Pembrolizumab with platinum-based chemotherapy for untreated advanced oesophageal cancer [ID3741]

A Single Technology Appraisal

Produced by

Peninsula Technology Assessment Group (PenTAG)
University of Exeter Medical School
South Cloisters
St Luke's Campus
Heavitree Road
Exeter
EX1 2LU

Authors

Dr Maxwell S. Barnish, Research Fellow[1]
Ms Rachael Slater, Associate[1] and Analyst[2]
Mr Justin Matthews, Research Fellow[1]
Dr Linda Long, Research Fellow[1]
Mr Ash Bullement, Associate[1] and Analyst[2]
Ms Naomi Shaw, Information Specialist[1]
Dr Nicole Dorey, Consultant Clinical Oncologist[3]
Dr Stephen Falk, Consultant Clinical Oncologist[4]
Ms Louise Crathorne, Senior Research Fellow[1]
Prof G.J. Melendez-Torres, Professor[1]

[1] Peninsula Technology Assessment Group (PenTAG), University of Exeter Medical School, Exeter
[2] Delta Hat Limited, Nottingham
[3] Royal Devon and Exeter NHS Foundation Trust, Exeter
[4] Bristol Haematology and Oncology Centre, Bristol

Correspondence to

Dr Maxwell Barnish
3.09 South Cloisters, St Luke's Campus, Heavitree Road, Exeter, EX1 2LU; m.s.barnish@exeter.ac.uk

Produced by	Peninsula Technology Assessment Group (PenTAG) University of Exeter Medical School South Cloisters St Luke's Campus Heavitree Road Exeter EX1 2LU
Date completed	20/04/2021
Source of funding	This report was commissioned by the NIHR Systematic Reviews Programme as project number 13/38/37.
Declared competing interests of the authors	None declared.
Acknowledgments	The authors acknowledge the administrative support provided by Mrs Sue Whiffin and Ms Jenny Lowe (both PenTAG).
Rider on responsibility for document	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	Barnish MS, Slater R, Matthews J, Long L, Bullement A, Shaw N, Dorey N, Falk S, Crathorne L, Melendez-Torres GJ. Pembrolizumab with platinum-based chemotherapy for untreated advanced oesophageal cancer [ID3741]: A Single Technology Appraisal. Peninsula Technology Assessment Group (PenTAG), 2021.

Author Contributions:

Maxwell S. Barnish	Project managed the ERG team and led the critique of the clinical effectiveness evidence. Contributed towards writing the report.
Rachael Slater	Acted as health economic project lead, critiqued the company's economic evaluation and contributed towards writing the report. Implemented the ERG's preferred settings and exploratory analyses within the company's economic model.
Justin Matthews	Critical appraisal of suitability of indirect treatment comparisons. Contributed towards writing the report.
Linda Long	Critical appraisal of clinical and cost effectiveness evidence. Contributed towards writing the report.
Ash Bullement	Critiqued the company's economic evaluation and contributed towards writing the report.
Naomi Shaw	Critical appraisal of the literature search strategies. Editorial input.
Nicole Dorey	Clinical advisor.
Stephen Falk	Clinical advisor.

Author Contributions:	
Louise Crathorne	Contributed to editing the final report. Editorial input.
G.J. Melendez-Torres	Critical appraisal of the company submission, writing and editorial input, and supervised the final report. Guarantor of the report.

Table of Contents

List of key issues

List of tables

List of figures

1. EXECUTIVE SUMMARY

This summary provides a brief overview of the key issues identified by the evidence review group (ERG) as being potentially important for decision making. It also includes the ERG's preferred assumptions and the resulting incremental cost-effectiveness ratios (ICERs).

Section 1.1 provides an overview of the key issues. Section 1.2 provides an overview of key model outcomes and the modelling assumptions that have the greatest effect on the ICER. Sections 1.3 to 1.6 explain the key issues in more detail. Background information on the condition, technology and evidence and information on non-key issues is provided in the main ERG report.

All issues identified represent the ERG's view, not the opinion of NICE.

1.1. Overview of the key issues in the clinical effectiveness evidence

Table 1: Key issues in the clinical effectiveness evidence

ID3741	Summary of issues	Report sections
Key Issue 1	The clinical evidence may not be generalizable to the UK population	3.2.2.2, 3.2.3.1
Key Issue 2	Clinical effectiveness evidence excluded probative estimates of effectiveness between standard of care regimens	3.3, 3.4, 6.2.4
Key Issue 3	The estimated overall survival projections have a large impact on cost-effectiveness	4.2.6.1, 6.2.1
Key Issue 4	The use of time-to-death utilities may overstate the QALYs accrued by patients	4.2.7.4, 6.2.3
Key Issue 5	The doublet used in the economic model does not reflect clinical practice in the UK	4.2.4, 6.2.5

Abbreviations: QALYs, quality-adjusted life years

The key differences between the company's preferred assumptions and the ERG's preferred assumptions are:

- Method of utility estimation, with the ERG preferring progression-based rather than time-to-death estimation;
- Distributions of subsequent treatments;
- Choice of progression-free survival extrapolation; and
- Implementation of a treatment waning effect.

1.2. Overview of key model outcomes

NICE technology appraisals compare how much a new technology improves length (overall survival) and quality of life in a quality-adjusted life year (QALY). An ICER is the ratio of the extra cost for every QALY gained.

Overall, the technology is modelled to affect QALYs by:

- Increasing overall survival
- Delaying disease progression

Overall, the technology is modelled to affect costs by:

- Drug acquisition costs for pembrolizumab
- Time on treatment for pembrolizumab

The modelling assumptions that have the greatest effect on the ICER are:

- Assumptions surrounding Overall Survival and the choice of utility method

1.3. The decision problem: summary of the ERG's key issues

The ERG did not identify any key issues with the company's interpretation of the decision problem.

1.4. The clinical effectiveness evidence: summary of the ERG's key issues

Key Issue 1: The clinical evidence may not be generalizable to the UK population

Report sections	3.2.2.2, 3.2.3.1
Description of issue and why the ERG has identified it as important	The pivotal trial, KEYNOTE-590, included a substantial number of patients from East Asian countries, where treatment guidelines for oesophageal cancer are considerably different from those applicable to the UK, and did not reflect the expected population composition of oesophageal squamous cell carcinoma and adenocarcinoma. This limits the ability to generalise findings from the trial to the UK context.
What alternative approach has the ERG suggested?	The ERG noted that the committee may wish to rely on analyses drawing on the 'rest of world' subgroup for decision-making. These are presented where available.
What is the expected effect on the cost-effectiveness estimates?	The impact on the ICER is unclear.
What additional evidence or analyses might help to resolve this key issue?	Greater clarity relating to consistency of estimates between Asian and rest of world populations, and with respect to the type of cancer, would support decision-making.

Abbreviations: ERG, Evidence Review Group

Key Issue 2: Clinical effectiveness evidence excluded probative estimates of effectiveness between standard of care regimens

Report sections	3.3, 3.4, 6.2.4
Description of issue and why the ERG has identified it as important	The company's search was carried out without the term 'gastric', and studies were excluded when subgroup results for oesophageal or oesophagogastric junction Siewert type I cancer patients could not be identified. As a consequence, certain evidence was not included such as doublet vs triplet effect estimates from network meta-analyses (NMAs) and the influential REAL-2 study. This led to the company's conclusion that no evidence could be assembled to compare doublet and triplet regimens. There was thus no comparison between pembrolizumab and triplet regimens. Clinical advice received is that systemic treatment is similar for oesophageal and gastric cancers. Including this wider evidence provides estimates

Report sections	3.3, 3.4, 6.2.4
	of doublet vs triplet efficacy in existing UK practice, from existing NMAs or meta-analyses.
What alternative approach has the ERG suggested?	The ERG has identified plausible estimates comparing doublet vs triplet regimens to inform cost-effectiveness analysis.
What is the expected effect on the cost-effectiveness estimates?	The ERG presented scenario analyses integrating this evidence to compare pembrolizumab plus doublet regimens against triplet regimens.
What additional evidence or analyses might help to resolve this key issue?	A precise conclusion could not be reached due to the need to apply effect estimates against summary Kaplan-Meier curves in the economic model. A more direct method of including triplet regimens in cost effectiveness modelling would resolve this uncertainty.

Abbreviations: ERG, Evidence Review Group

1.5. The cost effectiveness evidence: summary of the ERG's key issues

Key Issue 3: The estimated overall survival projections have a large impact on cost-effectiveness

Report sections	4.2.6.1, 6.2.1
Description of issue and why the ERG has identified it as important	The company's overall survival (OS) projections may overestimate the proportion of patients alive in the long term for the pembrolizumab in combination with chemotherapy arm. OS is a key driver of cost-effectiveness results, and projections are currently based on incomplete data from the KEYNOTE-590 study. Clinical advice provided to the ERG suggested that a range of alternative extrapolations appear to be clinically plausible, but each option has a notable impact on the ICER.
What alternative approach has the ERG suggested?	The ERG preferred the company's base case assumptions with a treatment waning effect scenario applied between the year 5-7. This approach made use of the company's base-case approach up until 5 years, after which extrapolations were adjusted such that by 7 years, the curves for both arms project identical hazards of death for the remainder of the modelled time horizon.
What is the expected effect on the cost-effectiveness estimates?	The expected effect on the cost-effectiveness estimates is to increase the ICER.
What additional evidence or analyses might help to resolve this key issue?	More mature KEYNOTE-590 OS data would help resolve the uncertainty inherent within the OS

Report sections	4.2.6.1, 6.2.1
	extrapolations. Clinical expert opinion may also support the selection of appropriate extrapolations, but as highlighted previously, clinical advice provided to the ERG suggested a range of extrapolations appeared plausible, producing a broad range of ICERs.

Abbreviations: ERG, Evidence Review Group

Key Issue 4: The use of time-to-death utilities may overstate the QALYs accrued by patients

Report sections	4.2.7.4, 6.2.3
Description of issue and why the ERG has identified it as important	The company's base case assigned utility values based on time to death, instead of based on progression. This led to markedly different estimates of average utility in each health state, with time-to-death utility generating a mean utility in the pre-progression health state above that of the general population.
What alternative approach has the ERG suggested?	The ERG preferred to use progression-based utility values, which may more appropriate capture expected QALY gains from pembrolizumab.
What is the expected effect on the cost-effectiveness estimates?	The expected effect on the cost-effectiveness estimates is to increase the ICER.
What additional evidence or analyses might help to resolve this key issue?	Additional justification for choice of time-to-death utilities, and evidence as to why the company's two approaches differ so widely in terms of average utilities in each health state, may support decision-making.

Abbreviations: ERG, Evidence Review Group

Key Issue 5: The doublet used in the economic model does not reflect clinical practice in the UK

Report sections	4.2.4, 6.2.5
Description of issue and why the ERG has identified it as important	The pivotal trial, KEYNOTE-590, used a doublet regimen as standard of care, cisplatin with 5-fluorouracil, that is only one of several doublet regimens available. Clinical advice to the ERG was that while doublet regimens are of exchangeable effectiveness (i.e. exhibit a class effect), doublet regimens with 5-fluorouracil are rarely used given the need for lengthy infusion time and only used when patients cannot swallow capecitabine tablets. The company also provided

Report sections	4.2.4, 6.2.5
	different chemotherapy regimens for the comparator arm, but the chemotherapy regimen in combination with pembrolizumab remained as the 5-fluorouracil plus cisplatin. This means that costs may not reflect what would be expected in clinical practice, and lack generalisability to the UK context.
What alternative approach has the ERG suggested?	Following on from the company's scenario results, the ERG has explored alternative costing assumptions for doublet regimens in combination with pembrolizumab and as the comparator, including a blended comparator with more clinically plausible UK market shares.
What is the expected effect on the cost-effectiveness estimates?	Impacts on the ICER vary by type of doublet used in the comparator and in combination with pembrolizumab in addition to varying the mix of treatments based on different market shares.
What additional evidence or analyses might help to resolve this key issue?	Specific market share evidence in the UK for doublet and triplet regimens may generate a more realistic costing assumption. Confirmation of what type of chemotherapies would be used in combination with pembrolizumab in clinical practice.

Abbreviations: ERG, Evidence Review Group

1.6. Other key issues: summary of the ERG's views

The ERG did not identify any other key issues that would be expected to affect decision-making.

1.7. Summary of ERG's preferred assumptions and resulting ICER

The ERG's preferred assumptions for the cost-effectiveness analysis of pembrolizumab in combination with chemotherapy compared to chemotherapy alone are outlined in Table 2.

Table 2: Summary of ERG's preferred assumptions and ICER

Scenario	Incremental cost	Incremental QALYs	ICER
Company's base case (post clarification questions)	£27,172	0.65	£41,688
ERG corrected company base case (see Section 6.1)	£27,173	0.65	£41,701
Remove half cycle correction	£27,172	0.65	£41,691
Administration costs using a day case setting	£27,402	0.65	£42,044

Scenario	Incremental cost	Incremental QALYs	ICER
Turning off stopping rules for treatments (i.e., just using the ToT KM estimates from KEYNOTE-590)	£27,630	0.65	£42,394
Re-distributing subsequent treatments	£27,439	0.65	£42,100
Progression-based utilities	£27,439	0.57	£48,108
PFS piecewise using 37-week cut-off and log-logistic extrapolation	£28,052	0.59	£47,270
Include treatment waning between 5-7 years	£28,007	0.54	£51,921
ERG's preferred base case (deterministic; see Section 6.3)	**£28,007**	**0.54**	**£51,921**

Abbreviations: ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life year

Modelling errors identified and corrected by the ERG are described in Sections 5.3 and 6.1. For further details of the exploratory and sensitivity analyses done by the ERG, see Section 6.

2. INTRODUCTION AND BACKGROUND

2.1. Introduction

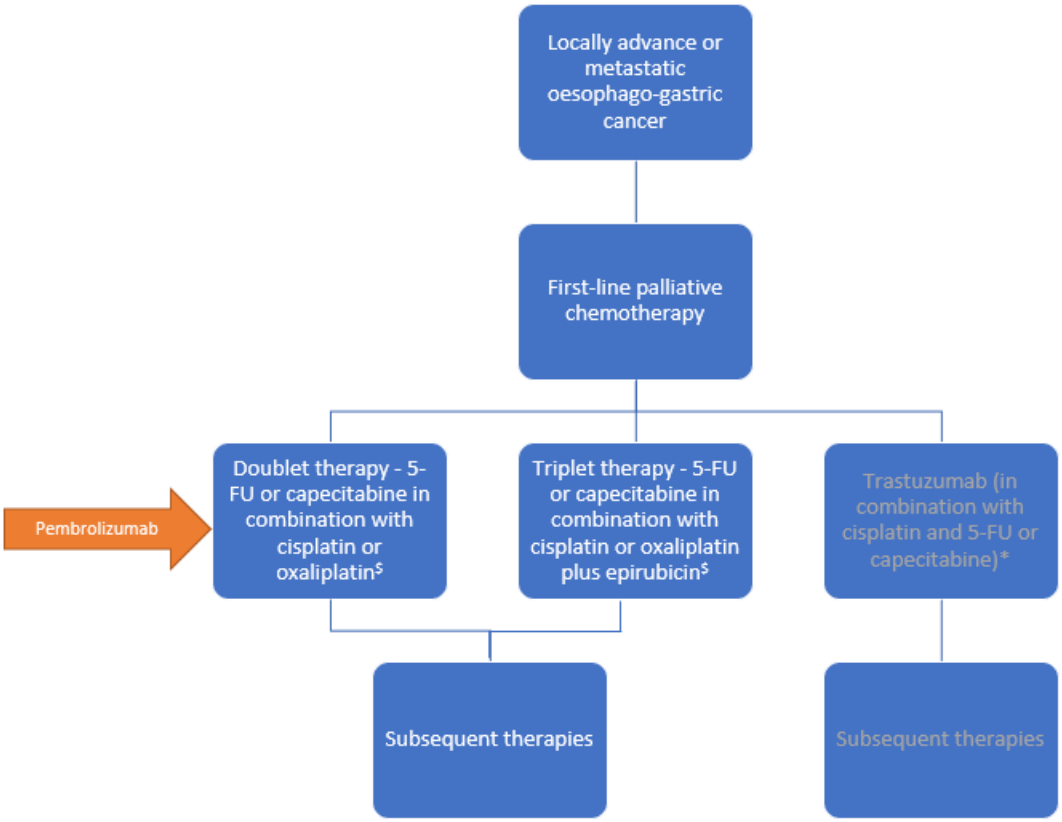
In this report, the ERG provides a review of the evidence submitted by Merck Sharp & Dohme in support of pembrolizumab with platinum-based chemotherapy, for the treatment of adults with untreated advanced oesophageal cancer.

2.2. Background and underlying health problem

Oesophageal cancer is believed to be the eighth most prevalent form of cancer worldwide.[1] The UK has the highest age-standardised incidence of oesophageal cancer in Europe.[2] Four in every five oesophageal cancers occur in adults aged 60 or over[3] with a greater prevalence in males than females.[4] Histologically, oesophageal cancer is subdivided into squamous and adenocarcinoma, the latter representing approximately two-thirds of UK cases and the former one-third.[5] Obesity, smoking and alcohol consumption have been identified as risk factors for oesophageal cancer. Survival prognosis for patients with oesophageal cancer is poor, with most living between three and 12 months after diagnosis and 4% living at least five years.[6] The Evidence Review Group (ERG) considered that the Company Submission (CS) offered an acceptable description of the condition; its pathophysiology, natural course and epidemiology; and the current treatment options available.

NICE Guideline 83[7] was identified in the CS as relevant to this appraisal. The company depicted the treatment pathway in this Guideline, and the proposed positioning of pembrolizumab, in a flowchart (Figure 1). Clinical advisors to the ERG indicated that this flowchart was an accurate depiction of current NHS clinical practice in England and Wales.

Figure 1 NICE pathway on locally advanced or metastatic oesophago-gastric cancer



Source: CS, p.16 – based on NICE Guideline 83. Trastuzumab combinations are used for HER2+ patients.

Pembrolizumab is a monoclonal antibody of the IgG4/Kappa isotope designed to exert dual ligand blockade of the programmed cell death protein 1 (PD-1) pathway by directly blocking the interaction between PD-1 and its ligands, programmed death-ligand 1 (PD-L1) and programmed death-ligand 2 (PD-L2), which appear on antigen-presenting or tumour cells. Pembrolizumab is used for a range of other cancer indications in current practice. The ERG considered that the company's intended positioning, as compared to current standard of care, was appropriate and generally well-described.

The company's intended positioning for pembrolizumab occupies the position in the treatment pathway for locally advanced or metastatic oesophageal-gastric cancer currently occupied by palliative care options. These typically take the form of doublet or triplet chemotherapy regimens. Clinical advice to the ERG indicated that while doublet and triplet regimens were the

appropriate comparators, and that class effects could generally be assumed, certain regimens mentioned in the CS would not be currently funded for NHS use:

- Cetuximab + cisplatin + fluorouracil (5-FU)
- Panitumumab + cisplatin + 5-FU
- Cisplatin + 5-FU + recombinant human lymphotoxin-α derivative (rhLTα-DA)
- Mitomycin + cisplatin + 5-FU.

2.3. Critique of company's definition of decision problem

The ERG considered that the company's definition of the decision problem generally matched the decision problem in the NICE scope.[8]

Table 3: Summary of decision problem

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG comment
Population	Adults with untreated, unresectable locally advanced or metastatic oesophageal cancer or gastroesophageal junction adenocarcinoma	Adults with locally advanced unresectable or metastatic carcinoma of the oesophagus or HER-2 negative gastroesophageal junction adenocarcinoma.	The population described by MSD reflects the anticipated licence indication wording	The ERG considered that the population considered in the company submission was generally well-matched to the NICE scope. However, it was narrower to reflect the anticipated licence indication wording
Intervention	Pembrolizumab with platinum-based chemotherapy	Pembrolizumab in combination with platinum and fluoropyrimidine based chemotherapy	The intervention described by MSD reflects the anticipated licence indication wording	The ERG considered that the intervention considered in the company submission was generally well-matched to the NICE scope. However, it was broader, including fluoropyrimidine based chemotherapy, to cover the full breadth of the anticipated licence indication wording
Comparator(s)	Platinum-based chemotherapy without pembrolizumab, such as: • doublet treatment with fluorouracil or capecitabine plus cisplatin or oxaliplatin	Platinum-based chemotherapy without pembrolizumab, such as: • doublet treatment with fluorouracil or capecitabine plus cisplatin or oxaliplatin • triplet treatment with fluorouracil or capecitabine	N/A	As per the scope for this appraisal

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG comment
	• triplet treatment with fluorouracil or capecitabine plus cisplatin or oxaliplatin epirubicin	plus cisplatin or oxaliplatin epirubicin		
Outcomes	• Overall survival • Progression-free survival • Response rate • Adverse effects of treatment • Health-related quality of life.	• Overall survival • Progression-free survival • Response rate • Adverse effects of treatment • Health-related quality of life.	N/A	As per the scope for this appraisal
Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared.	N/A	The ERG agreed that the economic analysis presented is aligned with the reference case

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG comment
	technologies being compared. Costs will be considered from an NHS and Personal Social Services perspective. The availability of any patient access schemes for the intervention or comparator technologies will be taken into account.	Costs will be considered from an NHS and Personal Social Services perspective. The availability of any patient access schemes for the intervention or comparator technologies will be taken into account.		
Subgroups	None specified	Pre-specified efficacy analyses were conducted in KEYNOTE-090 according to: • histology • geographic region • ECOG performance scale • disease status (locally advanced vs metastatic) • age • sex The economic model includes subgroup analysis in the CPS ≥ 10 sub-population	Subgroup analyses were pre-specified in the KEYNOTE-590 study protocol to determine whether the treatment effect was consistent across subgroups The company considered the CPS≥10 sub-population to be of particular clinical significance.	Although no subgroups were specified in the NICE scope, the ERG considered the pre-specified subgroups in KEYNOTE-590 to be generally appropriate

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG comment
Special considerations including issues related to equity or equality	Not stated	MSD does not envisage any equality issues with the use of pembrolizumab in combination with platinum-based chemotherapy for the first-line treatment of patients with locally advanced unresectable or metastatic carcinoma of the oesophagus or HER-2 negative gastroesophageal junction adenocarcinoma in adults.	Not applicable	No equity issues were identified

Abbreviations: CPS, combined positive score; ERG, Evidence Review Group; HER-2, human epidermal growth factor receptor 2; MSD, Merck Sharp & Dohme; N/A, not applicable; NICE, National Institute for Health and Care Excellence

3. CLINICAL EFFECTIVENESS

The sections below discuss the evidence submitted by the company in support of the clinical effectiveness of pembrolizumab for adults with locally advanced unresectable or metastatic carcinoma of the oesophagus or HER-2 negative gastroesophageal junction adenocarcinoma. The ERG reviewed the details provided on:

- Methods implemented to identify, screen, data extract and assess the risk of bias in relevant evidence
- Clinical efficacy of pembrolizumab
- Safety profile of pembrolizumab
- Assessment of comparative clinical effectiveness of pembrolizumab against relevant comparators

A detailed description of an aspect of the CS is only provided where the ERG disagreed with the company's assessment or proposal, or where the ERG identified a particular area of concern that the ERG considered necessary to highlight for the Committee.

The ERG identified two key issues relating to the clinical effectiveness evidence:

- The clinical evidence may not be generalizable to the UK population
- Clinical effectiveness evidence excluded probative estimates of effectiveness between standard of care regimens

3.1. Critique of the methods of review(s)

The company undertook a systematic review to identify relevant publications on the clinical efficacy and safety of pembrolizumab in combination with platinum and fluoropyrimidine based chemotherapy, as first line treatment in patients with locally advanced unresectable or metastatic carcinoma of the oesophagous (both squamous and adenosquamous) or HER-2 negative gastroesophageal junction adenocarcinoma in adults. The company considered direct and indirect comparisons between the intervention and comparators, with platinum-based chemotherapy without pembrolizumab, such as a) doublet treatment with fluorouracil or

capecitabine plus cisplatin or oxaliplatin or b) triplet treatment with flourouracil or capecitabine plus cisplatin or oxaliplatin epirubicin considered to be the most relevant comparators.

Table 4: Summary of ERG's critique of the methods implemented by the company to identify evidence relevant to the decision problem

Systematic review step	Section of CS in which methods are reported	ERG assessment of robustness of methods
Searches	Appendix D (page 82)	Broadly appropriate, however, the ERG noted the following limitations: specific searches for adverse effects were not completed; the RCT filter applied to database searches did not include terminology to retrieve single-arm studies; and database searches did not include all variant spellings for gastro-oesophageal junction. The ERG conducted additional searches with search terms for single-arm studies and variant spellings (see section 3.5.1) and did not identify any studies that should have been included with respect to the stated inclusion criteria. The ERG also noted that search terms for the drug S1 were not included, however, clinical advice to the ERG confirmed that this intervention is not relevant to current UK clinical practice.
Inclusion criteria	Appendix D (pages 82-83)	Broadly appropriate. Adults with locally advanced unresectable or metastatic adenocarcinoma or squamous cell carcinoma of the oesophagus or advanced/metastatic Siewert type 1 adenocarcinoma of the EGJ were included. Populations were eligible only if they had not received previous therapy. Subjects with human epidermal growth factor receptor-2 positive tumours were excluded.
Screening	Clarification response	Appropriate. No methodological details were provided in CS. However, appropriate methods were described during clarification (clarification question A10)[a]
Data extraction	Clarification response	Appropriate. No methodological details were provided in CS. However, appropriate methods were described during clarification (clarification question A10) [b]
Tool for quality assessment of included study or studies	Section B.2.5 (page 40)	Broadly appropriate. Study quality was assessed using the new Cochrane ROB2 instrument for included RCTs and the Newcastle Ottawa Scale for single-arm trials. Due to lack of reporting in the CS, it was unclear to the ERG if the ROB quality assessments were conducted rigorously i.e. if they were undertaken by two independent reviewers, and any discrepancies between the two reviewers were resolved by consensus or involvement of a third reviewer. The ERG note

Systematic review step	Section of CS in which methods are reported	ERG assessment of robustness of methods
		that the company used the Cochrane ROB2 tool to assess the quality of individual RCTs. This deviates from Cochrane's guidance to use the ROB2 tool to assess bias for individual outcome measures.
Evidence synthesis	Appendix D (pages 82 to 109)	Studies were conducted as set out above ('Searches'). The search criteria did not include gastric/stomach cancer: according to the company this 'would have introduced too many non-relevant studies in terms of location of the primary tumour'. Studies were also excluded when 'subgroup outcomes for oesophageal or esophagogastric junction Siewert type I cancer patients' could not be identified. The approach is reasonable, but filters out some relevant evidence, in particular the REAL-2 trial[9] and existing NMAs.[10,11] This led to a sparser evidence network and affected the suitability for NMA.

Abbreviations: CS, Company submission; EGJ, esophago-gastric junction; ERG, Evidence Review Group; NMA, network meta-analysis; RCT, randomised controlled trial; ROB, Risk of Bias

Notes:

[a] Abstracts were dual screened versus pre-defined eligibility criteria. Discrepancies were resolved with a third party. Potential full text articles were retrieved and screened in the same way. A list of excluded studies was provided in Appendix D, Section D1.1.3, Table 6 of the CS together with reasons for exclusion

[b] Data was extracted by a single reviewer using a pre-defined data extraction template, and data was checked by a second reviewer

While appropriate methods for study inclusion were employed by the company, poor reporting meant that the ERG could not evaluate the robustness of the screening and data extraction processes conducted by the company.

The ERG did not identify any studies that should have been included with respect to the stated criteria.

Appropriate tools for trial quality assessment were chosen by the company, but poor reporting meant that the ERG could not evaluate the robustness of the quality assessment process conducted by the company. The ERG did not consider the company's interpretation of the use of the Cochrane Risk of Bias 2 tool to be appropriate as it was used by the company to assess the quality of individual trials rather than individual outcomes (see Section 3.2.2.6 for more details).

In addition to the clinical effectiveness SLR, the company performed a targeted literature review of prognostic factors (see Appendix D) No studies were identified that reported on prognostic

factors in the target population of interest; however, 13 studies reporting multivariate analyses to identify prognostic factors in similar oesophageal-gastric cancer populations were reviewed and included. These publications and the identified prognostic factors are summarized in Section D1.2.6 of Appendix D. Disease stage was the most common prognostic factor identified (n=9), followed by age (n=7), gender (n=5), tumour size/length (n=5), weight loss/BMI (n=4), lymph node involvement (n=4), and grade (n=4). Clinical advisors to the ERG confirmed that these findings were in line with their clinical experience.

3.2. Critique of trials of the technology of interest, the company's analysis and interpretation (and any standard meta-analyses of these)

3.2.1. Studies included in the clinical effectiveness review

The CS described seven studies (Table 5). These comprised one blinded RCT,[12] one RCT with unknown blinding status,[13] four open-label RCTs[14-17] and one open-label single-arm trial.[18] Only one study (KEYNOTE-590[12]) reported evidence for pembrolizumab in combination with chemotherapy – and therefore forms the pivotal trial in the clinical effectiveness evidence. The remaining included studies assessed potentially relevant comparators but not pembrolizumab, and are addressed further in Section 3.3.

Table 5: Clinical evidence included in the CS

Study name and acronym	Study design	Population	Intervention	Comparator	Study type
KEYNOTE-590[12]	Double-blind placebo-controlled RCT	First-line patients with advanced or metastatic oesophageal carcinoma	Pembrolizumab in combination with cisplatin and 5-fluorouracil	Placebo in combination with cisplatin and 5-fluorouracil	Phase III
Lee 2008[18]	Open-label single-arm trial	First-line patients with advanced oesophageal squamous cell carcinoma	Capecitabine and cisplatin	N/A	Phase II
Lee 2015[14]	Open-label RCT	First-line patients with metastatic esophageal squamous cell carcinoma	Capecitabine in combination with cisplatin	Capecitabine in combination with paclitaxel	Phase II
Lorenzen 2009[17]	Open-label RCT	First-line patients with metastatic squamous cell carcinoma of the oesophagus	Cetuximab plus cisplatin–5-fluorouracil	Cisplatin– 5-fluorouracil	Phase II
POWER[16]	Open-label RCT	Patients with non-resectable, advanced or metastatic oesophageal squamous cell cancer	Cisplatin and 5-fluorouracil with epidermal growth factor receptor inhibition panitumumab	Cisplatin and 5-fluorouracil	Phase III
Ross 2002[13]	RCT (Blinding status N/S)	Patients with advanced oesophagogastric cancer	Mitomycin, cisplatin, and protracted venous-infusion fluorouracil (PVI 5-FU)	Epirubicin, cisplatin, and PVI 5-FU	N/S

Study name and acronym	Study design	Population	Intervention	Comparator	Study type
Wang 2017[15]	Open-label RCT	Patients with metastatic oesophageal squamous cell carcinoma	Recombinant human lymphotoxin-a derivative in combination with cisplatin and 5-fluorouracil (at two different doses)	Cisplatin and 5-fluorouracil	Phase IIb

Abbreviations: 5-FU, 5-fluorouracil; CS, company submission; N/A, Not applicable; N/S, Not stated; PVI, protracted venous infusion; RCT, randomized controlled trial

3.2.2. Description and critique of the design of the studies

3.2.2.1. Design of the studies

The key trial included in the company's SLR, and the only source of directly comparative evidence to inform the economic model, was a Phase III double-blind placebo-controlled RCT (KEYNOTE-590[12]) evaluating pembrolizumab in first-line patients with advanced or metastatic oesophageal carcinoma from 26 countries worldwide. There were 22 participants from three sites in the United Kingdom out of a total of 749 participants worldwide (2.9%). The majority of participants in KEYNOTE-590[12] (N=400, 53.4%) were from Asian sites. The population, intervention, comparator and outcomes in KEYNOTE-590[12] were broadly consistent with the NICE decision problem.

