Confidential until published

Pembrolizumab with trastuzumab and chemotherapy for untreated HER2-positive advanced gastric or gastro-oesophageal junction cancer. A Single Technology Appraisal

Produced by School of Health and Related Research (ScHARR), University of Sheffield

Authors Sarah Davis, Senior Research Fellow, ScHARR, University of Sheffield, Sheffield, UK

Emma Simpson, Senior Research Fellow, ScHARR, University of Sheffield, Sheffield, UK

Shijie Ren, Senior Research Fellow, ScHARR, University of Sheffield, Sheffield, UK

Andrew Metry, Research Fellow, ScHARR, University of Sheffield, Sheffield, UK

Sarah Ren, Research Associate, ScHARR, University of Sheffield, Sheffield, UK

Ruth Wong, Information Specialist, ScHARR, University of Sheffield, Sheffield, UK

Correspondence Author Sarah Davis, Senior Research Fellow, ScHARR, University of Sheffield, Sheffield, UK

Date completed Date completed (27/07/2023)

Source of funding: This report was commissioned by the NIHR Evidence Synthesis Programme as project number NIHR135983.

Declared competing interests of the authors

None of the authors has any conflicts of interest to declare.

Acknowledgements

We would like to thank Dr Cheng Boon and Dr Mark Harrison for clinical advice relating to this project. We would also like to thank Matt Stevenson, ScHARR, for providing comments on the draft report and Gill Rooney, Programme Manager, ScHARR, for providing administrative support and in preparing and formatting the report.

Rider on responsibility for report

The views expressed in this report are those of the authors and not necessarily those of the NIHR Evidence Synthesis Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:

Davis S, Simpson E, Ren S, Metry A, Ren S, Wong R. Pembrolizumab with trastuzumab and chemotherapy for untreated HER2-positive advanced gastric or gastro-oesophageal junction cancer. A Single Technology Appraisal: A Single Technology Appraisal. School of Health and Related Research (ScHARR), 2023.

Contributions of authors

Sarah Davis and Andrew Metry critiqued the health economic analysis submitted by the company. Emma Simpson summarised and critiqued the clinical effectiveness data reported within the company's submission. Shijie Ren and Sarah Ren critiqued the statistical aspects of the submission. Ruth Wong critiqued the company's search strategy. All authors were involved in drafting and commenting on the final report.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF BOXES

ABBREVIATIONS

5-FU	Fluorouracil
AEs	Adverse events
AIC	Akaike Information Criterion
ASCO	American Society of Clinical Oncology
BIC	Bayesian Information Criterion
BICR	Blinded independent central review
BID	Twice daily
BNF	British National Formulary
BSA	Body Surface Area
CAPOX	Capecitabine and oxaliplatin doublet chemotherapy
CEAC	Cost-Effectiveness Acceptability Curve
cLDA	Constrained longitudinal data analysis
CMU	Commercial medicines unit
CPS	Combined positive score
CR	Complete response
CS	Company Submission
CSR	Clinical study report
CT	Computerised tomography
DOR	Duration of response
DSU	Decision Support Unit
eMIT	electronic Market Information Tool
ECOG	Eastern Cooperative Oncology Group
EMA	European Medicines Agency
EORTC QLQ-C30	European Organisation for Research and Treatment of Cancer Quality-of-Life Questionnaire Core 30
EQ-5D-3L	EuroQol 5 dimensions 3 level
EQ-5D-5L	EuroQol 5 dimensions 5 level
EQ-VAS	EuroQol 5 dimensions Visual Analogue Scale
EAG	External Assessment Group
ERG	Evidence Review Group
ESMO	European Society of Medical Oncology
FISH	Fluorescent in-situ hybridization
FOLFOX	Fluorouracil and oxaliplatin doublet chemotherapy
FOLFIRI	Irinotecan with 5-FU and folinic acid
FP	Fluorouracil and cisplatin doublet chemotherapy

GOJ	Gastro-Oesophageal Junction Cancer
HCHS	Hospital & community health services
HER2	Human Epidermal Growth Factor Receptor 2
HR	Hazard Ratio
HRG	Health resource group
HRQoL	Health-Related Quality of Life
HTA	Health Technology Appraisal
IA	Interim analysis
ICER	Incremental Cost Effectiveness Ratio
IHC	Immunohistochemistry
ITT	Intention to treat
IV	Intravenously
KM	Kaplan-Meier
LSMean	Least square mean
MRI	Magnetic resonance imaging
MSI	Microsatellite instability
MUGA	Multigated acquisition scan
NICE	National Institute for Health and Care Excellence
NG83	NICE's guideline 83 on assessment and management in adults with oesophago-gastric cancer
NHS	National Health Service
NHSCII	NHS Cost Inflation Index
NR	Not Reported
ORR	Overall Response Rate
OS	Overall Survival
PAS	Patient Access Scheme
PD-L1	Programmed death-ligand-1
PET	Positron emission tomography
PFS	Progression-Free Survival
PR	Partial response
PSA	Probabilistic Sensitivity Analysis
PSS	Personal Social Services
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRO	Patient reported outcome
Q3W	Every three weeks
QALY	Quality-Adjusted Life Year

RECIST	Response Evaluation Criteria In Solid Tumours
RCT	Randomised Controlled Trial
RDI	Relative dose intensity
SAE	Serious adverse events
SLR	Systematic Literature Review
SoC	Standard of care
SOX	S-1 and oxaliplatin doublet chemotherapy
STA	Single Technology Appraisal
TA	Technology Appraisal
TA208	Technology Appraisal of trastuzumab for the treatment of HER2-positive metastatic gastric cancer
TA378	Technology Appraisal of ramucirumab for treating advanced gastric cancer or GOJ adenocarcinoma previously treated with chemotherapy
TA737	Technology Appraisal of pembrolizumab with platinum- and fluoropyrimidine-based chemotherapy for untreated advanced oesophageal and gastro-oesophageal junction cancer
TA852	Technology Appraisal of Trifluridine–tipiracil for treating metastatic gastric cancer or GOJ adenocarcinoma after 2 or more treatments
TSD	Technical Support Document
TTD	Time To Treatment Discontinuation
TWiST	Time without symptoms or toxicity
UK	United Kingdom
WHO	World Health Organisation
XP	Capecitabine and cisplatin doublet chemotherapy

1. EXECUTIVE SUMMARY

This summary provides a brief overview of the key issues identified by the External Assessment Group (EAG) as being potentially important for decision making in its review of the company submission (CS) for the appraisal of pembrolizumab in combination with trastuzumab and chemotherapy for the treatment of adult patients with previously untreated locally advanced unresectable or metastatic human epidermal growth factor receptor 2 (HER2) positive gastric or gastro-oesophageal junction (GOJ) adenocarcinoma whose tumours express programmed death-ligand-1 (PD-L1) with a combined positive score (CPS) ≥ 1. It also includes the EAG's preferred assumptions and the resulting incremental cost-effectiveness ratios (ICERs) for the comparison of pembrolizumab in combination with trastuzumab and chemotherapy against the current standard of care (SoC) which is trastuzumab and chemotherapy.

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 and Section 1.7 provides a summary of the EAG's preferred assumptions and its base case ICER. Background information on the condition, technology, evidence used and information on non-key issues are in the main EAG report.

All issues identified represent the EAG's view, not the opinion of the National Institute for Health and Care Excellence (NICE).

1.1 Overview of the EAG's key issues

Table 1 Summary of the key issues identified in the EAG critique

ID3742	Summary of issue	Report sections
1	The use of a *post hoc* analysis of the non-Asia cohort which excluded data from the Asia region, but combined data from two other regions	3.2.3 and 4.3.3.1
2	Method used to extrapolate overall survival (OS) and progression-free survival (PFS) in the economic model by applying a hazard ratio (HR) from the non-Asia (CPS ≥1) cohort to parametric curves fitted to the comparator (SoC) arm of the global (CPS ≥1) cohort	4.3.3.2, 4.4.2.2 & 4.4.2.3
3	Utilities based on time-to-death rather than using utilities based on progression status (i.e., progressed disease versus progression-free)	4.3.3.4
4	Severity modifier is not based on the expected quality-adjusted life-years (QALYs) predicted by the company's cost-effectiveness analysis because this incorporates data from the Asia cohort which the company considers not generalisable to England	5

The key differences between the company's preferred assumptions and the EAG's preferred assumptions are the EAG's approach to modelling OS and PFS using curves fitted separately to the intervention (pembrolizumab plus SoC) and comparator (SoC) arms of Kaplan-Meier data (KM) from the non-Asia (CPS≥1) cohort. In comparison, the company used parametric curves fitted to data to the comparator arm of the global (CPS≥1) cohort to model OS and PFS for SoC and then applied a HR from the non-Asia (CPS≥1) cohort to estimate OS and PFS for pembrolizumab plus SoC.

1.2 Overview of key model outcomes

NICE technology appraisals compare how much a new technology improves length (OS) 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 survival which also increases the time spent in health states with higher quality of life as quality of life is related to time-to-death in the model
- Marginally reducing quality of life at the beginning of treatment due to adverse events (AEs).

Overall, the technology is modelled to affect costs by:

- Increasing the costs required for drug acquisition and administration
- Increasing disease management costs, mainly by extending the period of PFS
- Marginally increasing costs required to manage AEs and provide subsequent treatment
- Marginally reducing end-of-life costs.

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

- Whether the data from the non-Asia region or the global cohort are used to model OS for SoC
- Whether OS for pembrolizumab plus SoC is modelled by fitting parametric curves to the data from the intervention arm of the KEYNOTE-811 study or by applying a HR to the curve fitted to the global cohort SoC arm
- The choice of parametric curves used to extrapolate OS and PFS
- The assumption that time-to-death rather than progression status best predicts quality of life.

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

The EAG did not have any key issues related to the decision problem, however, it wishes to briefly highlight several discrepancies between the decision problem addressed in the CS and that specified in the NICE scope which are further described in Section 2.

Firstly, in the pivotal KEYNOTE-811 study, which forms the primary evidence supporting the license, only a minority of patients in the SoC arm received a platinum–fluoropyrimidine doublet chemotherapy regimen that is compatible with the NICE recommendation for trastuzumab with chemotherapy (TA208) in this patient population. The company considers that all doublet chemotherapies which combine a platinum-containing agent with a fluoropyrimidine are clinically equivalent. The EAG's clinical experts considered this assumption to be broadly acceptable and for this reason the EAG does not consider this discrepancy to be a key issue.

Secondly, in TA208, trastuzumab is recommended in combination with chemotherapy for patients with HER2-postitive metastatic gastric or GOJ adenocarcinoma, as the marketing authorisation for trastuzumab did not include patients with locally advanced disease. However, the CS assumes that patients with unresectable locally advanced disease are treated like patients with metastatic disease and a comparison against trastuzumab without chemotherapy has not been presented for patients with locally advanced disease. The EAG's clinical advisors stated that the distinction between unresectable locally advanced disease and metastatic disease cannot always be made without invasive investigations to identify peritoneal metastases, and they would therefore want to offer trastuzumab to any HER2-positive patients who are not suitable for perioperative chemotherapy and surgery. In addition, any patient who is contraindicated for trastuzumab would also be contraindicated for pembrolizumab in combination with trastuzumab and chemotherapy. For these reasons, the lack of a comparison against doublet chemotherapy without trastuzumab is not considered a key issue.

Thirdly, the CS does not provide a comparison against triplet chemotherapy (i.e., a platinum–fluoropyrimidine chemotherapy with the addition of epirubicin). The EAG did not consider this to be a key issue because the EAG's clinical advisors stated that triplet chemotherapy is not usually used as a first-line palliative treatment in this patient population, as it is not thought to improve survival compared with offering doublet chemotherapy, but it does increase toxicity.

Finally, the CS focuses on data from the population with PD-L1 CPS ≥ 1 as this is in line with the anticipated marketing authorisation. The EAG considers this to be broadly acceptable as the CPS ≥ 1 group was a pre-specified subgroup and CPS status (i.e. 0 or ≥ 1) was a stratification factor for randomisation. The CS assumes that tests to determine PD-L1 status are already part of routine care in the NHS in England for this population because PD-L1 CPS score is used to determine eligibility for other treatments already recommended by NICE in the HER2-ngative subgroup and PD-L1 testing is undertaken concurrently with HER2 testing to avoid delays.

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

Issue 1 The use of a *post hoc* analysis to define the non-Asia cohort

Report section	3.2.4 and 4.3.3.1
Description of issue and why the EAG has identified it as important	The company claims that data from the Asia region are not generalisable to England as screening programmes for gastric cancer are common in many Asia region countries, but screening is not routinely performed in England. The company's cost-effectiveness analysis is informed by a HR from the non-Asia (CPS≥1) cohort which was generated in a *post hoc* analysis by combining data from the other two regions (Western Europe/Israel/North America/Australia cohort and Rest of the World cohort) for patients with CPS≥1. The EAG questions the validity of such a subgroup analysis (Asia vs. non-Asia) due to the *post hoc* nature.
What alternative approach has the EAG suggested?	The EAG agrees with excluding the Asia (CPS≥1) cohort but considers that the Western Europe/Israel/North America/Australia (CPS≥1) cohort may be more applicable to clinical practice in England and was a pre-specified subgroup and stratification factor.
What is the expected effect on the cost-effectiveness estimates?	Including data from the Rest of World (CPS≥1) cohort provides more favourable estimates for both OS and PFS. An increase in the ICER is expected using data from only the Western Europe/Israel/North America/Australia (CPS≥1) cohort.
What additional evidence or analyses might help to resolve this key issue?	Using data from the Western Europe/Israel/North America/Australia (in CPS ≥ 1) cohort only as a scenario analysis in cost-effectiveness modelling.

1.5 The cost-effectiveness evidence: summary of the EAG's key issues

Only those issues identified in the EAG's critique of the economic modelling which have an important impact on ICER, or are otherwise expected to materially impact on the conclusions regarding the cost-effectiveness of pembrolizumab plus SoC compared to SoC, are discussed in Section 1.5. A full discussion of the main issues identified in the EAG's critique of the company's economic analysis can be found in Section 4.3, including those factors which were found to have a more modest impact on the ICER in the EAG's exploratory analyses described in Section 4.4. Any ICERs discussed in this Section are those generated without applying any QALY weighting for severity.

Issue 2 Method used to extrapolate OS and PFS

Report section	4.2.6.1, 4.3.3.2, 4.4.2.2 & 4.4.2.3
Description of issue and why the EAG has identified it as important	The company used a proportional hazards modelling approach to extrapolate OS and PFS. A parametric survival model was fitted to the KM data for the comparator arm of the global (CPS≥1) cohort, and a HR estimate obtained from the non-Asia (CPS≥1) cohort was then applied to the extrapolated control arm to estimate the survival in the intervention arm. The use of OS and PFS data from the global (CPS≥1) cohort to extrapolate OS and PFS for the control arm contradicts the company's position that the non-Asia (CPS≥1) cohort is considered the most generalisable to the eligible population in England. The company's extrapolated curve for the intervention arm for both OS and PFS does not fit the intervention arm data from the non-Asia (CPS≥1) cohort.
What alternative approach has the EAG suggested?	Modelling OS and PFS using models fitted independently to the intervention and comparator arms of KM from the non-Asia (CPS≥1) cohort.
What is the expected effect on the cost-effectiveness estimates?	The EAG's exploratory analysis which uses OS curves fitted independently to both arms of the non-Asia (CPS≥1) cohort substantially increased the ICER from £[redacted] to £[redacted] per QALY. The analysis using the EAG's preferred approach for PFS had a smaller impact and decreased the ICER to £[redacted] per QALY. However, these analyses combined both the EAG's preference for using the non-Asia (CPS≥1) cohort to model the OS and PFS in the comparator arm and the EAG's preference to model the intervention arm separately instead of using a HR approach. The choice of data used to model OS in the comparator arm is expected to be the main driver of change in the ICER.
What additional evidence or analyses might help to resolve this key issue?	The EAG suggests that the company should also explore the use of OS and PFS data based on the Western Europe/Israel/North America/Australia (CPS ≥ 1) cohort as per Issue 1.

Issue 3 Utility analysis

Report section	4.2.6.2 & 4.3.3.4
Description of issue and why the EAG has identified it as important	Utility data used in the company's base case were based on the non-Asia (CPS≥1) cohort. Utility values were estimated based on a time-to-death approach with four categorical groups (<30 days; 30 to 179 days; 180 to 359 days, and ≥360 days) using descriptive statistics. The EAG considers that there is considerable uncertainty related to whether using a time-to-death approach for estimating utility is preferential to a progression-based approach that has historically been more widely used, and the company's estimated utility values lack face validity as the values for patients with a time-to-death >360 days are very similar to the age-adjusted utility values expected for the general population.
What alternative approach has the EAG suggested?	Analysing utility data using a linear mixed effect regression model for both time-to-death and progression-based approaches.
What is the expected effect on the cost-effectiveness estimates?	Using the time-to-death utility estimates from the linear mixed effect regression increased the ICER from £[redacted] to £[redacted] per QALY when applying these to the company's base case analysis. The ICER for the EAG's preferred base case scenario increased from £[redacted] to £[redacted] per QALY when switching from the time-to-death utilities to the progression-based utilities but still using the linear mixed effect regression approach.
What additional evidence or analyses might help to resolve this key issue?	The EAG suggests that the company should also explore using utility data from the Western Europe/Israel/North America/Australia (CPS ≥ 1) cohort as per Issue 1.

1.6 Other key issues: summary of the EAG's view

Issue 4 Severity modifier

Report section	Section 5
Description of issue and why the EAG has identified it as important	The company has applied a QALY weighting of 1.2 in its base case. The estimate of expected lifetime QALYs from the company's model for patients receiving current SoC provides estimates of absolute and proportional QALY shortfall, compared to age and sex matched members of the general population without gastric or GOJ cancer, which would support a QALY weighting of 1.0. However, the company argues that the estimate of expected lifetime QALYs in the SoC arm of the model is unrealistic because it is based on OS data that includes patients from the Asia region who have higher survival and which the company considers is not generalisable to clinical practice in England due to the widespread use of gastric cancer screening in Asia region countries. The company therefore prefers to use an estimate of the expected lifetime QALYs under current SoC from the appraisal of trastuzumab in combination with chemotherapy (TA208) which would support a QALY weighting of 1.2. The EAG argues that if the OS and PFS data from the Asia (CPS≥1) region are not considered generalisable to England, then the company should use data from the non-Asia (CPS≥1) region to estimate OS and PFS under SoC, but for consistency, these data should also be used to inform the QALYs used to estimate the ICERs in the cost-effectiveness analysis (see Issue 2).
What alternative approach has the EAG suggested?	The estimate of expected lifetime QALYs for the SoC arm from the EAG's preferred base case scenario is lower than predicted by the company's base case because the EAG has used data from the non-Asia (CPS≥1) cohort to estimate OS and PFS in the SoC arm. The estimate from the EAG's preferred base case would support a QALY multiplier of 1.2.
What is the expected effect on the cost-effectiveness estimates?	The committee's judgement regarding the cost-effectiveness of pembrolizumab plus SoC compared to SoC alone is dependent on both the ICER and the choice of QALY weight. These are both dependent on whether data from the non-Asia (CPS≥1) cohort are considered more generalisable to clinical practice in England than the data from the global (CPS≥1) cohort which include patients recruited in the Asia region.
What additional evidence or analyses might help to resolve this key issue?	The EAG suggests that the company should also explore what QALY weight would be supported by analyses using data from Western Europe/Israel/North America/Australia (CPS ≥ 1) cohort.

1.7 Summary of EAG's preferred assumptions and resulting ICER

A summary of the EAG's exploratory analyses is provided in Table 2. Each of the individual changes included in the EAG's preferred base case is presented as a single change applied to the company's base case model. The EAG's preferred base case, which combines all of these changes, is then presented. This is followed by scenario analyses which use the EAG's preferred base case as their starting point. All of the results presented have been generated using the deterministic model, with the exception of the EAG's preferred base case for which both deterministic and probabilistic results are provided. Full details on the methods used in EAG of the analyses conducted by the EAG is provided in Section 4.4.2. These results include the company's patient access scheme (PAS) price for pembrolizumab but do not include any confidential PAS prices or confidential prices from the commercial medicines unit (CMU) for any other drugs. These can be found in the confidential appendix. All ICERs discussed in the text below are those generated without a QALY weighting for severity. For reference, ICERs are provided in Table 2 both without a QALY weighting and when using a QALY weighting of 1.2.

The EAG's preferred estimate of the ICER is £[redacted] per QALY (£[redacted] when the probabilistic analysis). This is substantially higher than the company's base case estimate of £[redacted] per QALY (£[redacted] when using the probabilistic analysis). The main reason for this difference is that the EAG's preferred approach to modelling OS used parametric survival curves fitted to each arm of the KM data from the non-Asia (CPS≥1) cohort, whereas the company used data from the global (CPS≥1) cohort to estimate OS in the SoC arm, including data from the Asia region where OS survival was higher. The company then estimated OS in the pembrolizumab plus SoC arm by applying a HR from the non-Asia (CPS≥1) cohort to the parametric curve fitted to the SoC arm of the global (CPS≥1) cohort. The EAG's scenario analysis provide an ICER that ranges from £[redacted] to £[redacted] per QALY. The lower range is provided by a scenario applying alternative parametric survival curves (log-logistic for both PFS and OS) still fitted independently to each arm of the non-Asia (CPS≥1) cohort. The upper range is provided by the scenario using progression-based rather than time-to-death based utilities.

Table 2 Summary of the results of the EAG's exploratory analyses

Option	LYs	QALYs	Costs	Incremental		ICER (QALY weight x1)	ICER (QALY weight x1.2)
				QALYs	Costs		
Company base case – post-clarification							
SoC*	3.03	[redacted]	[redacted]	-	-		
Intervention**	4.94	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG exploratory analysis 1: correcting programming and implementation errors in the company's economic model							
SoC*	3.03	[redacted]	[redacted]	-	-	-	-
Intervention**	4.94	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG exploratory analysis 2: Using the EAG's preferred survival extrapolation for OS							
SoC*	1.59	[redacted]	[redacted]	-	-	-	-
Intervention**	2.17	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG exploratory analysis 3: Using the EAG's preferred survival extrapolation for PFS							
SoC*	3.03	[redacted]	[redacted]	-	-	-	-
Intervention**	4.94	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG exploratory analysis 4: Removing the cap for TTD of trastuzumab							
SoC*	3.03	[redacted]	[redacted]	-	-	-	-
Intervention**	4.94	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG exploratory analysis 5: Applying lower administration costs for trastuzumab when administered without pembrolizumab							
SoC*	3.03	[redacted]	[redacted]	-	-	-	-
Intervention**	4.94	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG exploratory analysis 6: Assuming subsequent therapy to include only taxanes and applying that to only a proportion of PFS events who get progressed (25% get paclitaxel and 25% get docetaxel)							
SoC*	3.03	[redacted]	[redacted]	-	-		
Intervention**	4.94	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG exploratory analysis 7: Limiting outpatient visits to 6 weekly after chemotherapy and adding CT scans 4 times per annum for patients on PFS							
SoC*	3.03	[redacted]	[redacted]	-	-	-	-
Intervention**	4.94	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG exploratory analysis 8: Increasing outpatient visits and CT scans to 4 times per annum for patients with progressed disease							
SoC*	3.03	[redacted]	[redacted]	-	-	-	-
Intervention**	4.94	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG exploratory analysis 9: Time-to-death utilities estimated using a linear mixed effects model							
SoC*	3.03	[redacted]	[redacted]	-	-	-	-
Intervention**	4.94	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Option	LYs	QALYs	Costs	Incremental		ICER (QALY weight x1)	ICER (QALY weight x1.2)
				QALYs	Costs		
EAG preferred base case scenario							
EAG base case applying analyses 1-9 (Deterministic)							
SoC*	1.59	[redacted]	[redacted]	-	-	-	-
Intervention**	2.17	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG base case applying analyses 1-9 (Probabilistic)							
SoC*	1.61	[redacted]	[redacted]	-	-	-	-
Intervention**	2.21	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Scenario analyses applying individual changes to the EAG base case							
EAG scenario 1 (Assuming a log-logistic curve for OS and PFS extrapolations)							
SoC*	1.84	[redacted]	[redacted]	-	-		
Intervention**	2.50	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG scenario 2 (Using restricted mean duration to estimate costs for first-line chemotherapy)							
SoC*	1.59	[redacted]	[redacted]	-	-	-	-
Intervention**	2.17	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG scenario 3 (Reducing the cap applied to TTD of first-line chemotherapy to 4 cycles)							
SoC*	1.59	[redacted]	[redacted]	-	-		
Intervention**	2.17	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG scenario 4 (Using utility values based on progression status)							
SoC*	1.59	[redacted]	[redacted]	-	-		
Intervention**	2.17	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG scenario 5 (Assuming 100% of doublet chemotherapy is with XP)							
SoC*	1.59	[redacted]	[redacted]	-	-		
Intervention**	2.17	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

CT – computerised tomography; EAG – external assessment group, ICER – incremental cost-effectiveness ratio; LYs - life-years; OS – overall survival, PFS – progression-free survival, QALYs - quality-adjusted life-years; TTD – time to treatment discontinuation; XP - cisplatin with capecitabine

** SoC: Trastuzumab plus chemotherapy*

*** Intervention: Pembrolizumab with SoC*

2 BACKGROUND

This section presents a brief summary and critique of the company's description of locally advanced unresectable or metastatic gastric or gastro-oesophageal junction adenocarcinoma which is classed as human epidermal growth factor receptor 2 (HER2) positive and the current treatment pathway in England for this disease. This is followed by a critique of the decision problem addressed in the company submission (CS).[1]

2.1 Critique of company's description of underlying health problem

The External Assessment Group (EAG) is broadly satisfied with the company's description of the underlying health problem. The CS (Section B.1.3)[1] describes gastric cancer as the fifth most common cancer worldwide, accounting for 2% of all new cancers in the United Kingdom (UK).[2, 3] Approximately half of all cases of gastric cancers are diagnosed in people aged 75 and over.[3] Gastric cancers are generally classified into those occurring where the oesophagus meets the stomach, referred to in the CS as cancers of the gastro-oesophageal junction (GOJ), and those arising elsewhere in the stomach, which the CS refers to as gastric cancer.[1] Adenocarcinomas, which are the type of gastric and GOJ cancer addressed in the CS, are the most common histological subgroup of gastric, GOJ and oesophageal cancer.[4] In addition to adenocarcinomas, there are other types of cancer which can occur in the stomach (for example gastrointestinal stromal tumours, neuroendocrine tumours, lymphomas), but these fall outside of the scope of this appraisal. The CS uses the Laurén classification, which is commonly used in clinical trials, and which categorises adenocarcinomas into four types: intestinal type, diffuse type, indeterminate type and unclassified type.[1, 5]

The CS describes the staging of cancer as dependent on whether the cancer is localised to the stomach (stage 1) or has spread beyond the stomach.[1] Locally advanced disease (stage 2 or 3) is when the cancer has spread to the surrounding tissues. The treatment for locally advanced disease is usually surgery to remove the affected area, which is described as surgical resection. However, in some cases this is not possible and then the disease is classified as unresectable locally advanced disease. If the cancer has spread beyond the stomach and surrounding tissues to the abdominal lining (peritoneum), bones or other organs, this is described as metastatic disease (stage 4). The proportion of patients diagnosed with metastatic disease was 44.9% in 2020/21.[6] The focus of the CS is on patients with unresectable locally advanced disease or metastatic disease, who have a poorer prognosis than those who can be treated by surgical resection.[1, 7] The EAG's clinical advisors stated that in practice the distinction between unresectable locally advanced and metastatic disease is not always easy to make as those classified as having unresectable locally advanced disease may also have undetected peritoneal metastases. Invasive investigations to distinguish metastatic from locally advanced disease, such as laparoscopy to identify peritoneal metastases, are only recommended if they will help guide ongoing management[8], and are

therefore not undertaken if the patient has already been deemed not suitable for perioperative chemotherapy and surgery and the decision has been made to offer first-line palliative chemotherapy.

The CS states that gastric and GOJ cancers are often diagnosed when the disease is as at an advanced stage, with 60% of people not eligible for curative treatment due to late presentation or comorbidities.[9] This is partly because the common symptoms - indigestion, poor appetite, weight loss, abdominal pain, and difficulty swallowing - can be vague, absent or not recognised as being potentially indicative of a serious health condition such as cancer.[1]

2.2 Critique of company's overview of current service provision

The CS describes the current treatment pathway for patients with locally advanced unresectable or metastatic gastric or GOJ HER2-positive adenocarcinoma, which is the population specified in the NICE scope. The treatment pathway for gastric cancer and the company's proposed positioning of pembrolizumab within the pathway is summarised in Figure 1. The CS states that there is currently no national screening programme for gastric cancer.[1] It describes palliative chemotherapy as being the first-line treatment for people with Eastern Cooperative Oncology Group (ECOG) performance status of 0 to 2 and no significant comorbidities. CS, Figure 2 (reproduced here as Figure 1)[1] highlights TA208 which recommends trastuzumab in combination with cisplatin and capecitabine or 5-fluorouracil as a treatment option for people with HER2-positive metastatic adenocarcinoma of the stomach or GOJ.[10] It also highlights TA191 which recommends capecitabine in combination with a platinum-based regimens for the first-line treatment of inoperable gastric cancer.[11]

The CS mentions NICE's guideline 83 on assessment and management in adults with oesophago-gastric cancer (NG83),[8] which recommends doublet chemotherapy as an option for first-line palliative combination chemotherapy.[1] However, rather than describing all the possible platinum–fluoropyrimidine doublet chemotherapy combinations that are possible under NG83, the company's description of the current treatment pathway (CS, pages 20-21) focuses on those doublet chemotherapy options that are recommended in combination with trastuzumab within TA208.[1] Those are cisplatin plus capecitabine (referred to as XP in the CS, where X is capecitabine and P is cisplatin) and cisplatin plus fluorouracil (referred to as FP in the CS, where F refers to the fluorouracil component - sometimes abbreviated to 5-FU). However, these are only two of the four possible platinum–fluoropyrimidine doublet chemotherapy options available. For clarity, the EAG has provided Table 3, which summarises the possible platinum–fluoropyrimidine doublet chemotherapy options available under NG83.[8] These also include capecitabine plus oxaliplatin (referred to as CAPOX in the submission) or oxaliplatin plus 5-FU, which is usually accompanied by folinic acid (referred to as FOLFOX in the submission). The CS states that, "*ESMO [European Society for Medical Oncology] guidelines and clinical opinion suggest that doublet chemotherapies (cisplatin and oxaliplatin; 5FU and capecitabine) are clinically*

equivalent[12, 13].”[1] The EAG’s clinical advisors agreed that there was no strong evidence to suggest that one particular doublet chemotherapy combination was more clinically effective than another. However, they stated that capecitabine, which is an oral treatment, was generally preferred to 5-FU which needs to be given intravenously over an extended period. The exception was when patients were unable to swallow oral medication, in which case 5-FU was a useful alternative. In addition, oxaliplatin was generally considered to have a better side-effect profile than cisplatin but clinicians in England were restricted by TA208 to using cisplatin when giving chemotherapy alongside trastuzumab.

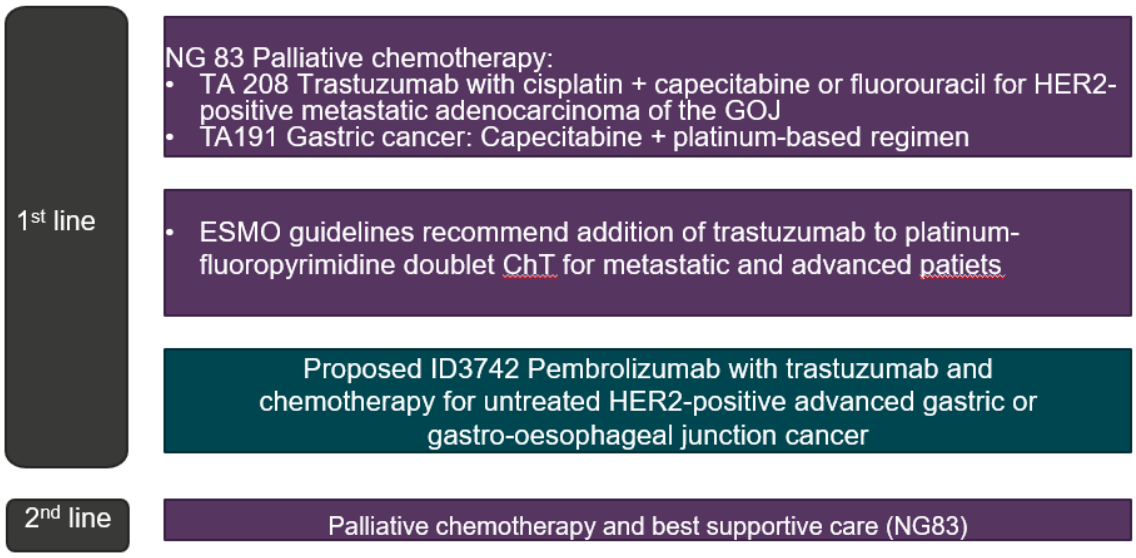
The CS also states that the ESMO guidelines recommend platinum–fluoropyrimidine doublet chemotherapy with trastuzumab as a standard of care in patients with advanced metastatic HER2-positive gastric cancer or GOJ adenocarcinoma.[1, 7] The company’s description of current practice in CS, Section B.1 does not highlight the fact that the wording in TA208 specifically states that trastuzumab with chemotherapy is recommended for metastatic cancer.[1,10] Instead, the CS highlights that treatment of locally advanced patients with unresectable disease is similar in clinical practice to those with metastatic disease (CS, pages 73 & 75).[1] The EAG’s clinical advisors agreed that they would want to use trastuzumab in HER2-positive patients regardless of whether the patient’s disease was classified as metastatic disease or unresectable locally advanced disease, given that the distinction between the two may not be clear when there may be undetected peritoneal metastases and they would expect a similar treatment response. The EAG’s clinical advisors also stated that there is a small group of patients HER2-positive disease, in whom trastuzumab may be contraindicated, usually due to cardiac comorbidities. Doublet chemotherapy without trastuzumab would be an option in some of these patients, provided they are not also contraindicated for chemotherapy. However, the EAG notes that any patient who is ineligible for trastuzumab would also not be eligible for pembrolizumab given in combination with trastuzumab and chemotherapy.

The EAG notes that NG83 recommends both doublet and triplet chemotherapy regimens as options for first-line palliative combination chemotherapy in people with advanced oesophago-gastric cancer.[8] Triplet chemotherapy as defined within NG83 comprises of epirubicin in combination with doublet chemotherapy (see Table 3).[8] However, the CS states that triplet chemotherapy regimens, “*do not have a role in treating HER2 positive metastatic or locally advanced GC [gastric cancer] or GOJ adenocarcinoma due to increased toxicity and lack of added clinical effect*.” The EAG’s clinical advisors stated that in current practice in England, triplet chemotherapy was not usually used as a first-line palliative treatment in patients with metastatic or locally unresectable advanced disease, despite being an option under NG83. This is because it is not thought to improve survival compared with offering doublet chemotherapy, but it does increase toxicity. The clinical advisors said that it is sometimes used in a small minority of patients with locally advanced disease with the aim of reducing tumour size to allow surgical resection, but that this is different from using it palliatively for those with

unresectable disease. However, the triplet chemotherapy combination recommended by ESMO in this downstaging neoadjuvant indication is a taxane containing regimen rather than the epirubicin containing triplet regimen recommended in NG83.[7, 8]

CS, Section B.1 does not specify doses or duration of treatment for any of the chemotherapy regimens that form part of the current standard treatment pathway.[1] The relevance of the treatment regimens received in the comparator arm of the pivotal KEYNOTE-811 study and the comparator treatments assumed in economic analysis to current clinical practice in discussed further in Sections 3 and 4 respectively.

