

Ramucirumab for Treating Advanced Gastric Cancer or Gastro-Oesophageal Junction Adenocarcinoma Previously Treated with Chemotherapy: An Evidence Review Group Perspective of a NICE Single Technology Appraisal

Nasuh C. Büyükkaramikli¹ · Hedwig M. Blommestein¹ · Rob Riemsma² · Nigel Armstrong² · Fiona J. Clay^{3,4} · Janine Ross² · Gill Worthy² · Johan Severens¹ · Jos Kleijnen^{2,5} · Maiwenn J. Al¹

Published online: 27 June 2017

© The Author(s) 2017. This article is an open access publication

Abstract The National Institute for Health and Care Excellence (NICE) invited the company that manufactures ramucirumab (Cyramza[®], Eli Lilly and Company) to submit evidence of the clinical and cost effectiveness of the drug administered alone (monotherapy) or with paclitaxel (combination therapy) for treating adults with advanced gastric cancer or gastro-oesophageal junction (GC/GOJ) adenocarcinoma that were previously treated with chemotherapy, as part of the Institute's single technology appraisal (STA) process. Kleijnen Systematic Reviews Ltd (KSR), in collaboration with Erasmus University Rotterdam, was commissioned to act as the Evidence Review Group (ERG). This paper describes the company's submission, the ERG review, and NICE's subsequent decisions. Clinical effectiveness evidence for ramucirumab monotherapy (RAM), compared with best supportive care (BSC), was based on data from the REGARD trial. Clinical effectiveness evidence for ramucirumab combination therapy (RAM + PAC), compared with paclitaxel monotherapy (PAC), was based on data from the

RAINBOW trial. In addition, the company undertook a network meta-analysis (NMA) to compare RAM + PAC with BSC and docetaxel. Cost-effectiveness evidence of monotherapy and combination therapy relied on partitioned survival, cost-utility models. The base-case incremental cost-effectiveness ratio (ICER) of the company was £188,640 (vs BSC) per quality-adjusted life-year (QALY) gained for monotherapy and £118,209 (vs BSC) per QALY gained for combination therapy. The ERG assessment indicated that the modelling structure represented the course of the disease; however, a few errors were identified and some of the input parameters were challenged. The ERG provided a new base case, with ICERs (vs BSC) of £188,100 (monotherapy) per QALY gained and £129,400 (combination therapy) per QALY gained and conducted additional exploratory analyses. The NICE Appraisal Committee (AC), considered the company's decision problem was in line with the NICE scope, with the exception of the choice of comparators for the combination therapy model. The most plausible ICER for ramucirumab monotherapy compared with BSC was £188,100 per QALY gained. The Committee considered that the ERG's exploratory analysis in which RAM + PAC was compared with PAC by using the direct head-to-head data (including utilities) from the RAINBOW trial, provided the most plausible ICER (i.e. £408,200 per QALY gained) for ramucirumab combination therapy. The Committee concluded that end-of-life considerations cannot be applied for either case, since neither failed to offer an extension to life of at least 3 months. The company did not submit a patient access scheme (PAS). After consideration of the evidence, the Committee concluded that ramucirumab alone or with paclitaxel could not be considered a cost-effective use of National Health Service resources for treating advanced GC/GOJ patients that were previously treated with

✉ Nasuh C. Büyükkaramikli
buyukkaramikli@imta.eur.nl

¹ Institute for Medical Technology Assessment (iMTA),
Institute of Health Policy and Management (iBMG), Erasmus
University Rotterdam, Rotterdam, The Netherlands

² Kleijnen Systematic Reviews Ltd, York, UK

³ Department of Psychiatry, University of Melbourne,
Melbourne, Australia

⁴ Department of Forensic Medicine, Monash University,
Melbourne, Australia

⁵ Department of Family Medicine, School for Public Health
and Primary Care (CAPHRI), Maastricht University,
Maastricht, The Netherlands

chemotherapy, and therefore its use could not be recommended. We might wonder if a complete STA process is necessary for treatments without a PAS, which are, according to the company's submission, already associated with ICERs far above the currently accepted threshold in all (base-case, sensitivity and scenario) analyses.

Key Points for Decision Makers

Ramucirumab (RAM) monotherapy and combination therapy seem to be clinically effective for treating advanced gastric cancer or gastro-oesophageal junction adenocarcinoma patients who were previously treated by chemotherapy, as both treatments (monotherapy and combination therapy) lead to limited increase in overall survival and progression-free survival compared with best supportive care (BSC) and paclitaxel alone, respectively, from REGARD and RAINBOW trials.

The evidence review group (ERG) considered that other comparators than BSC and docetaxel, which were mentioned in the scoping document (e.g. paclitaxel, irinotecan etc.) should also have been included. Furthermore, the ERG mentioned that the results of the indirect treatment comparison (ITC) should be interpreted with caution, due to considerable heterogeneity between the trials and the exclusion of potentially relevant trials from the ITC by the company. Finally, the ERG had concerns on the generalizability of the evidence to the UK patient population.

The Appraisal Committee (AC) considered that paclitaxel is the most appropriate comparator for the combination therapy, as its comparative effectiveness is based on direct evidence. The AC also considered some of the inputs should be adjusted for the UK population.

In the end, it was concluded that ramucirumab alone or with paclitaxel could not be considered a cost-effective use of National Health Service resources.

