new ERG base case. However in the CE scatter plot, an upwards shift for RAM+PAC costs can be observed. This shift is due to the use of region 1 specific weight/BSA instead of the ITT population weight/BSA. This increased the RAM+PAC treatment costs, since an additional vial of RAM is needed at each administration. For monotherapy, both the CE scatter plot and CEACs of the new ERG base case are more or less the same as the base case of the company and thus not presented here.

Figure 5.20: Cost-effectiveness plane for combination therapy (QALYs)

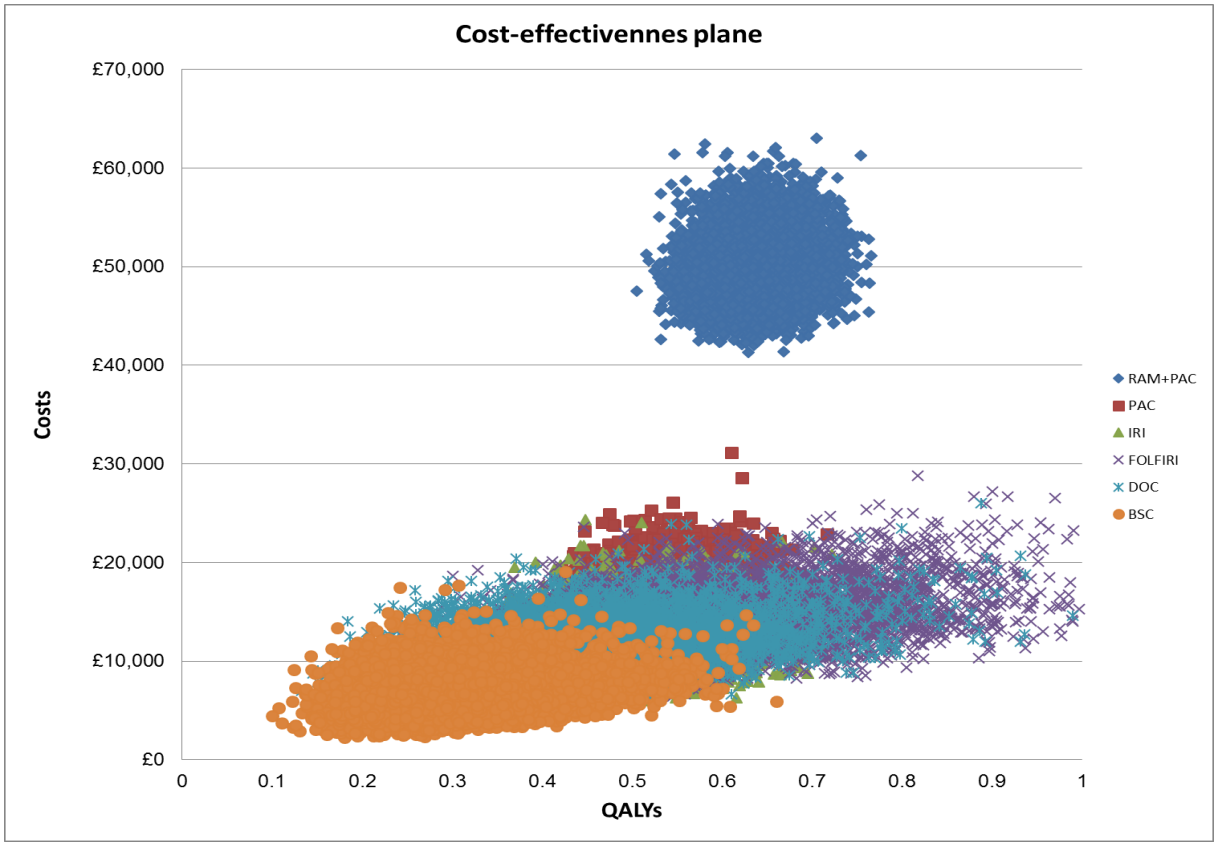
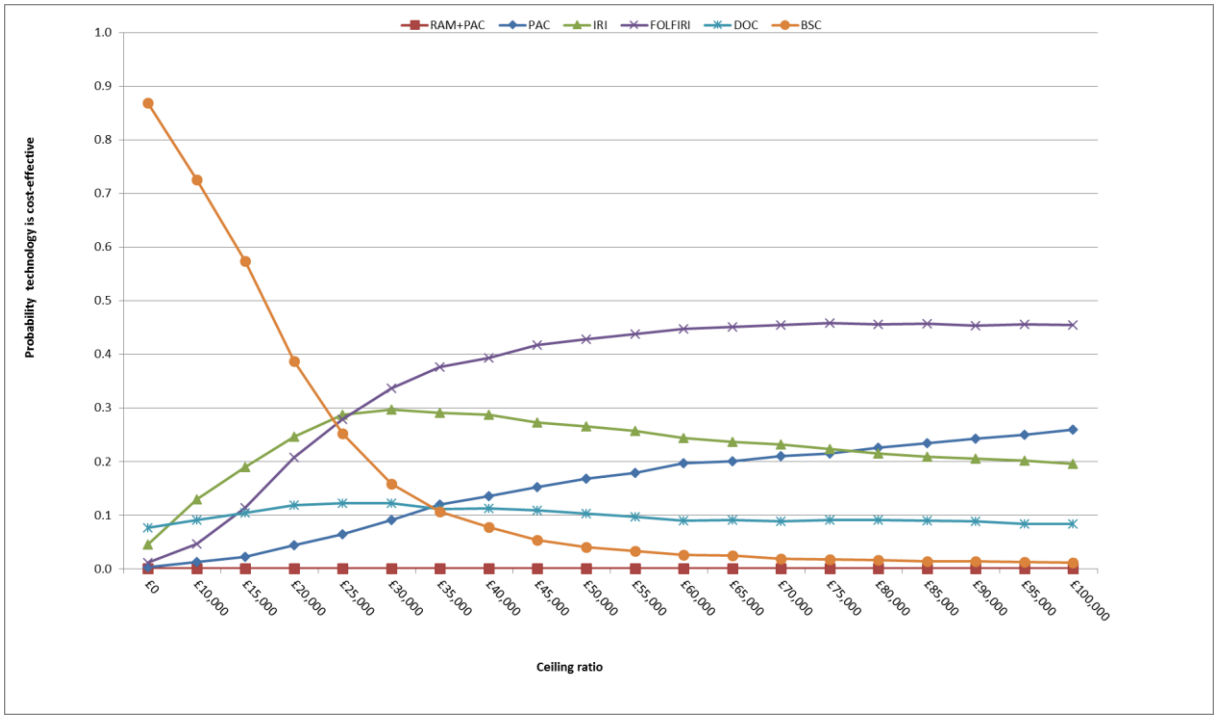


Figure 5.21: Cost-effectiveness acceptability curves for combination therapy



5.3.3 Exploratory scenario analyses undertaken by ERG

Scenario analysis1: incorporating Roy et al[28]

In the cost-effectiveness analysis in the CS, the OS HRs derived from an indirect treatment comparison without Roy et al 2013 (3 arm trial: IRI vs DOC vs PEP02)[28] were used. Previously, in Section 4.5 of this report, the reasons why Roy et al[28] should have been included were explained. In this scenario, we use the NMA results from the evidence network in which Roy et al was included and applied this to the new ERG base case (DOC versus BSC £38,280, RAM+PAC versus BSC £129,431) for the combination therapy model. The corresponding cost-effectiveness results can be seen in Table 5.33.