Figure 1 Gastric cancer treatment pathway and proposed pembrolizumab positioning (reproduced from CS, Figure 2)[1]



Abbreviations: ChT, chemotherapy; ESMO, European Society of Medical Oncology; HER2, human epidermal growth factor receptor 2; GOJ, gastro-oesophageal junction

The CS states that HER2 testing and programmed death-ligand-1(PD-L1) testing have become part of routine care for patients with gastric and GOJ cancer,[1] because HER2 testing is necessary in order to determine whether patients are eligible for trastuzumab under TA208,[10] and PD-L1 testing is necessary to determine eligibility for nivolumab in patients with HER2-negative disease under TA857.[14] The EAG noted that PD-L1 testing would also be required in patients with HER2-negative GOJ cancer under TA737.[15] The EAG's clinical advisors agreed that testing for both HER2 and PD-L1 was standard clinical practice and both tests would be requested at the same time to expedite treatment, rather than PD-L1 testing being requested only in those who are HER2-negative. The current summary of produce characteristics (SmPC) for pembrolizumab states, "*when assessing the PD-L1 status of the tumour, it is important that a well-validated and robust methodology is chosen to minimise false negative or false*

positive determinations."[16] The EAG noted that the PD-L1 testing assay used in KEYNOTE-811 (PD-L1 1HC 22c3 pharmDx™),[17] was different from the assay used in CheckMate 649 (PD-L1 IHC 28-8 pharmDx™),[18] the main clinical trial considered in TA857. In the company's response to clarification question A4,[19] the company stated that both of these testing assays were used routinely in National Health Service (NHS) clinical practice and that published studies concluded that "*PD-L1 22C3 and 28-8 pharmDx assays were highly comparable at CPS cut-offs of 1, 10, and 50 in gastric cancer. These results provide evidence for the potential interchangeability of the two PD-L1 assays in gastric cancer".*[20] Based on this response, the EAG was satisfied that either assay could be used in clinical practice to assess suitability for pembrolizumab or nivolumab. The EAG also noted advice from clinical experts that some centres have access to only one assay, whereas other centres have access to multiple assays and clinicians may request both assays for patients with GOJ cancer.

Table 3 Summary of different treatment combinations for first-line treatment

Combination abbreviation	PD-L1 targeted therapy	HER2 targeted therapy	Doublet chemotherapy		Addition for triplet chemotherapy	NICE recommended	Included KEYNOTE -811	Included in model
			Platinum	Fluoro-pyrimidine				
Pembrolizumab & trastuzumab & XP	Pembrolizumab	Trastuzumab	Cisplatin	Capecitabine	NA	Subject of current appraisal [4]	No	Yes, as scenario
Pembrolizumab & trastuzumab & FP	Pembrolizumab	Trastuzumab	Cisplatin	5-FU	NA		Yes	Yes
Pembrolizumab & trastuzumab & CAPOX	Pembrolizumab	Trastuzumab	Oxaliplatin	Capecitabine	NA		Yes	Yes
Pembrolizumab trastuzumab & FOLFOX	Pembrolizumab	Trastuzumab	Oxaliplatin	5-FU	NA		No	No
Trastuzumab & XP	None	Trastuzumab	Cisplatin	Capecitabine	NA	TA208: Untreated HER2-positive metastatic cancer of stomach or GOJ[10]	No	Yes, as scenario
Trastuzumab & FP	None	Trastuzumab	Cisplatin	5-FU	NA		Yes	Yes
Trastuzumab & CAPOX	None	Trastuzumab	Oxaliplatin	Capecitabine	NA	No	Yes	Yes
Trastuzumab & FOLFOX	None	Trastuzumab	Oxaliplatin	5-FU	NA	No	No	No
XP	None	None	Cisplatin	Capecitabine	NA	NG83: Doublet chemotherapy as first-line palliative treatment of locally advanced or metastatic oesophago-gastric cancer[8]	No	No
FP	None	None	Cisplatin	5-FU	NA		No	No
CAPOX	None	None	Oxaliplatin	Capecitabine	NA		No	No
FOLFOX	None	None	Oxaliplatin	5-FU	NA		No	No
ECX	None	None	Cisplatin	Capecitabine	Epirubicin	NG83: Triplet chemotherapy as first-line palliative treatment of locally advanced or metastatic oesophago-gastric cancer[8]	No	No
ECF	None	None	Cisplatin	5-FU	Epirubicin		No	No
EOX	None	None	Oxaliplatin	Capecitabin	Epirubicin		No	No
EOF	None	None	Oxaliplatin	5-FU	Epirubicin		No	No

Abbreviations: 5-FU, fluorouracil; GOJ, Gastro-Oesophageal Junction Cancer; HER2, Human Epidermal Growth Factor Receptor; PD-L1, Programmed death-ligand-1

2.3 Critique of company's definition of the decision problem

2.3.1 Population

The population addressed in the CS is patients with previously untreated locally advanced unresectable or metastatic HER2-positive gastric or GOJ adenocarcinoma whose tumours express PD-L1 with a combined positive score (CPS)≥ 1 (CS, Table 1).[1] This is described by the company as being in-line with the proposed marketing authorisation.[17] The proposed indication is for first-line treatment, and the KEYNOTE-811 study, which forms the primary evidence supporting the license, was restricted to patients who have not had previous treatment for metastatic or locally advanced unresectable disease, but included those who had received prior neoadjuvant or adjuvant therapy completed at least 6 months prior to randomization provided there was no evidence of progression within that timeframe[21]. This specification of the population does not exclude patients who received adjuvant or neoadjuvant treatment for localised disease earlier in their treatment pathway. The proposed marketing authorisation in the draft SmPC is also restricted to adults as per the population specified in the NICE scope.[4, 17]

The NICE scope specified that subgroups should be considered according to whether the patient has metastatic disease or locally advanced disease.[4] The CS states that the subgroup with locally advanced unresectable disease was not a pre-specified subgroup in the KEYNOTE-811 study, and it makes up only 3% of the trial population (CS, Table 1).[1] For these reasons, the CS does not present results for the locally advanced subgroup, however, results are provided for a limited set of outcomes (PFS and OS) for the metastatic subgroup (CS, Appendix E).[1] The EAG's clinical advisors considered that this approach was reasonable given the small numbers of patients with unresectable locally advanced disease in the KEYNOTE-811 trial and the fact that they would expect them to respond similarly to the treatment. In addition, the EAG's clinical advisors stated that in practice the distinction between locally advanced unresectable and metastatic disease is not always easy to make as those classified as having locally advanced unresectable disease may also have undetected peritoneal metastases. The NICE scope did not specify that subgroup analyses should be considered for gastric versus GOJ cancer.[4] The EAG's clinical advisors noted that the distinction between gastric and GOJ cancer may not always be clear cut given that gastric cancer can spread to the GOJ. In addition, the distinction between gastric and GOJ cancer is more relevant when considering surgical treatment options and less relevant when considering first-line palliative chemotherapy options as the palliative management of gastric and GOJ cancer would be similar for patients with HER2-positive disease.

The NICE scope also specified that subgroup analyses should be provided by PD-L1 status, without specifying what level of PD-L1 expression should be used to categorise patients The CS focuses on presenting results for the subgroup with PD-L1 CPS ≥ 1 as this restriction is included in the anticipated marketing authorisation.[1, 4] The EAG considers that it is reasonable for the CS to present results for the CPS ≥ 1 subgroup as PD-L1 status (CPS 0 versus CPS ≥ 1) was used as a stratification factor for

randomisation.[21] The EAG notes that other CPS scores have been used to defined PD-L1 status in other indications for pembrolizumab. For example, in the pembrolizumab indication for HER2-negative GOJ cancer, treatment with pembrolizumab is restricted using a PD-L1 CPS cut-off score of 10.[16] This was because a more pronounced treatment benefit was demonstrated for CPS ≥ 10 than CPS <10, in the KEYNOTE-590 study, although it should be noted that PD-L1 status was not a stratification factor in that study.[22] The EAG noted that the study protocol for KEYNOTE-811 stated in its study rationale section that based on previous studies in advanced gastric cancer, pembrolizumab demonstrated a high level of tumour response regardless of PD-L1 status (KEYNOTE-811 protocol, page 37).[21] This may explain why patients with a PD-L1 CPS score of zero were included when pembrolizumab is known to specifically target PD-L1 receptors. Overall, the EAG is satisfied that the CS has focused on data from the CPS ≥ 1 subgroup, given that this is consistent with the anticipated marketing authorisation.

In the CS the company argues that data from patients recruited in the Asia region, where screening programmes for upper gastrointestinal cancer are more widespread, may not be as appliable to countries such as England which do not have a screening programme.[1] Therefore, the CS focusses on presenting data from the 'non-Asia' region. However, it should be noted that this relates to the country in which the patients were recruited and not to the race/ethnicity of the individual. A separate subgroup analysis for race is presented for Asian and non-Asian patients, but this subgroup is not a focus of the submission, and this subgroup analysis should not be confused with the subgroup analysis for the non-Asia region.

2.3.2 Intervention

The intervention given in the NICE scope is pembrolizumab in combination with trastuzumab and chemotherapy.[4] The draft SmPC is slightly more specific and states that pembrolizumab is indicated for use in combination with trastuzumab, fluoropyrimidine and platinum-containing chemotherapy.[17] The exact combinations of chemotherapy agents that pembrolizumab can be given with are discussed further under Section 2.3.3 where comparator treatments are described.

The dose of pembrolizumab specified in the draft SmPC is 200 mg by intravenous infusion on day 1 of a 3-week cycle or 400 mg by intravenous infusion given on day 1 of a 6-week cycle.[17] The draft SmPC states that pembrolizumab should be continued until [redacted]

The EAG notes that CS, Table 2 states that that the method of administration is 200 mg every 3 weeks up to a maximum duration of 35 cycles,[1] but [redacted].[17] Treatment duration in the KEYNOTE-811 trial was until disease progression or unacceptable toxicities, up to a maximum of 35 doses.[21] There was also the option for a second course of pembrolizumab (up to 17 cycles) in

KEYNOTE-811, but only in patients who met specific criteria and this was a rare occurrence (see Section 3.2.1.2).

2.3.3 Comparators

The comparator intervention addressed in the CS is described as trastuzumab with cisplatin plus capecitabine or fluorouracil (CS, Table 1).[1] The EAG notes that in TA208, trastuzumab is only recommended in combination with either XP or FP (see Table 3).[10] Therefore, the company's stated comparator contains the two doublet chemotherapy regimens that are most applicable to clinical practice in England. However, the comparator in the KEYNOTE-811 trial, which is they key source of evidence for the CS, only included CAPOX and FP, with a minority of patients receiving FP.[21] Also, the draft SmPC for pembrolizumab does not specify an exact combination of fluoropyrimidine and platinum-containing chemotherapy,[17] and therefore it could be interpreted as being indicated in combination with trastuzumab and any of the four possible doublet chemotherapy regimens, as shown in Table 3. The EAG also notes that various biosimilar versions of trastuzumab are now available and the EAG uses the term trastuzumab to refer to any medicine licensed as being biosimilar to the reference medicine for trastuzumab, which was Herceptin.

Doublet chemotherapy and triplet chemotherapy without trastuzumab are also included as comparators in the NICE scope,[4] but these are not addressed as comparators in the CS.[1] The company's rationale for excluding doublet chemotherapy without trastuzumab is that ESMO guidelines and clinical opinion suggest that locally advanced unresectable and metastatic gastric or GOJ adenocarcinoma are treated like metastatic disease.[1] The EAG considered that this was reasonable based on the advice from its clinical advisors (see Section 2.1 and 2.2) that it is not always possible to distinguish between metastatic and unresectable locally advanced disease due to the potential presence of undetected peritoneal metastases, and that they would want to offer trastuzumab to any HER2-positive patients. Although doublet chemotherapy treatment without trastuzumab may be offered when trastuzumab is contraindicated, any patient contraindicated for trastuzumab will also be contraindicated for pembrolizumab as this is given in combination with trastuzumab. Therefore, the group eligible for doublet chemotherapy is unlikely to be eligible for pembrolizumab under its proposed marketing authorisation and the EAG accept that it is reasonable for the company not to have provided a comparison against doublet chemotherapy. The EAG also accepts the company's rationale for excluding triplet chemotherapy (see Table 3 for examples of triplet chemotherapy) based on advice from its clinical advisors (see Section 2.2) that triplet chemotherapy is likely to increase toxicity without improving survival.

2.3.4 Outcomes

The CS states that it has addressed all outcomes specified in the NICE Scope (see Table 4).[1] The EAG agrees that the CS addresses both overall survival (OS) and progression-free survival (PFS). For the primary outcome in KEYNOTE-811, the progression component of PFS was reported by blinded independent central review (BICR) using Response Evaluation Criteria In Solid Tumours (RECIST) 1.1 criteria.[1] PFS using investigator assessment of progression using a modified version of RECIST 1.1 for immune therapies (iRECIST) was an exploratory objective.[1] Response was reported using overall response rate (ORR) with ORR defined as the proportion having either complete or partial response. Adverse events (AEs) and health-related quality of life (HRQoL) were both addressed in the CS, however, the EAG noted that the reporting of HRQoL in Doc B, Section 2, was limited to EuroQoL Visual Analogue Scale (EQ-VAS) outcomes and other patient reported outcomes measures (European Organisation for Research and Treatment of Cancer [EORTC] Quality-of-Life Questionnaire Core 30 [QLQ-C30] and EORTC QLQ-STO22) were not reported.[1] Utility outcomes based on EuroQol 5 dimensions 5 level (EQ-5D-5L) by trial arm were not summarised in the original CS,[1] however, these were provided in response to the clarification letter (question A22).[19]

2.3.5 Other relevant factors

The company does not report any equality considerations in CS, Section B.1.4.[1] The company has provided an assessment of the severity modifier and has applied a quality-adjusted life-year (QALY) weighting of 1.2 based on its assessment of the absolute and proportional QALY shortfall (CS, Section B.3.6).[1] The company also notes in CS Section B1.3, that recent previous appraisals of treatments for this indication met the now superseded End-of-Life criteria.[1] The EAG's critique of the company's assessment of the appropriate severity modifier is provided in Section 5.

Table 4 The decision problem (adapted from CS, Table 1[1] with minor amendments and comments from the EAG)

	Final scope issued by NICE[4]	Decision problem addressed in the CS and rationale if different from NICE scope	EAG comments
Population	Adults with untreated locally advanced unresectable or metastatic HER2-positive gastric or GOJ adenocarcinoma	Patients with untreated locally advanced unresectable or metastatic HER2-positive gastric or GOJ adenocarcinoma whose tumours express PD-L1 with a CPS≥ 1. Population is based on the proposed marketing authorisation wording.	The CS focuses on presenting trial outcomes for the PD-L1 positive subgroup of the KEYNOTE-811 trial, defined as those with a CPS ≥ 1, who made up 85% of the global cohort.[1] The company claims that PD-L1 status is already being routinely assessed in current practice.
Intervention	Pembrolizumab with trastuzumab and chemotherapy	In line with final scope	The EAG notes that the draft SmPC specifies that pembrolizumab is indicated in combination with trastuzumab, fluoropyrimidine and platinum-containing chemotherapy[17]
Comparators	• Chemotherapy only, which includes: o doublet treatment with fluorouracil or capecitabine in combination with cisplatin or oxaliplatin o triplet treatment with fluorouracil or capecitabine in combination with cisplatin or oxaliplatin plus epirubicin Trastuzumab with cisplatin plus capecitabine or fluorouracil	Trastuzumab with cisplatin plus capecitabine or fluorouracil KEYNOTE-811 trial results provide direct evidence between: • pembrolizumab with trastuzumab and CAPOX or FP vs. • trastuzumab plus CAPOX or FP for locally advanced unresectable or metastatic HER2-positive gastric or GOJ adenocarcinoma. Based on previous appraisals in this setting, ESMO guidelines and clinical opinion received, doublet chemotherapy regimens are considered to be clinically equivalent.	The EAG considers that the company's assumption that all doublet chemotherapy regimens are clinically equivalent is broadly acceptable. However, it notes that in TA208, trastuzumab is only recommended in combination with cisplatin and either capecitabine or 5-FU (i.e., XP or FP).[10] Adherence to this guidance in clinical practice restricts the choice of doublet chemotherapy given in combination with trastuzumab for patients with metastatic disease. In addition, whilst TA208 is strictly speaking restricted to patients with HER2-postivie metastatic disease, in clinical practice trastuzumab with doublet chemotherapy is usually offered to those with HER2-positive

	Final scope issued by NICE[4]	Decision problem addressed in the CS and rationale if different from NICE scope	EAG comments
		ESMO guidelines and clinical opinion suggest that locally advanced unresectable and metastatic gastric or GOJ adenocarcinoma are treated like metastatic disease; therefore, a comparison versus chemotherapy without trastuzumab has not been conducted and is not presented in this submission.	locally advanced unresectable disease. This is because in practice the distinction between locally advanced unresectable and metastatic disease is not always easy to define as those classified as having locally advanced unresectable disease may have undetected peritoneal metastases. The EAG's clinical experts agreed that triplet chemotherapy is not usually used as a first-line palliative treatment in this patient group because it is likely to increase toxicity without improving survival compared with offering doublet chemotherapy.
Outcomes	• overall survival • progression-free survival • response rate • adverse effects of treatment • health-related quality of life.	In line with final scope	The EAG notes that for HRQoL outcomes, the clinical efficacy section of the CS (CS, Document B, Section B.2) only summarises trial outcomes for the EQ-VAS but data were also collected for both EORTC QLQ-C30 and EORTC QLQ-STO22.[1] Utilities by trial arm based on EQ-5D were provided in response to the clarification request.
Subgroups to be considered	• PD-L1 status • Locally advanced unresectable • Metastatic	• PD-L1 status KEYNOTE-811 included less than 3% of locally advanced unresectable population which was not pre-specified subgroup of patients, therefore	PD-L1 status (CPS 0 versus CPS ≥ 1) was used as a stratification factor for randomisation, making it reasonable to assess effectiveness in the CPS ≥ 1 subgroup.[1] However, given that patients with a higher CPS score have been shown to have a higher response in other indications for

	Final scope issued by NICE[4]	Decision problem addressed in the CS and rationale if different from NICE scope	EAG comments
		analysis in this subgroup of patients was not performed and not included in this submission. Clinical efficacy results in the metastatic population are available and have been provided in Appendix E.	pembrolizumab,[22] the EAG requested that the company provide any analyses already conducted that were stratified by CPS score. In response the company provided results for the CPS score ≥ 10 subgroup, but no results were provided for CPS <10 or CPS 1 to 9. An economic analysis was not conducted for the CPS score ≥ 10 subgroup. The EAG's clinical advisors considered that it was reasonable not to provide results for the subgroup of patients with locally advanced unresectable disease as this group is a small proportion of the population covered by the licensed indication. In addition, they stated that it was sometimes difficult to distinguish these patients from those with metastatic disease and they would be expected to respond similarly to patients with metastatic disease. The company has focused its presentation of efficacy outcomes in the submission on data from two of the three geographical regions used as stratification factors, which it has combined into a single 'non-Asia region' subgroup. The EAG notes that this trial subgroup relates to geographical region for the recruitment site rather than the ethnicity of the patient. The company's rationale for this subgroup relates to the widespread use of screening programmes in Asian countries which would be expected to lead to earlier

	Final scope issued by NICE[4]	Decision problem addressed in the CS and rationale if different from NICE scope	EAG comments
			diagnosis and would make the results from Asian countries less applicable to countries such as England where there is no screening programme.
Special considerations including issues related to equity or equality	None identified in the scope	The company does not report any equality considerations in CS, Section B.1.4.[1] The company has applied a QALY weighting of 1.2 based on its assessment of disease severity using absolute and proportional QALY shortfall	The EAG's agrees that there is evidence to support a QALY weighting of 1.2 but its assessment is based on a different approach to the company's (see Section 5).

Abbreviations: EAG - external assessment group; HER2 - Human epidermal growth factor receptor 2; HorR - hormone receptor; NHS - National Health Service; NICE - National Institute for Health and Care Excellence; PSS - personal social services; TPC - treatment of physician's choice;

3 CLINICAL EFFECTIVENESS

This chapter presents a summary and critique of the clinical effectiveness evidence contained within the CS for pembrolizumab with trastuzumab and chemotherapy for untreated HER2-positive advanced gastric or GOJ cancer. Section 3.1 describes the company's systematic review of clinical and safety evidence. Section 3.2 provides a summary of the clinical effectiveness and safety results.

3.1 Methods of review of clinical evidence

The systematic review methods for the clinical evidence are detailed in Section B.2.2 of the CS and CS Appendix D.[1] The company undertook a systematic literature review (SLR) to identify randomised controlled trials (RCTs) relating to the efficacy and safety of pembrolizumab in combination with trastuzumab and chemotherapy in patients with locally advanced unresectable or metastatic HER2-positive gastric or GOJ adenocarcinoma.

3.1.1 Searches

The company performed one clinical effectiveness search to identify all clinical effectiveness and safety studies of pembrolizumab or comparator treatments of adult patients with HER2-positive advanced gastric and GOJ adenocarcinoma in previously untreated settings.

The company searched several electronic bibliographic databases in January 2023 (Appendix D.1 Identification and selection of relevant studies): MEDLINE [via Ovid], EMBASE [via Ovid], Cochrane Central Register of Controlled Trials [via Ovid].

According to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Flow Diagram (Appendix D.1 Figure 1, page 66), the company reported records identified from other sources (N=3,625): from Conference Proceedings (n=388), ClinicalTrials.gov Registry (n=2,773), citation searches from published SLRs and conference abstracts (n=462) and expert recommendations (n=2). The company did not report the search strategies for the searches from the conference proceedings sources. The EAG would expect the following sources to be searched in recent years: American Society of Clinical Oncology (ASCO), ESMO. The company only searched the clinical trials registry. A cross-sectional study by Banno, Tsujimoto & Kataoka (2020) compared the coverage of the two trial registry records, ClinicalTrials.gov, World Health Organisation (WHO) International Clinical Trials Registry Platform and CENTRAL and concluded that all three sources should also be searched to identify unpublished trials.[23]

The company searched for the interventions that were listed in Appendix D1.1.1 Table 1. Interventions that are not listed were excluded. However, the EAG sought clarification (question A1) for the following interventions that were included in the search strategy (for example, Appendix D1 Tables 2 and Table

3, pages 5 and 26) but not listed in the inclusion criteria: doxetaxel; paclitaxel; s1-tegafur-oxonate; ipilimumab; avelumab; bevacizumab; leucovorin; carboplatin; sorafenib; ramucirumab; pralatrexate; irinotecan; cediranib; golvatinib; and epirubicin. The company acknowledged in the response that a broader list of interventions was included in the search strategy because it was designed for the global market, but the interventions were excluded at the stage of UK submission and thus not included in the feasibility assessment (not in scope).[19] Therefore, the records retrieved from database searches (identification stage) are not a representation of the number of records that are specifically for the UK context, as found in Appendix D Figure 1 PRISMA Flow Diagram. While the company acknowledges that the search was developed for the global market, the search strategy is restricted to English language publication only. Consequently, there may be studies that are missed in countries where a particular intervention is more common than in other countries, resulting in language bias.[24]

The EAG having reviewed the search strategies considers them to be comprehensive.

3.1.2 Inclusion criteria

The inclusion criteria for the company's SLR are described in CS Appendix D1.1.1. The inclusion criteria in the company's SLR for population were adult patients (≥18 years old) with previously untreated, locally advanced unresectable or metastatic gastric or GOJ adenocarcinoma who received no prior systemic therapy for treatment of advanced or metastatic disease. This was in line with the NICE final scope.[4] The intervention included in the company's SLR was pembrolizumab + trastuzumab + fluoropyrimidine (5-FU or capecitabine) ± leucovorin + platinum agent (oxaliplatin or cisplatin). This was consistent with the NICE final scope. The inclusion criteria in the company's SLR for comparators (trastuzumab + fluoropyrimidine [5-FU or capecitabine] ± leucovorin + platinum agent [oxaliplatin or cisplatin]; or fluoropyrimidine [5-FU or capecitabine] ± leucovorin + platinum agent [oxaliplatin or cisplatin]). Although epirubicin was included in the search strategy, epirubicin-containing triplet therapy was not listed as an eligible comparator for the SLR (CS Appendix D Table 1). Placebo-controlled or best supportive care controlled studies were eligible as trial comparators. All the outcomes in the NICE final scope[4] were included in the outcomes inclusion criteria in the company's SLR.

Eligibility was restricted to English language publications, which introduces the risk that relevant data not published in the English language may have been missed, however the EAG does not anticipate that key RCTs would have been missed. The included study design was limited to RCTs (Section B.2.1 of CS). This is standard practice to restrict to high quality study designs where they are available. It was not clear from Appendix D of the CS if study selection was conducted by one or more reviewers.[1]

While the inclusion criteria for the company's SLR were generally consistent with the NICE final scope,[4] except for epirubicin-containing triplet therapy not being included as a comparator, the inclusion criteria in the company's decision problem differed in terms of population and comparators, based on

the company's KEYNOTE-811 trial. The population was restricted to patients whose tumours express PD-L1 with a CPS≥ 1. The company's explanation for this was that this was in line with the anticipated marketing authorisation.[17]

PD-L1 status CPS<1 vs CPS≥1 was a randomisation stratification factor, and as such, it should be expected that participants be balanced between treatment groups, and it was also one of the pre-planned subgroup analyses. Given this and the anticipated marketing authorisation, the EAG considered it reasonable to present results for the subgroup of CPS≥1 patients. Comparators were restricted to trastuzumab with cisplatin plus capecitabine or fluorouracil. The draft SmPC[17] recommends for this population, pembrolizumab "*in combination with trastuzumab, fluoropyrimidine and platinum-containing chemotherapy*". Clinical advisors to the EAG stated that in clinical practice, trastuzumab would be used in HER2-positive patients, unless contraindicated (such as in patients with cardiac conditions), as such the EAG considered it reasonable to restrict to therapy including trastuzumab. Clinical advisors to the EAG stated that in clinical practice doublet (fluorouracil or capecitabine in combination with cisplatin or oxaliplatin), rather than triplet, chemotherapy would be given, as triplet therapy increased toxicity without much improvement in effectiveness, and many patients would be too frail for triplet chemotherapy.[25, 26] There may be preference for capecitabine over fluorouracil, as the mode of administration is oral, including via feeding tubes. This avoids the need for a central line, which is required for 5-FU and is associated with risks such as line infections. Dose interruptions to manage AEs are also simpler when using an oral treatment. The company's decision problem is discussed in EAG report Section 2.3.

Two trials were included in the CS SLR: KEYNOTE-811; and ToGA (Table 5). KEYNOTE-811 is described in Section 3.2. The company assessed the feasibility of using the KEYNOTE-811 and ToGA studies to provide an indirect comparison of pembrolizumab with trastuzumab and CAPOX/FP against trastuzumab and XP (CS, Section B2.9.1). It concluded that this comparison could only be made by assuming equivalence between doublet chemotherapy regimens, and the results would mirror the KEYNOTE-811 trial results. ToGA was not used further in the CS, and is therefore not described in detail in the EAG report, but some details on study design, risk of bias, and results are provided for reference in EAG report Appendix 1.

Table 5 KEYNOTE-811 and ToGA overview of study characteristics

Trial names and references	Trial design	Population	Intervention	Comparator	Primary outcomes
KEYNOTE-811 NCT03615326 EudraCT 2018-000224-34 MK-3475-811 Janjigian *et al.* 2021 [27] 2021[28] CSR[29] KEYNOTE-811 clinical trials registry [30]	Phase III Randomised Placebo-Controlled Trial, Double-Blind	Adults with HER2 positive participants with previously untreated, locally advanced unresectable or metastatic advanced gastric or GOJ adenocarcinoma	Trastuzumab and pembrolizumab plus either cisplatin plus 5-FU (FP) or oxaliplatin plus capecitabine (CAPOX)	Trastuzumab and placebo plus FP or CAPOX	PFS per RECIST 1.1 assessed by BICR - Time to Event OS - Time to Event
ToGA NCT01041404 BO18255 Bang *et al.* 2010[31] ToGA clinical trials registry [32]	Phase III RCT, open-label	Adults with HER2- positive locally advanced, recurrent, and/or metastatic cancer gastric or GOJ adenocarcinoma	Trastuzumab plus capecitabine plus cisplatin (XP) or fluorouracil plus cisplatin (FP)	Capecitabine plus cisplatin or fluorouracil plus cisplatin	OS - Percentage of participants With an Event OS - Time to Event

Abbreviations: FP= cisplatin plus 5-FU; CAPOX= oxaliplatin plus capecitabine; OS=overall survival; PFS=progression-free survival; HER2= human epidermal growth factor receptor 2; GOJ=gastro-oesophageal junction; BICR=blinded independent central review.

Neither the EAG nor clinical advisors to the EAG are aware of any additional studies of pembrolizumab within the scope of this appraisal.

3.1.3 Data extraction

No detail was reported in the CS Appendix D about the process of data extraction, and thus it is not clear by how many reviewers this was done, if it was checked, how any disagreements were resolved, or which fields were extracted.

Data extracted in the CS for the KEYNOTE-811 trial were checked by the EAG against the trial registry. The main publication for the KEYNOTE-811 trial[27] provided data from an earlier interim analysis than that in the CS, so was not relevant for checking. Following clarification questions, the clinical study report (CSR) [29] was provided and so data were checked by the EAG against the CSR.

3.1.4 Risk of bias assessment

Risk of bias was assessed based on the Cochrane Risk of Bias tool 2.0,[33] which is widely regarded as a robust tool for the assessment of bias in RCTs. It was not clear from Appendix D of the CS if risk of bias assessment was conducted by one or more reviewers.[1]

Risk of bias assessment of the included study, KEYNOTE-811, as undertaken by the company and the EAG, is presented in Section 3.2.3.

3.2 Trial of the technology of interest

3.2.1 Included pembrolizumab trial

The CS (CS Section B.2.2) included one study that examined the effectiveness of pembrolizumab in combination with trastuzumab and chemotherapy, KEYNOTE-811.

KEYNOTE-811 is a phase III double-blind RCT, ongoing at the time of writing. It is a multicentre RCT, with the global cohort recruiting from 192 centres in 20 countries[30]: Australia, Brazil, Chile, China, France, Germany, Guatemala, Ireland, Israel, Italy, Japan, New Zealand, Poland, Russia, South Korea, Spain, Turkey, UK, Ukraine, USA (CS Section B.2.3). There were 29 subjects from 10 UK centres. It consisted of two cohorts, global and Japan-specific SOX (S-1 + oxaliplatin) treated cohort, of which only the global cohort is considered in the CS.[1] SOX was not a comparator included in the NICE final scope,[4] the trial only planned to recruit 40 patients for this cohort,[28] [redacted][29] so the EAG considered it was appropriate to exclude the Japan-specific SOX cohort from the CS.

KEYNOTE-811 study characteristics are shown in

Table 6. Patients in the global cohort were randomised to pembrolizumab in combination with trastuzumab and chemotherapy, or placebo in combination with trastuzumab and chemotherapy. Randomisation was stratified by geographic region (1 Europe [note this refers to Western Europe, CS Clarification response A6] /Israel/North America/Australia, 2 Asia, 3 Rest of the World including South America [note this includes Eastern Europe, CS Clarification response A6]); and PD-L1 status CPS<1 versus CPS≥1; and chemotherapy treatment (FP or CAPOX), which was chosen by the investigating physician prior to randomisation.[19]

Table 6 KEYNOTE-811 (NCT03615326) study characteristics

Population	Intervention	Comparator	Primary outcomes
Adults with HER2 positive participants with previously untreated, locally advanced unresectable or metastatic advanced gastric or GOJ adenocarcinoma	Pembrolizumab 200mg i.v. and trastuzumab 8 mg/kg loading dose, then 6 mg/kg plus either cisplatin 80 mg/m^2 plus 5-FU (FP) or oxaliplatin plus capecitabine (CAPOX)	Placebo (normal saline, i.v.) and trastuzumab plus FP or CAPOX (doses as for intervention)	PFS per RECIST 1.1 assessed by BICR OS

Abbreviations: FP= cisplatin plus 5-FU; CAPOX= oxaliplatin plus capecitabine; OS=overall survival; PFS=progression-free survival; HER2= human epidermal growth factor receptor 2; GOJ=gastro-oesophageal junction; BICR=blinded independent central review; i.v.=intravenous.

PD-L1 status was assessed by PD-L1 IHC 22C3 pharmDx assay (Agilent Technologies, CA, USA) at a central laboratory facility.[28] CS clarification response A4 references Ahn and Kim 2021 [20] which reported that the PD-L1 22C3 and PD-L1 1HC 22c3 pharmDx assays produced comparable results in gastric cancer at CPS cut-offs of 1, 10, and 50. CS clarification response A4 states that both these assays are "*routinely used within NHS clinical practice*".[19] According to clinical advice to the EAG, many units may have access to one or the other of these, however there are multiple assays available depending on the pathology laboratory used, and other assays may give different results. There may also be inherent tumour heterogeneity, meaning even in the same tumour there are regions of positive and negative PD-L1.

KEYNOTE-811 was ongoing at the time of writing. Data in the CS were from interim analysis 2 (IA2) which had a data cut-off date of May 2022.[1] CS Clarification response A13 explained that the database lock for interim analysis (IA3) had occurred but analyses were ongoing, and so were unavailable at the time of writing.[19]

IA2 had been scheduled to be performed after approximately 542 PFS events, and approximately nine months after the last participant had been randomised, with allowance made for conducting the analysis with up to 3 months of additional follow-up, if events accrued slower than expected (CS Section B.2.4). In practice, IA2 occurred after 484 PFS events (of which 414 in CPS>=1 participants) (CS Section B.2.6).

Power is reported for the global cohort, that is, not the subgroup of CPS>=1 participants, or restricted to subgroups by region of the world. CS Section B2.4 gives power for ORR at IA1, for which the planned sample size had been reached[27] and so had approximately *"90% power for detecting a 25%*

difference in ORR (73% vs 48%) at an initially assigned 0.002 (1-sided) significance level" (CS Section B2.4). CS section B.2.4 gives power for PFS at IA3, and OS at the final analysis (CS Section B2.4) but does not report power for PFS or OS at IA2. [redacted]

There were pre-specified subgroup analyses for: PD-L1 Positive versus Negative; Region 1 Europe/Israel/North America/Australia versus 2 Asia versus 3 Rest of World (including South America); age <65 versus ≥65 years; sex female versus male; race Asian versus non-Asian; Microsatellite instability (MSI) status; primary location stomach versus GOJ; histological subtype diffuse versus intestinal versus indeterminate; tumour burden ≥median versus <median; number of metastatic sites ≤2 versus ≥3; prior gastrectomy yes versus no (CS Appendix E) (KEYNOTE-811 protocol amendment 2022).[21]

3.2.1.1 Patients

Eligibility criteria for the KEYNOTE-811 study were presented in CS Section B.2.3 (pages 30-33). The population met the specification of the NICE final scope,[4] in being adults (aged 18 or older) with untreated locally advanced unresectable or metastatic HER2-positive gastric or GOJ adenocarcinoma.[28] Diagnosis was histologically or cytologically confirmed,[30] measurable disease as defined by RECIST v1.1 (Response Evaluation Criteria in Solid Tumors version 1.1).[34] HER2-positive status was defined as either immunohistochemistry (IHC) 3+ or IHC 2+ in combination with in-situ hybridization positive (ISH+), or fluorescent in-situ hybridization (FISH), as assessed by central review on primary or metastatic tumour.[30]

The trial population was narrower than the NICE final scope,[4] in being restricted to Eastern Cooperative Oncology Group (ECOG) Performance Scale 0 or 1, with a life expectancy >6 months, and adequate organ function, and excluding a range of co-morbidities (CS Section B.2.3 pages 30-33). Clinical advice to the EAG suggested that patients with ECOG >1 are often excluded from RCTs. In practice, most participant have ECOG 1 or 2. ECOG 2 patients may be considered too frail for immunotherapy, however with nutrition and chemotherapy they may recover to ECOG 1.

3.2.1.2 Intervention

The intervention group were to receive pembrolizumab in combination with trastuzumab and chemotherapy.

Doses were:[28]

pembrolizumab 200mg i.v. Q3W (day 1 of each cycle);

trastuzumab 8 mg/kg loading dose, then 6 mg/kg maintenance dose i.v. Q3W (day 1 of each cycle);
FP – cisplatin 80 mg/m^2 i.v. Q3W (day 1 of each cycle), plus 5-fluorouracil 800 mg/m^2/day i.v. Q3W (day 1-5 of each cycle);
CAPOX – oxaliplatin 130 mg/m^2 i.v. Q3W (day 1 of each cycle), plus capecitabine 1000 mg/m^2 oral BID (day 1-14 of each cycle).

