NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Nivolumab in combination for untreated advanced unresectable recurrent or metastatic oesophageal squamous cell carcinoma cancer

Committee Papers

NICE National Institute for Health and Care Excellence

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Nivolumab in combination for untreated advanced unresectable recurrent or metastatic oesophageal squamous cell carcinoma cancer [ID2712]

Contents:

The following documents are made available to consultees and commentators:

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

1. **Company submission** from **Bristol-Myers Squibb Pharmaceuticals Ltd (BMS)**

2. **Clarification questions and company responses**
   a. Company Response
   b. Additional queries

3. **Patient group, professional group and NHS organisation submission** from:
   - GUTS UK

4. **Evidence Review Group report** prepared by Kleijnen Systematic Reviews (KSR)

5. **Evidence Review Group – factual accuracy check**

6. **Technical engagement response** from **Bristol-Myers Squibb Pharmaceuticals Ltd (BMS)**

7. **Technical engagement response from consultees and commentators:**
   - MSD

8. **Evidence Review Group critique of company response to technical engagement** prepared by Kleijnen Systematic Reviews (KSR)

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

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Nivolumab in combination with fluoropyrimidine- and platinum-based combination chemotherapy for the first-line treatment of adult patients with unresectable advanced, recurrent, or metastatic oesophageal squamous cell carcinoma with tumour cell PD-L1 expression ≥1%

[ID2712]

Document B
Company evidence submission

March 2022

File name	Version	Contains confidential information	Date
		Yes	

Contents

List of Tables

List of Figures

List of abbreviations

Abbreviation	Definition
AE	adverse event
AIDS	acquired immunodeficiency syndrome
AJCC	American Joint committee on Cancer
AOC	advanced oesophageal cancer
BICR	blinded independent central review
BMS	Bristol-Myers Squibb UK Ltd.
BOR	best overall response
CHMP	Committee for Medicinal Products for Human Use
CI	confidence interval
CMH	Cochran–Mantel–Haenszel
CR	complete response
CSR	clinical study report
CTC	common terminology criteria
DOR	duration of response
ECOG	Eastern Cooperative Oncology Group
ECS	oesophageal cancer subscale
EMA	European Medicines Agency
EQ-5D-3L	EuroQol questionnaire comprising 5 dimensions, with each dimension having 3 levels
ESCC	oesophageal squamous cell cancer
ESMO	European Society for Medical Oncology
EWB	emotional well-being
FACT-E	Functional Assessment of Cancer Therapy - Oesophageal
FACT-G7	Functional Assessment of Cancer Therapy – General
FWB	functional well-being
GEJ	gastro-oesophageal junction
HIV	human immunodeficiency virus
HR	hazard ratio
HRQoL	health-related quality of life
IgG4	immunoglobulin G4
IMAE	immune-mediated adverse event
IRT	interactive response technology
IV	intravenous
KM	Kaplan Meier
MAA	Marketing Authorisation Application
MMR	measles, mumps and rubella
NCI CTCAE	National Cancer Institute Common Terminology Criteria for Adverse Events
NICE	National Institute for Health and Care Excellence
OC	oesophageal cancer
OESI	other event of special interest
ORR	objective response rate
OS	overall survival
OSCC	oesophageal squamous cell carcinoma
PD	progressed disease

PD-1/2	programmed cell death protein 1/2
PD-L1/2	programmed death ligand 1/2
PFS	progression free survival
PK	pharmacokinetic
PR	partial remission
PWB	physical well-being
Q2W	every 2 weeks
Q4W	every 4 weeks
RANK-L	receptor activator of nuclear factor kappa-B ligand
RECIST	Response Evaluation Criteria in Solid Tumours
RoW	rest of world
SAE	serious adverse event
SAP	statistical analysis plan
SD	standard deviation
SLR	systematic literature review
SWB	social/family well-being
TMN	tumour/metastasis/node staging system
TSST	time to second subsequent therapy
TRAE	treatment-related adverse events
UK	United Kingdom
USA	United States of America
VAS	visual analogue scale
WHO	World Health Organisation

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

B.1.1 Decision problem

This submission covers the full anticipated marketing authorisation for [redacted]. The decision problem addressed is consistent with the final NICE scope and the NICE reference case as outlined in Table 1.

Table 1. The decision problem

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from NICE
Population	Adults with unresectable advanced, recurrent or metastatic, previously untreated OSCC	[redacted]	The evidence provided in this submission is derived from the pivotal CheckMate 648 trial, which demonstrates that the [redacted]. This population is also in line with the expected European Marketing Authorisation (EMA) licensing.
Intervention	Nivolumab in combination with fluoropyrimidine- and platinum-based chemotherapy	As per NICE scope	Not applicable; as specified in the draft Summary of Product Characteristics (SmPC)

Comparator(s)	Platinum-based chemotherapy without nivolumab, such as: • Doublet treatment with fluorouracil or capecitabine plus cisplatin or oxaliplatin • Triplet treatment with fluorouracil or capecitabine plus cisplatin	Platinum-based chemotherapy without nivolumab, such as: • Doublet treatment with fluorouracil or capecitabine plus cisplatin or oxaliplatin For tumours that express PD-L1 with a combined positive score (CPS) of 10 or more: • Pembrolizumab with platinum- and fluoropyrimidine-based chemotherapy	It should be noted that epirubicin-based triplet therapy is not commonly used in UK clinical practice. During TA737, the clinical expert stated that triplet therapy is no longer standard of care as it does not provide additional efficacy and increases toxicity.[1] The committee concluded that a dual chemotherapy regimen would be the appropriate comparator for TA737.[1] This aligns with expert advice provided to BMS.[2] Hence, assessment of epirubicin-based triplet therapy may not be relevant to decision making for this appraisal. Further, it should also be noted that pembrolizumab was only recently recommended by NICE (October 2021) and is hence not yet standard of care.

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from NICE
	or oxaliplatin plus epirubicin For tumours that express PD-L1 with a combined positive score (CPS) of 10 or more: • Pembrolizumab with platinum- and fluoropyrimidine-based chemotherapy		

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from NICE
Outcomes	The outcome measures to be considered include: • Overall survival • Progression-free survival • Response rate • Adverse effects of treatment • Health-related quality of life	As per NICE scope	Not applicable; additional relevant clinical outcomes are presented, including duration of response, objective response rate, complete response rate and partial response rate.

Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared.	As per NICE scope	Not applicable

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from NICE
	Costs will be considered from an NHS and Personal Social Services perspective. The availability of any commercial arrangements for the intervention, comparator and subsequent treatment technologies will be taken into account.		

Other considerations	If evidence allows subgroups by degree of PD-L1 expression and cancer histology will be considered. Guidance will only be issued in accordance with the marketing authorisation. Where the wording of the therapeutic indication does not include specific treatment combinations, guidance will be issued only in the context of the evidence that has underpinned the marketing authorisation	Pre-defined subgroups are presented for PD-L1 ≥1% and all randomised patients, in line with the NICE scope. The costs for PD-L1 screening are included.	As 98% of the patients included in CheckMate 648 study histologically have OSCC, no further subgroup analysis was conducted for the purpose of cost-effectiveness modelling.

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from NICE
	granted by the regulator.		
Special considerations including issues related to equity or equality	Not applicable	No equality issues have been identified or are anticipated.	Not applicable

B.1.2 Description of the technology being appraised

Details of the technology being appraised in this submission are summarised in Table 2 and detailed in the following subsections. Additionally, the Summary of Product Characteristics for nivolumab (Opdivo®) is presented in Appendix C.

Table 2. Technology being appraised

UK approved name and brand name	**Nivolumab (Opdivo®)**
Mechanism of action	Nivolumab is a human immunoglobulin G4 (IgG4) monoclonal antibody, which binds to the programmed death-1 (PD-1) receptor and blocks its interaction with PD-L1 and PD-L2. The PD-1 receptor is a negative regulator of T-cell activity that is involved in controlling the T-cell immune response. Engagement of PD-1 with the ligands PD-L1 and PD-L2, which are expressed on the surface of antigen presenting cells and may be expressed by tumours or other cells in the tumour microenvironment, results in the inhibition of T-cell proliferation and cytokine secretion. Nivolumab potentiates T-cell responses, including anti-tumour responses, through the blockade of PD-1 binding to PD-L1 and PD-L2 ligands.[3]
Marketing authorisation/ CE mark status	A Marketing authorisation application was submitted to the European Medicines Agency (EMA) for [redacted] • Regulatory submission was in [redacted]. • CHMP opinion was adopted on 24th February 2022. • Regulatory approval and marketing authorisation are expected in [redacted].
Indications and any restriction(s) as described in the summary of product	[redacted]

characteristics (SmPC)	
Method of administration and dosage	The recommended dose of nivolumab is 240 mg every 2 weeks or 480 mg every 4 weeks administered intravenously over 30 minutes in combination with fluoropyrimidine- and platinum-based chemotherapy. Treatment with nivolumab is recommended until disease progression, unacceptable toxicity, or up to 24 months in patients without disease progression.[4]
Additional tests or investigations	PD-L1 testing is required using a validated assay.[4]
List price and average cost of a course of treatment	**List price:** Nivolumab: £2,633 per 240 mg vial; £1,097 per 100 mg vial; £439.00 per 40 mg vial. **Patient access scheme price** Nivolumab: [redacted] per 240 mg vial; £[redacted] per 100 mg vial; £[redacted] per 40 mg vial.
Patient access scheme (if applicable)	There is a confidential simple discount PAS for nivolumab, which applies to all current and future indications.