3.2.2.2. Population

The KEYNOTE-590[12] study considered a population of patients with locally advanced unresectable or metastatic adenocarcinoma or squamous cell carcinoma of the oesophagus or advanced/metastatic Siewert type 1 adenocarcinoma of the oesophageal-gastric junction. Eligible patients had an ECOG score of 0 or 1, no active central nervous system metastases and/or carcinomatous meningitis, and no active infection or autoimmune disease that required systemic therapy. Detailed inclusion and exclusion criteria are provided in CS Section B.2.3.1, pp.22-24. The population for the pivotal trial was narrower than the NICE scope[8] for this appraisal. However, this was in accordance with the proposed marketing authorisation and was therefore considered appropriate.

3.2.2.3. Intervention

3.2.2.4. The intervention in the KEYNOTE-590[12] study was pembrolizumab 200 mg intravenously every three weeks in combination with cisplatin 80 mg/m^2 intravenously every three weeks and 5-FU 800 mg/m^2/day continuous intravenous infusion (4000 mg/m^2 per three-week cycle). The pembrolizumab dosing regimen was in accordance with the draft SmPC. Clinical advice to the ERG indicated that 5-FU was not the optimal comparator in light of UK clinical practice, where capecitabine-based doublet regimens would be more commonly used. While clinical advice to the ERG was that the choice of comparator regimen would be unlikely to have a substantial impact

upon relative efficacy, the choice of comparator could have cost implications (discussed further in Section 4.2.4). Comparator

The comparator in the KEYNOTE-590[12] study was saline placebo intravenously every three weeks in combination with cisplatin 80 mg/m^2 intravenously every three weeks and 5-FU 800 mg/m^2/day continuous intravenous infusion (4000 mg/m^2 per three-week cycle). Clinical advisers to the ERG indicated that the comparator used in KEYNOTE-590[12] would be considered 'old-fashioned' in the context of routine NHS clinical practice, where oxaliplatin and capecitabine would typically be used in preference (discussed further in Section 4.2.4). Moreover, the method of administration of chemotherapy in the trial, requiring inpatient or PICC line would be considered dated, where typical NHS practice is to provide chemotherapy treatments for oesophageal cancer in a day case setting.

3.2.2.5. Outcomes

The outcomes covered in the KEYNOTE-590[12] study were summarised in the CS (Section B.2.2., Table 3, p.19). Data for the five outcomes specified in the NICE scope[8] were available, and are outlined below. Time to deterioration, duration of response, patient reported outcomes and disease control rate were also available.

Overall survival

Overall survival was defined as the time from randomisation to death by any cause.

Progression-free survival

Progression-free survival was defined as the time from randomisation to the first documented disease progression per RECIST 1.1 criteria,[19] or death due to any cause, whichever occurs first.

Response rate

The measure of response rate used was objective response rate. This was determined per RECIST 1.1 criteria.

Adverse effects of treatment

The safety and tolerability of pembrolizumab was assessed. Total and cause-specific adverse events were profiled.

Health-related quality of life

Health-related quality of life (HRQoL) was assessed using EuroQoL EQ-5D-5L,[20] as well as disease-specific European Organization for the Research and Treatment of Cancer (EORTC) Quality of Life Questionnaire C30 (QLQ-C30)[21] and the EORTC Quality Of Life Questionnaire Oesophageal Module (QLQ-OES18)[22] measures. EuroQoL EQ-5D-5L[20] was mapped to EQ-5D-3L[23] for use in the economic model following NICE recommendations.

3.2.2.6. Critical appraisal of the design of the studies

The company reported no notable quality issues in relation to the KEYNOTE-590[12] RCT. The complete quality assessment is available in Appendix D of the CS (Section D1.4. Table 3, p.125). The company evaluated KEYNOTE-590[12] using the Cochrane Risk of Bias 2 tool, which the ERG considers an appropriate critical appraisal tool for RCTs. The ERG noted that the company had not followed Cochrane guidance on the correct use of the tool and used the tool to assess individual trial quality rather than the quality of assessment of individual outcomes. However, the ERG did not identify any concerns for risk of bias specifically for the outcomes reported in the CS that informed the decision problem/economic model (primarily overall survival [OS], response (ORR), progression free survival [PFS], HRQoL and also adverse effects [AEs] of treatment).

3.2.3. Description and critique of the results of the studies

3.2.3.1. Baseline characteristics

Baseline characteristics for patients included in the KEYNOTE-590[12] study were reported in the CS (Section 2.3.3, Table 6, p.31) for the ITT population. Considering the ITT population, the ERG agreed with the company's assertion that the pembrolizumab and control arms were generally well balanced for baseline characteristics and reasonably representative of the target population, with an important exception. The ERG noted an important departure from the expected UK clinical practice population with regard to histology. In the CS (Section 1.3, p.15), the company cited evidence[5] that adenocarcinoma accounts for approximately two thirds of UK cases of oesophageal cancer, while squamous cell carcinoma accounting for approximately one third. However, in the KEYNOTE-590[12] ITT population, patients with squamous cell carcinoma accounted for 73.2% of all participants. This substantial overrepresentation of patients with squamous cell carcinoma compared to the UK clinical practice population may have implications for the generalisability of the trial evidence to NHS clinical practice settings in England and

Wales. Moreover, it is important to note that the company used population characteristics from European patients (CS Section B.3.2, Table 41, p.81) in KEYNOTE-590[12] to populate the economic model, although clinical effectiveness inputs were from the global ITT population. The ERG, however, considered that the international population may be more suitable, given the population of Europe as a whole is less ethnically diverse than the UK population.

3.2.3.2. Clinical effectiveness results

Data in the target population were presented for overall survival, progression-free survival, objective response rate, adverse events and health-related quality of life. Statistical analyses were broadly appropriate. The primary efficacy population was the global intention to treat (ITT) population. The primary safety population was the global all subjects as treated (ASaT) population. Efficacy analyses were performed using the July 2020 Interim Analysis dataset, at which the median (range) duration of follow-up was 12.6 (0.1 to 33.6) months in the pembrolizumab arm and 9.8 (0.1 to 33.6) months in the control group, with the exception of patient-reported outcomes, such as HRQoL. These were assessed in the patient reported outcome full analysis set (PRO FAS), which comprised participants who had received at least one dose of study medication and had completed at least one patient-reported outcome assessment.

Overall survival

In the ITT population, the median overall survival was 12.4 months (95% CI 10.5-14.0 months) for the pembrolizumab arm compared to 9.8 months (95% CI 8.8-10.8 months) for the control arm. There was a 27% reduction in the risk of death for people in the pembrolizumab arm compared to the control arm (HR = 0.73, 95% CI 0.62-0.86, p<0.0001).

Progression-free survival

In the ITT population, using the primary PFS censoring rule, the median progression-free survival was 6.3 months (95% CI 6.2-6.9 months) for the pembrolizumab arm compared to 5.8 months (95% CI 5.0-6.0 months) for the control arm. There was a 35% reduction in the risk of progression or death for people in the pembrolizumab arm compared to the control arm (HR = 0.65, 95% CI 0.55-0.76, p<0.001).

The company conducted two sensitivity analyses of PFS using different censoring rules, as outlined in the CS Section B 2.4.1, Table 9, as well as the primary analysis. Results were only

provided in the CS using the primary censoring rule for PFS. Therefore, the ERG asked the company at the Clarification stage for the PFS results using the alternative censoring rules. The HRs for people in the pembrolizumab arm compared to the control arm were [redacted] (95% CI [redacted]) using sensitivity censoring rule 1 and [redacted] (95% CI [redacted]).using sensitivity censoring rule 2. The ERG was satisfied that the choice of PFS censoring rule had little material impact on the PFS results.

Objective response rate

In the ITT population, the objective response rate was 45.0% for the pembrolizumab arm compared to 29.3% for the control arm. This 15.8% difference was considered clinically and statistically ($p<0.0001$) significant.

Health-related quality of life

In the PRO FAS population, there were no clinically meaningful changes in EQ-5D VAS scores from baseline to week 18 for either the pembrolizumab or control arm, and there was no statistically significant difference in change scores from baseline to week 18 between the arms (Baseline mean (SD) pembrolizumab arm 72.59 (18.65); Baseline mean (SD) control arm 74.43 (17.14); week 18 mean (SD) pembrolizumab arm 72.41 (18.55); week 18 mean (SD) control arm 74.04 (16.59)).

Subgroup analyses

Pre-specified subgroup analyses were conducted according to the following stratification factors:

- Histology (adenocarcinoma vs squamous cell carcinoma)
- Geographic region (Asia vs Rest of World)
- ECOG performance status (0 vs 1)
- Disease status (Locally advanced vs metastatic)
- Age category (binary split at 65)
- Sex (male vs female)

The ERG considered histology and geographic region to be particularly important stratification factors, given that the majority of patients were Asian and the balance between adenocarcinoma and squamous cell carcinoma patients was markedly different than would be found in a UK population. Arm-level subgroup results for the two primary efficacy outcomes of overall survival and progression-free survival (CS, Appendix E) were presented solely in terms of numbers of events rather than median survival estimates, which the ERG considered unhelpful in terms of interpreting these results in the context of the headline ITT results. Hazard ratios for overall survival (adenocarcinoma 0.74 (95% CI 0.54, 1.02); squamous cell carcinoma 0.72 (95% CI 0.60, 0.88) and progression-free survival (adenocarcinoma 0.63 (95% CI 0.46, 0.87); squamous cell carcinoma 0.65 (95% CI 0.54, 0.78)) were comparable for adenocarcinoma and squamous cell carcinoma patients, although this comparison must be interpreted with caution due to unequal numbers of patients in the two histological groups. Moreover, the different treatment pathways associated with these two histological groups may have an impact in terms of resource use and costs. However, the relative benefit of pembrolizumab versus control on both overall survival (Asia HR [redacted]); Rest of the World HR [redacted]) and progression-free survival (Asia HR [redacted] Rest of the World HR [redacted]) was stronger for Asian patients than in the Rest of the World. This may mean that the effect estimates in the KEYNOTE-590[12] global ITT population overestimate the clinical effectiveness of pembrolizumab compared to the UK clinical practice setting.

3.2.3.3. Safety results

Adverse effects

Adverse events (AEs) in the KEYNOTE-590[12] study were reported in the CS (Section B.2.10). AEs were considered in the ASaT population, which formed the primary safety analysis population. Overall, the ERG agreed with the company that pembrolizumab had an acceptable safety profile and that AE rates were comparable between the pembrolizumab and control arms. However, the ERG noted that AEs were very common, with all participants in the pembrolizumab arm and 99.5% of participants in the control arm encountering at least one AE. The profile of AE types was comparable between arms, the most common being nausea, anaemia, and decreased appetite. Serious AEs were encountered by the majority of participants, although rates were comparable between arms (55.4% pembrolizumab vs 55.1% control).

Mortality

Death rates were lower overall on pembrolizumab (7.6%) than control (10.3%). However, deaths due to drug-related adverse events were more common on pembrolizumab (2.4%) than control (1.4%). The ERG asked at the clarification stage for information regarding death rates at one month, three months, six months, 12 months, 18 months and two years from randomisation in the pembrolizumab and control arms, and was satisfied that there was no evidence of an elevated risk of early deaths on pembrolizumab.

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

3.3.1. Summary

The company search criteria are defined in CS Appendix D Table 1. The search identified '7 studies relevant to the UK context' (CS doc B table 26) including the pivotal trial KEYNOTE-590,[12] and the company sifted further to reach a final total of three studies. Examining these, the company rejected use of a NMA (small, disconnected network) and MAIC (differences between populations and with target population). The ERG notes that a relaxation of the search and exclusion criteria or disease definition increases the available evidence. No further trial comparisons relating specifically to the most typical UK doublet and triplet regimens were found, but some further evidence is available for regimens used previously and believed to have similar efficacy.

3.3.2. Company approach

The company identified seven trials (CS, Appendix D, Figure 1), six of which were in patients with oesophageal cancer only, and these are listed in the CS (Doc B, Table 26). Assessing these for use within a network, the company stated that three trials were excluded for using the comparison (cisplatin + 5FU) 'which is already captured in the population of interest in the index trial, KEYNOTE-590.' (CS, Doc B, p61).

It was unclear to the ERG why network connections to a node representing a common, indeed central node (in this case cisplatin + 5FU) led to exclusion by the company of these studies. On the other hand, clinicians advised the ERG that the interventions used in the three excluded trials (cetuximab, panitumumab and rhLTα-DA) were irrelevant to UK practice, and for this

reason the ERG agreed with their exclusion. A further study was then excluded 'due to a lack of reported patient characteristics for the population of interest'.

The final set of three studies included a trial with IPD (KEYNOTE-590[12]) and two studies with aggregate outcomes.[14,18] Therefore, the company examined them for potential analysis by unanchored MAIC, but this was rejected because of key differences between study populations, including that the comparator studies were limited to South Korean patients with ESCC only.

The company indicated (in response to clarification A6), that the search for evidence excluded the term 'gastric', and the ERG believed that studies covering patients with more general advanced gastric/ stomach cancer (but including relevant information for the decision problem) might have been sifted out. Furthermore (in response to clarification A5) studies were excluded when results specifically for oesophageal or oesophagogastric junction Siewert type I cancer patients could not be identified. The ERG accepted that this was a coherent approach with regard to the decision problem, but noted that:

- relevant patients with junctional/oesophageal cancer that has metastasised to or from the stomach might then be excluded, if the site of the primary tumour is unclear or unreported;
- this rule was not set out in the exclusion criteria (CS Appendix D Table 1);
- Siewart type I is not specified in the scope (CS Doc B Table 1); and,
- the trial by Cunningham (REAL-2 trial)[9] was consequently rejected; yet this influential study provided evidence of the noninferiority of oxalipatin and capecitabine, and the ERG understood that it underpinned the most typical UK clinical practice of substituting these for cisplatin and 5-FU respectively.

3.3.3. Further trials

The ERG was also aware of several existing NMAs and meta-analyses in advanced gastric cancer (ter Veer, 2016; Guo, 2019; Wagner, 2006; Wang, 2017; Okines, 2009; GASTRIC, 2013),[10,11,24-27] albeit often with broader disease definition than the CS. These are termed hereon 'the reviews'. The reviews effectively supply some results of searches for head-to-head trials without excluding the terms 'stomach' or 'gastric'. The ERG focused on further evidence for doublet vs triplet regimens relevant to the UK.

According to clinical advisors to the ERG, the most typical UK treatment/ UK standard of care doublet is (oxaliplatin + capecitabine) and most typical triplet is (oxaliplatin + capecitabine + epirubicin). No further direct comparisons of these were found within the reviews. Ter Veer et al. (2016)[10] remarked in particular that '… although conventional anthracycline-, platinum-, and fluoropyrimidine-based triplets, as defined in the REAL-2 study are used frequently in clinical practice, head-to-head RCTs are missing between these triplets and fluoropyrimidine-based doublets (i.e., fluoropyrimidine/oxaliplatin).'

Some head-to-head trials between relevant UK regimens used more frequently in the past (cisplatin + 5-FU) vs (cisplatin + 5-FU + epirubicin) were found by the reviews but not by the company's search. These were Kim et al. (2001),[28] and KRGCGC (1992).[29] Another relevant doublet vs triplet trial cited is Yun et al. (2010)[30] between (cisplatin + capecitabine + epirubicin) vs (cisplatin + capecitabine). With regard to their omission, the company explained (response to clarification A8) that: 'All of the above studies were conducted in gastric cancer patients and did not include oesophageal or esophagogastric junction Siewert type I patients, therefore are not relevant to the current decision problem'. But, somewhat contrary to this statement, the company also refers to the results of these three studies in CS Doc B Table 43 (comparison number 2 ; these same studies were identified by NG83[7] Section 9.2.2). The ERG reappraised the studies and while it agreed that they were carried out in gastric cancer patients, it was not apparent that oesophageal or oesophagogastric junction cancer patients were not included, as was stated by the company. The ERG further noted that the scope does not specify Siewert type I patients.

There are several further caveats on the use of data from these three trials including:

- The populations are wholly Asian, precluding generalisability to a UK context.
- No support (or rejection) of the constant HR assumption is shown in Kim et al. (2001)[28] or KRGCGC (1992).[29] Indeed, no HRs are given at all, and the HRs adopted in the NMA appear to be based on the reported medians with an assumption of an exponential survival distribution.
- Yun et al. (2010)[30] gives crossing KM curves (conflicting with the constant HR assumption).

3.4. Critique of the indirect comparison and/or multiple treatment comparison

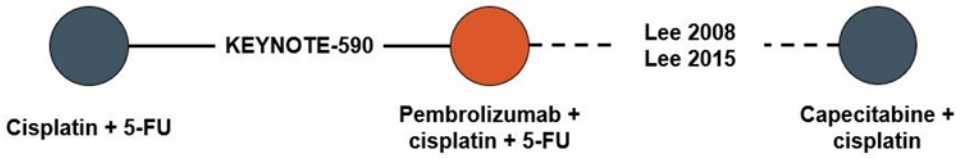
3.4.1. Summary

The company did not carry out an NMA, because the network formed under its search and exclusion criteria was uninformative. The company then considered but rejected a MAIC, primarily because of differences between study populations. While the ERG agreed with these decisions, it considered that the strict exclusion of stomach/gastric cancer, while coherent, limits the available evidence (and in particular excludes an important and influential study, REAL-2 ; Cunningham et al. (2008) [9]). The ERG is aware of several existing meta-analyses/NMAs,[10,11,24-27] and has summarised the most relevant results for UK doublet versus triplet comparisons.

3.4.2. Company approach

The company did not carry out a network meta-analysis, because the network formed under its search criteria (CS Appendix D Table 1 and response to clarification A6) and other exclusion criteria (response to clarification A5) was very minimal. The company's network is reproduced here in Figure 2. Even this small network is disconnected, since two trials shown do not include a common comparator with KEYNOTE-590[12] (improper connection shown by dotted line).

Figure 2. Network diagram of studies identified through SLR



Source: CS Doc B Figure 3

The CS stated that a NMA of (pembrolizumab + chemotherapy) versus competing interventions (including capecitabine plus cisplatin and epirubicin with cisplatin and 5-FU) was not feasible

'*because these interventions have generally only been evaluated in non-comparative studies*.' (CS Doc B Section B.2.9).

The ERG's view differed from this, insofar as there is evidence of existing comparative studies and NMAs, but with a looser definition of the disease and perhaps an assumption of class effects. This is discussed further below.

Having rejected an NMA, the feasibility process was 'adapted to the context of an unanchored MAIC'. The company provided further details of this process in response to clarification A9. The company did not carry out an (unanchored) MAIC on the final three studies because of differences between populations (further details in Section 3.3.2). The ERG agreed with this decision.

3.4.3. Existing indirect comparisons

The ERG is aware of two relevant published NMAs (Section 3.3.3). The disease definition in Wang (2017)[11] on esophagogastric junctional adenocarcinoma overlaps with the decision problem, but its results are presented in terms of broad drug classes that the ERG judged are not suitable for the decision problem (in particular, grouping together first- and second-line drugs: docetaxel, epirubicin and irinotecan). Results obtained from another NMA (ter Veer et al. 2016)[10] are potentially useful because the drug comparisons are apt, but against this, were targeted at less specific disease ('patients with pathologically proven metastatic, unresectable, or recurrent adenocarcinoma of the esophagus, gastroesophageal junction (GEJ), *or stomach'* [ERG italics]).

Clinical experts advised the ERG that advanced oesophageal cancer and gastric cancers are similar in terms of systemic therapy, but also that squamous carcinoma tends to occur in the upper two-thirds of the oesophagus while junctional cancer, like gastric cancer, is predominantly adenocarcinoma.

From the NMA, ter Veer et al. (2016)[10] give results for 'ACF' vs 'CF' (A=anthracycline, C=cisplatin, F= fluoropyrimidine): For OS, HR = 0.86 (95% CI: 0.71 to 1.02; ter Veer et al. Figure 3) and for PFS, HR = 0.85 (95% CI: 0.68 to 1.05; ter Veer et al.[10] Figure. The ERG has briefly critiqued (Section 3.3.3) the three underlying studies in the NMA that provide direct comparative evidence of ACF vs CF (Kim, 2001; KRCGGC, 1992; Yun 2010)[28-30] and some caveats were

noted. On the other hand, individual comparisons are supported by indirect as well as direct evidence in this large NMA (17 regimens and 37 direct comparisons for OS).

The ERG found that several reviews pool direct evidence on the comparison of (cisplatin + 5-FU) vs (cisplatin + 5-FU + anthracyline). For example, NG83 provides an estimate of HR 0.70, 95% CI: 0.43-1.15 (CS Doc B table 43, comparison 2) based on direct evidence alone. The REAL-2 trial (Cunningham et al. 2008)[9] provided evidence that in terms of efficacy, oxaliplatin is noninferior to cisplatin, and capecitabine is noninferior to 5-FU. Taken together these provide direct evidence for the current UK SoC (oxaliplatin + capecitabine) vs (oxaliplatin + capecitabine + epiribucin) on the assumption of exchangeability/class effects, with oxaliplatin substituting for cisplatin and capecitabine substituting for 5-FU. A broader disease definition is adopted when admitting evidence from these trials, but the ERG noted that REAL-2 is likewise premised on a broad disease definition ('carcinoma of the esophagus, gastroesophageal junction, *or stomach* that was locally advanced (inoperable) or metastatic').

3.5. Additional work on clinical effectiveness undertaken by the ERG

3.5.1. Searches

The ERG conducted searches of Ovid MEDLINE (1st March 2021) and Embase (8th March 2021) to confirm that the company's literature searches had identified all relevant studies. These searches used additional free-text search terms for gastro-oesophageal junction adenocarcinoma (and alternate spellings) and single-arm study designs, but did not include search terms for 'gastric' or 'stomach' neoplasms. (Full search strategies are available in Appendix A). The titles and abstracts of search results were screened by one reviewer. The ERG identified two studies[31,32] to review at full-text. The study by Lordick et al. (2013)[31] included patients with HER-2 positive advanced gastric cancer so was excluded on the basis of population. Shah et al. (2017)[32] was excluded as onartazumab was not considered a relevant comparator. The ERG did not identify additional single-arm studies or other relevant trials that should have been included with respect to the stated criteria. However, as described in Section 3.3, the ERG identified further trials from existing NMAs with a broader disease definition of stomach/gastric cancer.[10,11]

3.5.2. Network evidence

The ERG carried out exploratory survival and cost-effectiveness analysis using an effect size estimate on doublet versus triplet treatment, applied to the doublet arm of the KEYNOTE-590[12] trial. The results of applying the NMA-derived hazard ratio from ter Veer et al. (2016)[10] (Section 3.4) within the economic model to the preferred OS and PFS distributions under doublet therapy are described in Section 6.2.4.

3.6. Conclusions of the clinical effectiveness section

The ERG considered that the company had identified all relevant clinical evidence for this appraisal. All key outcomes from the NICE final scope[8] were covered in the CS. Requisite information regarding the methodology and outcomes for clinical effectiveness was available in the CS and clarification responses provided by the company, and was generally reasonably described.

There was one RCT (KEYNOTE-590[12]) comparing pembrolizumab 200 mg intravenously every three weeks in combination with cisplatin 80 mg/m^2 intravenously every three weeks and 5-FU 800 mg/m^2/day continuous intravenous infusion (4,000 mg/m^2 per three-week cycle) to saline placebo intravenously every three weeks in combination with cisplatin 80 mg/m^2 intravenously every three weeks and 5-FU 800 mg/m^2/day continuous intravenous infusion (4,000 mg/m^2 per three-week cycle) that could provide directly comparative evidence for the base case economic model. All other studies included in the company's SLR did not assess pembrolizumab, but rather potentially relevant comparator treatments. The ERG was satisfied that KEYNOTE-590[12] was generally a high quality trial. The ERG's concerns about the trial related to generalisability rather than internal validity. The ERG was satisfied that the trial showed a benefit for pembrolizumab over placebo in terms of OS and PFS. The company considered an NMA or a MAIC analysis given that the comparator regimen in the pivotal trial did not encompass the range of eligible comparators in the NICE final scope.[8] However, the options of conducting an NMA and MAIC were both rejected by the company due to a small, disconnected network and differences between populations and with target population respectively. The ERG identified existing NMAs in gastric cancer and considered, given the similarity of treatment pathways, that data from this broader population could potentially be informative.

The key issues in the clinical effectiveness evidence identified by the ERG were as follows:

- The clinical evidence may not be generalizable to the UK population
- Clinical effectiveness evidence excluded probative estimates of effectiveness between standard of care regimens.

4. COST-EFFECTIVENESS

4.1. ERG comment on company's review of cost-effectiveness evidence

The company carried out a SLR to identify existing cost-effectiveness evidence, HRQoL evidence, and cost and healthcare resource use evidence in adults with advanced, unresectable or metastatic oesophageal cancer, including carcinoma of the gastro-oesophageal junction. A summary of the ERG's critique of the methods implemented by the company to identify relevant cost-effectiveness evidence is presented in Table 6.

Table 6. Summary of ERG's critique of the methods implemented by the company to identify cost-effectiveness evidence

Systematic review step	Section of CS in which methods are reported	ERG assessment of robustness of methods
Searches	Appendix G	Appropriate
Inclusion criteria	Appendix G (Table 1, p.160)	Appropriate. Broad criteria were applied. Full economic evaluations of interventions aimed at managing advanced, unresectable or metastatic OC (including carcinoma of the gastro-oesophageal junction) published in English language from data inception to Year 2020 were included as per NICE scope
Screening	Appendix G	Appropriate [a]
Data extraction	Appendix G	No details provided in Appendix G. However, no cost-effectiveness studies relevant to the UK population were identified during screening.
QA of included studies	Appendix G	

Abbreviations: CS, Company Submission; ERG, Evidence Review Group; OC, oesophageal cancer; QA, quality assessment

Notes:

[a] Abstracts were dual screened versus pre-defined PICOS selection criteria. Discrepancies were resolved with a third party. Potential full text articles were retrieved and screened in the same way. A list of excluded studies was provided in Appendix G Table 2, p.165 of the CS together with reasons for exclusion

The ERG was satisfied with the company's review of the cost-effectiveness literature. Ten economic evaluations were identified. The ERG agreed with the company's judgment that none of these ten studies were relevant to the UK population and were hence correctly not summarised in the CS.

Table 7. Summary of ERG's critique of the methods implemented by the company to identify health related quality of life (in terms of utilities)

Systematic review step	Section of CS in which methods are reported	ERG assessment of robustness of methods
Searches	Appendix H	Broadly appropriate. The ERG noted that the search strategies used to identify studies reporting HRQoL or utility values did not include terms for specific measures (e.g. EQ-5D), however, the ERG was satisfied that all relevant HRQoL literature was identified.
Inclusion criteria	Appendix H (Table 20, p.179)	Appropriate. Broad criteria were applied. Studies reporting HRQoL or utility values related to advanced, unresectable or metastatic OC, including carcinoma of the gastro-oesophageal junction, published in English language from data inception to Year 2020 were included.
Screening	Appendix H	No detail provided. It was unclear to the ERG if screening was performed independently by two reviewers. Study selection was documented in a PRISMA flow diagram (CS, Appendix H, Figure 4).
Data extraction	Appendix H	No detail provided. The company summarised details for the identified studies (CS, Appendix H, Table 21)
QA of included studies	Appendix H	No detail provided. No formal critical appraisal of the studies was conducted, however the company did provide an assessment of the consistency of each study with the reference case (CS, Appendix H, Table 22)

Abbreviations: CS, Company Submission; ERG, Evidence Review Group; HRQoL, health-related quality of life; OC, oesophageal cancer; QA, quality assessment

Notes:

[a] Abstracts and full text articles were screened versus pre-defined eligibility criteria (Appendix G, Table 20, p.178) with no further details provided in the CS.

b Data was extracted using a pre-defined data extraction template (Appendix G (Table 22)), with no further details provided in the CS.

The ERG was broadly satisfied with the company's review of the literature reporting health effects (health-related quality of life and utilities). The company identified nine studies[12,33-40] reporting utility estimates in people with OC which are summarised in Appendix H (Table 21 and

Table 22) of the CS. The ERG noted the absence of methodological reporting for screening and data extraction. While no formal critical appraisal of studies was conducted, the company provided an assessment of the consistency of each study with the reference case. The ERG noted that none of the nine studies identified in the review of utilities were used in the model. The ERG was satisfied that the incorporation of utilities data from KEYNOTE-590[12] only into the model was appropriate.

Table 8. Summary of ERG's critique of the methods implemented by the company to identify healthcare resource use and costs

Systematic review step	Section of CS in which methods are reported	ERG assessment of robustness of methods
Searches	Appendix I	Appropriate
Inclusion criteria	Appendix I (Table 40, p.202)	Appropriate. Broad criteria were applied. The company included studies reporting healthcare costs and/or resource use in the treatment and on-going management of advanced unresectable or metastatic oesophageal cancer (including carcinoma of the gastro-oesophageal junction) in order to evaluate the economic burden of oesophageal cancer in the United Kingdom. Studies published in English language from data inception to Year 2020 were included.
Screening	Appendix I	Appropriate [a]
Data extraction	Appendix I	No detail provided. The company summarised details for one study which they judged to meet the criteria of the UK population (Appendix I, Table 41)
QA of included studies	Appendix I	No details provided. No formal critical appraisal was provided.

Abbreviations: CS, Company Submission; ERG, Evidence Review Group; HRQoL, health-related quality of life; QA, quality assessment

Notes:

[a] Abstracts were dual screened versus pre-defined PICOS selection criteria (CS, Appendix I, Table 40).Discrepancies were resolved with a third party. Potential full text articles were retrieved and screened in the same way. A mapping of excluded studies together with reasons for exclusion were provided in a PRISMA flow diagram (CS, Appendix I, Figure 5).

The ERG was broadly satisfied with the company's review of the literature reporting healthcare resource use and costs. The company identified 16 studies which reported cost or resource use

associated with advanced, unresectable or metastatic oesophageal cancer, of which only one study was judged by the company to meet the criteria of the UK population.[41] However, there was no discussion of the applicability of the identified study to the economic model within the CS.

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

4.2.1. NICE reference case checklist

Table 9: NICE reference case checklist

Attribute	Reference case	ERG comment on company's submission
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	✓
Perspective on costs	NHS and PSS	✓
Type of economic evaluation	Cost–utility analysis with fully incremental analysis	✓ Pairwise comparison of pembrolizumab in combination with chemotherapy versus trial comparator or non-trial blended comparator
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	✓ Time horizon of 20 years was originally used. The ERG noted that 3% of patients were still alive in the pembrolizumab + chemotherapy arm. Company amended time horizon to 30 years post clarification questions
Synthesis of evidence on health effects	Based on systematic review	✓ Systematic review undertaken to identify relevant evidence. However none of the findings were used to inform the submission.
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of health-related quality of life in adults.	✓ EQ-5D utility values used to inform the model using a time-to-death approach. Sensitivity analysis presented where utility values based on progression status.
Source of data for measurement of health-related quality of life	Reported directly by patients and/or carers	✓ Reported directly by patients in the KEYNOTE-590 trial
Source of preference data for valuation of changes in health-related quality of life	Representative sample of the UK population	✓ Based on van Hout et al. (2012)[42] cross walk value set

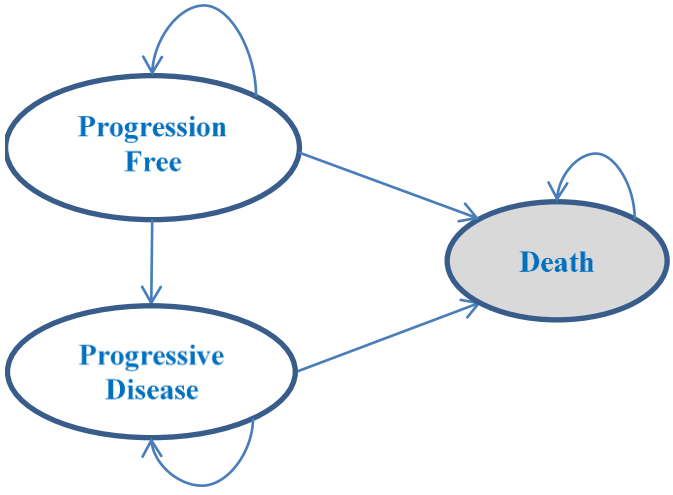
Attribute	Reference case	ERG comment on company's submission
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	✓
Evidence on resource use and costs	Costs should relate to NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS	✓
Discounting	The same annual rate for both costs and health effects (currently 3.5%)	✓

Key: EQ-5D, EuroQol 5 dimension; HRQoL: health-related quality of life; NHS, National Health Service; PSS, Personal Social Services; QALY: quality-adjusted life year; TA: technology appraisal

4.2.2. Model structure

The company developed a partitioned survival analysis (PartSA) model to estimate health outcomes and costs for pembrolizumab in combination with chemotherapy versus chemotherapy. The company's model schematic is shown in Figure 4.