Treatment was continued for up to 35 cycles, or until disease progression, unacceptable toxicity, or noncompliance, or if the investigator or patient decided to withdraw a participant from the study.[28] Patients with complete response could discontinue treatment after eight or more doses of study treatment.[28] Investigating physicians could choose to continue treating those with disease progression if they were clinically stable.[28] [redacted] Patients with stable disease or better may be eligible for a second course of pembrolizumab (17 doses) if their disease progresses while they are off study treatment.[28] According to local guidelines, some regions discontinued cisplatin at six cycles.[28] CS clarification response A7 states there were up to six cycles for cisplatin, and 6-8 cycles of oxaliplatin.[19]

3.2.1.2.1. Comparator

The comparator group were to receive placebo in combination with trastuzumab and chemotherapy. Doses of trastuzumab and chemotherapy were as for the intervention group.[28]

Clinical advisors to the EAG stated that in clinical practice, trastuzumab would be used in HER2-positive patients, unless contraindicated (such as in patients with cardiac conditions). In cases where trastuzumab is contraindicated, it is unusual for patients to be fit for chemotherapy. The advisors also stated that in clinical practice, doublet, rather than triplet, chemotherapy would be given for 4-8 cycles, in general.[25]

Concomitant treatments (across both treatment groups) were allowed at the physician's discretion, with the exception of the following excluded treatments: antineoplastic systemic chemotherapy; biologic therapy; immunotherapy or chemotherapy not specified in the protocol; other investigational agents; radiotherapy; live vaccines within 30 days prior to (and throughout) trial treatment; systemic glucocorticoids (unless to treat AE or for cisplatin or 5-FU supportive care); inhibitors of the enzyme dihydropyrimidine dehydrogenase for participant given 5-FU therapy (CS Section B2.3).

3.2.1.2.2. Outcomes

The primary outcomes of KEYNOTE-811 were PFS and OS.[30] [28]

PFS was defined as the time from randomisation to the first documented disease progression per RECIST 1.1 as assessed by blinded independent central review (BICR) or death due to any cause, whichever occurred first.[28,30] As in RECIST 1.1, progressive disease was defined as at least a 20% increase in the sum of diameters of target lesions, taking as reference the smallest sum on study.[30] In addition to the relative increase of 20%, the sum must also demonstrate an absolute increase of at least 5 mm.[30] The appearance of one or more new lesions was also considered progression.[30] Assessment of tumour status was performed every six weeks,[28] until progressive disease, death, start of new anticancer treatment, or withdrawal of consent. The use of established response evaluation criteria (i.e. RECIST) is recommended practice by the European Medicines Agency (EMA) for chemotherapy trials, although not established for immunotherapy.[35,36] iRECIST is available for immunotherapy.[37] According to clinical advice to the EAG, immunotherapy may lead to "pseudo-progression" whereby there is an increase in size of the target lesion within a few weeks/months of starting immunotherapy. Scanning that is more frequent than in UK practice (that is, three months following treatment initiation) may result in a patient erroneously being regarded as having progressed disease.

OS was defined as the time from randomisation to death due to any cause.[28,30] For patients no longer being monitored every six weeks, follow-up for survival was conducted every twelve weeks. EMA research recommendations [38] advise that OS should be considered a secondary outcome in Phase III trials where PFS is the primary outcome, and should demonstrate or show a trend towards superiority.

Secondary outcomes were: Objective Response Rate (ORR); Duration of Response (DOR); AEs [28,30] and treatment discontinuation due to AEs.[30] ORR was defined as the percentage of participants who have a complete response (CR) that is disappearance of all evidence of disease, or partial response (PR) that is regression of measurable disease and no new sites), per RECIST 1.1 as assessed by BICR.[30] DOR was defined as the time from first response (CR or PR) to subsequent disease progression or death from any cause, whichever occurs first, per RECIST 1.1 as assessed by BICR.[30] Adverse events were defined in the protocol,[28] in accordance with the National Cancer Institute Common Terminology Criteria for Adverse Events version 4.0.[39] Serious adverse events (SAE) were defined as: fatal; life-threatening; requiring hospitalisation or prolonged existing hospitalisation; resulting in persistent or significant disability/incapacity; congenital anomaly/birth defect; other important medical event according to medical or scientific judgement (KEYNOTE-811 protocol amendment 2022).[21]

Exploratory outcomes were HRQoL, utilities, molecular biomarkers, and PFS and ORR per immune-related RECIST as assessed by investigating physicians.[28]

Effectiveness outcomes were analysed in the intent to treat (ITT) population, that is all randomly assigned patients in the group they were assigned to.[28] The safety population included all randomly assigned patients who received ≥1 dose of study treatment, analysed by treatment received.[28]

Three interim analyses and final analyses were planned according to project milestones (CS Section B.2.4 Table 12).[1] At time of writing, the results of IA2 were provided in the CS but had not been published, and the results of IA1 had been published (Janjigan *et al.* 2021).[27]

3.2.1.2.3. Ongoing studies

KEYNOTE-811 was ongoing at time of writing. CS Section B.2.11 states that there are no other ongoing trials of pembrolizumab in patients with previously untreated locally advanced unresectable or metastatic HER2-positive gastric or GOJ adenocarcinoma. Neither the EAG nor clinical advisors to the EAG are aware of any additional studies of pembrolizumab within the scope of this appraisal.

3.2.2 Details of relevant RCTs not included in the submission

Neither the EAG nor clinical advisors to the EAG were aware of any additional RCTs within the scope of this appraisal. According to clinical advice there are ongoing trials of novel HER2 inhibitors in this condition, however these would not meet the final NICE scope criteria.

3.2.3 Risk of bias assessment KEYNOTE-811

The company provided a risk of bias was assessed based on the Cochrane Risk of Bias tool 2.0. A summary of the risk of bias in the KEYNOTE-811 study undertaken by the company alongside the EAG's independent quality assessment (from the publications of Janjigan *et al,.*[27] and Chung *et al.*,[28] clinical trials registry[30] and the CSR[29]) is presented in Table 7. The company's critical appraisal and the EAG's critical appraisal of the KEYNOTE-811study were similar.

Randomisation allocation concealment was adequate.[28] Baseline characteristics at IA2 appear balanced between treatment groups in [redacted] CS Section B2.3.2). Participants and clinicians were blind to treatment, and PFS was assessed by BICR (clinicaltrials.gov). Outcome measurement was the same for both treatment groups.[28] ITT analyses were planned for effectiveness measures and all participants were included in the ITT analyses for the global cohort (CS Appendix M) (CSR)[29].

Subgroup analyses were pre-specified for the randomisation stratification factors of: geographic region (1 Europe/Israel/North America/Australia, 2 Asia, 3 Rest of the World including South America); PD-L1 status CPS<1 vs CPS≥1. However, CS presents data for *post hoc* analyses of the subgroup of CPS≥1, and within that CPS≥1 subgroup, the non-Asian participants; that is the combination of the two

subgroups Europe/Israel/North America/Australia, and Rest of the World including South America. Both CPS and region were stratification factors, and so treatment groups should be balanced. However, the exclusion of the Asia region subgroup was a *post hoc* analysis. Presenting results for the subgroup of CPS≥1 was in line with the anticipated marketing authorisation,[17] and so were relevant for this submission. Clinical practice in the Asia region is dissimilar to clinical practice in England, and so data from the Asia region may not be generalisable to the population likely to be treated England. However, it is unclear if the Western Europe population would be more generalisable to England than the Rest of the World region (which had more favourable results to pembrolizumab then the Europe population). The Rest of the World region included South American and Eastern Europe, and according to clinical advice there may be less access to care in these settings than in England.

The results of IA2 of KEYNOTE-811 had not yet been published at the time of writing; as such, it cannot be assessed if the authors measured more outcomes than they reported. However, data for outcomes of relevance to this review were provided by the company in the CS and clarification response.[19]

Overall, the KEYNOTE-811 was well-designed to give a low risk of bias. Data from IA2 had lower numbers of PFS events than was anticipated, and there is some uncertainty about statistical power of the PFS analysis, and the trial was not powered for subgroups. There is some concern about the *post hoc* analysis of CPS≥1 excluding the Asia region, as it is uncertain whether the Western Europe subgroup alone would be more generalisable to England, than the grouping together with Rest of World subgroup.

Table 7 Risk of bias assessment KEYNOTE-811

Type of bias	KEYNOTE 811			
	Review authors' judgement	**Support for judgement**	**EAG judgement**	**Support for judgement**
Bias arising from the randomization process	Low risk	Double-blind study; participants were randomly assigned 1:1 to pembrolizumab or placebo via an integrated interactive voice- and web-response system and assignment was masked to both participants and investigators.	Low risk	Allocation sequence – stratified randomisation via interactive voice/web response system implies computer-generated random numbers [28] Allocation concealment - Randomisation performed centrally using an interactive voice/web response system [28] Baseline characteristics – balanced across treatment groups for CPS≥1 participants, and for whole global cohort at IA1 [redacted]
Bias due to deviations from intended interventions	Low risk	Double blind study; no deviations from the intended interventions arose because of trial context and appropriate analysis methods were employed to estimate treatment effects.	ITT population - low risk. CPS≥1 and non-Asian participants – some concerns	Participant awareness of assigned intervention – blinded, placebo controlled (unclear if side effects alerted some participants to intervention) Clinician/carer awareness of assigned intervention – blinded, placebo controlled (unclear if side effects alerted some clinicians to participants' interventions) Trial context – no strong reason to believe that the trial context led to failure to implement the protocol interventions Appropriate analyses - ITT analyses planned for effectiveness measures.[28] However, CS presents *post hoc* data for CPS≥1 and non-Asia region participants, which excludes eligible trial participants.

Type of bias	KEYNOTE 811			
	Review authors' judgement	**Support for judgement**	**EAG judgement**	**Support for judgement**
				Impact of excluding eligible participants from analyses – potential for impact of CPS≥1 mitigated by randomisation stratification. Regional subgroups were stratified, however subgroups had different treatment effects, and the excluded Asia region had a less favourable effect for pembrolizumab.
Bias due to missing outcome data	Low risk	Data for outcomes available represented all or nearly all randomized participants.	Low risk	Available outcome data – OS and PFS ITT analyses; also subgroup CPS≥1 ITT
Bias in measurement of the outcome	Low risk	Appropriate method used to measure outcomes.	Low risk	Method of measuring outcomes – appropriate Outcome measurement for treatment groups – same measurements and same time points (every 6 weeks for PFS; every 12 weeks following progression or treatment change for OS) [28]
Bias in selection of the reported result	Low risk	Analysis was in accordance with a pre-specified analysis plan that was finalized before the outcome data were available for analysis.	ITT population - low risk. CPS≥1 and non-Asian participants – some concerns	Analyses pre-specified – subgroup analyses prespecified for stratification factors, however *post hoc* to combine Europe and Rest of World (excluding Asia) Multiple outcome measurements – effectiveness outcomes were pre-defined with one clear measurement for the outcome domain
Overall bias	Low risk	Low risk of bias across all domains.	ITT population - low risk. CPS≥1 and non-Asian participants – some concerns	

3.2.4 KEYNOTE-811 results

Data are from KEYNOTE-811 IA2. Data from IA3 were not available at time of writing (CS clarification response A13).[19] Data are from the global cohort of KEYNOTE-811 (the Japan specific SOX cohort was excluded throughout the CS). Median follow-up at IA2 was 16.1 months for the pembrolizumab group, and 14.8 months for the comparator group (CS Clarification response A23)(CSR).[19,29]

The CS concentrated on data from the subgroup of patients with PD-L1 with a CPS≥ 1. This was in line with the anticipated marketing authorisation. This was a randomisation stratification factor, and as such, it should be expected that participants be balanced between treatment groups. Within the CPS≥1 subgroup, the CS concentrated on the non-Asian participants, that is the combination of the two subgroups: Europe/Israel/North America/Australia; and Rest of the World including South America. Region was a randomisation stratification factor. The combinations of the two subgroups were *post hoc* analyses.

The global cohort of KEYNOTE-811 randomised 698 patients. Of these, 594 had CPS ≥1, and 104 had CPS<1. There were 224 patients recruited in Europe/Israel/North America/Australia; 237 recruited in Asia; and 237 in Rest of the World including South America (CS Appendix M). Within the CPS ≥1 subgroup, there were 193 patients recruited in Europe/Israel/North America/Australia; 192 recruited in Asia; and 209 in Rest of the World including South America (CS Appendix E).

CS Appendix D.1.2 reports participant flow for CPS ≥1 participants. All but one patient (randomised to comparator) started study treatment. At the time of database cut-off for IA2, 71.8% of the pembrolizumab group, and 83.1% of the comparator group, had discontinued study treatment. In the majority of cases, discontinuation of study treatment was due to progressive disease or death, 50.7% of the pembrolizumab group, and 63.1% of the comparator group. Adverse events led to discontinuation for 10.4% of the pembrolizumab group, and 8.8% of the comparator group (CS Appendix Table 19). Discontinuation rates were similar for the Western Europe/Israel/North America/Australia and Rest of World regions. For Western Europe/Israel/North America/Australia at time of database cut-off for IA2, there was discontinuation of treatment for 79.4% of the pembrolizumab group (10.3% due to AE), and 88.5% of the comparator group (14.6% due to AE) (CS Clarification response A14).[19] For the Rest of World region, there was discontinuation of treatment for 68.6% of the pembrolizumab group (14.3% due to AE), and 85.4% of the comparator group (9.7% due to AE) (CS Clarification response A14).[19]

Baseline characteristics of the CPS ≥1 participants are reported in CS Table 6. Characteristics appear balanced between the treatment groups.

According to clinical advice, trial participants were younger than would be seen in clinical practice in England (by about ten years). It is common for trial participants to be younger or fitter than would be seen in practice.[40] The difference in the average age between patients in the RCT and in England may result in a different treatment effect. There is some uncertainty as KEYNOTE-811 was not powered for subgroups, however patients under 65 years appeared to have more favourable treatment effects for pembrolizumab for PFS and OS, than older patients (CS Appendix E). Most English patients are 70 years or older according to clinical advice.[41]

In terms of primary location at diagnosis, and the mix of locally advanced versus metastatic disease, clinical advisors thought this was representative of the eligible population in England.[41] The eligible population in England would be likely to comprise a higher proportion of black patients, however clinical advice suggested this was unlikely to alter the treatment effect. The company's preferred population (CPS≥1 and non-Asia region) contains only 2 Asian participants. The eligible population in England would be likely to comprise a higher proportion of Asian patients, however clinical advice suggested this was unlikely to alter the treatment effect. Note that ethnicity and region are separate subgroups (EAG report Section 2.3.1).

At IA2, for CPS ≥1 participants, the median duration of follow up for the pembrolizumab group was 17.0 months (range: 0.6 to 41.6 months), and in the comparator group 13.9 months (range: 0.3 to 41.2 months) (CS Appendix D1.2)

Drug exposure for CPS ≥1 participants (in treated participants: pembrolizumab n=298; comparator n=295) was reported in CS Tables 23 and 24. Exposure to study drug was longer in the pembrolizumab group (median 10.2 months, range 0.3 to 36.6) compared with the comparator group (median 7.1 months, range 0.0 to 36.1) (CS Section B2.10)(Keytruda (MK-3475) HTA report 2022).[42] Only three patients received a second course of pembrolizumab (CS Clarification response A8).[19]

Of the CPS ≥1 participants, 504 were prescribed CAPOX (n=251 pembrolizumab group, n=253 comparator group), and 90 were prescribed FP (n=47 pembrolizumab group, n=43 comparator group) (CS Section B2.3). Mean number of chemotherapy cycles for CPS ≥1 non-Asia region participants were reported in CS Clarification response B35 (

Table 8).[19]

Table 8 Mean number of chemotherapy cycles administered per treatment arm in KEYNOTE-811 (non-Asia CPS≥1 cohort) (reproduced from clarification response, Table 45[19]; supersedes CS, Table 32[1])

	Pembrolizumab with trastuzumab plus chemotherapy; mean (SD)	Trastuzumab plus chemotherapy; mean (SD)
Capecitabine (in CAPOX)	13.3 (10.7)	9.6 (8.3)
Oxaliplatin (in CAPOX)	7.3 (4.6)	6.8 (4.6)
Cisplatin (in FP)	5.3 (1.8)	5.6 (1.9)
5-FU (in FP)	9.5 (6.7)	11.2 (9.5)

Abbreviations: SD, standard deviation

3.2.4.1 PFS

IA2 was conducted after the occurrence of 484 PFS events in the global cohort (CS Section B2.6.1).

At IA2, of the 594 participants with CPS≥1, 414 had PFS events (CS Section B2.6.1). Median follow-up at IA2 was 17 months for the pembrolizumab group, and 13.9 months for the comparator group (CS Appendix D1.2, CS Clarification response A23).[19] In the pembrolizumab group, 199/298 (66.8%) patients had a PFS event (n=29 death, n=170 progression), and median PFS was 10.8 months (95% confidence interval [CI] 8.5, 12.5). In the comparator group, 215/296 (72.6%) patients had a PFS event (n=30 death, n=185 progression), and median PFS was 7.2 months (95%CI 6.8, 8.4). For CPS≥1 patients, the hazard ratio (HR) for PFS significantly favoured the pembrolizumab group, HR 0.70 (95% CI 0.58, 0.85, p = 0.0001) (CS Section B2.6.1 Table 15).

Within CPS≥1, subgroup data were reported (CS Appendix E). [redacted]

Within the stratified CPS≥1, subgroup data for the RCT's other stratification factors (region and chemotherapy type) were reported (CS Appendix E). Hazard ratios for pembrolizumab with reference comparator group were similar for CAPOX (HR 0.69 [95%CI 0.56; 0.85] and FP (HR 0.69 [95%CI 0.43; 1.12]), with a wider confidence interval for FP which had a smaller sample size (n=90) (CS Appendix E Table 22 and Figure 7). Data by region varied considerably. Western Europe/Israel/North America/Australia had a HR of 0.69 [95%CI 0.50; 0.97]; Asia had a HR of 0.85 [95%CI 0.59; 1.22]; Rest of the World had a HR of 0.56 [95%CI 0.41; 0.78] (CS Clarification response A10).[19] This implied a more favourable effect for pembrolizumab on PFS for the Rest of the world region than for the other regions; and also a less favourable effect for pembrolizumab on PFS for the Asia region. However, the interaction across the three regions did not reach statistical significance (CS Clarification response A10).[19]

Other pre-planned subgroup data, for non-stratified subgroups, were reported in CS Appendix E Table 22 and Figure 7). There appeared to be a more favourable effect for pembrolizumab on PFS for age<65years than for older patients, however the interaction did not reach statistical significance (for this, or any of the other pre-planned subgroups within the CPS≥1 subgroup) (CS Clarification response A10).[19]

A *post hoc* analysis of the combined subgroup of non-Asia regions, within CPS≥1 patients, (the company's preference of population). Median PFS for the pembrolizumab group was 9.9 months (95%CI 8.3, 11.3), and for the comparator group 6.3 months (95%CI 5.6, 7.3) (CS Clarification response A15). The HR for PFS favoured the pembrolizumab group, HR 0.62 ([95% CI: 0.49; 0.78] <0.0001) (CS Section B2.6.1 Table 14, and CS Clarification response A15).[19]

In the global cohort, using a different cut-off for CPS, a *post hoc* analysis of the subgroup CPS≥10 did not find a significant treatment group difference for PFS, HR 0.72 (95% CI: 0.52, 1.01) (CS Clarification response A5).[19] When restricted to non-Asia region participants, CPS≥10, PFS HR was [redacted] (CS Clarification response A5).[43]

3.2.4.2 OS

At IA2, of the 594 participants with CPS≥1, 350 had OS events (CS Section B2.6.1). In the pembrolizumab group, 167/298 patients had an OS event, and median time to death was 20.5 months (95%CI 18.2, 24.3). In the comparator group, 183/296 patients had an OS event, and median time to death was 15.6 months (95%CI 13.5, 18.6). For CPS≥1 patients, the HR for OS favoured the pembrolizumab group, HR 0.79 (95% CI 0.64, 0.98, p = 0.0143) (CS Section B2.6.1 Table 17).

Within the CPS≥1 subgroup, the HR for pembrolizumab with reference comparator group for CAPOX was HR 0.82 (95%CI 0.65, 1.03), and for FP was HR 0.71 (95%CI 0.43, 1.18), neither being statistically significant, FP appearing more favourable for pembrolizumab but with wide confidence interval (CS Appendix E Table 23 and Figure 8).

Within the CPS≥1 subgroup, data by region varied significantly (interaction effect p=0.0317) (CS Clarification response A17).[19] Direction of effect favoured pembrolizumab for the regions: Western Europe/Israel/North America/Australia had a HR of 0.81 (95%CI 0.57, 1.15); and Rest of the World had a HR of 0.57 (95%CI 0.40, 0.80) (CS Clarification response A17). Median OS for CPS≥1 participants, Western Europe/Israel/North America/Australia, the pembrolizumab group (n=97) was 18.8 months (95%CI 14.6, 24.2), and for the comparator group (n=96) median OS was 12.2 months (95%CI 10.4, 15.7). Median OS for CPS≥1 participants, Rest of the World, the pembrolizumab group

(n=105) was 20.3 months (95%CI 14.8, 27.9), and for the comparator group (n=104) median OS was 13.4 months (95%CI 10.4, 15.5).

For the Asia region, direction of effect favoured the comparator group, HR for pembrolizumab (reference comparator group) HR 1.15 (95%CI 0.76, 1.76) (CS Clarification response A17).[19]

PFS is generally, but not always, a suitable surrogate for OS, and EMA recommends trials with PFS as a primary endpoint include OS as an outcome.[35, 38] PFS may not be a suitable surrogate for OS due to treatment subsequent to study-treatment, or may be more pronounced where detection of progressed disease is improved, thus leading to a longer duration of post-progression survival.[44] For frailer patients with upper gastro-intestinal cancers, there are few subsequent treatment options available, and so PFS may be more reflective of OS. However, according to clinical advice, there is also the issue of pseudo-progression in immunotherapy, whereby there is an increase in size of the target lesion within a few weeks/months of starting immunotherapy, and so may overestimate progression rates at early measurements.

One meta-analysis found that treatment effects for PFS and OS were only moderately correlated in gastric cancer, although this included Asia region and non-Asia region trials, was not restricted by HER2 status (thus differing from the population in the company's model), and included trials with second-line treatments (which may act to dilute the effect of first-line treatment.[45]

Within the CPS≥1 subgroup, for patients aged <65 years there was a more favourable effect for pembrolizumab for OS (HR 0.63 [95%CI 0.48; 0.84]), than for patients aged 65 years or older (HR 1.06 [0.77; 1.47]), with a significant interaction effect p=0.0174 (CS Appendix E Table 23).

In a *post hoc* analysis of the combined subgroup of non-Asia regions, within CPS≥1 patients, (the company's preference of population) median time to death for the pembrolizumab group was 18.8 months (95%CI 15.5, 24.3), and for the comparator group 12.6 months (95%CI 11.1, 14.9) (CS clarification response A15).[19] The HR for OS favoured the pembrolizumab group, HR 0.67 ([95% CI: 0.52; 0.85], p=0.0006) (CS Section B2.6.1 Table 16 and CS Clarification response A15).[19]

Looking at a different cut-off for CPS, a *post hoc* analysis of the subgroup CPS≥10 did not find a significant treatment group difference for OS, HR 0.93 (95% CI: 0.66, 1.32) (CS Clarification response A5).[19] When restricted to non-Asia region participants, CPS≥10, OS HR for pembrolizumab (n=71) with reference comparator (n=64) was [redacted] (CS Clarification response A5).[43]

3.2.4.3 Response rate

At IA2, of the participants with CPS≥1, the pembrolizumab group had an ORR of 218/298 (73.2% [95%CI 67.7, 78.1]) (CS Section B2.6.1 Table 18). This comprised of 42 patients with CR, and 176 with PR (CS Section B2.6.1 Table 19). The comparator group had an ORR of 173/296 (58.4% (95%CI 52.6, 64.1)). This comprised of 29 patients with CR, and 144 with PR (CS Section B2.6.1 Table 19). The difference in ORR favoured the pembrolizumab group, estimate 14.7% (95%CI 7.1%, 22.2%) p=0.00008 (CS Section B2.6.1 Table 18). ORR for the ITT population were reported in CS Clarification response A18.[19]

The median DOR, of the participants with CPS≥1, was 11.3 months in the pembrolizumab group, and 9.5 months in the comparator group (CS Section B2.6.1 Table 20) and CS clarification response A23).[19]

3.2.4.4 Adverse events

The safety population included all randomly assigned patients who received ≥1 dose of study treatment, analysed by treatment received.[28]

In the global cohort there were four physician-assessed, drug-related, AEs resulting in death in the pembrolizumab group (pneumonitis, hepatitis, sepsis, and cerebral infarction) and three in the comparator group (myocarditis, pulmonary embolism, and cholangitis) (CS Section B2.10).

The safety population, for participants with CPS≥1, included 298 patients in the pembrolizumab group, and 295 in the comparator group (CS Section B2.10). For participants with CPS≥1, there were two physician-assessed. drug-related AEs resulting in death in the pembrolizumab group (pneumonitis, hepatitis) and one in the comparator group (myocarditis) (CS clarification response A20).[19]

At IA2, of the participants with CPS≥1, AEs led to discontinuation for 10.4% of the pembrolizumab group, and 8.8% of the comparator group (CS Appendix D1.2 Table 19). AEs of special interest led to discontinuation for 7.0% of the pembrolizumab group, and 4.1% of the comparator group (CS B2.10 Table 26).

For participants with CPS≥1, 97.0% of pembrolizumab treated patients experienced one or more AEs, and 96.3% of comparator treated patients (CS Section B2.10). Grade ≥3 AEs were experienced by 73.2% of the pembrolizumab treated patients, and 65.1% of the comparator treated patients (CS Appendix F). The most common grade 3-5 AEs were anaemia, diarrhoea, neutropenia, vomiting, nausea, fatigue, thrombocytopenia, neutrophil count decreased, platelet count decreased and peripheral sensory neuropathy (CS Section B3 Table 55 and CS Clarification response).[1,19]

An overview of AEs of special interest in the subgroup of participants with CPS≥1 is provided in CS Section B2.10 Table 26 and CS Appendix F. For participants with CPS≥1, 37.6% of pembrolizumab treated patients experienced one or more AEs of special interest, and 23.1% of comparator treated patients (CS Section B2.10). Grade ≥3 AEs of special interest were experienced by 10.4% of the pembrolizumab treated patients, and 3.4% of the comparator treated patients (CS Section B2.10).

3.2.4.5 Health-related quality of life (HRQoL)

HRQoL was analysed in participants with at least one dose of study treatment and at least one patient-reported outcome (PRO) assessment (CS Section B2.10). Change from baseline was based on a constrained longitudinal data analysis cLDA model with the PRO scores as the response variable with covariates for treatment by study visit interaction and trial stratification factors. At IA2 in the CPS≥1 global cohort, change from baseline to week 24 in EQ-5D-5L was least square mean (LSMean) 1.20 (95%CI -0.81, 3.21) in the pembrolizumab group (n=292), and LSMean 1.36 (95%CI -0.81, 3.53) in the comparator group (n=290), with no significant change for either group, and no significant difference between groups (CS Section B2.6 Table 21; LSMean reported as analysis adjusted for covariates and stratification factors).

In the CPS≥1 non-Asia region participants, change from baseline to week 24 in EQ-5D-5L was LSMean [redacted] in the pembrolizumab group (n=[redacted]), and LSMean [redacted] in the comparator group (n=[redacted]) [redacted] (CS clarification response A21).[19]

3.3 Conclusions of the clinical effectiveness section

The EAG believes that no RCTs of pembrolizumab meeting the inclusion criteria of the NICE final scope have been missed. The company's search for clinical evidence reflected the decision problem in the NICE final scope, although the company's decision problem was limited by population (limited to the subgroup with PD-L1 CPS ≥ 1 in line with the anticipated marketing authorisation) and comparators (limited to trastuzumab and doublet chemotherapy). These restrictions were thought to be acceptable by the EAG.

One RCTs of pembrolizumab in previously untreated HER2- positive locally advanced, recurrent, and/or metastatic cancer gastric or GOJ cancer was included in the CS SLR: KEYNOTE-811. The key clinical evidence for pembrolizumab was based on the KEYNOTE-811 RCT. KEYNOTE-811 is a phase III, multi-centre, double-blind RCT, ongoing at the time of writing. Data were provided for interim analysis 2 (IA2). The study randomised 698 patients to either pembrolizumab in combination

with trastuzumab and chemotherapy, or placebo in combination with trastuzumab and chemotherapy. Chemotherapy was CAPOX or FP, with CAPOX given to the majority of patients in both treatment groups.

The CS concentrated on the subgroup of PD-L1 CPS ≥ 1 patients, and within that a subgroup by region that combined the two regions of Western Europe/Israel/North America/Australia, and Rest of World (including South America); that is excluding the Asia region. Both of these subgroup variables had been randomisation stratification factors. Patients with PD-L1 CPS ≥ 1 was a pre-planned subgroup analysis, the exclusion of the Asia region was a *post hoc* analysis. Although combining Western Europe/Israel/North America/Australia and Rest of the world does not break the randomisation as randomisation was preserved within each region, the EAG questions the validity of such subgroup analysis (Asia vs. non-Asia) due to the *post hoc* nature.

The KEYNOTE-811 RCT was well-designed to give a low risk of bias, however there is some uncertainty about statistical power of the PFS analysis. The trial was not powered for subgroups. There is some concern about the *post hoc* analysis of CPS≥1 excluding the Asia region, as it is uncertain whether the Western Europe subgroup alone would be more generalisable to England, than the grouping together with Rest of World subgroup.

According to clinical advice, patients in KEYNOTE-811 RCT were younger, with a higher proportion of white patients than would be seen in clinical practice in England, but were generally representative in terms of primary location of disease at diagnosis, and the mix of locally advanced versus metastatic disease. Age may influence effectiveness, as patients under 65 years appeared to have more favourable treatment effect toward pembrolizumab for PFS and OS than older patients, however there is uncertainty in this as KEYNOTE-811 was not powered for subgroups.

The primary outcomes were OS and PFS. At IA2, for CPS≥1 participants excluding the Asia region, the HR for OS favoured the pembrolizumab group, HR 0.67 ([95% CI: 0.52; 0.85], p=0.0006). Median OS for the pembrolizumab group was 18.8 months (95%CI 15.5, 24.3), and for the comparator group 12.6 months (95%CI 11.1, 14.9). At IA2, for CPS≥1 participants excluding the Asia region, the HR for PFS favoured the pembrolizumab group, HR 0.62 ([95% CI: 0.49; 0.78] <0.0001). Median PFS for the pembrolizumab group was 9.9 months (95%CI 8.3, 11.3), and for the comparator group 6.3 months (95%CI 5.6, 7.3).

For CPS≥1 participants, Grade ≥3 AEs were experienced by 73.2% of the pembrolizumab treated patients, and 65.1% of the comparator treated patients. For CPS≥1 participants, there was no significant

change for either group in HRQoL as measured by EQ-5D-5L, and no significant treatment group difference.

4 COST EFFECTIVENESS

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

4.1.1 Objective of cost effectiveness review

The objective of the company's review of published cost-effectiveness studies is not entirely clear from the submission. CS, Appendix G, Section G.1 describes the review question as being to understand the economic burden of patients with untreated locally advanced unresectable gastric or GOJ adenocarcinoma, in addition to identifying studies reporting economic evaluations and health care resource use in this population. This objective is much broader than reviewing cost-effectiveness studies that match the decision problem in the NICE scope. However, the reporting of the review in CS, Section B.3.1 is more focused on the decision problem specified in the NICE scope, but this more focused objective is not clearly stated in the CS.

4.1.2 Searches

The company performed systematic literature searches in April 2023 for published cost-effectiveness studies, economic burden, and healthcare resource use (including cost data) of patients with untreated locally advanced unresectable gastric or GOJ adenocarcinoma (CS Appendix G). These searches were also used to identify studies that reported data on utilities associated with gastric cancer (section G.5.7., page 201) although supplementary HRQoL searches are also found in CS Appendix H.

In the cost-effectiveness studies, economic burden, cost, and resource use study search strategies were combined into one search strategy and the following sources were searched: MEDLINE [via Embase.com]; MEDLINE In-Process [PubMed]; Embase [via Embase.com]; NHS Economic Evaluation Database [via CRD databases]; and Tufts Cost-effectiveness Analysis Registry. The company has also undertaken searches of the bibliographies of the included studies and reviews.

The company searched several conference abstract websites in the last five years (2018-2023): ASCO; ASCO-Society for Immunotherapy of Cancer; ASCO-Gastrointestinal; American Society for Radiation Oncology; European Society for Medical Oncology; European Society for Medical Oncology; European Society for Medical Oncology-Gastrointestinal; European Society for Medical Oncology-Immuno-Oncology Congress; Asia-pacific Gastroesophageal Cancer Congress; American Association for Cancer Research; Japanese Society of Medical Oncology; Society for Immunotherapy of Cancer; European Cancer Congress; International Society for Pharmacoeconomics and Outcomes Research (Europe and International); Annual Meeting of Academy of Managed Care Pharmacy; and NEXUS.

Additionally, the company searched several country-specific and international Health Technology Appraisal websites: NICE; Scottish Medicines Consortium; Institute for Quality and Efficiency in Health Care; Haute Autorité de Santé; Canadian Agency for Drugs and Technologies in Health; Pharmaceutical Benefits Advisory Committee; International Network of Agencies for Health Technology Assessment; International Society for the promotion of health technology assessment (htai.org); and European Network for Health Technology Assessment. The EAG considers that the search is comprehensive.

The company conducted supplementary HRQoL and outcome searches for patients with locally advanced unresectable gastric or GOJ adenocarcinoma and their carers (CS Appendix H). The searches were undertaken in April 2023 in the following sources: MEDLINE [via Embase.com]; MEDLINE In-Process [PubMed]; Embase [via Embase.com]; the Central Register of Controlled Trials; and the Cochrane database of systematic reviews [via Wiley]. The company also searched several conference proceedings sources in the last five years, as listed in Appendix G of the CS (pages 149-150). There were no consequential errors in the search, and the EAG considers that the search is comprehensive.

While the searches in Appendix G.2 encompass the searches for cost and healthcare resource use (Appendix I), in addition to data from NICE HTAs, the company carried out additional searches of the excluded studies list to find disease management costs to reflect the current practice (Appendix I, pages 291-292). The strategies for the additional searches were not reported in the submission.

4.1.3 The inclusion and exclusion criteria used in the study selection

The target population for the review is described in Appendix G as "*adult (≥18 years) patients with previously untreated, locally advanced, unresectable gastric or GOJ adenocarcinoma*" (CS, Appendix G, Section G.3), with the table of inclusion/exclusion criteria defining this further as "*stage II-III-IVa*". The review search criteria suggest that the company also intended to identify studies in patients with stage 4b and 4c disease, and the review does appear to have included studies in patients with metastatic disease, despite not using this terminology when describing the target population. Based on this, the EAG's interpretation is that the review intended to cover both metastatic and unresectable locally advanced disease as per the population specified in the NICE scope.

The review is described as not limited to the treatment combinations listed in the NICE scope or whether treatment is being given as first-line treatment despite describing the target population as "*previously untreated*" (Document B, Appendix G). The review was not restricted to any country or geographical region. The review was not restricted to cost-effectiveness studies and also includes cost minimisation studies, budget impact studies, cost of illness studies and resource use studies. Therefore, the

inclusion/exclusion criteria for the review described in Appendix G appears to be much broader than the decision problem specified in the NICE scope.

4.1.4 Findings of the cost effectiveness review

Section B.3.1 of the CS states that no studies were identified which evaluated pembrolizumab in combination with trastuzumab in the specified population. The CS then goes on to describe details of the model that informed the NICE appraisal of trastuzumab (TA208). This is the only study described in Section B.3.1 despite 62 published studies being included in the broader review described in Appendix G.

4.1.5 Conclusions of the cost effectiveness review

The conclusion of the review in CS, Section B.3.1, appears to be that no cost-effectiveness studies evaluating pembrolizumab in combination with trastuzumab in the specified population were identified.

4.1.6 EAG critique of the company's review of cost effectiveness

The company's reporting of the results of the review implies that the objective of the review was to identify cost-effectiveness studies of either pembrolizumab in combination with trastuzumab or of the comparator strategies described in the scope, one of which is the treatment combination assessed in TA208.[10] It is unclear to the EAG why the company has conducted a review with such a broad remit in Appendix G, and has then restricted its reporting of the results of the review to a subset of these studies without providing a specific objective for the review which would justify this restriction. However, the EAG considered that it is unlikely that the company's review has failed to identify any cost-effectiveness analyses that are directly relevant to the decision problem specified in the NICE scope.