In this appraisal, all incremental cost-effectiveness ratios (ICERs) from the company submission (base-case, sensitivity/scenario analyses) were far above the currently accepted threshold. For these cases, a faster procedure appraisal might be more efficient.

1 Introduction

The National Institute for Health and Care Excellence (NICE) is an independent organization providing national guidance on promoting good health and preventing and treating ill health [1]. The single technology appraisal (STA) process is designed to provide recommendations on a single product, device or other technology with a single indication. The process covers new technologies and enables NICE to produce guidance shortly after the technology is introduced into the UK. The NICE Appraisal Committee (AC) obtains relevant evidence from several sources: the company submission (CS), a report from the appointed independent Evidence Review Group (ERG) and advice from consultees (i.e. patients, experts and other stakeholders). The CS includes a written report and a mathematical model that describes the clinical and cost effectiveness of the technology under investigation. The ERG, an external organization independent of the NICE, reviews the CS and produces a summary report and provides a critique of the submitted evidence. After consideration of all the relevant evidence, the AC formulates preliminary guidance in the form of the Appraisal Consultation Document (ACD) as to whether or not to recommend the intervention. The stakeholders are invited to comment on this ACD and the submitted evidence. A subsequent ACD may be produced or a Final Appraisal Determination (FAD) is issued. Once published, NICE technology guidance provides a legal obligation for NHS providers to reimburse technologies that have been approved. This paper presents a summary of the ERG report and the development of NICE guidance based on the findings of the AC for the STA of ramucirumab (alone or in combination with paclitaxel), for treating advanced gastric cancer or gastro-oesophageal junction (GC/GOJ) adenocarcinoma previously treated with chemotherapy. Full details of all the relevant appraisal documents can be found on the NICE website [2].

2 The Decision Problem

The indication GC/GOJ refers to cancers that originate in the lining of the stomach and the gastro-oesophageal junction. GC/GOJs are rare (designated orphan status by the European Medicines Agency [EMA]) and aggressive cancers. The annual incidence of GC/GOJ is low in the UK; in 2012, there were 5637 new cases of GC and 3085 new cases of GOJ [3]. Of patients diagnosed with GC,

approximately 80% are diagnosed with advanced, metastatic disease [4]. The prognosis in this group is very poor with a 5-year survival rate of approximately 5% [5].

For inoperable patients with advanced GC/GOJ, chemotherapy is administered. The standard first-line treatment in the UK is a regimen comprising a fluoropyrimidine and a platinum agent, with or without an anthracycline [6]. For GC/GOJ, there are two NICE guidance documents available, capecitabine in combination with a platinum-based regimen [7], and trastuzumab in combination with fluoropyrimidine [8], both recommended for first-line treatment.

There is currently no established second-line treatment and as a consequence, second-line treatments for GC/GOJ patients vary [6]. Docetaxel, paclitaxel (PAC) and irinotecan are among the therapies that are administered.

Ramucirumab is a human receptor-targeted monoclonal antibody that specifically binds to the vascular endothelial growth factor receptor. It is approved and has been designated orphan status by the EMA [10, 11] for the treatment of adult patients with advanced GC/GOJ with disease progression after prior platinum or fluoropyrimidine chemotherapy:

- in combination with paclitaxel (RAM + PAC);
- as monotherapy (RAM), in patients for whom treatment in combination with PAC is not appropriate.

The remit of this appraisal was specified by NICE's final scope [2], which was to assess the clinical and cost effectiveness of ramucirumab within its licensed indication. In line with its license, the scope addressed the use of RAM + PAC as well as RAM; the latter option being for patients in which treatment in combination with PAC is inappropriate. The comparators listed in the scope were docetaxel monotherapy, irinotecan monotherapy, irinotecan- and fluorouracil-based therapy (FOLFIRI), PAC monotherapy and best supportive care (BSC).

3 The Independent Evidence Review Group (ERG) Report

Kleijnen Systematic Reviews Ltd (KSR), in collaboration with Erasmus University Rotterdam, acted as the ERG, and reviewed the evidence on the clinical and cost effectiveness of RAM + PAC and RAM among adults with advanced GC/GOJ, who were previously treated with chemotherapy as submitted by the company (Eli Lilly and Company).

The review embodied three aims:

- to assess whether the CS conformed to the methodological guidelines issued by NICE [1];
- to assess whether the company's interpretation and analysis of the evidence were appropriate;

- to indicate the presence of other sources of evidence or alternative interpretations of the evidence that could inform NICE guidance.

The ERG critically reviewed the evidence in the CS, in the response to clarification questions and evidence provided after the publication of the ACD. Furthermore, it conducted additional literature searches, explored the impact of assumptions on the incremental cost-effectiveness ratio (ICER), revised the economic model and explored additional scenario analyses.

3.1 Summary of the Clinical Evidence

The CS included a systematic review of the literature on the clinical/cost effectiveness of ramucirumab.

3.1.1 Ramucirumab Monotherapy

Estimation of the efficacy of RAM compared with BSC relied on the REGARD trial (Ramucirumab monotherapy for previously treated advanced gastric or gastro-oesophageal junction adenocarcinoma) [9]. In this global, multicentre trial, adult patients with advanced GC/GOJ, who progressed after chemotherapy ($n = 355$), were randomized (2:1) to receive BSC and either ramucirumab (8 mg/kg administered intravenously [IV] every 2 weeks) or placebo.