Figure 4: Company's model schematic



Source(s): CS Figure 10

The three mutually exclusive health states; progression-free, progressive disease and death, are informed by the overall survival (OS) and progression-free survival (PFS) curves. The area under the OS curve is used to estimate the proportion of patients alive over time, and the area

under the PFS curve is used to estimate the proportion of patients who are in the progression-free health state over time. The difference between OS and PFS is used to estimate the proportion of patients in the progressive-disease health state. KEYNOTE-590[12] data were used to generate the PFS and OS curves either by using the Kaplan-Meier estimates or from a parametric distribution.

The company justified its choice of model structure based on its extensive use in previous advanced cancer NICE submissions[43-45] and a recent advanced oesophageal cancer in the second-line setting.[46] The ERG considered the choice of model structure to be appropriate and suitable for decision making in this disease area. Nevertheless, there are several limitations with the PartSA approach which are important to note when interpreting results and model functionality:

- The proportion of patients who progress per model cycle is not explicitly modelled. Thus, there are limitations when needing to assign costs to the exact proportion of patients who progress. Within this context, patients who progress are assigned a one-off subsequent treatment cost which is applied to the proportion of patients who leave the progression-free state and includes those leaving due to death as well as progression. However, it should be noted that post-clarification questions, the company calculated the one-off subsequent treatment cost based on the proportions from the total number of progression events and not just those who progressed and therefore the ERG did not consider this to be of great concern.

- The use of an overarching OS curve impacts the relationship between progressive-disease costs versus efficacy (e.g., subsequent treatment inputs). The company's base case uses the same subsequent treatment distribution as the modelled efficacy, however, should changes to the subsequent treatment distribution be needed to account for a new treatment, this would lead to higher costs with no direct impact on efficacy. At the time of writing, nivolumab monotherapy is not currently available in UK clinical practice, as NICE ID1249 is still under consultation. At clarification stage the company provided further information concerning subsequent treatments, highlighting that a small proportion of patients on both treatment arms went on to receive nivolumab after progression (3.2% for pembrolizumab in combination with chemotherapy versus 4.6% chemotherapy). Therefore, the ERG did not consider this limitation to be of great concern.

In the company's base case, the PFS outcomes only impacted costs associated with resource use and subsequent treatment. Drug costs are calculated separately using a Kaplan-Meier estimates of time-on-treatment (ToT) to estimate the proportion of patients on treatment on either pembrolizumab in combination with chemotherapy or chemotherapy. Quality-adjusted life-years (QALYs) were estimated from utilities using the time-to-death approach and are therefore not impacted by ToT or PFS. Therefore, although the model structure was described as progression-based, progression itself had no impact on quality of life or life-years in the company's base case analysis.

4.2.3. Population

The company stated that the model considered patients with untreated, unresectable locally advanced or metastatic oesophageal cancer or HER-2 negative gastroesophageal junction adenocarcinoma (CS Section B.3.2.2). This population aligned with the anticipated licence and NICE final scope.[8] KEYNOTE-590[12] was used to inform the population and efficacy model parameters, and was reflective of patients with *"locally advanced unresectable or metastatic adenocarcinoma or squamous cell carcinoma of the oesophagus or advanced/metastatic Siewert type 1 adenocarcinoma of the esophagogastric junction"* (CS Section B.2.3.1). The company also included a subgroup population considering patients with untreated, unresectable locally advanced or metastatic oesophageal cancer or HER-2 negative gastroesophageal junction adenocarcinoma who had a combined positive score (CPS) greater than or equal to 10 (CPS≥10). This subgroup was considered by the company to be of clinical significance.

Patient characteristics (age, sex, body weight, and body surface area [BSA]) in the model were based on European patients from KEYNOTE-590.[12] In its submission, the company did not provide an explicit justification for using only the European subgroup to inform its base-case analysis. However, the ERG noted that the full population patient characteristics did not differ greatly from the European patients with the exception of body weight and BSA (see Table 10). Based on the characteristics of the full KEYNOTE-590[12] study population, the ERG expected that the lower average body weight/ BSA in the full ITT population is driven by the inclusion of Asian patients (for whom body weight and BSA are typically lower than a European population). The ERG considered the use of European patient characteristics to inform body weight and BSA in particular to be most appropriate to inform the base-case analysis.

Table 10: Baseline characteristics of patients included in the model (European versus all patients)

Patient characteristic	All patients		CPS ≥ 10	
	European patients	**All patients**	**European patients**	**All patients**
Average age	61.4	62.4	60.8	61.9
Proportion male	80.7%	83.4%	71.9%	81.7%
Average patient weight (kg)	71.2	63.1	68.4	62.6
Body surface area (m^2)	1.84	1.70	1.79	1.70

Abbreviations: CPS, combined positive score

Source(s): CS Table 41; CS Appendix M Table 1; company model (KEYNOTE-590)

The ERG noted some additional important features of the KEYNOTE-590[12] study which have implications concerning the generalisability of the modelled patient population versus those patients that would be treated in UK clinical practice (see Section 3.2.3.1).

Given the inclusion criteria of KEYNOTE-590[12] specifying patients with ECOG status of 0 or 1, it is likely that these patients in the trial are fitter than the UK oesophageal cancer population which includes those with an ECOG score >1. While this is a common feature of many advanced cancer trials, this means that the KEYNOTE-590[12] study does not provide information concerning the safety or efficacy of pembrolizumab in combination with chemotherapy in an ECOG 2+ population.

Over half of the KEYNOTE-590[12] study population comprised of patients from Asia (52.5%, versus 47.5% from the rest of the world [ROW]) (ITT population CS Table 6). Region has an apparent impact on the HR for OS (0.64 [95% CI 0.51-0.81] for Asian patients versus 0.83 [95% CI 0.66-1.05] for the ROW patients).[47] Clinical advice provided to the ERG highlighted that there could be a difference in fitness, screening and treatment approach between Asian and ROW patients (each of which may contribute to differences in outcome). Advice provided to the ERG suggested that patients from Asia tend to be treated more aggressively than their European counterparts, although it is expected that treatment pathways are broadly similar when comparing practices in Asian countries and the UK. Clinical experts also explained that obesity increases the risk of adenocarcinomas which reflects the UK population more than the Asian populations. The ERG considered the high proportion of patients from an Asian region was not reflective of the UK patient population, and noted with concern the impact this appears to have

on OS. The impact of region on OS could have been caused by several reasons, including those highlighted above, as well as which treatments patients receive after progression. The ERG requested that the company provide a scenario analysis removing the Asian region population, however the company did not provide this subgroup in the model stating that *"KEYNOTE-590 was not powered to detect differences by region…and feedback from clinicians MSD has consulted that there is no clinical rationale for the difference"* (see response to clarification question B4). Therefore, the ERG was unable to consider any further analysis for this subgroup.

The ERG also noted the histology in KEYNOTE-590[12] (26.8% adenocarcinoma versus 73.2% squamous cell carcinoma) (ITT population CS Table 6). Clinical advice provided to the ERG confirmed that in UK practice, the proportion of patients by histology would be expected to be approximately two-thirds being adenocarcinoma and one-third being squamous cell carcinoma (i.e., the opposite proportionate split versus the KEYNOTE-590 study). Clinicians also advised that histology is an important factor given the differences in disease and potential treatment (see Section 3.2.3.1).

4.2.4. Interventions and comparators

The company's model considered pembrolizumab in combination with platinum-based chemotherapy in line with the dosing schedule in KEYNOTE-590:

- Pembrolizumab administered intravenously at a fixed dose of 200 mg over 30 minutes every three weeks (Q3W) for up to 35 cycles.
- Cisplatin administered intravenously at a dose of 80 mg/m^2 Q3W for six doses.
- 5-FU administered as a continuous infusion on Days 1 to 5 at a dose of 800 mg/m^2/day (4,000 mg/m^2 in total) Q3W for up to 35 cycles.

The NICE scope identified the relevant platinum-based comparators as doublet treatment (fluorouracil or capecitabine plus cisplatin or oxaliplatin) and triplet treatment (fluorouracil or capecitabine plus cisplatin or oxaliplatin plus epirubicin). The company identified evidence to support the assumption of similar efficacy between doublet treatments and little benefit with the addition of epirubicin supported by data from the NICE Guideline in the assessment and management of oesophago-gastric cancer in adults (NG83)[7] and clinical opinion. This evidence was used to justify the use of the comparator arm from KEYNOTE-590[12] to inform the efficacy of

the chemotherapy arm in the model regardless of treatment regimen selected. The ERG would like to note that the evidence from NG83 presented by the company, reports a HR of 0.7 for the comparison of 5-FU+cisplatin with or without an anthracycline (e.g. epirubicin). Although the confidence intervals cross 1, this should not be used solely as evidence that there is no difference between treatment regimens when the mean is so far from no difference (i.e. 1). Clinician experts advised the ERG that if there is a benefit of triplets versus doublets, it would be small therefore assuming equivalent efficacy is reasonable and that in clinical practice epirubicin is generally dropped as there is concern over the increased toxicity with little benefit. Therefore, the ERG considers the company's assumption to be reasonable, however given the uncertainty, the ERG presented scenarios exploring the efficacy of triplet regimens based on the NMA's discussed in Section 3.5.2. Results of these scenarios are presented in Section 6.2.

The main comparator in the model reflects the KEYNOTE-590[12] trial comparator; cisplatin plus 5-FU with the dosing schedules as per the trial:

- Cisplatin administered intravenously at a dose of 80 mg/m^2 Q3W for six doses
- 5-FU administered as a continuous infusion on Days 1 to 5 at a dose of 800 mg/m^2/day (4,000 mg/m^2 in total) Q3W for up to 35 cycles

Clinical experts advised the ERG that these dosing schedules are slightly different to those commonly used in UK practice, with cisplatin usually given at a dose of 60 mg/m^2 for up to six to eight cycles. The five-day infusion of 5-FU is no longer considered the standard of care in UK clinical practice, and is now mainly replaced by capecitabine (a different fluoropyrimidine that is administered orally, and the body converts to 5-FU) or a two-day infusion of 5-FU (instead of five-day). However, the clinical experts confirmed that the efficacy of 5-FU would not be impacted by these dosing differences. As such, the ERG explored scenarios changing the dose to reflect UK practice (see Section 6.2).

The model also includes other regimens as a pairwise comparison versus pembrolizumab in combination with chemotherapy and as a blended comparison assuming equal market share in scenario analysis (see Table 11).

Table 11: Comparator treatments included in the company's economic model

Type	Platinum	Fluoropyrimidine	Other
Doublets	Cisplatin	5-FU	-
	Cisplatin	Capecitabine	-
	Oxaliplatin	Capecitabine	-
Triplets	Oxaliplatin	5-FU	Leucovorin (folinic acid)*
	Cisplatin	5-FU	Epirubicin (anthracycline)
	Oxaliplatin	5-FU	Epirubicin (anthracycline)
	Cisplatin	Capecitabine	Epirubicin (anthracycline)
	Oxaliplatin	Capecitabine	Epirubicin (anthracycline)

Key: 5-FU, 5-fluorouracil

Note: *The combination of oxaliplatin + 5-FU + leucovorin is also known as FOLFOX and is considered a doublet regimen

At clarification stage, the ERG requested further information from the company to justify the assumption of equal market share. The company claimed that their market share data *"lacked face validity versus the comparators outlined in the NICE Final Scope".* In addition, clinicians were unable to provide market share expectations at their advisory board, as such the company *"chose to include all therapies listed in the final NICE scope and distribute them evenly with respect to market shares"* (see company clarification response B16). The ERG was unable to validate the company's justification as no information on the market share data or clinical input from the advisory board was provided by the company within the submission or in response to clarification questions. The ERG found the company's approach of assuming equal market share inadequate to reflect UK practice. However, the ERG acknowledged that the company ran scenario analysis amending the comparator arm to each of the chemotherapy regimens individually to investigate the impact of comparator therapies. Based on the company's revised base case post clarification questions, the ICER ranged from £39,812 to £42,172.

Clinical advice provided to the ERG noted that not all of these treatments are used in UK practice and certainly do not have equal market shares. Capecitabine (administered orally) is used more than 5-FU as 5-FU is only used in the small number of patients who cannot tolerate tablets or who experience dysphagia. Doublet treatments are more common in UK practice but there is still a small usage of triplet regimens, mainly capecitabine plus oxaliplatin plus epirubicin (a combination also known as EOX). In addition, based on the results on the REAL-2[9] study, oxaliplatin should have largely replaced cisplatin in clinical practice given no decline in efficacy,

reduced toxicity and reduced infusion time, however the decision can also depend on other factors such as comorbidity in patients, histology and capacity of chemotherapy in the day unit. Thus, the ERG considered the trial comparator in KEYNOTE-590[12] was not the most relevant comparator for this decision problem. In addition, some of the other comparators included in the model were considered irrelevant. The ERG ran scenarios using a more clinically plausible distribution of market shares based on clinical expert opinion provided to the ERG. Results of this scenario are presented in Section 6.2.

4.2.5. Perspective, time horizon and discounting

The company discounted costs and outcomes (life-years [LYs] and QALYs) at 3.5% per annum and the model adopted an NHS and PSS perspective. The ERG was satisfied that the perspective and discounting adopted by the company's model are aligned with the NICE reference case.

The model included half cycle correction in their base case; however, the ERG considered this was unnecessary given that the cycle length was only seven days.

A time horizon of 20 years was used to inform the company's base case to reflect a lifetime horizon as specified in the NICE reference case. However, the ERG noted that in the pembrolizumab in combination with chemotherapy arm at 20 years [redacted] of patients were estimated to still be alive using the company's base case survival settings (CS, Doc B, Table 46). The ERG requested justification for this time point at clarification stage. The company stated that their choice of 20 years was *"informed by the current estimates of survival of patients treated within UK clinical practice*", though subsequently amended their time horizon to 30 years in their revised base case "*to fully capture costs and benefits of pembrolizumab in combination with chemotherapy"*. The ERG noted that at this timepoint <1% of patients were estimated to be alive at this timepoint and therefore considered the change appropriate based on the company's base case choice of OS modelling.

4.2.6. Treatment effectiveness and extrapolation

Data from the KEYNOTE-590[12] trial constituted the primary evidence base from which estimates of treatment effectiveness are made to inform the economic model. In terms of treatment effectiveness, two outcomes from the KEYNOTE-590[12] trial are used to inform the model: OS

and PFS. Data from KEYNOTE-590[12] concerning the estimated duration of treatment are discussed separately within Section 4.2.8.3.

For clarity, the descriptions of the time-to-event outcomes used to inform the model are provided below:

- **OS:** the proportion of patients who were alive at each model cycle, regardless of disease progression status. This was calculated as the time from randomisation until the last known date of survival
- **PFS:** the proportion of patients who were alive with non-progressed disease at each model cycle. The proportion of patients with progression-free disease was less than or equal to the proportion of patients alive at each model cycle. Therefore, any extrapolations of PFS were not permitted to "cross" the OS curve

For both OS and PFS, survival modelling methods were used to extrapolate over the lifetime horizon of the model, given that the follow-up period for data collected in the KEYNOTE-590[12] trial was shorter than the time horizon of the economic model (20 years in the company's original base-case analysis, 30 years in the company's revised base-case analysis, and up to a maximum of 33.6 months in the KEYNOTE-590 trial). The CS explained that NICE DSU TSD 14[48] guidance was followed in determining the most suitable survival extrapolations to inform the model.

The model also included the cost and utility implications associated with the occurrence of adverse events. The included adverse events are highlighted in this section, with the impacts on utility and costs discussed later in Sections 4.2.7.3 and 4.2.8, respectively.

4.2.6.1. Overall survival

As described in NICE DSU TSD 14, assuming patient-level data are available for analysis, a comparison of suitable plots should be undertaken to allow initial selection of appropriate models.[48] In the CS however, only cumulative hazard plots and log-cumulative hazard plots (LCHPs) were presented (CS Figures 13 and 14, for the pembrolizumab in combination with chemotherapy versus SoC arms, respectively). These plots allowed for an assessment of whether the proportional hazards (PH) assumption holds, and the potential suitability of PH models, such as the exponential, Gompertz, or Weibull model. However, these plots did not allow for an assessment of whether other types of model are potentially suitable – for example,

a quantile-quantile (Q-Q) plot would allow an assessment of whether a jointly-fitted accelerated failure time (AFT) model would be suitable (such as a lognormal or log-logistic model, with a covariate for treatment arm).

The company concluded in its submission that the PH assumption does not hold (on the basis of non-parallel lines seen in the LCHP, CS Figure 14), and so a joint parametric model fitted within a PH framework (e.g., with a covariate for treatment arm) was deemed inappropriate. In response to a further clarification question concerning the LCHP (clarification question B6), the company provided further justification:

- NICE DSU TSD 14 guidance[48] states: *"Generally, when patient-level data are available, it is unnecessary to rely upon the proportional hazards assumption and apply a proportional hazards modelling approach"*. Also mentioned is that when individual-level patient data are available, fewer assumptions are required when fitting separate versus jointly-fitted parametric models. The company explained that both aspects of guidance presented in TSD 14 apply here, suggesting a jointly-fitted model would be less appropriate than separate models.

- The mechanism of action of pembrolizumab (an immune checkpoint inhibitor) given in combination with chemotherapy is purported to differ substantially from that of chemotherapy given alone. Accordingly, the company considered a modelling approach wherein a 'two-dimensional' treatment effect (i.e., an impact on both shape and scale parameters) to be more appropriate when considering alternative modelling approaches.

The ERG agreed that a jointly-fitted, single, PH parametric model is unlikely to provide a good fit to the KEYNOTE-590[12] trial data. However, the ERG disagreed with the company's decision to reject *all* jointly-fitted models on the basis of inspecting only cumulative hazard-based plots. Other types of jointly-fitted models, such as a model that assumes a constant time ratio (i.e., a jointly-fitted AFT model), may be appropriate and this possibility requires further exploration.

At clarification stage, a Q-Q plot was requested, as well as several other plots to further explore the suitability of different parametric models to estimate OS. The Q-Q plot was provided (see company response to clarification question B5), which is re-produced in Figure 5. To justify the use of a jointly-fitted AFT model, the Q-Q plot should show a straight line extending from the origin (shown in Figure 5 as the red dashed line).

Figure 5: Q-Q plot of overall survival from KEYNOTE-590

Key: SOC, standard of care.
Source: Image re-produced from company's response to clarification question B4

In response to clarification question B4, the company states that the Q-Q plot for OS suggests *"that the observed data was bending away from the straight line (slope became smaller over time)"* and that this *"suggests that the hazards of death for the pembrolizumab + chemotherapy arm was decreasing faster than the SOC arm and the trend cannot be captured by an AFT model"*. The ERG acknowledges that the Q-Q plot does not demonstrate an 'perfect' straight line extending from the origin but would not reject the use of a jointly-fitted AFT model on the basis of this Q-Q plot alone, as the plot does not show a clear violation of a constant time ratio (in the view of the ERG). Furthermore, as the number of patients at risk decreases substantially over time, the robustness of the Q-Q plot towards the end of follow-up is especially uncertain.

The CS explained that given the availability of patient-level data from the KEYNOTE-590[12] trial, the assessment of the LCHP (CS Figure 14), and the different mechanistic properties of pembrolizumab in combination with chemotherapy versus chemotherapy, separately-fitted parametric models were preferred over jointly-fitted models. The ERG agreed that separate models required fewer assumptions but noted that these models required additional parameters to be estimated. By specifying separate models, a multi-dimensional effect of pembrolizumab is

implicitly modelled (which is especially important to consider if different model distributions are selected). Nevertheless, considering the visual fit of the independent models (discussed later in this section), the ERG considered it appropriate to exclude jointly-fitted models (both PH and AFT) for the outcome of OS.

Following inspection of KEYNOTE-590[12] trial data, two different modelling approaches were implemented within the economic model for the outcome of OS:

- **Fully-fitted modelling approach (henceforth termed "single parametric model"):** A single parametric model was fitted to the OS data from KEYNOTE-590 (separately for each treatment arm), from $time = 0\ weeks$
  - Six parametric models were considered: exponential, Weibull, Gompertz, lognormal, log-logistic, and generalised gamma
- **Piecewise models:** The Kaplan–Meier estimate of OS was used to directly inform OS within the economic model outcome up until a given cut-off point, after which the remainder of the OS was informed by a parametric model fitted to a reduced set of data from KEYNOTE-590
  - Like the single parametric models, these models were fitted independently by treatment arm, and the same six models were considered per the single model approach
  - Events and censored observations before the cut-off point were not included in the parametric component (i.e., the parametric models were fitted to 're-based' data, where $time_{re-based} = time_{original} - "cut-off"\ point$)

At clarification stage, the ERG asked the company to explain why other alternative modelling approaches were not explored (clarification question B11). In response, the company explained that diagnostic plots for the hazard function (LCHP, smoothed hazard plots, etc.) showed *"a relatively simple trend"*, and that given the maturity of the data from KEYNOTE-590, piecewise models were capable of providing reasonable estimates (company response to clarification question B11). The company went on to explain that other flexible methods require *"additional assumptions"*, highlighting that when fitting spline-based models it is necessary to specify the number and location of knots, which can have an important effect on the resultant extrapolation.

The ERG observed that other possible methods may have also yielded reasonable extrapolations. For example, spline-based models have been used in a range of previous NICE appraisals of immune-checkpoint inhibitor treatments. The ERG noted in particular that while spline-based models require certain assumptions related to the number of knots and their location(s), piecewise models also require similar assumptions (for example, the number and location of cut-off points). Nevertheless, the ERG was satisfied that the range of models provided by the company within its submission was sufficient to inform decision making.

For the piecewise models, it was necessary to specify where the cut-off point should be imposed. To select a cut-off point, the CS explains that Chow tests were conducted to *"identify structural changes"* where *"higher Chow test statistics indicating a higher likelihood of structural change"* (CS Section B.3.3). Plots of the Chow test statistics based on a range of different cut-off points are presented in CS Figure 15 for the pembrolizumab in combination with chemotherapy arm, but a corresponding plot for the chemotherapy arm was not provided, as *"the Chow test statistics for the SOC arm proved inconclusive for determining an appropriate cut-off"* (CS Section B.3.3). At clarification stage, a plot of the Chow test statistics for the chemotherapy arm was requested and provided (see company's response to clarification question B7).

Based on the Chow test statistics for the pembrolizumab in combination with chemotherapy arm, the company selected a base-case cut-off point of Week 40, with an alternative cut-off point of Week 32 explored within scenario analysis. The same cut-off point was selected for both treatment arms. The ERG noted that a Chow test can be used to assess whether a single structural break occurs at a given time point, assuming that the time point is known. For example, the test was illustrated by Chow[49] via an example to explore the demand for automobiles in the United States, and if there was evidence this changed 1922 to 1953 and 1954 to 1957. However, it was the ERG's understanding that the Chow test was not designed to detect the timepoint at which a structural break may occur. In the example presented by Chow, the timepoint was selected based on when data were reported and was not chosen following inspection of Chow test statistics. Accordingly, the ERG did not consider it statistically sound to choose a cut point based on the Chow test statistics alone.

Visual and statistical goodness of fit scores (Akaike's and Bayesian information criterion [AIC and BIC, respectively]), were used by the company to inform its determination of the best-fitting estimate of OS (focusing on fit to the Kaplan-Meier estimate). Based on AIC and BIC scores,

the log-logistic model was highlighted as the best fitting for the pembrolizumab in combination with chemotherapy arm, whereas a lognormal was the best-fitting model for the chemotherapy arm. The company also highlighted that the Gompertz model led to *"clinically-implausible"* outcomes (where the hazard of death after the end of follow-up approached zero, leading to an indefinite plateau in the OS extrapolation, see CS Figure 16).

To aid with model selection, the company undertook a targeted search of the published literature to identify studies that reported longer-term OS estimates in an advanced and metastatic oesophageal cancer population. Three studies were identified, which are summarised below:

- **Gavin et al. (2012):**[50] A European study based on data collected from 66 cancer registries, with patient five-year survival rate of 3.8% for patients with distant stage oesophageal cancer. The ERG highlights that the reported value of 3.8% is an estimate of *relative* survival. In this example, relative survival should be interpreted as the ratio of the proportion of observed survivors in the Gavin et al. (2012) study to the proportion of expected survivors in a relatively healthy population. A five-year relative survival of 3.8% implies that the five-year OS is lower than 3.8% (when taking into account death from other causes), though five-year OS is not reported in the Gavin et al. 2012[50] study)
- **Surveillance, Epidemiology, and End Results (SEER) database:** Several different studies were cited, all based on data collected from the SEER database.

  The American Cancer Society[51] reports a five-year relative survival rate for patients with distant oesophageal cancer of 5%, based on people diagnosed with oesophageal cancer between 2010 and 2016. Again, the ERG highlights that this is an estimate of *relative* survival, and so the true estimate of 5-year OS is expected to be less than 5% based on this study.

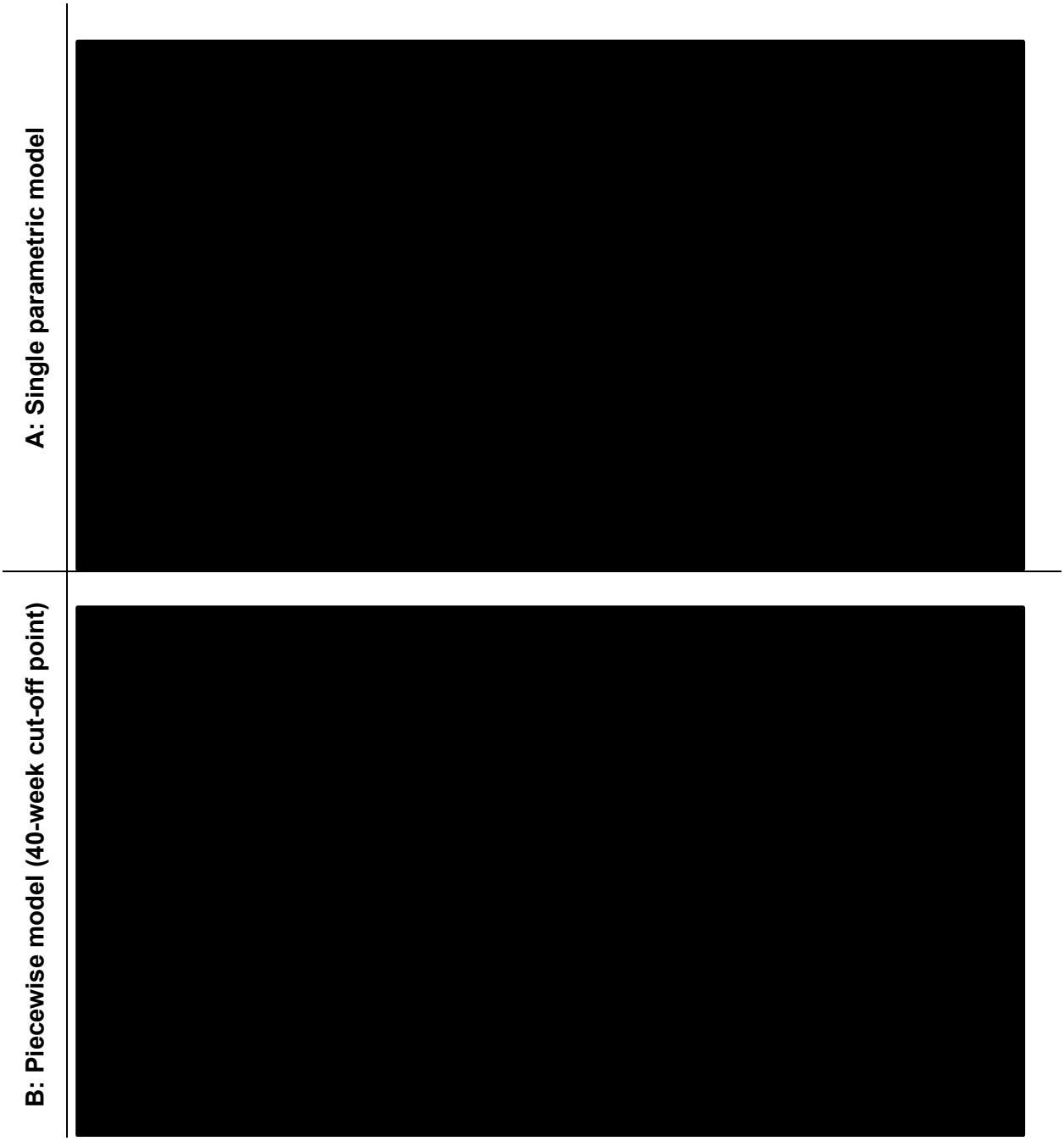
  - In 2016, Wu et al. (2016)[52] reported five- and 10-year OS rates of 5.4% and 3.5%, respectively, for patients with metastatic oesophageal cancer. While not stated in the CS, it should be noted that median OS for this population was 10 months, and that one- and two-year OS rates were 40.5% and 14.6%, respectively. In KEYNOTE-590,[12] median OS for the chemotherapy arm was [redacted] months, with one- and two-year OS rates of [redacted] and [redacted], respectively, demonstrating that the estimates reported by Wu *et al.* [redacted].

  - Another study by Wu et al.[53] published in 2017 demonstrated median OS estimates of between five and six months, depending on whether patients had a single site of distant metastasis, or multiple sites of distant metastasis. However, five-year OS was not possible to robustly estimate from this study. The ERG speculated that median OS is lower in the later study by Wu et al.[53] (published in 2017) versus the earlier study by Wu et al.[52] (published in 2016) due to the latter study likely including more patients with multiple sites of distant metastasis, but this is unclear.

- **Tanaka et al. (2010):**[54] A single-centre, Japanese study of n=80 patients with oesophageal squamous cell cancer and distant organ metastasis. The median OS was 6.4 months, with one- and two-year OS rates of 23.7% and 11.2%, respectively. Five-year OS was not reported, but from the Tanaka *et al.* paper, it can be inferred through inspection of the Kaplan-Meier estimates (Fig 1 in the paper) that five-year OS is likely less than 5% (as the end of the Kaplan-Meier estimate evaluated at approximately five years suggests around 1–2% of patients were still alive, and the OS estimate falls to less than 5% at around three years). The ERG noted, however, that this is a purely squamous cell carcinoma population, from a single centre in Japan. Therefore, this study was unlikely to serve as a useful validation source for the KEYNOTE-590[12] trial (which included adenocarcinoma patients, and patients from outside of Asia)

At clarification stage, the ERG requested that the company provided additional plots to help with selecting a given model for the outcome of OS (clarification question B5 part c). The company provided a range of additional plots, including smoothed hazard plots (company response to clarification question B5 part c). Figures 13 and 14 of the company's response to clarification questions present a comparison of the fitted models to the estimated smoothed hazard functions for the pembrolizumab in combination with chemotherapy arm. These plots are presented together in Figure 6.