4.2 **Summary of the company's submitted economic evaluation**

4.2.1 Population

The population for the economic evaluation is described in CS, Table 28 as, "*adult patients with untreated locally advanced unresectable or metastatic HER2 positive gastric or gastroesophageal junction adenocarcinoma whose tumours express PD-L1 with a CPS≥1 (based on non-Asia region cohort).*"[1] The company's rationale is that this is aligned with the anticipated marketing authorisation and is informed by a trial population that is representative of NHS patients. As previously discussed, this is narrower than the population specified in the NICE scope which was not restricted by PD-L1 status. In addition, the company has chosen to use data from the non-Asia region, to populate the model, including the baseline characteristics summarised in Table 9.

Table 9 Baseline patient characteristics of base case model cohort (non-Asia CPS≥1 patients) (reproduced from CS, Table 29[1])

Characteristics	CPS≥1 (non-Asia region)
Age (years), mean	60.2
Male (%)	79.1
Body weight (kg), mean	72.0
Body weight (kg), standard deviation	16.3
BSA (m^2), mean	1.8
BSA (m^2), standard deviation	0.2

Abbreviations: BSA, body surface area; CPS, combined positive score.
Source: KEYNOTE-811 (database cut-off date: May 25, 2022).

4.2.2 Interventions and comparators

The intervention arm in the model is pembrolizumab plus trastuzumab plus chemotherapy. In the base case pembrolizumab is assumed to be given once every 3 weeks at a dose of 200mg by intravenous infusion. A scenario analysis explores the impact of 6-weekly dosing at a dose of 400mg. Pembrolizumab is assumed to be given until progression or unacceptable toxicity, up to a maximum of 35 cycles. As previously discussed in Section 2.3.2, [redacted] but otherwise the usage of pembrolizumab is consistent with the draft SmPC.

The chemotherapy given alongside pembrolizumab and trastuzumab in the intervention arm is assumed to be either CAPOX or FP as these were the two chemotherapy regimens used in KEYNOTE-811. The proportions receiving CAPOX and FP in the pembrolizumab are assumed to be 77.2% and 22.8% respectively based on data from the pembrolizumab arm of the KEYNOTE-811 study.

The comparator in the company's economic analysis is trastuzumab plus chemotherapy, again assumed in the company's base case to be CAPOX or FP, with the proportions receiving each treatment based on usage in the KEYNOTE-811 study (78.5% and 21.5% respectively). The company also provided a scenario analysis in which XP was also an option in the comparator arm. The company's response to clarification question B34 stated that this presumed that 80% of patients received XP, with the remaining proportion being distributed equally between CAPOX and FP.[19] This distribution was based on clinical expert advice.

The doses for trastuzumab, CAPOX, FP and XP are summarised in Table 10. All treatments are given using a 3-week cycle, with capecitabine taken orally on days 1 to 14, 5-FU given by continuous infusion

on days 1-5 and the other treatments given on day 1 of the cycle intravenously. Trastuzumab is the only treatment to include a higher loading dose in the first cycle. The maximum number of cycles was assumed to be 35 for trastuzumab and 6 for the double chemotherapy agents, however, the number of cycles actually received is determined by the time on treatment data from KEYNOTE-811, which is then capped at the maximum value stated in Table 10. The EAG notes that the caps on the duration of treatment for each drug shown in Table 10 were not strictly applied in KEYNOTE-811 and this discrepancy is discussed further in Section 4.3.3.2. Scenario analyses were provided by the company in which the treatment durations were not capped.

Table 10 Dosing schedules assumed in the model (reproduced from CS, Table 48[1])

Regimen	Drug	Frequency	Dosage	Maximum treatment cycles	Source for dosage
	Pembrolizumab	Q3W	200mg IV	35 [29]	SmPC[16]
Loading dose	Trastuzumab	NA	8 mg/kg IV on Day 1	35 [29]	NICE TA208[10]
Maintenance dose		Q3W	6 mg/kg IV on Day 1		
CAPOX	Capecitabine	Q3W	1000 mg/m2 orally BID on Days 1–14	6	KEYNOTE-811[27]
	Oxaliplatin		130 mg/m2 IV on Day 1	6	
FP	5-FU	Q3W	800 mg/m2 IV on Days 1–5	6	
	Cisplatin		80 mg/m2 IV on Day 1	6	
XP (Scenario analysis only)	Capecitabine	Q3W	1000 mg/m2 orally BID on Days 1–14	6	NICE Guideline NG83[8]
	Cisplatin		80 mg/m2 IV on Day 1	6	

Abbreviations: 5-FU, fluorouracil; BID, twice daily; IV, intravenous; NA, not applicable; NICE, National Institute for Health and Care Excellence; Q2W, ever 2 weeks; Q3W, every 3 weeks; SmPC, summary of product characteristics.
Note: all chemotherapy regimens are capped at a maximum of 6 cycles based on UK clinical expert opinion

4.2.3 Perspective, time horizon and discounting

The company's economic analysis is described in the CS as taking an NHS and Personal Social Services (PSS) perspective (CS, Table 28).[1] The source used to estimate health care costs for progressed disease did not specifically capture palliative care and did not assess costs falling on community, hospice or social care services. However, the end-of-life costs included did cover community nursing and hospice costs for patient dying outside of a hospital setting. This is further discussed in Section 4.3.3.9.

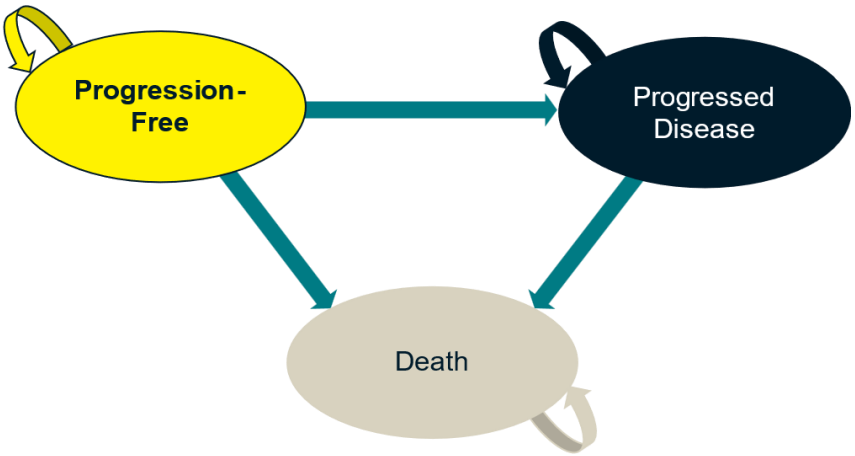
The company's base case uses a time horizon of 40 years with shorter time horizons of 8 and 20 years explored in scenario analyses. The company's model discounts future costs and benefits at 3.5% per annum.

4.2.4 Model structure

The general structure of the company's economic model is described on pages 95-97 of the CS[1] as a partitioned survival model based on three health states: (1) progression-free and alive; (2) post- disease progression and alive, and (3) dead (see Figure 2).

The EAG notes that occupancy of these health states influences only costs in the company's base-case analysis as HRQoL outcomes are modelled using a time-to-death approach rather than being based on the patient's progression status. However, the structure of the model allows the use of utilities by progression status which is explored by the company in a scenario analysis.

Figure 2: Company's model structure (reproduced from CS, Figure 12)[1]



In the company's base case analysis, patients enter the model in the progression-free state and receive first-line treatment with either pembrolizumab plus trastuzumab and doublet chemotherapy or trastuzumab and doublet chemotherapy; trastuzumab and doublet chemotherapy has been denoted standard of care (SoC) when describing the economic modelling.

The allocation of patients amongst the health states are determined by two chosen distributions, one for OS, and one for PFS. At any time, the probability of being alive and progression-free is given by the cumulative PFS survival curve. The probability of being alive following disease progression at any time is calculated as the cumulative probability of survival minus the cumulative probability of PFS. The probability of being dead at any time is the complement of the cumulative probability of survival. A partition survival approach does not explicitly model transitions between health states. Time on first-

line treatment is estimated directly from the treatment-specific time to treatment discontinuation (TTD) Kaplan-Meier (KM) data from KEYNOTE-811 study as explained in Section 4.2.6.1.3.

For the SoC arm, the cumulative probabilities of OS and PFS in each time interval are modelled using parametric distributions fitted to time-to-event data from the global (CPS≥ 1) cohort from KEYNOTE-811.[27] The OS and PFS curves for pembrolizumab plus trastuzumab and doublet chemotherapy (pembrolizumab plus SoC) are then modelled by applying constant HRs for OS and PFS from KEYNOTE-811 to the respective survival probabilities chosen for SoC. In contrast to the time-to-event data used for modelling OS and PFS for SoC, the HRs applied for OS and PFS were the ones reported for the non-Asia subpopulation.

The survivor functions and the evidence sources used to derive these functions are summarised in Table 11, with further detail provided in Section 4.2.6.1. Within each treatment group, the model applies two structural constraints: (i) that PFS must be less than or equal to OS, and (ii) that the OS risk for the modelled population must be at least as high as the mortality risk of the age- and sex-matched general population of the UK.

The EAG notes that the company in its clarification response stated that the global cohort data were used for modelling the survival curves for SoC as it was *"the most complete and quality-assured data set"* at the time of submission. However, it acknowledges that the company consider the non-Asia CPS ≥ 1 cohort, *"to be the most relevant to the England and Wales population"* (clarification response to question B2).[19]

HRQoL is assumed to be independent of treatment received and determined by the patient's time to death, based on four categorical groups (<30 days; ≥30 to 180 days; ≥180 to 360 days, and ≥360 days) with utility declining as patients approach death. Health utilities used in the model are based on the EQ-5D-5L data collected from the CPS ≥ 1 non-Asia region in KEYNOTE-811. Health utilities are adjusted to reflect reducing utilities with age across the life-time horizon.[46] In addition, the model explicitly includes QALY loss associated with Grade ≥3 AEs for pembrolizumab plus SoC and SoC alone. HRQoL inputs are further discussed in Section 4.2.6.2.

The model includes costs associated with: (i) drug acquisition; (ii) drug administration; (iii) subsequent treatment received; (iv) disease management; (v) management of AEs and (vi) end-of-life (terminal care) costs. Costs related to PD-L1 testing were not included as these *"tests are administered to all patients in both treatment arms of the model"*. Cost details are discussed in Section 4.2.6.4

The incremental health gains, costs and cost-effectiveness of pembrolizumab plus SoC versus SoC are modelled over a time horizon of 40 years using 1-week cycles. Half-cycle correction is applied only as a scenario analysis.

4.2.5 Key assumptions employed in the company's model

The company's model employs the following key assumptions for its base case:

- OS and PFS estimates from the trastuzumab and chemotherapy (CAPOX or FP) arm of the KEYNOTE-811 global cohort with PD-L1 CPS ≥1 are representative of expected OS and PFS under current standard care in England;
- The HRs for OS and PFS from the company's reported results for the non-Asia subgroup of patients with PD-L1 CPS ≥1 from KEYNOTE-811 are representative of the treatment effect expected from adding pembrolizumab to SoC in England;
- The HRs estimated during the trial period of KEYNOTE-811 (see Section 3.2.4) are expected to be constant over time both during and after treatment with pembrolizumab whereas waning of treatment effect was explored in scenario analyses;
- The model includes a general population mortality constraint to ensure that the risk of death for the modelled population is never lower than for the age-sex matched general population;
- The occupancy of the progression-free health state is constrained to ensure that there can never be more people in the progression-free health state than are alive;
- The rates of treatment discontinuation observed in the non-Asia subgroup of patients with PD-L1 CPS ≥1 from the respective treatment arms in KEYNOTE-811, for each separate component of treatment, are representative of the expected rates of discontinuation when these treatments are used for this patient group in England;
- Total number of treatment cycles given is constrained by a maximum number of treatment cycles that is specific to each component of the treatment combination (see Section 4.2.2), but is not constrained by progression status, meaning that the model allows patients to still get first-line treatment after progression;
- HRQoL is modelled according to the patients' time to death with utility declining as a patient approaches death and is therefore independent of treatment or progression status;
- A single administration cost is applied each cycle and this is based on the administration cost for the treatment component with the highest cost;
- Drug costing assumes no vial sharing for any intravenous drugs
- The proportions of patients receiving subsequent lines of treatment in each arm and are based on treatment arm specific data from KEYNOTE-811 but the durations of subsequent treatments are assumed to be the same across arms;

- The frequency of clinical follow-up visits and cardiac monitoring are assumed independent of treatment, but dependent on progression status with lower costs applied post-progression;
- A cost associated with terminal care was assumed in the model which was the same for all treatments evaluated;
- Only grade ≥3 AEs that occurred in ≥3% of all non-Asia CPS≥1 patients in either treatment group of KEYNOTE-811 are included in the company's model and these are assumed to occur at the start of treatment;
- All grade ≥3 AEs included in the model are assumed to have the same impact on HRQoL, but the utility decrement is applied for different durations for each AE
- All grade ≥3 AEs are assumed to require a hospital admission

The EAG notes in particular that the company has stated that they assume that the data from the non-Asia cohort are more clinically relevant and applicable to patients receiving treatment in England and therefore the majority of the model inputs were updated to use data from the non-Asia (CPS≥1) cohort (clarification response, B40).[19] The key exception was the OS and PFS data applied in the SoC arm which was based on the global (CPS≥1) cohort. These estimates also informed the OS and PFS in the pembrolizumab with SoC arm as these were estimated by applying a HR to the data for the SoC arm, although the HR was estimated from the non-Asia (CPS≥1) cohort.

4.2.6 Evidence used to inform the company's model parameters

Table 11 summarises the evidence sources used to inform the model's parameters in the company's updated base case analyses following the clarification process.[19] These are discussed in detail in the subsequent sections.

Table 11: Summary of evidence used to inform the company's base case analyses

Parameter group	Source
Patient characteristics (age, BSA, weight, proportion of females)	Based on characteristics of trial participants with PD-L1 CPS≥1 from the non-Asia region enrolled in KEYNOTE-811[27]
OS – SoC	A 2-knot odds spline model separately fitted to observed comparator* group OS data from KEYNOTE-811 (global cohort population with CPS ≥1).
OS – pembrolizumab plus SoC	The HR for OS for intervention** versus control group* estimated from KEYNOTE-811 (non-Asia subgroup with CPS ≥1) is applied to the OS survival function for SoC.
PFS – SoC	A 2-knot hazard spline model separately fitted to observed comparator* group PFS data from KEYNOTE-811 (global cohort population with CPS ≥1).
OS – pembrolizumab plus SoC	The HR for PFS for the intervention** versus control group* estimated from KEYNOTE-811 (non-Asia subgroup with CPS ≥1) is applied to the PFS survival function for SoC.

Parameter group	Source
TTD – pembrolizumab	Observed intervention group** TTD KM data from KEYNOTE-811 (non-Asia cohort with CPS ≥1) (truncated at 35 cycles).
TTD – trastuzumab and each component of either doublet chemotherapy (capecitabine, oxaliplatin, cisplatin, 5-FU)	Observed intervention group** and comparator group* TTD KM data from KEYNOTE-811 (non-Asia cohort with CPS ≥1). Separate KM data applied for each component of treatment in each arm. Trastuzumab capped at 35 cycles, chemotherapy capped at 6 cycles.
HRQoL	EQ-5D-5L data collected in KEYNOTE-811 (non-Asia subgroup with CPS ≥1) and mapped onto the 3L value set. Data analysed according to time to death (<30 days; ≥30 to 180 days; ≥180 to 360 days, and ≥360 days).
Frequency of AEs	AE frequencies for either treatment arm based on Grade ≥3 AEs with incidence of ≥3% from KEYNOTE-811 (non-Asia CPS≥1 analysis). Event frequencies were treatment arm specific and were adjusted to account for multiple AE episodes per patient.
QALY loss resulting from AEs	Estimated disutility was calculated based on analyses of EQ-5D-5L data from the KEYNOTE-811 as the difference between the "During Grade ≥3 AE" value and the "*without AE value*". This was the same irrespective of treatment arm. The duration for each AE was sourced from KEYNOTE-811 (non-Asia CPS≥1) and was assumed the same between treatment arms. QALY losses therefore only differ between arms due to differing frequencies of specific AEs
Probability of receiving subsequent therapy	Arm-specific proportions receiving each agent of subsequent treatments in KEYNOTE-811 (non-Asia cohort with CPS ≥1).
Mean duration of subsequent therapy	Agent-specific mean duration in KEYNOTE-811 (non-Asia cohort with CPS ≥1).
Drug acquisition costs	Electronic Market Information Tool (eMIT) and British National Formulary (BNF).[47, 48]
Drug administration costs	National Schedule of NHS Costs 2021/22[49]
RDI	Based on KEYNOTE-811 study (non-Asia cohort with CPS≥1) for first-line treatments but assumed to be 100% for subsequent therapies.
Disease management costs	Based on NICE TA208,[10] National Schedule of NHS Costs 2021/22,[49] and Gomez-Ulloa *et al*.[50]
Costs associated with AEs	Unit costs based on previous NICE TAs,[10, 14, 51] National Schedule of NHS Costs 2021/22.[49]
End of life care costs	Based on a previous NICE appraisal (TA522),[52] inflated to 2021/22 costs using the HCHS pay & prices and the NHSCII indices.[53]

5-FU - 5-fluorouracil; AE - adverse event; BSA - body surface area; CSP - combined positive score; EQ-5D-5L - EuroQol EQ-5D 5-level; HCHS - hospital & community health services; HR - hazard ratio; HRQoL - health-related quality of life; KM - Kaplan-Meier; NHSCII - NHS Cost Inflation Index; OS - overall survival; PFS - progression-free survival; QALY - quality-adjusted life year; RDI - relative dose intensity; TA - technology appraisal, TTD - time to treatment discontinuation
**Control group corresponds to the placebo plus trastuzumab and CAPOX/FP arm in KEYNOTE-811 study*
***Intervention group corresponds to the pembrolizumab plus trastuzumab and CAPOX/FP arm in KEYNOTE-811 study.*

4.2.6.1 Time-to-event parameters

The company's approach used for each individual endpoint and each arm is described in further detail in the subsequent sections. Time-to-event outcomes for the SoC and pembrolizumab plus SoC groups are based on data from the comparator and intervention arms of KEYNOTE-811.[27]

The EAG notes that based on the company's response to clarification questions, the non-Asia CPS ≥1 subgroup from the trial should be considered the most relevant to the population of England.[19] However, the company's base case still uses the KM data for PFS and OS from the global cohort CPS ≥1 of the KEYNOTE-811 to model the survival outcomes for the SoC arm.

4.2.6.1.1 Overall survival (OS)

The company used a proportional hazards modelling approach to extrapolate OS "*in accordance with the non-rejection of the proportional hazards assumption*". The proportional hazards modelling approach consists of two steps. The first step is to fit both standard parametric and Royston-Parmer spline models to the individual patient-level data (IPD) from the SoC arm of the global cohort with CPS≥1 KEYNOTE-811 (trastuzumab with CAPOX or FP [N= 296]). The second step is to apply the HR (0.67) calculated from a Cox regression model using the non-Asia CPS≥1 subgroup to the selected comparator model to derive the survival for the intervention arm.

In response to clarification question B2, the company clarified the reason for using the global (CPS≥1) cohort for the control arm as follows "*In the submitted model, OS and PFS extrapolations for the SoC arm were informed by data from the Global CPS≥1 cohort. Analysis of the population submitted for regulatory approval was prioritised with analysis of subgroups completed in succession. At the time of submission, the most complete and quality-assured data set was presented*."[19]

The company considered six standard parametric survival models: exponential, Weibull, Gompertz, log-logistic, log-normal and generalized gamma. In addition, various spline models with different assumptions (modelling the log cumulative hazard [hazard], the log cumulative odds [odds], or the inverse normal distribution of the survival function [normal] as a spline function) and different numbers of knots were also investigated.

The CS states that the candidate models were assessed for inclusion in the base case analysis through consideration of relative goodness-of-fit statistics (the Akaike Information Criterion [AIC] and the Bayesian Information Criterion [BIC]); visual inspection of the fitted distributions to the observed data; examination of the Schoenfeld residual and the cumulative hazard functions to judge the proportional hazard assumption; and expert opinion.

Among the standard parametric models, the company selected the log-logistic model based on AIC/BIC for the SoC arm which also had a reasonable visual fit to the hazard plot and to the KM curve. Similarly, the 2-knot odds model was selected as the best fit among the spline models. The latter was preferred over the log-logistic model based on *"visual comparison"*. Figure 3 presents the KM survival functions and modelled OS survival functions for both arms. In its base case the company assumed no treatment waning effect, therefore the survival benefit associated with adding pembrolizumab to SoC is sustained for the entire modelled time horizon. The company also presented a scenario analysis implementing a treatment waning from 7 years to 9 years.

Figure 3: OS survival functions included in company's base case analysis (adapted from CS, Figure 23)

In response to clarification question B6, the company performed alternative extrapolation approaches (jointly modelling approach with treatment as a covariate and independent modelling approach) using the non-Asia (CPS≥1) subgroup.[19] However, the results from these survival analyses have not been applied in the updated economic model. This is discussed in detail in in Section 4.3.3.2 of the EAG's critique, with the EAG's preferred approach to modelling OS described in Section 4.4.2.2.

4.2.6.1.2 Progression-free survival (PFS)

As with the OS analysis, the analysis of PFS was based on a proportional hazards modelling approach "*in accordance with the non-rejection of the proportional hazards assumption*". The first step is to fit a survival model to the SoC arm from the global cohort with CPS≥1 in KEYNOTE-811 (trastuzumab and CAPOX or FP [N=296]). The company fitted the same range of standard parametric models (exponential, Weibull, Gompertz, log-logistic, log-normal, and generalised gamma) and spline models.

The second step is to apply the HR (0.62) calculated from a Cox regression model using the non-Asia CPS≥1 subgroup to the selected comparator model to derive the survival for the intervention arm.

Among the standard parametric models, the log-logistic model was selected as the best fit for the SoC arm based on the AIC/BIC and having a reasonable visual fit to the hazard plot and to the KM curve. Similarly, the 2-knot hazard model was selected as the best fit among the spline models. The latter was preferred over the log-logistic model based on *"visual comparison"*. Treatment waning was not considered by the company due to the maturity of the trial data. A constraint is applied to the model to ensure that PFS must be less than or equal to OS at any given time. Figure 4 presents the KM survival functions and modelled PFS survival functions for both arms.

Figure 4: PFS survival functions included in company's base case analysis (adapted from CS, Figure 31)

In response to clarification question B6, the company performed alternative extrapolation approaches (jointly modelling approach with treatment as a covariate and independent modelling approach) using the non-Asia (CPS≥1) subgroup.[19] However, the results from these survival analyses have not been applied in the updated economic model. This is discussed in detail in in Section 4.3.3.2 of the EAG's critique, with the EAG's preferred approach described in Section 4.4.2.3.

4.2.6.1.3 Time to treatment discontinuation (TTD)

In KEYNOTE-811, TTD data for all CPS ≥1 global cohort patients were relatively mature, and the KM data were used directly to inform the company's base case model without the need for parametric extrapolation which the company deemed to *"introduce additional uncertainty to a dataset which is deemed reasonably informative"*. Arm-specific TTD data were used to inform treatment acquisition

and administration costs and were available for each single agent involved (i.e., pembrolizumab, trastuzumab, capecitabine, oxaliplatin, 5-FU, and cisplatin) as depicted in the CS, Figures 32 to 37.[1]

In addition, all treatments had maximum durations after which all patients discontinue this treatment. These were 35 cycles for pembrolizumab and trastuzumab in line with the trial protocol, and 6 cycles for all chemotherapy agents, in line with NHS clinical practice, as confirmed by the company's clinical experts. The EAG has commented on the appropriateness of capping the number of cycles for each treatment in the combination in Section 4.3.3.3.

The EAG notes, however, that a constraint to ensure that TTD does not exceed PFS is not included in the base case analyses. This leads to the assumption that patients can receive first-line treatment after disease progression. However, after examining the TTD and PFS curves the EAG did not think this was likely to be a significant issue.

4.2.6.2 Health-related quality of life (HRQoL)

HRQoL data used in the company's model are based on EQ-5D-5L data collected in KEYNOTE-811 from the non-Asia subgroup with CPS≥1. Within the study, the questionnaire was administered at baseline, every 3 weeks for the first 5 treatment cycles (weeks 1, 4, 7, 10 and 13), then every 6 weeks until week 52 or end of treatment, whichever was earlier; in the case of treatment discontinuation, the questionnaire was also applied at the treatment discontinuation and 30-day post-treatment safety follow-up visits. The utility values were then mapped to the 3L value set using the mapping function developed by the Decision Support Unit (DSU).[54]

Utility values in the base case analysis were estimated for the pooled treatment arms by proximity to death, based on four categorical groups (<30 days; 30 to 179 days; 180 to 359 days, and ≥360 days). The utilities for each time-to-death category are assumed to be independent of initial treatment.

Within the model, the proportion of patients in the time-to-death categories at each time *t* were calculated as follows:

- < 30 days from death: calculated as the probability of dying during the interval *t*+0 cycles and *t*+4 cycles;
- 30 days to 179 days from death: calculated as the probability of dying during the interval *t*+5 cycles and *t*+25 cycles;
- 180 days to 359 days from death: calculated as the probability of dying during the interval *t*+26 cycles and *t*+50 cycles;

- ≥360 days from death: calculated as the 1 minus of the sum of the probabilities of being in the other three states.

The EAG notes that the description of the time-to-death categories do not align with the implementation in the model. In the model, the four categories are: <4 weeks (28 days); ≥4 to 24 weeks (28 to 175 days); ≥25 to 51 weeks (175 to 357 days), and ≥51 weeks (357 days). The EAG notes, however, that this is unlikely to noticeably affect the ICER and fitted in with the weekly time cycle in the model.
The use of a time-to-death approach for modelling HRQoL is justified by the company on the basis that it would overcome the problem of limited questionnaire availability to inform the post-progression health state utility estimates, which is a consequence of the EQ-5D questionnaire collection not being collected after treatment discontinuation or beyond 30-days after disease progression. Therefore, the estimates of utility data for post-progression health state may not be representative of the patient's quality of life in the whole post progression state. The estimates for utility data applied in the company's model are summarised in Table 12. Additionally, the company provided a scenario analysis where the EQ-5D data were analysed by the progression status using data pooled across both trial arms, and this resulted in mean utility values of [redacted] and [redacted] for the progression-free and progressed-disease health states respectively. A second scenario was also conducted which used the baseline utility value from the trial for the PFS state ([redacted]), and then utility for the progressed disease state was estimated using the difference in utilities between progression-free and progressed disease patients in the trial, giving a utility of 0.706 ([redacted]).

Table 12: Mean EQ-5D utilities used in the company's base case analyses (reproduced from CS Table 40)

Time-to-death (days)	N	Mean	SE
<30	[redacted]	[redacted]	[redacted]
30 to 180	[redacted]	[redacted]	[redacted]
180 to 360	[redacted]	[redacted]	[redacted]
≥ 360	[redacted]	[redacted]	[redacted]

N, number of participants with non-missing score; SE, standard error

Health utilities are adjusted for aging by using utility multipliers for each age. This was achieved by estimating general population utility values at the baseline starting age of the model and subsequent ages using the DSU database.[54] The multiplier was then calculated by dividing the utility value at any specific age by the baseline utility value, this value was then multiplied by the QALYs calculated for each of the time-to-death categories per cycle.

Table **13** shows a selection of multipliers used at certain ages. The removal of the health utilities age-adjustment was explored in the company's scenario analyses.

Table 13: Utility multipliers used in the company's base case to adjust for utility decline by age*

Age	General population utility	Utility multiplier
60.2 (starting age)	0.845	1.00
65	0.828	0.98
70	0.810	0.96
75	0.789	0.93
80	0.768	0.91
85	0.744	0.88
90	0.718	0.85
95	0.689	0.82
100	0.656	0.78

** annual declines implemented but only 5-year values presented here*

4.2.6.3 Adverse events

The company's model included all Grade 3+ AEs and those that occurred in at least 3% of all patients in either arm of the KEYNOTE-811 trial. These data were initially based on the global cohort, as presented in CS, Table 55.[1] However, in response to clarification question B40, these data were updated in the model to reflect the non-Asia subgroup for the non-Asia (CPS≥1) cohort.[19] The updated data, extracted from the model by the EAG, are presented in Table 14 were further adjusted by the company to account for the mean number of each AE per patient.

Table 14: AE frequency per treatment arm used in the company's base case model (non-Asia subgroup CPS ≥1 in KEYNOTE-811; supersedes global cohort data in CS, Table 55[1])

Adverse event	% of patients experiencing the event		Mean number of events per patient		Adjusted % of patients experiencing the event*	
	Pembrolizumab + SoC†	SoC†	Pembrolizumab + SoC†	SoC†	Pembrolizumab + SoC†	SoC†
Anaemia	[redacted]					
Neutropenia						
Thrombocytopenia						
Diarrhoea						
Nausea						
Vomiting						
Asthenia						
Fatigue						
Neutrophil count decreased						
Platelet count decreased						
Decreased appetite						
Hypokalaemia						
Peripheral sensory neuropathy						

**defined as the total number of AE episodes (considering that some patients experienced multiple AE episodes) divided by the total patient number*

† SoC = trastuzumab and chemotherapy

The disutility for modelled AEs ([redacted]), which was assumed to be the same for all modelled AEs, was estimated based on analyses of EQ-5D data from the KEYNOTE-811 trial, as the difference between the "*During Grade 3+ AE*" utility value and the "*without AE*" utility value. In response to clarification question B15, that company stated that the difference between the "*During Grade 3+ AE*" value and the "*without Grade 3+ AE*" value was similar ([redacted]) and using this alternative estimate had only a minor impact on the ICER.[19]

Mean duration of AE per affected patient in KEYNOTE-811 were reported in Table 44 of the CS.[1] The company stated in response to clarification B40, that the AE duration had been updated to use data from the non-Asia region,[19] but the EAG noted that the data appeared to be identical to those provided prior to clarification.

QALY loss due to AEs was incorporated in the model for the modelled cohort by multiplying the disutility by AE-specific mean duration and by AE incidence (specific to both the treatment arm and the individual AE) and applying this as a one-off QALY loss in the first cycle of the model. This accounted to QALY losses of [redacted] and [redacted] for the pembrolizumab and the SoC arms respectively.

4.2.6.4 Resource use

4.2.6.4.1 Drug acquisition and administration costs

Drug acquisition costs have been calculated based on the dosing schedules provided in Table 10. There is a patient access scheme (PAS) in place for pembrolizumab. The cost per cycle provided in Table 15 incorporates this PAS. None of the other treatments included in the first-line treatment regimens covered in Table 15 have a confidential PAS. The company obtained used an NHS indicative price for trastuzumab and eMIT prices for all other first-line therapies. NICE has provided the EAG with confidential prices for trastuzumab and capecitabine from the commercial medicines unit (CMU) and the impact of including these is explored in a confidential appendix. The prices cited by the company have been used in the EAG analyses reported in Section 4.4.

The company has applied a simple relative dose intensity (RDI) approach to account for missed or delayed doses between the first and last dose received in KEYNOTE-811 (see clarification response B22).[19] This is calculated as the actual number of cycles administered divided by the expected number of cycles administered based on the time between the first and last dose (multiplied by 100 to convert the proportion into a %). For drugs received once per cycle, this correlates to the proportion of doses received, whereas for drugs where more than one dose is administered per cycle, a cycle is counted as having been administered provided a single dose has been received. Therefore, this approach does not account for missed doses for patients self-administering oral capecitabine. It also does not account for any administrations where the dose was reduced but still administered even though dose modifications

were permitted in KEYNOTE-811.[29] As the company considers the RDI to be confidential, the EAG has summarised the drug costs when assuming 100% RDI in Table 15. The actual RDI values applied in the model can be found in CS, Table 49.[1]

For pembrolizumab the dose is not dependent on weight or body surface area (BSA) and the 200mg dose can be achieved using a whole number of vials. For all other intravenous drugs, the company's base case analysis includes drug wastage, assuming in its calculations that no vial sharing across patients occurs and any partially used vials are discarded. The company uses a method of moments approach to estimate the average number of vials required based on assuming a lognormal distribution for patient weight. The company assumes no wastage when estimating the costs for treatments administered orally. The company also provides a scenario analysis in which vial sharing is assumed to occur in 100% of intravenous administrations resulting in zero drug wastage.

The CS assumes that all combinations of intravenous treatments can be given within a single session covered by a single reference cost. Treatment combinations including 5-FU are assumed to be covered by SB14Z (Deliver complex chemotherapy, including prolonged infusion treatment, at first attendance) at a cost of £474.94, due to the requirement for prolonged infusion over 5 days. All other treatment combinations are assumed to be covered by SB13Z (Deliver more complex parenteral chemotherapy at first attendance) at a cost of £353.64. The company states that both trastuzumab and pembrolizumab are considered to be complex treatments with short infusion times. It argues that it is appropriate to include only a single reference cost for complex chemotherapy (SB13Z) when these drugs are given after chemotherapy is completed (i.e after CAPOX or XP) whether trastuzumab is given alone or in combination with pembrolizumab. However, the company did provide a scenario in which a lower cost was applied when trastuzumab is given alone, in response to clarification question B27, and the further critique is provided on this issue in Section 4.3.3.6. Administration of oral capecitabine is assumed to incur no additional cost when given alongside an intravenous treatment.

Table 15 Drug acquisition and administration costs per cycle (when assuming 100% RDI, base case assumes wastage)

Treatment	Dose per administration	Administrations per cycle	Drug dose in mg (no wastage)	Drug dose in mg (with wastage)	Drug cost per cycle, £ (no wastage)	Drug cost per cycle, £ (with wastage)	Admin cost per cycle, £
Pembrolizumab	200 mg	1	200	200	[redacted]	[redacted]	0
Trastuzumab	6 mg/kg*	1	576	665	1,056	1,267	/ 354**
Capecitabine (CAPOX)	1000 mg/m^2	28	1800	1800	31	31	354
Oxaliplatin (CAPOX)	130 mg/ m^2	1	234	384	25	41	
5FU (FP)	800 mg/ m^2	5	1440	2500	12	20	475
Cisplatin (FP)	80 mg/ m^2	1	144	200	16	22	
Capecitabine (XP)	1000 mg/m^2	28	1800	1800	31	31	354
Cisplatin (XP)	80 mg/ m^2	1	144	200	16	22	

* 8 mg/kg loading dose has a drug cost of £1,408 without wastage and £1,625 with wastage;

** £354 when given after chemotherapy, either together or alone – zero additional cost when given with chemotherapy

4.2.6.4.2 Subsequent treatments

The company has estimated subsequent treatment by combining treatments received in any treatment line after completing or discontinuing the study drug. The company has restricted its analysis to the top eight subsequent treatments received but has increased the usage of these to incorporate the usage of other less frequently received treatments. It has also redistributed the proportion receiving paclitaxel with ramucirumab as this treatment combination is not available in clinical practice in England despite being one of the eight most common subsequent treatments in KEYNOTE-811. This redistribution is assumed to affect only the cost of the subsequent therapies with clinical outcomes being assumed to be unchanged by the distribution of subsequent therapies. The resultant proportions are summarised in Table 16, alongside the durations of subsequent treatment which have been estimated across both study arms. When calculating the drug acquisition and administration costs for subsequent treatments, the company has used as similar approach to that taken for first-line therapies but has assumed an RDI of 100%. Combining the information on the distribution of subsequent therapies, their duration and their costs provides a total cost for subsequent therapies of £5,556 and £3,683 for the intervention and comparator arms (these supersede the figures given in CS, Table 58 of £8,283 and £8,549 respectively[1]). The EAG notes that the company's analysis does not include the confidential PAS price for trastuzumab deruxtecan as this information was not available to the company. The EAG has provided a confidential appendix which includes the impact of incorporating the confidential PAS for trastuzumab deruxtecan. The price for trastuzumab deruxtecan cited by the company has been used by the EAG in the analyses reported in Section 4.4.