Randomization was stratified by weight loss, geographic location and location of the primary tumour. Overall survival (OS) was the primary outcome. Key secondary outcomes included progression-free survival (PFS), overall response rate (ORR) and quality of life (QoL). All analyses were performed using the intention-to-treat (ITT) population. The OS and PFS results are given in Table 1.

Health-related QoL in REGARD was assessed using the European Organisation for Research and Treatment of Cancer (EORTC) QLQ-C30 instrument [10]. At 6 weeks, the proportion of patients with improved or stable QoL was higher for the ramucirumab arm (34.1%) than for the placebo arm (13.7%); but the difference between these two arms was not statistically significant ($p = 0.23$).

Overall safety results from REGARD showed that 45% of the patients in the ramucirumab arm had at least one serious adverse event (AE), compared with 44% of the patients in the placebo arm. The proportion of patients who stopped treatment was 10.5% in the ramucirumab arm and 6% in the placebo arm.

3.1.2 Ramucirumab plus Paclitaxel

Estimation of the efficacy of RAM + PAC compared with PAC relied on one global, multi-centre trial (RAINBOW;

Table 1 OS and PFS results from the REGARD study

Median OS RAM (months)	Median OS BSC (months)	HR (RAM vs BSC)
5.2 (95% CI 4.4–5.7)	3.8 (95% CI 2.8–4.7)	0.776 (95% CI 0.603–0.998)
Median PFS RAM (months)	Median PFS BSC (months)	HR (RAM vs BSC)
2.1 (95% CI 1.5–2.7)	1.3 (95% CI 1.3–1.4)	0.483 (95% CI 0.376–0.620)

BSC Best supportive care, *HR* hazard ratio, *OS* overall survival, *PFS* progression-free survival, *RAM* ramucirumab monotherapy

Table 2 OS and PFS results from the RAINBOW study

Median OS RAM + PAC (months)	Median OS PAC (months)	HR (RAM + PAC vs PAC)
9.63 (95% CI 8.6–10.8)	7.36 (95% CI 6.3–8.4)	0.807 (95% CI 0.678–0.962)
Median PFS RAM + PAC (months)	Median PFS PAC (months)	HR (RAM + PAC vs PAC)
4.4 (95% CI 4.2–5.3)	2.9 (95% CI 2.8–3.0)	0.635 (95% CI 0.536–0.752)

HR Hazard ratio, *OS* overall survival, *PAC* paclitaxel monotherapy, *PFS* progression-free survival, *RAM + PAC* ramucirumab combination therapy

Ramucirumab plus paclitaxel versus placebo plus paclitaxel in patients with previously treated advanced gastric or gastro-oesophageal junction adenocarcinoma [11], in which patients with advanced GC/GOJ who had disease progression after chemotherapy ($n = 665$) were randomized (1:1) to receive RAM 8 mg/kg plus PAC 80 mg/kg, or placebo plus PAC 80 mg/kg administered IV. RAM was given every 2 weeks whereas PAC was given on Days 1, 8 and 15 of each 28-day cycle.

Randomization was stratified by geographic location, time to progression from the start of first-line chemotherapy (<6 months or not) and disease measurability (measurable or not). The primary outcome was OS, key secondary outcomes included PFS, ORR and QoL. All analyses were performed on the ITT population. The OS and PFS results are given in Table 2.

Health-related QoL in RAINBOW was also assessed using the EORTC QLQ-C30 instrument [10]. RAM + PAC was associated with statistically significant improved outcomes for two symptom scales (i.e. emotional function and nausea and vomiting), compared with PAC.

A similar percentage of patients in both arms stopped treatment because of AEs (11.8% in the RAM + PAC arm and 11.3% in the PAC arm). The most frequently reported treatment-emergent serious AE was neutropenia, 54.4% in the RAM + PAC arm and 31.0% in the PAC arm.

3.1.3 Network Meta-Analysis (NMA)

The company carried out a NMA to compare the OS of RAM + PAC with BSC and docetaxel. The company identified 23 trials for inclusion in the systematic review; however, only the five trials listed in Table 3 were included in the NMA of OS. The other 18 trials (out of the identified

Table 3 Trials that were included in the network meta-analysis

Trial	Treatment arms
RAINBOW [11]	RAM + PAC vs PAC
COUGAR-02 [13]	Docetaxel vs BSC
Roy et al. [14] ^a	Docetaxel vs irinotecan
Hironaka et al. [15]	PAC vs irinotecan
Thuss-Patience et al. [16]	Irinotecan vs BSC
Sym et al. [12]	Irinotecan vs FOLFIRI

BSC Best supportive care, *FOLFIRI* fluorouracil-based therapy, *PAC* paclitaxel monotherapy, *RAM + PAC* ramucirumab combination therapy

^a Roy et al. [14] was included only in one of the sensitivity analyses, as it was not a phase III trial

23) were not included as they were considered to compare treatments that were not in the scope [2]. However, among these excluded studies, Sym et al. [12] compared irinotecan with FOLFIRI, both of which were listed as comparators in the scope. Upon request from the ERG, in the company's response to clarification letter, FOLFIRI was incorporated into the evidence network using Sym et al. [12].

The hazard ratio (HR) results from the OS NMA are given in Table 4.