Figure 6: Smoothed hazard plots: single versus piecewise models



Key: OS, overall survival; SOC, standard of care.

Source: Figures re-produced from company response to clarification question B5.

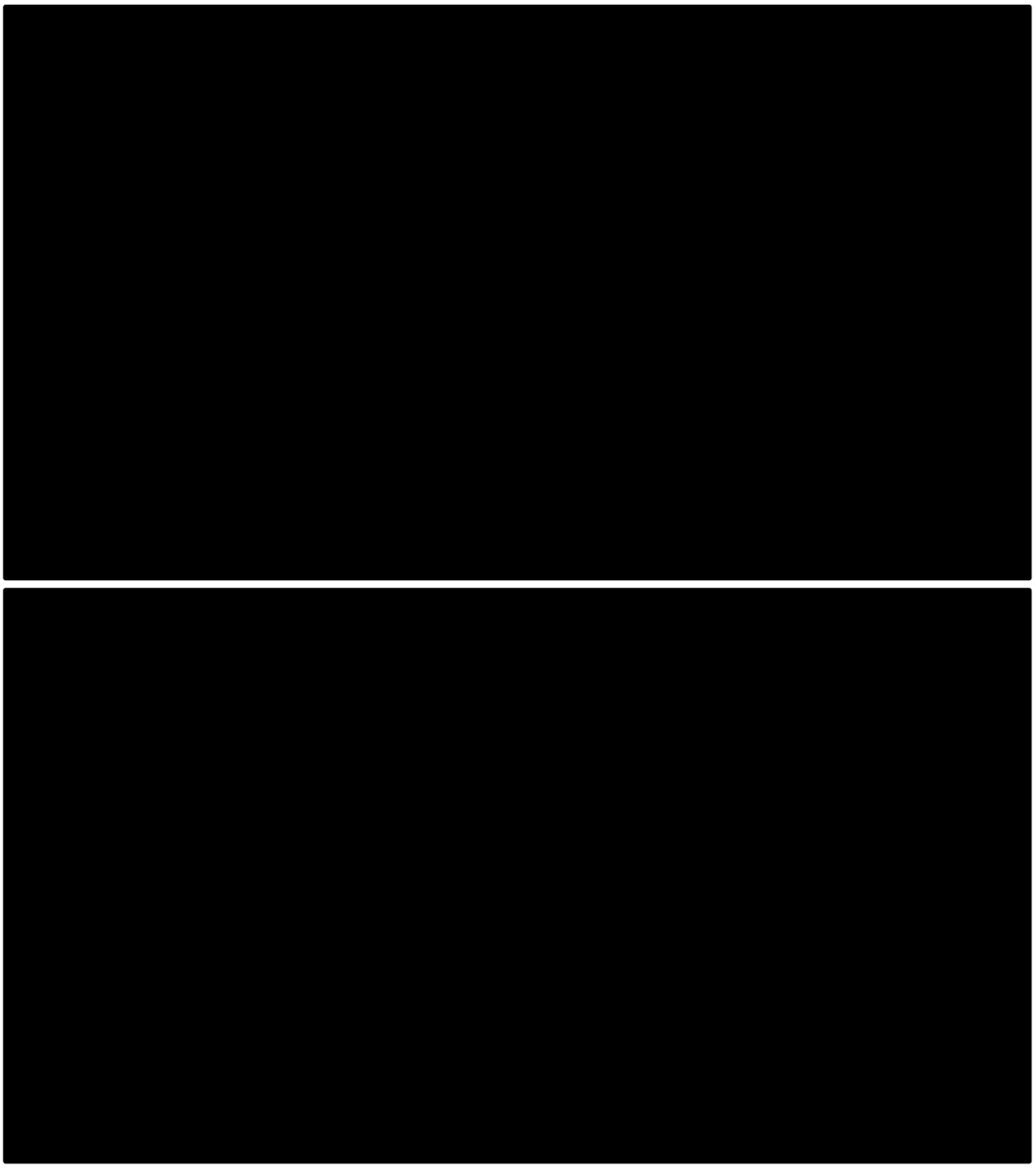
The plots for the single parametric models (Figure 6, panel A) illustrate that none of the seven[1] models were fully capable of reflecting the shape of the hazard function – that is, a gradual increase in the estimated hazard of death until approximately 40 weeks, followed by a consistent decline. As noted by the company, each of the models appear to overestimate the hazard of death from approximately 75 weeks onwards. The plots for the piecewise models with a 40-week cut-off point (Figure 6, panel B) appear to provide estimated hazards that lay closer to the smoothed hazard estimate. However, from these plots, it cannot be readily determined how the estimated hazard of death in the longer-term may differ, and how these may compare to the age- and sex-adjusted general population.

The company selected a piecewise Kaplan-Meier + log-logistic model for the outcome of OS for both treatment arms. These models were selected owing to their statistical goodness-of-fit scores (with the models providing either the best, or second-best fit according to AIC and BIC, see CS Table 44), their visual fit, as well as input from clinical experts concerning the expectation of a percentage of patients deriving a *"long-term survival benefit from treatment with pembrolizumab in combination with chemotherapy"* (CS Section B.3.3). For the chemotherapy arm, the company noted that the log-logistic model produced five- and 10-year OS estimates of [redacted] and [redacted] respectively (CS Table 45); and that these estimates were broadly in keeping with the external studies identified via a targeted review of the literature. However, the CS did not describe the plausibility of using the same model for the chemotherapy arm, given that this regimen does not contain pembrolizumab.

The company's base-case projections of OS for both arms are provided in Figure 7. At the end of the modelled time horizon (20 years), the company's base-case analysis predicts that [redacted] of the chemotherapy arm, and [redacted] of the pembrolizumab in combination with chemotherapy arm, are still alive (CS Tables 45 and 46).

[1] The ERG notes that an additional seventh model is presented in this plot – a 'gamma' model. The ERG understands this to be a two-parameter Gamma model, though this is not presented in the company's submitted economic model and is therefore not discussed further.

Figure 7: Company projections of overall survival (5- and 20-year time horizon)



Key: KM, Kaplan-Meier; SOC, standard of care.

Note(s): This figure is a re-formatted version of CS Figure 19. Figure re-produced for ease of interpretation related to time axis. Company base-case analysis includes cut-off point at week 40 (switch from KM to log-logistic model).

Source(s): Produced based on information provided in the company-submitted economic model.

The ERG considered the company's choice of OS models to inform its preferred base-case analysis to be broadly appropriate – the models provide a reasonable fit to the Kaplan-Meier estimate (helped in part by setting OS to be equal to the Kaplan-Meier estimate until 40 weeks) and appear to provide plausible estimates to inform the cost-effectiveness analysis (see CS Section B.3.3 for company's full rationale).

However, given the uncertainty in the extrapolated tail and the limited information available concerning long-term outcomes in this patient population, a range of alternative extrapolations may be suitable to aid decision making. To provide a range of plausible extrapolation options, the ERG has focused on four scenarios:

1. **Kaplan-Meier + log-logistic tail:** This is the company's base-case estimate of OS, used for both treatment arms. The Kaplan-Meier estimate of OS is applied up until 40 weeks, after which a log-logistic model is used to extrapolate OS for the remainder of the model time horizon
2. **Kaplan-Meier + log-logistic tail + assume treatment waning effect applies linearly between 5 and 7 years:** This approach adjusts the company's base-case estimate of OS by assuming the projection of OS for the pembrolizumab in combination with chemotherapy arm applies up until 5 years, at which point the projected hazard of death gradually approaches that of the chemotherapy arm until 7 years, after which the projected hazards are identical across both arms (equal to the projection for the chemotherapy arm)
    - The ERG highlighted here that while this is termed a "treatment waning effect" within the context of the model, the ERG has used this functionality purely within the interest of exploring how influential the projected tail is on the estimated ICER
3. **Single log-logistic parametric model:** This approach constitutes the best-fitting of the single parametric models, with the second-most optimistic estimate of OS for the pembrolizumab in combination with chemotherapy arm (most optimistic was the lognormal model)
    - The ERG considered this approach important to demonstrate the impact on estimated OS by considering a piecewise approach, per the company's base-case analysis

4. **Kaplan-Meier + generalised gamma tail:** This approach selects a less optimistic/ more pessimistic extrapolation for consideration after 40 weeks versus the company's base-case analysis

    - The generalised gamma model was selected as it represented a mid-range estimate of OS, in consideration of the range presented in the company's model (e.g., five-year OS estimates for the pembrolizumab in combination with chemotherapy arm were [redacted], with the generalised gamma model providing an estimate of [redacted] [see CS Table 46])

The results from these four scenarios are discussed further in Section 6.2.1.

The CS explained that clinical expert opinion was obtained in selecting the base-case approach to modelling OS (CS Section B.3.3). At the clarification stage, the ERG asked for further information to be provided about the process taken to elicit expert opinion (clarification question B23). The company explained that informal interviews were held with four clinical oncologists, separately, working in the treatment of oesophageal cancer. However, additional information was not provided, such as which questions were asked, and if there were any disagreements between experts.

4.2.6.2. Progression-free survival

Unlike OS, for the outcome of PFS diagnostic plots were not provided, and only piecewise models were presented in the original CS. The company explained that PFS data from the KEYNOTE-590[12] trial were relatively complete, with over 90% of patients having reached the PFS endpoint (CS Section B.3.3). The ERG interpreted this statement to be a comment on the proportion of patients still at risk for a PFS event at the end of the Kaplan-Meier estimate, as based on the KEYNOTE-590 CSR, [redacted] of the pembrolizumab in combination with chemotherapy arm, and [redacted] of the chemotherapy arm were recorded with a PFS event by the end of follow up ([redacted]).[47] Nevertheless, the ERG agreed with the statement made by the company that the PFS data are near complete.

The company also described how protocol-scheduled tumour imaging assessment scans had an impact on PFS outcomes. Based on the KEYNOTE-590 trial protocol, the first planned scan was scheduled to take place between Weeks 8 and 10 (at Week 9, ± 1 week either side of this time point). In the Kaplan-Meier estimate of PFS (CS Figure 6), drops in the PFS curve can be

seen at this assessment point, as well as later assessments in nine-weekly intervals. Figure 8 shows an overlay of the Kaplan-Meier estimates of PFS and these nine-weekly scans.

Figure 8: Visualisation of imaging assessments versus drops in PFS curve

Key: KM, Kaplan-Meier; SOC, standard of care.

Note(s): Imaging assessments shown every nine weeks, with a width of two weeks (i.e., ± 1 week either side).

Source(s): Produced based on information provided in the company-submitted economic model.

Based on the ERG's understanding of the KEYNOTE-590[12] study, the Kaplan-Meier estimates shown in Figure 8 may be interpreted as optimistic estimates of PFS for both treatment arms. This is because patients may have progressed prior to a scan, but were only recorded as having progressed at the time the scan was conducted. This feature of the trial is not unique to KEYNOTE-590,[12] but has important implications for the interpretation of the results of the study, and how to most appropriate inform the economic model.

A further consideration of the company's use of PFS to inform its model is the range of censoring rules for PFS within the KEYNOTE-590[12] study (i.e., allocation of individual observations as events or censors). At clarification stage, the company explained the differences between the censoring rules used in the primary analysis, and two sensitivity analyses (company responses to clarification questions A23 and B10). The reasons for the different analyses were related to missed doses and/or initiation of a new anticancer treatment.

The company's model made use of the primary analysis censoring rule. In this analysis, patients were considered to have a PFS event if a progression or death event occurred after either 0 or 1 missed disease assessment, and before new anti-cancer therapy (if applicable). However, patients were censored if either (i) a death or progression event occurred after ≥2 consecutive missed disease assessments without further valid non-PD disease assessments, or after new anti-cancer therapy; (ii) no death or progression event occurred, and new anticancer treatment was not initiated; or (iii) no death or progression event occurred, and new anticancer treatment was initiated.

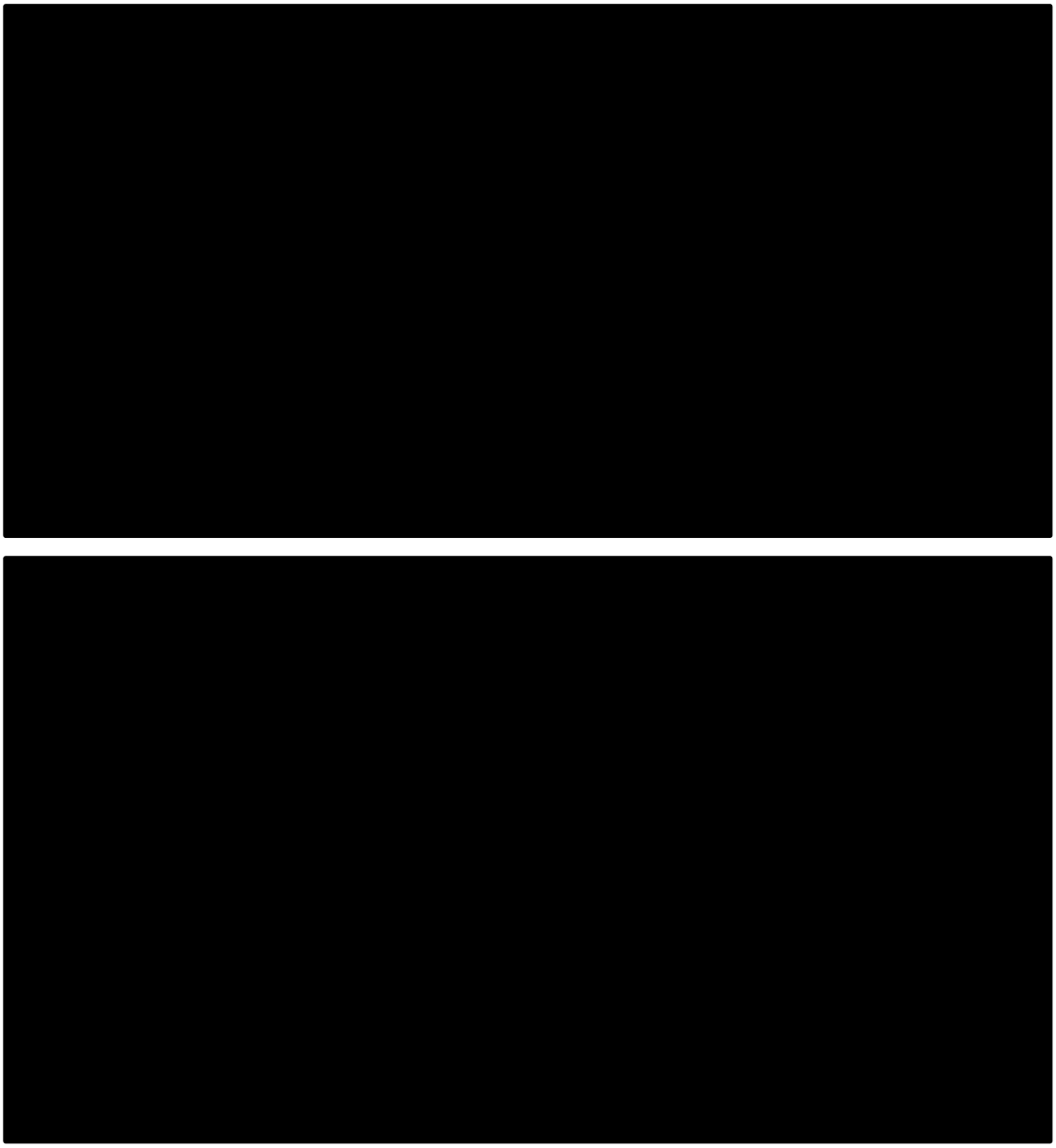
In sensitivity analysis 1, the company considered the first of the three censoring reasons (an event after ≥2 consecutive missed disease assessments without further valid non-PD disease assessments, or after new anti-cancer therapy) to be an event instead of a censored observation. In sensitivity analysis 2, the company considered discontinuation of treatment due to reasons other than complete response or initiation of new anticancer treatment, whichever occurs later, to be an event for subjects without documented PD or death. The analyses provided by the company showed little difference in the estimates of PFS by censoring rule, and so the ERG is satisfied with the use of the primary analysis to inform the model.

The company explained in its submission that through visual inspection of the Kaplan-Meier estimates, a *"steep drop"* is seen at around Week 9 for both arms, which is the time at which the first scan is performed (CS Section B.3.3). The company opted for a piecewise model using a cut-off point of 10 weeks – the time after which all patients are expected to have had their first scan. No explanation is provided in the CS concerning the choice to deviate from a single parametric model, and only consider piecewise models. A scenario analysis is presented in the CS using an alternative cut-off point of 37 weeks, but no explicit rationale for the choice of this alternative cut-off point was provided in the CS.

In choosing a model for PFS, the CS states guidance from NICE DSU TSD 14[48] was followed (CS Section B.3.3). The log-logistic model was determined to provide the best statistical goodness-of-fit (lowest AIC and BIC scores, see CS Table 47 for scores), with the second best-fitting model being the generalised gamma. Owing to PFS being near complete, the company focused mostly on the fit of the models within the observed follow-up period, versus plausibility of long-term extrapolation. Therefore, the piecewise log-logistic model was selected by the company for both arms, given that it provided the best statistical goodness of fit. This is the same modelling approach as used for OS, with the exception that the cut-off point was set at 10

weeks, instead of 40 weeks (as was used for OS). The company's base-case projections of PFS for both arms are provided in Figure 9.

Figure 9: Company projections of progression-free survival (5- and 20-year time horizon)



Key: KM, Kaplan-Meier; SOC, standard of care.

Note(s): This figure is a re-formatted version of CS Figure 22. Figure re-produced for ease of interpretation related to time axis. Company base-case analysis includes cut-off point at week 10 (switch from KM to log-logistic model).

Source(s): Produced based on information provided in the company-submitted economic model.

While the ERG did not consider the company's base-case choice of models for PFS to provide a particularly poor fit to the Kaplan-Meier estimates from the KEYNOTE-590[12] trial, alternative estimates models for PFS were explored for completeness, especially when considering progression-based instead of time-to-death-based utility values (see Section 4.2.7). The ERG highlighted that in the company's base-case analysis, 10-year PFS for the pembrolizumab in combination with chemotherapy arm was <[redacted], yet [redacted] of patients are still projected to be alive at this time. This means that approximately [redacted] of patients alive at 10 years were assumed to have progressed disease, which the ERG considered to have questionable face validity in light of the limited treatment options available for patients that progress following treatment with pembrolizumab in combination with chemotherapy. The results from the PFS scenarios are discussed further in Section 6.2.2.

4.2.6.3. Adverse events

Adverse events were included in the model at Grade 3+ if they occurred in at least 5% of patients on either treatment arm. The observed incidence of AEs in KEYNOTE-590[12] by treatment arm was provided by the company within its submission (see Table 48 of the CS). The most common Grade 3+ AEs (occurring in either treatment arm) were anaemia, neutropenia, hyponatraemia, and decrease neutrophil count. The ERG highlighted that AE rates were similar between arms.

At clarification stage, the company provided scenario analyses exploring the use of alternative AE rates for the blended comparator (clarification question B12). The scenarios led to a small increase in the ICER, driven by fewer AEs resulting in more QALYs at a reduced cost for the comparator arm. The ERG agreed with the company that the published studies suffer from a number of limitations and prefers the use of the KEYNOTE-590[12] derived AE rates to inform the base-case analysis.

4.2.7. Health-related quality of life

4.2.7.1. Summary of data available from KEYNOTE-590

The KEYNOTE-590[12] study collected information concerning patient-reported health-related quality of life (HRQoL) via the EQ-5D-5L questionnaire. The -5L version of the EQ-5D allows respondents to describe each dimension using five different levels: no problems, slight problems, moderate problems, severe problems and extreme problems.[20] However, as of October 2019, NICE no longer recommends the use of utility values generated using published

EQ-5D-5L value sets[55]; instead, NICE recommends use of a mapping function developed by van Hout et al. (2012)[42] for reference-case analyses, to produce corresponding -3L version utility values. The company has used the mapping function by van Hout et al. (2012)[42] to inform its economic model. For brevity, the remainder of this section describes the EQ-5D-5L data collected from the study and the EQ-5D-3L values produced using the van Hout et al. (2012)[42] mapping function, both as EQ-5D.

In KEYNOTE-590, the EQ-5D questionnaire was administered to patients at the following time points:

- Day 1 of each treatment cycle, for Cycle 1 (baseline) to Cycle 9 (Week 24)
- Day 1 of every third treatment cycle, from Cycle 10 (Week 33) to either one year (Week 51) or the end of treatment, whichever was first
- At the time of discontinuation (including time points after one year)
- 30 days after discontinuation of treatment at a follow-up visit.

Per the KEYNOTE-590 trial protocol, a visit window of ±7 days either side of the planned assessment times was permitted for the EQ-5D assessments and factored into the analyses of EQ-5D data. The CS stated that analyses of EQ-5D utilities were based on the Full Analysis Set (FAS) population from KEYNOTE-590 (a total of n=713 participants). In order to be included within the FAS population, patients were required to have been randomised, receive a study treatment, and completed at least one EQ-5D questionnaire. At clarification stage, the ERG asked the company to clarify if the completion of at least one questionnaire was in addition to a baseline measure, or if a baseline measure alone was sufficient for inclusion (clarification question B13). In response, the company confirmed that a total of n=40 patients included within the utility analysis only had a baseline EQ-5D measure reported.

The ERG also asked whether baseline utility was accounted for within the utility analyses conducted (clarification question B14). The company stated in response: *"Baseline utility has been accounted for within the mixed linear effects model for both approaches."*. However, as further information was not provided, it was unclear to the ERG exactly how baseline utility was accounted for. More specifically, the ERG was not able to verify if baseline utilities were

included within the data set analysed, or if post-baseline utility were modelled controlling for baseline utility at the patient level.

The completion rate of the EQ-5D throughout the KEYNOTE-590[12] trial is provided in CS Table 52. At baseline, [redacted] of patients completed the EQ-5D questionnaire, and compliance was consistently greater than [redacted] up until one year after initiation of treatment (range: [redacted]). However, after one year, compliance reduced markedly, and by Week 69 only [redacted] completed the EQ-5D questionnaire out of [redacted] participants that were expected to complete it (equivalent to an overall compliance level of [redacted], CS Table 52).

For each of the utility approaches, mean EQ-5D utility scores by health status were estimated per treatment arm (pembrolizumab in combination with chemotherapy and chemotherapy arms), and pooled for both arms. In addition, 95% CIs were obtained for each estimated EQ-5D utility and the statistical significance of the differences between treatment arms was tested. The level of EQ-5D compliance through time is presented in Table 52.

4.2.7.2. Analysis of data from KEYNOTE-590

Two different approaches to analysing the EQ-5D data from the KEYNOTE-590[12] trial were considered within the company's economic model. The first of these was to calculate utility values for patients according to their progression status, and apply these within the model as health state utility values (a '**progression**' approach). The second was to instead calculate utilities according to the hypothetical time until death associated with each measure of utility, and group these to demonstrate the deterioration in utility expected as patients experience disease progression and get closer to death (a '**time-to-death**' approach). Both approaches are discussed in turn below.

Progression approach

In the progression approach, health states were defined according to disease progression status (i.e., the presence or absence of progression). The company highlighted that a limitation of the progression approach is that due to the distribution and collection of EQ-5D questionnaires within KEYNOTE-590,[12] it is challenging to calculate an average utility value for the progressed disease health state. This is because the utility for patients with newly-progressed disease is expected to be higher than the utility those who progressed earlier,

ceteris paribus. Accordingly, the utility values estimated for the progressed disease state are expected to be greater than the values for the "true" progressed disease state were EQ-5D data collected for progressed patients at time points after the final Day 30 post discontinuation visit (assuming patients discontinued due to progression). For patients who discontinued for reasons other than progression, per the KEYNOTE-590 trial protocol, and did not progress within 30 days following discontinuation, no value for the progressed disease state would be recorded.

To address the concerns highlighted with the progressed disease state, it may be important to consider external data sources. However, as acknowledged by the company, the NICE reference case states a preference for using utility data collected from the relevant clinical trial(s) to inform the model, where available. In addition, use of external data sources requires the availability of relevant sources that would be suitable for consideration within the economic model. From the company's SLR, nine potentially relevant studies were identified (CS Appendix H). However, in its submission, the company explains that *"due to the paucity of data within this disease area, it is not possible to substitute utility values from the literature to alleviate this issue"* (CS Section B.3.4).

A progression approach has often, but not always, been adopted within models for a range of previous evaluations of cancer therapies, particularly those that make use of a three-state PartSA model (as has been used by the company). Therefore, in spite of its limitations, utility values were generated using a progression approach and included as an option within the economic model. Utility observations were separated by progression status using the date of progression recorded in the KEYNOTE-590 trial.[12]

The company used a linear mixed-effects regression model to estimate average utility values for each health state. The model included covariates for progression as well as the presence/ absence of any Grade 3+ AEs. A covariate for treatment arm was not included within the mixed-effects regression model, as utility values were assumed to be the same across all treatment arms.

In the economic model, the utility values by progression status were applied to the proportion of patients expected to reside within each health state over the course of the model time horizon. However, to account for the fact that utility was expected to decrease over time as patients aged, the economic model also includes a multiplier to account for patients aging over the

course of the model time horizon. The age adjustment is based on a study by Ara and Brazier et al. (2010).[56]

The AE-related disutility was used to generate a total QALY loss associated with each treatment arm based on the calculated AE rates and durations over which the disutility was expected to apply. The calculation of AE rates and expected durations over which utility decrements are assumed to apply is described in Section 4.2.7.3. The total QALY loss attributable to AEs was applied within the model as a lump sum in the first model cycle.

Time-to-death approach

In the time-to-death approach, utility values are generated using groupings of utility observations based on how close they were reported versus the patient's OS time. In the CS, a number of applications of a time-to-death approach are cited, including a number of previous NICE technology appraisals of treatments for non-small-cell lung cancer and melanoma.

The CS stated that a time-to-death approach more accurately captures the decrease in HRQoL over time (versus standard progression-based utilities) for patients with advanced oesophageal cancer, but does not provide a source for this justification. A study by Hatswell *et al*. (2014)[57] was highlighted within the CS, which considered an analysis of utility values in melanoma. The authors of this study commented that disease progression alone may not fully capture *"all predictive factors of patient utility"*, and that time-to-death may, as death approaches, be *"as or more important"* than progression status in predicting utility.

In order to consider a time-to-death analysis, patients with a censored OS time which is within the time period of the grouping furthest from death cannot be assigned to a group. In the company's time-to-death analysis, the upper bound was set at 360 days. Therefore, utility values recorded for patients that were censored for the outcome of OS at a time point within 360 days until this censoring time could not be included in the time-to-death analysis. However, in its submission, the company explained that approximately [redacted] of all utility values captured could not be appropriately assigned to a time-to-death category, meaning that [redacted] of recorded values were not affected by this limitation.

In the company's time-to-death analysis, a linear mixed-effects regression model was fitted, with five different groupings. The model included a covariate for the presence/ absence of any Grade

3+ AEs (as was also captured in the progression-based analysis). The groupings used for time-to-death categories were as follows:

- <30 days until death
- 30≤ days until death <90
- 90≤ days until death <180
- 180≤ days until death <360
- ≥360 days until death

The rationale behind the choice of these groupings was not provided within the CS. At clarification stage, the ERG asked for further information concerning these groupings (clarification question B14). The company explained that the groupings were consistent with several previous NICE assessments of pembrolizumab in other indications, including lung and renal cancer. The ERG would have preferred the groupings to have been informed through further inspection of the utility data available from the KEYNOTE-590[12] study.

In addition, it is unclear why these groupings were not aligned with the model cycle length used (seven days). This means that for a practical application of these utility values within the model, each of the time points would need to be rounded to a seven-day model cycle. The application of the utility values within the economic model is shown below, versus the label used to describe utility analysis performed:

● <30 days until death	→	≤28 days until death
● 30≤ days until death <90	→	28< days until death ≤84
● 90≤ days until death <180	→	84< days until death ≤175
● 180≤ days until death <360	→	175< days until death ≤357
● ≥360 days until death	→	>357 days until death

The ERG would have preferred for the categories used to fully align with the model cycle length used, or for a different cycle length to be used such that these categories could be applied as intended (with a preference for the former of these options, given that the remainder of model is aligned with a weekly model cycle length). Nevertheless, the impact of changing these groupings is not expected to have a large impact on the overall utility values produced. In

addition, when the ERG performed a check of the model calculation, an apparent error in the application of the time-to-death utilities was identified (discussed further in Section 5.3, with the impact of resolving this error on results discussed in Section 6.1).

As per the progression-based analysis, age adjustment was applied to account for decreasing utility as patients aged. In addition, the AE-related disutility was used to generate a total QALY loss for each arm, and applied as a lump sum in the first model cycle (also as per the progression-based approach).

4.2.7.3. Impact of adverse events

To include the impact of AEs on utility within the model, the company included covariates within the utility regressions for the presence of a Grade 3+ AE. Then, to estimate the total loss in QALYs due to Grade 3+ AEs, the company extracted the mean duration of the all-cause AEs from the KEYNOTE-590[12] patient-level data and multiplied this by the loss in utility. The duration of AEs was assumed equal between arms, and the average duration was calculated to be 8.24 weeks (see CS Table 49).

4.2.7.4. Utility values used in the model

The utility values generated from both analyses are presented in Table 12. Of note, these values are representative of the utility values used to populate the model in the first cycle only, as future cycles are affected by age-related disutility. The company's preferred base-case analysis made use of the time-to-death derived utility values, with progression-based values explored within sensitivity analysis.

Table 12: Utilities calculated based on KEYNOTE-590 trial data

Health state	Progression approach	Time-to-death approach
Progression-free	[redacted]	-
Progressed disease	[redacted]	-
≥360 days to death	-	[redacted]
180 to 360 days to death	-	[redacted]
90 to 180 days to death	-	[redacted]
30 to 90 days to death	-	[redacted]
0 to 30 days to death	-	[redacted]

Health state	Progression approach	Time-to-death approach
Presence of Grade ≥3 AE	[redacted]	[redacted]

Abbreviations: AE, adverse event.

Source(s): Produced based on CS Tables 50 and 51, as well as information provided in the company-submitted economic model. Time-to-death approach used in base-case analysis, progression approach explored in sensitivity analysis.

Based on the mean age at baseline (61.4 years, for the European population), and the proportion of patients that are male (80.7%, also for the European population), the average utility expected in the general population (based on the study by Ara and Brazier, 2010)[56] is estimated to be 0.829. The average utility value used for the progression-free health state is slightly lower than this value ([redacted]), suggesting that the impact of disease on HRQoL is relatively small (see Table 12). Furthermore, the average utility in the ≥360 days to death grouping ([redacted]) was greater than the equivalent value in the general population (0.829). In response to clarification questions, the company capped the utility values applied within the model to be equal to general population, should the value estimated from KEYNOTE-590 exceed this. Although this addresses the issue of the utility being greater than general population, the capping still assumes that patients with advanced oesophageal cancer over a year away from death have the same quality of life as the general population, most of whom would be expected to have a life expectancy greater than 1 year. As such, these results are misaligned with the expectation of relatively low utility for patients with metastatic cancer undergoing intensive chemotherapy with a relatively poor prognosis (versus 'healthy' individuals in the general population). Therefore, the ERG has concerns with the generalisability of the utility values produced based on analysis of KEYNOTE-590[12] data (regardless of which approach is used), as the outputted values imply that patients have a similar, or potentially better utility than the age- and sex-adjusted UK general population.

To further explore the utility values, the ERG calculated the average utility value for patients on the pembrolizumab in combination with chemotherapy combination arm using both approaches. To do this, the total undiscounted QALYs were divided by the total undiscounted LYs. This crude calculation allows for further exploration of how QALYs are accrued within the company's model.