The average cost of subsequent treatment per patient completing or discontinuing their study drug has been applied to patients leaving the progression-free health state. This means it is applied to patients at the time of either progression or death, rather than at the time of completing or discontinuing study drug. The company has also provided a scenario analysis in which the costs of subsequent treatment are increased to account for the fact that only a proportion of the cohort have progressed at the time of the study follow-up. The proportions for this scenario are also presented in Table 16. In this scenario the subsequent treatment costs are £9,739 and £5,892 for intervention and comparator arms respectively. Based on clinical advice, a scenario analysis has also been provided assuming that only 50% of patients receive subsequent therapies and these are evenly split between docetaxel and platinum rechallenge with CAPOX, with the intention of reflecting treatments received in current clinical practice in England rather than those received in KEYNOTE-811. The subsequent treatment costs in this scenario are £902 for both arms. Further critique of the company's estimation of subsequent therapies is provided in Section 4.3.3.7.

Table 16 Proportions of patients receiving subsequent treatments per treatment arm (non-Asia CPS≥1 cohort)[*]

Subsequent treatments[†]	Company's base case		Scenario where % is uplifted to account for proportion who have not progressed		Mean duration of subsequent treatment across both arms (weeks)	Drug acquisition cost per week, £	Drug administration costs per week, £
	Pembrolizumab with SoC[‡], %	SoC[‡], %	Pembrolizumab with SoC[‡], %	SoC[‡], %			
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	8.04	88.41
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	2,234.88	95.57
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	0.26	143.36
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	3.19	237.47
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	4.41	95.57
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	0.85	237.47
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	351.98	95.57
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	NA	NA	NA

[*] *adapted from clarification response Table 43,[19] which supersedes CS Table 57; mean duration has been extracted from the model as data from CS Table 63 have been superseded by data for the non-Asia CPS≥1 cohort but these were not presented in the clarification response[19]*

[†] *Most common treatment combinations or monotherapies excluding paclitaxel with ramucirumab - accounted for 6.7% and 12.8% across intervention and control arms – % receiving this combination and % receiving any other treatments were redistributed to give correct total % receiving subsequent therapies*

[‡] *SoC = trastuzumab with chemotherapy*

4.2.6.4.3 Disease management by health state

Resource use for the progression-free health state (excluding administration of first-line treatment) is based on information from the appraisal of trastuzumab in TA208. The company assumes an overall cost of £176 per week, (see CS Table 52 for details) which covers oncology outpatient attendances and cardiac monitoring. The CS applies two difference reference costs for follow-up oncology appointments, with one applied once per three weeks and the other applied once every six weeks, giving a total of 26 follow-up visits per annum during PFS. The CS assumes cardiac monitoring 4 times a year with one third of this monitoring being by multigated acquisition scan (MUGA) scan and the other two thirds by echocardiogram. The monitoring costs are being applied for the whole of the PFS duration rather than for the duration of trastuzumab, although the EAG notes that the intention would be to continue trastuzumab until disease progression, with treatment being only stopped before then only due to unacceptable toxicity. Therefore, the duration of trastuzumab treatment is likely to be similar to the progression free duration in the majority of patients.

Resource use for the progressed disease health state were based on a retrospective chart review of patients receiving second line treatment for confirmed metastatic or unresectable gastric or GOJ adenocarcinoma in one of 5 countries, including the UK. Each patient's charts were reviewed for 12 months after starting second line treatment or until death, which ever occurred first. The majority of patients UK (92%) patients had received HER2 status testing and 20% were HER2-postive. The majority had received triplet chemotherapy at 1st line (73%). The paper reports the percentage of patients receiving different types of health care resources including hospital admission, emergency room visits and outpatient visits. The mean observation period was 6.6 months and the CS states that each patient reported as having used a particular type of resource is assumed to have used it once in that period. The overall cost is £2,132 per annum (£42 per week, see CS, Table 53 for details).[1] The EAG has provided further commentary on the appropriateness of the resource use estimates based on this study in Section 4.3.3.8.

4.2.6.4.4 Adverse event management

The company model includes resource use for hospital admission for a non-elective short stay for each of the grade 3+ AEs included in the model, and these are reproduced in Table 17. These are applied as a one-off cost assuming that AEs occur mainly during the first cycle of treatment. The proportion of patients experiencing one or more AE of each type is increased to account for the mean number of AEs per patient (see Table 14). Overall, this results in a cost of £565 for the pembrolizumab arm and £394 for the comparator arm (these supersede values in Table 56 of the CS[1], which was based on the frequency of AEs in the global CPS >=1 cohort rather than the non-Asia CPS>=1 cohort). The EAG's clinical advisors stated that sepsis and diarrhoea were the main AEs that result in admission in this

patient population. They agreed that AEs related to chemotherapy would be likely to occur early in the model, however, they noted that rare but severe AEs can occur with ongoing pembrolizumab treatment. In response to clarification question B30, which asked the company to explore the potential impact of rare but severe immune-related AEs for patients receiving pembrolizumab, the company conducted an analysis which increased the cost due to AEs by 10% for the intervention arm only.[19] Based on this analysis which showed limited impact on the ICER, the company concluded that its base case approach to modelling AEs was robust.

Table 17 Adverse event unit costs (reproduced from CS, Table 54[1])

Adverse event	Unit cost (£)	Notes
Anaemia	770	Weighted average of SA01G-K non-elective short stay: based on ERG criticism in TA737
Neutropenia	2,257	Weighted average of SA35A-E; note that TA208 currency codes for febrile neutropenia have been discontinued
Thrombocytopenia	993	Weighted average of SA12G-K: consistent with TA857
Diarrhoea	522	FD10M non-elective short stay; consistent with TA857, TA208 codes have been discontinued
Nausea	522	Assumed equal to diarrhoea
Vomiting	522	Assumed equal to diarrhoea
Asthenia	780	Assumed equal to fatigue
Fatigue	780	SA01G - Aplasia or Other Aplastic Anaemia, with CC Score 8+. Non-elective short stay (consistent with TA737). NR in TA208
Neutrophil count decreased	445	Non-elective short stay. WJ11Z Other disorders of immunity (consistent with TA737). NR in TA208
Platelet count decreased	993	Assumed equivalent to thrombocytopenia
Decreased appetite	561	Weighted average of Non-elective short stay FD04B-E; NR in TA208
Hypokalaemia	2,257	Assumed equivalent to neutropenia
Peripheral sensory neuropathy	607	Weighted average of AA26C-H, Acute setting

Abbreviations: ERG, Evidence Review Group; NR, not reported; TA, technology appraisal

4.2.6.4.5 End of life costs

The company's estimate of end-of-life costs was taken from the appraisal of pembrolizumab for untreated PD-L1-positive locally advanced or metastatic urothelial cancer (TA522).[52] This CS reported this cost as £7,253 (2015/2016 value) for hospital care in the last three months of life.[1] This was increased by the company to £8,169 to reflect current prices.[1, 55]

4.2.6.4.6 PD-L1 testing costs

The company has assumed in its base case that no additional testing will be required to determine eligibility for pembrolizumab in this cohort based on PD-L1 status because patients are already routinely tested for HER2 and PD-L1 status concurrently in order to determine eligibility for nivolumab which is already recommended by NICE in the HER2-negative cohort (TA857).[1, 14] In its response to clarification question B26, the company provided information on the potential PD-L1 testing costs that would apply if testing was not already being carried out.[19] They estimate a cost of £424 per patient eligible to receive pembrolizumab based on a cost of £53 per test and an estimate that eight patients would need to be testing to identify a single eligible patient. However, when these data are incorporated in the model inputs, this cost is applied to both treatment arms and therefore it has no impact on the ICER.

4.2.7 Model validation and face validity check

The company describes its model validation process as including quality checks, clinical expert opinion and comparison with external data sources. The quality checks on the model (verification) were conducted by an independent health economist using the TechVER checklist.[56] The company sought expert opinion from two clinical experts who are experienced in the management of HER2-positive advanced gastric or GOJ cancer patients in England. It said that these discussions were used to ensure that the base case reflects current UK practice and to assess the face validity of the outcomes predicted by the model. The CS states that the comparison with other trial data was limited by the paucity of trials available and its comparison of model outcomes was limited to the model used to inform the appraisal of trastuzumab in TA208.

4.2.8 Cost effectiveness results

The probabilistic and deterministic results presented in this section are based on the updated version of the company's model submitted in response to the clarification process. The results presented in this section include the company's agreed PAS for pembrolizumab whilst excluding price discounts available for any other drugs used in subsequent treatments. The results incorporating the confidential PAS discount for trastuzumab deruxtecan and the CMU prices for trastuzumab and capecitabine are provided in a confidential appendix to this EAG report. The company has presented evidence to support a QALY weight of 1.2, based on its assessment of the severity modifier. The company's evidence to support this severity modifier is further discussed in Section 5. The EAG has presented company results both with, and without, this QALY weight.

Central estimates of cost-effectiveness

The company's base case cost-effectiveness results are presented in Table 18, which shows the probabilistic estimates of the company's base case estimated using the average costs and QALYs across 1,000 probabilistic sensitivity analysis (PSA) samples when the model was rerun by the EAG. Total costs, QALYs and ICERs were judged to have converged after running the PSA 1,000 iterations.

The probabilistic version of the model suggests that the pembrolizumab arm is expected to generate an additional [redacted] QALYs at an additional cost of £[redacted] per patient compared to the SoC arm resulting in an ICER of £[redacted] per QALY gained (£[redacted] when the QALY weight is 1.2). The deterministic version of the model produces a slightly lower ICER (£[redacted] per QALY gained without QALY weight). QALY gains predominantly relate to differences in survival as utility is related to time-to-death rather than progression status (1.94 additional life years gained on the pembrolizumab arm in the probabilistic model).

Table 18 The company's base case results

Technology	LYs	QALYs accrued	Total costs incurred	Incremental			ICER	ICER with 1.2x QALY weight
				LYs	QALYs	Costs		
Probabilistic model (1000 runs by the EAG)								
SoC*	3.12	[redacted]	[redacted]	-	-	-		
Intervention**	5.06	[redacted]	[redacted]	1.94	[redacted]	[redacted]	[redacted]	[redacted]
Deterministic model								
SoC*	3.03	[redacted]	[redacted]	-	-	-		
Intervention**	4.94	[redacted]	[redacted]	1.91	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: ICER, incremental cost-effectiveness ratio; LYs, life years; QALYs, quality-adjusted life years
** SoC: Trastuzumab plus chemotherapy*
*** Intervention: Pembrolizumab with SoC*

The company's model presents disaggregated outcomes for the deterministic model in terms of costs accrued by different elements and QALYs accrued in different time-to-death categories. These results are presented in Table 19. The differences in costs are primarily associated with the acquisition cost of pembrolizumab whilst the additional QALY gain is mainly a consequence of additional time spent on the pembrolizumab arm in the over-360-day time to death category compared to the SoC arm, and the higher utility value associated with such category.

Table 19 Base case disaggregated outcomes for company's base case (deterministic model)

Description	Intervention**	SoC*	Incremental
Disaggregated costs (discounted)			
Drug acquisition costs	[redacted]	[redacted]	[redacted]
Drug administration costs	[redacted]	[redacted]	[redacted]
Subsequent treatment costs	[redacted]	[redacted]	[redacted]
AE related costs	[redacted]	[redacted]	[redacted]
Disease management costs	[redacted]	[redacted]	[redacted]
End of life costs	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]
Disaggregated QALYs (discounted)			
Time to death <30 days	[redacted]	[redacted]	[redacted]
Time to death 30-179 days	[redacted]	[redacted]	[redacted]
Time to death 180-359 days	[redacted]	[redacted]	[redacted]
Time to death ≥360 days	[redacted]	[redacted]	[redacted]
QALYs gained with AEs	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]

Abbreviations: AE, adverse event; QALY, quality-adjusted life-years.
** SoC: Trastuzumab plus chemotherapy*
*** Intervention: Pembrolizumab with SoC*

Figure 5 presents the cost-effectiveness plane for the company's base case PSA, and Figure 6 shows the corresponding cost-effectiveness acceptability curve (CEAC) (both based on the EAG's re-run of 1,000 PSA samples). The EAG's re-run of the company's PSA suggests that the probability that the pembrolizumab arm generates more net monetary benefit than the SoC arm at a WTP threshold of £20,000 and £30,000 per QALY gained is approximately [redacted] and [redacted] respectively. The same probabilities are [redacted] and [redacted] respectively when a QALY has 1.2x weight.

Figure 5: Company's base case PSA scatterplot with the QALY weight of 1x (run by the EAG)

Figure 6: Company's base case CEAC with the QALY weight of 1x (run by the EAG)

4.2.9 Company's deterministic sensitivity analyses

The company's deterministic sensitivity analyses were rerun by the EAG post-clarification and are presented using a tornado plot (Figure 7 and Figure 8 for a QALY weight of 1 and 1.2 respectively). The analyses are performed by using the lower and upper bounds of 95% confidence intervals assuming that the standard error was set as 20% of the mean if not reported.

The company's results show that the parameters which had the biggest impact on the ICER were: the HR value used to extrapolate the OS survival curve for the pembrolizumab arm (ICER difference of ~£[redacted] between when using the lower bound and upper bound values); relative dose intensity associated with pembrolizumab (ICER difference of ~£[redacted]); the HR value used to extrapolate the PFS survival curve for the pembrolizumab arm (ICER difference of less than £[redacted]); relative dose intensity associated with trastuzumab (ICER differences of [redacted]), and the unit cost and frequency of clinician's visits while still progression-free (ICER differences of less than £[redacted]). None of the other parameter ranges explored produced an ICER difference above £[redacted] per QALY gained.

Figure 7: One-way scenario analysis results for the company's post-clarification base case at a QALY weight of 1

Figure 8: One-way scenario analysis results for the company's post-clarification base case at a QALY weight of 1.2

4.2.10 Sensitivity analyses

Updated results for scenario analyses for the pembrolizumab arm versus the SoC arm are provided in the clarification response to question B42 with, and without, using the QALY weight (CS, Tables 54 and Table 55 respectively). The EAG requested additional scenarios which were provided in the clarification response to questions B12 and B36. All the ICERs reported within the text of this section are without the QALY weight. The scenarios with the largest impact that increased the ICER were limiting the model's time horizon to 8 years (which increases the ICER from £[redacted] to £[redacted]), the treatment effect on OS waning gradually between 7 and 9 years from the model start (i.e. applying a gradual increase to the OS HR till it approaches 1) (increases the ICER to ~£[redacted]), limiting the model's time horizon to 20 years (ICER increases to ~£[redacted]), and using a utility value set based on the progression status (which increases the ICER to between £[redacted] when progression-based utilities are used from KEYNOTE-811 and ~£[redacted] when using baseline utilities from KEYNOTE-811 for PFS and maintaining the proportionate difference between PFS and progressed-disease from KEYNOTE-811). The only scenarios that had a large impact but decreased the ICER were those assuming no discounting or 1.5% discounting (ICERs between ~£[redacted] with no discounting and ~£[redacted] with 1.5% discounting).

The following scenarios had less impact on the ICER (less than £[redacted]) compared with the above mentioned scenarios: assuming an RDI of unity for pembrolizumab, trastuzumab, and all first-line chemotherapy; not applying age-related disutilities; different assumptions regarding subsequent therapy in the UK and the proportions receiving it; pembrolizumab administered at a dose of 400 mg every 6 weeks instead of a 200 mg 3-week cycle; assuming vial sharing; excluding end-of-life costs; excluding disutility attributed to AEs; first-line chemotherapy distribution informed by clinical experts instead of using trial data; using the mean number of cycles as observed in the trial to decide treatment duration; and removing the treatment duration cap.

4.3 Critique of company's submitted economic evaluation by the EAG

4.3.1 Model verification

The EAG believes the company's updated version of the model to be generally well programmed with two exceptions. The way the model is coded means it is not possible to select the choice of cohort (global versus non-Asia) used to inform the OS and PFS in the comparator arm separately from the choice between using a HR approach or using separately fitted curves to model the intervention arm. The EAG also identified a minor error related to the administration costs for paclitaxel when used as a subsequent therapy, which is described in Section 4.3.3.6. The impact of correcting this error is explored in Section 4.4.2.1. The EAG also experienced issues when working with the company's model which sometimes froze or closed unexpectedly without saving a recovery backup version.

4.3.2 Adherence of the company's model to the NICE reference case

The EAG has summarised the adherence of the company's model to the NICE reference case in Table 20. The main issues identified related to the choice of relevant comparators, in particular the choice of doublet chemotherapy given in combination with trastuzumab. However, these issues have been previously covered in detail in Section 2.3.3 and 4.2.2.

Table 20 Adherence of the company's economic analysis to the NICE reference case

Element	Reference case	EAG comments
Population	The scope developed by NICE	The population in the company's economic model is narrower than the population in the NICE scope because it is restricted to patients with PD-L1 CPS ≥ 1. The EAG accepts that this is appropriate because it is aligned with the anticipated marketing authorisation. The company has also assumed that the non-Asia region cohort from KEYNOTE-811 are most representative of the patients likely to be offered treatment in clinical practice in England and has therefore used this cohort to define the starting characteristics in the model and to source the majority of the model parameters.
Intervention	As listed in the scope developed by NICE	The intervention is pembrolizumab in combination with trastuzumab and chemotherapy (pembrolizumab plus SoC). The chemotherapy offered in the intervention arm is assumed to be the same as offered in the comparator arm and is therefore discussed below.
Comparator(s)	As listed in the scope developed by NICE	The comparator is trastuzumab with chemotherapy (SoC), with chemotherapy assumed to be either CAPOX or FP in the company's base case analysis. The proportions receiving either doublet chemotherapy is informed by the treatment regimens offered in KEYNOTE-811. The EAG notes that only FP or XP are used in combination with trastuzumab in current clinical practice in England and XP is preferred when patients are able to tolerate oral treatments. The company has also explored a scenario analysis in which the majority of patients receive XP, which the company considers better reflects chemotherapy usage in England based on clinical expert advice. The company has not included either triplet chemotherapy or doublet chemotherapy without trastuzumab as comparators in the economic model. The EAG's comments on this have been given previously in Table 4, but in summary, the EAG considers this to be reasonable as current practice is to offer trastuzumab with doublet chemotherapy, rather than doublet chemotherapy alone, in any HER2-positive patient where trastuzumab is not contraindicated. In addition, triplet chemotherapy is not widely used in the indication.
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	The company's approach is consistent with the NICE reference case. Health gains accrued by patients are valued in terms of QALYs gained. Health impacts on carers are not included.

Element	Reference case	EAG comments
Perspective on costs	NHS and PSS	The company's base case analysis adopts an NHS and PSS perspective. This is therefore consistent with the NICE reference case. However, the EAG notes that costs for social care do not appear to have been included except in the context of end-of-life costs.
Type of economic evaluation	Cost-utility analysis with fully incremental analysis	The company has not provided a fully-incremental analysis against each of the comparators specified in the NICE scope because they have argued that triplet chemotherapy and doublet chemotherapy without trastuzumab are not relevant comparators. They have therefore only provided a single comparison against trastuzumab with doublet chemotherapy. The company has also not provided an incremental comparison against trastuzumab combined with each possible combination of doublet chemotherapies (see Table 3 for the possible combinations). Instead, it has assumed that each fluoropyrimidine and platinum-containing doublet chemotherapy is clinically equivalent, and has used the mix of doublet chemotherapy treatments offered in KEYNOTE-811 in its base case. The company has explored the impact of altering the mix of doublet chemotherapies offered but this only affects estimates of costs.
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	A 40-year horizon has been adopted. This is considered by the EAG to be consistent with the NICE reference case in this population.
Synthesis of evidence on health effects	Based on systematic review	The company conducted a systematic review, but only one study, KEYNOTE-811, was identified to inform the clinical outcomes in the model. The company considered the feasibility of conducting an indirect comparison against doublet chemotherapy alone (i.e., without trastuzumab) using data from the TOGA study, but concluded that this was not feasible.
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of HRQoL in adults.	Health gains are valued in terms of QALYs. Utility values obtained from the EQ-5D-5L in the KEYNOTE-811 study have been incorporated in the company's economic analysis. These have been mapped, using an appropriate approach, to a UK general population valuation set for the EQ-5D-3L.
Source of data for measurement of HRQoL	Reported directly by patients and/or carers	
Source of preference data for valuation of changes in HRQoL	Representative sample of the UK population	

Element	Reference case	EAG comments
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit, except in specific circumstances	The company has presented evidence to support a severity modifier of 1.2x using estimates of expected QALYs in people treated with trastuzumab with chemotherapy based on estimates from TA208. In response to clarification, the company has presented ICERs both with and without the severity modifier applied. The company's evidence in support of the severity modifier is commented on in Section 5.
Evidence on resource use and costs	Costs should relate to NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS	The company's base case cost-effectiveness analysis generally used appropriate estimates of resource use and unit costs that were consistent with the NICE reference case.
Discount rate	The same annual rate for both costs and health effects (currently 3.5%)	Costs and health effects are discounted at a rate of 3.5% per annum. This is consistent with the NICE reference case.

4.3.3 Key issues identified from the EAG's critical appraisal

The main issues identified from the EAG's critical appraisal are summarised in Box 1 with cross references provided to the subsections where these are discussed in more detail. Items numbered 1, 2, 3, and 5 in Box 1 were identified as key issues in Section 1.

Box 1 Summary of the main issues identified within the company's health economic model

1. The model has been populated with data from the non-Asia CPS ≥1 cohort, which is a *post hoc* combination of data from two regions, and it is unclear if data from the Western Europe/Israel/North America/Australia region would be more generalisable to England than data from the Rest of World region (**Key Issue 1** –Section 3.2.3 and 4.3.3.1)
2. PFS and OS survival curves for the SoC arm use data from the global CPS ≥1 cohort instead of data from the non-Asia CPS ≥1 cohort, despite the company claiming that data from the non-Asia region are more generalisable to England (**Key Issue 2** – Section 4.3.3.2)
3. A proportional hazards modelling approach has been used to extrapolate OS and PFS in the pembrolizumab plus SoC arm and a constant HR has been applied life-long for both OS and PFS (**Key Issue 2** – Section 4.3.3.2)
4. Duration of treatment for each component of the intervention and control arms is capped by a maximum number of cycles resulting but this does not always correspond with usage in the KEYNOTE-811 study or expected usage in clinical practice in England (Section 4.3.3.3)
5. Use of utilities based on time-to-death utilities rather than a progression-based approach (**Key Issue 3** – Section 4.3.3.4)
6. Administration costs for trastuzumab are the same when given alone or in combination with pembrolizumab (Section 4.3.3.5)
7. Subsequent therapies based on KEYNOTE-811 do not reflect current practice (Section 4.3.3.6)
8. Disease management costs for patients who are progression-free overestimate follow-up visits and exclude costs for routine staging scans (Section 4.3.3.7)
9. Disease management costs for patients with progressed disease underestimate follow-up visits and costs for routine staging scans (Section 4.3.3.8)
10. End-of-life (terminal care) costs are based on estimates from an appraisal of urothelial cancer (Section 4.3.3.9)

4.3.3.1 Generalisability of the non-Asia cohort to patients being treated in England

As previously discussed in Section 3.2.3, the company considered that data from the Asia region were less generalisable to a UK setting and therefore used data from a non-Asia cohort which was generated by combining data from two regions: Europe/Israel/North America/Australia and Rest of World. Whilst each of the regions was prospectively defined as a subgroup of interest and region was a stratification factor in the randomisation, the combination of data from two regions into a non-Asia cohort was *post hoc*. The EAG agrees with excluding the Asia cohort from the analysis, but the EAG considers that the company has not adequately justified whether patients from the Rest of World region are as generalisable to eligible patients in England as patients from the Europe/Israel/North America/Australia region. The EAG considers that the Western Europe/Israel/North America/Australia cohort may be more applicable to clinical practice in England, and as this was a pre-specified subgroup and a stratification factor it would be valid to populate the model with data exclusively from the Europe/Israel/North America/Australia in CPS ≥ 1) cohort. However, the EAG were unable to do this using the data provided by the company.

4.3.3.2 Approach to modelling OS and PFS

The EAG disagrees with the company's survival extrapolation approach for the following reasons: (i) the extrapolated curve for the intervention arm for both OS and PFS does not fit the intervention arm data from the KEYNOTE-811 trial (see Figure 3 and Figure 4); (ii) data from the global cohort were used in the extrapolation when the company has claimed that data from the non-Asia cohort are more generalisable to the UK; (iii) a constant HR was assumed for a life-time which has not been justified by the company; (iv) a HR generated from a separate Cox model was applied.

In response to clarification question B6, the company provided survival extrapolation for OS and PFS based on the non-Asia CPS≥1 subgroup and joint modelling approach (i.e., with treatment as a covariate).[19] The company determined the most plausible model for OS is a 2-knot hazard spline model and for PFS is a log-logistic model. However, the results from these survival analyses have not been applied in the updated economic model. The EAG disagrees with the use of a joint modelling approach because either a constant HR was assumed for a life-time which has not been justified by the company, or a constant acceleration factor was assumed for a life-time which has not been justified by the company.

The EAG notes that in response to clarification question B6, the company also provided the results (estimated model coefficients, AIC and BIC) from the independent modelling approach (i.e., fitting a model to each arm independently).[19] Again these results have not been applied in the updated economic model and the company did not provide its view in terms of the most plausible model for OS and PFS when using an independent modelling approach.

As described in Section 4.3.1, the company's updated model does not allow the user to select options related to the choice of cohort (global versus non-Asia) separately from options related to the choice between using a proportional hazards modelling approach or independent parametric modelling approach.

The EAG's preferred approach to modelling OS and PFS, which incorporates curves fitted to data from the non-Asia (CPS ≥1) cohort using the independent modelling approach, is described in Sections 4.4.2.2 and 4.4.2.3.

4.3.3.3 Duration of treatment for each component of the intervention and control arms

Whilst the company's base case analysis assumed that the treatment duration for each drug would be capped as described in Section 4.2.2, this was inconsistent with the treatment durations permitted in the KEYNOTE-811 study for some drugs. The company clarified (response to question A8) that a second course of pembrolizumab (up to 17 cycles) was allowed following disease progression in patients who had either stopped treatment after 35 administrations for reasons other than disease progression or toxicity (i.e. they had completed the course), or had stopped treatment after attaining a complete response.[19] However, this only occurred in 1% (3/298) patients in the CPS≥1 global cohort. As the company also stated that all these patients were in the non-Asia region, the EAG infers that this occurred in 1.5% (3/202) patients in the CPS≥1 non-Asia cohort. Also, the mean duration of treatment for patients starting a second course was 8.8 weeks, with a range of 6.4 to 10.9 weeks suggesting that most patients only received 2 or 3 doses in their second course. As the second course was only taken up by a small proportion of patients and typically lasted a short duration, the EAG is satisfied with this not contributing to the costs of pembrolizumab in the model and has therefore kept the company's assumption that the maximum duration of treatment for pembrolizumab is 35 cycles.

The duration of trastuzumab is capped in the company's base case analysis at 35 cycles, although the CS notes that there is no restriction on duration of treatment in TA208, other than for disease progression or unacceptable toxicity.[10] The CS is inconsistent in reporting whether the duration of trastuzumab was restricted in KEYNOTE-811, stating on page 154 that a maximum of 35 doses could be given in the trial, and stating on page 35 that it could be given for up to a year after the 35 doses. The EAG's clinical advisors stated that they would continue to offer trastuzumab up to disease progression or unacceptable toxicity in clinical practice, but in practice most patients stopped before reaching 35 cycles. Based on this, the EAG preferred to use the TTD KM data to determine treatment duration for trastuzumab (see Section 4.4.2.4).

The EAG notes that the number of cycles for cisplatin and oxaliplatin was based on local guidance in KEYNOTE-811, with this being 6 cycles for cisplatin and 6 to 8 cycles for oxaliplatin (clarification

response A7).[19] The company states that the majority of patients in England receive 6 cycles of oxaliplatin based on clinical advice. The EAG's clinical advisors agreed that stopping chemotherapy at six cycles was a reasonable assumption for the majority of this patient group, although one clinical advisor noted that they often started patients on 4 cycles of chemotherapy, with the option to reassess and extend to 8 cycles, rather than starting with a 6-cycle course.

The duration of capecitabine and 5-FU was not restricted in KEYNOTE-811 to the same duration as the platinum-containing agent, with both being allowed up to 1 year after the 35 cycles of either pembrolizumab or placebo had been completed. The mean duration of each chemotherapy agent was shown previously in

Table 8. The EAG's clinical advisors noted that they do not generally extend capecitabine or 5-FU beyond the duration of the platinum-containing treatment in clinical practice because extended use is not thought to improve outcomes in this indication but is associated with AEs that might require hospital admission. The company has provided a scenario analysis in which the mean number of cycles was applied for each chemotherapy agent (see clarification response B35) and a scenario in which the actual TTD curve for each treatment given in KEYNOTE-811 was applied unrestricted (see clarification response B36).[19] As the EAG's clinical advisors advised that extended use of chemotherapy is unlikely to improve outcomes, and may overestimate the costs of treatment relative to clinical practice, the EAG prefers to assume in its base case that the duration of chemotherapy is capped, as in the company's base case. However, it has explored the impact of applying the mean number of chemotherapy cycles administered in KEYNOTE-811 as a scenario analysis. It has also explored the impact of restricting the maximum duration of chemotherapy to 4 cycles to determine how sensitive the model is to duration of chemotherapy given in standard care (see Section 4.4.2.4).

4.3.3.4 Utilities based on time-to-death instead of progression

There is considerable uncertainty related to whether using a time-to-death approach for estimating utility is preferential to a progression-based approach that has historically been more widely used. The EAG comments that neither approach overcomes the main limitation that the data collected have been heavily censored, either at the point of progression, or at treatment discontinuation.

The EAG's clinical advisors disagreed with the use of a time-to-death approach. They suggest that progression symptoms and AEs are key drivers for utility, and an analysis based on time without symptoms or toxicity (TWiST) may be a better approach to use.

Patients with a time-to-death ≥360 days or 180 to 360 days are assigned utility scores of [redacted] and [redacted], respectively. These values are very similar to the general population utility value for individuals aged [redacted] years and [redacted] years respectively (estimated general population utilities are [redacted] and [redacted] in the model at these ages respectively). The model may therefore overestimate HRQoL for patients in these time-to-death categories, given that the population has advanced gastric or GOJ cancer.

The company's utility analysis was based on descriptive statistics rather than modelling the data using a mixed effects model to consider the fact that data were repeatedly measured and to adjust for covariates which may be important confounders. In response to clarification question B16, the company investigated analysing utility data using a linear mixed effects regression model for both time-to-death based and progression based approaches.[43]

For the utility analysis with the time-to-death approach, the company explored the inclusion of age, sex, grade 3+ AEs and time-to-death as fixed effect covariates and concluded that both age and sex are not statistically significant, and these variables are not included in the final model. The mean and standard error based on the time-to-death approach are presented in Table 21.

Table 21 Mean (Standard Error) of EQ-5D utilities by time-to death using linear mixed effects model (reproduced from the company's additional analysis[43])

Time-to-death (days)	Without Grade 3+ AE	During Grade 3+ AE
<30	[redacted]	[redacted]
30-180	[redacted]	[redacted]
180-360	[redacted]	[redacted]
>360	[redacted]	[redacted]

Abbreviations: AE, adverse event.

For the utility analysis with progression status, the company also explored to include age, sex, grade 3+ AE and progression status as fixed effect covariates and concluded that both age and sex are not statistically significant, and these variables not are not included in the final model. The mean and standard error based on the progression-based approach are presented in Table 22.

Table 22 Mean (Standard Error) of EQ-5D utilities by progression status using linear fixed effects model (reproduced from the company's additional analysis [43])

Health state	Without Grade 3+ AE	During Grade 3+ AE
Progression-free	[redacted]	[redacted]
Progressed disease	[redacted]	[redacted]

Abbreviations: AE, adverse event.

The EAG notes that the company's additional analyses also provide descriptive statistics for both time-to-death and progression-based approaches.[43] However, the number of patients and the estimated mean in each category are slightly different from the values presented in the CS. The EAG is unclear about the reasons for such discrepancy.

The EAG prefers to use the utility values estimated using a linear mixed effect model instead of descriptive statistics because the mixed effect modelling approach takes into account of the effect of covariates and correlations within a patient, and provides estimates with more face validity. The EAG therefore uses the data from Table 21 in its base case but has explored the use of data from Table 22 as scenario analysis.

4.3.3.5 Administration costs for trastuzumab alone versus pembrolizumab with trastuzumab

The company's base case applies the reference cost for health resource group (HRG) code SB13Z to trastuzumab whether given alone or with pembrolizumab after completion of CAPOX/XP. In response to clarification question B27, the company stated that it had explored a scenario analysis in which patients receiving trastuzumab monotherapy (i.e. those receiving SoC after completion of either CAPOX/XP) have an administration cost of SB12Z (Delivery simple chemotherapy at first attendance), whereas those receiving pembrolizumab in combination with trastuzumab after completion of CAPOX/XP continue to have SB13Z.[19] The results for this scenario are not provided by the company although it states that it resulted in a minor increase in the ICER which it described as negligible. The EAG considers that there should be some difference in administration costs for trastuzumab given alone versus trastuzumab given in combination with pembrolizumab. The EAG accepts that both are complex treatments but believes the company's scenario which applies a cost of £286.71 (HRG code SB12Z) for administering trastuzumab alone and £353.64 (HRG code SB13Z) for administering trastuzumab in combination with pembrolizumab is more appropriate when trying to capture the incremental impact of adding pembrolizumab to the existing treatment pathway (see Section 4.4.2.6).

4.3.3.6 Subsequent therapies

The EAG is concerned that the company has estimated subsequent therapies per patient completing or discontinuing study drug but has then applied the costs to those leaving the PFS state for any reason. Those leaving the PFS state due to death rather than progression are unlikely to incur costs for subsequent therapies. Equally, those stopping treatment for reasons other than progression or death, will incur subsequent therapy costs within the PFS state rather than at the time of exiting the PFS state. The company has implemented a scenario analysis, in which they increase the subsequent therapies costs to account for the fact that not all patients have either progressed or died at the time of the data cut. However, for this they have used the proportion of patients starting treatment who have progressed and not the proportion of patients completing or discontinuing treatment who have progressed.

In addition, the EAG was unable to verify the company's estimates of subsequent treatments from the drug utilisation report provided.[42] The data used in the company's model appear to analyse subsequent treatments according to the treatment combination received e.g. paclitaxel is separate from ramucirumab with paclitaxel, whereas the drug utilisation report only provides total usage of individual agent regardless of whether they were used alone or in combination.[42] The EAG understands that this means that the usage reported in the drug utilisation report and usage implemented in the model may not correlate exactly if some combinations were not frequent enough to be included in the top eight treatments. This means that some more commonly used drugs may have been excluded from the model where they were combined with other drugs in a combination that was used infrequently. This also

makes it impossible for the EAG to verify the data used in the model from the data provided in the drug utilisation report.

The EAG's clinical advisors said that there was no standard second line chemotherapy option for this patient group. One clinical advisor said that their most common treatment was docetaxel, but they were aware that some larger centres used paclitaxel, and some centres offered irinotecan. The other clinical advisor said that their preferred option was irinotecan with 5-FU and folinic acid (FOLFIRI), particularly if the patient had received a taxane containing regimen in the adjuvant or neoadjuvant setting. The EAG's clinical advisors noted that trastuzumab deruxtecan and nivolumab are not available outside of clinical trials. One clinical advisor said they would sometimes rechallenge with platinum but only if they had used less than the maximum dose at first-line and if the patient had been progression-free for more than a year. The other clinical advisor commented that they would not rechallenge with platinum if there had been progression on platinum-based chemotherapy. The EAG considered that these responses were not supportive of the company's scenario in which platinum rechallenge was used as commonly as docetaxel. The EAG also noted the conclusions of the committee in TA378 (TA of ramucirumab for treating advanced gastric cancer or GOJ adenocarcinoma previously treated with chemotherapy) which considered appropriate comparators for patients previously treated with chemotherapy.[57] In that appraisal, the committee heard from professional group submissions that taxanes are routinely used with irinotecan and FOLFIRI used less frequently. The committee concluded that both docetaxel and paclitaxel were relevant comparators with FOLFIRI and irinotecan not considered relevant because they are not in established use.[57] In addition, in TA852 (TA of trifluridine–tipiracil for treating metastatic gastric cancer or GOJ adenocarcinoma after 2 or more treatments), the committee heard that paclitaxel was generally used as second line treatment.[57] The EAG has conducted an exploratory analysis in which further treatment is equally split between docetaxel and paclitaxel and has included this assumption in their base case (see Section 4.4.2.7).