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

Summary

- Oesophageal cancer (OC) is a malignant tumour developing from the cells lining the oesophagus (Figure 1)[5]
- In the UK, OC is often diagnosed at a late stage (70-80% of patients with OC are diagnosed with either lymph node or distant metastasis),[6] and 37-42% of cases have metastases at the point of diagnosis.[7]
- Squamous cell carcinoma (SCC) and adenocarcinoma are the two major histology types of OC and account for over 95% of cases.[8] However, there is a notable global variation in the distribution of histological types of OC where in Western countries, such as the UK, the majority (two-thirds) of OC cases are adenocarcinomas, while approximately a third are SCC.[9,10]
- The prognosis for unresectable OC is poor. In England, less than half of patients diagnosed with OC (46.5%) remain alive at 12 months.[11]
- Management of patients with unresectable advanced, recurrent or metastatic OSCC is limited, and aims to keep the disease under control for as long as possible and relieve any symptoms.[6,12]
- Nivolumab in combination with chemotherapy provides a significant and clinically meaningful improvement in median OS for patients with OSCC vs. chemotherapy alone ([redacted] months vs. [redacted] months, based on the most recent data available).[13] The improvement is more marked in the subgroup of patients with PD-L1 ≥1%, where median OS benefit is [redacted] months ([redacted] months vs. [redacted] months).
- Similarly, nivolumab with chemotherapy provides a significant improvement in median PFS for patients with PD-L1 ≥1% ([redacted] months vs. [redacted] months).[13]
- Nivolumab with chemotherapy would represent an additional first-line treatment option for patients with unresectable advanced, recurrent or metastatic, previously untreated OSCC, addressing the significant unmet need for patients with tumours cell PD-L1 ≥1%.

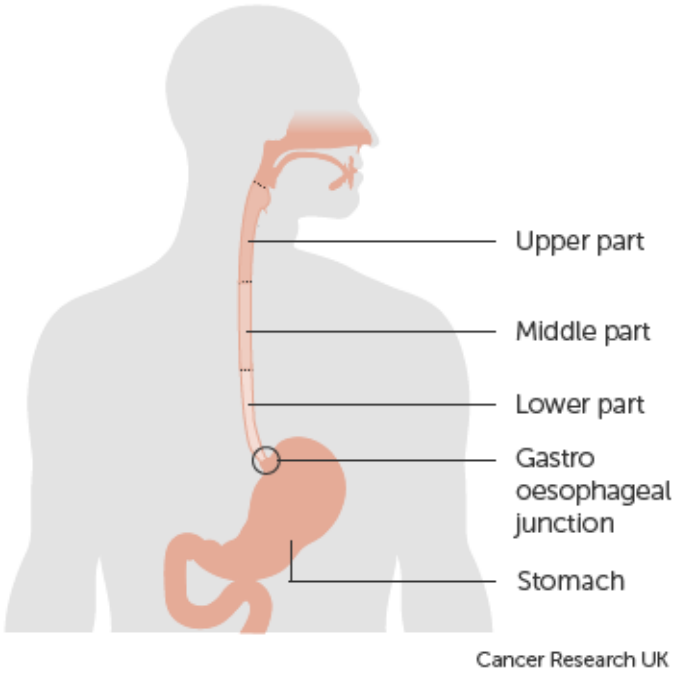
B.1.3.1 Disease Background

Oesophageal cancer (OC) is a malignant tumour developing from cells lining the oesophagus (Figure 1).[5] There are two main histological subtypes of OC: oesophageal squamous cell carcinoma (OSCC) and oesophageal adenocarcinoma (OAC),[8] which account for more than

95% of cases of OC and can be considered epidemiologically and pathologically distinct diseases that share an anatomical site.

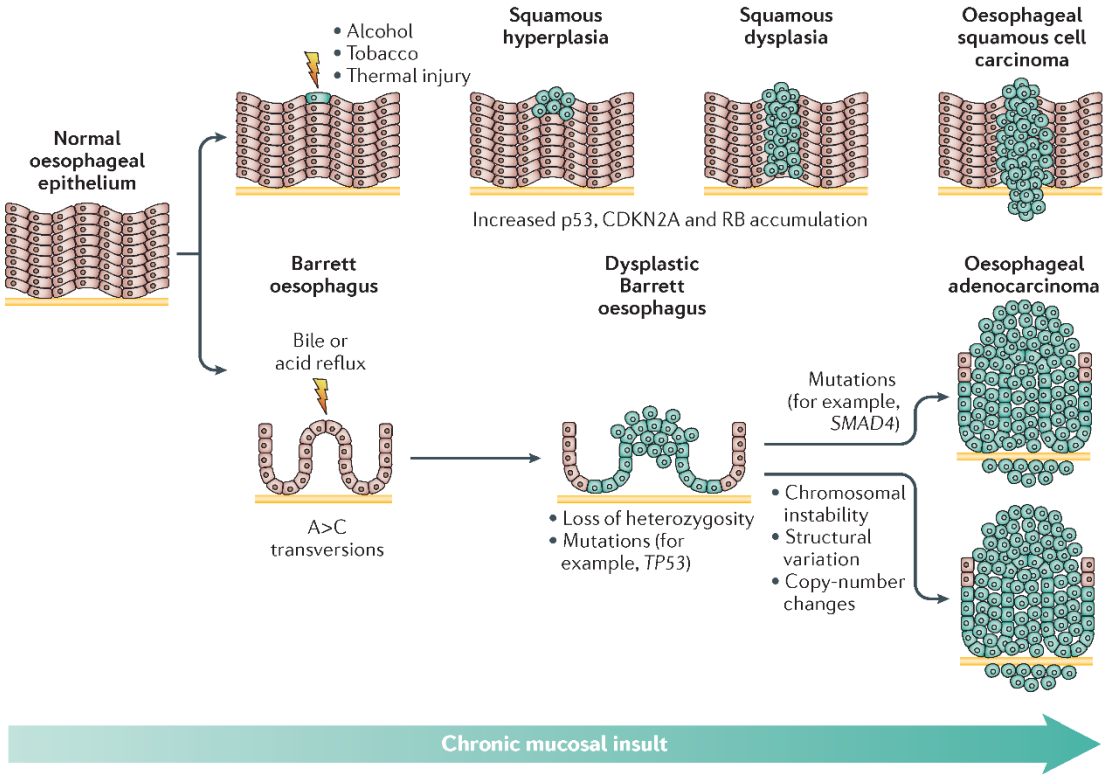
OSCC develops from the squamous epithelial cells that make up the inner lining of the oesophagus, as outlined in Figure 2. Risk factors include recurrent chemical or physical insults to the oesophageal mucosa, such as tobacco smoking and alcohol consumption, as presented in Figure 2.[14] By contrast, adenocarcinomas typically arise from Barrett's oesophagus, a condition where tissue that is similar to the lining of the intestine replaces the tissue that lines the oesophagus. Risk factors for adenocarcinomas include excess body weight and gastro-oesophageal reflux.[6,14] OSCC is more common in the upper and middle third of the oesophagus, while adenocarcinomas are more common the distal (lower) section of the oesophagus (Figure 1).[6]

Figure 1. Oesophageal cancer locations[5]



The upper part, middle part and lower part refer to the sections of the oesophagus where OSCC develops.

Figure 2. Pathogenesis of OC (reproduced from Figure 2 of Smyth et al., 2017)[6]



B.1.3.1.1 Symptoms, diagnosis and staging

Early OC often causes no signs or symptoms. Patients with OC commonly present at an advanced stage of the disease.[15] Solid food dysphagia is the primary symptom causing patients with OSCC to seek medical attention.[6] As well as dysphagia, patients may experience weight loss, pain and/or fatigue.[16]

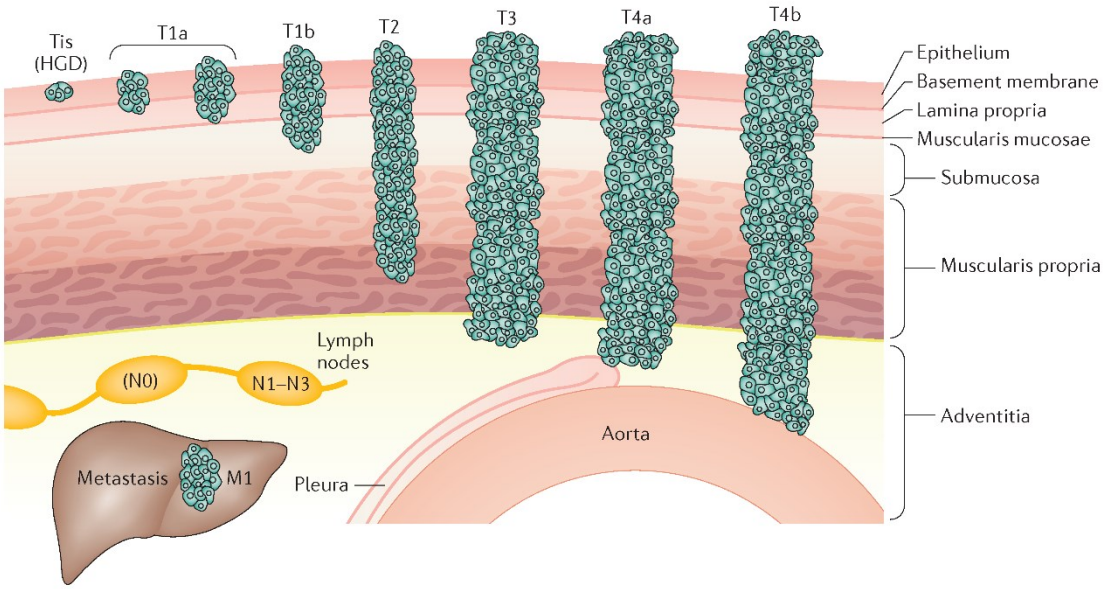
OC, including OSCC, is frequently diagnosed at an advanced stage, especially in Western countries, where early screening and prevention programs are not widely implemented.[6] OC is diagnosed by endoscopic evaluation and diagnostic imaging.[17] Differentiation between OSCC and OAC is based on histological variations identified by immunohistochemical staining of biopsy samples taken from the oesophagus.[18]