Table 5.33: Combination therapy results with HRs calculated from the evidence network including Roy et al

Technologies	Total costs (£)	Total LYs	Total QALYs	Incremental costs (£) vs BSC	Incremental LYs vs BSC	Incremental QALYs vs BSC	Incremental cost/LYs vs BSC	ICER (£) incremental (QALYs) vs BSC
BSC	£6,954	0.51	0.32	-	-	-	-	
DOC	£11,121	0.68	0.44	£4,166	0.17	0.12	£24,370	£35,230
RAM+PAC	£50,050	0.94	0.62	£43,095	0.42	0.30	£101,473	£143,577

As it can be seen, when Roy et al[28] is included the ICER of RAM+PAC versus BSC increases from £129,431 to £143,577. However, it should be noted that the comparison of RAM+PAC versus BSC relies on four different studies from various regions, contributing to serious heterogeneity between the studies. On top of that, considering the heterogeneity of the studies

used in the PFS NMA and the assumptions made on the PFS HRs, all these results should be interpreted with great caution.

Scenario analysis 2: RAM+PAC versus PAC (only data from RAINBOW trial is used)
Due to the heterogeneity between studies and the assumptions in the ITC and NMA, in this scenario, we use only direct evidence as obtained from the RAINBOW trial to compare RAM+PAC vs PAC. The corresponding results were given in Table 5.34.

Table 5.34: Combination therapy RAM+PAC vs PAC results when only RAINBOW trial data is used

Technologies	Total costs (£)	Total LYs	Total QALYs	Incremental costs (£) vs PAC	Incremental LYs vs PAC	Incremental QALYs vs PAC	Incremental cost/LYs vs PAC	ICER (£) incremental (QALYs) vs PAC
PAC	£14,938	0.83	0.53	-	-	-	-	
RAM+PAC	£50,050	0.94	0.62	£35,112	0.11	0.09	£323,590	£392,108

Note that the PAC results in Table 5.34 were calculated from the RAINBOW IPD data (independent modelling of the RAINBOW OS KM data until the last event observation and exponential extrapolation afterwards). This is different from the ERG/CS base cases, in which the OS HR (PAC versus RAM+PAC) was found from the indirect treatment comparison, and that HR is applied on the RAM+PAC OS curve under the PH assumption.

Since in the monotherapy model, only REGARD trial data is being used in the RAM versus BSC comparison, no additional analyses are conducted.

Scenario analysis 3: utility values directly from RAINBOW EQ-5D results
In this scenario analysis, mean EQ-5D scores collected at different time points in the RAINBOW trial are implemented. These EQ-5D results were given in Table 21 in the CS. The mean EQ-5D score from a time point is applied in the model until the next time point at which the next EQ-5D score is reported. It is assumed that the scores from the PAC+PBO arm are representative for the other comparators of RAM+PAC, as well (BSC, DOC). No utility increment due to response or decrements due to AEs were applied in this analysis since it was assumed that this was already captured by the RAINBOW data. The results of this scenario are tabulated in Table 5.35.

Table 5.35: Combination therapy results when EQ-5D scores collected in different time points in the RAINBOW trial are implemented.

Technologies	Total costs (£)	Total LYs	Total QALYs	Incremental costs (£) vs BSC	Incremental LYs vs BSC	Incremental QALYs vs BSC	ICER (£) incremental (QALYs) vs BSC
BSC	£6,695	0.45	0.28	-	-	-	
DOC	£10,523	0.59	0.388	£3,827	0.14	0.11	£36,196
RAM+PAC	£50,050	0.94	0.62	£43,354	0.48	0.34	£129,384

The implementation of the EQ-5D scores collected at different time points in the model had a small impact on the ICER.

When the RAM+PAC vs PAC comparison in scenario 2 is conducted with the utility values directly taken from the RAINBOW EQ-5D, the corresponding results can be seen in Table 5.36.

Table 5.36: Combination therapy RAM+PAC vs PAC results when only RAINBOW trial data is used and utility values from RAINBOW EQ-5D are directly used

Technologies	Total costs (£)	Total LYs	Total QALYs	Incremental costs (£) vs PAC	Incremental LYs vs PAC	Incremental QALYs vs PAC	Incremental cost/LYs vs PAC	ICER (£) incremental (QALYs) vs PAC
PAC	£14,938	0.83	0.53	-	-	-	-	
RAM+PAC	£50,050	0.94	0.62	£35,112	0.11	0.09	£323,590	£408,223

It can be seen that, compared to the ICER in scenario 2, the ICER has increased in this scenario slightly. This is due to the fact that the EQ-5D trial data, when applied directly to the model give slightly more favourable results for PAC.

5.4 Conclusions of the cost-effectiveness section

The economic model described in the CS is considered by the ERG to meet the NICE reference case to a reasonable extent, but is not fully in line with the decision problem specified in the scope. Reviewing the overall evidence, the ERG confirmed that there was no existing cost-effectiveness model for ramucirumab for the current indication.

In terms of population, it was not clear to what extent the patients in the REGARD trial can be seen as representative for patients who are not appropriate for paclitaxel treatment. The inclusion criteria of the REGARD and RAINBOW trials were the same and in the RAINBOW trial, patients all received PAC. Thus, it is likely that the REGARD trial actually consisted mostly of patients for whom paclitaxel treatment *was* appropriate. Thus, only if patients for whom paclitaxel is appropriate have the same treatment response as patients for whom it is not will the presented analysis by the company reflect the true cost-effectiveness of monotherapy.

In terms of comparators, in the combination therapy model, paclitaxel, irinotecan, and FOLFIRI, which were mentioned in the final scope, were excluded based on a threshold of >10% (indicative of routine use in the NHS England). Similarly, in the monotherapy model, docetaxel, irinotecan and FOLFIRI, which were mentioned in the final scope, were excluded based on the >10% threshold. The ERG thinks that in the NICE scoping process, the aspect of routine use was already taken into account, hence applying a second filter by the company may have excluded important comparators. Also, this 10% threshold was not always consistently implemented, as docetaxel was included as an exception in the combination therapy and excluded in the monotherapy.

The ERG assessment indicated that the model was generally well presented and reported. However, a few issues regarding the electronic model were identified that altered the cost-

effectiveness results for the combination therapy model. By correcting these issues and changing a few input parameters (to region specific values), an ERG base case was defined. The company base case ICER for RAM+PAC vs BSC was £118,209 per QALY gained for combination therapy whilst the corresponding ERG base case ICER amounted to £129,431 per QALY gained. For monotherapy, both the company and ERG base case ICERs for RAM vs BSC were around £188,000 per QALY gained for monotherapy. The reason of the change in combination therapy is mostly because of the additional vial required at each RAM administration due to the increase in the weight/BSA of patients in the ERG base case.