The EAG notes that the company's costing of paclitaxel does not capture the requirement for weekly intravenous infusions due to an error in the implementation. It appears that the company had intended to model second-line paclitaxel as infusions on days 1, 8, and 15 of a 28-day cycle but has instead only included one administration cost per 28-day cycle by selecting none for the resource use of doses given on days 8 and 15. Therefore, the EAG has also corrected this within its exploratory analysis (see Section 4.4.2.1). This increases the admin cost for paclitaxel from £88.41 per week to £273.27 per week.

4.3.3.7 Disease management for the progression-free state

The company stated at clarification that it was unclear how to interpret the expert opinion cited in the appraisal of trastuzumab with chemotherapy (TA208) regarding the frequency of follow-up visits during PFS.[19] It therefore considered that it was more conservative to apply both 3-weekly and 6-weekly

follow-up costs. However, based on the EAG report for TA208 (page 66) the EAG believes that the company applied follow-up visits once per 3-week cycle whilst receiving chemotherapy and once every other cycle (i.e., 6-weekly) for the remainder of their PFS regardless of whether they received trastuzumab or no further therapy.[58] Therefore, including both 3-weekly and 6-weekly follow up visits concurrently for the duration of PFS as in the company's approach is incorrect. The EAG therefore prefers to exclude the 3-weekly follow-up visits from the PFS costs, leaving the 6-weekly follow-up visits included, but has allowed for additional visits to account for 3-weekly follow-up during chemotherapy (see Section 4.4.2.8).

One clinical advisor stated that they would see patients 3-monthly after they completed chemotherapy, whilst the other stated that they would see patients 6-weekly after completing chemotherapy. Therefore, the PFS costs, which assumed continued 6-weekly throughout the period of PFS, may be overestimated relative to clinical practice in some NHS centres, although it is not possible to know the degree to which this might bias the ICER without having more comprehensive information from a range of centres.

The EAG notes that the CS does not include any routine computerised tomography (CT) scans for detecting progression whereas the EAG's clinical advisors stated that they would do 3-monthly CT scans with additional CT scans or endoscopies requested if patients had new symptoms. The EAG has included 3-monthly CT scans in its exploratory analysis (see Section 4.4.2.8)

The EAG's clinical advisors stated that they use cardiac monitoring less frequently than every 3 months in this population as this patient group are unlikely to be on trastuzumab for many years. They therefore use cardiac monitoring every 3 to 6 months during trastuzumab treatment and echocardiogram is used more than MUGA. Therefore, cardiac monitoring costs may also be overestimated in the company's base case, although the EAG expects the impact of this on the ICER to be small and therefore has not amended this.

The EAG's clinical advisors noted that specific blood tests are required to monitor patients receiving pembrolizumab to detect immune-related hepatitis, nephritis and endocrinopathies. This involves requesting blood tests for full blood counts, liver function tests and urea and electrolytes each cycle and cortisol tests every 8 weeks. Costs for these are not included in the company's model but were included for patients in the progression free health state in TA737.[51] However, the EAG notes that the costs for these blood tests are likely to be low, with costs in the progressed disease health state of £4.70 and £1.54 applied for full blood counts and biochemistry tests respectively. The EAG has therefore not explored this issue further.

4.3.3.8 Disease management for the progressed disease costs

The EAG considers that it is unlikely that resource use after disease progression should be less than resource use prior to disease progression. The EAG considers that this discrepancy is likely due to the company's assumption that only one incidence of resource use occurred per patient which potentially underestimates resource use, particularly for activities such as routine outpatient follow-up which may not occur as one-off outcomes. The company has excluded resource use types that were used in less than 5% of patients meaning that higher cost imaging tests such as magnetic resonance imaging (MRI) and positron emission tomography (PET) scans were excluded despite being reported in the UK cohort. The EAG notes that this study was restricted to patients starting second-line treatment options and therefore it does not capture typical resource use for patients who were too unwell to receive second-line chemotherapy treatment. Also, in many patients, the period of observation will overlap with the 3 months before death which the company has included as a separate end of life cost. The authors state that palliative care costs may be underestimated because they were not specifically captured in this study which was restricted to capturing "*hospitalizations, outpatient and emergency room visits, and laboratory and imaging tests performed*".[50] In addition, this paper is reporting resource use for 5 UK sites in the period 2013 to 2015 and may therefore not reflect current UK practice. The EAG's clinical advisors stated that whilst progressed patients who are not eligible for further treatment would be seen less by hospital oncologists, they may be seen more frequently in a community rather than a hospital setting to receive palliative care. The EAG notes that this community based palliative care would not have been captured in the Gómez-Ulloa *et al*. study.[50] Also, whilst the company's estimates of subsequent treatment costs capture administration costs for ongoing treatments, they do not capture the follow-up care required in patients with progressed disease who are receiving subsequent treatments and these appear to be potentially underestimated by Gómez-Ulloa *et al*. due to the assumption that each patient who reported a specific category of care only received that type of care once during the 6 months of follow-up. The EAG notes that in the appraisal of trastuzumab with chemotherapy (TA208) it was assumed that patients with progressed disease would receive supportive care at a cost of £542 per month in addition to the costs of subsequent treatment, although this was based on an estimate from a guideline for breast cancer rather than gastric cancer.[58] In addition, TA737 assumed 3-monthly consultations for patients with progressed disease which is higher than the 1.5 outpatient visits per annum applied by the company.[51] In response to clarification question B25, the company explored using the cost from TA208, (£679 when inflating the cost of £542 to current prices) and it reported that this increased the ICER by £3000.[19] Clinical advice to the EAG was that regular CT scans would not be used in patients no longer receiving active treatment and in these patients CT scans would only be required if there was an acute problem that might need intervention. However, as the model does not distinguish between progressed patients who are receiving subsequent therapy and those receiving only supportive care, it was not possible to properly reflect this advice in the model. The EAG has explored an assumption of applying 4 outpatient visits and 4 CT scans per year for progressed disease to see if

the cost-effectiveness is sensitive to assumptions regarding resource use post-progression (see Section 4.4.2.9).

4.3.3.9 End of life costs

The EAG identified the source of the terminal care cost from TA522 cited by the company.[52] It noted that while the company describes this as "*hospital care in the last 3 months of life*", according to Table 27 of the EAG report for TA522, the costs included both hospital and community care.[59] This included GP home consultations, community nursing hours, Macmillan nursing hours for terminal care at home as well as terminal care in a hospice or hospital. The majority of the resource use was based on a Marie Curie funded report which estimated the costs over 14 days of dying at home, in a hospital or hospice setting which was not specific to any type of cancer.[60] The estimate from TA522 also included costs for radiotherapy which amounted to 45% of the overall terminal care costs and these radiotherapy sessions were based on TA272 which was an appraisal of a treatment for advanced or metastatic urothelial cancer.[59] The EAG did not understand why the cost of radiotherapy sessions estimated for patients with urothelial cancer should be included in the cost of terminal care for patients with gastric or GOJ cancer. In TA737, the EAG queried the applicability of terminal care costs from TA522 to a population with gastro-oesophageal cancer patients and explored the impact of excluding radiotherapy costs.[51] The EAG in that appraisal also explored the impact of implementing the terminal care costs used in TA707 which were £8,974 over 3 months in 2019 prices.[51] However, as it was reported in the EAG report for TA737 that the ICER was not particularly sensitive to the end-of-life cost, the EAG has noted that this is an area of potential uncertainty but has not amended its base case analysis.

4.4 Exploratory analyses undertaken by the EAG

4.4.1 Overview of EAG's exploratory analyses

The methods for the EAG's exploratory analyses are provided in Section 4.4.2 with results provided in Section 4.4.3. The EAG has indicated in each case which changes are included in its base case and which are included only in its scenario analyses.

4.4.2 EAG's exploratory analyses – methods

4.4.2.1 Correction of errors in the company's model

The EAG corrected the company's implementation of administration costs for paclitaxel to match the intended weekly administration schedule. This was achieved by setting the resource use selection in N125 of the 'Subsequent Tx Costs' worksheet to "Deliver more complex parenteral chemotherapy at first attendance" so that this resource use is applied for each of the 3 doses given on days 1, 8, and 15, rather than just for the dose on day 1. This increased the administration cost for paclitaxel from £88.41 per week to £265.23 per week.

4.4.2.2 EAG's preferred survival extrapolation for OS

The EAG prefers to use an independent modelling approach for both OS and PFS, which avoids assuming either constant HR or constant acceleration factor for a life-time. The EAG uses the results from the company's independent modelling approach in response to clarification question B6 to determine its base case and scenario analysis.[19] These analyses use data from the non-Asia (CPS≤1) cohort. The EAG's choice of model was based on measures of statistical goodness-of-fit (AIC and BIC), visual inspection of the fitted curves in comparison to the KM data, the assessment of the empirical hazard function, and the assessment of long-term plausibility beyond the trial period using clinical expert opinion. The EAG notes that all models were fitted by the company (see clarification response B6).

For OS, the statistical goodness-of-fit of the standard parametric models and the spline models are summarised in Table 23. For the intervention arm, the log-logistic provides the lowest AIC and BIC scores. The log-normal, generalised gamma, one-knot hazard spline model, one-knot and three knots odds spline model all provide similar AIC scores (within three-point difference) to the log-logistic model, which indicates that these models fit the data equally well. The log-normal model provides similar BIC score to the log-logistic model, but the generalised gamma, one-knot hazard spline model, one-knot and three-knot odds spline model all have slightly worse BIC scores compared with the log-logistic model because these models are associated with a greater number of model parameters and BIC penalises more for the number of parameters in the model than AIC.

Visual assessments of the KM data and fitted models show that all spline models and standard parametric models fit the data well except for exponential model and Gompertz model (Figure 9 and Figure 10). The smoothed hazard function shows a unimodal shape (Figure 11), which indicates that the log-normal, log-logistic, generalised gamma and all spline models may be appropriate. The EAG notes that the shape of the unsmoothed hazard function is unclear as only part of the unsmoothed hazard function is presented by the company's empirical hazard plot.

The long-term predictions for the intervention arm using different models are summarised in

Table 24. The Weibull and Gompertz model provide 5 years and 10 years survival probabilities within the range provided by clinical experts (Table 25). The one-knot hazard spline model provides slightly higher 5 years survival probability (11% vs. <10%) but 10 years survival probability was within the range provided by the clinical experts (0%-1%, Table 25). The predictions from the other models are all higher than the range provided by clinical experts.

Based on the assessments above, the EAG's base case model for OS for the intervention arm is the one-knot hazard spline model with the log-logistic model (lowest AIC/BIC model) as a scenario analysis.

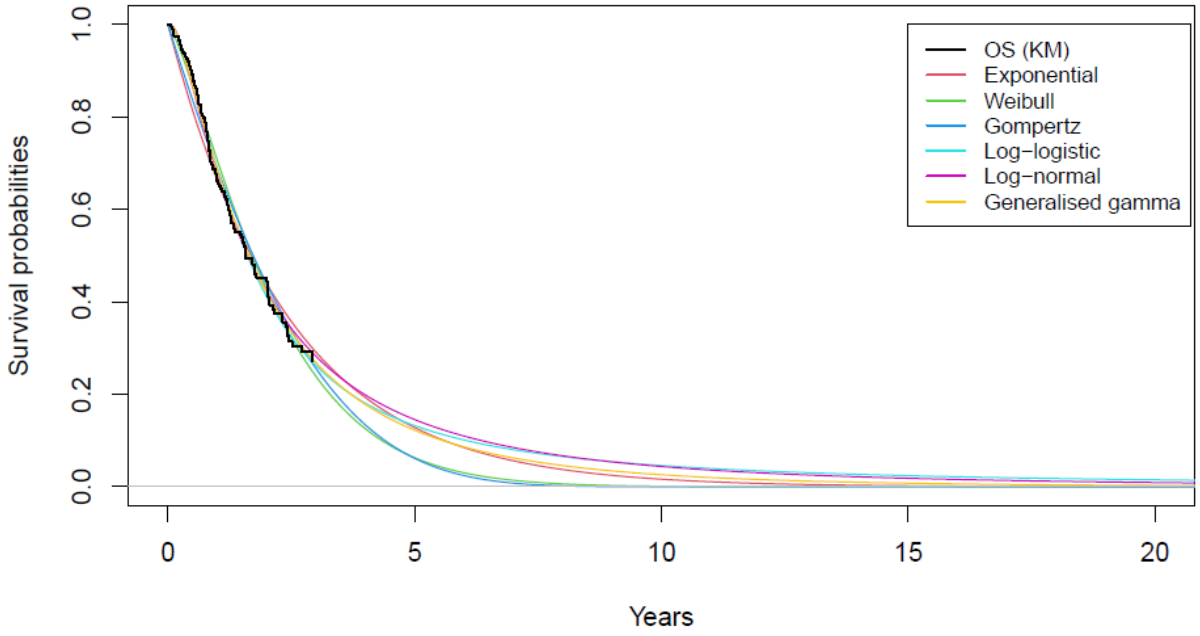
Table 23 Fit statistics of OS extrapolation in non-Asia CPS≥1 subgroup

Model	**Intervention**		**Comparator**	
	AIC	**BIC**	**AIC**	**BIC**
Standard parametric models				
Exponential	1404.19	1407.50	1549.44	1552.74
Weibull	1396.51	1403.12	1545.82	1552.41
Gompertz	1403.85	1410.46	1551.01	1557.61
Log-logistic	**1390.06**	**1396.68**	**1538.89**	**1545.49**
Log-normal	**1391.19**	**1397.81**	1546.38	1552.98
Generalised gamma	**1392.53**	1402.46	1543.73	1553.63
Spline models				
Hazard, 1 knot	**1392.47**	1402.39	1544.48	1554.37
Hazard, 2 knots	1393.41	1406.64	**1539.31**	1552.5
Hazard, 3 knots	1394.11	1410.65	**1539.52**	1556.02
Odds, 1 knot	**1391.61**	1401.54	**1540.12**	1550.01
Odds, 2 knots	1393.18	1406.41	**1540.05**	1553.24
Odds, 3 knots	**1392.86**	1410.41	**1539.35**	1555.84
Normal, 1 knot	1394.11	1402.22	**1541.81**	1551.71
Normal, 2 knots	1393.87	1406.09	**1540.70**	1553.89
Normal, 3 knots	1394.02	1410.56	**1539.01**	1555.50

Abbreviations: AIC, Akaike Information Criterion; BIC, Bayesian Information Criterion.

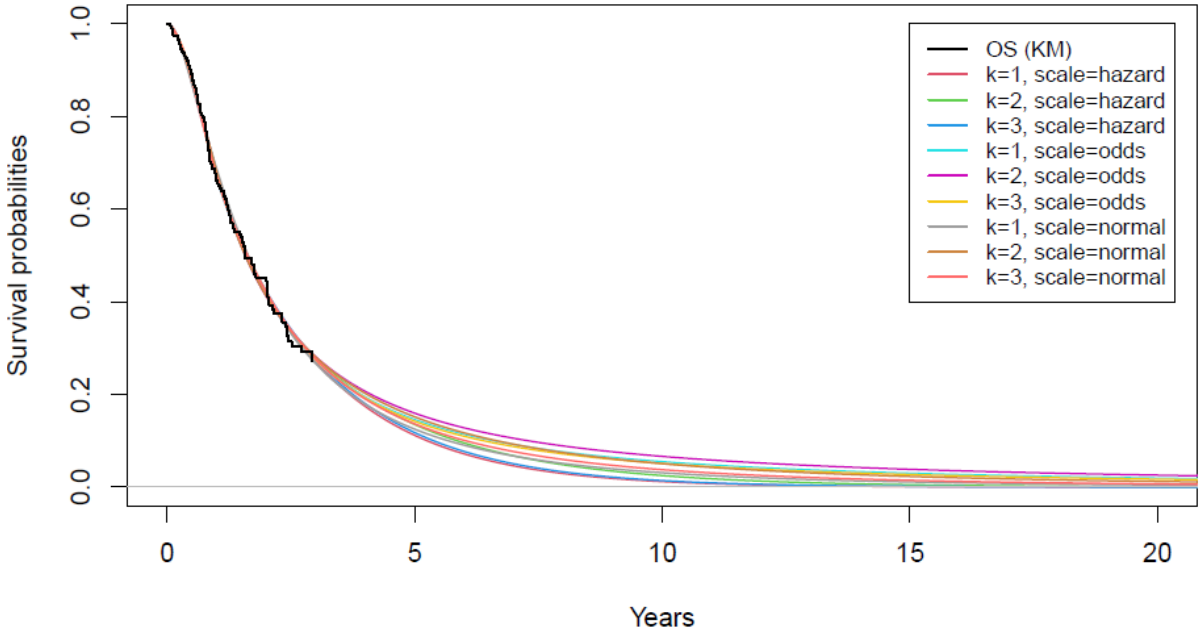
Bold: models with the lowest AIC/BIC (within three-point difference)

Figure 9 OS for the intervention arm, independently fitted standard parametric models



Abbreviations: OS, overall survival; KM, Kaplan-Meier.

Figure 10 OS for the intervention arm, independently fitted spline models



Abbreviations: OS, overall survival; KM, Kaplan-Meier.

Figure 11 Unsmoothed hazards versus smoothed hazards for OS, the intervention arm (reproduced from clarification response, Figure 8[19])

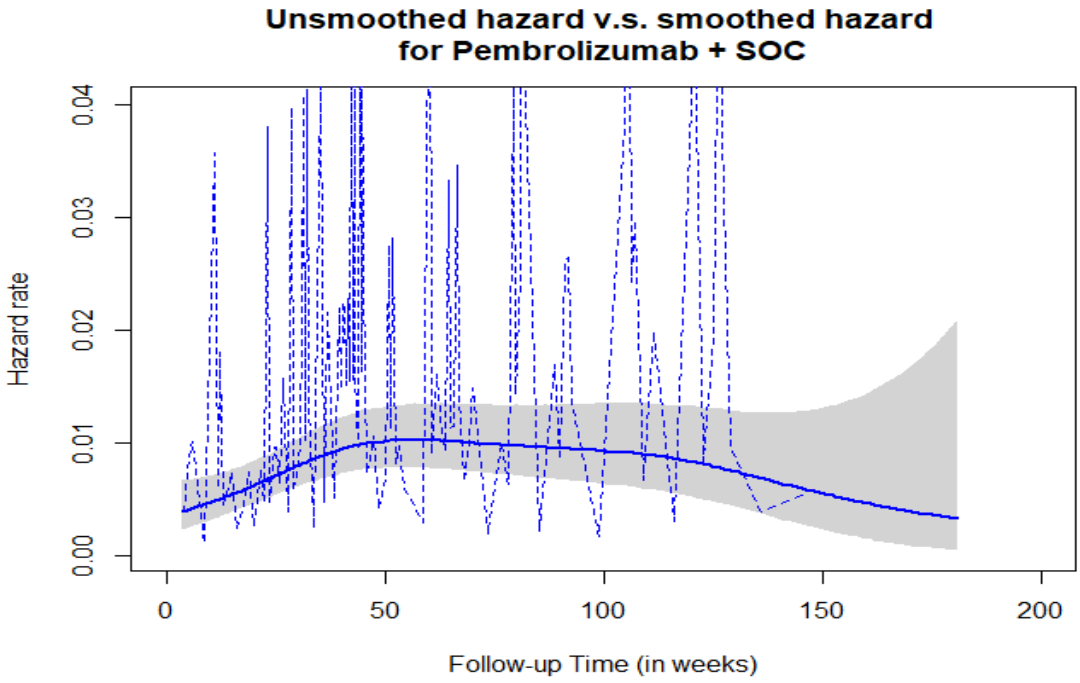


Table 24 OS predictions for the intervention in non-Asia CPS≥1 subgroup

Model	1-Year (KM estimation: 0.66)	2-Year (KM estimation: 0.44)	5-Year	10-Year	20-Year
Standard parametric models					
Exponential	0.66	0.44	0.13	0.02	0.00
Weibull	0.71	0.43	0.06	0.00	0.00
Gompertz	0.69	0.44	0.06	0.00	0.00
Log-logistic	0.69	0.41	0.13	0.05	0.02
Log-normal	0.68	0.42	0.15	0.04	0.01
Generalised gamma	0.68	0.42	0.12	0.03	0.00
Spline models					
Hazard, 1 knot	**0.68**	**0.42**	**0.11**	**0.01**	**0.00**
Hazard, 2 knots	0.68	0.41	0.14	0.02	0.00
Hazard, 3 knots	0.67	0.42	0.12	0.01	0.00
Odds, 1 knot	0.68	0.41	0.15	0.05	0.02
Odds, 2 knots	0.68	0.41	0.16	0.07	0.03
Odds, 3 knots	0.67	0.42	0.14	0.05	0.02
Normal, 1 knot	0.69	0.42	0.13	0.03	0.00
Normal, 2 knots	0.68	0.41	0.15	0.05	0.01
Normal, 3 knots	0.67	0.42	0.15	0.04	0.01

Bold: EAG's base case

Table 25 OS long-term plausibility informed by clinical expert opinion

	Expected survival probability for the intervention arm			
Timepoint	**Company's expert 1**	**Company's expert 2**	**EAG's expert 1**	**EAG's expert 2**
5 years	NA	NA	5-10%	0%
10 years	NA	NA	1%	0%
	Expected survival probability for the control arm			
5 years	5%	2-5%	≤5%	0%
10 years	2%	0-1%	0%	0%

For the control arm, the log-logistic provides the lowest AIC and BIC scores. All spline models apart from one-knot hazard spline model have similar AIC scores to the log-logistic model (within three-point difference, Table 23), indicating that these models fit the KM data equally well. No other models provide BIC scores which are within three-point difference to the log-logistic model BIC score.

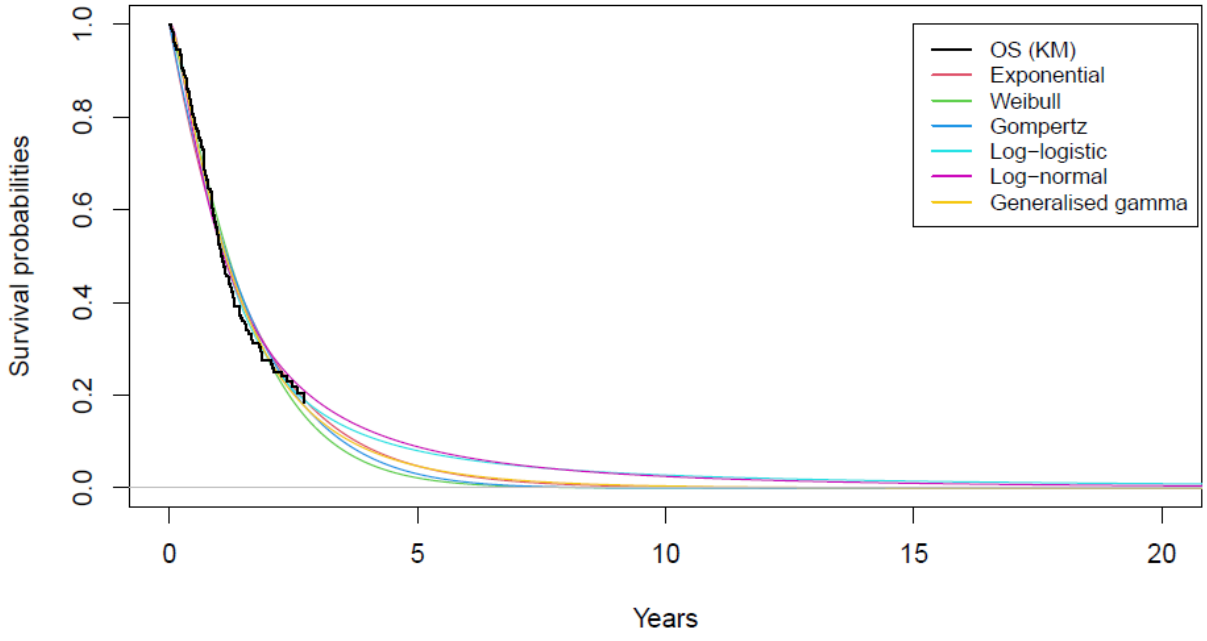
Visual assessments of the KM data versus the fitted models shows that log-logistic and all spline models provide a reasonable visual fit the observed KM data (Figure 12 and Figure 13). The smoothed hazard function shows a unimodal shape (Figure 14), which indicates that the log-normal, log-logistic, generalised gamma and all spline models may be appropriate. The EAG notes that the shape of the unsmoothed hazard function is unclear as only part of the unsmoothed hazard function is presented by the company's empirical hazard plot.

The long-term predictions for the control arm using different models are summarised in

Table 26. The exponential, Weibull, Gompertz, generalised gamma, one-knot hazard spline and one-knot normal spline models provide 5 years and 10 years survival probabilities within the range provided by clinical experts (Table 25). The predictions from the other models are all higher than the range provided by clinical experts either at 5 years or both 5 and10 years.

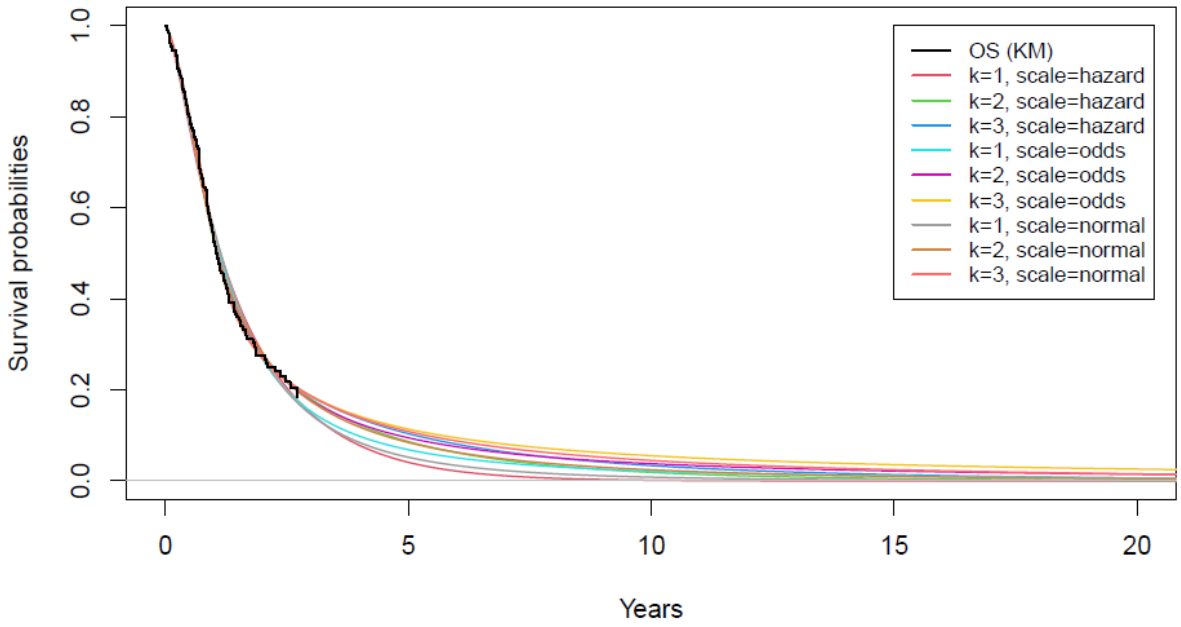
Based on the assessments above, the EAG's base case model for OS for the control arm is the one-knot normal spline model with the log-logistic model (lowest AIC/BIC model) as a scenario analysis.

Figure 12 OS for the comparator arm, independently fitted standard parametric models



Abbreviations: OS, overall survival; KM, Kaplan-Meier.

Figure 13 OS for the comparator arm, independently fitted spline models



Abbreviations: OS, overall survival; KM, Kaplan-Meier.

Figure 14 Unsmoothed hazards versus smoothed hazards for OS, the comparator arm (reproduced from clarification response, Figure 9)[19]

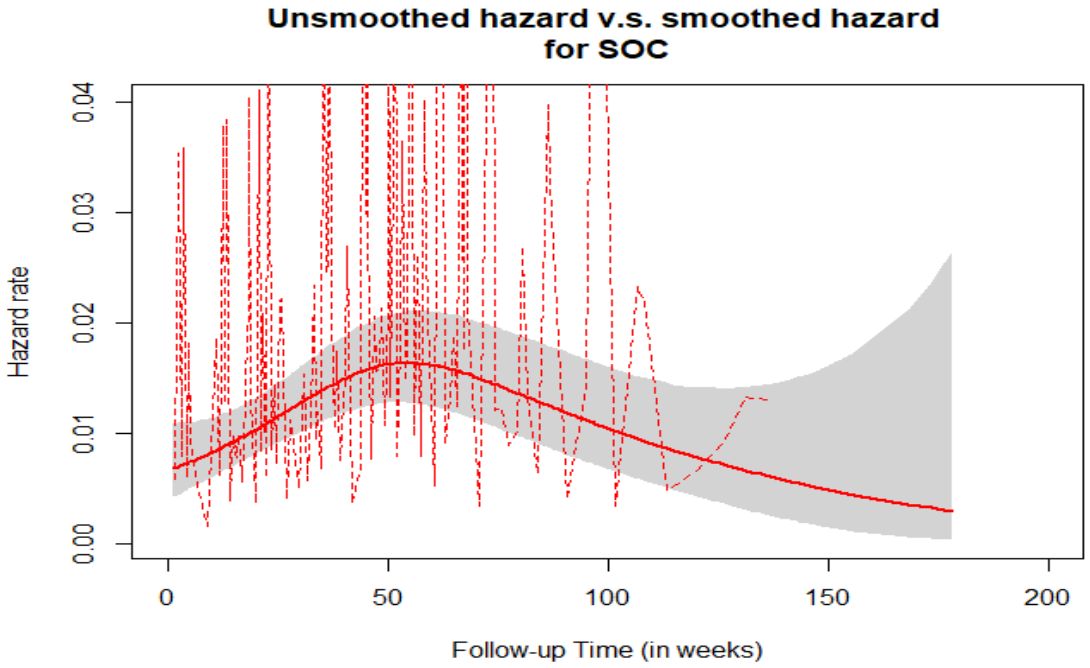


Table 26 OS predictions for the comparator in non-Asia CPS≥1 subgroup

Model	1-Year (KM estimation: 0.53)	2-Year (KM estimation: 0.28)	5-Year	10-Year	20-Year
Standard parametric models					
Exponential	0.54	0.30	0.05	0.00	0.00
Weibull	0.57	0.28	0.02	0.00	0.00
Gompertz	0.55	0.29	0.03	0.00	0.00
Log-logistic	0.54	0.28	0.08	0.03	0.01
Log-normal	0.53	0.30	0.09	0.02	0.00
Generalised gamma	0.55	0.28	0.05	0.00	0.00
Spline models					
Hazard, 1 knot	0.55	0.28	0.04	0.00	0.00
Hazard, 2 knots	0.53	0.28	0.09	0.02	0.00
Hazard, 3 knots	0.54	0.27	0.10	0.03	0.01
Odds, 1 knot	0.54	0.27	0.07	0.02	0.01
Odds, 2 knots	0.53	0.28	0.09	0.04	0.01
Odds, 3 knots	0.53	0.27	0.11	0.06	0.03
Normal, 1 knot	**0.55**	**0.27**	**0.05**	**0.01**	**0.00**
Normal, 2 knots	0.53	0.28	0.08	0.02	0.01
Normal, 3 knots	0.53	0.27	0.11	0.04	0.02

Bold: EAG's base case

4.4.2.3 EAG's preferred survival extrapolation for PFS

For PFS, the statistical goodness-of-fit of the fitted standard parametric models and spline models are summarised in Table 27. For the intervention arm, the three-knot normal spline model has the lowest AIC score. The spline models apart from the one-knot normal spline model all have similar AIC scores to the three-knot spline model (within three-point difference). In terms of BIC scores, the log-normal has the lowest BIC score of all the fitted models, and the log-logistic and one-knot odds spline model have similar BIC scores to the log-normal model (within three-point difference). Those models with similar AIC/BIC scores fit the KM data equally well.

The plots showing the intervention PFS KM curve versus the fitted curves using standard parametric models and spline models are presented in Figure 15 and Figure 16 separately. The EAG notes that there is a noticeable change in the gradient of the PFS KM curve around 1.5 years. All the standard

parametric models do not seem to fit the KM data well after around 1.5 years. The spline models fit the KM data better when compared with standard parametric models. The smoothed hazard function in Figure 17 has a unimodal shape, which indicates that the log-normal, log-logistic, generalised gamma and all spline models may be appropriate. The EAG notes that the shape of the unsmoothed hazard function is unclear as only part of the unsmoothed hazard function is presented in the company's empirical hazard plot.

The long-term predictions for the intervention arm using different models are summarised in

Table 29. All the standard parametric models, one-knot hazard spline, one-knot odds spline and one-knot normal spline models provide 5 years survival probabilities within the range provided by clinical experts (<10%,

Table 28). The exponential and Weibull models provide 10 years survival probabilities within the range provided by clinical experts (0%,

Table 28). The log-normal model provides slightly higher 10 years survival probability (1%).

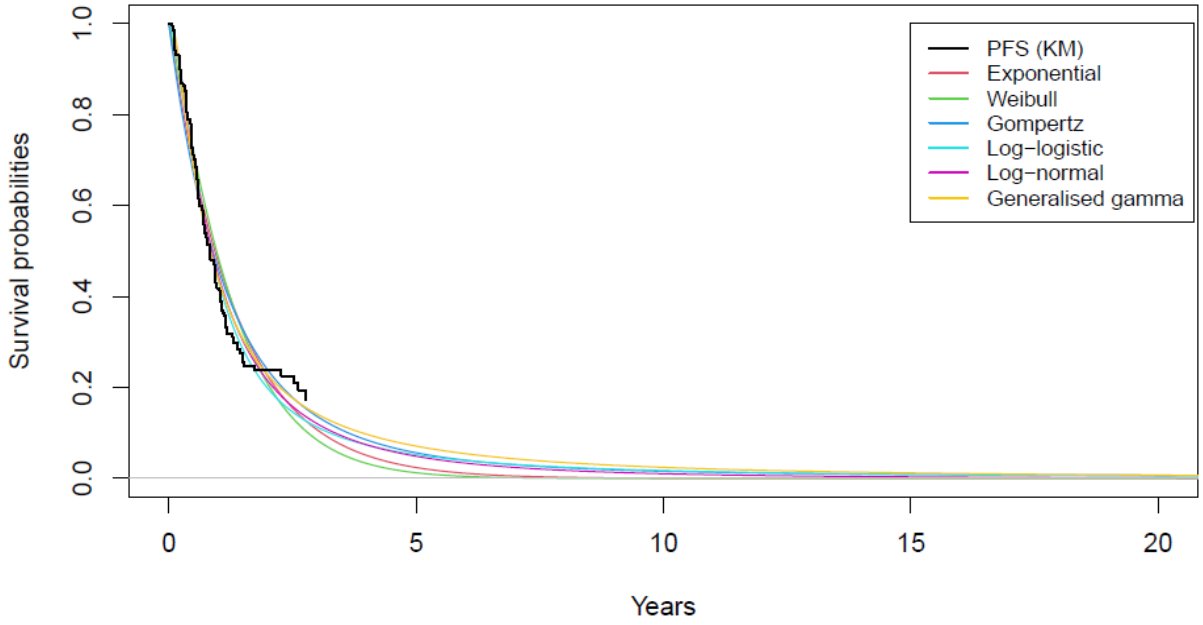
Based on the assessments above, the EAG's base case model for PFS for the intervention arm is the log-normal model with the log-logistic model as a scenario analysis.