Using the company's corrected base case model (see Section 6.1), for the progression approach, the average utility was estimated to be [redacted]. Switching to the time-to-death approach,

the average utility was estimated to be [redacted]. Using the ERG's preferred assumptions (see Section 6.3), the average utility values are [redacted] and [redacted], respectively. When setting the utility value for all health states to be 0.829 (average utility expected at baseline for the age- and sex-adjusted general population), and disabling AE-related QALY losses, the equivalent average utility for the general population was estimated to be 0.808. The ERG is concerned that the two utility analysis approaches lead to a substantially different estimation of the "average" utility experienced over the course of the model time horizon. This means that the incremental QALY gain attributable to pembrolizumab in combination with chemotherapy estimated for both utility analyses also varies markedly: The time-to-death approach (company base-case analysis post corrections) yields an incremental QALY gain of 0.652, versus 0.570 for the progression approach. The ERG considers the progression approach to yield a more realistic "average" utility for this patient population, especially given that the time-to-death approach yields an "average" utility that is close to the estimate for the general population.

The ERG has explored several alternative utility values within exploratory analyses, which are described in Section 6.2.3.

In the company's base-case time-to-death analysis, the total QALY loss attributable to Grade 3+ AEs is estimated at [redacted] for the pembrolizumab in combination with chemotherapy combination arm, versus [redacted] for the chemotherapy arm. In the progression approach, the total QALY losses are estimated at [redacted] versus [redacted] (for pembrolizumab in combination with chemotherapy versus chemotherapy). The ERG notes that it is unclear why the estimated QALY loss due to AEs is greater for the progression versus the time-to-death approach. In addition, the ERG questioned the face validity of a near-negligible impact in terms of toxicity for the addition of pembrolizumab to the combination of fluorouracil and cisplatin. Taking these two observations together, the ERG found it strange that the method of analysing the utility data appears to have a notably larger impact on the total estimated loss in QALYs due to AEs versus the introduction of a third treatment.

Further to the concerns raised with the impact of AEs on utility discussed above, the ERG had concerns with the estimated utility decrement associated with a Grade 3+ AE. It was the ERG's understanding that as utility measures are most likely to be taken at the start of each treatment cycle, any AEs resulting in hospitalisation (as implied by only including those at Grade 3 or above) were likely to be recorded after the EQ-5D questionnaire was completed, or that patients currently experiencing a particularly severe AE are likely to have not completed the EQ-5D

questionnaire. Therefore, the magnitude of the estimated decrement associated with all occurrences of Grade 3+ AEs ([redacted] in the company's base-case analysis) may be under-estimated.

4.2.8. Resources and costs

The company's model included costs relating to pembrolizumab, chemotherapy treatments, medial resource use, subsequent treatments, and the resolution of adverse events; each of which are discussed below.

4.2.8.1. Drug acquisition costs

The list price of a 100 mg vial for pembrolizumab is £2,630 resulting in a cost of £5,260 per administration for two 100 mg vials every three weeks, for a maximum of 35 cycles. The cost per administration is [redacted] when the company's patient access scheme (PAS) of [redacted] is applied.

The costs of the chemotherapy treatments were sourced from the NHS drugs and pharmaceutical electronic market information tool (eMIT), and dosing of the regimens were based on the KEYNOTE-590 trial protocol (see Section 4.2.4) or based on the summary of product characteristics (SmPC). The company's base case assumed vial sharing for treatments based on body surface area (BSA) using the minimum cost/mg, therefore costs for wastage are not included. The model included relative dose intensity (RDI) to account for missed doses and dose reductions which the company interprets as *"a proportion of the protocol dose that participants actually received"* (CS Section B.3.5). The RDI from KEYNOTE-590[12] was used assuming that oxaliplatin has the same RDI as cisplatin, and leucovorin, capecitabine and epirubicin has the same RDI as 5-FU. The ERG checked these assumptions with clinical advisors who did not think the assumption of epirubicin being equivalent to 5-FU was appropriate. However, given the triplet regimens are exploratory and epirubicin had a relatively low cost the ERG did not explore this further. In response to clarification questions, the company confirmed that treatment compliance in KEYNOTE-590 is considered different between pembrolizumab and chemotherapy. For pembrolizumab this is defined as the percentage of actual non-zero dose treatment cycles versus the expected number of treatment cycles per subject. For 5-FU or cisplatin, this is defined as total dosage received versus the total dosage expected (company response to clarification questions B18).

The costs and dosing used for each treatment in the company's model are presented in Table 13. As discussed in Section 4.2.4, clinical advice given to the ERG indicated that some of the chemotherapy dosages used in the CS were not reflective of doses currently received by patients in UK clinical practice. As such, the ERG explored alternative doses in scenario analysis (see Section 6.2).

Table 13: Costs and dosing of treatments included in the company's model

Regimen	Drug	Dose per administration	Unit cost	Pack size	RDI	Cost per administration
Pembrolizumab + 5-FU + cisplatin	Pembrolizumab	200 mg Q3W	£2,630	100 mg vial	[redacted]	[redacted]
	5-FU	800 mg days 1-5 Q3W	£2.84	2,500 mg vial	[redacted]	[redacted]
	Cisplatin	80 mg/m^2 Q3W	£6.66	100 mg vial	[redacted]	[redacted]
5-FU + cisplatin	5-FU	800 mg days 1-5 Q3W	£2.84	2,500 mg vial	[redacted]	[redacted]
	Cisplatin	80 mg/m^2 Q3W	£6.66	100 mg vial	[redacted]	[redacted]
5-FU + oxaliplatin + leucovorin	5-FU	2,600 mg/m^2 Q2W	£2.84	2,500 mg vial	[redacted]	[redacted]
	Oxaliplatin	85 mg/m^2 Q2W	£8.67	100 mg vial	[redacted]	[redacted]
	Leucovorin	200 mg/m^2 Q2W	£7.19	500 mg vial	[redacted]	[redacted]
Capecitabine + cisplatin	Capecitabine	2,000 mg/m^2 days 1-14 Q3W	£7.29	60 x 300 mg tablets	[redacted]	[redacted]
	Cisplatin	80mg/m^2 Q3W	£6.66	100 mg vial	[redacted]	[redacted]
Capecitabine + oxaliplatin	Capecitabine	2,000 mg/m^2 days 1-14 Q3W	£7.29	60 x 300 mg tablets	[redacted]	[redacted]
	Oxaliplatin	130 mg/m^2 Q3W	£8.67	100 mg vial	[redacted]	[redacted]
5-FU + cisplatin + epirubicin	5-FU	200 mg/m^2 days 1-21 Q3W	£2.84	2,500 mg vial	[redacted]	[redacted]
	Cisplatin	60 mg/m^2 Q3W	£6.66	100 mg vial	[redacted]	[redacted]
	Epirubicin	50 mg/m^2 Q3W	£19.29	200 mg vial	[redacted]	[redacted]
5-FU + oxaliplatin + epirubicin	5-FU	200 mg/m^2 days 1-21 Q3W	£2.84	2,500 mg vial	[redacted]	[redacted]
	Oxaliplatin	130 mg/m^2 Q3W	£8.67	100 mg vial	[redacted]	[redacted]
	Epirubicin	50 mg/m^2 Q3W	£19.29	200 mg vial	[redacted]	[redacted]
Capecitabine + cisplatin + epirubicin	Capecitabine	625 mg/m^2 days 1-21 Q3W	£7.29	60 x 300 mg tablets	[redacted]	[redacted]
	Cisplatin	60 mg/m^2 Q3W	£6.66	100 mg vial	[redacted]	[redacted]
	Epirubicin	50 mg/m^2 Q3W	£19.29	200 mg vial	[redacted]	[redacted]

Regimen	Drug	Dose per administration	Unit cost	Pack size	RDI	Cost per administration
Capecitabine + oxaliplatin + epirubicin	Capecitabine	625 mg/m^2 days 1-21 Q3W	£7.29	60 x 300 mg tablets	[redacted]	[redacted]
	Oxaliplatin	130mg/m^2 Q3W	£8.67	100 mg vial	[redacted]	[redacted]
	Epirubicin	50mg/m^2 Q3W	£19.29	200 mg vial	[redacted]	[redacted]

Key: 5-FU, fluorouracil; Q2W, every 2 weeks; Q3W, every 3 weeks; RDI, relative dose intensity

Notes: *List price

4.2.8.2. Treatment administration

The majority of treatments included in the company's model are administered via intravenous infusion with the exception of capecitabine which is an oral treatment. The CS stated that administration costs are included in the model based on NHS reference costs 2018-19, however the model inputs actually use NHS reference costs 2018-19 and uplift them to 2019/20 costs using indices from the NHS cost inflation index.[58] All treatment regimens were assigned the cost of £322.88 (uplifted from £317.73 SB14Z, deliver complex chemotherapy, including prolonged infusion treatment at first attendance), except for oxaliplatin plus capecitabine which has the cost of £263.28 assigned (uplifted from £259.08 SB13Z, deliver more complex parenteral chemotherapy at first attendance). At clarification stage, the company confirmed that choice of administration HRG code was based on the National Tariff Chemotherapy Regimens List 2017-18[59] (see company response to clarification questions B19).

The pembrolizumab in combination with chemotherapy does not have a specific code according to the National Tariff Chemotherapy Regimens List. Pembrolizumab would add an additional 30 minutes according to the SmPC, thus it is unclear if another code would be more appropriate. However, given that administration costs have little impact on the ICER, the ERG considered the administration cost codes to be broadly satisfactory.

Clinical advice provided to the ERG was that treatment administration would be given in a day case setting, whereas the company used the outpatient cost code. In addition, clinical experts explained that 5-FU would require more visits to the day case unit and there would be additional charges for pump disconnections and a peripherally inserted central catheter (PICC) line, as such, the administration and monitoring costs of patients receiving 5-FU could have been underestimated. Though the ERG noted that the company's administration code for 5-FU is in line with the National Tariff Chemotherapy Regimens List.

Consequently, the ERG's preferred assumptions for administration costs included costs based on a day case setting. These costs are provided in Table 14 in comparison to the costs used by the company. The ERG's preferred administration costs are factored into the ERG's preferred base case (see Section 6.3).

Table 14: Comparison of company and ERG preferred administration costs

Administration cost	Company	ERG
Deliver more complex parenteral chemotherapy at first attendance	£263.28 (uplifted from £259.08 SB13Z OP)	£319.46 (uplifted from £314.39 SB13Z DCRDN)
Deliver complex chemotherapy, including prolonged infusional treatment, at first attendance	£322.88 (uplifted from £317.73 SB14Z OP)	£391.52 (uplifted from £385.28 SB14Z DCRDN)

Key: OP, outpatient; DCRDN, day case/regular day/night

Source(s): NHS reference costs 2018-19[60]

4.2.8.3. Time on treatment

The company applied maximum treatment durations for pembrolizumab, 5-FU and cisplatin in line with the stopping rules in the KEYNOTE-590 protocol. That is, 35 cycles for both pembrolizumab and 5-FU, and 6 cycles for cisplatin. These stopping rules were applied to the pembrolizumab in combination with chemotherapy arm and the trial-based comparator (5-FU plus cisplatin). No stopping rules were formally applied to the other chemotherapy regimens included in the blended chemotherapy arm.

To estimate the proportion of patients on treatment per cycle, the company used ToT Kaplan-Meier estimates from KEYNOTE-590.[12] Parametric curves were fit to the ToT data separately for each the five component treatments in KEYNOTE-590; pembrolizumab, 5-FU and cisplatin (in the pembrolizumab combination arm), 5-FU and cisplatin (in the chemotherapy control arm), though given the maturity of the ToT data, the Kaplan-Meier estimates were used directly in the base case. At clarification stage, the ERG requested further information on the maturity of the ToT data from the KEYNOTE-590 trial. The company presented the number of events for ToT in each arm by the July 2020 data cut off which showed that [redacted] had discontinued pembrolizumab, [redacted] had discontinued cisplatin (in both arms) with [redacted] and [redacted] having discontinued 5-FU in the pembrolizumab in combination with chemotherapy arm and chemotherapy arm, respectively. The ERG noted that using the Kaplan-Meier curve directly may underestimate treatment costs given not all patients have discontinued treatment (especially in the pembrolizumab arm). However, given the trial stopping rules at 2-years, more mature data is not expected to have a large impact on the ToT Kaplan-Meier data.

These are presented for the full population in the CS Figures 23 and 24, and [redacted]10 and [redacted]11 below. As per the RDI assumptions, for those blended chemotherapy treatments not

included in the trial it is assumed that oxaliplatin had the same ToT as cisplatin, and leucovorin, capecitabine and epirubicin had the same ToT as 5-FU.

The ERG agrees that using the ToT Kaplan-Meier estimates directly is appropriate to account for treatment discontinuations and disruptions for various reasons. However, given that the ToT data from KEYNOTE-590 already incorporate the protocol driven stopping rules, the ERG noted that it was not necessary to apply the maximum treatment durations in addition to using the Kaplan-Meier estimates and RDI. The maximum treatment durations are applied in the model as a hard stop at those time points, whereas the ToT Kaplan-Meier estimates show that some patients were still on treatment after those time points (due to dose interruptions), therefore the model base case does not currently capture these (see [redacted]10 and [redacted]11). At clarification stage, the ERG requested a scenario where the treatment stopping rules are disabled and ToT Kaplan-Meier data is used directly. This resulted in only a slight impact on the ICER (see Section 5.2.3); however, the ERG considered this to be more reflective of clinical practice and therefore, the ERG's preferred base-case analysis removes the maximum treatment durations (see Section 6.3).

[redacted]10

[redacted]

Key: 5-FU, fluorouracil; KM, Kaplan-Meier; ToT, time on treatment

[redacted]11[redacted]

Key: 5-FU, fluorouracil; KM, Kaplan-Meier; ToT, time on treatment

4.2.8.4. Health state unit costs and resource use

Disease management costs associated with the progression-free and progressed health states are included in the model based on resource frequencies derived from clinical expert opinion and a previous NICE appraisal for previously treated advanced gastric cancer or gastro-oesophageal junction adenocarcinoma (TA378),[61] respectively.

Progression-free

The resource frequencies and unit costs were presented in the CS in Table 58. The progression-free health state included a full blood count, a renal function test and a hepatic function test every three weeks, with a consultation visit every four weeks and a CT scan every 12 weeks. Based on clinical advice provided to the ERG, patients are currently monitored every three weeks whilst on platinum-based chemotherapy then every three months whilst continuing treatment with a fluoropyrimidine. If patients are still receiving pembrolizumab after discontinuation of platinum-based chemotherapy, then monitoring would be every six weeks, instead of every three months. At clarification stage, the ERG requested the company to provide revised resource use based on the increased monitoring for those patients remaining on pembrolizumab after discontinuing chemotherapy. In response, the company stated that

*"Disease management costs applied in the model are linked to the progression free survival curve, not to the pembrolizumab time on treatment curve. As such, treatmen*t status does not impact the monitoring costs" (see response to clarification questions B20). As the company did not provide revised estimates post clarification questions, the ERG has explored the impact of increased monitoring within Section 6.2.

Progressed disease

For progressive disease, the only resource use included was a consultation visit every 12 weeks. The ERG registered several concerns with the estimates of progressed resource use frequencies. The progressed disease state resource use appears implausibly reduced compared to progression-free patients. Given this indication is in a first-line setting, some patients will receive subsequent treatment when they progress and are therefore expected to require more than a consultation visit every 12 weeks.

The company states that their progressed disease resource use was based on TA378, however in TA378,[61] patients on (2L+) treatment are assumed to have a full blood count, a renal function test and a hepatic function test every treatment cycle, with a consultation visit every 4 weeks and a CT scan every 12 weeks. Upon progression (or patients off treatment), patients are then assumed to have consultation visits every 12 weeks. This means that the resource use from TA378 has been assumed to apply to an earlier line, but without accounting for resource use needed for patients receiving active therapy in a second-line setting.

In the recent nivolumab appraisal for previously treated advanced oesophageal cancer (ID1249),[46] resource use was determined from a clinical survey for those patients on treatments consisting of consultations, imaging scans, blood tests, liver function tests, kidney function tests, hospitalisations and palliative care specialist nurses. The mean weekly visits were calculated from the possible options of every three months, monthly, biweekly, weekly and never.

The ERG requested more information post clarification questions in which the company revised its base case to include a one-off cost within the subsequent treatment cost to account for extra monitoring for those patients receiving treatment. This includes the monitoring frequencies described in TA378 for those patients on 2L+ treatment. The ERG agreed with the company's revisions on progressed disease monitoring.

PD-L1 testing costs

For the CPS ≥ 10 sub-population PD-L1 testing costs were included to reflect the costs incurred by the NHS when testing for PD-L1 expression. The unit cost per test was based on the cost used in the previous NICE submission TA522[62] (£40.50) uplifted to 2019/20 costs using indices from the NHS cost inflation index.[58] It is assumed that 51.1% of patients tested will be classed as PD-L1 positive based on KEYNOTE-590[12] resulting in a total cost of £85.12.

It is not clear from the company submission or TA522 where the PD-L1 testing cost comes from so the ERG are unable to assess its appropriateness. Clinical experts advised the ERG that PD-L1 will be conducted at the same time at other histologically tests on the biopsy samples, and that the cost of including PD-L1 testing is not expected to be resource intensive. As such the ERG does not envision the PD-L1 test to be expensive and considers the cost used by the company to be appropriate (though not possible to verify).

4.2.8.5. Adverse events unit costs and resource use

The company included management costs associated with each adverse event (discussed in Section 4.2.6.3) based on NHS reference costs 2018-19 uplifted to 2019/20 costs using indices from the NHS cost inflation index.[58] The company justifies its choice of cost codes assigned to each adverse event based on previous pembrolizumab NICE submissions in advanced urothelial cancer and metastatic squamous cell head and neck cancer.[43,63] The ERG did not consider it good practice to refer to previous submissions alone (especially those only conducted by the company itself) in different indications to justify the most appropriate model inputs.

The ERG noted some differences between the cost codes used in the previous submissions cited by the company and the ones used for this submission, mainly due to choosing one cost code over a weighted average as presented in Table 15. There were no justifications for these differences or choices of cost codes for each of the adverse events within the CS.

Table 15: Comparison of adverse event costs used in previous submissions

Adverse event	ID3741 Company submission	TA519	TA661
Anaemia	SA01G - Acquired Pure Red Cell Aplasia or Other Aplastic Anaemia, with CC Score 8+. Non-elective short stay	SA01G-K- Acquired Pure Red Cell Aplasia or Other Aplastic Anaemia. Weighted cost of non-elective long stay, short stay and day case	SA01G-K- Acquired Pure Red Cell Aplasia or Other Aplastic Anaemia. Weighted cost of non-elective long stay, short stay and day case
	£623.25	£1,315.94	£631.88
Decreased Appetite	FD04A - Nutritional Disorders with Interventions, with CC Score 2+. Day case	NR	SPHMSEDSAAPC – Adult Specialist Eating Disorder Services: Admitted patient
	£301.33	-	£461.74
Dysphagia	A13A1 - Speech and Language Therapist, Adult, One to One. Community health services	NR	Assumed to be £0
	£108.24	-	£0.00
Fatigue	SA01G - Acquired Pure Red Cell Aplasia or Other Aplastic Anaemia, with CC Score 8+. Non-elective short stay	WH52A – Follow-Up Examination for Malignant Neoplasm, with Interventions. Non-elective long stay 8-9 days	SA01G-K- Acquired Pure Red Cell Aplasia or Other Aplastic Anaemia. Weighted cost of non-elective long stay, short stay and day case
	£623.25	£2,499.99	£631.88
Hypokalaemia	KC05H - Fluid or Electrolyte Disorders, with Interventions, with CC Score 0-4. Non-elective short stay	NR	KC05G-H - Fluid or Electrolyte Disorders, with Interventions, with CC Score 0-5+
	£963.30	-	£1,104.28
Hyponatraemia	KC05H - Fluid or Electrolyte Disorders, with Interventions, with CC Score 0-4. Non-elective short stay	NR	KC05G-H - Fluid or Electrolyte Disorders, with Interventions, with CC Score 0-5+
	£963.30	-	£1,104.28
Nausea	FD10M - Non-Malignant Gastrointestinal Tract Disorders without	NR	FZ91M - Non-Malignant Gastrointestinal Tract Disorders without

Adverse event	ID3741 Company submission	TA519	TA661
	Interventions, with CC Score 0-2. Non-elective short stay		Interventions, with CC Score 0-2: Non elective short stay (two hospital admissions)
	£418.64	-	£894.04
Neutropenia	SA35A - Agranulocytosis with CC Score 13+. Non-elective short stay	WJ11Z – Other disorders of immunity. Weighted average of non-elective long and short stay and day case. (10% require treatment)	WJ11Z – Other disorders of immunity. Weighted average of non-elective long and short stay and day case
	£728.33	£70.80	£78.69
Neutrophil Count Decreased	WJ11Z - Other Disorders of Immunity. Total HRGs	Assumed same as neutropenia	WJ11Z – Other disorders of immunity. Weighted average of non-elective long and short stay and day case
	£474.18	£70.80	£78.69
Platelet count decrease	SA12 - Thrombocytopenia with CC Score 8+. Non-elective short stay	NR	Assumed to be £0
	£620.79	-	£0.00
Pneumonia	DZ11P - Lobar, Atypical or Viral Pneumonia, with Single Intervention, with CC Score 8-12. Total HRGs	DZ11K-V - Lobar, Atypical or Viral Pneumonia. Weighted average of non-elective long and short stay and day case	DZ11K-V - Lobar, Atypical or Viral Pneumonia. Weighted average of non-elective long and short stay and day case
	£3,449.89	£1,751.08	£495.81
Stomatitis	CB02A - Non-Malignant, Ear, Nose, Mouth, Throat or Neck Disorders, with Interventions, with CC Score 5+. Non-elective short stay	NR	Assumed to be £0
	£669.91	-	£0.00
Vomiting	FD10M - Non-Malignant Gastrointestinal Tract Disorders without Interventions, with CC Score 0-2. Non-elective short stay	NR	FZ91M - Non-Malignant Gastrointestinal Tract Disorders without Interventions, with CC Score 0-2: Non

Adverse event	ID3741 Company submission	TA519	TA661
			elective short stay (two hospital admissions)
	£418.64	-	£894.04
Weight decrease	N16AF - Specialist Nursing, Enteral Feeding Nursing Services, Adult, Face to face. Community health services	NR	NR
	£108.15	-	-
White blood cell count decrease	WJ11Z - Other Disorders of Immunity. Total HRGs	Assumed to be the same as neutropenia	WJ11Z - Other Disorders of Immunity. Weighted average of non-elective long and short stay and day case
	£474.18	£70.80	£78.69

Key: NR, not reported

Source(s): NHS reference costs 2018-19;[60] TA519;[43], TA661[63]

Some of the choices for adverse events are questionable; for example, the cost code used for dysphagia (A13A1 - Speech and Language Therapist, Adult, One to One. Community health services) does not seem appropriate. Dysphagia affects the patient's ability to swallow foods or liquids and although speech and language therapy could be used to learn new swallowing techniques,[64] this may not be appropriate for a Grade 3 or 4 adverse event, especially for gastro-oesophageal cancer patients. Clinical experts confirmed that these patients are usually treated with a stent or possibly tube fed until fit enough to go onto chemotherapy. Therefore, the ERG thought that a more suitable NHS reference cost could have been utilised (e.g., FE10A-D, Endoscopic Insertion of Luminal Stent into Gastrointestinal Tract).

However, given the similar adverse event profiles of pembrolizumab in combination with chemotherapy and chemotherapy using the KEYNOTE-590[12] data (see CS Table 48 and Section 3.2.3.3), the adverse event rates for dysphagia have a difference of <1%, therefore the resulting total adverse event costs are similar between arms and hence the unit costs have very little impact on cost-effectiveness results. Thus, the ERG did not explore this further.

4.2.8.6. Miscellaneous unit costs and resource use

Subsequent therapy costs

Within the economic model, subsequent treatment costs are applied as a one-off cost when a patient leaves the 'progression-free' health state. The distribution of subsequent treatments and mean durations are based on the KEYNOTE-590[12] data. The company applied an arbitrary cut-off of excluding all subsequent treatments in which less than 5% of patients received. In addition, the distribution was equally re-weighted to exclude ramucirumab as this was not recommended for use in NHS for previously treated advanced gastro-oesophageal junction adenocarcinoma patients.[61] The resulting distributions are presented in the CS (Table 61) with the dosing schedules and costs for each subsequent treatment presented in the CS Table 62. The distributions show a high usage of paclitaxel followed by fluorouracil, whereas the most common treatment in clinical practice is generally docetaxel based on clinical opinion. However, paclitaxel is also commonly used in practice therefore the subsequent treatments included in the model look reflective of UK practice.

Estimates of RDI were not included for subsequent treatments and vial sharing was assumed which the company claims is: *"constituting a conservative approach"* (CS Section B.3.5). The ERG was unclear why this is considered a conservative approach.

At clarification stage, the ERG requested to see the full subsequent treatment list received by patients in the KEYNOTE-590[12] trial before the company applied the arbitrary 5% cut-off. The company provided the full list in addition to exploring an alternative approach to calculate the distribution using the total number of progression events (progression and death) as the denominator instead of the number of patients within the as-treated population.

After reviewing the full subsequent treatment table, the ERG noted with grave concern with the application of the 5% cut-off within the subsequent treatment costing. In KEYNOTE-590, [redacted] and [redacted] patients received subsequent treatments in the pembrolizumab in combination with chemotherapy arm and chemotherapy arm, respectively. Of the [redacted] patients who had subsequent treatments, there was a total of [redacted] subsequent treatments covering multiple treatment lines. When applying the 5% cut-off, the company only includes [redacted] subsequent treatments and excludes the others. This results in an underestimation of subsequent treatment costs and the unnecessary removal of data. In order to account for all subsequent treatments received by patients in KEYNOTE-590, the ERG re-distributed the remaining treatments into the most common treatments received (using the company's 5% threshold). The redistribution and resulting costs are presented in Table 16 compared to the company's estimates. The ERG included this re-distribution within their preferred assumptions (see Section 6.3)

Table 16: Subsequent treatment re-distributions

Subsequent treatment	KEYNOTE-590		Company's re-distribution [a]		ERG's re-distribution [b]	
	Pembrolizumab in combination with chemotherapy	Chemotherapy	Pembrolizumab in combination with chemotherapy	Chemotherapy	Pembrolizumab in combination with chemotherapy	Chemotherapy
N	370	370	370	370	370	370
Cisplatin	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Docetaxel	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
5-FU	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Irinotecan hydrochloride	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Oxaliplatin	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Paclitaxel	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Others	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total [c]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total cost [d]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Notes: [a] Remove any below 5% in both treatment arms and re-distribute removing ramucirumab. [b] Remove any below 5% and ramucirumab and re-distribute the remainder between the included treatments. [c] Including multiple subsequent treatment lines. [d] Includes, drug administration and disease monitoring costs.

Terminal care costs

The economic model includes a terminal care cost to reflect the costs associated with death applied as a one-off cost when patients enter the death health state. In the CS (Document B, page 117), the company states the end-of-life cost to be £7,630.19, however the economic model uses £7,795.01 which is the correct value. The terminal care cost was derived from a previous pembrolizumab NICE submission in advanced or metastatic urothelial cancer (TA522)[62] and uplifted to 2019/2020 costs using the indices from the NHS cost inflation index.[58]

The original cost of £7,252.82 used in TA522[62] was derived from a variety of sources based on resource frequencies and unit costs derived from PSSRU 2015. These resources were based on what was previously accepted in TA519 (pembrolizumab for previously treated advanced or metastatic urothelial cancer)[43] which in turn was derived from other previous submissions (TA272, TA374 and TA277).

The ERG noted a few concerns with this cost. Firstly, the ERG considered the use of uplifting the total cost from TA522 to be unnecessary given the breakdown of resources and unit costs are presented and as such the company could have used the latest unit costs per resource to update the costs. Secondly, this cost has been derived from a chain of previous submissions and thus, does not take into consideration the assumptions surrounding the individual resources and whether they are appropriate for gastro-oesophageal cancer patients. For example, in TA519 the company submission states that *"Clinical advice suggested that due to their propensity to bleed, patients with urothelial cancer receive radiotherapy at end of life; therefore, this cost has also been included."*[43] This cost of radiotherapy was included within the terminal care cost used in TA522 and subsequently in this submission. However, the radiotherapy cost may not be appropriate to consider for gastro-oesophageal cancer patients as it was specifically included for patients with urothelial cancer and as such the terminal care costs may be over-estimated. On the other hand, the cost used in the company's model is lower than that used in the recent nivolumab appraisal (ID1249)[46] which used £8,973.61 from the literature (inflated from £7,827.00) estimating the per-patient costs in the last three months of life.[65]

It was not clear from the company's submission or from tracing back through the previous submissions what period of time the terminal care costs covers making it difficult to assess the appropriateness of a one-time cost, but includes resource use associated with 28 hours

community nurse, seven GP home visits, 50 hours of Macmillan nurse time, and terminal care in hospital or a hospice for a proportion of patients per resource.

In conclusion, the company should have provided more information about the origins of this cost and assessment of how reflective these assumptions were to gastro-oesophageal cancer patients, in addition to recalculating the total cost on first principles as rather than inflating the previous total cost. However, the ERG acknowledged that terminal care costs have little influence on cost-effectiveness results given that the model covers almost a lifetime horizon with the majority of patients dead in both arms. Exploring different values (i.e., removing radiotherapy costs or using the same cost as per the ID1249 submission) resulted in minor differences in the ICER (see Section 6.2).

5. COST-EFFECTIVENESS RESULTS

5.1. Company's cost-effectiveness results

5.1.1.1. Base case results

All patients

The company revised its base case following clarification questions, therefore only the revised results are presented in this section. The changes included:

- Changing the time horizon to 30-years
- Including a utility cap for the time-to-death category '≥360 days' based on general population utility values
- Including disease management costs for those patients who receive subsequent treatment after progression

The revised results reported by the company are shown in Table 17 for the comparison against the trial comparator as per KEYNOTE-590.[12] The deterministic and probabilistic results are consistent with incremental cost-effectiveness ratios (ICERs) of £41,688 and £42,303 per QALY gained respectively. Of note, the company's base case analysis incorporated a PAS discount of [redacted] applied to the list price of pembrolizumab.

Table 17: Company base case results – all patients

	Discounted costs	Discounted QALYs	Incremental discounted costs	Incremental discounted QALYs	Cost per QALY gained
Company deterministic base case					
Pembrolizumab + chemotherapy	[redacted]	[redacted]	-	-	-
5-FU + cisplatin	[redacted]	[redacted]	£27,172	0.65	£41,688
Company probabilistic base case					
Pembrolizumab + chemotherapy	[redacted]	[redacted]	-	-	-
5-FU + cisplatin	[redacted]	[redacted]	£27,253	0.64	£42,303

Key: QALYs, quality adjusted life years

Source(s): Company response to clarification questions, Appendix C Table 1 and Table 4

The results reported by the company for the non-trial comparators are shown in Table 18. As discussed in Section 4.2.4, the company assumed the same efficacy as the trial comparator using the control arm from KEYNOTE-590[12] therefore only the drug costs influenced ICER differences. The CS did not provide fully incremental analysis; however, based on the assumption of equal efficacy and safety, the chemotherapy regimen with the lowest overall costs would be cost-saving versus the other chemotherapy regimens. This means that capecitabine + oxaliplatin is cost-saving versus the other chemotherapy regimens listed in Table 18. The resulting ICER for pembrolizumab in combination with chemotherapy versus capecitabine plus oxaliplatin is £42,172.