The input for the model was derived from RAINBOW/REGARD trial data and literature. For some input values, such as AE costs, non-systematic literature searches were performed or other sources, such as clinical expert opinion, were used.

The utility for pre- and post- progression states were derived from the EQ-5D data collected in the RAINBOW study for both models. For disutilities due to AEs and utility increments due to clinical response, values from published literature were used.

The cost effectiveness results were generally robust under the scenario analyses conducted, though a few scenarios impacted the ICER noticeably.

In the combination therapy, the ICER was mostly influenced by the source for drug costs (BNF instead of eMIT prices), length of hospitalisation stay (stratification by region 1 or by treatment or one single length of stay for all patients), relative dose intensity (ON/OFF) and source for BSA and body-weight data (region 1 instead of all ITT patients).

In the monotherapy model, the ICER was sensitive to the hospitalisation rate (observed, unadjusted or observed adjusted based on treatment), length of hospitalisation stay (stratification by region 1 or by treatment or one length of stay for all patients), the relative risk of post progression survival (RR= 3 or 1) and wastage (vial sharing or vial wastage). The ICER was also influenced by the method for estimating OS but to a lesser extent.

In the exploratory ERG analyses, first additional comparators as defined in the final scope were included. Based on these analyses, it seems that additional comparators IRI and FOLFIRI are on the efficient frontier for both combination therapy and monotherapy model analyses. However these analyses should be interpreted with care since several assumptions were made. The probability that RAM+PAC is the most cost-effective is negligibly small for thresholds below £100,000 in the combination therapy model. This holds true for RAM in the monotherapy model, as well.

The ERG defined scenario analyses also revealed that the ICER is sensitive to changes in the inclusion of Roy et al 2013 study in the network meta-analysis for the OS.[28] In another analysis, a comparison in which the treatment efficacy is only based on data from the RAINBOW trial, the ICER of RAM+PAC vs PAC was almost £400,000 per QALY gained.

In another analysis, it was observed that the direct implementation of RAINBOW EQ-5D trial data did not change the ICER significantly.

The ICER of RAM+PAC and RAM vs any one of the comparators was not been below £100,000 for any of the various scenario analyses and sensitivity analyses.

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

In section 5.3 the ERG base case was presented, which was based on various changes compared to the manufacturer base case. Table 6.1, 6.2 and 6.3 show how each individual change impacts the ICER plus the combined effect of all changes simultaneously.

The following changes are implemented on the model. They are explained in Appendix 3 of this report in detail.

- Correction of the docetaxel treatment cost calculation coding error
- Hospitalisation stratification is based on treatment and region 1
- BSA/weight based on region 1 are used instead of ITT population
- Double counting of hospitalisations due to AEs are corrected

Table 6.1: Revised base case cost-effectiveness analysis, incorporating corrections and amendments identified by the ERG: combination therapy RAM+PAC vs BSC

	Best supportive care		**RAM+PAC**		**Incremental**		**ICER**
	Cost	QALY	Cost	QALY	Cost	QALY	Cost per QALY gained
Manufacturer's base case analysis	**£13,400**	**0.29**	**£52,996**	**0.62**	**£39,595**	**0.33**	**£118,209**
Corrected confirmed programming errors	£13,419	0.29	£53,003	0.62	£39,584	0.33	£118,174
Corrected docetaxel treatment cost calculation coding error	£13,418	0.29	£53,003	0.62	£39,585	0.33	£118,177
Double counting of hospitalisations due to AEs are corrected	£6,642	0.29	£46,945	0.62	£40,303	0.33	£120,321
Hospitalisation rate is based on treatment and region 1	£6,693	0.29	£47,060	0.62	£40,367	0.33	£120,512
BSA/weight based on region 1 are used instead of ITT population	£6,695	0.29	£50,050	0.62	£43,354	0.33	£129,431
ERG revised base case	**£6,695**	**0.29**	**£50,050**	**0.62**	**£43,354**	**0.33**	**£129,431**

Table 6.2: Revised base case cost-effectiveness analysis, incorporating corrections and amendments identified by the ERG for combination therapy RAM+PAC vs DOC

	DOC		**RAM+PAC**		**Incremental**		**ICER**
	Cost	QALY	Cost	QALY	Cost	QALY	Cost per QALY gained
Manufacturer's base case analysis	**£18,779**	**0.39**	**£52,996**	**0.62**	**£34,217**	**0.23**	**£148,769**
Corrected confirmed programming errors	£18,849	0.39	£53,003	0.62	£34,153	0.24	£145,302
Corrected docetaxel treatment cost calculation coding error	£18,824	0.39	£53,003	0.62	£34,179	0.24	£145,412
Double counting of hospitalisations due to AEs are corrected	£10,980	0.39	£46,945	0.62	£35,965	0.24	£153,008
Hospitalisation rate is based on treatment and region 1	£10,518	0.39	£47,060	0.62	£36,542	0.24	£155,466
BSA/weight based on region 1 are used instead of ITT population	£10,523	0.39	£50,050	0.62	£39,527	0.24	£168,164
ERG revised base case	**£10,523**	**0.39**	**£50,050**	**0.62**	**£39,527**	**0.24**	**£168,164**

Monotherapy

Table 6.3: Revised base case cost-effectiveness analysis, incorporating corrections and amendments identified by the ERG: monotherapy RAM vs BSC

	Best supportive care		**RAM**		**Incremental**		**ICER**
	Cost	QALY	Cost	QALY	Cost	QALY	Cost per QALY gained
Manufacturer's base case analysis	**£14,137**	**0.31**	**£36,678**	**0.43**	**£22,542**	**0.12**	**£188,640**
Corrected confirmed programming errors	£14,168	0.31	£36,685	0.43	£22,517	0.12	£188,437
Corrected docetaxel treatment cost calculation coding error	£14,168	0.31	£36,685	0.43	£22,517	0.12	£188,437
Double counting of hospitalisations due to AEs are corrected	£6,796	0.31	£29,408	0.43	£22,612	0.12	£189,232
Hospitalisation rate is based on treatment and region 1	£6,852	0.31	£29,324	0.43	£22,472	0.12	£188,055
BSA/weight based on region 1 are used instead of ITT population	£6,853	0.31	£29,325	0.43	£22,472	0.12	£188,055
ERG revised base case	**£6,853**	**0.31**	**£29,325**	**0.43**	**£22,472**	**0.12**	**£188,055**

7 END OF LIFE

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

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

It is very unlikely that these criteria are fulfilled for ramucirumab monotherapy. In fact the company makes no claim at all that ramucirumab monotherapy should be considered an end of life treatment. This is mainly because it is unlikely that ramucirumab monotherapy offers an extension to life of at least an additional three months over best supportive care. As shown in the REGARD trial, the difference in median survival between ramucirumab monotherapy and best supportive care was 1.4 months; therefore, it is unlikely that mean survival gain will be more than three months. Based on the economic model mean additional survival of ramucirumab over best supportive care is 2.06 months.