In order to compare the PFS of RAM + PAC with BSC and docetaxel, additional assumptions were needed since some studies did not report PFS HRs [13, 14, 16]. Using the PFS HRs reported [9, 11, 15] and assuming HR = 1 (standard error [SE] = 0.01) between RAM, irinotecan and docetaxel, the results from the NMA suggested that RAM + PAC was associated with a statistically significantly improved PFS compared with BSC (HR 0.27; 95% CI 0.14–0.53) and docetaxel (HR 0.56; 95% CI 0.41–0.76).

Table 4 Hazard ratio results from the overall survival network meta-analysis

	Hazard ratio (95% CI)				
	BSC	Docetaxel	PAC	Irinotecan	FOLFIRI
RAM + PAC	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)

BSC Best supportive care, FOLFIRI fluorouracil-based therapy, PAC paclitaxel monotherapy, RAM + PAC ramucirumab combination therapy

3.2 Critique of the Clinical Evidence

3.2.1 Ramucirumab Monotherapy

REGARD was deemed to be a good-quality randomized controlled trial. The uncertainty about long-term efficacy was considered to be small, because both OS and PFS data were mature.

In the final scope, RAM monotherapy was indicated for patients who were not suitable for PAC. However, in REGARD's eligibility criteria, nothing was mentioned on patients' suitability for PAC treatment. Therefore, the clinical evidence for ramucirumab monotherapy was not necessarily based on the indicated population (patients not suitable for PAC).

Furthermore, in the submission, the company assumed that if a patient is not suitable for PAC, s/he is not suitable for any other cytotoxic chemotherapy either. If this is plausible, the comparison of RAM versus BSC is in line with the NICE final scope. Otherwise, comparisons with cytotoxic chemotherapy other than PAC (i.e. docetaxel, irinotecan and FOLFIRI) should have been included.

An indirect comparison with docetaxel, using the COUGAR-02 [13] trial (Docetaxel versus active symptom control for refractory oesophagogastric adenocarcinoma), showed that the hazard ratio of OS of monotherapy ramucirumab versus docetaxel was not significantly different from 1 (HR 1.16; 95% CI 0.77–1.73).

3.2.2 Ramucirumab Combination Therapy

RAINBOW was deemed to be a good-quality randomized controlled trial. Uncertainty about long-term efficacy was considered to be small, because both OS and PFS data were mature. There were few ($n = 15$) UK patients in the trial, in the region 1 stratum (includes patients from Europe, Israel, USA and Australia).

The ERG concluded that the NMA results should be interpreted with caution, due to considerable heterogeneity resulting from the inclusion of predominantly Asian studies (e.g. the Hironaka et al. [15] study was based only on Japanese patients). There are substantial differences between Asian and Western countries in terms of gastric cancer incidence, histology, screening and treatment approaches [17–19].

Furthermore, the study by Thuss-Patience et al. [16], which was included in the NMA, was underpowered ($N = 40$) since it was closed prematurely because of poor recruitment. The ERG also stated that the NMA would have been more reliable if the results from Roy et al. [14], which also included an irinotecan arm, were included in the base case.

3.3 Summary of the Cost-Effectiveness Evidence

The company submitted two separate partitioned survival models to assess the cost effectiveness of RAM and RAM + PAC. The structures of both models were the same, comprising three states: pre-progression, post-progression and death. Patients entered the model in the pre-progression state. The cycle length was 1 week, and half-cycle correction was applied. Lifetime horizon was used in both models. Both models adopted a National Health Service (NHS) perspective and costs and benefits were discounted at a rate of 3.5%.

The only comparator in the monotherapy economic model was BSC since the company claimed that people who are ineligible for PAC are not eligible for any cytotoxic chemotherapy.

The comparators in the combination therapy model were BSC and docetaxel. According to the company, a comparator was eligible for inclusion if it was used sufficiently (i.e. above an arbitrary threshold of 10% use). PAC was only included as a means of validating the model results by comparing model outcomes with the clinical evidence from RAINBOW.

Transition probabilities between the health states for either RAM or RAM + PAC were determined from parametric survival functions fitted to the OS and PFS data from the RAINBOW and REGARD trials. Transition probabilities for the comparators docetaxel and BSC were estimated using HRs from the NMA.

In the monotherapy model, the company used the gamma distribution to model OS and the interval-censored log-normal distribution to model PFS.

For the combination therapy model, the company used the OS Kaplan–Meier curve from RAINBOW until the end of the follow-up period and afterwards an exponential extrapolation was assumed. An interval censored Weibull distribution was chosen for PFS.

Utility values for the pre-progression and the post-progression health states were derived from EQ-5D data from RAINBOW in both models. For the pre-progression state, the average utility of the patients at baseline was used, whereas for the post-progression state, the average utility of the patients who discontinued treatment due to progressive disease was utilized. According to the company, data from the monotherapy trial (REGARD) could not be used since only EORTC-QLQ-C-30 data were collected and post-baseline data was insufficient due to rapid disease progression in both arms.

Utility decrements were applied for AEs in both models. AEs were included based on grade (i.e. 3 and 4) and occurrence (i.e. >5% in any of the relevant trials). Utility decrements were taken from the literature in other cancer areas [20–23]. Duration of AEs was taken from the NICE STA for pixantrone in non-Hodgkin's B cell lymphoma [24].