Table 18: Company base case results versus the non-trial comparators – all patients

	Discounted costs	Discounted QALYs	Incremental discounted costs	Incremental discounted QALYs	Cost per QALY gained
Pembrolizumab + chemotherapy	[redacted]	[redacted]	-	-	-
5-FU+cisplatin	[redacted]	[redacted]	27,172	0.65	41,688
5FU + oxaliplatin + leucovorin	[redacted]	[redacted]	25,949	0.65	39,812
Capecitabine + cisplatin	[redacted]	[redacted]	27,072	0.65	41,535
Capecitabine + oxaliplatin	[redacted]	[redacted]	27,487	0.65	42,172
5-FU + cisplatin + epirubicin	[redacted]	[redacted]	27,115	0.65	41,601
5-FU + oxaliplatin + epirubicin	[redacted]	[redacted]	27,073	0.65	41,536
Capecitabine + cisplatin + epirubicin	[redacted]	[redacted]	27,036	0.65	41,480
Capecitabine + oxaliplatin + epirubicin	[redacted]	[redacted]	26,994	0.65	41,415
Blended comparator*	[redacted]	[redacted]	26,988	0.65	41,405

Key: QALYs, quality adjusted life years; SOC, standard of care

Source(s): Company response to clarification questions, Appendix C Table 2 and Table 3

Notes: *Weighted costs assuming equal market share (~12.5% for each treatment)

CPS ≥10

The results reported by the company for the CPS ≥10 sub-population are shown in Table 19 for the comparison against the trail comparator as per KEYNOTE-590.[12] The deterministic results

gave an ICER of £32,995 per QALY gained. Of note, the company's base case analysis incorporated a PAS discount of [redacted] applied to the list price of pembrolizumab.

Table 19: Company base case results = CPS ≥10

	Discounted costs	Discounted QALYs	Incremental discounted costs	Incremental discounted QALYs	Cost per QALY gained
Company deterministic base case					
Pembrolizumab + chemotherapy	[redacted]	[redacted]	-	-	-
5-FU + cisplatin	[redacted]	[redacted]	£30,293	0.92	£32,995

Key: CPS, combined positive score; QALYs, quality adjusted life years

Source(s): Company response to clarification questions, Appendix C Table 8

The results reported by the company for the non-trial comparators are shown in Table 20. As for the full population, the CS did not provide fully incremental analysis; however, based on the assumption of equal efficacy, the chemotherapy regimen with the lowest overall costs would be cost-saving versus the other chemotherapy regimens. Accordingly, capecitabine + oxaliplatin is cost-saving versus the other chemotherapy regimens listed in Table 20. The resulting ICER for pembrolizumab in combination with chemotherapy versus capecitabine plus oxaliplatin was £33,337.

Table 20: Company base case results versus the non-trial comparators – CPS ≥10

	Discounted costs	Discounted QALYs	Incremental discounted costs	Incremental discounted QALYs	Cost per QALY gained
Pembrolizumab + chemotherapy	[redacted]	[redacted]	-	-	-
5-FU+cisplatin	[redacted]	[redacted]	£30,293	0.92	£32,995
5FU + oxaliplatin + leucovorin	[redacted]	[redacted]	£29,059	0.92	£31,650
Capecitabine + cisplatin	[redacted]	[redacted]	£30,189	0.92	£32,881
Capecitabine + oxaliplatin	[redacted]	[redacted]	£30,608	0.92	£33,337
5-FU + cisplatin + epirubicin	[redacted]	[redacted]	£30,231	0.92	£32,927
5-FU + oxaliplatin + epirubicin	[redacted]	[redacted]	£30,191	0.92	£32,883
Capecitabine + cisplatin + epirubicin	[redacted]	[redacted]	£30,154	0.92	£32,843

	Discounted costs	Discounted QALYs	Incremental discounted costs	Incremental discounted QALYs	Cost per QALY gained
Capecitabine + oxaliplatin + epirubicin	[redacted]	[redacted]	£30,113	0.92	£32,798
Blended comparator*	[redacted]	[redacted]	£30,105	0.92	£32,789

Key: CPS, combined positive score; QALYs, quality adjusted life years; SOC, standard of care

Source(s): Company model post clarification questions

Notes: *Weighted costs assuming equal market share (~12.5% for each treatment)

5.2. Company's sensitivity analyses

The company reported a number of sensitivity analyses to explore the impact of alternative settings and assumptions, as well as the role of parameter uncertainty within the model results. These analyses are discussed in turn below. Of note, no sensitivity analysis was presented in the CS for the CPS ≥10 sub-population.

5.2.1. One-way sensitivity analysis

The company conducted one-way sensivity analysis (OWSA) on various parameters listed in the CS Section B.3.6. Each variable was varied using the lower and upper bounds of the 95% confidence intervals. The majority of the confidence intervals were calculated from their assigned distributions (see CS Table 64) and assuming the standard error is 10% of the mean. Exceptions to this were the patient characteristics (upper and lower bounds calculated from the data), ToT Kaplan-Meier hazard ratio (upper and lower bounds assumed to be ±10% of the mean), and duration of Grade 3+ adverse events (upper and lower bounds calculated from the data).

A tornado plot was used to present the OWSA results in the CS Figure 27 for pembrolizumab in combination with chemotherapy versus the trial based comparator. The company's results showed that the OS parameters, relative dose intensity and annual discount rate for effectiveness had the greatest influence on the ICER ranging from £27,746 to £60,344.

The ERG identified a number of errors associated with the parameters included in sensivity analysis. Firstly, the company assigned gamma distributions to costs based on an average cohort which should have been assigned normal distrbutions. Secondly, the company included parameters which have a multivariate distribution in the OWSA. As these parameters are linked to other parameters (e.g., survival distribution paramaters shape and scale), they should not be

varied individually. Thirdly, drug costs sourced from eMIT were excluded from sensivity analysis when they should be included using the provided standard errors from eMIT. Finally, the company incorporated total costs into the sensivity analysis instead of the individual components (e.g., Total calculated adverse event cost per treatment are varied instead of adverse event rates and unit costs per adverse event individually varied). The ERG noted that such an approach would mask the individual impact these parameters could have on the results, as they may act in opposite directions or apply only to one treatment arm. In addition, they could be assigned different distributions if done seperately (e.g., adverse event rates assigned the beta distribution and unit costs assigned the normal distribution). Conversely, by grouping these parameters together and assuming a large standrd error, the uncertainty may be substantially over estimated.

The ERG flagged these to the company at clarification stage, and the company subsequently made the following changes:

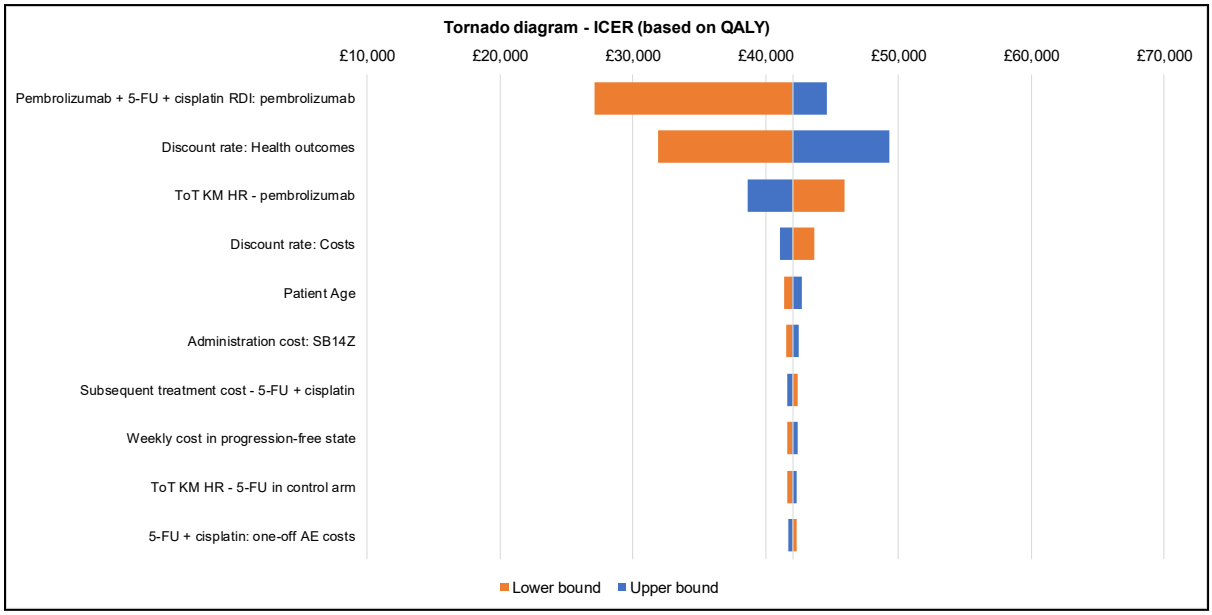
- Removed utility coefficients and survival curve coefficiants from OWSA.
- Revised some cost distributions from Gamma to normal (administration, subsequent treatments, disease management and adverse event costs).
- Included drug costs sourced from eMIT within OWSA and PSA with as assigned Gamma distribution.

The changes made by the company are considered appropriate, however the ERG would like to note that eMIT costs should have been assigned a normal distribution instead of Gamma. In addition, the company did not separate parameters to include them individually stating that *"this would require substantial modification of the model programming, and MSD are confident that the impact on the sensitivity analysis results would be minimal, and unimpactful on the deterministic base case result. Indeed, this was taken into account in the original model programming, and overall health state costs were accordingly used as the input"* (see company's response to clarification questions B22). The ERG believe that this way of incorporating parameters does not meet modelling best practice and as stated previously, grouping paramaters could under or over estimate the uncertainty.

The company's revised OWSA based on the revised base case (see Section 5.1.1.1) and changes described above is presented in Figure 12. This results show that the RDI for

pembrolizumab, annual discount rates and pembrolizumab's duration of treatment had the greatest influence on the ICER ranging from £26,764 to £48,930.

Figure 12: Company's tornado diagram presenting the results of the deterministic sensitivity analysis for the 10 most sensitive variables



Key: HR, hazard ratio; ICER, incremental cost-effectiveness ratio; KM, Kaplan-Meier; QALY, quality-adjusted life-year; RDI, relative dose intensity; ToT, time on treatment

Source: Company response to clarification questions, Appendix C Figure 3

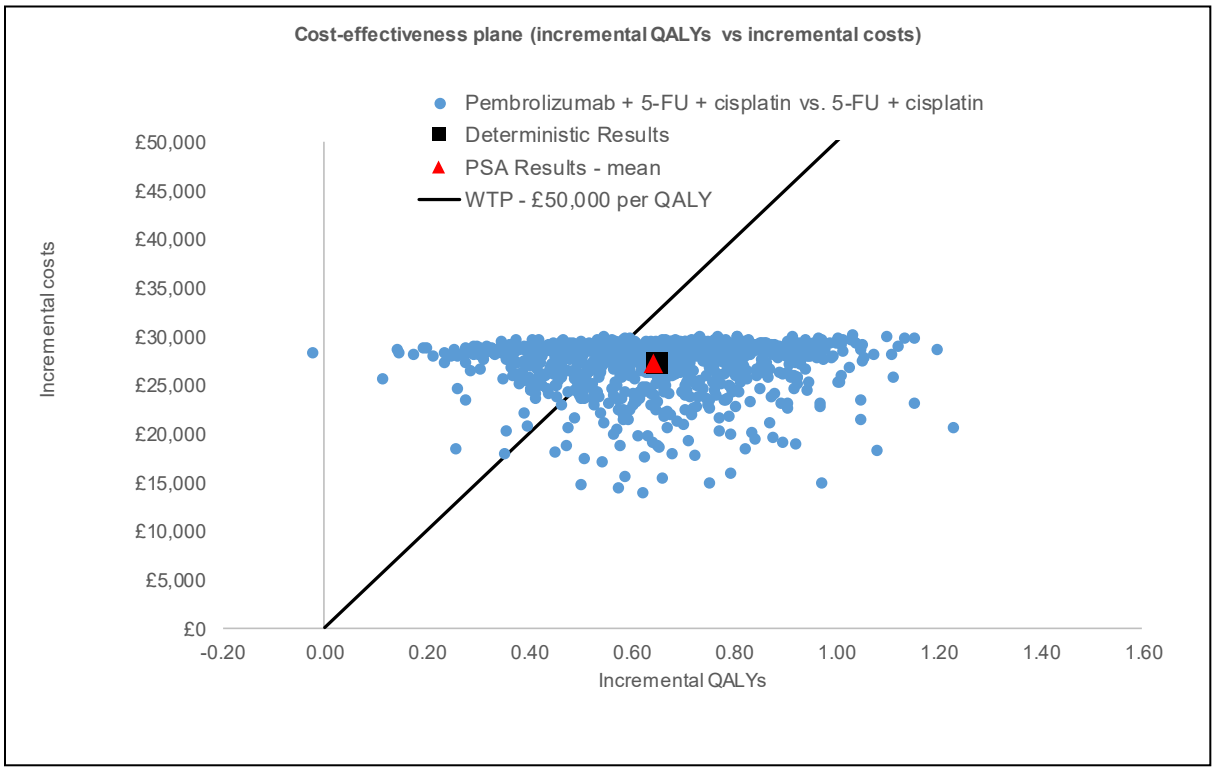
5.2.2. Probabilistic sensitivity analysis

The company conducted probabilstic sensitivity analysis (PSA) to explore the impact of parameter uncertainty. PSA results are presented in the CS Table 69 for pembrolizumab in combination with chemotherapy versus the trial based comparator. This showed consistant results to the detemrinistic ICER and demonstrated a 69.8% chance of pembrolizumab in combination with chemotherapy being cost-effective compared to the trial based comparator at the £50,000 per QALY threshold.

As discussed above in Section 5.2.1, the ERG identified a number of errors associated with the parameters included in sensivity analysis. The ERG requested the company make changes and re-run their analysis. Following clarification questions, the company updated their base case (see Section 5.1.1.1) and made changes to the sensivity analysis as described in Section 5.2.1. The updated PSA results showed consistent results to the deterministic ICER and demonstrated

a 68.5% chance of pembrolizumab in combination with chemotherapy being cost-effective compared to the trial based comparator at the £50,000 per QALY threshold. The revised cost-effectiveness plane and cost-effectiveness acceptability curve are presnted in Figure 13 and Figure 14, respectively.

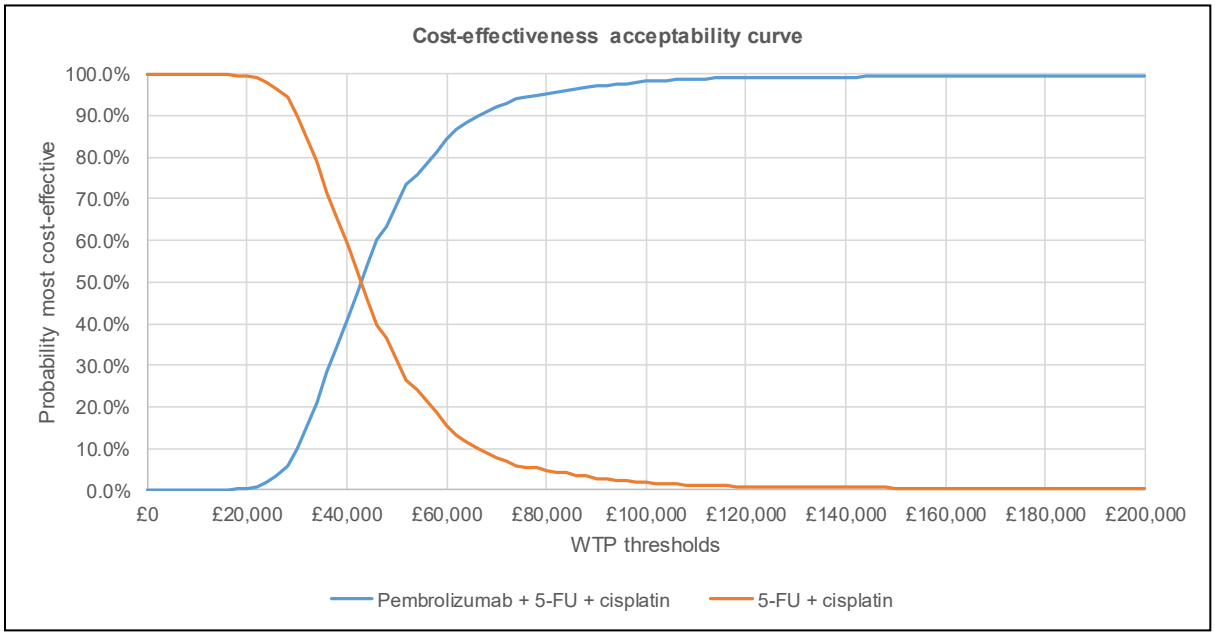
Figure 13: Company's scatter plot of PSA simulations



Key: PSA, probabilistic sensitivity analysis; QALY, quality-adjusted life-year

Source: Company response to clarification questions, Appendix C Figure 1

Figure 14: Company's cost-effectiveness acceptability curve



Key: 5-FU, five fluoruracil; WTP, willingness to pay

Source: Company response to clarification questions, Appendix C Figure 2

5.2.3. Scenario analyses

The company conducted a number of scenario analyses to assess the impact of structual uncertanties and alternative settings and assumptions on the base ase results versus the trial comparator. These scenarios include:

- Alternative parametric distributions for OS (log-normal and Weibull).
- Alternative cut-off for the OS piece-wise modelling of 32-weeks.
- Exploring a treatment waning effect starting at five years finishing at seven years.
- Alternative parametric distributions for PFS (log-normal).
- Alternative cut-off for the PFS piece-wise modelling of 37-weeks.
- Alternative apporach to model ToT using fully parametric fitted models (generalised gamma for pembrolizumab, Weibull for 5-FU and KM for cisplatin).
- Exploring the removal of relative dose intensity.

- Alternative time horizons (10, 30 and 40 years).
- Assuming all patients go receive nivoumab as subsequent treatment following chemotherapy.
- Disutility scenarios (removing AE related disutity and age-adjusted disutility).
- Exploring the assumption of vial sharing (i.e., no wastage).
- Removing half-cycle correction.

Following clarification questions, the revised results are provided in Table 5 of the company's response to clarifiction questions, Appendix C. Following the requests at the clarification stage, additional scenarios were explored in relation to clarification questions B11, B17, B19 and B21:

- Using a fully parametric model for PFS curves (log-logistic).
- Removing treatment stopping rules.
- Administration costs based on a day case setting.
- Distribution of subsequent treatments based on PFS events.

Based on the company's presented scenarios versus the trial comparator (company's response to clarifiction questions, Appendix C Table 5), the scenario with the largest impact was assuming all patients after chemotherapy receive nivolumab. As this added a large increase in costs to the comparator arm, the ICER was reduced to £8,318. The scenarios which resulted in the highest ICER were due to the alternative OS parametric distribution (Weibull) and alternative OS cut-off point for the piece-wise modelling (32-weeks) resulting in ICERs of £71,729 and £52,790, respectively.

The additional scenarios requested by the ERG at clarification stage are presented in Table 21. All demonstrated minor impacts on the company's base case ICER.

Table 21: Additional scenarios post clarification questions

Scenario	Description	Pembrolizumab in combination with chemotherapy			Chemotherapy			Incremental		
		Total costs	Total LYs	Total QALYs	Total costs	Total LYs	Total QALYs	Costs	QALYs	ICER
Versus trial comparator										
Base Case	-	[redacted]	2.21	[redacted]	[redacted]	1.39	[redacted]	£27,172	0.65	£41,688
B11	Fully fitted parametric modelling approach for PFS using log-logistic distribution	[redacted]	2.21	[redacted]	[redacted]	1.39	[redacted]	£27,130	0.65	£41,623
B17	Removing treatment stopping rules	[redacted]	2.21	[redacted]	[redacted]	1.39	[redacted]	£27,396	0.65	£42,032
B19	Drug administration costs occurring in day-case setting	[redacted]	2.21	[redacted]	[redacted]	1.39	[redacted]	£27,402	0.65	£42,041
B21	Alternative subsequent therapy approach (PFS events)	[redacted]	2.21	[redacted]	[redacted]	1.39	[redacted]	£27,280	0.65	£41,854
Versus blended comparator*										
Base Case	-	[redacted]	2.21	[redacted]	[redacted]	1.39	[redacted]	£26,988	0.65	£41,405
B12	AEs based on Yoon 2016	[redacted]	2.21	[redacted]	[redacted]	1.39	[redacted]	£27,834	0.65	£43,069
	AEs based on Cleary 2019	[redacted]	2.21	[redacted]	[redacted]	1.39	[redacted]	£27,641	0.65	£42,688
	AEs based on Waddell 2013	[redacted]	2.21	[redacted]	[redacted]	1.39	[redacted]	£27,692	0.65	£42,797

Key: AEs, adverse events; ICER, incremental cost-effectiveness ratio; QALYs, quality-adjusted life-years; PFS, progression-free survival

Notes: *Weighted costs assuming equal market share (~12.5% for each treatment)

5.3. Model validation and face validity check

The company internally validated clinical outcomes from the model with what was observed in KEYNOTE-590 (CS Appendix J). The ERG noted some discrepancy between the modelled PFS values compared to the KEYNOTE-590[12] observed data, particularly for the pembrolizumab in combination with chemotherapy arm which appeared to be underestimated in the model. At 6 months, the model predicts PFS to be 56.2% when observed data from KEYNOTE-590[12] is actually 62.4%. At two years, PFS is 8.4% versus 11.8% for the model outcomes versus the KEYNOTE-590.[12] The modelled OS outcomes look reasonable compared to the KEYNOTE-590 trial data.

In additon to internal validation checks, the company stated that the modelling approach was validated by clinical experts, however, no information on how this validation step was conducted was provided in the CS. As such, the ERG reqested information at clarifiation stage. The company confirmed that separate informal interviews were conducted with four clinical oncologists working in the treatment of oesophageal cancer and held an advisory board on the 29th January. However, outputs of the advisory board were not used within the CS due to the close proximity to the submission date. No further information was shared by the company on the questions asked or topics discussed within the informal interviews stating that *"Due to the informal nature of the interviews with clinical experts, MSD consider that it would not be appropriate to share the outputs of these interviews"* (see company's response to clarification questions B23), therefore the ERG was not able to assess whether the clinical opinion sought was fairly executed.

The company also had the model validated through a comprehensive quality check by the economists who developed the model and by an external vendor who the company stated found no implementation errors or bugs.

The results of the model could not be compared to any publications as no studies assessing the cost-effectiveness of pembrolizumab in combination with chemotherapy versus standard of care were identified in the systematic literature review. The ERG replicated the company's model using a simple 'back of the envelope' type model in Excel and pasted values where necessary (e.g., survival curves) and was able to replicate the company's base case results. However, during this exercise, the ERG noted several errors within the model calculations:

- Firstly, when half-cycle correction is not applied in the company's model, the proportion of patients in each health state per cycle is moved to the next cycle (i.e., patients start the model in the progression-free health state at Cycle 1 (seven days) instead of Cycle 0. This misaligns the annual discount rate applied.
- Secondly, the way the company has calculated the life-years in the time-to-death health states is incorrect as they include those patients who die within that cycle. This impacts the QALYs which are accrued over time.

These errors only have minor impact on the results and are corrected in the ERG's base case (see Section 6.1).

6. EVIDENCE REVIEW GROUP'S ADDITIONAL ANALYSES

The ERG identified a number of limitations within the company's base case and explored the impact of parameter values, and assumptions, which the ERG believed were more plausible.

This section is organised as follows: Section 6.1 details the impact of errors identified in the ERG's validation of the executable model. Section 6.2 details a series of scenario analyses exploring the robustness of the cost-effectiveness results to specific assumptions and additional uncertainties identified by the ERG. These analyses were conducted within the company corrected base-case analysis.

The scenario analyses presented in Section 6.2 focus on exploring the following issues and uncertainties:

- Exploring progression-based utilities
- Alternative PFS and OS extrapolations
- Exploring efficacy of triplet therapy versus doublet therapy
- Exploring chemotherapy regimens based on UK clinical practice
    - Market share distributions based on clinical exert opinion
    - Alternative doses for some chemotherapies
- Removing half-cycle correction
- Removing treatment stopping rules
- Exploring treatment based monitoring
- Alternative adverse event costs
- Use of all subsequent treatment data from KEYNOTE-590
- Alternative terminal care costs

In Section 6.3, the ERG base-case is presented based on a combination of the exploratory analyses presented in Section 6.2.

6.1. ERG corrections and adjustments to the company's base case model

A small number of errors were identified during the face validity check of the cost-effectiveness model relating to the application of half-cycle correction and time-to-death utilities. These are described in detail within Section 5.3 and below.

- When half-cycle correction is switched-off in the company's model, the proportion of patients in each health state per cycle is moved to the next cycle (i.e., patients start the model in the progression-free health state at Cycle 1 (seven days) instead of Cycle 0. This mis-aligns the annual discount rate applied.
- The way the company has calculated the life-years in the time-to-death health states is incorrect as they include those patients who die within that cycle. Hence these patients are accruing utilities within the death health state. This impacts the QALYs which are accrued over time.

The ERG implemented the corrections within the company's economic model. The correction to the half cycle correction application only applies when half cycle correction is switched off. As such, this correction does not impact the company's base case.

Table 22 and Table 23 present the correct company base case for the full population and CPS ≥10 sub-population, respectively.

Table 22: ERG-corrected company base case results – all patients

	Discounted costs	Discounted QALYs	Incremental discounted costs	Incremental discounted QALYs	Cost per QALY gained
ERG corrected company deterministic base case					
Pembrolizumab + chemotherapy	[redacted]	[redacted]	-	-	-
5-FU + cisplatin	[redacted]	[redacted]	£27,173	0.65	£41,701
ERG corrected company probabilistic base case					
Pembrolizumab + chemotherapy	[redacted]	[redacted]	-	-	-
5-FU + cisplatin	[redacted]	[redacted]	£27,085	0.65	£41,669

Abbreviations: QALYs, quality adjusted life years

Table 23: ERG-corrected company base case results – CPS ≥10

	Discounted costs	Discounted QALYs	Incremental discounted costs	Incremental discounted QALYs	Cost per QALY gained
ERG corrected company deterministic base case					
Pembrolizumab + chemotherapy	[redacted]	[redacted]	-	-	-
5-FU + cisplatin	[redacted]	[redacted]	£30,293	0.92	£33,006
ERG corrected company probabilistic base case					
Pembrolizumab + chemotherapy	[redacted]	[redacted]	-	-	-
5-FU + cisplatin	[redacted]	[redacted]	£30,122	0.93	£32,526

Abbreviations: QALYs, quality adjusted life years

6.2. Exploratory and sensitivity analyses undertaken by the ERG

The ERG conducted a number of additional sensitivity analyses within the company's model, which are described in turn within each section below.

6.2.1. Overall survival

Given the broad range of options available within the model to inform OS, the ERG has focused on four key scenarios to model OS for both treatment arms. These are described in further detail within Section 4.2.6.1. Clinical advice to the ERG was that each of these four scenarios may be considered as broadly clinically-plausible but are not possible to robustly validate given that no long-term data are currently available for the use of pembrolizumab in combination with chemotherapy in this patient population.

Plots demonstrating the difference in projections for each of these models are provided in Table 24, alongside estimates of OS at key time points in the longer-term up until the end of the model time horizon. The corresponding impact of these extrapolations of OS on the ICER is provided within Section 6.2.8.

Table 24: Comparison of four scenarios considered for overall survival

[redacted]

[redacted]

6.2.2. Progression-free survival

Similar to OS, the ERG has considered a number of alternative extrapolations for PFS within the model. The ERG has chosen to focus on four scenarios:

1. The company's base-case analysis.
2. Changing the cut-point to 37 weeks.
3. Changing the extrapolated tail to generalised gamma.
4. Changing both the cut-point to 37 weeks and the extrapolated tail to generalised gamma.

6.2.3. Utilities

Section 4.2.7.4 describes the utility values used within the model, based on two approaches: a progression-based approach, and a time-to-death based approach. The ERG explored additional analyses varying the absolute health state utility values by subtracting 10% of their base values to explore the impact on results. This exploratory scenario was considered because of the utility values appearing relatively high relative to the general population, but is by definition an arbitrary variation of the KEYNOTE-590-derived utility values.

In addition, the ERG sought to identify any utility values identified by the company as part of its SLR that could be applied within the model. Based on the CS (Appendix H, Table 22), the only study that reported values either as a function of the time to death or by progression status was a study by Zhang et al. (2020).[40] While the study by Zhang *et al*. was based on the ATTRACTION-3 study of nivolumab versus chemotherapy for patients with advanced oesophageal squamous cell carcinoma refractory or intolerant to previous chemotherapy, the utility values themselves were taken from a different study by Saito et al. (2017)[66] – a cost-utility analysis of paclitaxel + ramucirumab for advanced gastric cancer. The study by Saito *et al.* cited utility values of 0.741 for progression-free disease, and 0.581 for progressed disease, but these were taken from two other studies – a study by Al-Batran *et al*. (2016)[67] and NICE TA378[61] of ramucirumab for advanced gastric cancer or gastro-oesophageal junction adenocarcinoma previously treated with chemotherapy. The ERG accepts these values are subject to several important limitations (including, but by no means limited to, the difference in disease area, potential concerns around generalisability, and a lack of reported information concerning the derivation of the utility values themselves). Nevertheless, as these were the only non-

KEYNOTE-590 utility values identified by the company that were possible to consider within the submitted model and acknowledging the ERG's prior comments concerning the magnitude of the utility values derived from KEYNOTE-590,[12] an exploratory analysis was conducted to apply these utility values within the model for both arms.

6.2.4. Efficacy of doublet chemotherapy versus triplet chemotherapy

As discussed in Section 3.5.2 and 4.2.4, the company assumes equivalent efficacy between doublet regimens and triplet regimens based on evidence from the NICE Guideline in the assessment and management of oesophago-gastric cancer in adults (NG83)[7] and clinical opinion. This evidence was used to justify the use of the comparator arm from KEYNOTE-590[12] to inform the efficacy of the chemotherapy arm in the model regardless of treatment regimen selected. In order to explore the sensitivity of this assumption, the ERG looked at scenarios whereby the triplet efficacy was estimated using results from the NMA reported in ter Veer *et al.*[10] which explored the efficacy and safety of first-line chemotherapy in advanced oesophageal cancer using an NMA. The NMA compares 'ACF' vs 'CF' (A=anthracycline, C=cisplatin, F=5FU) and reports a HR for OS and PFS; OS HR = 0.86 (95% CI: 0.71 to 1.02) and for PFS, HR = 0.85 (95% CI: 0.68 to 1.05). The confidence intervals cross 1 showing non- statistical differences and the authors conclude that "anthracycline-containing triplets…were not more effective than F-doublets" but notes they were associated with increased toxicity compared to doublets. As such, the ERG note that these scenarios are limited based on the evidence available and technical application so ICERs should be viewed with caution, however these are deemed necessary to explore the uncertainty associated with efficacy differences between triplet and doublet regimens.

In the ERG's analysis, these HRs were applied to the doublet OS and PFS curves from KEYNOTE-590 and using the blended comparator arm, the resulting OS and PFS curves were weighted based on the proportion of triplets and doublets.

Figure 15 and Figure 16 present the company's base case OS and PFS curves compared to the estimated triplet chemotherapy OS and PFS curves using the NMA HRs, respectively.

Figure 15: Company's base case OS curves with the estimated triplet OS curve

Key: OS, overall survival

Figure 16: Company's base case PFS curves with the estimated triplet PFS curve

Key: OS, overall survival

Using these curves for the triplet regimens, the ERG explored ICERs based on the blended chemotherapy arm with the following assumptions:

- Using the company's estimated market shares (12.5% per regimen resulting in a mix of 37.5% doublets versus 62.5% triplets).
- Using UK estimated market shares (see Section 6.2.5.2, resulting in a mix of 68.8% doublets versus 31.3% triplets).
- Pairwise comparisons versus each triplet therapy individually.

In all scenarios, time on treatment data was assumed to equivalent to the data from KEYNOTE-590[12] due to lack of data to inform otherwise. The resulting ICERs ranged from £46,832 to £68,512 which is an increase of between £5,131 to £26,811 compared to the company's corrected base case.