In the UK, the severity of OC is assessed using the American Joint committee on Cancer (AJCC) tumour/node/metastasis (TNM) staging system, which classifies tumours according to the amount of tumour invasion (T), involvement of the lymph nodes (N) and distant metastasis (M), as outlined in Figure 3 and Wu et al (2017).[19] Tumours can be classified by pathological stage following surgery or clinical stage after a physical exam, biopsy and imaging.[6] Patients with cT3-T4 or cN1-3 M0 disease are classified as having locally advanced disease, while M1

signifies distant metastasis.[18] Tumours are often advanced at the time of diagnosis and accurate staging is important for prognosis and treatment planning.[8]

The most common sites of metastasis include liver, distant lymph nodes, lung, bone and brain, with lung metastases more frequent in patients with OSCC and liver, bones and brain more common in patients with adenocarcinoma.[19-21] Survival in patients with advanced OC varies dependant on the site of metastasis and histological subtype, with distant lymph nodes associated with greater survival compared to those with liver, bone or lung metastases in OSCC.[19]

Figure 3. TNM staging in oesophageal cancer (reproduced from Figure 4 of Smyth et al., 2017)6



B.1.3.2 Epidemiology of OC

OC is a significant health issue worldwide. While OC is relatively rare, with 9,272 new OC diagnoses in the UK between 2016-2018, of which 7,680 cases were in England,[22] it is the seventh most common cause of cancer death in the UK and was responsible for an estimated 7,990 deaths in the UK between 2016 and 2018.[23] This highlights that survival rates for OC are extremely poor, with only ~15% of people diagnosed with OC surviving for five years or more (2013-2017).[11]

Globally, most OC cases are OSCC, however, in Western countries most OC cases are adenocarcinomas. A recent study reported that, in the UK, approximately two-thirds of OC cases are adenocarcinomas and roughly a third are OSCC.[9]

In the UK, 70-80% of patients are diagnosed with either lymph nodes or distant metastases, and 37-42% have distant metastases at diagnosis.[7]

B.1.3.3 Life expectancy

In 2020, OC caused the sixth-highest cancer death toll globally.[24] In England, fewer than half of patients diagnosed with OC (46.5%) remain alive at 12 months.[25]

The prognosis of patients with OC worsens with tumour stage, where patients with unresectable, advanced OC have poorer outcomes than those diagnosed with localised disease. In OC patients diagnosed with regional and distant disease, five-year survival is 25% and 5%, respectively, and median survival in patients diagnosed with metastatic OC is 10 months.[10,26] Several studies investigating the outcomes of patients with advanced OSCC after first-line chemotherapy demonstrated that median overall survival did not exceed one year.[27-32] Survival is also impacted by histological type, where patients with OSCC have worse survival outcomes than those with adenocarcinoma.[33] Thus, there is significant unmet need for effective therapies to improve outcomes in this patient population.

B.1.3.4 Burden of OC and unmet need

Before patients are diagnosed with OC, most patients experience dysphagia, eating difficulties and appetite loss, resulting in considerable weight loss and fatigue, impacting patient's quality of life (QoL).[6] Patients with OC have worse QoL than the general population, and compared with patients with other common cancer types, including lung, breast, liver and stomach.[34] In addition to the burden of OC symptoms, treatment of metastatic OC can cause serious toxicity and morbidity that can significantly impact patients' QoL.[35] A global retrospective study demonstrated that over half of patients with advanced OSCC who receive first-line chemotherapy reported nausea (70.9%), fatigue (63.1%), anaemia (56.4%), and/or neutropenia (55.5%).[36] Results are similar for advanced patients with OSCC treated with fluoropyrimidine-based chemotherapy; the toxicity composite endpoint (TCE, defined as the first occurrence of grade 3 or 4 diarrhoea, neutropenia, febrile neutropenia, fever, infection, nausea, and vomiting, or grade ≥2 renal or neurotoxicity) has also been shown to be high, affecting 44% (95% CI, 0.35–0.53) of patients.[37] These studies emphasise the need for additional treatment options with better improvement in patient HRQoL and limited toxicity.

OC is one of the most aggressive forms of cancer. OSCC is also more chemo-resistant than OAC. In a study pooling 973 patients with advanced, untreated gastroesophageal adenocarcinoma or OSCC (841 with OAC and 132 with OSCC) predominantly from the UK and treated with fluoropyrimidine-based chemotherapy, the overall response rate (ORR) was

44% for patients with OAC versus 33% for patients with OSCC (p=0.01).[37] Survival differed between the patients with OAC and OSCC, with median OS of 9.5 months versus 7.6 months, respectively (HR=0.85; 95% CI, 0.70–1.03, p=0.09) and one-year survivals of 38.8% vs. 28.2%, respectively.[37] A greater proportion of patients with OSCC also progressed while under treatment: 29% versus 19% of patients with OAC.[37] Guidelines from the European Society for Medical Oncology (ESMO) also support this and state that chemotherapy is less effective for OSCC than for OAC.[38] These data reinforce the need for innovative therapies, beyond chemotherapy, for patients with OSCC to improve outcomes.

OC is a major cause of disease burden worldwide. OC caused 11.7 million (95% CI, 10.4–12.9) disability-adjusted life years (DALYs) globally in 2019.[39] The majority of DALYs were attributable to years of life lost (YLL), amounting to 11.5 million (95% CI, 10.2–12.8), with years lost to disability (YLD) amounting to only 150,000 (95% CI, 107,000–196,000).[39] OC has a substantial economic burden across all disease stages and histological subgroups, due to high healthcare resource utilisation, disease morbidity, and mortality.[40,41] In the EU, a report from The Swedish Institute for Health Economics (IHE) estimated the total cost of OC across 31 European countries, including 27 EU member states, to be €3.6 billion in 2018.[41]

B.1.3.5 Current pathway of care

The stage of the patient's disease is a critical factor for treatment decisions. Patients diagnosed with early stage OC may be offered surgery, which is potentially curative; other treatments, including chemotherapy and radiotherapy, may also be appropriate depending on the extent of disease and the patient's fitness.[14] However, most patients in the UK are diagnosed at an advanced disease stage (70-80% diagnosed with either lymph nodes or distant metastases), by which time surgery may no longer be a viable treatment option.[7,14]

Globally, a retrospective analysis has shown that nearly half of patients undergoing systemic treatment for OSCC in the first-line setting do not respond to their treatment and over a third of patients progress to the next line of treatment.[42]

There is currently a high unmet need for effective first-line treatments for patients with advanced OC, with doublet palliative chemotherapy options being the current standard of care. Pembrolizumab in combination with platinum and fluoropyrimidine-based chemotherapy has recently been recommended for the treatment of untreated, locally advanced unresectable or metastatic oesophageal carcinoma in adults whose tumours express PD-L1 with a combined positive score (CPS) ≥10.[1] However, this indication does not include patients with CPS <10, who are covered by the indication for nivolumab with chemotherapy, which includes patients

with PD-L1 CPS ≥1, and therefore, there remains an unmet need for therapeutic options covering this patient population.

A summary of the National Institute for Health and Care Excellence (NICE) guidelines for the treatment of locally advanced or metastatic OC is described below:

Locally advanced or metastatic (first-line)

- For patients who have a performance status of 0 to 2 and no significant comorbidities, palliative chemotherapy with doublet (5-fluorouracil or capecitabine in combination with cisplatin or oxaliplatin) or triplet (5-fluorouracil or capecitabine in combination with cisplatin or oxaliplatin plus epirubicin) regimens is recommended for first-line systemic treatment.[12,43,44]

- Pembrolizumab in combination with platinum and fluoropyrimidine-based chemotherapy has recently been recommended for the treatment of untreated, locally advanced unresectable or metastatic oesophageal carcinoma in adults whose tumours express PD-L1 with CPS ≥10.[1]

Locally advanced or metastatic (second-line)

- Second-line palliative chemotherapy is recommended for patients who have progressed on the first-line therapy; however, specific chemotherapy regimens are not defined in the NICE clinical guidelines in the second-line setting.[12,43,44]

- Nivolumab monotherapy is also used at second line for the treatment of adult patients with unresectable advanced, recurrent or metastatic OSCC after prior fluoropyrimidine- and platinum-based combination chemotherapy.