A utility increment was applied in the combination therapy model to the responders. For RAM + PAC, the ORR was taken from RAINBOW (27.9%). The company assumed the ORR for docetaxel was similar to PAC in RAINBOW (16.1%). Since ORR for BSC and RAM in REGARD was very low (2.6 and 3.4%), no response was assumed for BSC and RAM in the monotherapy model.

The costs included in the model were for drug acquisition, drug administration, monitoring/testing and follow-up care. All costs used in the model calculations were based on their 2014 values. The prices of generic chemotherapies were taken from the electronic market information tool (eMIT) [25]. Non-generic drug prices were taken from the British National Formulary (BNF) [26]. The drug dosages, treatment durations and relative dose intensity data for RAM, RAM + PAC and docetaxel were derived from REGARD, RAINBOW and COUGAR-02 [27]. In the base case, drug wastage from open vials was assumed and pre-medication for RAM + PAC and DOC treatments were based on their summary of product characteristics. The cost components of BSC were identified from a review of hospital medical records [6]. Drug administration costs were based on NHS reference costs.

Costs for tests and monitoring were based on expert opinion. Costs of grade 3 or 4 AEs were based on NHS reference costs. Hospitalization costs were based on trial data. It was assumed that only 12% of patients receive a third-line therapy and relevant costs (drug acquisition, administration and follow-up care) were applied in the first cycle after progression. Inflation-adjusted terminal care costs from Coyle et al. [23] were applied in the base case.

The base-case deterministic ICER for RAM versus BSC was £188,640 per QALY gained. Deterministic sensitivity analysis showed that the ICER was most sensitive to hospital admission rates, length of hospital stay, assumptions

on vial waste and extrapolation of post-progression survival. Assuming a log-normal distribution (the distribution with a better fit using the goodness-of-fit diagnostic tests) instead of gamma distribution for OS reduced the ICER to £174,485 per QALY gained.

The base-case deterministic ICER for RAM + PAC versus BSC was £118,209 per QALY gained. The deterministic sensitivity analysis showed that the ICER was most sensitive to the assumptions surrounding source of drug prices (eMIT vs BNF), length of hospital stay, relative dose intensity and body surface area/weight. Assuming a Weibull distribution for OS gave similar results to the base-case analysis (£117,236 per QALY gained), whereas the log-logistic distribution reduced the ICER to £96,103 per QALY gained. In a scenario analysis, the OS, PFS and time on treatment were adjusted based on region 1 specific data, which led to similar ICER results.

3.4 Critique of the Cost-Effectiveness Evidence

The ERG considered excluding comparators due to an arbitrary threshold (10%) not appropriate and found that treatments listed in the scope for the combination therapy model (PAC, irinotecan and FOLFIRI) should have been included. Furthermore, even under the 10% threshold rule, the ERG found that PAC should have been included since 10.5% of second-line GC/GOJ patients received PAC [6].

In general, the process for extrapolating survival curves was clear, but the choice of the survival regression model was not always consistent. It was also not clear which approach was followed for interval-censoring adjustments. In the model, mortality in the pre-progression state was neglected while calculating the number of 'newly progressed patients'. However, its impact on incremental results was expected to be low because further-line treatment costs, derived from the number of 'newly progressed patients', constitute a minor part of the total costs.

There were also some issues concerning the generalizability of results from RAINBOW and REGARD for UK patients. For instance, the ERG believed that region 1 data better reflects the UK population weight and body surface area for calculating the drug costs for RAM + PAC, as the whole RAINBOW population included a lot of patients from Asian countries, who had relatively lower weight and body surface areas.

Also, double counting of hospitalization costs was identified because, in addition to the modelled costs based on observed hospitalization rates, Health Resource Groups' codes referring to AEs in the model also included hospitalizations. In its response to the clarification letter, the company provided a scenario that reduced the rate of hospitalizations by an estimate of the proportion of hospitalizations due to AEs. The ERG used these adjusted

hospitalization rates in its exploratory analyses for its base case. In addition, the ERG found a few minor programming errors in the original company model; correcting those had negligible impact on the ICER.

3.5 Additional Exploratory Analyses Conducted by the ERG

3.5.1 Additional Comparators

In exploratory analyses, the ERG included the comparators defined in the final scope for the combination therapy model. These analyses were presented using the company’s base-case assumptions (except confirmed programming errors). The results of these exploratory analyses are presented in Table 5. They should be interpreted with caution because they relied on the NMA, which was associated with significant uncertainty as a result of heterogeneity between the studies.

3.5.2 ERG Base-Case and Scenario Analyses

The ERG base case included the adjustments listed in Table 6, and the results are given in Table 7. In addition, the ERG explored three different scenarios in the combination therapy model, as listed in Table 8.

3.5.3 End-of-Life Considerations

NICE end-of-life (EOL) supplementary advice at the time applied in the following circumstances and when all criteria referred to in Table 9 are satisfied.

For RAM + PAC, the company claimed that EOL criteria should be applied based on the mean additional survival in comparison with BSC (6.03 months) and docetaxel (4.13 months). However, the additional survival of RAM + PAC versus other comparators in the scope was 1.44 months for PAC, 2.27 months for irinotecan and 1.1 months for FOLFIRI. Therefore, the ERG argued the EOL criteria were not fulfilled.