For ramucirumab plus paclitaxel, the company does claim that the supplementary advice should be applied to this appraisal (CS, Chapter 1, pages 19-20 & Chapter 4.13, pages 131-133). Therefore, we will discuss the criteria separately:

- Life expectancy less than 24 months. This is probably correct for this population. Patients in the placebo arm of the REGARD trial had a median overall survival of 3.8 months (95% CI: 2.8 to 4.7) and the Kaplan-Meier curve shows that only two out of 117 patients are still at risk after 18 months follow-up.
- Extension to life of at least three months. This depends on the definition of current NHS treatment. The company presents additional survival in comparison with best supportive care (mean additional survival: 6.03 months) and docetaxel (mean additional survival: 4.13 months). In the response to the clarification letter (Question A9), the company also presents additional survival in comparison with paclitaxel (mean additional survival: 1.44 months).

  Based on the economic model, we were able to compute additional survival of ramucirumab plus paclitaxel versus the other comparators mentioned in the scope: irinotecan (mean additional survival: 2.27 months) and (m)FOLFIRI (mean additional survival: 1.1 months).

  Depending on the definition of current NHS treatment the criteria can be regarded fulfilled or not. If current NHS treatment only includes best supportive care and docetaxel, there is a good argument that the criteria are fulfilled. However, if current NHS treatment also includes paclitaxel, irinotecan and (m)FOLFIRI, then the criteria are not fulfilled.

  It should also be taken into account that any comparisons with treatments other than paclitaxel are based on indirect comparisons using a network of studies with considerable heterogeneity. Therefore, any estimates of additional survival gain are very uncertain.

With this in mind, the comparison with paclitaxel is most reliable; and should probably be used as the main comparator representing current NHS treatment.

- Small patient populations. This is probably correct for this population. As reported in the CS approximately 657 patients diagnosed with advanced GC/GOJ living in the United Kingdom are estimated to be eligible for second-line treatment following the first-line chemotherapy (based on 2012 data).

Based on this assessment, we do not think ramucirumab plus paclitaxel fulfils the end of life criteria.

8 OVERALL CONCLUSIONS

The ERG assessment indicated that the model was generally well presented and reported and the model structure was in general considered to be appropriate. However, a single post-progression health state is a simplification of the treatment pathway.

The ERG has no major methodological concerns on the selection of data used within the model, /application of survival modelling. However the inclusion criteria of the REGARD trial does not reflect the approved market authorisation indication for RAM monotherapy, which is the patients for whom treatment in combination with PAC is not appropriate.

Several assumptions were made in the health economic model, but the impact of assumptions was extensively explored in the deterministic sensitivity analysis. These analyses illustrate that several of the assumptions had an impact on the ICER.

As a major issue for the comparators, according to the ERG team all relevant comparators as defined in the NICE scope should be included in the analysis. The choice of comparator has a considerable impact on the ICER.

Also, the generalisability of results from the REGARD trial to the licenced indication of RAM monotherapy is highly uncertain since ineligibility for combination therapy with PAC was not an inclusion criterion in the trial.

The uncertainty around the effectiveness of the comparators is not limited to the statistical uncertainty that was explored in the univariate and probabilistic sensitivity analysis. Interval censoring adds an additional level of uncertainty. Also heterogeneity in between studies makes the NMA/ITC results less reliable. Several assumptions were made in the NMA for the HR (=1) of PFS.

8.1 Implications for research

In order to improve the robustness of the health economic outcome for UK, a direct comparison among a western population is necessary of RAM and RAM+PAC with the comparators as defined in the NICE scope. This would provide more generalisable input parameters for both effectiveness (i.e. PFS, OS, AEs) and costs (resource use of drugs, hospitalisation and subsequent therapies). In addition, quality of life data for RAM monotherapy and comparators among the targeted population (i.e. patients ineligible for PAC combination therapy) would be of great value. For the “PAC-ineligible” indication, the corresponding ineligibility should be reflected in the future clinical trials.

9 REFERENCES

[1] National Institute for Health and Care Excellence. *Single Technology Appraisal: Ramucirumab for treating advanced gastric cancer or gastro-oesophageal junction adenocarcinoma previously treated with chemotherapy. Final scope [Internet]*. London: NICE, 2015 [accessed 26.6.15]. 3p. Available from: https://www.nice.org.uk/guidance/gid-tag500/documents/gastric-cancer-metastatic-ramucirumab-after-chemotherapy-id741-final-scope2

[2] European Medicines Agency. *Public summary of opinion on orphan designation: Ramucirumab for the treatment of gastric cancer [Internet]*. London: EMA, 2015 [accessed 13.7.15]. 4p. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Orphan_designation/2012/07/WC500130310.pdf

[3] Liepa AM, Brown J, Bapat B, Kaye JA. Real-world treatment patterns of previously treated advanced gastric and gastroesophageal junction adenocarcinoma (GC) in the United Kingdom (UK). Poster presented at: Gastrointestinal Cancers Symposium; 15 – 17 Jan 2015; San Francisco: CA. *J Clin Oncol* 2015;33(suppl 3):P184.

[4] Kim R, Tan A, Choi M, El-Rayes BF. Geographic differences in approach to advanced gastric cancer: is there a standard approach? *Crit Rev Oncol Hematol* 2013;88(2):416-26.

[5] Kim J, Sun CL, Mailey B, Prendergast C, Artinyan A, Bhatia S, et al. Race and ethnicity correlate with survival in patients with gastric adenocarcinoma. *Ann Oncol* 2010;21(1):152-60.

[6] Ohtsu A, Yoshida S, Saijo N. Disparities in gastric cancer chemotherapy between the East and West. *J Clin Oncol* 2006;24(14):2188-96.

[7] NHS England. *Are older people receiving cancer drugs? An analysis of patterns in cancer drug delivery according to the age of patient [Internet]*. London: NHS England, 2013 [accessed 7.7.15]. 54p. Available from: http://www.england.nhs.uk/wp-content/uploads/2013/12/old-people-rec-cancer-drugs.pdf

[8] Eli Lilly and Company Ltd. Ramucirumab for treating metastatic gastric cancer or gastro-oesophageal junction adenocarcinoma after chemotherapy: Submission to National Institute of Health and Clinical Excellence. Single technology appraisal (STA). Eli Lilly and Company Ltd, 2015: 269.

[9] Wilke H, Muro K, Van Cutsem E, Oh SC, Bodoky G, Shimada Y, et al. Ramucirumab plus paclitaxel versus placebo plus paclitaxel in patients with previously treated advanced gastric or gastro-oesophageal junction adenocarcinoma (RAINBOW): a double-blind, randomised phase 3 trial. *Lancet Oncol* 2014;15(11):1224-35.

[10] Fuchs CS, Tomasek J, Yong CJ, Dumitru F, Passalacqua R, Goswami C, et al. Ramucirumab monotherapy for previously treated advanced gastric or gastro-oesophageal junction adenocarcinoma (REGARD): an international, randomised, multicentre, placebo-controlled, phase 3 trial. *Lancet* 2014;383(9911):31-9.