Similar to UK guidance, guidelines from the European Society for Medical Oncology (ESMO) recommend palliative chemotherapy in the management of advanced or metastatic OSCC.[18] However, due to a lack of evidence of effectiveness, specific chemotherapy regimens are not specified.[18] During an advisory board held by BMS, there was consensus that most UK clinicians used doublet chemotherapy regimens, including 5-fluorouracil or capecitabine in combination with cisplatin or oxaliplatin. It was agreed that very few clinics currently offer triplet therapy. Oxaliplatin with capecitabine was considered standard of care, as it is better tolerated than cisplatin with 5-fluorouracil. Epirubicin is also no longer considered standard of care.[2]

B.1.3.6 Nivolumab: Mechanism of action

PD-1 is an immune checkpoint protein receptor expressed at high levels on activated T-cells. This receptor has been shown to control the inhibition of T-cell response at the effector stage of the immune response, in the setting of human malignancy.[45-49] Tumour cells can exploit this pathway by up-regulating proteins that engage PD-1 (programmed death ligand-1 (PD-L1) and programmed death ligand-2 (PD-L2)), to limit the activity of T-cells at the tumour site. In one study, 18 (43.9%) of the 41 oesophageal tumours evaluated were positive for PD-L1 or PD-L2 gene expression.[50] A similar proportion is seen in Checkmate 648, where 473 patients (48.8%) among all randomised subjects (n=970) had tumour cell PD-L1 expression of ≥1%.[51]

Through exploitation of the PD-1 immune checkpoint inhibitor pathways, OC cells are able to escape immune surveillance. Hence, PD-1 and its ligands may be considered as therapeutic targets for immune-mediated therapies in OC.

Nivolumab contains the humanised, monoclonal immunoglobulin antibody (IgG4). Nivolumab binds to the PD-1 receptor and blocks its interaction with PD-L1 and PD-L2. Nivolumab is a checkpoint inhibitor of the PD-1 mediated T-cell response pathway.[52]

B.1.3.6.1 Nivolumab: Pseudo-progression in response to checkpoint inhibitor therapy

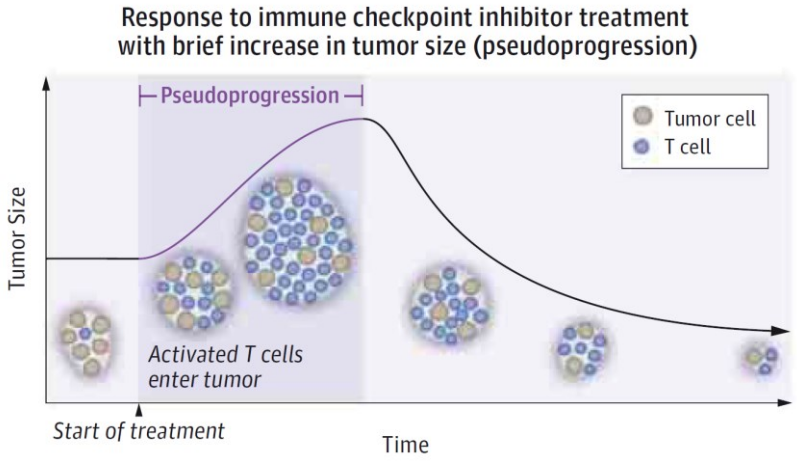
Conventional anti-cancer therapies typically aim to reduce the tumour burden through direct disruption of tumour cell proliferation or induction of apoptosis. By contrast, the novel mechanism of action of immunotherapies like nivolumab can result in different patterns of response, including pseudo-progression.[53]

Due to the indirect anti-tumour mechanism associated with immunotherapies, where host immune cells are recruited to the tumour site, the initial effect of immunotherapy may present as increased size of existing lesions or formation of new lesions that result from the infiltration of tumour-specific immune cells and other inflammatory cells (“pseudo-progression,”

Figure 4).54-56 This brief initial enlargement of the tumour may be followed by tumour shrinkage or eradication.[54,55]

Due to the delayed clinical response to immunotherapies, the “time to response” from immunotherapy treatment may differ from that seen after conventional chemotherapy.[56] These differences in response patterns after immunotherapy may be prematurely misclassified as disease progression under the WHO or RECIST criteria.[55,56]

Figure 4. Pseudo-progression response to immune checkpoint inhibitor treatment[54]

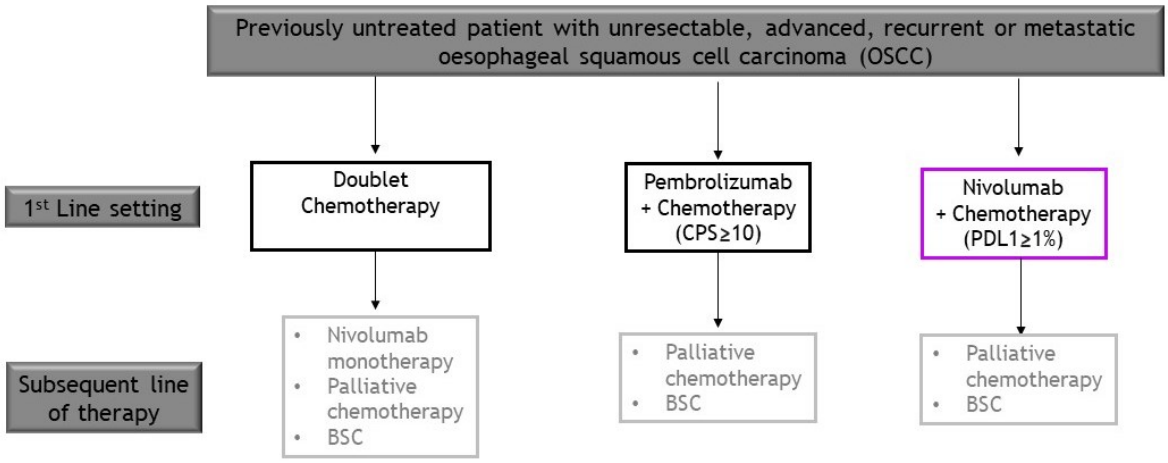


B.1.3.7 Nivolumab with chemotherapy within the current clinical pathway

While chemotherapy is the standard of care at first-line for unresectable, advanced recurrent or metastatic OSCC, it has drawbacks including lack of durable efficacy and toxicity.[57] There is no clear evidence indicating that, in the first-line setting, chemotherapy prolongs the survival of patients with advanced OC compared with best supportive care.[58] An opportunity exists to redefine the clinical pathway for patients with OSCC by offering a therapy targeted to specific molecular mechanisms of the tumour pathology. Nivolumab with chemotherapy would provide a better treatment option for patients with unresectable, advanced, recurrent or metastatic OSCC with improved survival outcomes compared to chemotherapy and a manageable tolerability profile.

The proposed positioning of nivolumab with chemotherapy for patients with OSCC is presented in Figure 5 with efficacy and safety supporting this change outlined in the subsequent sections.

Figure 5. Anticipated positioning of nivolumab with chemotherapy in the current treatment pathway



B.1.4 Equality considerations

It is not considered that this appraisal will exclude any people protected by equality legislation; or lead to a recommendation that would have a different impact on people protected by equality legislations than on the wider population; or lead to recommendations that would have an adverse impact on people with a particular disability.

B.2 Clinical effectiveness

Key Points

- In the Checkmate 648 trial, nivolumab with chemotherapy (NIVO-CHEMO) significantly extended the median OS of patients with OSCC to [redacted] months vs. [redacted] months) with chemotherapy (CHEMO) alone in all randomised patients. At 18 months, among all randomised patients, the OS rate in the NIVO-CHEMO arm was [redacted] compared to [redacted] in patients who received CHEMO only.
- The survival benefit was more marked in the PD-L1 ≥1% population, where there was a 6-month improvement in median OS in the NIVO-CHEMO group vs. CHEMO ([redacted] months vs. [redacted] months).
- The safety profiles for patients treated with NIVO-CHEMO were consistent with those of the individual agents, with no new safety signals identified. [redacted]% of patients (n=56) in the NIVO-CHEMO arm discontinued due to serious grade 3-4 adverse events compared to [redacted]% (n=33) of patients in the CHEMO arm.
- Patients treated with NIVO-CHEMO reported a similar increase from baseline at most-on treatment assessments for EQ-5D-3L utility index, EQ-5D-3L VAS, FACT-E and FACT-G compared to patients treated with CHEMO alone.
- NIVO-CHEMO meets the end-of-life criteria in the patient group that would be eligible for treatment under the proposed indication.