Table 27 Fit statistics of PFS extrapolation in non-Asia CPS≥1 subgroup

Model	Intervention		Comparator	
	AIC	BIC	AIC	BIC
Standard parametric models				
Exponential	1481.71	1485.02	1491.32	1494.62
Weibull	1481.07	1487.69	1487.98	1494.58
Gompertz	1482.12	1488.73	1493.23	1499.83
Log-logistic	1458.93	**1465.55**	**1469.57**	**1476.17**
Log-normal	1458.17	**1464.79**	**1471.57**	**1478.16**
Generalised gamma	1458.50	1468.43	1473.34	1483.23
Spline models				
Hazard, 1 knot	**1457.96**	1467.88	1473.34	1483.23
Hazard, 2 knots	**1455.38**	1468.61	**1472.43**	1485.62
Hazard, 3 knots	**1456.18**	1472.72	1474.15	1490.64
Odds, 1 knot	**1457.40**	**1467.33**	**1471.56**	1481.45
Odds, 2 knots	**1456.58**	1469.82	1473.08	1486.27
Odds, 3 knots	**1456.57**	1473.11	1473.63	1490.12
Normal, 1 knot	1459.18	1469.11	1473.04	1482.94
Normal, 2 knots	**1455.66**	1468.89	**1472.43**	1485.63
Normal, 3 knots	**1455.14**	1471.68	1472.8	1489.29

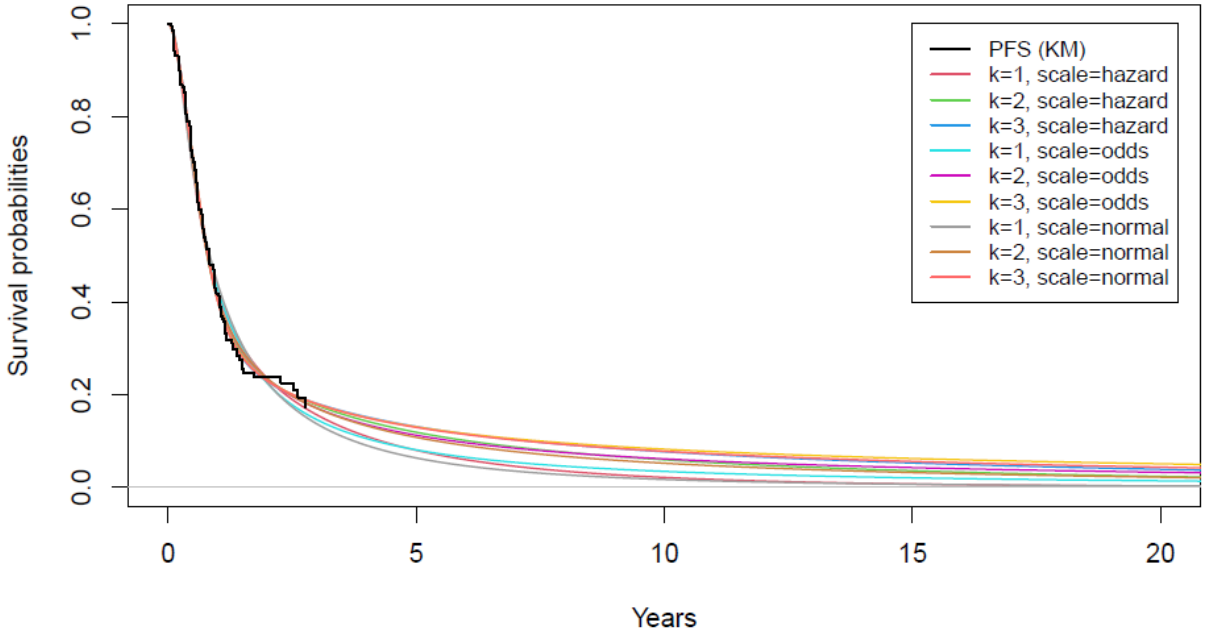
Abbreviations: AIC, Akaike Information Criterion; BIC, Bayesian Information Criterion.
Bold: models with the lowest AIC/BIC (within three-point difference)

Figure 15 PFS for the intervention arm, independently fitted standard parametric models



Abbreviations: PFS, progression-free survival; KM, Kaplan-Meier.

Figure 16 PFS for the intervention arm, independently fitted spline models



Abbreviations: PFS, progression-free survival; KM, Kaplan-Meier.

Figure 17 Unsmoothed hazards versus smoothed hazards for PFS, the intervention arm (reproduced from clarification response, Figure 18[19])

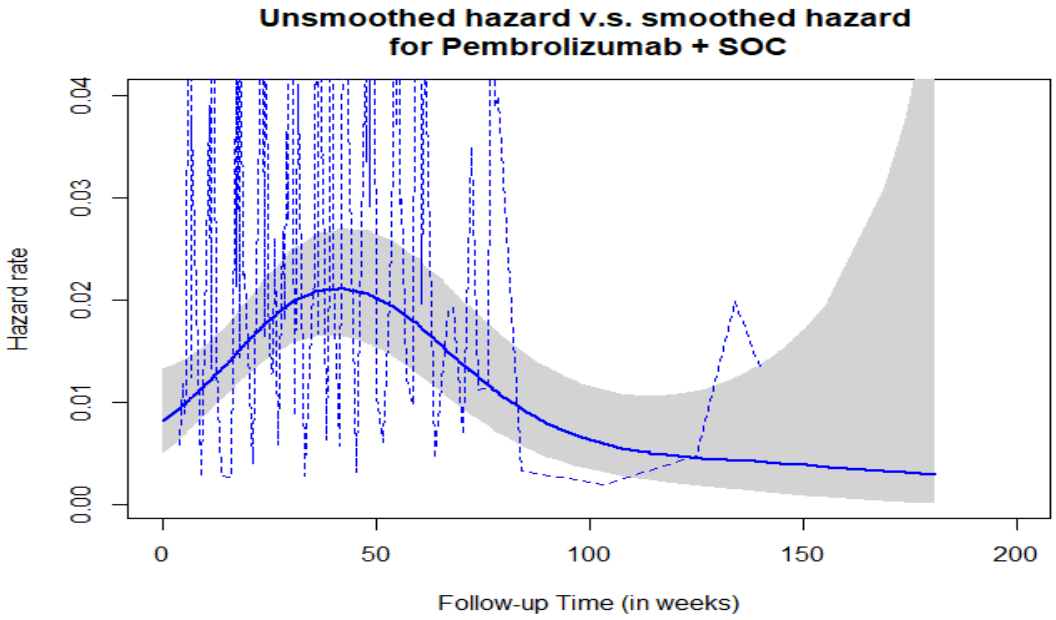


Table 28 PFS long-term plausibility informed by clinical expert opinion

	Expected survival probability for the intervention arm			
Timepoint	**Company's expert 1**	**Company's expert 2**	**EAG's expert 1**	**EAG's expert 2**
5 years	NA	NA	5-10%	0%
10 years	NA	NA	0%	0%
	Expected survival probability for the control arm			
5 years	NA	NA	0%	0%
10 years	NA	NA	0%	0%

Table 29 PFS predictions for the intervention in non-Asia CPS≥1 subgroup

Model	1-Year (KM estimation: 0.42)	2-Year (KM estimation: 0.24)	5-Year	10-Year	20-Year
Standard parametric models					
Exponential	0.47	0.22	0.02	0.00	0.00
Weibull	0.49	0.21	0.01	0.00	0.00
Gompertz	0.46	0.24	0.06	0.02	0.01
Log-logistic	0.43	0.20	0.05	0.02	0.01
Log-normal	**0.45**	**0.21**	**0.05**	**0.01**	**0.00**
Generalised gamma	0.44	0.23	0.07	0.02	0.01
Spline models					
Hazard, 1 knot	0.42	0.24	0.08	0.02	0.00
Hazard, 2 knots	0.40	0.23	0.12	0.06	0.02
Hazard, 3 knots	0.40	0.23	0.13	0.08	0.04
Odds, 1 knot	0.43	0.22	0.08	0.03	0.01
Odds, 2 knots	0.41	0.23	0.11	0.06	0.03
Odds, 3 knots	0.40	0.23	0.13	0.08	0.05
Normal, 1 knot	0.44	0.23	0.06	0.02	0.00
Normal, 2 knots	0.41	0.23	0.11	0.05	0.02
Normal, 3 knots	0.40	0.23	0.13	0.08	0.04

Bold: EAG's base case

For the control arm, the log-logistic model has the lowest AIC and BIC scores. The log-normal, two-knot hazard spline, one-knot odds spline and two-knot normal spline models have similar AIC scores to the log-logistic model (within three-point difference). The log-normal provides similar BIC score to the log-logistic model (within three-point difference).

Visual assessments of the KM data and fitted models show that all spline models and standard parametric models fit the data well except for Weibull and Gompertz models (

Figure 18 and

Figure 19). The smoothed hazard function shows a unimodal shape (see

Figure 20), which indicates that the log-normal, log-logistic, generalised gamma and all spline models may be appropriate. The EAG notes that the shape of the unsmoothed hazard function is unclear as only part of the unsmoothed hazard function is presented in the company's empirical hazard plot.

The long-term predictions for the comparator PFS using different models are summarised in

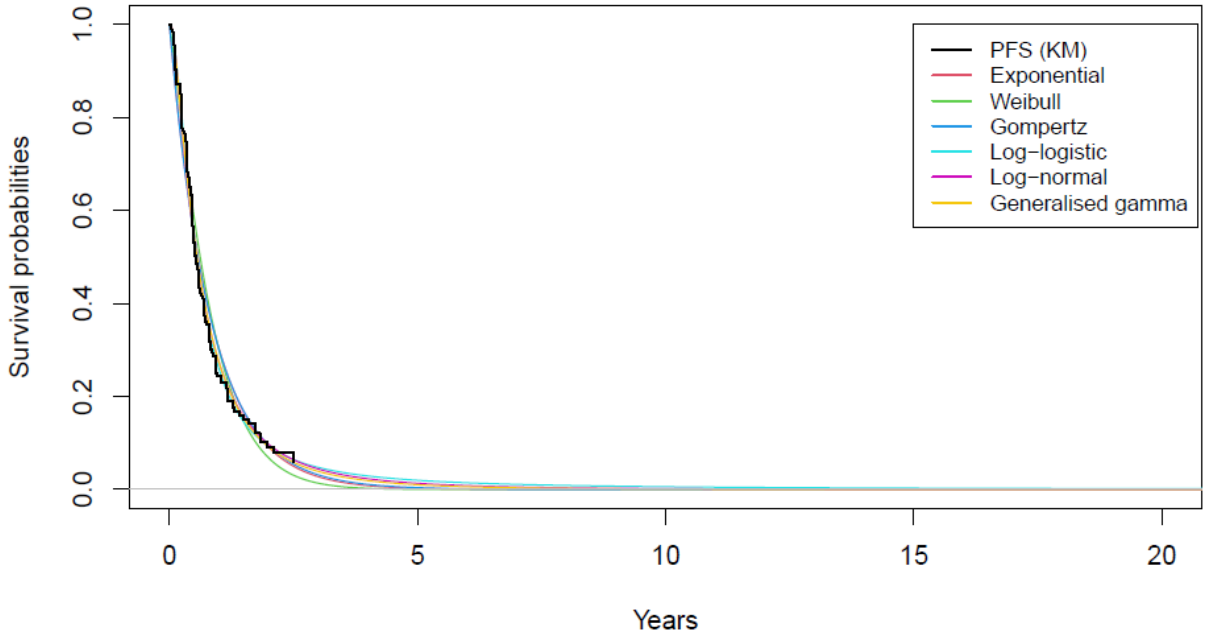
Table 30. The exponential, Weibull and Gompertz models provide 5 years and 10 years survival probabilities within the range provided by clinical experts (0%,

Table 28). The log-normal, generalised gamma, one-knot hazard spline, one-knot normal spline models provide slightly higher 5 years survival probability (1%), but 10 years survival probability was within the range provided by the clinical experts (0%,

Table 28). The predictions from the other models are all higher than the range provided by clinical experts.

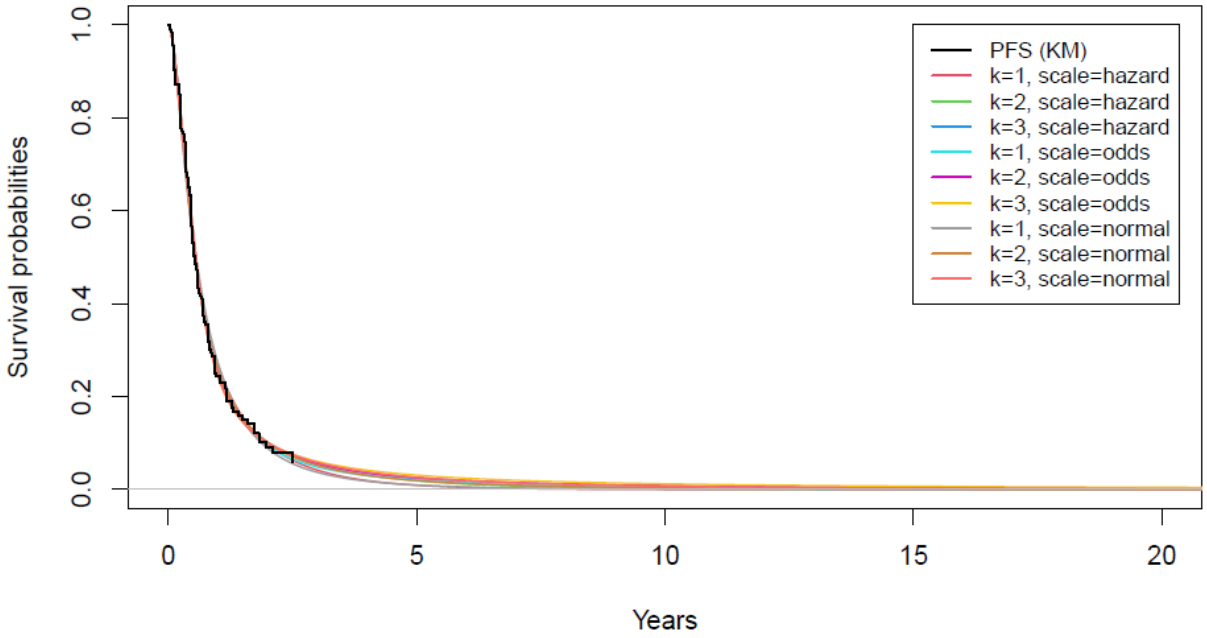
Based on the assessments above, the EAG's base case model for PFS for the comparator arm is the log-normal model with the log-logistic model (lowest AIC/BIC model) as a scenario analysis.

Figure 18 PFS for the comparator arm, independently fitted standard parametric models



Abbreviations: PFS, progression-free survival; KM, Kaplan-Meier.

Figure 19 PFS for the comparator arm, independently fitted spline models



Abbreviations: PFS, progression-free survival; KM, Kaplan-Meier.

Figure 20 Unsmoothed hazards versus smoothed hazards for PFS, the comparator arm (reproduced from clarification response, Figure 19[19])

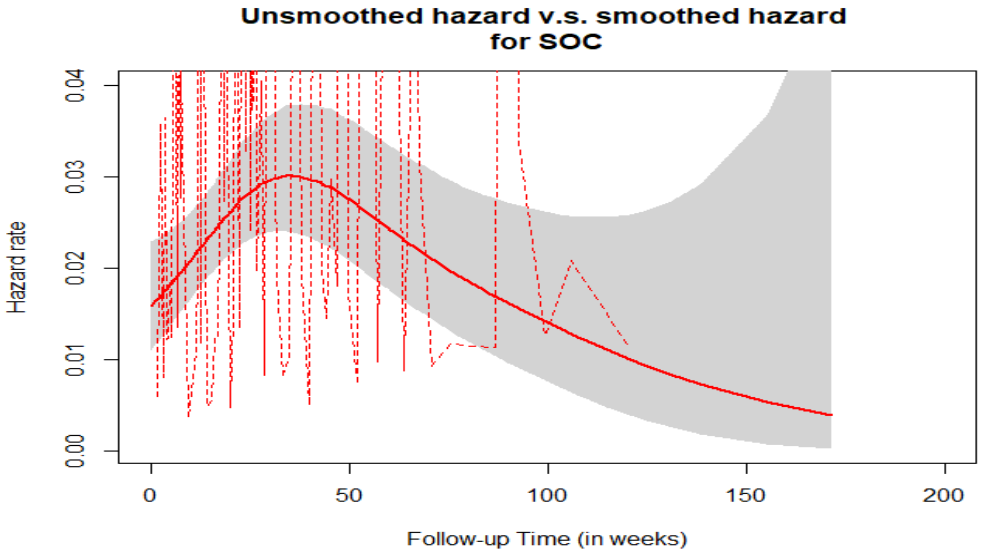


Table 30 PFS predictions for the comparator in non-Asia CPS≥1 subgroup

Model	1-Year (KM estimation: 0.24)	2-Year (KM estimation: 0.09)	5-Year	10-Year	20-Year
Standard parametric models					
Exponential	0.30	0.09	0.00	0.00	0.00
Weibull	0.30	0.07	0.00	0.00	0.00
Gompertz	0.30	0.10	0.00	0.00	0.00
Log-logistic	0.26	0.09	0.02	0.01	0.00
Log-normal	**0.27**	**0.10**	**0.01**	**0.00**	**0.00**
Generalised gamma	0.27	0.09	0.01	0.00	0.00
Spline models					
Hazard, 1 knot	0.27	0.10	0.01	0.00	0.00
Hazard, 2 knots	0.25	0.10	0.02	0.00	0.00
Hazard, 3 knots	0.25	0.10	0.02	0.00	0.00
Odds, 1 knot	0.26	0.09	0.02	0.01	0.00
Odds, 2 knots	0.26	0.10	0.03	0.01	0.00
Odds, 3 knots	0.24	0.10	0.03	0.01	0.00
Normal, 1 knot	0.27	0.09	0.01	0.00	0.00
Normal, 2 knots	0.26	0.10	0.02	0.00	0.00
Normal, 3 knots	0.24	0.10	0.03	0.01	0.00

Bold: EAG's base case

The KM curves and the extrapolated curves informed by the most plausible model as well as the model used in the scenario analysis are presented in

Figure 21 and Figure 22 below.

Figure 21 EAG's choices of extrapolations for OS in the non-Asia CPS≥1 subgroup

Figure 22 EAG's choices of extrapolations for PFS in the non-Asia CPS≥1 subgroup

4.4.2.4 Duration of treatments

The EAG has applied the company's assumption that the maximum number of cycles of chemotherapy is 6 cycles in its base case analysis. However, it has explored scenario analyses using the average number of cycles administered and using a maximum number of 4 cycles. The EAG has restricted the maximum number of cycles for pembrolizumab in its base case to 35 cycles, but it has allowed trastuzumab to be used according to the TTD curve from KEYNOTE-811 without any limit applied.

4.4.2.5 Utilities using a time-to-death approach

The EAG has included the company's time-to-death approach for utilities in its base case but has explored the impact of using the progression-based approach in its scenario analysis. The EAG prefers to use the company's utility analyses which used a linear mixed effect regression (provided in response to clarification question B16). The EAG therefore applied the data in Table 21 for its preferred base case and has explored using the data in Table 22 in its scenario analysis.

4.4.2.6 Resource use – administration costs

The EAG has incorporated the scenario analysis in which a lower HRG cost has been applied for administering trastuzumab alone (£286.71 for HRG code SB12Z)) versus trastuzumab with pembrolizumab (£353.64 for HRG code SB13Z) in the period after the chemotherapy element of the treatment has been completed. This is in addition to the correction described in Section 4.4.2.1 for the administration costs for paclitaxel as a subsequent therapy. This change has been included in the EAG's preferred base case.

4.4.2.7 Subsequent therapies

In any scenario in which the mix of subsequent therapies is based on KEYNOTE-811 (EAG exploratory analysis 10), the EAG has recalculated the proportions receiving subsequent therapies so they are estimated as a proportion of the progressed patients rather than as a proportion of those completing or discontinuing first-line therapy.

However, in the EAG's preferred base case it has assumed that 50% of progressed patients receive subsequent treatment and that subsequent treatment consists of either paclitaxel or docetaxel in equal proportions. This means that in the EAG's base case the distribution of subsequent treatment is not based on treatments received in KEYNOTE-811.

The EAG has also amended the model so that subsequent treatment costs are only applied to the proportion of patients leaving the progression-free state whose PFS event was progression rather than death. This is in keeping with the subsequent treatment costs having been calculated per progressed patient. This is applied both when using subsequent treatments based on KEYNOTE-811 and when using the EAG's assumption that subsequent treatment consists only of taxanes (EAG exploratory analyses 6 and 10).

4.4.2.8 Disease management for progression-free state

The EAG prefers to assume 3 weekly follow-up visits during doublet chemotherapy, with 6 weekly follow-up for the remainder of PFS to align with its understanding of what was modelled in TA208 (see Section 4.3.3.7). This was achieved by excluding the 3-weekly follow-up visits implemented by the

company for the duration of PFS, keeping the 6-weekly follow-up visits implemented by the company for the duration of PFS, and adding additional follow-up visits during the doublet chemotherapy phase of treatment to achieve an average of 3-weekly visits during the first 18 weeks.

The EAG has not updated the disease management costs for the progression-free health state to account for additional blood tests required for patients receiving pembrolizumab with trastuzumab relative to trastuzumab alone or to explore a different frequency for cardiac monitoring. This is because any changes to the ICER based on these are expected to be small. The EAG has included costs for 4 CT staging scans per year in the progression-free state.

4.4.2.9 Disease management for progressed-disease state

The EAG was concerned that the company's estimate of resource use in the progressed-disease state did not account for the frequency of activities that might occur more than once in the follow-up period (see Section 4.3.3.8). The EAG has therefore adjusted the resource use for the progressed-disease state to include 4 outpatient visits and 4 CT scans per year to determine the sensitivity of the model to changes in the costs for the progressed-disease state.

4.4.3 Results of the EAG's exploratory analyses

The EAG's exploratory analyses showing the impact of making individual changes to the company's base case model are provided in Table 31. The ICERs discussed in the following text are those generated when applying no QALY weighting, but the results when applying a QALY weighting of 1.2 are provided in Table 31 for reference. The exploratory analysis that has the most significant impact on the ICER is implementing the EAG's preferred survival extrapolation for OS, which increases the ICER from £[redacted] to £[redacted] per QALY. This change is driven by the EAG's preference for using the OS data from the non-Asia region, as the OS is lower when using data from the non-Asia region (see Figure 3). Implementing the EAG's preferred approach to modelling PFS does not have a large impact on the ICER because utilities are based on time-to-death and not progression-status. Therefore, changing PFS only affects costs and in this case it reduces the incremental cost resulting in a reduction in the ICER from £[redacted] to £[redacted] per QALY. Reducing the frequency of follow-up visits in the PFS and including 3 monthly CT scans also reduced the ICER bringing it down to [redacted] per QALY. Conversely increasing the frequency of outpatient visits and including 3 monthly CT scans in the progressed disease state increased the ICER to £[redacted] per QALY. Assuming that subsequent treatment is only received by 50% and consists only of taxanes reduced the ICER to £[redacted] per QALY. The scenarios which allowed trastuzumab treatment to extend beyond 35 cycles and allowed for a lower cost when it is given alone rather than combined with pembrolizumab both increased the ICER marginally. Implementing the time-to-death utilities from the linear mixed effects regression marginally increased the ICER to £[redacted] per QALY. The EAG's adjustment to the calculation of

subsequent therapies using the data from KEYNOTE-811 did not have a substantial impact on the ICER and was not included in its base case because the EAG preferred to assume that subsequent treatment consists of taxanes in its base case. The ICER for the EAG's preferred base case, which combined EAG's exploratory analyses 1 to 9, was substantially increased at £[redacted] per QALY, mainly due to the impact of the EAG's preferred OS extrapolation. The probabilistic ICER for the EAG's preferred base case was £[redacted] per QALY. Pembrolizumab with SoC had an ICER under £30,000 per QALY in [redacted]% of PSA runs both when using a QALY weight of 1.0 and when using a QALY weight of 1.2.

Table 31 EAG's exploratory analyses

Option	LYs	QALYs	Costs	Incremental		ICER (QALY weight x1)	ICER (QALY weight x1.2)
				QALYs	Costs		
Company base case – post-clarification (Deterministic)							
SoC*	3.03	[redacted]	[redacted]	-	-		
Intervention**	4.94	[redacted]	[redacted]	[redacted]	[redacted]	*******	*******
EAG exploratory analysis 1: correcting programming and implementation errors in the company's economic model							
SoC*	3.03	[redacted]	[redacted]	-	-	-	-
Intervention**	4.94	[redacted]	[redacted]	[redacted]	[redacted]	*******	*******
EAG exploratory analysis 2: Using the EAG's preferred survival extrapolation for OS							
SoC*	1.59	[redacted]	[redacted]	-	-	-	-
Intervention**	2.17	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG exploratory analysis 3: Using the EAG's preferred survival extrapolation for PFS							
SoC*	3.03	[redacted]	[redacted]	-	-	-	-
Intervention**	4.94	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG exploratory analysis 4: Removing the cap for TTD of trastuzumab							
SoC*	3.03	[redacted]	[redacted]	-	-	-	-
Intervention**	4.94	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG exploratory analysis 5: Applying lower administration costs for trastuzumab when administered without pembrolizumab							
SoC*	3.03	[redacted]	[redacted]	-	-	-	-
Intervention**	4.94	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG exploratory analysis 6: Assuming subsequent therapy to include only taxanes and applying that to only a proportion of PFS events who get progressed (25% get paclitaxel and 25% get docetaxel)							
SoC*	3.03	[redacted]	[redacted]	-	-		
Intervention**	4.94	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG exploratory analysis 7: Limiting outpatient visits to 6 weekly after chemotherapy and adding CT scans 4 times per annum for patients on PFS							

Option	LYs	QALYs	Costs	Incremental		ICER (QALY weight x1)	ICER (QALY weight x1.2)
				QALYs	Costs		
SoC*	3.03	[redacted]	[redacted]	-	-	-	-
Intervention**	4.94	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG exploratory analysis 8: Increasing outpatient visits and CT scans to 4 times per annum for patients with progressed disease							
SoC*	3.03	[redacted]	[redacted]	-	-	-	-
Intervention**	4.94	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG exploratory analysis 9: Time-to-death utilities estimated using a linear mixed effects model							
SoC*	3.03	[redacted]	[redacted]	-	-	-	-
Intervention**	4.94	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG exploratory analysis 10: Subsequent therapies are estimated as a proportion of those progressing using KEYNOTE-811 and only applied to patient leaving PFS due to progression†							
SoC*	3.03	[redacted]	[redacted]	-	-		
Intervention**	4.94	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG base case applying analyses 1-9 (Deterministic)							
SoC*	1.59	[redacted]	[redacted]	-	-	-	-
Intervention**	2.17	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG base case applying analyses 1-9 (Probabilistic)							
SoC*	1.61	[redacted]	[redacted]	-	-	-	-
Intervention**	2.21	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

CT – computerised tomography; EAG – external assessment group, ICER – incremental cost-effectiveness ratio, OS – overall survival, LYs - life-years; PD – progressed disease, PFS – progression-free survival, QALYs- quality-adjusted life-years; TTD – time to treatment discontinuation

** SoC: Trastuzumab plus chemotherapy*

*** Intervention: Pembrolizumab with SoC*

† Not included in EAG base case because the EAG prefers subsequent therapies as described in exploratory analysis 6

The EAG has also conducted deterministic scenario analyses, shown in Table 32, using its preferred base case scenario as the starting point. The probabilistic model was not run for the scenario analyses as the results for the EAG's base case suggest that the deterministic ICER provides a close estimate of the expected probabilistic ICER. The ICER in these scenarios ranged from £[redacted] per QALY when using the log-logistic extrapolation for OS and PFS (still fitted separately to each arm of the non-Asia cohort) to £[redacted] per QALY when basing utilities on progression status rather than time-to-death. The scenario analyses also suggest that the ICER is not particularly sensitive to the choice of chemotherapy agent (scenario 5) or the duration of chemotherapy treatment (scenarios 2 and 3), although all of these scenarios alter only the cost of treatment and have no impact of clinical outcomes which are assumed to remain as observed in KEYNOTE-811.

Table 32 EAG's scenario analyses

Option	LYs	QALYs	Costs	Incremental		ICER (QALY weight x1)	ICER (QALY weight x1.2)
				QALYs	Costs		
EAG base case (Deterministic)							
SoC*	1.59	[redacted]	[redacted]	-	-	-	-
Intervention**	2.17	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG scenario 1 (Assuming a log-logistic curve for OS and PFS extrapolations)							
SoC*	1.84	[redacted]	[redacted]	-	-		
Intervention**	2.50	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG scenario 2 (Using restricted mean duration to estimate costs for first-line chemotherapy)							
SoC*	1.59	[redacted]	[redacted]	-	-	-	-
Intervention**	2.17	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG scenario 3 (Reducing the cap applied to TTD of first-line chemotherapy to 4 cycles)							
SoC*	1.59	[redacted]	[redacted]	-	-		
Intervention**	2.17	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG scenario 4 (Using utility values based on progression status)							
SoC*	1.59	[redacted]	[redacted]	-	-		
Intervention**	2.17	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
EAG scenario 5 (Assuming 100% of doublet chemotherapy is with XP)							
SoC*	1.59	[redacted]	[redacted]	-	-		
Intervention**	2.17	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

EAG – external assessment group; ICER – incremental cost-effectiveness ratio; LYs - life-years, OS – overall survival, PFS – progression-free survival, QALYs- quality-adjusted life-years, TTD – time to treatment discontinuation; XP - cisplatin with capecitabine

** SoC: Trastuzumab plus chemotherapy*

*** Intervention: Pembrolizumab with SoC*

4.4.4 The EAG's estimate of the ICER

The EAG's exploratory analyses demonstrate that the ICER is highly sensitive to whether the data from the global cohort or the non-Asia cohort are used to estimate OS. The EAG's preferred approach of using parametric survival curves fitted separately to each arm of the non-Asia cohort provides a much higher ICER than the company's base case analysis. However, the EAG's preferred ICER was fairly robust to the choice of parametric curve when considering only curves fitted to the non-Asia cohort. The EAG considers that its base case ICER is somewhat uncertain due to uncertainty regarding the most appropriate method of capturing changes in utility over time as modelling utility based on progression status increased the ICER compared to the company's approach of using time-to-death.

5 SEVERITY MODIFIER

The company has presented an estimate of the proportionate and absolute QALY shortfall based on a comparison of the discounted QALYs generated in the control arm of the model (i.e., for trastuzumab with chemotherapy) and the discounted QALY expected for members of the general population who do not have gastric or GOJ cancer but otherwise have the same starting characteristics (i.e., age 60 and 21% females). Based on these data, the appropriate QALY multiplier would be x1 (see Table 33) because the proportional shortfall is [redacted] and the absolute short fall is [redacted]. The company originally stated that the Schneider *et al*. tool had been used to estimate the QALYs in the general population (see CS, page 149),[1] but it later stated, in response to clarification question B37, that this was not in fact the case and the model had been used instead to generate these estimates.[19] However, the company stated that they agreed with the principle of the Schneider *et al*. tool being used to provide consistency across appraisals. The EAG notes that when using the Schneider *et al*. tool, with the utility set described in the tool as being the reference case (Measurement and Valuation of Health [MVH] value set + Health Survey for England [HSE] 2014 Adjusted Limited Dependent Variable Mixture Model [ALDVMM] model, Hernandez Alava *et al*.), the discounted QALYs for the general population would be 12.62 rather than the 12.28 estimate provided by the company's model.[61] Combining this with the discounted QALY from the trastuzumab with chemotherapy arm estimated by in the company's base case model results in the same QALY multiplier as estimated by the company.

The company argues that the estimates of QALY in the SoC arm of the model are unrealistic because they are based on the whole CPS>=1 population which includes patients from the Asia region. It notes that the non-Asia region has shorter OS than the Asia region in KEYNOTE-811 and consider it implausible for the proportional QALY loss to be less than 85% for patients treated in current NHS practice. To address this concern, it prefers to use the QALYs estimated for trastuzumab with chemotherapy from TA208 which they report as being 0.980. When using this estimate of the QALYs under current practice, the appropriate QALY multiplier would be 1.2 (see Table 33) because the proportionate short fall is between 0.85 and 0.95. However, the EAG notes that the QALYs for the HER2-positive (IHC3+) subgroup range from 1.089 to 1.194 depending on whether trastuzumab is combined with XP or FP (ERG addendum dated 5th August 2010, Table 3c).[62] Whilst both of these provide smaller proportional QALY shortfalls than the company' estimate, they do still provide a figure compatible with a 1.2 QALY multiplier.

Overall, the EAG considers that a 1.2 multiplier is supported by the evidence if it is accepted that the data from the Asia region are not generalisable to the UK, whereas data from the other two regions are generalisable to the UK. This seems reasonable given the company's explanation that outcomes are expected to be better in the Asia region due to the widespread implementation of screening.

The EAG notes that both the ICER and the severity modifier are dependent on whether the data for OS from the non-Asia region are more generalisable to the population likely to receive treatment in England than the data from the global cohort. The EAG considers that it would be inconsistent to use the company's base case ICER, which is based on OS data from the global cohort including patients from the Asia region, and to then apply a severity modifier that has been calculated assuming that only data from the non-Asia region are applicable to patients receiving treatment in England

Table 33 Severity modifier calculations for various company and EAG scenarios

Analysis	Lifetime expected QALYs for the general population	Lifetime expected QALYs under current SoC	Absolute QALY shortfall	Proportional QALY shortfall	QALY weight
Company - modelled comparator arm and modelled general population*	12.277	[redacted]	[redacted]	[redacted]	1.0
Company - comparator arm from TA208 and modelled general population*	12.277	0.980	11.297	92.0%	1.2
EAG - modelled comparator arm for company base case and Schneider tool for general population	12.62	[redacted]	[redacted]	[redacted]	1.0
EAG - modelled comparator arm for EAG base case and Schneider tool for general population.	12.62	[redacted]	[redacted]	[redacted]	1.2
EAG - TA208 for comparator arm (trastuzumab +FP) and Schneider *et al.* tool for general population	12.62	1.089	11.53	91.4%	1.2
EAG - TA208 for comparator arm (trastuzumab +XP) and Schneider *et al.* tool for general population	12.62	1.194	11.43	90.6%	1.2

**CS, Table 60.*[1]

***The EAG notes that this figure appears to be based on a third comparator arm included in the model and not the trastuzumab with chemotherapy arm, however, the difference in absolute discounted QALYs between this estimate and the one for the trastuzumab with chemotherapy arm is -0.003 meaning that this error is unlikely to alter the conclusion regarding the appropriate QALY multiplier*

6 OVERALL CONCLUSIONS

The clinical evidence for pembrolizumab was based on one ongoing RCT, KEYNOTE-811. The EAG believes that no RCTs of pembrolizumab meeting the inclusion criteria of the NICE final scope have been missed. KEYNOTE-811 randomised 698 patients to either pembrolizumab in combination with trastuzumab and chemotherapy (CAPOX or FP), or placebo in combination with trastuzumab and chemotherapy (CAPOX or FP). According to clinical advice, patients in KEYNOTE-811 RCT were younger than would be seen in clinical practice in England but were generally representative in terms of primary location of disease at diagnosis, and the mix of locally advanced versus metastatic disease. The non-Asia cohort had a higher proportion of white patients than would be seen in clinical practice in England. Age may influence effectiveness, as patients under 65 years appeared to have more favourable treatment effect toward pembrolizumab for PFS and OS than older patients, however there is uncertainty in this as KEYNOTE-811 was not powered for subgroups.

The CS focused on the subgroup of patients with PD-L1 CPS ≥ 1 (in line with the anticipated marketing authorisation). For CPS≥1 participants, Grade ≥3 AEs were experienced by 73.2% of the pembrolizumab treated patients, and 65.1% of the comparator treated patients. Within the CPS ≥ 1 subgroup, the CS effectiveness outcomes concentrated on region, reporting the combined two subgroup regions of Western Europe/Israel/North America/Australia, and Rest of World (including South America); that is excluding the Asia region. The KEYNOTE-811 RCT was well-designed to give a low risk of bias. There is some concern that a *post hoc* analysis was used, that is combining West Europe/Israel/North America/Australia and Rest of the world, although it is noted that randomisation was preserved within each region.

At interim analysis 2 (IA2), for CPS≥1 participants excluding the Asia region, median OS for the pembrolizumab group (n=202) was 18.8 months (95%CI 15.5, 24.3), and for the comparator group (n=200) median OS was 12.6 months (95%CI 11.1, 14.9). The HR for OS significantly favoured the pembrolizumab group, HR 0.67 (95% CI 0.52, 0.85, p=0.0006). If considering region subgroups separately, OS for CPS≥1 participants, Western Europe/Israel/North America/Australia had a HR of 0.81 (95%CI 0.57, 1.15) whilst Rest of the World had a HR of 0.57 (95%CI 0.40, 0.80).