[11] Eli Lilly and Company Limited. Ramucirumab for treating metastatic gastric cancer or gastro-oesophageal junction adenocarcinoma after chemotherapy – Response to request for clarification from the ERG. Eli Lilly and Company Limited, 2015: 30.

[12] Canadian Agency for Drugs and Technologies in Health. *CADTH peer review checklist for search strategies [Internet]*. Ottawa: CADTH, 2013 [accessed 17.7.13]. 3p. Available from: http://www.cadth.ca/en/resources/finding-evidence-is

[13] National Institute for Health and Care Excellence. *Single Technology Appraisal: company evidence submission template [Internet]*. London: NICE, 2015 [accessed 11.6.15]. 28p. Available from: http://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisals/specification-for-company-submission-of-evidence-2015-version.docx

[14] National Institute for Health and Care Excellence. *Guide to the methods of technology appraisal 2013 [Internet]*. London: NICE, 2013 [accessed 7.7.15]. 93p. Available from: http://publications.nice.org.uk/pmg9

[15] Higgins JPT, Green S, eds. *Cochrane handbook for systematic reviews of interventions [Internet]*. Version 5.1.0 [updated March 2011]: The Cochrane Collaboration, 2011 [accessed 23.3.11]. Available from: http://www.cochrane-handbook.org/

[16] Al-Batran SE, Van Cutsem E, Oh Sang C, Bodoky G, Shimada Y, Hironaka S, et al. RAINBOW: A global, phase III, randomized, double-blind study of ramucirumab plus paclitaxel versus placebo plus paclitaxel patients with previously treated gastric or gastroesophageal junction (GEJ) adenocarcinoma: Quality-of-life (QoL) results. *J Clin Oncol* 2014;32(15 suppl):A4058.

[17] Al-Batran SE, Oh Sang C, Kim TY, Ohtsu A. RAINBOW: a global, phase 3, randomized, double-blind study of ramucirumab plus paclitaxel versus placebo plus paclitaxel with previously treated gastric of gastroesophageal junction (GEJ) adenocarcinoma - patient reported outcomes and performance status. *Ann Oncol* 2014;25(Suppl 2):ii105-ii117.

[18] Hironaka S, Shimada Y, Sugimoto N, Komatsu Y, Nishina T, Yamaguchi K, et al. RAINBOW: A global, phase III, randomized, double-blind study of ramucirumab (RAM) plus paclitaxel (PTX) versus placebo (PL) plus PTX in the treatment of metastatic gastroesophageal junction and gastric adenocarcinoma (mGC) following disease progression on first-line platinum- and fluoropyrimidine-containing combination therapy - efficacy analysis in Japanese and Western patients. *J Clin Oncol* 2014;32(15 suppl):A4005.

[19] Wilke H, Cunningham D, Ohtsu A, Nuber U, Bruns R, Schmitt-Bormann B. A randomized, multicenter, double-blind, placebo (PBO)-controlled phase III study of paclitaxel (PTX) with or without ramucirumab (IMC-1121B; RAM) in patients (pts) with metastatic gastric adenocarcinoma, refractory to or progressive after first-line therapy with platinum (PLT) and fluoropyrimidine (FP). *J Clin Oncol* 2012;30(Suppl):TPS4139.

[20] Wilke H, Ko A, Emig M, Chandrawansa K, Clingan P, Ananda S, et al. RAINBOW: A global, phase 3, double-blind study of ramucirumab plus paclitaxel versus placebo plus

paclitaxel in the treatment of gastric cancer following disease progression: Western population subgroup. *Ann Oncol* 2014;25(suppl 2):ii105-ii117.

[21] Wilke H, Van Cutsem E, Oh SC, Bodoky G, Shimada Y, Hironaka S, et al. RAINBOW: A global, phase 3, randomized, double-blind study of ramucirumab plus paclitaxel versus placebo plus paclitaxel in the treatment of metastatic gastric adenocarcinoma following disease progression on first-line platinum- and fluoropyrimidine-containing combination therapy: results of a multiple Cox regression analysis adjusting for prognostic factors. *J Clin Oncol* 2014;32(15 suppl):A4076.

[22] Wilke H. RAINBOW: A global, phase III, randomized, double-blind study of ramucirumab plus paclitaxel versus placebo plus paclitaxel in the treatment of metastatic gastroesophageal junction (GEJ) and gastric adenocarcinoma following disease progression on first-line platinum-and fluoropyrimidine-containing combination therapy rainbow IMCL CP12-0922 (I4T-IE-JVBE) [Internet] [Oral presentation]. 2014 [accessed 9.7.15]. Available from: http://meetinglibrary.asco.org/content/90231

[23] Chau I, Passalacqua R, Zalcberg JR, Fuchs CS, Liepa AM, Hsu Y, et al. Tolerability and quality-of-life (QoL) results from the phase 3 REGARD study: ramucirumab versus placebo in patients with previously treated gastric or gastroesophageal junction (GEJ) adenocarcinoma. Poster presented at European Cancer Congress 2013; 27 Sept - 1 Oct 2013; Amsterdam: Netherlands. 2013.

[24] Fuchs CS, Tomasek J, Cho JY, Tomasello G, Goswami C, Vieira dos Santos L, et al. REGARD: A phase 3, randomized, double-blind trial of ramucirumab (RAM) and best supportive care (BSC) versus placebo (PL) and BSC in the treatment of metastatic gastric or gastroesophageal junction (GEJ) adenocarcinoma following disease progression (PD) on first-line platinum- and/or fluoropyrimidine-containing combination therapy: age subgroup analysis. *J Clin Oncol* 2014;32(15 suppl):A4057.

[25] Fuchs CS, Tomasek J, Cho JY, Dumitru F, Passalacqua R, Goswami C, et al. REGARD: A phase III, randomized, double-blinded trial of ramucirumab and best supportive care (BSC) versus placebo and BSC in the treatment of metastatic gastric or gastroesophageal junction (GEJ) adenocarcinoma following disease progression on first-line platinum- and/or fluoropyrimidine-containing combination therapy. *Journal Clinical Oncology* 2013;31(Suppl 4):LBA5.

[26] Sym SJ, Hong J, Park J, Cho EK, Lee JH, Park YH, et al. A randomized phase II study of biweekly irinotecan monotherapy or a combination of irinotecan plus 5-fluorouracil/leucovorin (mFOLFIRI) in patients with metastatic gastric adenocarcinoma refractory to or progressive after first-line chemotherapy. *Cancer Chemother Pharmacol* 2013;71(2):481-8.

[27] Ford HE, Marshall A, Bridgewater JA, Janowitz T, Coxon FY, Wadsley J, et al. Docetaxel versus active symptom control for refractory oesophagogastric adenocarcinoma (COUGAR-02): an open-label, phase 3 randomised controlled trial. *Lancet Oncol* 2014;15(1):78-86.

[28] Roy AC, Park SR, Cunningham D, Kang YK, Chao Y, Chen LT, et al. A randomized phase II study of PEP02 (MM-398), irinotecan or docetaxel as a second-line therapy in patients with locally advanced or metastatic gastric or gastro-oesophageal junction adenocarcinoma. *Ann Oncol* 2013;24(6):1567-73.