B.2.1 Identification and selection of relevant studies

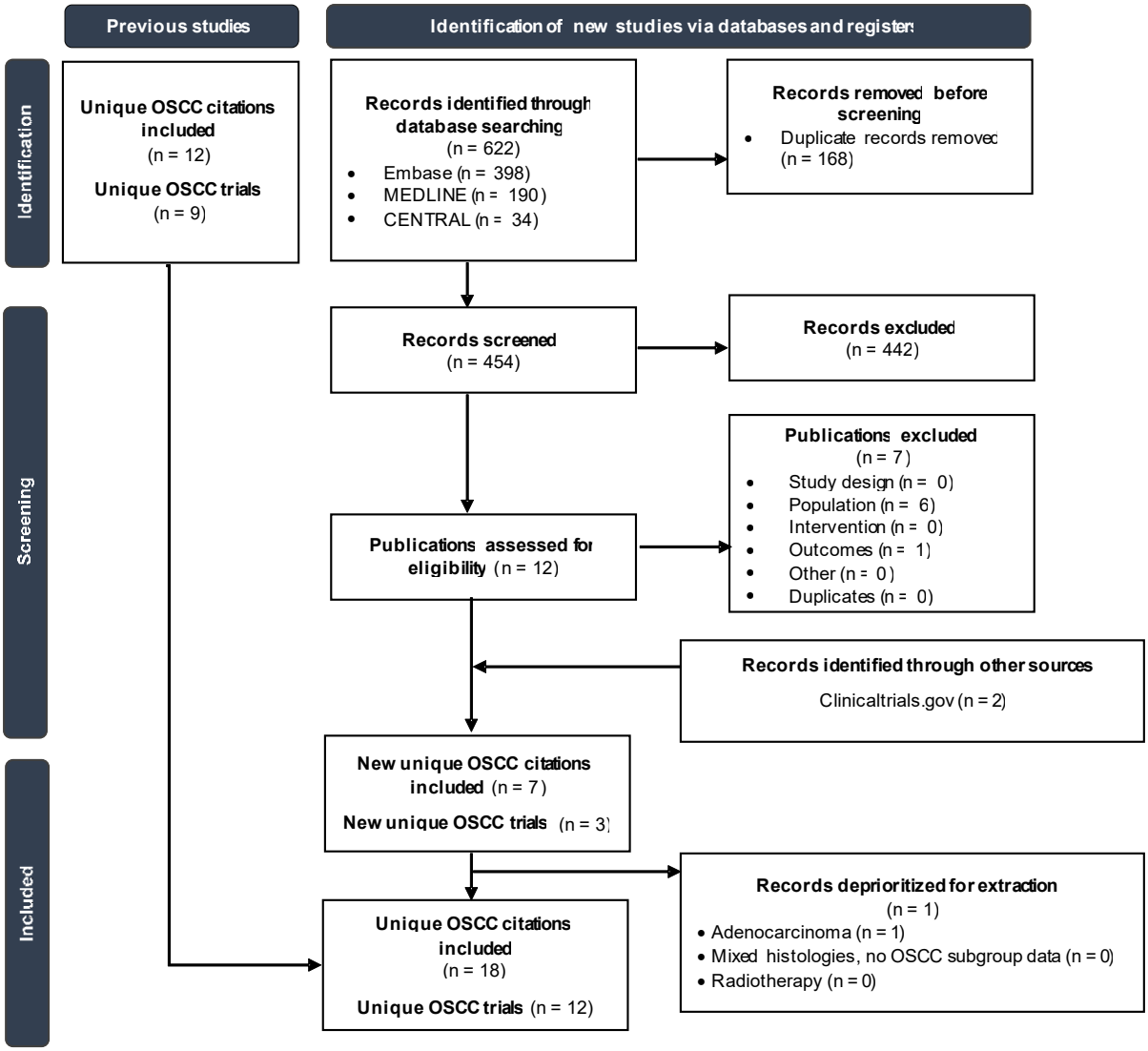
A systematic literature review (SLR) was undertaken to identify the clinical effectiveness evidence (efficacy and safety) of interventions for the treatment of unresectable advanced recurrent or metastatic previously untreated OSCC. Searches were originally run on January 14, 2021, and updated searches were run on October 4, 2021. Relevant studies were identified by searching the following databases: Medical Literature Analysis and Retrieval System Online (MEDLINE, via Ovid), Excerpta Medica dataBASE (Embase, via Ovid) and Cochrane Central Register of Controlled Trials (via Cochrane Library). Conference proceedings from 2019–2021 were searched using Northern Lights to identify relevant publications from the following conferences: American Society of Clinical Oncology (ASCO),

American Society of Clinical Oncology Gastrointestinal (ASCO GI), European Society for Medical Oncology (ESMO) and European Society for Diseases of the Esophagus (ESDE). When records were not indexed on Northern lights, conference proceedings were hand searched. Full details of the methods and processes employed to identify and select the relevant clinical evidence are summarised in Appendix E.

The SLR identified 39 unique randomised controlled trials, across 57 publications. Of these, 45 publications representing 30 trials were excluded as they did not report outcomes for patients with OSCC or evaluated radiotherapy. Therefore, a total of 18 publications describing 12 trials were included in the clinical SLR. The selection process is outlined in Figure 6.

All 12 included studies were randomised controlled trials (RCTs), five of which were phase II and six of which were phase III (one did not report study phase). Six trials were open label, four were double-blind and two did not report the blinding details. Three trials were international, two trials were conducted in multiple European countries and seven trials were conducted in a single country (China, Germany, South Korea, and US). None of the identified studies were conducted in the UK.

Figure 6. PRISMA flow diagram for the clinical SLR



B.2.2 List of relevant clinical effectiveness evidence

Evidence to support the effectiveness of nivolumab in combination with chemotherapy for the treatment of unresectable advanced, recurrent or metastatic, previously untreated OSCC is derived primarily from CheckMate-648 (NCT03143153), shown in Table 3.

Checkmate 648 is ongoing, and future analyses will provide long-term efficacy and safety evidence for nivolumab with chemotherapy in OSCC.

Table 3. Clinical effectiveness evidence

Study	CheckMate 648		
Study design	Phase III, multicentre, randomised, open-label study		
Population	Adult patients with unresectable advanced, recurrent or metastatic, previously untreated oesophageal squamous cell carcinoma.		
Intervention(s)	NIVO-CHEMO: Nivolumab at a dose of 240 mg administered intravenously (IV) over 30 minutes every two weeks, with fluorouracil and cisplatin administered every four weeks		
Comparator(s)	CHEMO: Fluorouracil with cisplatin administered every four weeks		
Indicate if trial supports application for marketing authorisation	Yes	**Indicate if trial used in the economic model**	Yes
Rationale for use/non-use in the model	Source of direct comparative evidence evaluating the efficacy of nivolumab with chemotherapy versus chemotherapy alone in the indicated patient population.		

Reported outcomes specified in the decision problem	• **Overall survival per BICR for all patients and patients with tumour cell PD-L1 ≥1%**
	• **Progression-free survival per BICR for all patients and patients with tumour cell PD-L1 ≥1%**
	• Overall survival per investigator for all patients and patients with tumour cell PD-L1 ≥1%
	• Progression-free survival per investigator for all patients and patients with tumour cell PD-L1 ≥1%
	• Response rate (ORR, DOR, PFS2/TSST)
	• **Adverse events**
	• Health-related quality of life
	  ○ **EQ-5D-3L utility index**
	  ○ EQ-5D-3L VAS
	  ○ FACT-E		

DOR: duration of response; IV: intravenous; OC: oesophageal cancer; ORR: objective response rate; PFS2: time to second progression; TSST: time to second subsequent therapy; VAS: visual analogue scale.
Note: Outcomes in bold are included in the economic model.
Source: Clinicaltrials.gov,[59] CheckMate 648 study protocol,[4]

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

B.2.3.1 Study design

A summary of the methodology for CheckMate-648 is provided in Table 4, with further details provided in the study protocol.[4]

Table 4. Summary of trial methodology

Trial number (acronym)	CheckMate 648
Location	USA, Argentina, Austria, Brazil, Canada, Chile, China, Colombia, Czechia, Denmark, France, Hong Kong, Italy, Japan, Korea, Mexico, Peru, Poland, Romania, Russia, Singapore, Spain, Taiwan, Turkey, UK (5 centres in the UK, included 34 randomised patients)
Trial design	Phase III, multicentre, randomised, open-label trial (ongoing)
Eligibility criteria for participants	Adult patients with unresectable advanced, recurrent or metastatic, previously untreated oesophageal squamous cell carcinoma.
Trial drugs	Intervention 1 (n = 321): NIVO+CHEMO: nivolumab 240 mg Q2W IV + fluorouracil 800 mg/m2/day IV on Day 1 through Day 5 + cisplatin 80 mg/m2 IV on Day 1 of a 4-week cycle Intervention 2 (n = 325): NIVO+IPI nivolumab 3 mg/kg every 2 weeks (Q2W) intravenously (IV) + ipilimumab 1 mg/kg every 6 weeks (Q6W) IV **Please note that intervention 2 is not part of this submission.**
	Comparator arm (n = 324*): CHEMO: fluorouracil 800 mg/m2/day IV Day 1 through Day 5 + cisplatin 80 mg/m2 IV on Day 1 of a 4-week cycle
Permitted and disallowed concomitant medications	Disallowed: The following medications are prohibited during the study (unless utilised to treat a treatment-related adverse event): • Immunosuppressive agents • Immunosuppressive doses of systemic corticosteroids (some exemptions – see “Permitted”) • Any concurrent anti-neoplastic therapy (i.e., chemotherapy, hormonal therapy, immunotherapy, extensive, non-palliative radiation therapy, or standard or investigational agents for the treatment of OC). • Botanical formulations with an approved indication for cancer treatment [e.g., traditional Chinese medicines]; these should be discontinued (if used) at least 2 weeks prior to randomisation. • Any live / attenuated vaccine (e.g., varicella, zoster, yellow fever, rotavirus, oral polio and measles, mumps, rubella [MMR]) during treatment and until 100 days post last dose.
	Permitted: • Topical, ocular, intra-articular, intranasal and inhalational corticosteroids, with minimal systemic absorption. • Adrenal replacement steroid doses (>10 mg daily prednisone). • A brief (< 3 weeks) course of corticosteroids for prophylaxis or for treatment of non-autoimmune conditions is permitted. • Regular concomitant use of bisphosphonates and RANK-L inhibitors for prevention or reduction of skeletal-related events in patients with bone metastasis is allowed if initiated prior to first dose of study therapy. Palliative radiotherapy was permitted for patients without evidence of progression

	per RECIST 1.1 provided the lesions were non-target lesions and this was discussed and approved by the BMS Clinical Trial Physician (Medical Monitor). Patients with evidence of progression per RECIST 1.1 must have met criteria to continue treatment beyond progression in order to resume immunotherapy after palliative local therapy.
Pre-planned subgroups	• Age (< 65, ≥ 65 and ≥ 75) • Sex • Region (Asia and non-Asia) • ECOG PS (0 and 1) • Number of organs with metastasis (≤ 1 and ≥ 2) • Disease stage at current diagnosis • Smoking status • Alcohol use PD-L1 CPS subgroups: • ≥1%, ≥5% and ≥10%

* Patient numbers are number of patients randomised and not the number who received treatment
AE: adverse event; BMS: Bristol-Myers Squibb; ECOG: Eastern Oncology Cooperative Group; IV: intravenous; MMR: measles, mumps, rubella; RANK-L: Receptor activator of nuclear factor kappa-B ligand; RECIST: Response Evaluation Criteria in Solid Tumours; OC: oesophageal cancer; PD-L1: programmed death ligand 1; PFS2: time to second progression; TSST: time to second subsequent therapy
Source: Clinicaltrials.gov,[59] CheckMate 648 study protocol,[4] Chau et al. 2021[60]

CheckMate 648

Study design

CheckMate 648 is a phase III controlled study of nivolumab with fluorouracil plus cisplatin (NIVO-CHEMO) vs. fluorouracil plus cisplatin (CHEMO) in patients with unresectable advanced, recurrent or metastatic previously untreated OSCC (NCT03143153).[59] The objective of the study was to evaluate the efficacy and safety of NIVO-CHEMO in this patient population. The trial was initiated on June 29, 2017 and recruited patients at centres in multiple countries. There were [redacted] UK centres in the trial that recruited [redacted] randomised patients.