6.2.5. Chemotherapy regimens

6.2.5.1. UK based chemotherapy regimen

As discussed in Section 4.2.4, the chemotherapy regimen included in KEYNOTE-590[12] and subsequently used to form the company's base case is rarely used in UK clinical practice. Clinical advice provided to the ERG suggested that capecitabine plus oxaliplatin was more commonly used out of the chemotherapy regimens available. The company provided options and scenarios in the model which changed the comparator arm to each individual chemotherapy regimen however, the platinum-based chemotherapy in combination with pembrolizumab remained as 5-FU plus cisplatin. The proposed license states "KEYTRUDA, in combination with platinum and fluoropyrimidine based chemotherapy" (CS Appendix C) which is therefore not specific to 5-FU and cisplatin. However, clinical advice provided to the ERG stated that as the evidence for pembrolizumab is specifically with 5-FU and cisplatin, then there could be a change in practice with more cisplatin and 5-FU use, unless NICE guidance is clear that oxaliplatin can be substituted. Nevertheless, the ERG included additional functionality in the model to change pembrolizumab's combination chemotherapy to oxaliplatin plus capecitabine as per clinical opinion. Amending the chemotherapy regimen to capecitabine plus oxaliplatin (both in combination with pembrolizumab and as the comparator) resulted in an ICER of £42,400 per QALY gained which is a slight increase compared to the company's ICER.

6.2.5.2. UK based market shares for the blended comparator

In the company's comparison to the 'blended chemotherapy arm', the company assumed an equal market share between treatments which was considered by the ERG to be implausible and not reflective of the treatments given in clinical practice (see Section 4.2.4). The ERG requested market shares from clinical experts to explore more clinically plausible options.

Table 25 presents the expected usage in clinical practice versus the usage assumed by the company. Although expected usage varies and can be difficult to estimate, there is a general consensus that capecitabine + oxaliplatin is most commonly used, with a small usage of cisplatin (instead of oxaliplatin) and 5-FU (instead of capecitabine) and still a proportion using triplets instead of doublets.

Table 25: Market shares of chemotherapy regimens

Treatment regimen	Company base case	Expected usage
Capecitabine + oxaliplatin	12.5%	60%
Capecitabine + oxaliplatin + epirubicin	12.5%	15%
Capecitabine + cisplatin	12.5%	5%
Capecitabine + cisplatin + epirubicin	12.5%	5%
5-FU + cisplatin	12.5%	3.75%
5-FU + oxaliplatin + epirubicin	12.5%	3.75%
5-FU + cisplatin + epirubicin	12.5%	3.75%
5-FU + oxaliplatin + leucovorin	12.5%	3.75%

5-FU, five fluorouracil

Using the expected usage increases the ICER to £41,853 per QALY gained (see Table 26).

6.2.5.3. UK based chemotherapy dosing

In addition to including the most appropriate UK based chemotherapies, the ERG explored alternative dosing based on clinical expert opinion. As discussed in Section 4.2.4, clinical experts advised the ERG that some of the chemotherapy dosing schedules are slightly different to those commonly used in UK practice, with cisplatin usually given at a dose of 60 mg/m^2 for up to 6 to 8 cycles. In addition, the two-day infusion of 5-FU is considered the standard of care in UK clinical practice instead of the five-day infusion used in KEYNOTE-590. However, the clinical experts confirmed that the efficacy of 5-FU would not be impacted by these dosing differences. The ERG also considers the administration cost code used in the company base case (SB14Z) to still apply to 5-FU based regimens so no changes are required. The ERG amended the dose

of cisplatin to be 60 mg/m^2 (instead of 80 mg/m^2 used by the company). This change had minimal impact on the results (see Table 26), nevertheless this reflects UK clinical practice more than the company's base case.

6.2.6. Half cycle correction

The company's model cycle length of one week does not warrant the use of half cycle correction, therefore the ERG has explored the impact of removing this. As discussed in Section 5.3 and Section 6.1, the ERG noted an error when half cycle correction is removed which mis-aligns the annual discount rate applied. This error has been fixed in this scenario.

6.2.7. Resources and costs

6.2.7.1. Treatment stopping rules

The company included treatment stopping rules which caps treatment costs at a certain time points in addition to using ToT Kaplan-Meier data estimated directly from KEYNOTE-590.[12] The ERG noted in Section 4.2.8.3 that the ToT data from KEYNOTE-590[12] already incorporates the protocol driven stopping rules, therefore is not necessary to apply the maximum treatment durations in addition to using the Kaplan-Meier estimates and RDI. This is demonstrated in [redacted]10 and [redacted]11. At clarification stage, the ERG requested the company to provide a scenario in which only the ToT Kaplan-Meier estimates were used to inform treatment costs with the removal of treatment stopping rules. The company provided this scenario (see company's response to clarification questions B17), which slightly increases the ICER to £42,045 per QALY gained.

6.2.7.2. Administration costs

The company included administration based on the outpatient setting in their base case; however, as discussed in Section 4.2.8.2, clinical advice provided to the ERG suggested that administration would be given in a day case setting. At clarification stage, the ERG requested the company to provide a scenario in which administration costs were based on the day case setting. The company provided this scenario (see company's response to clarification questions B19), which slightly increases the ICER to £42,054 per QALY gained.

6.2.7.3. Treatment specific monitoring

The CS based the disease monitoring costs on progression status (i.e., progression-free or progressed). Clinical experts stated that monitoring frequency would differ by treatment and whether patients were on treatment or had discontinued (see Section 4.2.8.4). The ERG performed exploratory analysis which amended the progression-free monitoring based on treatment status, i.e., patients are assumed to be monitored every three weeks whilst on platinum-based chemotherapy (e.g., cisplatin) then every three months while continuing treatment with a fluoropyrimidine (e.g., fluorouracil). If patients are still receiving pembrolizumab after discontinuation of platinum-based chemotherapy, then monitoring would be every six weeks. For those patients who discontinued all treatments but remain progression-free, the ERG assumed disease monitoring was the same as the progressed disease state which costs a consultation visit every 12 weeks. The company's progression-based disease monitoring assumes patients are monitored every three with a consultation visit every four weeks which in comparison to the treatment-based monitoring assumes more resource use. Therefore, despite the pembrolizumab in combination with chemotherapy arm having increased frequencies in the treatment-based monitoring compared to chemotherapy, the overall disease monitoring costs are reduced in both arms in this scenario and as such applying the treatment based monitoring reduces the ICER to £41,173.

6.2.7.4. Subsequent treatments

The company applied an arbitrary cut-off of excluding all subsequent treatments received by less than 5% of patients received. As discussed in Section 4.2.8.6, after reviewing the full subsequent treatment table, the ERG noted that applying the 5% cut-off results in an underestimation of subsequent treatment costs and the unnecessary removal of data. In order to account for all subsequent treatments received by patients in KEYNOTE-590,[12] the ERG have re-distributed the remaining treatments into the most common treatments received (using the company's 5% threshold). The redistribution and resulting costs are presented in Table 16 compared to the company's estimates. Re-distributing the subsequent treatments including all incidences reduces the ICER to £41,434.

6.2.7.5. Terminal care costs

The ERG noted that the source of the terminal care cost has been derived from a chain of previous submissions and thus does not take into consideration the assumptions surrounding

the individual resources and whether they are appropriate for gastro-oesophageal cancer patients (see Section 4.2.8.6). The ERG performed some exploratory analysis using different terminal care values.

First, the ERG removed the radiotherapy cost which was included in the previous TA519[43] submission specifically for urothelial cancer and may not be appropriate for gastro-oesophageal cancer patients. The cost used in TA519 for radiotherapy was £3,232.43. The ERG removed this from the company's terminal care cost before this was uplifted to 2020 values, resulting in a cost of £4,320.93. This scenario increased the ICER to £41,864.

Another scenario exploring the cost used in the recent nivolumab appraisal (ID1249)[46] which used £8,973.61 from the literature (inflated from £7,827.00) estimating the per-patient costs in the last three months of life.[65] This scenario reduced the ICER to £41,646.

6.2.8. Impact on the ICER of additional clinical and economic analyses undertaken by the ERG

The ERG made the changes described in Sections 6.2.3 to 6.2.7. Each change has been made individually. The results of the ERG's exploratory analyses are provided in Table 26.

The majority of the scenarios when considered in isolation had only a minor impact on the ICER. The key scenarios conducted were the alternative extrapolations to OS (increasing the ICER by between £4,646 to £32,342), exploring efficacy of triplet regimens (increasing the ICER by between £5,908 to £26,811) and different utility options (increasing the ICER by between £3,249 to £12,248).

Table 26: ERG's exploratory analyses

Preferred assumption	Section in ERG report	Pembrolizumab in combination with chemotherapy		Chemotherapy		Incremental		ICER £/QALY	+/- company base case
		Total costs	Total QALYs	Total costs	Total QALYs	Costs	QALYs		
ERG corrected company base-case	*6.1*	[redacted]	[redacted]	[redacted]	[redacted]	£27,173	0.65	£41,701	-
OS: Assume treatment waning effect applies between 5 and 7 years	*6.2.1*	[redacted]	[redacted]	[redacted]	[redacted]	£27,128	0.59	£46,347	+£4,646
OS: Single log-logistic parametric model		[redacted]	[redacted]	[redacted]	[redacted]	£26,970	0.36	£74,043	+£32,342
OS: Change to generalised gamma tail		[redacted]	[redacted]	[redacted]	[redacted]	£27,067	0.50	£54,447	+£12,746
PFS: Change cut-point to 37 weeks	*6.2.2*	[redacted]	[redacted]	[redacted]	[redacted]	£27,792	0.65	£42,653	+£952
PFS: Change to generalised gamma tail		[redacted]	[redacted]	[redacted]	[redacted]	£27,174	0.65	£41,703	+£2
PFS: Change cut-point to 37 weeks and to generalised gamma tail		[redacted]	[redacted]	[redacted]	[redacted]	£27,134	0.65	£41,643	-£58
Utilities: KEYNOTE-590 progression-based utility values	*6.2.3*	[redacted]	[redacted]	[redacted]	[redacted]	£27,172	0.57	£47,661	+£5,960
Utilities: Reduce magnitude of all health state utility values by 10%		[redacted]	[redacted]	[redacted]	[redacted]	£27,172	0.60	£44,950	+£3,249
Utilities: Apply published utility values (by progression status)		[redacted]	[redacted]	[redacted]	[redacted]	£27,172	0.50	£53,949	+£12,248
Triplet efficacy vs doublet efficacy – company market share	*6.2.4*	[redacted]	[redacted]	[redacted]	[redacted]	£26,690	0.52	£51,394	+£9,693
Triplet efficacy vs doublet efficacy – UK expected market share		[redacted]	[redacted]	[redacted]	[redacted]	£27,107	0.58	£46,832	+£5,131

Preferred assumption	Section in ERG report	Pembrolizumab in combination with chemotherapy		Chemotherapy		Incremental		ICER £/QALY	+/- company base case
		Total costs	Total QALYs	Total costs	Total QALYs	Costs	QALYs		
Triplet efficacy vs doublet efficacy – 5-FU + cisplatin + epirubicin		[redacted]	[redacted]	[redacted]	[redacted]	£26,520	0.39	£68,512	+£26,811
Triplet efficacy vs doublet efficacy – 5-FU + oxaliplatin + epirubicin		[redacted]	[redacted]	[redacted]	[redacted]	£26,478	0.39	£68,403	+£26,702
Triplet efficacy vs doublet efficacy – capecitabine + oxaliplatin + epirubicin		[redacted]	[redacted]	[redacted]	[redacted]	£26,398	0.39	£68,198	+£26,497
Triplet efficacy vs doublet efficacy – capecitabine + cisplatin + epirubicin		[redacted]	[redacted]	[redacted]	[redacted]	£26,441	0.39	£68,307	+£26,606
Pembrolizumab in combination with capecitabine plus oxaliplatin versus capecitabine plus oxaliplatin	*6.2.5.1*	[redacted]	[redacted]	[redacted]	[redacted]	£27,628	0.65	£42,400	+£699
Blended comparator based on UK expected market shares	*6.2.5.2*	[redacted]	[redacted]	[redacted]	[redacted]	£27,271	0.65	£41,853	+£152
Cisplatin dosed as 60 mg/m^2	*6.2.5.3*	[redacted]	[redacted]	[redacted]	[redacted]	£27,173	0.65	£41,702	+£1
Remove half-cycle correction	*6.2.6*	[redacted]	[redacted]	[redacted]	[redacted]	£27,172	0.65	£41,691	-£10
Remove treatment stopping rules	*6.2.7.1*	[redacted]	[redacted]	[redacted]	[redacted]	£27,396	0.65	£42,045	+£344
Administration based on the day case setting	*6.2.7.2*	[redacted]	[redacted]	[redacted]	[redacted]	£27,402	0.65	£42,054	+353
Include treatment-based monitoring	*6.2.7.3*	[redacted]	[redacted]	[redacted]	[redacted]	£26,829	0.65	£41,173	-£528
Re-distribute subsequent treatments	*6.2.7.4*	[redacted]	[redacted]	[redacted]	[redacted]	£26,998	0.65	£41,434	-£267
Alternative terminal care costs	*6.2.7.5*								
- Removing radiotherapy		[redacted]	[redacted]	[redacted]	[redacted]	£27,279	0.65	£41,864	+£163
- Based on ID1249						£27,136	0.65	£41,646	-£55

Key: ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; OS, overall survival; PFS, progression-free survival; QALY, quality adjusted life year

6.3. ERG's preferred assumptions

The ERG's preferred base case analysis comprises several alternative model settings and assumptions which are discussed in Section 6.2. The cumulative impact of these changes is presented in Table 27 with the final base case presented in

Table 28 compared to the company's base case. The ERG preferred base case ICER is £51,921.

Although the ERG's preferred extrapolation for OS is to use the company's approach with the treatment waning adjustment, the ERG would like to highlight that the other OS extrapolation scenarios listed in Section 4.2.6.1 and Section 6.2.1 are all considered plausible. Therefore, considering all of these the most plausible ICER (incorporating other ERG preferred settings) lies between £47,270 to £77,722.

Table 27: ERG's preferred model assumptions – all patients

Preferred assumption	Section in ERG report	Cumulative ICER £/QALY
ERG-corrected company base-case	6.1	£41,701
Remove half cycle correction	6.2.6	£41,691
Administration costs using a day case setting	6.2.7.2	£42,044
Turning off stopping rules for treatments (i.e., just using the ToT KM estimates from KEYNOTE-590)	6.2.7.1	£42,394
Re-distributing subsequent treatments	6.2.7.4	£42,100
Progression-based utilities	6.2.3	£48,108
PFS piecewise using 37-week cut-off and log-logistic extrapolation	6.2.2	£47,270
Include treatment waning between 5-7 years	6.2.1	£51,921

Key: ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life year

Table 28: Comparison of company's and ERG's preferred base case – all patients

	Discounted costs	Discounted QALYs	Incremental discounted costs	Incremental discounted QALYs	Cost per QALY gained
Company base case (deterministic)					
Pembrolizumab + chemotherapy	[redacted]	[redacted]	-	-	-
5-FU + cisplatin	[redacted]	[redacted]	£27,173	0.65	£41,701
ERG base case (deterministic)					
Pembrolizumab + chemotherapy	[redacted]	[redacted]	-	-	-
Chemotherapy	[redacted]	[redacted]	£28,007	0.54	£51,921

Key: QALYs, quality adjusted life years

6.4. Conclusions of the cost-effectiveness section

The company's model is appropriate for decision making

The company's PartSA model is considered appropriate for decision making and consistent with previous NICE submissions in similar disease areas. Overall, the ERG found the company's model to be clear and well-constructed. Where the ERG identified errors, resolving these had little influence on the estimated ICER.

The systematic literature review was satisfactory; however, there was no discussion of the applicability of the identified study to the economic model within the CS.

The ERG was satisfied with the company's review of the cost-effectiveness literature. The ERG agreed with the company's judgment that none of the studies identified were relevant to the UK population. The ERG was broadly satisfied with the company's review of the literature reporting health effects (HRQoL and utilities), health care resource use, and costs. The ERG noted an absence of methodological reporting for screening and data extraction regarding health effects. While no formal critical appraisal of utility studies was conducted, the company provided an assessment of the consistency of each study with the reference case. The ERG noted that none of the studies identified in the review of utilities were used in the model and no discussion of the applicability of the one identified study for health care resource use. However, the ERG was satisfied that the incorporation of utilities data from KEYNOTE-590[12] into the model was

appropriate to inform the base-case analysis and was generally satisfied with the sources used for resource use.

The generalisability of KEYNOTE-590 to UK patients is unclear

Over half of the KEYNOTE-590[12] study population were from Asia (52.5%, versus 47.5% from the ROW), and region was shown to have an apparent impact on the HR for OS. The ERG considered the high proportion of patients from Asia was not reflective of the UK patient population and had concerns with the impact this appears to have on OS. The ERG requested that the company provide a scenario analysis removing Asian patients, however the company declined to provide this subgroup analysis. Therefore, the ERG was unable to consider any further analysis for the ROW population specifically.

The ERG also noted the histology split between adenocarcinoma and squamous cell carcinoma in KEYNOTE-590[12] (26.8% adenocarcinoma versus 73.2% squamous cell carcinoma) was the opposite of the proportionate split expected within the UK population. Histology is an important factor given the differences in disease and potential treatment.

The comparator treatment given in KEYNOTE-590 was not considered the most reflective of current NHS practice

The main comparator considered by the company in its economic model was per the comparator used within the KEYNOTE-590[12] study (5-FU + cisplatin). This was considered by the ERG to not reflect the most common chemotherapy regimen used within NHS practice. In addition, the ERG found the company's approach to reflect NHS practice including multiple chemotherapy doublets and triplets to be inadequate by assuming equal market share for all possible alternatives. However, the ERG acknowledged that the company ran a scenario analysis amending the comparator arm to each of the chemotherapy regimens individually to investigate the impact of comparator therapies.

Clinical advice provided to the ERG noted that of the chemotherapy regimens included within the company's model, not all are used in NHS practice, and by extension do not have equal market shares. Based on advice provided to the ERG, the main chemotherapy used in practice is capecitabine plus oxaliplatin. Thus, the ERG considered the KEYNOTE-590[12] comparator regimen was not the most relevant comparator for this decision problem. The ERG accepted the company's approach of using the KEYNOTE-590[12] efficacy to inform the chemotherapy OS and

PFS within the economic model; however, costs based on the trial comparator do not reflect standard NHS practice.

Estimation of OS is a key driver of cost-effectiveness

Clinical advice to the ERG was that the company's base-case extrapolation was plausible, but that several alternative extrapolations were also plausible. The ERG's base-case used the same extrapolation per the company's base-case analysis with an adjustment for the long-term extrapolation after five years. This adjustment assumes that between five and seven years, the projected hazard of death for the pembrolizumab in combination with chemotherapy arm gradually tends to that of the chemotherapy arm. Hence, from seven years onwards, the projected hazard of death is assumed equal between arms. It was not possible for the ERG to assess with available data the plausibility of a lifetime treatment effect, or a treatment effect that would eventually dissipate by seven years. The choice of OS model remains a key uncertainty of the cost-effectiveness analysis, and considers that a range of scenarios may be informative for decision making.

Concerns were identified concerning the generalisability of the utilities derived from KEYNOTE-590, in particularly using a time-to-death approach

The ERG had concerns with the generalisability of the utility values produced based on analysis of KEYNOTE-590[12] data (regardless of which approach is used), as the outputted values imply that patients have a similar, or potentially better utility than the age- and sex-adjusted UK general population. The ERG was concerned that the two approaches to utility analysis lead to a substantially different estimation of the "average" utility experienced over the course of the model time horizon. This meant that the incremental QALY gain attributable to pembrolizumab in combination with chemotherapy estimated for both utility analyses also varied markedly. The ERG considered the progression approach to yield a more realistic "average" utility for this patient population, especially given that the time-to-death approach yields an "average" utility that is close to the estimate for the general population.

It is inappropriate to justify cost inputs based predominantly on prior company submissions of pembrolizumab in other disease areas

The majority of the company's model cost inputs were justified on the basis of being used in previous company submissions of pembrolizumab in different advanced cancer populations.

Although ultimately no major concerns were identified with the values used, it would be remiss of the ERG not to highlight the shortcomings of this approach to identifying model inputs. The ERG would have preferred that values be identified systematically, including reference where necessary to submissions made by different companies in similar disease areas (i.e., not restricting to those made only by the submitting company for pembrolizumab). Should values be taken from previous company submissions, appropriate clinical validation should be undertaken, and amendments be made as required (with justification presented). The ERG has attempted to correct for some differences in scenario analysis based on expert opinion or flagged the impact on the ICER, however as previously highlighted, no major issues were identified.

The majority of subsequent treatment instances were excluded from the company's calculations

The ERG highlighted concerns with the application of the 5% cut-off within the subsequent treatment costing resulting in [redacted] of subsequent treatments instances being excluded from the model. This results in an underestimation of subsequent treatment costs and the unnecessary removal of data. Consequently, the ERG re-distributed the remaining treatments into the most common treatments received (using the company's 5% threshold) within its preferred assumptions.

The ERG's preferred base case analysis yields an ICER slightly greater than the company's base case ICER and is just over the £50,000 per QALY gained threshold

The ERG's preferred base case analysis includes alternative OS and PFS assumptions, a different utility approach, the removal of half cycle correction and treatment stopping rules, using all subsequent treatment usage and assuming day case setting for administration. When combined, these changes result in larger total costs and fewer QALYs, causing an increase in the ICER from £41,701 to £51,921. The ERG highlights that other OS extrapolation scenarios listed in Section 4.2.6.1 and 6.2.1 are considered plausible. Considering these alternative OS extrapolations, the most plausible ICER (incorporating other ERG preferred settings) lies between £47,270 to £77,722. Accordingly, assuming a willingness-to-pay threshold of £50,000 per QALY gained, there is uncertainty as to whether pembrolizumab in combination with chemotherapy would be a cost-effective use of NHS and PSS resources.

7. END OF LIFE

The company stated that pembrolizumab in combination with chemotherapy meets end of life criteria for this indication, and summarises the basis for this assertion in Table 40 of the CS with respect to the ITT and CPS≥10 populations. The ERG regarded that the company's representations were generally appropriate with respect to the whole trial population, but noted that the strength of evidence was greater for the CPS ≥10 population, and noted that specific evidence for the rest of world subgroup did not substantiate the required increase in life expectancy.

The company noted that in the ITT population, the difference in median OS was 2.6 months, less than the three months required, though the estimated difference in mean months from the economic model was 7.5 months. However, in the rest of world subgroup specifically, the difference in median OS was approximately [redacted] (clarification response appendices, Table 3). For the CPS≥10 population, both the difference in median OS (4.1 months) and the difference in model-estimated means (10.6 months) were above the requisite threshold, but these were not presented for the rest of world subgroup.

Moreover, in the KEYNOTE-590[12] trial, the ITT population that received standard of care had median survival of 9.8 months, whereas the CPS≥10 population that received standard of care had median survival of 9.4 months. This suggests that with respect to the entire trial population, the short life expectancy criterion was met. In the rest of world subgroup, the median OS for the standard of care arm was 44.2 weeks for the ITT population; specific rest of world estimates for the CPS≥10 population and this subgroup were not provided.

References

1. Rashid N, Elshaer M, Kosmin M, Riaz AA. Current management of oesophageal cancer. British Journal of Medical Practitioners. 2015;8(1):15-24.
2. Gatenby PAC, Hainsworth A, Caygill C, Watson A, Winslet M. Projections for oesophageal cancer incidence in England to 2033. European Journal of Cancer Prevention. 2011;20(4):283-6.
3. Cancer Research UK. Oesophageal cancer. Risks and causes, 2019. Available from: https://www.cancerresearchuk.org/about-cancer/oesophageal-cancer/causes-risks.
4. National Health Service (NHS). What is oesophageal cancer? 2020. Available from: https://www.nhs.uk/conditions/oesophageal-cancer/.
5. World Cancer Research Fund. Oesophageal cancer, 2020. Available from: https://www.wcrf-uk.org/uk/preventing-cancer/cancer-types/oesophageal-cancer.
6. Cancer Research UK. Oesophageal cancer. Survival, 2019. Available from: https://www.cancerresearchuk.org/about-cancer/oesophageal-cancer/survival.
7. National Institute for Health and Care Excellence (NICE). Oesophago-gastric cancer: Assessment and management in adults (NG83), 2018. Available from: http://www.nice.org.uk/guidance/ng83.
8. National Institute for Health and Care Excellence (NICE). Pembrolizumab with platinum-based chemotherapy for untreated advanced oesophageal or gastroesophageal junction cancer - Final Scope, 2018.
9. Cunningham D, Starling N, Rao S, Iveson T, Nicolson M, Coxon F, et al. Capecitabine and oxaliplatin for advanced esophagogastric cancer. New England Journal of Medicine. 2008;358(1):36-46.
10. ter Veer E, Haj Mohammad N, van Valkenhoef G, Ngai LL, Mali RMA, Anderegg MC, et al. The efficacy and safety of first-line chemotherapy in advanced esophagogastric cancer: A network meta-analysis. JNCI: Journal of the National Cancer Institute. 2016;108(10).
11. Wang C, Song DJ, Xu ZL, Xie SP, Hu JH. A network meta-analysis of the short-term efficacy of five chemotherapy regimens based on cisplatin and fluorouracil for esophagogastric junctional adenocarcinoma. Experimental & Molecular Medicine. 2017;49(9):e383.
12. Kato K, Shah MA, Enzinger P, Bennouna J, Shen L, Adenis A, et al. KEYNOTE-590: Phase III study of first-line chemotherapy with or without pembrolizumab for advanced esophageal cancer. Future Oncology. 2019;15(10):1057-66.
13. Ross P, Nicolson M, Cunningham D, Valle J, Seymour M, Harper P, et al. Prospective randomized trial comparing mitomycin, cisplatin, and protracted venous-infusion fluorouracil (PVI 5-FU) with epirubicin, cisplatin, and PVI 5-FU in advanced esophagogastric cancer. Journal of Clinical Oncology. 2002;20(8):1996-2004.
14. Lee SJ, Kim S, Kim M, Lee J, Park YH, Im YH, et al. Capecitabine in combination with either cisplatin or weekly paclitaxel as a first-line treatment for metastatic esophageal squamous cell carcinoma: A randomized phase II study. BMC Cancer. 2015;15(1):1-7.
15. Wang FH, Wang Y, Sun GP, Chen JH, Lin YC, Liu W, et al. Efficacy and safety of recombinant human lymphotoxin-α derivative with cisplatin and fluorouracil in patients with metastatic esophageal squamous cell carcinoma: A randomized, multicenter, open-label, controlled, phase 2b trial. Cancer. 2017;123(20):3986-94.
16. Moehler M, Maderer A, Thuss-Patience PC, Brenner B, Meiler J, Ettrich TJ, et al. Cisplatin and 5-fluorouracil with or without epidermal growth factor receptor inhibition panitumumab for patients with non-resectable, advanced or metastatic oesophageal squamous cell cancer: a prospective, open-label, randomised phase III AIO/EORTC trial. Annals of Oncology. 2020;31(2):228-35.

17. Lorenzen S, Schuster T, Porschen R, Al-Batran SE, Hofheinz R, Thuss-Patience P, et al. Cetuximab plus cisplatin-5-fluorouracil versus cisplatin-5-fluorouracil alone in first-line metastatic squamous cell carcinoma of the esophagus: A randomized phase II study of the Arbeitsgemeinschaft Internistische Onkologie. Annals of Oncology. 2009;20(10):1667-73.
18. Lee J, Im YH, Cho EY, Hong YS, Lee HR, Kim HS, et al. A phase II study of capecitabine and cisplatin (XP) as first-line chemotherapy in patients with advanced esophageal squamous cell carcinoma. Cancer Chemotherapy and Pharmacology. 2008;62(1):77-84.
19. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). European Journal of Cancer. 2009;45(2):228-47.
20. EuroQol Research Foundation. EQ-5D-5L. About. Available from: https://euroqol.org/eq-5d-instruments/eq-5d-5l-about/.
21. EORTC Quality of Life. EORTC Quality of Life of Cancer Patients Questionnare (QLQ-C30), 1995. Available from: https://qol.eortc.org/questionnaire/eortc-qlq-c30/.
22. EORTC Quality of Life. EORTC Quality of Life Questionnaire Oesophageal Module (QLQ-OES18), 1999. Available from: https://qol.eortc.org/questionnaire/qlq-oes18/.
23. EuroQol Research Foundation. EQ-5D-3L. About. Available from: https://euroqol.org/eq-5d-instruments/eq-5d-3l-about/.
24. Guo X, Zhao F, Ma X, Shen G, Ren D, Zheng F, et al. A comparison between triplet and doublet chemotherapy in improving the survival of patients with advanced gastric cancer: a systematic review and meta-analysis. BMC Cancer. 2019;19(1):1125.
25. Wagner AD, Grothe W, Haerting J, Kleber G, Grothey A, Fleig WE. Chemotherapy in advanced gastric cancer: a systematic review and meta-analysis based on aggregate data. Journal of Clinical Oncology. 2006;24(18):2903-9.
26. Okines AFC, Norman AR, McCloud P, Kang YK, Cunningham D. Meta-analysis of the REAL-2 and ML17032 trials: evaluating capecitabine-based combination chemotherapy and infused 5-fluorouracil-based combination chemotherapy for the treatment of advanced oesophago-gastric cancer. Annals of Oncology. 2009;20(9):1529-34.
27. Oba K, Paoletti X, Bang YJ, Bleiberg H, Burzykowski T, Fuse N, et al. Role of chemotherapy for advanced/recurrent gastric cancer: an individual-patient-data meta-analysis. European Journal of Cancer. 2013;49(7):1565-77.
28. Kim TW, Choi SJ, Ahn JH, Bang HS, Chang HM, Kang YK, et al. A prospective randomized phase III trial of 5-fluorouracil and cisplatin (FP) versus epirubicin, cisplatin, and 5-fu (ECF) in the treatment of patients with previously untreated advanced gastric cancer (AGC). European Journal of Cancer. 2001;37:S314.
29. Kyoto Research Group for Chemotherapy of Gastric Cancer (KRGCGC). A randomized, comparative study of combination chemotherapies in advanced gastric cancer: 5-fluorouracil and cisplatin (FP) versus 5-fluorouracil, cisplatin, and 4'-epirubicin (FPEPIR). Kyoto Research Group for Chemotherapy of Gastric Cancer (KRGCGC). Anticancer Research. 1992;12(6b):1983-8.
30. Yun J, Lee J, Park SH, Park JO, Park YS, Lim HY, et al. A randomised phase II study of combination chemotherapy with epirubicin, cisplatin and capecitabine (ECX) or cisplatin and capecitabine (CX) in advanced gastric cancer. European Journal of Cancer. 2010;46(5):885-91.
31. Lordick F, Kang YK, Chung HC, Salman P, Oh SC, Bodoky G, et al. Capecitabine and cisplatin with or without cetuximab for patients with previously untreated advanced gastric cancer (EXPAND): a randomised, open-label phase 3 trial. Lancet Oncology. 2013;14(6):490-9.
32. Shah MA, Bang YJ, Lordick F, Alsina M, Chen M, Hack SP, et al. Effect of fluorouracil, leucovorin, and oxaliplatin with or without onartuzumab in HER2-Negative, MET-positive gastroesophageal adenocarcinoma: The METGastric randomized clinical trial. JAMA Oncology. 2017;3(5):620-7.