Table 5 Pairwise base-case results for additional comparators compared with ramucirumab plus paclitaxel using the company’s base-case assumptions^a

Intervention	Comparator	Incremental QALY	Incremental cost (£)	ICER (per QALY gained) (£)
RAM + PAC	BSC	0.33	39,584	118,174
	Docetaxel	0.24	34,153	145,302
	Irinotecan	0.15	31,238	213,015
	PAC	0.1	26,790	273,657
	FOLFIRI	0.1	28,166	294,362

BSC Best supportive care, *FOLFIRI* fluorouracil-based therapy, *ICER* incremental cost-effectiveness ratio, *PAC* paclitaxel monotherapy, *QALY* quality-adjusted life-year, *RAM + PAC* ramucirumab combination therapy

^a Company’s base-case assumptions were used except for a corrected programming error

Table 6 List of adjustments on company submission base case

No.	Change on the company base case
1	Removal of programming errors (half cycle correction and docetaxel price calculations)
2	Weight/body surface area and treatment-specific hospital length of stay of patients based on region 1, not all patients in the trials
3	Correction of double counting for hospitalizations due to adverse events

Table 7 The results of the ERG and CS base-case ICERs (Cost per QALY gained)

	ERG base-case ICER (per QALY gained) (£)	CS base-case ICER (per QALY gained) (£)
Monotherapy model		
RAM vs BSC	188,055	188,640
Combination therapy model		
RAM + PAC vs BSC	129,431	118,209
RAM + PAC vs docetaxel	168,164	148,769
RAM + PAC vs PAC	359,794	NA

BSC Best supportive care, *CS* Company submission, *ERG* evidence review group, *FOLFIRI* fluorouracil-based therapy, *ICER* incremental cost-effectiveness ratio, *NA* not available, *PAC* paclitaxel monotherapy, *QALY* quality-adjusted life-year, *RAM* ramucirumab monotherapy, *RAM + PAC* ramucirumab combination therapy

3.6 Conclusions of the ERG Report

In REGARD, RAM was associated with a slightly higher OS and PFS compared with BSC, and RAINBOW showed more favourable OS and PFS results for RAM + PAC in comparison with PAC. The NMA suggested some gains in OS and PFS for RAM + PAC compared with BSC, and gains in PFS for RAM + PAC in comparison with docetaxel. The ERG considered that the results based on indirect comparisons should be interpreted with caution due to significant heterogeneity between studies. For instance,

Table 8 Description and impact of the ERG scenarios

Description of the ERG scenario	Impact of the scenario
Including the results from Roy et al. [14] to the OS NMA	Increase in ICER around £14,000 per QALY gained for RAM + PAC vs BSC compared with the ERG base case
Using direct evidence from the RAINBOW trial and not NMA for the efficacy data	ICER for RAM + PAC compared with PAC increased from the ERG base case of £359,794 per QALY gained to £392,108 per QALY gained
Implementation of time-varying utility values that are directly derived from the RAINBOW trial in the model, instead of a single, uniform utility value for the pre-progression state in all cycles	This resulted in an ICER of £408,223 per QALY gained for RAM + PAC vs PAC

BSC Best supportive care, *ERG* evidence review group, *ICER* incremental cost-effectiveness ratio, *NMA* network meta-analysis, *OS* overall survival, *PAC* paclitaxel monotherapy, *QALY* quality-adjusted life-year, *RAM* ramucirumab monotherapy, *RAM + PAC* ramucirumab combination therapy

Table 9 The end-of-life criteria for the National Institute for Health and Care Excellence (NICE)

Short life expectancy, normally <24 months on current National Health Service (NHS) treatment
The intervention offers an extension to life, normally of at least an additional 3 months, compared with current NHS treatment
The treatment is licensed or otherwise indicated for small patient populations (this criterion has subsequently been removed) [28]

some studies were predominantly Asian, and the histology and treatment pattern of gastric cancer between Western patients and Asian patients might be very different [17–19].

By correcting issues in the model and changing a few input parameters, a new ERG base case was defined for both monotherapy and combination therapy models. The ERG conducted some exploratory analyses in which additional comparators from the NICE scope were included. Furthermore, additional scenario analyses were conducted in the combination therapy model. In all analyses, the ICER of RAM + PAC or RAM compared with any one of the comparators was never below £90,000 per QALY gained. Similarly, the probability of RAM + PAC or RAM becoming the most cost-effective therapy was negligible for thresholds below £100,000 per QALY gained in all analyses.

In order to improve the robustness of the health economic results for the UK, a direct comparison of RAM and RAM + PAC with all of the relevant comparators among a predominantly western patient population would be necessary. In addition, QoL data for RAM monotherapy and for comparators among the targeted population (i.e. patients ineligible for PAC combination therapy) would reduce the uncertainty around utilities.

Nevertheless, regardless of any problems with the evidence and the model, the ICERs from even the company base case far exceeded the usual threshold of £20,000 to

£30,000 per QALY gained and even the EOL threshold of £50,000 per QALY gained.

3.6.1 Key Methodological Issues

For many of the key cost-related inputs (i.e. body weight, hospitalization rate, length of stay, etc.), data from all patients in the REGARD and RAINBOW trials were used, but average estimates from these trials may not be generalizable to the UK population as the body weight/surface as well as the gastric cancer histology and treatment patterns in the Asian population might differ from those in the Western population [17–19].