[29] Hironaka S, Ueda S, Yasui H, Nishina T, Tsuda M, Tsumura T, et al. Randomized, open-label, phase III study comparing irinotecan with paclitaxel in patients with advanced gastric cancer without severe peritoneal metastasis after failure of prior combination chemotherapy using fluoropyrimidine plus platinum: WJOG 4007 trial. *J Clin Oncol* 2013;31(35):4438-44.

[30] Thuss-Patience PC, Kretzschmar A, Bichev D, Deist T, Hinke A, Breithaupt K, et al. Survival advantage for irinotecan versus best supportive care as second-line chemotherapy in gastric cancer--a randomised phase III study of the Arbeitsgemeinschaft Internistische Onkologie (AIO). *Eur J Cancer* 2011;47(15):2306-14.

[31] Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin Epidemiol* 1997;50(6):683-91.

[32] Kang JH, Lee SI, Lim do H, 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.

[33] International Agency for Research on Cancer. GLOBOCAN 2012: estimated cancer incidence, mortality and prevalence worldwide in 2012, stomach cancer 2012 [Internet]. 2012 [accessed 7.7.15]. Available from: http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx

[34] Leung WK, Wu MS, Kakugawa Y, Kim JJ, Yeoh KG, Goh KL, et al. Screening for gastric cancer in Asia: current evidence and practice. *Lancet Oncol* 2008;9(3):279-87.

[35] Han H, S., Lauwers GY. Her 2 in Gastric Carcinoma [Internet]. 2010 [accessed 7.7.15]. Available from: http://www.dako.com/28830_14dec10_connection_15_her2_in_gastric_cancer.pdf.

[36] Shiroiwa T, Fukuda T, Shimozuma K. Cost-effectiveness analysis of trastuzumab to treat HER2-positive advanced gastric cancer based on the randomised ToGA trial. *Br J Cancer* 2011;105(9):1273-8.

[37] Wu B, Ye M, Chen H, Shen JF. Costs of trastuzumab in combination with chemotherapy for HER2-positive advanced gastric or gastroesophageal junction cancer: an economic evaluation in the Chinese context. *Clin Ther* 2012;34(2):468-79.

[38] National Institute for Health and Care Excellence. *Trastuzumab for the treatment of HER2-positive metastatic gastric cancer. NICE technology appraisal guidance 208 [Internet]*. London: NICE, October 2009, 2010 [accessed 26.6.15]. 47p. Available from: https://www.nice.org.uk/guidance/ta208/resources/guidance-trastuzumab-for-the-treatment-of-her2positive-metastatic-gastric-cancer-pdf

[39] Philips Z, Ginnelly L, Sculpher M, Claxton K, Golder S, Riemsma R, et al. Review of guidelines for good practice in decision-analytic modelling in health technology assessment. *Health Technol Assess* 2004;8(36):iii-iv, ix-xi, 1-158.

[40] Latimer N. *NICE DSU Technical Support Document 14: Undertaking survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data*, 2011 [last updated March 2013] Available from: http://www.nicedsu.org.uk

[41] Cancer Research UK. Stomach Cancer Symptoms [Internet]. 2014 [accessed 17.6.15]. Available from: http://www.cancerresearchuk.org/about-cancer/type/stomach-cancer/about/stomach-cancer-symptoms

[42] Al-Batran SE, Ajani JA. Impact of chemotherapy on quality of life in patients with metastatic esophagogastric cancer. *Cancer* 2010;116(11):2511-8.

[43] 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.

[44] Aultman R, Urspruch A. PCN116 analysis of the EQ-5D questionnaire for patients with advanced HER2-positive gastric cancer (HER2 AGC) based on the trastuzumab for gastric cancer (TOGA) study. *Value Health* 2010;13(7):A273.

[45] O'Gorman P, McMillan DC, McArdle CS. Impact of weight loss, appetite, and the inflammatory response on quality of life in gastrointestinal cancer patients. *Nutr Cancer* 1998;32(2):76-80.

[46] Dolan P. Modeling valuations for EuroQol health states. *Med Care* 1997;35(11):1095-108.

[47] National Institute for Health and Care Excellence. *Pixantrone monotherapy for treating multiply relapsed or refractory aggressive non-Hodgkin's B cell lymphoma. NICE technology appraisal guidance 306 [Internet]*. London: NICE, October 2009, 2014 [accessed 26.6.15]. 76p. Available from: http://www.nice.org.uk/guidance/ta306

[48] Swinburn P, Lloyd A, Nathan P, Choueiri TK, Cella D, Neary MP. Elicitation of health state utilities in metastatic renal cell carcinoma. *Curr Med Res Opin* 2010;26(5):1091-6.

[49] Lloyd A, Nafees B, Narewska J, Dewilde S, Watkins J. Health state utilities for metastatic breast cancer. *Br J Cancer* 2006;95(6):683-90.

[50] Nafees B, Stafford M, Gavriel S, Bhalla S, Watkins J. Health state utilities for non small cell lung cancer. *Health Qual Life Outcomes* 2008;6:84.

[51] Coyle D, Small N, Ashworth A, Hennessy S, Jenkins-Clarke S, Mannion R, et al. Costs of palliative care in the community, in hospitals and in hospices in the UK. *Crit Rev Oncol Hematol* 1999;32(2):71-85.

[52] Pettengell R, Coiffier B, Narayanan G, de Mendoza FH, Digumarti R, Gomez H, et al. Pixantrone dimaleate versus other chemotherapeutic agents as a single-agent salvage treatment in patients with relapsed or refractory aggressive non-Hodgkin lymphoma: a phase 3, multicentre, open-label, randomised trial. *Lancet Oncol* 2012;13(7):696-706.

[53] Joint Formulary Committee. *British National Formulary. 67*. 67th ed. London: BMJ Group and Pharmaceutical Press, 2014.

[54] Eli Lilly and Company Ltd. Summary of Product Characteristics: Cyramza 10 mg/ml concentrate for solution for infusion [Internet]. 2015 [accessed 19.1.15]. Available from: http://www.medicines.org.uk/emc/medicine/29765.

[55] Ltd HU. Summary of Product Characteristics: Paclitaxel 6mg/ml concentrate for solution for infusion [Internet]. 2014 [accessed 7.7.15]. Available from: https://www.medicines.org.uk/emc/medicine/15842#

[56] Medicines Complete. British National Formulary [Internet]. 2015 [accessed 24.4.2015]. Available from: https://www.medicinescomplete.com/mc/bnf/current/

[57] Eli Lilly and Company Ltd. Market Research Data [Company Confidential]. 2014.

[58] National Institute for Health and Care Excellence. *Aflibercept in combination with irinotecan and fluorouracil-based therapy for treating metastatic colorectal cancer that has progressed following prior oxaliplatin-based chemotherapy. NICE technology appraisal guidance 307 [Internet]*. London: NICE, 2014 [accessed 15.7.15]. 65p. Available from: http://www.nice.org.uk/guidance/ta307

Appendix 1: ERG Search Strategies

Clinical effectiveness

The ERG also noted that only a limited number of free text terms were used in the search strategy, the search could have perhaps been more sensitive with the use of additional text terms for the condition or the site of the condition such as:

"GEJ", "GOJ", "esophagogastric junction", "oesophago gastric junctional cancer", "cardia".