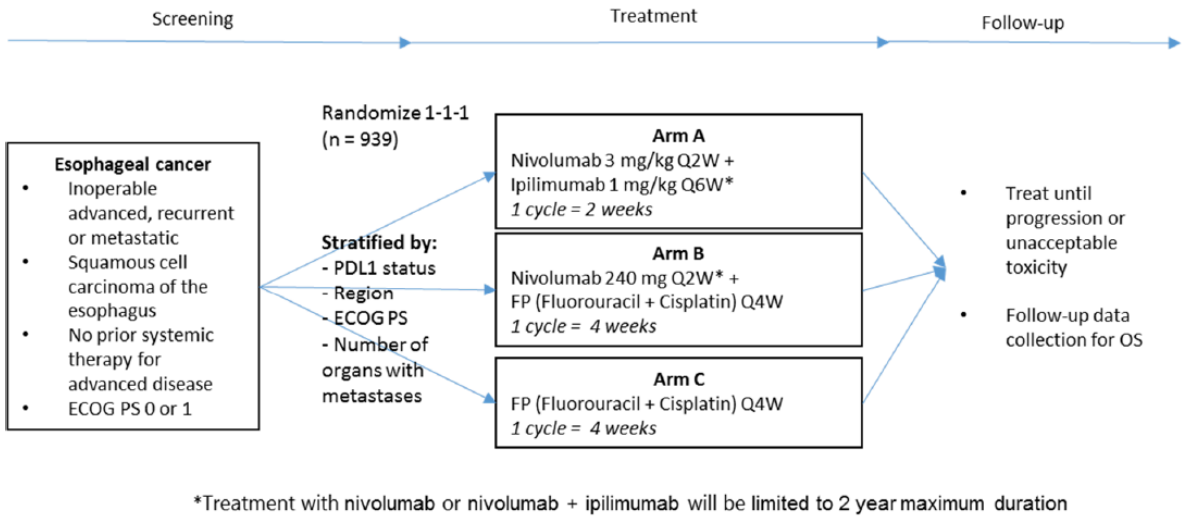
Patients were randomised in a 1:1:1 ratio to treatment with:

- Nivolumab (240 mg every two weeks IV) with fluorouracil plus cisplatin every four weeks,
- Nivolumab (3 mg/kg every two weeks) with ipilimumab (1 mg/kg every six weeks),
- Fluorouracil plus cisplatin every four weeks

Please note, that while CheckMate 648 also included a cohort who received nivolumab plus ipilimumab, this is outside the scope of the proposed indication. As such, results are only presented for the cohorts relevant to the proposed indication: the NIVO-CHEMO and CHEMO only arms of the CheckMate 648 study.

Patients were treated with nivolumab and chemotherapy for up to 24 months in the absence of disease progression or unacceptable toxicity.[4] Chemotherapy was given until disease progression or unacceptable toxicity. Randomisation was stratified by tumour cell PD-L1 status (≥ 1% vs < 1% or indeterminate/non-evaluable), region (East Asia [Japan, Korea, Taiwan] versus the rest of Asia versus the rest of the world), ECOG performance status (0 versus 1) and the number of organs with metastasis (≤1 versus ≥2). The study design of CheckMate 648 is provided in Figure 7.

Figure 7. Study design of CheckMate 648



ECOG: Eastern Cooperative Oncology Group; OS: overall survival; PD-L1: programmed death ligand 1; PS: performance status; Q2W: every 2 weeks; Q4W: every 4 weeks, Q6W: every 6 weeks.
Source: CheckMate 648 study protocol,[4]

B.2.3.2 Eligibility criteria

Patients with unresectable advanced, recurrent or metastatic, previously untreated OSCC were enrolled and randomised post-selection. Key eligibility criteria for patients in CheckMate 648 are provided in Table 5; please see the trial protocol for a full list of inclusion and exclusion criteria.[4]

Table 5. Key inclusion and exclusion criteria for CheckMate 648

Key inclusion criteria	Key exclusion criteria
• Male or female at least 18 years of age • Must have histologically confirmed squamous cell carcinoma or adenosquamous cell carcinoma of the oesophagus (predominant squamous differentiation) • Patients must have unresectable advanced, recurrent or metastatic OSCC • Patients must not be amenable to curative approaches such as definitive chemoradiation and/or surgery • No prior systemic or anticancer therapy given as primary therapy for advanced, metastatic disease • ECOG performance status of 0 or 1 • Patients must have at least one measurable lesion by CT or MRI per RECIST 1.1 criteria (radiographic tumour assessment must be performed within 28 days prior to randomisation) • Tumour tissues must be provided for biomarker analyses • Patient must have PD-L1 expression classification ≥1% or <1% or indeterminate as determined by the central lab.	• Patients must have recovered from the effects of major surgery or significant traumatic injury at least 14 days before randomisation • Prior malignancy requiring active treatment within the previous 3 years except for locally curable cancers that have been apparently cured, such as basal or squamous cell skin cancer, superficial bladder cancer or carcinoma in situ of the prostate, cervix or breast • Patients with active, known or suspected autoimmune disease. Patients with Type I diabetes mellitus residual hypothyroidism due to autoimmune thyroiditis only requiring hormone replacement, skin disorders not requiring systemic treatment are permitted to enrol • Patients with a condition requiring systemic treatment with either corticosteroids (>10 mg daily prednisone equivalent) or other immunosuppressive medications within 14 days of start of study treatment. Inhaled or topical steroids and adrenal replacement steroid doses >10 mg daily prednisone equivalent are permitted in the absence of active autoimmune disease • Prior treatment with an anti-PD-1, anti-PD-L1, anti-PD-L2, anti-CD137 or anti-CTLA-4 antibody or any other antibody or drug specifically targeting T-cell co-stimulation or checkpoint pathways
Source: CheckMate 648 study protocol,[4]	

B.2.3.4 Study endpoints and assessments

The primary, secondary and exploratory endpoints of CheckMate-648 are provided in Table 6. Co-primary endpoints were overall survival (OS) and progression-free survival (PFS) by blinded independent central review (BICR) in patients with tumour-cell PD-L1 expression ≥1%. OS was defined as the time between the date of randomisation and the date of death due to any reason. For patients without documentation of death, OS was censored on the last date the patient was known to be alive.

PFS, as assessed by BICR, was defined as the time from randomisation to the date of the first documented progressed disease (PD) or death due to any cause. Patients who died without

a reported prior PD per BICR, and died without start of subsequent therapy, were considered to have progressed on the date of death. Patients who did not have documented PD per BICR per RECIST 1.1 criteria and who did not die, were censored at the date of the last evaluable tumour assessment on or prior to initiation of the subsequent anti-cancer therapy. Patients who did not have any on-study tumour assessments and did not die (or died after initiation of the subsequent anti-cancer therapy) were censored at the randomisation date. Patients who started any subsequent anti-cancer therapy without a prior reported PD per BICR were censored at the last tumour assessment on or prior to initiation of the subsequent anti-cancer therapy.