If direct comparisons based on phase III randomized controlled trials for all relevant comparators from the scope are not available, STAs often rely on indirect treatment comparisons. As was the case in this STA, sometimes these analyses involve assumptions that are not evidence based (e.g. assuming PFS HR = 1 for RAM vs docetaxel). Also, this STA revealed that decisions regarding exclusion of trials, (e.g. Roy et al. [14]) might influence outcomes substantially. Finally, results of indirect treatment comparisons should be interpreted with extreme caution if these are obtained by pooling data from heterogeneous trials.

4 National Institute for Health and Care Excellence Guidance

4.1 Preliminary Guidance

The Committee considered the company's decision problem, and noted that it was in line with the NICE scope, with the exception of the choice of comparators.

While BSC would be the only comparator for RAM, the Committee concluded that for people for whom

RAM + PAC is appropriate, PAC and docetaxel are both relevant comparators.

The Committee concluded that the REGARD and RAINBOW trials formed suitable evidence on which it could base its decision on the clinical efficacy of RAM and RAM + PAC.

For combination therapy, the Committee considered that there was no reason to use the NMA results, rather than using relevant head-to-head data from a good-quality, international, randomized controlled trial (RAINBOW) with mature OS and PFS data.

For cost effectiveness, the Committee agreed with the error corrections and adjustments carried out by the ERG to use region 1 data for body surface area and body weight, to correct for double counting of hospitalizations, and to adjust length of hospitalization stay for region 1. It concluded that the model submitted by the company was robust and suitable for the purposes of its decision making and that the ERG's suggested amendments to the model were appropriate.

The Committee considered the most robust analysis was the ERG's exploratory analysis, which used RAINBOW trial data for RAM + PAC compared with PAC, and used time-varying utility values from RAINBOW, collected during the pre-progression period. This analysis provided the most plausible ICER of RAM + PAC versus PAC for people with GC/GOJ for whom treatment in combination with cytotoxic chemotherapy is appropriate: £408,200 per QALY gained (with incremental costs of £35,100 and incremental QALYs of 0.09).

The Committee concluded that the most plausible ICER for people with GC/GOJ adenocarcinoma for whom further cytotoxic chemotherapy is not appropriate (i.e. RAM vs BSC) was £188,100 per QALY gained (representing incremental costs of £22,500 and incremental QALYs of 0.12).

According to the Committee, EOL criteria were not met, and the overall conclusion was that, for the treatment of adults with advanced GC/GOJ adenocarcinoma with disease progression after platinum and fluoropyrimidine chemotherapy, neither RAM + PAC nor RAM (where paclitaxel is not appropriate) were a cost-effective use of NHS resources at the usual range of ICERs (£20,000–£30,000 per QALY gained).

4.2 Response to the Preliminary Guidance

The response of the company to the ACD first focused on the unwarranted variation and inequity in the provision of second-line GC/GOJ treatments in clinical practice, which might be aggravated by the lack of a licensed second-line treatment. Secondly, the company expressed that the weaknesses of the NMA were overestimated (arguing that

the Thuss-Patience et al. study [16] should be considered as systematically unbiased, and that the Hironaka et al. study [15] results are transferable to the UK), and that the NMA results should be considered sufficiently plausible to permit its use. Thirdly, concerning the comparators, the company did not agree with the Committee's decision that the comparison of PAC with RAM + PAC based on RAINBOW data provided a good basis for assessing the EOL criteria and the most plausible ICER estimates, because the company considered BSC and docetaxel were more relevant and commonly used comparators. Finally, the company discussed a number of potential factual inaccuracies and inconsistencies, which were clarified before the final guidance.

4.3 Final Guidance

The Committee considered the comments raised by the company in its response to the ACD. However, this did not lead to any change in the ACD and the Committee concluded that RAM or RAM + PAC could not be recommended within the market authorization.

5 Conclusions

This appraisal demonstrated that the selection of the relevant comparators is crucial and might have a substantial impact on the ICER. Even though the comparators were identified in the scope, the company and the decision maker had different views on the relevance of some of the comparators for the inclusion in the cost-effectiveness analysis.

Similarly, the inclusion of available evidence for the NMA is another crucial decision that may have a considerable impact on ICER, and different parties may have differing views on this decision as well.

Another important issue is the generalizability of the evidence to the UK patient population. This STA showed that when differences are expected between the UK population and the population on which the evidence is built (e.g. weight/body surface area, histology and treatment patterns of the Asian patients), it might be important to adapt these analyses for the UK.

Finally, this appraisal shows that decision makers may prefer to base their decision on direct evidence rather than indirect evidence, especially if the latter has the potential to be heterogeneous and biased.

Despite these differing opinions between the company and decision maker, in all of the analyses from the company and the ERG, ICERs of monotherapy and combination therapy were always far above the currently accepted thresholds. This review described a STA process without a patient access

scheme (PAS), with unacceptably high ICERs, even in the best-case scenarios of the company. In the case of no PAS, if the best-case scenario from the company is not within the reach of an acceptable cost-effectiveness ratio, one may wonder to what extent the current STA process, in its full scale, is a cost-effective way of spending NHS resources.