33. Liu Q, Zeng H, Xia R, Chen G, Liu S, Zhang Z, et al. Health-related quality of life of esophageal cancer patients in daily life after treatment: A multicenter cross-sectional study in China. Cancer Medicine. 2018;7(11):5803-11.
34. Zhao ZM, Pan XF, Wen Y, Huang WZ, Chen F, Lan H, et al. Quality of life among patients with esophageal/cardiac precursor lesion or cancer: A one-year prospective survey. Value in Health. 2014;17(7):A738.
35. Doherty M, Naik H, Eng L, Patel D, Kong QQ, Xu W, et al. The relationship between health utility, quality of life, and symptom scores in Canadian patients with esophageal cancer. Journal of Clinical Oncology. 2015;33(3 Suppl):149-.
36. Doherty MK, Leung Y, Su J, Naik H, Patel D, Eng L, et al. Health utility scores from EQ-5D and health-related quality of life in patients with esophageal cancer: a real-world cross-sectional study. Diseases of the Esophagus. 2018;31(12).
37. Wildi SM, Cox MH, Clark LL, Turner R, Hawes RH, Hoffman BJ, et al. Assessment of health state utilities and quality of life in patients with malignant esophageal dysphagia. American Journal of Gastroenterology. 2004;99(6):1044-9.
38. Wang Y, Shi J, Du L, Huang H, Wang L, Zhu J, et al. Health-related quality of life in patients with esophageal cancer or precancerous lesions assessed by EQ-5D: A multicenter cross-sectional study. Thoracic Cancer. 2020;11(4):1076-89.
39. Janmaat VT, Bruno MJ, Polinder S, Lorenzen S, Lordick F, Peppelenbosch MP, et al. Cost-effectiveness of cetuximab for advanced esophageal squamous cell carcinoma. PloS One. 2016;11(4):e0153943-e.
40. Zhang P-F, Xie D, Li Q. Cost–effectiveness analysis of nivolumab in the second-line treatment for advanced esophageal squamous cell carcinoma. Future Oncology. 2020;16(17):1189-98.
41. Agus AM, Kinnear H, O'Neill C, McDowell C, Crealey GE, Gavin A. Description and predictors of hospital costs of oesophageal cancer during the first year following diagnosis in Northern Ireland. European Journal of Cancer Care. 2013;22(4):450-8.
42. Van Hout B, Janssen MF, Feng YS, Kohlmann T, Busschbach J, Golicki D, et al. Interim scoring for the EQ-5D-5L: Mapping the EQ-5D-5L to EQ-5D-3L value sets. Value in Health. 2012;15(5):708-15.
43. National Institute for Health and Care Excellence (NICE). Single Technology Appraisal (STA): Pembrolizumab for previously treated advanced or metastatic urothelial cancer - Committee Papers [TA519], 2018. Available from: https://www.nice.org.uk/guidance/ta519/documents/committee-papers.
44. National Institute for Health and Care Excellence (NICE). Single Technology Appraisal (STA): Pembrolizumab with axitinib for untreated advanced renal cell carcinoma [TA650] - Committee Papers, 2020. Available from: https://www.nice.org.uk/guidance/ta650/evidence.
45. National Institute for Health and Care Excellence (NICE). Single Technology Appraisal (STA): Pembrolizumab with pemetrexed and platinum chemotherapy for untreated lung cancer [TA557] - Committee Papers, 2018. Available from: https://www.nice.org.uk/guidance/ta557/evidence.
46. National Institute for Health and Care Excellence (NICE). Single Technology Appraisal (STA): Nivolumab for previously treated unresectable advanced oesophageal cancer [ID1249], 2020. Available from: https://www.nice.org.uk/guidance/indevelopment/gid-ta10222.
47. Merck Sharpe & Dohme. A randomized, double-blind, placebo-controlled phase III clinical trial of pembrolizumab (MK-3475) in combination with cisplatin and 5-fluorouracil versus placebo in combination with cisplatin and 5-fluorouracil as first-line treatment in subjects with advanced/metastatic esophageal carcinoma (KEYNOTE-590). Clinical Study Report. 2020.
48. Latimer N. NICE DSU technical support document 14: survival analysis for economic evaluations alongside clinical trials-extrapolation with patient-level data, 2011 Sheffield:

Decision Support Unit, ScHARR, University of Sheffield. Available from: http://nicedsu.org.uk/wp-content/uploads/2016/03/NICE-DSU-TSD-Survival-analysis.updated-March-2013.v2.pdf.
49. Chow GC. Tests of equality between sets of coefficients in two linear regressions. Econometrica. 1960;28(3):591-605.
50. Gavin AT, Francisci S, Foschi R, Donnelly DW, Lemmens V, Brenner H, et al. Oesophageal cancer survival in Europe: A EUROCARE-4 study. Cancer Epidemiology. 2012;36(6):505-12.
51. American Cancer Society. Survival rates for esophageal cancer, 2021. Available from: https://www.cancer.org/content/cancer/en/cancer/esophagus-cancer/detection-diagnosis-staging/survival-rates/.
52. Wu SG, Xie WH, Zhang ZQ, Sun JY, Li FY, Lin HX, et al. Surgery combined with radiotherapy improved survival in metastatic esophageal cancer in a surveillance epidemiology and end results population-based study. Scientific Reports. 2016;6:28280.
53. Wu SG, Zhang WW, He ZY, Sun JY, Chen YX, Guo L. Sites of metastasis and overall survival in esophageal cancer: A population-based study. Cancer Management and Research. 2017;9:781-8.
54. Tanaka T, Fujita H, Matono S, Nagano T, Nishimura K, Murata K, et al. Outcomes of multimodality therapy for stage IVB esophageal cancer with distant organ metastasis (M1-Org). Diseases of the Esophagus. 2010;23(8):646-51.
55. National Institute for Health and Care Excellence (NICE). Position statement on use of the EQ-5D-5L value set for England, 2019. Available from: https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/technology-appraisal-guidance/eq-5d-5l.
56. Ara R, Brazier JE. Populating an economic model with health state utility values: Moving toward better practice. Value in Health. 2010;13(5):509-18.
57. Hatswell AJ, Pennington B, Pericleous L, Rowen D, Lebmeier M, Lee D. Patient-reported utilities in advanced or metastatic melanoma, including analysis of utilities by time to death. Health and Quality of Life Outcomes. 2014;12(1):1-9.
58. Curtis LA, Burns A. Unit costs of health and social care, 2019 Kent, U.K.: Personal Social Services Research Unit. Available from: https://kar.kent.ac.uk/79286/11/UCFinalFeb20.pdf.
59. NHS Digital. National Tariff Chemotherapy Regimens List 2017-18. Version 1.0, 2018.
60. NHS England. National Schedule of NHS Costs - NHS Reference Costs 2018/19, 2019.
61. National Institute for Health and Care Excellence (NICE). Single Technology Appraisal (STA): Ramucirumab for treating advanced gastric cancer or gastro–oesophageal junction adenocarcinoma - Committee Papers [TA378], 2016. Available from: https://www.nice.org.uk/guidance/ta378/.
62. National Institute for Health and Care Excellence (NICE). Single Technology Appraisal (STA): Pembrolizumab for locally advanced or metastatic urothelial cancer where cisplatin is unsuitable - Committee Papers [TA522], 2018. Available from: https://www.nice.org.uk/guidance/ta522/resources/pembrolizumab-for-untreated-locally-advanced-or-metastatic-urothelial-cancer-when-cisplatin-is-unsuitable-pdf-82606837115077.
63. National Institute for Health and Care Excellence (NICE). Single Technology Appraisal (STA): Pembrolizumab for untreated metastatic or unresectable recurrent squamous cell head and neck cancer [TA661] - Committee Papers, 2020. Available from: https://www.nice.org.uk/guidance/ta661/evidence/appraisal-consultation-committee-papers-pdf-8948996605.
64. NHS. Dysphagia (swallowing problems), 2018. Available from: https://www.nhs.uk/conditions/swallowing-problems-dysphagia/.

65. Georghiou T, Bardsley M. Exploring the cost of care at the end of life, 2014: Nuffield Trust. Available from: https://www.nuffieldtrust.org.uk/files/2017-01/end-of-life-care-web-final.pdf.
66. Saito S, Muneoka Y, Ishikawa T, Akazawa K. Cost-effectiveness of paclitaxel + ramucirumab combination therapy for advanced gastric cancer progressing after first-line chemotherapy in Japan. Clinical Therapeutics. 2017;39(12):2380-8.
67. Al-Batran SE, Van Cutsem E, Oh SC, Bodoky G, Shimada Y, Hironaka S, et al. Quality-of-life and performance status results from the phase III RAINBOW study of ramucirumab plus paclitaxel versus placebo plus paclitaxel in patients with previously treated gastric or gastroesophageal junction adenocarcinoma. Annals of Oncology. 2016;27(4):673-9.

Appendix A: Search strategies for Ovid MEDLINE and Embase

Search strategies for additional work completed by the ERG reported in Section 3.5.1.

Search strategy for Ovid MEDLINE

1 exp Esophageal Neoplasms/ (51854)
2 exp esophagus cancer/ (51854)
3 (Cancer of the esophagus or esophageal adenocarcinoma or esophageal squamous cell carcinoma).mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms] (15585)
4 ((Esophageal or esophagus) and (cancer or carcinom? or tumour? or tumor? or neoplasm?)).mp. (76667)
5 (("adenocarcinoma of the esophagus" or "squamous cell carcinoma of the esophagus" or "Siewert type 1 adenocarcinoma of the esophageal junction") and (stage 3 or stage III or stage 3b or stage IIIb or stage 4 or stage IV or metasta? or advanced)).mp. (425)
6 or/1-5 (77194)
7 (pembrolizumab or MK-3475 or MK3475 or lambrolizumab or keytruda).mp. (5154)
8 exp Nivolumab/ (2975)
9 (nivolumab or ONO-4538 or ONO 4538 or BMS-936558 or BMS 936558 or MDX-1106 or MDX 1106 or MDX1106 or opdivo).mp. (6205)
10 exp ipilimumab/ (2071)
11 ("Anti-CTLA-4 MAb" or "Anti CTLA 4 MAb Ipilimumab" or "Ipilimumab, Anti-CTLA-4 Mab" or Yervoy or "MDX 010" or "MDX010" or "MDX-010" or "MDX-CTLA-4" or "MDX CTLA 4").mp. (197)
12 exp epirubicin/ (5260)
13 (epirubicin or epiadriamycin or epidoxorubicin).mp. (7561)
14 exp trastuzumab/ (7251)
15 exp paclitaxel/ (27506)
16 (paclitaxel or nab-paclitaxel or "abi 007" or "abi007" or abraxane or anzatax or asotax or "bms 181339" or "bms181339" or britaxol or coroxane or "mbt 0206" or "mbt0206 or nab paclitaxel or nsc 125973 or nsc125973" or pacitaxel or praxel or paxene or taxol).mp. (41285)
17 (ramucirumab or cyramza or "imc 1121 b" or "imc 1121b" or "imc1121 b" or "imc1121b" or "ly3009806" or "ly 3009806").mp. (903)
18 exp docetaxel/ (10907)
19 (docetaxel or docetaxol accord or daxotel or dexotel or "lit 976" or "lit976" or "nsc 628503" or "nsc628503" or "rp 56976" or "rp56976" or taxoter or taxotere).mp. (17519)
20 exp irinotecan/ (7199)
21 (irinotecan or camptosar or campto or "cpt 11" or "cpt11" or irinotecan hydrochloride or irinotel).mp. (11863)
22 exp capecitabine/ (4738)
23 (capecitabine or apecitab or ecansya or "ro 09 1978" or "ro 091978" or "ro091978" or "ro09 1978" or xeloda).mp. (7570)
24 exp carboplatin/ (11915)
25 (carboplatin or paraplatin).mp. (18281)
26 exp leucovorin/ (10289)

27 exp folinic acid/ (10289)
28 (leucovorin or folinic acid or Wellcovorin or Citrovorum Factor or leucovorin or leucovoran or leukovorin).mp. (13829)
29 exp 5-FU/ (47430)
30 exp fluorouracil/ (47430)
31 (5-fluorouracil or "5 fluorouracil" or adrucil or "5-FU" or "5 FU" or fluoroblastin or fluorolex or "fluorouracil 5" or "nsc 18913" or "nsc18913" or "nsc 19893" or "nsc19893").mp. (39372)
32 exp cisplatin/ (53179)
33 (cisplatin or cisplatinum or cis-platinum or cis platinum or platamin or neoplatin or cismaplat or cis-maplat or "mpi 5010" or "mpi5010" or "nk 801" or platinol or platinex or platamine).mp. (79657)
34 exp oxaliplatin/ (6803)
35 (Oxaliplatin or eloxatin).mp. (12612)
36 (mFOLFOX6 or mFOLFOX-6 or m-FOLFOX6 or m-FOLFOX-6 or modified FOLFOX6 or modified FOLFOX-6).mp. (811)
37 (FOLFOX* or FOLFOX6 or FOLFOX-6 or "folfox regimen").mp. (3553)
38 (FOLFIRI or FOLinic acid-Fluorouracil-IRInotecan).mp. (1529)
39 (CAPOX or XELOX or capecitabine-oxaliplatin).mp. (1165)
40 capecitabine-carboplatin.mp. (4)
41 or/7-40 (211467)
42 Randomized Controlled Trials as Topic/ (141253)
43 randomized controlled trial/ (524786)
44 Random Allocation/ (104805)
45 Double Blind Method/ (162861)
46 Single Blind Method/ (29846)
47 clinical trial/ (527778)
48 clinical trial, phase i.pt. (21375)
49 clinical trial, phase ii.pt. (34350)
50 clinical trial, phase iii.pt. (18066)
51 clinical trial, phase iv.pt. (2060)
52 controlled clinical trial.pt. (94093)
53 randomized controlled trial.pt. (524786)
54 multicenter study.pt. (289732)
55 clinical trial.pt. (527778)
56 exp Clinical Trials as topic/ (353660)
57 or/42-56 (1413298)
58 (clinical adj trial$).tw. (391486)
59 ((singl$ or doubl$ or treb$ or tripl$) adj (blind$3 or mask$3)).tw. (178444)
60 PLACEBOS/ (35369)
61 placebo$.tw. (222942)
62 randomly allocated.tw. (30522)
63 (allocated adj2 random$).tw. (33923)
64 or/58-63 (668871)
65 57 or 64 (1699891)
66 6 and 41 and 65 (1558)
67 limit 66 to english language (1422)
68 limit 67 to yr=2000-current (1171) [original search reported in CS]
69 exp Esophageal Neoplasms/ (51854)
70 exp esophagus cancer/ (51854)

71 (Cancer of the esophagus or esophageal adenocarcinoma or esophageal squamous cell carcinoma).mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms] (15585)
72 ((Esophag* or oesophag* or gastroesophag* or gastrooesophag* or gastro-esophag* or gastro-oesophag* or siewert*) and (adenocarcinoma* or cancer or carcinom? or tumour? or tumor? or neoplasm? or metastas* or metastatic)).mp. (87890)
73 (("adenocarcinoma of the esophagus" or "squamous cell carcinoma of the esophagus" or "Siewert type 1 adenocarcinoma of the esophageal junction") and (stage 3 or stage III or stage 3b or stage IIIb or stage 4 or stage IV or metasta? or advanced)).mp. (425)
74 or/69-73 (87890)
75 (pembrolizumab or MK-3475 or MK3475 or lambrolizumab or keytruda).mp. (5154)
76 exp Nivolumab/ (2975)
77 (nivolumab or ONO-4538 or ONO 4538 or BMS-936558 or BMS 936558 or MDX-1106 or MDX 1106 or MDX1106 or opdivo).mp. (6205)
78 exp ipilimumab/ (2071)
79 ("Anti-CTLA-4 MAb" or "Anti CTLA 4 MAb Ipilimumab" or "Ipilimumab, Anti-CTLA-4 Mab" or Yervoy or "MDX 010" or "MDX010" or "MDX-010" or "MDX-CTLA-4" or "MDX CTLA 4").mp. (197)
80 exp epirubicin/ (5260)
81 (epirubicin or epiadriamycin or epidoxorubicin).mp. (7561)
82 exp trastuzumab/ (7251)
83 exp paclitaxel/ (27506)
84 (paclitaxel or nab-paclitaxel or "abi 007" or "abi007" or abraxane or anzatax or asotax or "bms 181339" or "bms181339" or britaxol or coroxane or "mbt 0206" or "mbt0206 or nab paclitaxel or nsc 125973 or nsc125973" or pacitaxel or praxel or paxene or taxol).mp. (41285)
85 (ramucirumab or cyramza or "imc 1121 b" or "imc 1121b" or "imc1121 b" or "imc1121b" or "ly3009806" or "ly 3009806").mp. (903)
86 exp docetaxel/ (10907)
87 (docetaxel or docetaxol accord or daxotel or dexotel or "lit 976" or "lit976" or "nsc 628503" or "nsc628503" or "rp 56976" or "rp56976" or taxoter or taxotere).mp. (17519)
88 exp irinotecan/ (7199)
89 (irinotecan or camptosar or campto or "cpt 11" or "cpt11" or irinotecan hydrochloride or irinotel).mp. (11863)
90 exp capecitabine/ (4738)
91 (capecitabine or apecitab or ecansya or "ro 09 1978" or "ro 091978" or "ro091978" or "ro09 1978" or xeloda).mp. (7570)
92 exp carboplatin/ (11915)
93 (carboplatin or paraplatin).mp. (18281)
94 exp leucovorin/ (10289)
95 exp folinic acid/ (10289)
96 (leucovorin or folinic acid or Wellcovorin or Citrovorum Factor or leucovorin or leucovoran or leukovorin).mp. (13829)
97 exp 5-FU/ (47430)
98 exp fluorouracil/ (47430)
99 (5-fluorouracil or "5 fluorouracil" or adrucil or "5-FU" or "5 FU" or fluoroblastin or fluorolex or "fluorouracil 5" or "nsc 18913" or "nsc18913" or "nsc 19893" or "nsc19893").mp. (39372)
100 exp cisplatin/ (53179)

101 (cisplatin or cisplatinum or cis-platinum or cis platinum or platamin or neoplatin or cismaplat or cis-maplat or "mpi 5010" or "mpi5010" or "nk 801" or platinol or platinex or platamine).mp. (79657)
102 exp oxaliplatin/ (6803)
103 (Oxaliplatin or eloxatin).mp. (12612)
104 (mFOLFOX6 or mFOLFOX-6 or m-FOLFOX6 or m-FOLFOX-6 or modified FOLFOX6 or modified FOLFOX-6).mp. (811)
105 (FOLFOX* or FOLFOX6 or FOLFOX-6 or "folfox regimen").mp. (3553)
106 (FOLFIRI or FOLinic acid-Fluorouracil-IRInotecan).mp. (1529)
107 (CAPOX or XELOX or capecitabine-oxaliplatin).mp. (1165)
108 capecitabine-carboplatin.mp. (4)
109 or/75-108 (211467)
110 Randomized Controlled Trials as Topic/ (141253)
111 randomized controlled trial/ (524786)
112 Random Allocation/ (104805)
113 Double Blind Method/ (162861)
114 Single Blind Method/ (29846)
115 clinical trial/ (527778)
116 clinical trial, phase i.pt. (21375)
117 clinical trial, phase ii.pt. (34350)
118 clinical trial, phase iii.pt. (18066)
119 clinical trial, phase iv.pt. (2060)
120 controlled clinical trial.pt. (94093)
121 randomized controlled trial.pt. (524786)
122 multicenter study.pt. (289732)
123 clinical trial.pt. (527778)
124 exp Clinical Trials as topic/ (353660)
125 or/110-124 (1413298)
126 (clinical adj trial$).tw. (391486)
127 ((singl$ or doubl$ or treb$ or tripl$) adj (blind$3 or mask$3)).tw. (178444)
128 PLACEBOS/ (35369)
129 placebo$.tw. (222942)
130 randomly allocated.tw. (30522)
131 (allocated adj2 random$).tw. (33923)
132 (single arm or "single arm").ti,ab. (8516)
133 or/126-132 (675347)
134 125 or 133 (1703120)
135 74 and 109 and 134 (2209)
136 limit 135 to english language (2044)
137 limit 136 to yr=2000-current (1673) [search strategy amended by ERG]
138 137 not 68 (502) [additional studies identified by ERG search strategy]

Search strategy for Ovid Embase

1 exp Esophageal Neoplasms/ (86939)
2 exp esophagus cancer/ (73510)
3 exp esophagus carcinoma/ (40124)
4 exp esophagus metastasis/ (552)
5 exp esophagus tumor/ (86939)
6 exp esophageal adenocarcinoma/ (11913)

7 exp esophageal squamous cell carcinoma/ (14800)
8 (Cancer of the esophagus or esophageal adenocarcinoma or esophageal squamous cell carcinoma).mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word, candidate term word] (31062)
9 ((Esophageal or esophagus) and (cancer or carcinom? or tumour? or tumor? or neoplasm?)).mp. (122919)
10 (("adenocarcinoma of the esophagus" or "squamous cell carcinoma of the esophagus" or "Siewert type 1 adenocarcinoma of the esophageal junction") and (stage 3 or stage III or stage 3b or stage IIIb or stage 4 or stage IV or metasta? or advanced)).mp. (682)
11 or/1-10 (124826)
12 exp pembrolizumab/ (18704)
13 (pembrolizumab or "mk-3475" or "mk3475" or "mk 3475" or lambrolizumab or keytruda).mp. (19814)
14 exp Nivolumab/ (20909)
15 (nivolumab or ONO-4538 or ONO 4538 or BMS-936558 or BMS 936558 or MDX-1106 or MDX 1106 or MDX1106 or opdivo).mp. (22048)
16 exp ipilimumab/ (15724)
17 ("Anti-CTLA-4 MAb" or "Anti CTLA 4 MAb Ipilimumab" or "Ipilimumab, Anti-CTLA-4 Mab" or Yervoy or "MDX 010" or "MDX010" or "MDX-010" or "MDX-CTLA-4" or "MDX CTLA 4").mp. (1334)
18 exp epirubicin/ (29473)
19 (epirubicin or epiadriamycin or epidoxorubicin).mp. (30167)
20 exp trastuzumab/ (40894)
21 exp paclitaxel/ (111046)
22 (paclitaxel or nab-paclitaxel or "abi 007" or "abi007" or abraxane or anzatax or asotax or "bms 181339" or "bms181339" or britaxol or coroxane or "mbt 0206" or "mbt0206 or nab paclitaxel or nsc 125973 or nsc125973" or pacitaxel or praxel or paxene or taxol).mp. (117528)
23 exp ramucirumab/ (3030)
24 (ramucirumab or cyramza or "imc 1121 b" or "imc 1121b" or "imc1121 b" or "imc1121b" or "ly3009806" or "ly 3009806").mp. (3307)
25 exp docetaxel/ (61159)
26 (docetaxel or docetaxol accord or daxotel or dexotel or "lit 976" or "lit976" or "nsc 628503" or "nsc628503" or "rp 56976" or "rp56976" or taxoter or taxotere).mp. (63326)
27 exp irinotecan/ (39113)
28 (irinotecan or camptosar or campto or "cpt 11" or "cpt11" or irinotecan hydrochloride or irinotel).mp. (41110)
29 exp capecitabine/ (30495)
30 (capecitabine or apecitab or ecansya or "ro 09 1978" or "ro 091978" or "ro091978" or "ro09 1978" or xeloda).mp. (32522)
31 exp carboplatin/ (72037)
32 (carboplatin or paraplatin).mp. (74583)
33 exp leucovorin/ (37701)
34 exp folinic acid/ (37701)
35 (leucovorin or folinic acid or Wellcovorin or Citrovorum Factor or leucovorin or leucovoran or leukovorin).mp. (39615)
36 exp 5-FU/ (141803)
37 exp fluorouracil/ (141803)
38 (5-fluorouracil or "5 fluorouracil" or adrucil or "5-FU" or "5 FU" or fluoroblastin or fluorolex or "fluorouracil 5" or "nsc 18913" or "nsc18913" or "nsc 19893" or "nsc19893").mp. (54115)

39 exp cisplatin/ (189824)
40 (cisplatin or cisplatinum or cis-platinum or cis platinum or platamin or neoplatin or cismaplat or cis-maplat or "mpi 5010" or "mpi5010" or "nk 801" or platinol or platinex or platamine).mp. (199050)
41 exp oxaliplatin/ (41594)
42 (Oxaliplatin or eloxatin).mp. (43970)
43 (mFOLFOX6 or mFOLFOX-6 or m-FOLFOX6 or m-FOLFOX-6 or modified FOLFOX6 or modified FOLFOX-6).mp. (1585)
44 (FOLFOX* or FOLFOX6 or FOLFOX-6 or "folfox regimen").mp. (6761)
45 (FOLFIRI or FOLinic acid-Fluorouracil-IRInotecan).mp. (3823)
46 (CAPOX or XELOX or capecitabine-oxaliplatin).mp. (2640)
47 capecitabine-carboplatin.mp. (8)
48 or/12-47 (517617)
49 clinical trial/ (1004398)
50 Randomized controlled trial/ (650791)
51 controlled clinical trial/ (466124)
52 multicenter study/ (282811)
53 Phase 3 clinical trial/ (52263)
54 Phase 4 clinical trial/ (4248)
55 exp RANDOMIZATION/ (90715)
56 Single Blind Procedure/ (42240)
57 Double Blind Procedure/ (182483)
58 Crossover Procedure/ (66463)
59 PLACEBO/ (364703)
60 randomi?ed controlled trial$.tw. (253069)
61 rct.tw. (41261)
62 (random$ adj2 allocat$).tw. (46065)
63 Single blind$.tw. (26659)
64 Double blind$.tw. (218829)
65 ((treble or triple) adj blind$).tw. (1315)
66 Placebo$.tw. (323746)
67 Prospective study/ (671266)
68 or/49-67 (2477180)
69 11 and 48 and 68 (4143)
70 limit 69 to english language (3880)
71 limit 70 to yr=2000-current (3573) [original search reported in CS]
72 exp Esophageal Neoplasms/ (86939)
73 exp esophagus cancer/ (73510)
74 exp esophagus carcinoma/ (40124)
75 exp esophagus metastasis/ (552)
76 exp esophagus tumor/ (86939)
77 exp esophageal adenocarcinoma/ (11913)
78 exp esophageal squamous cell carcinoma/ (14800)
79 (Cancer of the esophagus or esophageal adenocarcinoma or esophageal squamous cell carcinoma).mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word, candidate term word] (31062)
80 ((Esophageal or esophagus) and (cancer or carcinom? or tumour? or tumor? or neoplasm?)).mp. (122919)

81 (("adenocarcinoma of the esophagus" or "squamous cell carcinoma of the esophagus" or "Siewert type 1 adenocarcinoma of the esophageal junction") and (stage 3 or stage III or stage 3b or stage IIIb or stage 4 or stage IV or metasta? or advanced)).mp. (682)
82 ((Esophageal or esophagus or esophagogastric* or oesophageal or esophagus or esophagogastric* or gastroesophag* or gastrooesophag* or gastro-esophag* or gastro-oesophag* or siewert*) and (adenocarcinoma* or cancer or carcinom? or tumour? or tumor? or neoplasm? or metastas* or metastatic)).mp. (134020)
83 or/72-82 (134189)
84 exp pembrolizumab/ (18704)
85 (pembrolizumab or "mk-3475" or "mk3475" or "mk 3475" or lambrolizumab or keytruda).mp. (19814)
86 exp Nivolumab/ (20909)
87 (nivolumab or ONO-4538 or ONO 4538 or BMS-936558 or BMS 936558 or MDX-1106 or MDX 1106 or MDX1106 or opdivo).mp. (22048)
88 exp ipilimumab/ (15724)
89 ("Anti-CTLA-4 MAb" or "Anti CTLA 4 MAb Ipilimumab" or "Ipilimumab, Anti-CTLA-4 Mab" or Yervoy or "MDX 010" or "MDX010" or "MDX-010" or "MDX-CTLA-4" or "MDX CTLA 4").mp. (1334)
90 exp epirubicin/ (29473)
91 (epirubicin or epiadriamycin or epidoxorubicin).mp. (30167)
92 exp trastuzumab/ (40894)
93 exp paclitaxel/ (111046)
94 (paclitaxel or nab-paclitaxel or "abi 007" or "abi007" or abraxane or anzatax or asotax or "bms 181339" or "bms181339" or britaxol or coroxane or "mbt 0206" or "mbt0206 or nab paclitaxel or nsc 125973 or nsc125973" or pacitaxel or praxel or paxene or taxol).mp. (117528)
95 exp ramucirumab/ (3030)
96 (ramucirumab or cyramza or "imc 1121 b" or "imc 1121b" or "imc1121 b" or "imc1121b" or "ly3009806" or "ly 3009806").mp. (3307)
97 exp docetaxel/ (61159)
98 (docetaxel or docetaxol accord or daxotel or dexotel or "lit 976" or "lit976" or "nsc 628503" or "nsc628503" or "rp 56976" or "rp56976" or taxoter or taxotere).mp. (63326)
99 exp irinotecan/ (39113)
100 (irinotecan or camptosar or campto or "cpt 11" or "cpt11" or irinotecan hydrochloride or irinotel).mp. (41110)
101 exp capecitabine/ (30495)
102 (capecitabine or apecitab or ecansya or "ro 09 1978" or "ro 091978" or "ro091978" or "ro09 1978" or xeloda).mp. (32522)
103 exp carboplatin/ (72037)
104 (carboplatin or paraplatin).mp. (74583)
105 exp leucovorin/ (37701)
106 exp folinic acid/ (37701)
107 (leucovorin or folinic acid or Wellcovorin or Citrovorum Factor or leucovorin or leucovoran or leukovorin).mp. (39615)
108 exp 5-FU/ (141803)
109 exp fluorouracil/ (141803)
110 (5-fluorouracil or "5 fluorouracil" or adrucil or "5-FU" or "5 FU" or fluoroblastin or fluorolex or "fluorouracil 5" or "nsc 18913" or "nsc18913" or "nsc 19893" or "nsc19893").mp. (54115)
111 exp cisplatin/ (189824)

112 (cisplatin or cisplatinum or cis-platinum or cis platinum or platamin or neoplatin or cismaplat or cis-maplat or "mpi 5010" or "mpi5010" or "nk 801" or platinol or platinex or platamine).mp. (199050)
113 exp oxaliplatin/ (41594)
114 (Oxaliplatin or eloxatin).mp. (43970)
115 (mFOLFOX6 or mFOLFOX-6 or m-FOLFOX6 or m-FOLFOX-6 or modified FOLFOX6 or modified FOLFOX-6).mp. (1585)
116 (FOLFOX* or FOLFOX6 or FOLFOX-6 or "folfox regimen").mp. (6761)
117 (FOLFIRI or FOLinic acid-Fluorouracil-IRInotecan).mp. (3823)
118 (CAPOX or XELOX or capecitabine-oxaliplatin).mp. (2640)
119 capecitabine-carboplatin.mp. (8)
120 or/84-119 (517617)
121 clinical trial/ (1004398)
122 Randomized controlled trial/ (650791)
123 controlled clinical trial/ (466124)
124 multicenter study/ (282811)
125 Phase 3 clinical trial/ (52263)
126 Phase 4 clinical trial/ (4248)
127 exp RANDOMIZATION/ (90715)
128 Single Blind Procedure/ (42240)
129 Double Blind Procedure/ (182483)
130 Crossover Procedure/ (66463)
131 PLACEBO/ (364703)
132 randomi?ed controlled trial$.tw. (253069)
133 rct.tw. (41261)
134 (random$ adj2 allocat$).tw. (46065)
135 Single blind$.tw. (26659)
136 Double blind$.tw. (218829)
137 ((treble or triple) adj blind$).tw. (1315)
138 Placebo$.tw. (323746)
139 Prospective study/ (671266)
140 "single arm".mp. (18467)
141 or/121-140 (2484235)
142 83 and 120 and 141 (4856)
143 limit 142 to english language (4583)
144 limit 143 to yr=2000-current (4266) [search strategy amended by ERG]
145 144 not 71 (693) [additional studies identified by ERG search strategy]