Due to time constraints, the ERG has not had the opportunity to investigate what effect using these additional text terms might have had on recall.

An alternative search approach could have been to just search for the intervention drug, this line in the company search strategy retrieved 640 records for all the resources selected, and this is not an unreasonable number to screen for relevant studies. Searching for the drug alone could have also informed the AE and non-RCT sections of the submission.

Cost-effectiveness

No host or specific database date span is given for the Embase or NHS Economic Evaluation Database searches. The line numbering for the Embase search seems have been altered by formatting actions, making line combinations a little confusing.

Measurement and value of health effects

The NHS EED search (CS Appendix 10, Table 28) date span is up to 1st Quarter 2013, although the search was run in September 2013; this is quite outdated and may have benefited from being updated.

Appendix 2: Phillips et al checklist

Results of assessing the manufacturers report based on the checklist by Phillips et al.

1. Is there a clear statement of the decision problem?

Yes, the decision problem is clearly stated.

2. Is the objective of the evaluation and model specified and consistent with the stated decision problem?

The objective of the evaluation and model is the cost-effectiveness of second-line treatment with ramucirumab in adult patients with advanced GC/GOJ both in combination with PAC and as monotherapy. Monotherapy (RAM) is licensed for patients for whom further treatment in combination with [PAC is not appropriate.

3. Is the primary decision-maker specified?

The term is not used, but implicitly the NHS is assumed

4. Is the perspective of the model stated clearly?

Yes, it is the perspective NHS.

5. Are the model inputs consistent with the stated perspective?

No, utility decrements for AEs were obtained from the literature.

6. Has the scope of the model been stated and justified?

Yes

7. Are the outcomes of the model consistent with the perspective, scope and overall objective of the model?

Yes apart from the deviation of the included comparators (PAC, IRI, FOLFIRI).

8. Is the structure of the model consistent with a coherent theory of the health condition under evaluation?

Yes, although a single post-progression health state is a simplification of the treatment pathway.

9. Are the sources of data used to develop the structure of the model specified?

Yes

10. Are the causal relationships described by the model structure justified appropriately?

Yes

11. Are the structural assumptions transparent and justified?

Yes

12. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?

Yes.

13. Is there a clear definition of the options under evaluation?

Yes

14. Have all feasible and practical options been evaluated?

No, not all possible options have been evaluated. Several of the comparators from the NICE scope were excluded.

15. Is there justification for the exclusion of feasible options?

Yes, a justification was given that the proportion of patients receiving these treatments was considered too small. However, the ERG does not consider these exclusions defensible.

16. Is the chosen model type appropriate given the decision problem and specified causal relationships within the model?

Yes

17. Is the time horizon of the model sufficient to reflect all important differences between options?

Yes, it is 7.23 years and considering the median OS ranges from 3.8-9.63 months, this seems to reflect a lifetime horizon.

18. Are the time horizon of the model, the duration of treatment and the duration of treatment effect described and justified?

Yes

19. Do the disease states (state transition model) or the pathways (decision tree model) reflect the underlying biological process of the disease in question and the impact of interventions?

Yes

20. Is the cycle length defined and justified in terms of the natural history of disease?

Yes, it is defined, but justified based on the monitoring frequency of ramucirimab.

21. Are the data identification methods transparent and appropriate given the objectives of the model?

Yes

22. Where choices have been made between data sources, are these justified appropriately?

Yes

23. Has particular attention been paid to identifying data for the important parameters in the model?

Yes

24. Has the quality of the data been assessed appropriately?

Yes

25. Where expert opinion has been used, are the methods described and justified?

No, in the CS, details on the clinical expert consultation were absent. However, additional information was provided in the clarification letter.

26. Is the data modelling methodology based on justifiable statistical and epidemiological techniques?

Yes

27. Is the choice of baseline data described and justified?

Yes

28. Are transition probabilities calculated appropriately?

Yes

29. Has a half-cycle correction been applied to both cost and outcome?

Yes (a small programming error required correction for a few cost elements)

30. If not, has this omission been justified?

N/A

31. If relative treatment effects have been derived from trial data, have they been synthesised using appropriate techniques?

Yes

32. Have the methods and assumptions used to extrapolate short-term results to final outcomes been documented and justified?

Mostly yes, Interval censoring method is not clear.

33. Have alternative extrapolation assumptions been explored through sensitivity analysis?

Yes

34. Have assumptions regarding the continuing effect of treatment once treatment is complete been documented and justified?

N/A

35. Have alternative assumptions regarding the continuing effect of treatment been explored through sensitivity analysis?

N/A

36. Are the costs incorporated into the model justified?

Yes

37. Has the source for all costs been described?

Yes, although the ERG could not verify the incidence of all AEs

38. Have discount rates been described and justified given the target decision-maker?

Yes

39. Are the utilities incorporated into the model appropriate?

Yes, for combination therapy the baseline utilities are appropriate. The utility decrements were obtained from the literature.

No, for monotherapy, the baseline utilities were obtained from a potentially different population receiving other treatments.

40. Is the source for the utility weights referenced?

Yes

41. Are the methods of derivation for the utility weights justified?

Yes

42. Have all data incorporated into the model been described and referenced in sufficient detail?

Yes

43. Has the use of mutually inconsistent data been justified (i.e. are assumptions and choices appropriate)?

N/A

44. Is the process of data incorporation transparent?

Yes

45. If data have been incorporated as distributions, has the choice of distribution for each parameter been described and justified?

Yes

46. If data have been incorporated as distributions, is it clear that second order uncertainty is reflected?

Yes

47. Have the four principal types of uncertainty been addressed?

No

Methodological uncertainty is not discussed. This is acceptable given the fact that the analysis should be performed according to the NICE reference case

Structural uncertainty is explored through the deterministic sensitivity analysis (i.e. different scenarios were explored).

Heterogeneity: Heterogeneity of the trials included in the ITC/NMA has not been addressed sufficiently. In the combination therapy model, a scenario analysis for Region 1 (i.e. patients from North America, Europe, Australia and Israel) was undertaken where OS, PFS and time on treatment were based on Region 1 patients only. No subgroup analysis were performed for the monotherapy model.

Parameter uncertainty has been assessed in the PSA.

48. If not, has the omission of particular forms of uncertainty been justified?

No

49. Have methodological uncertainties been addressed by running alternative versions of the model with different methodological assumptions?

No. This is acceptable given the fact that the analysis should be performed according to the NICE reference case

50. Is there evidence that structural uncertainties have been addressed via sensitivity analysis?

Yes

51. Has heterogeneity been dealt with by running the model separately for different subgroups?

No, in the combination therapy model, only OS, PFS and treatment duration were based on region 1 and not for example body weight or hospitalisations.

No subgroup analysis were performed for the monotherapy model.