The CS provides an analysis of the cost-effectiveness of pembrolizumab plus SoC against SoC alone, where SoC is assumed to comprise of trastuzumab with chemotherapy. The addition of pembrolizumab to SoC is estimated to increase lifetime costs, largely through the increase in drug acquisition costs, but also through increased time spent in the progression-free health state. The addition of pembrolizumab to SoC is estimated to increase OS, resulting in additional time spent in the health state where time-to-

death is more than 1 year, which is associated with a greater health utility than time spent in health states with lower expected survival. This combination of increased survival and additional time spent in a state with higher utility results in an expected QALY gain for pembrolizumab plus SoC compared to SoC alone. The company's base case analysis provides a deterministic ICER of £[redacted] per QALY and a probabilistic ICER of £[redacted] per QALY when no QALY weighting is applied. The company's base case ICER when applying a QALY weighting of 1.2 is £[redacted] for the deterministic analysis and £[redacted] for the probabilistic analysis.

The EAG's primary concern regarding the company's economic analysis relates to the modelling of OS and PFS. The company claims that data from the Asia region of the KEYNOTE-811 study are less generalisable to clinical practice in England due to the widespread use of gastric cancer screening in Asia region countries which is not routinely offered in England. Therefore, the majority of the company's model inputs are informed by data from the non-Asia (CPS≥1) cohort. However, the company's approach to modelling OS and PFS uses data from the global (CPS≥1) cohort (including data from the Asia region) to model OS and PFS in the SoC arm. HRs from the non-Asia (CPS≥1) cohort are then applied to estimate OS and PFS for the pembrolizumab plus SoC arm. The EAG prefers to use parametric OS and PFS curves fitted separately to data from both the intervention and comparator arms of the non-Asia (CPS≥1) cohort. The EAG's exploratory analyses demonstrate that its alternative approach to modelling OS has a substantial impact on the ICERs, increasing it to £[redacted] (without QALY weighting) when applied as a single change to the company's base case. This is because OS was higher in the Asia region than in the two other regions which were combined to generate the non-Asia cohort. This has implications both for generating an appropriate estimate of the ICER and for determining the appropriate QALY weighting to account for the severity of the condition.

The EAG also noted that the time-to-death approach used by the company to model utilities provided a utility estimate for people with expected survival of over 1 year that was very similar to age-adjusted utility values in the general population. The EAG preferred to use the utility estimates from the company's linear mixed effects model, but this did not have a large impact on the ICER. The EAG also explored the impact of using a progression-based approach to estimate utilities and this demonstrated that the ICER is somewhat sensitive to the choice between a time-to-death and a progression-based approach to model utilities.

Overall, the EAG's preferred base case ICER was £[redacted] per QALY for the deterministic analysis and £[redacted] per QALY for the probabilistic analysis, when not applying any QALY weighing to account for severity. Using an alternative parametric extrapolation, but still modelling OS and PFS using parametric curves fitted separately to each arm of the non-Asia cohort, reduced the deterministic

ICER to £[redacted]. Using the progression-based approach to estimate utilities from the linear mixed effects model increased the EAG's preferred estimate of the ICER to £[redacted].

The company argues that the appropriate QALY weighting is 1.2 based on the lifetime QALYs estimated for trastuzumab with chemotherapy in the trastuzumab appraisal (TA208). The EAG prefers to use the QALYs from the SoC arm of the model to estimate the absolute and proportionate QALY shortfall and determine the corresponding QALY weighting. This approach would support a QALY weighting of 1.0 when using the company's base case analysis, and 1.2 when using the EAG's preferred base case analysis. The EAG's base case ICER when applying a QALY weighting of 1.2 is £[redacted] when using the deterministic analysis and £[redacted] when using the probabilistic analysis.

Overall, the EAG's estimate of the ICER is much higher than the company's estimate and is above £30,000 per QALY even when applying a QALY weighting of 1.2. This is largely due to the EAG preferring to exclude data from the Asia region when estimating OS and PFS in the cost-effectiveness analysis, which is consistent with the company's claim that this data is not generalisable to the UK. However, the EAG notes that the non-Asia cohort is a *post hoc* analysis combining data from two regions. The EAG considers that data from the Western Europe/Israel/North America/Australia region could be more applicable to England than data from the Rest of World region. The EAG also expects that populating the model with data from the Western Europe/Israel/North America/Australia region could have a substantial impact on the ICER, but this could not be explored by the EAG with the data provided by the company. This is therefore an additional uncertainty that is not captured within the EAG's exploratory analysis.

7 REFERENCES

1. Merck Sharp & Dohme Ltd. Pembrolizumab with trastuzumab and chemotherapy for untreated HER2 positive advanced gastric or gastro-oesophageal junction cancer [ID3742]: Company submission. London, UK: National Institute for Health and Care Excellence,; 2023.
2. Yang L, Ying X, Liu S, Lyu G, Xu Z, Zhang X, *et al.* Gastric cancer: Epidemiology, risk factors and prevention strategies. *Chin J Cancer Res* 2020;32:695-704.
3. Cancer Research UK. Stomach cancer statistics. 2022. https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/stomach-cancer#heading-Zero (Accessed 26/04/2023).
4. National Institute for H, Care E. Pembrolizumab with trastuzumab and chemotherapy for untreated locally advanced unresectable or metastatic HER2-positive gastric or gastrooesophageal junction adenocarcinoma [ID3742]: Final scope. 2023. https://www.nice.org.uk/guidance/gid-ta10615/documents/final-scope (Accessed May 2023).
5. Smyth EC, Nilsson M, Grabsch HI, van Grieken NCT, Lordick F. Gastric cancer. *The Lancet* 2020;396:635-48.
6. Park Mh WMAHMNCTTBTNGJNRCD. National Oesophago-Gastric Cancer Audit. 2022 Annual Report. 2023. https://www.nogca.org.uk/reports/2022-annual-report/ (Accessed May 2023).
7. Lordick F, Carneiro F, Cascinu S, Fleitas T, Haustermans K, Piessen G, *et al.* Gastric cancer: ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up. *Annals of Oncology* 2022;33:1005-20.
8. National Institute for H, Care E. Oesophago-gastric cancer: assessment and management in adults: NICE guideline [NG83]. 2018. https://www.nice.org.uk/guidance/ng83 (Accessed May 2023).
9. Allum W, Lordick F, Alsina M, Andritsch E, Ba-Ssalamah A, Beishon M, *et al.* ECCO essential requirements for quality cancer care: Oesophageal and gastric cancer. *Critical Reviews in Oncology/Hematology* 2018;122:179-93.
10. National Institute for H, Care E. Trastuzumab for the treatment of HER2-positive metastatic gastric cancer: Technology appraisal guidance [TA208]. 2010. https://www.nice.org.uk/guidance/ta208 (Accessed May 2023).
11. National Institute for Health and Care Excellence. Capecitabine for the treatment of advanced gastric cancer: TA191. London, UK: National Institute for Health and Care Excellence,; 2010.
12. Cunningham D, Starling N, Rao S, Iveson T, Nicolson M, Coxon F, *et al.* Capecitabine and oxaliplatin for advanced esophagogastric cancer. *N Engl J Med* 2008;358:36-46.
13. Al-Batran SE, Hartmann JT, Probst S, Schmalenberg H, Hollerbach S, Hofheinz R, *et al.* Phase III trial in metastatic gastroesophageal adenocarcinoma with fluorouracil, leucovorin plus either oxaliplatin or cisplatin: a study of the Arbeitsgemeinschaft Internistische Onkologie. *J Clin Oncol* 2008;26:1435-42.
14. National Institute for H, Care E. Nivolumab with platinum- and fluoropyrimidine-based chemotherapy for untreated HER2-negative advanced gastric, gastro-oesophageal junction or oesophageal adenocarcinoma: Technology appraisal guidance [TA857]. 2023. https://www.nice.org.uk/guidance/ta857 (Accessed April 2023).
15. National Institute for H, Care E. Pembrolizumab with platinum- and fluoropyrimidine-based chemotherapy for untreated advanced oesophageal and gastro-oesophageal junction cancer: Technology appraisal guidance [TA737]. 2021. https://www.nice.org.uk/guidance/ta737 (Accessed May 2023).
16. Medicines, Healthcare products Regulatory A. KEYTRUDA® Summary of Product Characteristics. https://products.mhra.gov.uk/search/?search=keytruda&page=1&doc=Spc&rerouteType=0. (Accessed May 2023).

17. Merck Sharp & Dohme Ltd. KEYTRUDA® Summary of Product Characteristics (draft update).
18. European Medicines Agency. OPDIVO: Summary of Product Characteristics. https://www.ema.europa.eu/en/documents/product-information/opdivo-epar-product-information_en.pdf (Accessed 6th July 2023).
19. Merck Sharp & Dohme Ltd. Pembrolizumab with trastuzumab and chemotherapy for untreated HER2 positive advanced gastric or gastro-oesophageal junction cancer [ID3742]: Company response to the clarification letter. London, UK: National Institute for Health and Care Excellence,; 2023.
20. Ahn S, Kim KM. PD-L1 expression in gastric cancer: interchangeability of 22C3 and 28-8 pharmDx assays for responses to immunotherapy. *Mod Pathol* 2021;34:1719-27.
21. Merck Sharp & Dohme Ltd. KEYNOTE-811 Protocol. Data on File,: Merck Sharp & Dohme Ltd,; 2022.
22. Committee for Medicinal Products for Human Use (CHMP). Assessment Report: Keytruda - Procedure No. EMEA/H/C/003820/II/0097. Amsterdam, Netherlands: European Medicines Agency,; 2021.
23. Banno M, Tsujimoto Y, Kataoka Y. Using the Cochrane Central Register of Controlled Trials to identify clinical trial registration is insufficient: a cross-sectional study. *BMC Med Res Methodol* 2020;20:200.
24. Lefebvre C, Glanville J, Briscoe S, Featherstone R, Littlewood A, Marshall C*, et al.* Chapter 4: Searching for and selecting studies. In: Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M*, et al.*, editors. *Cochrane Handbook for Systematic Reviews of Interventions version 6.3 (updated February 2022). Cochrane, 2022*: Cochrane; 2022.
25. Waddell T, Chau I, Cunningham D, Gonzalez D, Okines AF, Okines C*, et al.* Epirubicin, oxaliplatin, and capecitabine with or without panitumumab for patients with previously untreated advanced oesophagogastric cancer (REAL3): a randomised, open-label phase 3 trial. *Lancet Oncol* 2013;14:481-9.
26. Hall PS, Swinson D, Cairns DA, Waters JS, Petty R, Allmark C*, et al.* Efficacy of Reduced-Intensity Chemotherapy With Oxaliplatin and Capecitabine on Quality of Life and Cancer Control Among Older and Frail Patients With Advanced Gastroesophageal Cancer: The GO2 Phase 3 Randomized Clinical Trial. *JAMA Oncol* 2021;7:869-77.
27. Janjigian YY, Kawazoe A, Yañez P, Li N, Lonardi S, Kolesnik O*, et al.* The KEYNOTE-811 trial of dual PD-1 and HER2 blockade in HER2-positive gastric cancer. *Nature* 2021;600:727-30.
28. Chung HC, Bang YJ, C SF, Qin SK, Satoh T, Shitara K*, et al.* First-line pembrolizumab/placebo plus trastuzumab and chemotherapy in HER2-positive advanced gastric cancer: KEYNOTE-811. *Future Oncol* 2021;17:491-501.
29. Merck S, Dohme. Pembrolizumab (MK-3475): clinical study report (P811V02MK3475). In: A Phase III, Randomized, Double-blind Trial Comparing Trastuzumab Plus Chemotherapy and Pembrolizumab With Trastuzumab Plus Chemotherapy and Placebo as First-line Treatment in Participants With HER2 Positive Advanced Gastric or Gastroesophageal Junction Adenocarcinoma (KEYNOTE-811). Data on file; 2022.
30. ClinicalTrials.gov. Pembrolizumab/Placebo Plus Trastuzumab Plus Chemotherapy in Human Epidermal Growth Factor Receptor 2 Positive (HER2+) Advanced Gastric or Gastroesophageal Junction (GEJ) Adenocarcinoma (MK-3475-811/KEYNOTE-811). NIH National Library of Medicine; 2022. https://www.clinicaltrials.gov/study/NCT03615326?term=%20NCT03615326&rank=1 (Accessed 29/06/2023).
31. Bang YJ, Van Cutsem E, Feyereislova A, Chung HC, Shen L, Sawaki A*, et al.* Trastuzumab in combination with chemotherapy versus chemotherapy alone for treatment of HER2-positive

advanced gastric or gastro-oesophageal junction cancer (ToGA): a phase 3, open-label, randomised controlled trial. *Lancet* 2010;376:687-97.

32. ClinicalTrials.gov. ToGA Study - A Study of Herceptin (Trastuzumab) in Combination With Chemotherapy Compared With Chemotherapy Alone in Patients With HER2-Positive Advanced Gastric Cancer. NIH National Library of Medicine; 2014. https://www.clinicaltrials.gov/study/NCT01041404?term=NCT01041404&rank=1 (Accessed 29/06/2023).
33. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I*, et al.* RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019;366:l4898.
34. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R*, et al.* New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer* 2009;45:228-47.
35. European Medicines A. Appendix 1 to the guideline on the evaluation of anticancer medicinal products in man - methodological consideration for using progression-free survival (PFS) or disease-free survival (DFS) in confirmatory trials - Scientific guideline. 2013. https://www.ema.europa.eu/en/appendix-1-guideline-evaluation-anticancer-medicinal-products-man-methodological-consideration-usinghttps://www.ema.europa.eu/en/appendix-1-guideline-evaluation-anticancer-medicinal-products-man-methodological-consideration-using (Accessed June 2023)
36. Belin L, Tan A, De Rycke Y, Dechartres A. Progression-free survival as a surrogate for overall survival in oncology trials: a methodological systematic review. *British Journal of Cancer* 2020;122:1707-14.
37. Seymour L, Bogaerts J, Perrone A, Ford R, Schwartz LH, Mandrekar S*, et al.* iRECIST: guidelines for response criteria for use in trials testing immunotherapeutics. *Lancet Oncol* 2017;18:e143-e52.
38. European Medicines A. Guideline on the evaluation of anticancer medicinal products in man. 2017. https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-evaluation-anticancer-medicinal-products-man-revision-5_en.pdf (Accessed
39. National Cancer Institute. Common Terminology Criteria for Adverse Events (CTCAE) Version 4.0; 2009.
40. Editor. Editorial. *Nature* 2021;592:659-60.
41. Commissioned by the Healthcare Quality Improvement Partnership. National Oesophago-Gastric Cancer Audit 2022; 2023.
42. Msd Bards HTA. Keytruda (MK-3475) HTA report: Drug Utilization. In: Protocol 811: A Phase III, Randomized, Double-blind Trial Comparing Trastuzumab Plus Chemotherapy and Pembrolizumab With Trastuzumab Plus Chemotherapy and Placebo as First-line Treatment in Participants With HER2 Positive Advanced Gastric or Gastroesophageal Junction Adenocarcinoma. Data on file; 2022.
43. Merck Sharp & Dohme Ltd. Pembrolizumab with trastuzumab and chemotherapy for untreated HER2 positive advanced gastric or gastro-oesophageal junction cancer [ID3742]: Company additional response to the clarification questions A5 and B16. London, UK: National Institute for Health and Care Excellence,; 2023.
44. Sjoquist KM, Lord SJ, Friedlander ML, John Simes R, Marschner IC, Lee CK. Progression-free survival as a surrogate endpoint for overall survival in modern ovarian cancer trials: a meta-analysis. *Ther Adv Med Oncol* 2018;10:1758835918788500.
45. Paoletti X, Oba K, Bang YJ, Bleiberg H, Boku N, Bouché O*, et al.* Progression-free survival as a surrogate for overall survival in advanced/recurrent gastric cancer trials: a meta-analysis. *J Natl Cancer Inst* 2013;105:1667-70.
46. Ara RWAJ. NICE DSU Technical Support Document 12: The use of health state utility values in decision models. 2011. https://www.sheffield.ac.uk/nice-dsu/tsds/full-list (Accessed April 2023).

47. Joint Formulary C. British National Formulary (BNF) (online). 2023. https://bnf.nice.org.uk (Accessed May 2023).
48. Department of H, Social C. Drugs and pharmaceutical electronic market information tool (eMIT). 2023. https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit (Accessed May 2023).
49. England NHS. National Schedule of NHS Costs 2021/22. 2023. https://www.england.nhs.uk/costing-in-the-nhs/national-cost-collection/ (Accessed May 2023).
50. Gomez-Ulloa D, Amonkar M, Kothari S, Cheung WY, Chau I, Zalcberg JR, *et al.* Real-world treatment patterns, healthcare resource use and clinical outcomes of patients receiving second line therapy for advanced or metastatic gastric cancer. *BMC Gastroenterol* 2020;20:133.
51. Barnish MS, Slater R, Matthews J, Long L, Bullement A, Shaw N, *et al.* Pembrolizumab with platinum-based chemotherapy for untreated advanced oesophageal cancer [ID3741]: A Single Technology Appraisal. London, UK: National Institute for Health and Care Excellence,; 2021.
52. National Institute for H, Care E. Pembrolizumab for untreated PD-L1-positive locally advanced or metastatic urothelial cancer when cisplatin is unsuitable: Technology appraisal guidance [TA522]. 2018. https://www.nice.org.uk/guidance/ta522 (Accessed
53. Curtis LA BA. Unit Costs of Health and Social Care 2022 - PSSRU: Unievrsity of Kent; 2022.
54. Wailoo MHASPA. NICE DSU: Estimating EQ-5D by age and sex for the UK. 2022. https://www.sheffield.ac.uk/nice-dsu/methods-development/estimating-eq-5d (Accessed April 2023).
55. Jones KC, Weatherly H, Birch S, Castelli A, Chalkley M, Dargan A, *et al.* Unit Costs of Health and Social Care 2022 Manual. Personal Social Services Research Unit (University of Kent) & Centre for Health Economics (University of York), Kent, UK; 2023.
56. Buyukkaramikli NC, Rutten-van Molken M, Severens JL, Al M. TECH-VER: A Verification Checklist to Reduce Errors in Models and Improve Their Credibility. *Pharmacoeconomics* 2019;37:1391-408.
57. National Institute for H, Care E. Ramucirumab for treating advanced gastric cancer or gastro–oesophageal junction adenocarcinoma previously treated with chemotherapy: Technology appraisal guidance [TA378]. 2016. https://www.nice.org.uk/guidance/ta378 (Accessed April 2023).
58. Norman G, Rice S, Spackman E, Stirk L, Danso-Appiah A, Suh D, *et al.* Trastuzumab for the treatment of HER2 positive metastatic adenocarcinoma of the stomach or gastro-oesophageal junction: ERG report. London, UK: National Institute for Health and Care Excellence,; 2010.
59. Ren S, Squires H, Kaltenthaler E, Hock E, Rawdin A, Clowes M, *et al.* Pembrolizumab for locally advanced or metastatic urothelial cancer where cisplatin is unsuitable: A Systematic Review. London, UK.: National Institute for Health and Care Excellence,; 2018.
60. Taylor R, Carter S. Valuing choice - dying at home: A care for the more equitable provision of high quality support for people who wish to die at home. An economic and social policy opinion commissioned by Marie Curie Cancer Care. London, UK: School of Pharmacy, University of London.; 2004.
61. Schneider P MSL-KJDJNG. QALY Shortfall Calculator. 2021. https://r4scharr.shinyapps.io/shortfall/ (Accessed April 2023).
62. Norman G, Rice S, Spackman E, Stirk L, Danso-Appiah A, Suh D, *et al.* Trastuzumab for the treatment of HER2 positive metastatic adenocarcinoma of the stomach or gastro-oesophageal junction: ERG revisions based on manufacturer response to ACD. London, UK: National Institute for Health and Clinical Excellence; 2010.
63. Centre for Review and Dissemination. Systematic reviews. Guidance for Undertaking Reviews in Health Care: CRD, University of York; 2009.
64. Satoh T, Bang YJ, Gotovkin EA, Hamamoto Y, Kang YK, Moiseyenko VM, *et al.* Quality of life in the trastuzumab for gastric cancer trial. *Oncologist* 2014;19:712-9.

65. Sawaki A, Ohashi Y, Omuro Y, Satoh T, Hamamoto Y, Boku N*, et al.* Efficacy of trastuzumab in Japanese patients with HER2-positive advanced gastric or gastroesophageal junction cancer: a subgroup analysis of the Trastuzumab for Gastric Cancer (ToGA) study. *Gastric Cancer* 2012;15:313-22.

8 APPENDICES

Appendix 1: Trial of comparator, ToGA

Study characteristics for ToGA trial

ToGA was a Phase III open-label RCT that compared chemotherapy with or without trastuzumab, for previously untreated HER2- positive locally advanced, recurrent, and/or metastatic cancer gastric or GOJ cancer (Bang *et al.*2010).[31] It was a multicentre study in 24 countries[31] across Europe (including the UK), Asia, North America, South America, African and Australia.[32] Randomisation was stratified by ECOG performance status, chemotherapy regimen, extent of disease, primary cancer site, and measurability of disease.[31] Cancers were histologically confirmed, and HER2 status was assessed with immunohistochemistry (HercepTest, Dako, Denmark]) and fluorescence in-situ hybridisation (FISH; HER2 FISH pharmDx, Dako).[31] Recruitment occurred between 2005 and 2008.[31]

The chemotherapy regimen[31] for both treatment groups was given every three weeks for six cycles, and consisted of cisplatin (80 mg/m² i.v. infusion day 1) plus either:
Capecitabine (1000 mg/m² orally twice a day for 14 days followed by one-week rest):
or 5-fluorouracil (800 mg/m² per day was given by continuous intravenous infusion on days 1–5 of each cycle).
The intervention group additionally received trastuzumab (8 mg/kg i.v. infusion day 1 of the first cycle, followed by 6 mg/kg every 3 weeks).[31]
Study treatment was continued until disease progression or unacceptable toxicity.[31] Dose adjustments for chemotherapy, or interruptions of trastuzumab were allowed.[31] The chemotherapy received by the majority of patients in both treatment groups was capecitabine plus cisplatin (XP) (Table 34).

The primary outcome was OS, defined as time from randomisation until death from any cause (assessment schedule days 1, 8, 15, 22, 43, 64, 85, 106, 127, and then every 21 days).[31,32] Secondary outcomes included: PFS defined as time from randomisation to progression (at least a 20% increase for target lesion, or unequivocal progression of existing non-target lesion) or death; ORR (RECIST criteria); and safety.[31,32]

Table 34 ToGA (NCT01041404) overview of study characteristics[31, 32]

Trial design	Population	Intervention N=294 in analysis	Comparator N=290 in analysis	Primary outcome
Phase III RCT, open-label	Adults (age >18) with previously untreated HER2-positive, histologically confirmed locally advanced, recurrent, and/or metastatic gastric or GOJ adenocarcinoma	Trastuzumab plus: capecitabine plus cisplatin (XP) n=256 (87%); or 5-fluorouracil plus cisplatin (FP) n=38 (13%)	Capecitabine plus cisplatin (XP) n=255 (88%) or 5-fluorouracil plus cisplatin (FP) n=35 (12%)	Overall Survival (OS) - Time to Event

Abbreviations: FP= cisplatin plus 5-FU; OS=overall survival;; HER2= human epidermal growth factor receptor 2; GOJ=gastro-oesophageal junction; RCT, randomised controlled trial

Risk of bias for ToGA trial

Risk of bias was assessed based on the Cochrane Risk of Bias tool 2.0. There was some similarity between assessment by the CS and EAG (Table 35). The EAG and CS differed in assessment of bias arising from the randomization process. The CS thought this was high risk because ToGA was not a double-blind study, however allocation concealment refers to preventing bias in intervention assignment by preventing trial personnel from knowing, or altering, the allocation sequence before and until assignment. Lack of blinding is captured in bias in measurement of the outcome. Lack of blinding can lead to a risk of performance and detection bias. The ToGA trial was open-label, and it was unclear who assessed the outcome measures. If outcome assessors are not blinded, then the potential for bias needs to be considered for each outcome assessed. Patient-reported outcome measures are more likely to be biased than objective measures such as OS, as assessment of OS would not have been influenced by knowledge of intervention received.[63] Both PFS and safety outcomes might have an element of subjectivity, and so are subject to bias in the ToGA trial. However, PFS and AEs are well-defined, which should reduce the effect of bias.

Table 35 Risk of bias ToGA trial

Type of bias	ToGA			
	Review authors' judgement	**Support for judgement**	**EAG judgement**	**Support for judgement**[31, 32 64]
Bias arising from the randomization process	High risk	Open-label study; participants were randomly assigned 1:1 to trastuzumab + chemotherapy or chemotherapy alone using a central interactive voice recognition system and assignment was *not* masked to either patients or investigators.	Low risk	Allocation sequence – stratified, randomised block design via interactive voice recognition system Allocation concealment – allocation was concealed, as randomisation performed centrally using an interactive voice recognition system Baseline characteristics –balanced between treatment groups
Bias due to deviations from intended interventions	Some concerns	Open-label study; patients and investigators were aware of assigned interventions. There were no deviations from the intended intervention because of trial context and appropriate analysis methods were employed to estimate treatment effects.	Some concerns	Participant awareness of assigned intervention – open-label Clinician/carer awareness of assigned intervention – open-label Trial context – no strong reason to believe, that the trial context led to failure to implement the protocol interventions Appropriate analyses - analyses were appropriate, modified ITT for effectiveness and HRQoL included all patients who received study medication at least once

Type of bias	ToGA			
	Review authors' judgement	**Support for judgement**	**EAG judgement**	**Support for judgement[31, 32 64]**
Bias due to missing outcome data	Low risk	Data for outcomes available represented all or nearly all randomized participants.	Low risk	Available outcome data – OS and PFS nearly all participants provided data
Bias in measurement of the outcome	High risk	Open-label study; assessment of the outcome may have been influenced by knowledge of intervention received.	OS – low risk PFS and AEs – some concerns	Method of measuring outcomes – appropriate Outcome measurement for treatment groups – same measurements at same time points across treatment groups Outcome assessor awareness – unclear Could assessment of the outcome have been influenced by knowledge of intervention received? – OS no; PFS might have an element of subjectivity, however the outcome is well-defined; AEs prone to influence, however the outcomes are well-defined
Bias in selection of the reported result	Low risk	Analysis was in accordance with a pre-specified analysis plan that was finalized before the outcome data were available for analysis.	Low risk	Analyses pre-specified – analyses were prespecified for OS and PFS Multiple outcome measurements – outcomes definitions pre-specified
Overall bias	Some concerns	Some concerns regarding bias due to open-label design.	Some concerns	

ToGA trial results

Of 594 patients randomised, 584 received study treatment and provided data for the effectiveness analyses.[31] Analyses occurred after 18.6 months median follow-up in the trastuzumab group, and 17.1 months in the comparator group.[31] In the trastuzumab plus chemotherapy group, the median number of cycles of trastuzumab therapy was eight (range 1–49), and the median number of chemotherapy cycles was six for cisplatin, capecitabine and 5-fluorouracil.[31] In the chemotherapy group, the median number of cycles of was five for cisplatin and capecitabine, and four for 5-fluorouracil.[31]

Median PFS for the trastuzumab plus chemotherapy group was 6.7 months (95%CI 6, 8), and for the chemotherapy group 5.5 months (95%CI 5, 6), HR 0.71 (95%CI 0.59, 0.85) p=0.0002, significantly favouring the trastuzumab plus chemotherapy group.[31]

Median OS for the trastuzumab plus chemotherapy group (n=294) was 13.8 months (95%CI 12, 16), and for the chemotherapy group (n=290) 11.1 months (95% confidence interval 10, 13), HR 0.74 (95%CI 0.60, 0.91) p=0.0046, significantly favouring the trastuzumab plus chemotherapy group.[31] Pre-planned subgroup analyses were reported for OS. For region subgroups, HRs for trastuzumab plus chemotherapy with reference chemotherapy were: Central or South America (n=52) HR 0.44 (95%CI 0.21, 0.90); Europe (n=190) HR 0.63 (95%CI 0.44, 0.89); Asia (n=319) HR 0.82 (95%CI 0.67, 1.11); Other (n=23) HR 1.22 (95%CI 0.48, 1.46).[31] The results showed the same direction of effect for OS for all regions except “Other” which had a small sample size and wide confidence intervals. It appeared that in the Europe region there was a more favourable treatment response to trastuzumab than in the Asia region, and that the most favourable trastuzumab response was in Central or South America, however subgroups were relatively small and not powered to detect treatment differences.[31]

In the Japanese region subgroup of ToGA, in which chemotherapy type was XP for all patients, median OS for the trastuzumab plus chemotherapy group (n=51) was 15.9 months (95%CI 12, 25), and for the chemotherapy group (n=50) 17.7 months (95% confidence interval 12, 24), HR 1.00 (95%CI 0.59, 1.69).[65] The Japanese region subgroup of ToGA had a median PFS for the trastuzumab plus XP (n=51) was 6.2 months (95%CI 5, 7), and for the XP group (n=50) 5.6 months (95% confidence interval 5, 7), HR 0.92 (95%CI 0.60, 1.43).[65] These were non-significant treatment group differences, however the direction of effect favoured the intervention group for PFS, but favoured the comparator group for OS, reflecting the pattern of the Asia region results of KEYNOTE-811.

Adverse events led to non-completion of the trials for 35/294 (11.9%) of the trastuzumab plus chemotherapy group, and 45/290 (15.5%) of the chemotherapy group.[32] Grade 3 or 4 AEs were experienced by 201/294 (68%) of the trastuzumab plus chemotherapy group, and 198/290 (68%) of the

chemotherapy group.[31] The most common grade 3 or 4 AEs were anaemia, neutropenia, diarrhoea, nausea, vomiting, thrombocytopenia and asthenia.[31] Any AE was experienced by 99% of the trastuzumab plus chemotherapy group, and 98% of the chemotherapy group.[31]

HRQoL was measured with the European Organization for Research and Treatment of Cancer (EORTC) quality of life QLQ-C30 (version 3.0) Global Health status (GHS). [64] The median time to 10% deterioration in the GHS score of the QLQ-C30 questionnaire was 10.2 months in the trastuzumab plus chemotherapy group, and 6.4 months in the chemotherapy group, significantly favouring the trastuzumab plus chemotherapy group ($p<0.0001$).[64]

ToGA and KEYNOTE-811comparison

Both ToGA and KEYNOTE-811 included a trastuzumab plus chemotherapy group. Chemotherapy type could be FP in either trial, however a minority of patients received this, with the majority in KEYNOTE-811 receiving CAPOX, and in ToGA, XP.

ToGA recruited between 2005 and 2008 [31], whereas KEYNOTE-811 started recruiting in 2018.[30] Both trials recruited adults with previously untreated HER2-positive, histologically confirmed locally advanced, recurrent, and/or metastatic gastric or GOJ adenocarcinoma. Both trials had a majority metastatic population (Table 36). ToGA did not report PD-L1 status. ToGA included ECOG 2, unlike KEYNOTE-811.

Outcomes of the trials are shown in Table 37. The trastuzumab plus chemotherapy group of the KEYNOTE-811 trial had longer OS and PFS than the trastuzumab plus chemotherapy group of the ToGA trial. This may have been due to patient characteristics (e.g., inclusion of ECOG2 in ToGA) or the influence of subsequent therapies. Adverse event rates were similar across trials.

Table 36 Participant Baseline Characteristics by Treatment Group [adapted from CS Section B Table 6, and ToGA references]

	KEYNOTE-811 CPS≥1 Participants global cohort Pembrolizumab plus trastuzumab plus chemotherapy		**KEYNOTE-811 CPS≥1 Participants global cohort Trastuzumab plus Chemotherapy**		**ToGA Trastuzumab plus Chemotherapy**		**ToGA Chemotherapy**	
	n	**(%)**	**n**	**(%)**	**n**	**(%)**	**n**	**(%)**
Participants in population	298		296		294		290	
PD-L1 status CPS ≥1	298	100	296	100	NR	NR	NR	NR
Sex								
Male	240	(80.5)	237	(80.1)	226	(77%)	218	(75%)
Female	58	(19.5)	59	(19.9)	68	(23)	72	(25)
Age (Years)								
Mean	60.6		61.4		59·4 (SD 10·8)		58·5 (SD 11·2)	
SE	0.7		0.6					
Race								
American Indian Or Alaska Native	5	(1.7)	6	(2.0)	NR	NR	NR	NR
Asian	97	(32.6)	97	(32.8)	151	(51)	158	(54)
Black Or African American	2	(0.7)	2	(0.7)	1	(<1)	2	(1)
Multiple	5	(1.7)	4	(1.4)	NR	NR	NR	NR
White	188	(63.1)	184	(62.2)	115	(39)	105	(36)
Missing	1	(0.3)	3	(1.0)	NR	NR	NR	NR
Geographic Region of Enrolling Site								
Western Europe/Israel/North America/Australia	97	(32.6)	96	(32.4)	NR Across both groups			

	KEYNOTE-811 CPS≥1 Participants global cohort Pembrolizumab plus trastuzumab plus chemotherapy		**KEYNOTE-811 CPS≥1 Participants global cohort Trastuzumab plus Chemotherapy**		**ToGA Trastuzumab plus Chemotherapy**		**ToGA Chemotherapy**	
	n	**(%)**	**n**	**(%)**	**n**	**(%)**	**n**	**(%)**
					Europe n=190			
Asia	96	(32.2)	96	(32.4)	NR Across both groups Asia n=319			
Rest of the World	105	(35.2)	104	(35.1)	NR Across both groups Central or South America n=52			
ECOG Performance Scale								
0	127	(42.6)	121	(40.9)	NR, ECOG 0 or 1 264	(90)	NR, ECOG 0 or 1 263	(91)
1	171	(57.4)	174	(58.8)				
2	0	0	0	0	30	(10)	27	(9)
Missing	0	(0.0)	1	(0.3)	0	0	0	0
Primary Location at Diagnosis								
Adenocarcinoma of the GOJ	97	(32.6)	99	(33.4)	58	20	48	17
Adenocarcinoma of the stomach	201	(67.4)	197	(66.6)	236	80	242	83

	KEYNOTE-811 CPS≥1 Participants global cohort Pembrolizumab plus trastuzumab plus chemotherapy		KEYNOTE-811 CPS≥1 Participants global cohort Trastuzumab plus Chemotherapy		ToGA Trastuzumab plus Chemotherapy		ToGA Chemotherapy	
	n	**(%)**	**n**	**(%)**	**n**	**(%)**	**n**	**(%)**
Disease Status								
Locally advanced	8	(2.7)	6	(2.0)	10	3	10	3
Metastatic	290	(97.3)	290	(98.0)	284	97	280	97
Chemotherapy Regimen								
CAPOX	251	(84.2)	253	(85.5)	0	0	0	0
FP	47	(15.8)	43	(14.5)	38	13	35	12
XP	0	0	0	0	256	87	255	88

Table 37 Outcomes of ToGA and KEYNOTE-811

	KEYNOTE-811 CPS≥1 Participants global cohort Pembrolizumab plus trastuzumab plus chemotherapy	**KEYNOTE-811 CPS≥1 Participants global cohort Trastuzumab plus Chemotherapy**	**ToGA Trastuzumab plus Chemotherapy**	**ToGA Chemotherapy**
N patients	298	296	294	290
Median OS (95%CI)	20.5 months (18.2-24.3)	15.6 months (13.5-18.6)	13.8 months (12-16)	11.1 months (10-13)
Median PFS (95%CI)	10.8 months (8.5-12.5)	7.2 months (6.8-8.4)	6.7 months (6-8)	5.5 months (5-6)
ORR n (%)	218 (73.2)	173 (58.4)	139 (47)	100 (35)
AEs N patients	298	295	294	290
Any AE	97.0	96.3	99.0	98.0
Grade ≥3 AEs	73.2%	65.1%	68%	68%

Summary of ToGA trial

ToGA was a Phase III open-label RCT that compared chemotherapy (XP or FP) with or without trastuzumab. The open-label nature of the trials meant there were some concerns of risk of bias. Median time to death for the trastuzumab plus chemotherapy group (n=294) was 13.8 months (95%CI 12, 16), and 11.1 months (95%CI 10, 13) for the chemotherapy group (n=290), favouring the trastuzumab plus chemotherapy group (HR 0.74 (95%CI 0.60, 0.91) p=0.0046). Median PFS for the trastuzumab plus chemotherapy group was 6.7 months (95%CI 6, 8), and for the chemotherapy group 5.5 months (95%CI 5, 6), favouring the trastuzumab plus chemotherapy group (HR 0.71 (95%CI 0.59, 0.85) p=0.0002). Grade 3 or 4 AEs were experienced by 201/294 (68%) of the trastuzumab plus chemotherapy group, and 198/290 (68%) of the chemotherapy group.