Table 6. Study endpoints in CheckMate 648

CheckMate 648 study outcomes	
Primary endpoints	• Overall survival (OS) in patients with PD-L1 expression ≥1%. • Progression-free survival (PFS) in patients with PD-L1 expression ≥1%.
Secondary and exploratory endpoints	**Secondary endpoints**: • OS in all randomised patients • PFS in all randomised patients • Objective response rate (ORR) in patients with PD-L1 expression ≥1% and in all randomised patients **Key exploratory endpoints:** • Safety and tolerability: o Incidence of: ▪ Adverse events (AEs), ▪ Serious adverse events (SAEs), ▪ AEs leading to discontinuation ▪ AEs leading to dose modification ▪ Select AEs ▪ Immune-mediated adverse events ▪ Other events of special interest (OESI) ▪ Deaths o Laboratory abnormalities • PFS as assessed by investigators in patients with PD-L1 expression ≥1% and all randomised patients • ORR as assessed by investigators in patients with PD-L1 expression ≥1% and all randomised patients • Duration of response (DOR) as assessed by BICR and by investigators in patients with PD-L1 expression ≥1% and all randomised patients • PFS2/TSST in patients with PD-L1 expression ≥1% and all randomised patients. • Quality of life, measured using the E5-5D-3L descriptive system and VAS, as well as the FACT-E questionnaire (including the Esophageal Cancer Subscale [ECS] and FACT-G7) Please see the study protocol for further exploratory endpoints, including biomarker analysis, immunogenicity, and pharmacokinetics[4]

AE: adverse event; BICR: blinded independent central review; BOR: best overall response; CR: complete response; DOR: duration of response; ECOG PS: Eastern Cooperative Oncology Group Performance Score; EQ-5D-3L: EuroQol questionnaire comprising 5 dimensions, with each dimension having 3 levels; FACT-E: Functional Assessment of Cancer Therapy-Esophageal; FACT-G7: 7-item version of FACT-General; ORR: objective response rate; OS: overall survival; PD: progressive disease; PD-L1: programmed death ligand 1; PFS: progression-free survival; PFS2: progression free survival after the next line of the subsequent therapy; PR: partial response; SAE: serious adverse event; TSST: time to second subsequent therapy;
Source: CheckMate 648 protocol[4]

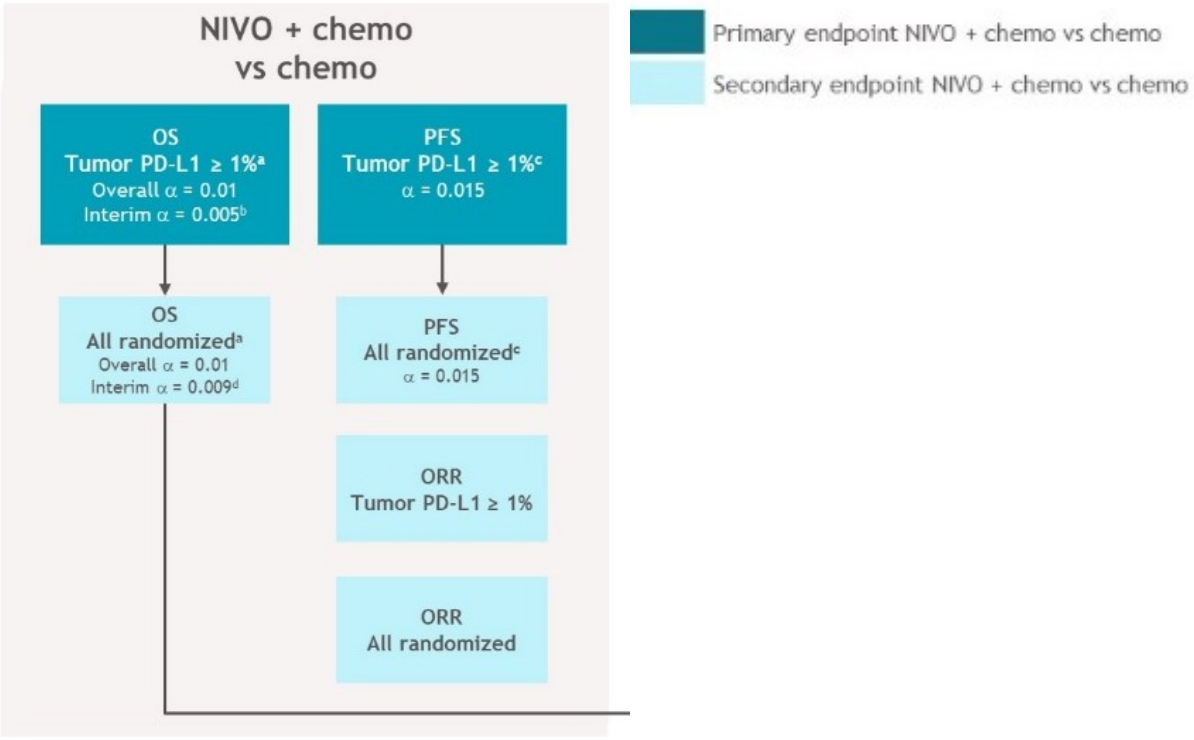
B.2.4 Statistical analysis and definition of study groups in the relevant clinical effectiveness evidence

B.2.4.1 Checkmate 648: Objectives and endpoints

The primary, secondary and exploratory outcomes of the CheckMate 648 trial are defined in

Table 6. An overview of the statistical testing is provided in Figure 8.

Figure 8. Summary of statistical testing of outcomes from CheckMate 648



All endpoints in the top row were tested first and in parallel. The secondary endpoints were tested hierarchically only if the corresponding primary endpoints above were significant. 100% of α was passed from successful hypotheses to next endpoint(s) as indicated by the arrows.

B.2.4.2 Sample size and power calculation

Sample size calculations assumed that the prevalence of patients with PD-L1 ≥1% was approximately 50%, and the proportion of subjects with (≥ 1%) or without (<1% or indeterminate) PD-L1 expression was monitored during enrolment.

The study sample size was based on the primary objectives. For both experimental arms, the same OS and PFS distributions were assumed. A piecewise mixture cure rate model was used for the design setup, with cure rates in the experimental arms of 15% for OS in PD-L1 ≥1%, 10% for OS in PD-L1 <1%, and 0% for PFS per BICR. As a result, for the NIVO-CHEMO vs CHEMO comparisons:

- [redacted] PFS events in approximately 313 subjects with PD-L1 ≥1% would provide approximately 90% power to detect an average hazard ratio (HR) of 0.62 with a Type I error of 1.5% (two-sided).

- [redacted] OS events in approximately 313 subjects with PD-L1 ≥1% would provide approximately 90% power to detect an average HR of 0.6 with a Type I error of 1% (two-sided).

In case the significance level from the corresponding primary endpoint in patients with PD-L1 ≥1% was passed to the secondary endpoint in all randomised subjects:

- [redacted] PFS events in approximately 626 patients (all comers) would provide approximately 90% power to detect an average HR of 0.72 with a Type I error of 1.5% (two sided);

- [redacted] OS events in approximately 626 patients (all comers) would provide approximately 94% power to detect an average HR of 0.68 with a Type I error of 1% (two sided).

To have approximately 313 randomised patients with PD-L1 ≥1% for each comparison, approximately 470 patients with PD-L1 ≥1% needed to be randomised in a 1:1:1 ratio in the 3 arms. This translated to a total of approximately 939 patients (with any PD-L1 result) to be randomised in a 1:1:1 ratio to the NIVO-IPI, NIVO-CHEMO or CHEMO arms. Assuming a piecewise constant accrual rate, it was estimated that these 939 patients would be accrued within 29 months.

B.2.4.2 Timing of analysis of primary endpoints

Although the same treatment effect was assumed for the comparison of NIVO-CHEMO with the control arm (CHEMO), observed treatment effects may vary. Therefore, the primary outcomes (OS, PFS) observed in the CHEMO arm only were used to determine the timing of the interim and final efficacy analyses.

Final PFS analysis was planned when [redacted] events by BICR were observed among the patients with PD-L1 expression ≥ 1% in the CHEMO arm. This was expected to be reached after approximately 33 months.

Final OS analysis was planned when [redacted] events were observed among the patients with PD-L1 expression ≥ 1% in the CHEMO arm. This was expected to be reached after approximately 49 months.

However, Revised Protocol 05 specified that if the planned number of PFS events per BICR was unlikely to be reached for unforeseen reasons, the final PFS per BICR analysis could

occur when at least 12 months minimum follow-up (defined as the time from the date when the last patient was randomised to the clinical cut-off date) was reached. Indeed, the primary analyses of final PFS per BICR and interim analysis of OS in all randomised subjects with tumour cell PD-L1 expression ≥ 1% were triggered on the basis of achieving 12 months minimum follow-up. Given the study outcomes at that time, the OS interim analysis (IA) is considered as the OS final analysis.

As planned, OS and PFS in all randomised subjects were tested formally only if significance level was to be passed on them.

B.2.4.3 Protection of Type I error across primary and secondary endpoints

The co-primary and secondary endpoints were tested using the Bonferroni-based graphical approach by Maurer and Bretz (2013).[61]

NIVO-CHEMO vs CHEMO:

- **For PFS:** since the primary endpoint of PFS in patients with PD-L1 ≥1% was significant at the 2-sided alpha level 0.015 (p-value: 0.0023), then the secondary endpoint of PFS in all randomised patients was tested with the 2-sided alpha level 0.015 passed from the primary endpoint. Since the secondary endpoint of PFS was not significant at the 2-sided alpha level 0.015 (p-value: 0.0355), the subsequent secondary endpoints ORR in all randomised patients with PD-L1 ≥1% and in all randomised patients were not formally tested.

- **For OS in patients with PD-L1≥1:** the observed number of OS events in patients with PD-L1 ≥1% at interim analysis was 219 [87.6% of the target of 250 OS events]. With the initial allocated overall alpha of 0.01, the significance level was 0.005 for OS IA using O'Brien-Fleming alpha spending function.

- **For OS in all randomised patients:** Since the primary endpoint of OS was significant at the IA 2-sided alpha level 0.005 (p-value<0.0001), then the secondary endpoint of OS in all randomised patients was tested with the overall 2-sided alpha level of 0.01 passed from the primary endpoint. The observed number of OS events in all randomised patients at IA was 441 [85.8% of the target of 514 OS events]. With the overall alpha of 0.01, the significance level was 0.009 for OS IA in all randomised patients using Pocock alpha spending function.

B.2.4.4 Analysis of primary endpoints

OS and PFS (as assessed by BICR) in all subjects with PD-L1 ≥ 1% were compared between NIVO-CHEMO and CHEMO using a two-sided log-rank test, stratified by the following stratification factors:

- ECOG performance status (0 vs. 1)
- Number of organs with metastases (≤ 1 vs. ≥ 2)

Note that although randomisation of the study population was stratified by region (East Asia vs Rest of Asia vs Rest of World), region was excluded from all stratified analyses due to small sample size in Rest of Asia.