Author Contributions Nasuh C. Büyükkaramikli acted as a health economist on this assessment, critiqued the manufacturer's economic evaluation, wrote the manuscript and acted as overall guarantor for the overall content of this article. Hedwig M. Blommestein and Nigel Armstrong acted as health economists on this assessment, critiqued the manufacturer's economic evaluation and contributed to the writing of the article. Rob Riemsma acted as project lead and systematic reviewer on this assessment, critiqued the clinical effectiveness methods and evidence. Maiwenn J. Al acted as health economic project lead and critiqued the manufacturer's economic evaluation. Fiona Clay acted as systematic reviewer and critiqued the clinical effectiveness methods and evidence. Gill Worthy acted as statistician and critiqued the analyses in the manufacturer's submission. Janine Ross critiqued the search methods in the submission. Hans Severens and Jos Kleijnen critiqued the manufacturer's definition of the decision problem and their description of the underlying health problem and current service provision. All authors contributed to the conception and planning of the work, and critically revised and approved the final submitted version of the manuscript. This summary has not been externally reviewed by PharmacoEconomics.

Compliance with Ethical Standards

Funding This project was funded by the National Institute for Health Research (NIHR) Health Technology Assessment (HTA) Program (project number 14/177/03 STA) [see the HTA program website for further information—<http://www.hta.ac.uk>]. This summary of the ERG report was compiled after the AC's review. The views and opinions expressed are the authors' and do not necessarily reflect those of the HTA Programme, NICE, NIHR, NHS, or the Department of Health.

Conflicts of interest Nasuh C. Büyükkaramikli, Hedwig M. Blommestein, Rob Riemsma, Nigel Armstrong, Fiona J. Clay, Janine Ross, Gill Worthy, Johan Severens, Jos Kleijnen and Maiwenn J. Al have no conflicts of interest that are relevant to the content of this review.

Open Access This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. NICE. Guide to the single technology appraisal process, London, 2015. [Online]. <https://www.nice.org.uk/process/pmg24/resources/single-technology-appraisal-user-guide-for-company-evidence-submission-template-pdf-72286715419333>. Accessed 08 Aug 2016.
2. NICE. Gastric cancer (metastatic)—ramucirumab (after chemotherapy) [ID741], NICE technology appraisal guidance [TA378], London, 2016. [Online]. <https://www.nice.org.uk/guidance/indevelopment/gid-tag500>. Accessed 08 Aug 2016.
3. ONS. Cancer Registration Statistics, London, 2012. [Online]. <https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/cancerregistrationstatisticsengland/2014-06-19#results>. Accessed 06 June 2017.
4. Cancer Research UK. Stomach Cancer Incidence Statistics, Cancer Research UK, 2012. [Online]. <http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/stomach-cancer/incidence>. Accessed 08 Aug 2016.
5. Cancer Research UK. Stomach Cancer Mortality Statistics. Cancer Research UK, 2012. [Online]. <http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/stomach-cancer/mortality>. Accessed 08 Aug 2016.
6. Liepa AM et al. Real-world treatment patterns of previously treated advanced gastric and gastroesophageal junction adenocarcinoma (GC) in the United Kingdom (UK). In: Gastrointestinal Cancers Symposium, 2015.
7. NICE. Capecitabine for the Treatment of Advanced Gastric Cancer (TA 191), London, 2010. [Online]. <https://www.nice.org.uk/Guidance/ta191>. Accessed 08 Aug 2016.
8. NICE. Trastuzumab for the treatment of HER2-positive metastatic gastric cancer, London, 2010. [Online]. <https://www.nice.org.uk/guidance/TA208>. Accessed 08 Aug 2016.
9. Fuchs CS, 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.
10. Aaronson NK, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst*. 1993;85(5):365–76.
11. Wilke H, 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.
12. Sym SJ, 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.
13. Ford HER, 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.
14. Roy AC, 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.
15. Hironaka S, 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.
16. Thuss-Patience PC, 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.

17. Kim J, et al. Race and ethnicity correlate with survival in patients with gastric adenocarcinoma. *Ann Oncol Off J Eur Soc Med Oncol*. 2010;21(1):152–60.
18. 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.
19. Ohtsu A, Yoshida S, Saijo N. Disparities in gastric cancer chemotherapy between the East and West. *J Clin Oncol*. 2006;24(14):2188–96.
20. 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.
21. Lloyd A, Nafees B, Narewska J, Dewilde S, Watkins J. Health state utilities for metastatic breast cancer. *Br J Cancer*. 2006;95(February):683–90.
22. 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.
23. Coyle D, 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.
24. NICE. Technology appraisal guidance [TA306], Pixantrone monotherapy for treating multiply relapsed or refractory aggressive non-Hodgkin's B-cell lymphoma, London, 2014. [Online]. <https://www.nice.org.uk/guidance/ta306>. Accessed 08 Aug 2016.
25. Commercial Medicines Unit. Drugs and pharmaceutical electronic Market Information (eMIT), 2014. [Online]. <https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit>. Accessed 23 Apr 2015.
26. Medicines Complete. British National Formulary [Internet]. 2015. <https://www.medicinescomplete.com/mc/bnf/current/>. Accessed accessed 24 Apr 2015.
27. Kang JH, 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.
28. National Institute for Health and Care Excellence. PMG19 Addendum A—Final amendments to the NICE technology appraisal processes and methods guides to support the proposed new Cancer Drugs Fund arrangements. Technology Appraisal Processes—CDF, London, 2016. [Online]. <https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisals/process-and-methods-guide-addendum.pdf>. Accessed 06 June 2016.