52. Are the methods of assessment of parameter uncertainty appropriate?

Yes

53. If data are incorporated as point estimates, are the ranges used for sensitivity analysis stated clearly and justified?

Yes

54. Is there evidence that the mathematical logic of the model has been tested thoroughly before use?

No

55. Are any counterintuitive results from the model explained and justified?

No counterintuitive results occurred

56. If the model has been calibrated against independent data, have any differences been explained and justified?

No, the model has not been calibrated

57. Have the results of the model been compared with those of previous models and any differences in results explained?

No prior models have been discussed in the submission.

Appendix 3: Details model changes implemented by the ERG

These corrections are applied for both combination and monotherapy

Issue	Sheet	Cell/ name	Old formula	New formula	Comment
Prog error	Drug Costs	Q21	“eMT”	"eMIT"	Has effect on the minimum DOC cost selection
Double counting of hospitalisations due to AEs are corrected	Developer	H73	“0”	"1”	
Hospitalisation rate is based on treatment and region 1	Developer	F50	“2”	“4”	
BSA/weight based on region 1 are used instead of ITT population	Developer	F42	“2”	“1”	

Appendix 4: PSA for monotherapy

Figure 1: Cost-effectiveness acceptability curves for monotherapy

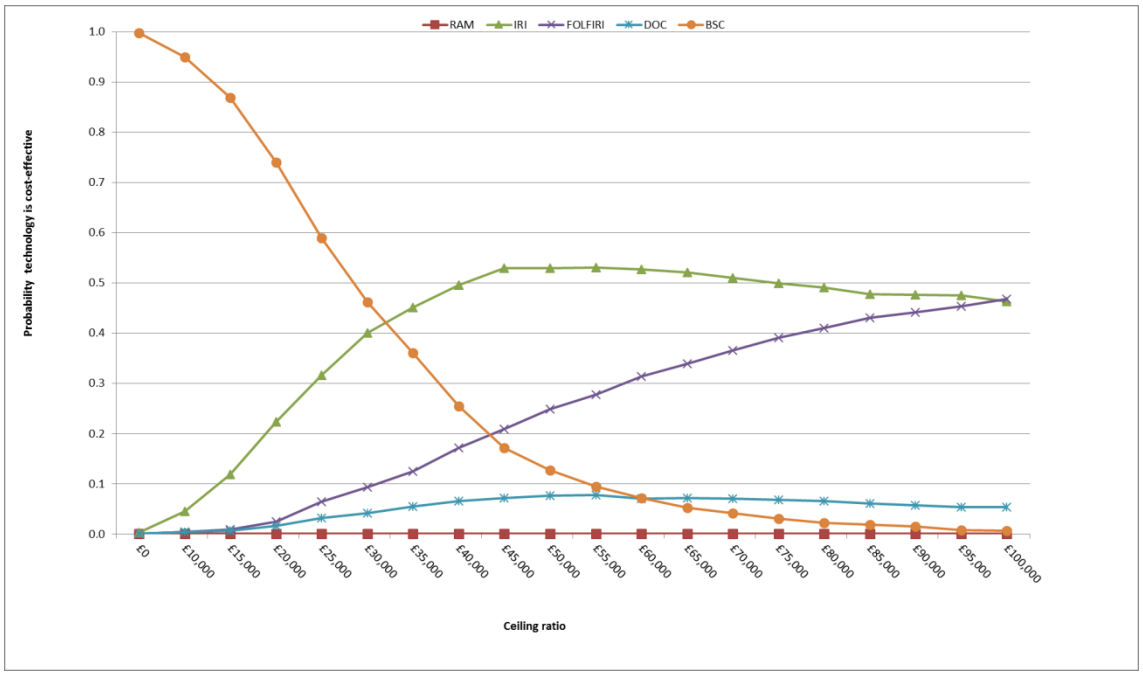
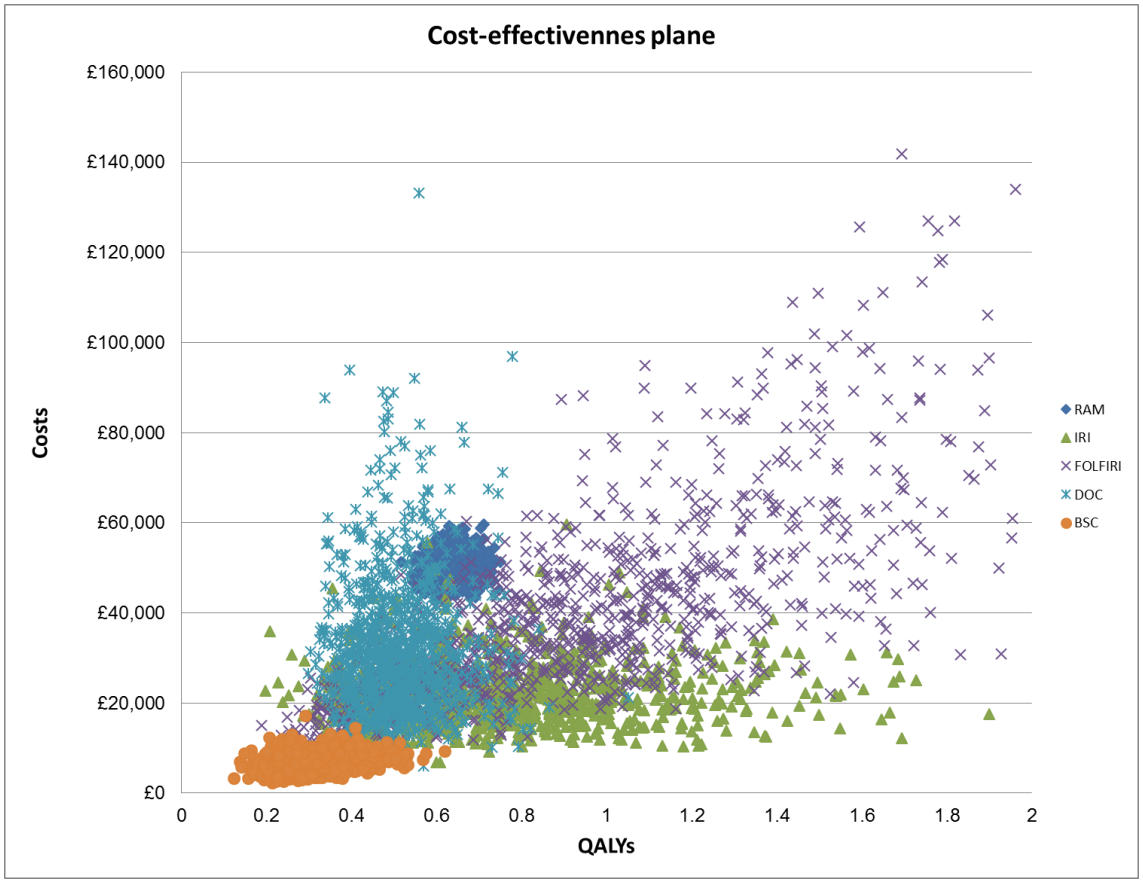



Figure 2: Cost-effectiveness plane for monotherapy (QALYs)