For each comparison, the HRs of PFS per BICR and OS with its associated two-sided 100(1-α)% confidence intervals (CIs) were estimated via a stratified Cox model with treatment arm as the only covariate in the model.

Median OS and PFS for each treatment arm were estimated and plotted using the Kaplan-Meier (KM) product-limit method. Median OS and PFS along with 95% CIs were constructed based on a log-log transformed CI for the survival function.

Additional analyses of OS and PFS

Additional analyses of OS and PFS included the following:

- Assessment of consistency of treatment effects in different subsets via a “forest” plot of the OS and PFS unstratified HR (and 95% CI) in the following subgroups: age category, sex, race, region, ECOG PS, weight category, disease stage at initial diagnosis, histologic grade at initial diagnosis, histological classification at initial diagnosis, location at initial diagnosis, disease status at current diagnosis, smoking status, alcohol use, number of organs with metastases at baseline, time from initial disease diagnosis to randomisation, prior surgery (excluding biopsy), and prior radiotherapy.
- OS and PFS rates at 3, 6, 9, and 12 months estimated using KM estimates on the OS and PFS curves for each randomised arm, with associated two-sided 95% CIs calculated using Greenwood’s formula. Minimum follow-up must have been approximately longer than or equal to the timepoint to generate the rate.

B.2.4.5 Analysis of secondary endpoints

There are four secondary endpoints:

- OS in all randomised patients
- PFS by BICR in all randomised patients
- ORR by BICR in all randomized patients with PD-L1 ≥ 1%
- ORR by BICR in all randomised patients

Analyses for each of these endpoints were performed by treatment group as randomised.

OS and PFS

If any of the primary endpoints was significantly superior, the corresponding secondary endpoint of OS and PFS per BICR in all randomised subjects was compared using a two-sided log-rank test at the allocated significance level, stratified by:

- All randomised patients: ECOG PS, number of organs with metastases, and PD-L1 expression (≥ 1% vs < 1% or indeterminate)

For each comparison, the HR with its associated two-sided 95% CI was estimated via a stratified Cox model with treatment arm as the only covariate in the model. OS and PFS for each treatment arm were estimated and plotted using the KM product-limit method. Median OS and PFS with associated two-sided 95% CI were constructed based on a log-log transformed CI for the survival function.

The same additional analyses were carried out for OS and PFS in all randomised patients as for OS and PFS in all randomised patients with PD-L1 ≥1%.

ORR

ORR (as assessed by BICR) in patients with PD-L1 ≥1% and in all randomised patients was to be tested only if significance level is passed on them. ORR was computed in each treatment group along with the exact 95% CI using Clopper-Pearson method. An estimate of the difference in ORRs and corresponding 95% CI were calculated using Cochran-Mantel-Haenszel (CMH) methodology and adjusted by the stratification factors. The stratified odds ratios (Mantel-Haenszel estimator) between the treatments were provided along with the 95% CI.

B.2.4.6 Safety analysis

Safety analyses were performed for all treated patients by treatment group, unless otherwise specified. AEs were coded using the Medical Dictionary for Regulatory Activities (MedDRA) Version 23.1. AEs and laboratory values were graded for severity according to the National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE) version 4.0. All on-study AEs, treatment-related AEs, SAEs, and treatment-related SAEs were tabulated using worst grade per NCI CTCAE version 4.0 criteria by System Organ Class (SOC) and Preferred Terms (PT). In the AE summary tables, unless otherwise specified, subjects were counted only once at the PT, only once at the SOC, and only once at the subject level for the counting of total number of subjects with an AE.

B.2.5 Quality assessment of the relevant clinical effectiveness evidence

The clinical effectiveness evidence provided in this submission is derived from a large phase III trial conducted in line with the requirements of regulatory bodies. The complete quality assessment of CheckMate 648 is summarised in

Table 7. A quality assessment of the trials identified during the clinical SLR was conducted based on the Centre for Reviews and Dissemination's (CRD's) guidance and used to inform the indirect treatment comparison (ITC); additional detail is provided in Appendix E.

Table 7. Quality assessment results for CheckMate 648

	CheckMate 648 (NCT03143153)
Was randomisation carried out appropriately?	Yes, all eligible patients were randomised in a 1:1:1 ratio using interactive response technology. Randomisation was stratified by PD-L1 status (≥1% or <1%), region (East Asia [Japan, Korea, Taiwan], rest of Asia and rest of world), ECOG performance status (0 or 1), and the number of organs with metastasis (≤1 or ≥2).
Was the concealment of treatment allocation adequate?	No, the study was open label as a safety measure, so that prompt and accurate assessment of the unique toxicities associated with study treatments could be conducted.
Were the groups similar at the onset of the study in terms of prognostic factors?	Yes, the baseline characteristics of the two treatment arms were generally balanced (see Table 9).
Were the care providers, participants and outcome assessors blind to treatment allocation?	No, the study was open label as a safety measure, so that prompt and accurate assessment of the unique toxicities associated with study treatments could be conducted.
Were there any unexpected imbalances in dropouts between groups?	No, a similar number of patients discontinued in both study arms (see Table 8).
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No, all measured outcomes have been reported.
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes, an appropriate ITT analysis was conducted and the methods to account for missing data were also appropriate.
Adapted from Systematic reviews: CRD's guidance for undertaking reviews in health care (University of York Centre for Reviews and Dissemination)[62]	

B.2.6 Clinical effectiveness results of the relevant trials

Evidence for the clinical efficacy of nivolumab with chemotherapy is derived from the CheckMate 648 study, a phase III randomised trial. The design and methodology for CheckMate 648 are described in Section B.2.3.

B.2.6.1 CheckMate 648: Patient disposition

A total of 1,358 patients were enrolled and 970 were randomised to receive either nivolumab with chemotherapy (n=321), nivolumab with ipilimumab (n=325) or chemotherapy alone (n=324). In the NIVO-CHEMO arm, 11 (3.4%) patients were randomised but not treated, compared to 20 (6.2%) in the CHEMO arm.[13] At the database lock in [redacted], 11 (4%) patients in the NIVO-CHEMO arm were continuing treatment, compared to 0 (0%) in the CHEMO arm.

A summary of patient disposition is provided in Table 8.

Table 8. CheckMate 648: patient exposure and disposition ([redacted])

	NIVO-CHEMO	CHEMO
Number of patients (randomised) n	321	324
Number of treated patients n	310	304
Discontinued treatment n (%)	[redacted]	[redacted]
Disease progression (%)	[redacted]	[redacted]
AE related to treatment (%)	[redacted]	[redacted]
AE not related to treatment (%)	[redacted]	[redacted]
Patient request (%)	[redacted]	[redacted]
Other* (%)	[redacted]	[redacted]
Median duration of treatment (range), months	[redacted]	[redacted]
>3 months	[redacted]	[redacted]
>6 months	[redacted]	[redacted]
>9 months	[redacted]	[redacted]
>12 months	[redacted]	[redacted]

AE: adverse event; CHEMO: chemotherapy; NIVO-CHEMO: nivolumab with chemotherapy
*Includes patients still on treatment and patients off treatment continuing in the follow-up period
Note: Percentages are given against the treated population
Source: CheckMate 648 [redacted] Summary data[13]

B.2.6.2 CheckMate 648: Baseline patient characteristics

The demographics and baseline characteristics of patients enrolled in CheckMate 648 are summarised in Table 9. A total of 970 patients were randomised. At database lock ([redacted]), the minimum follow-up was 20 months.

The median age in the NIVO-CHEMO was 64 (range: 40-90) compared to 64 (range: 26-81) in the CHEMO arm. There was similar proportion of patients aged above and below 65 years. Most patients were male (78.8% in the NIVO-CHEMO arm and 84.9% in the CHEMO arm). The predominant histological type was squamous cell carcinoma (96.9% in the NIVO-CHEMO arm and 98.1% in the CHEMO arm). Geographically, the largest proportion of patients came from East Asia ([redacted]% in the NIVO-CHEMO arm and [redacted]% in the CHEMO arm), followed by the rest of the world (29.9% and 30.2%, respectively) and the rest of the Asia ([redacted]% and [redacted]%, respectively). There was an equal distribution of patients with PD-L1 expression of <1% and ≥1% in all treatment groups. Patients randomised to receive NIVO-CHEMO were overall comparable to patients randomised to receive CHEMO in terms of baseline characteristics. Disease stage at initial entry as well as disease status were also similar between the groups.

Table 9. Characteristics of participants in the CheckMate 648 trial across treatment groups in all randomised patients

Baseline characteristic		NIVO-CHEMO	CHEMO
Cohort size		321	324
Age	Median (range), years	64 (40-90)	64 (26-81)
Sex	Male n (%)	253 (78.8)	275 (84.9)
Race, n (%)	White	85 (26.5)	84 (25.9)
	Black	1 (0.3)	6 (1.9)
	Asian	227 (71)	227 (70)
	Other	6 (1.9)	6 (1.9)
Geographic location, n (%)	Asia	225 (70)	226 (70)
	Rest of world	96 (29.9)	98 (30.2)
ECOG PS, n (%)	0	150 (46.7)	154 (47.5)
	1	171 (53.3)	170 (52.5)
Histological type, n (%)	Squamous cell carcinoma	311 (96.9)	318 (98.1)
	Adenosquamous cell carcinoma	[redacted]	[redacted]
	Other	[redacted]	[redacted]
Tumour cell PD-L1 expression, n (%)*	≥ 1 %	158 (49.2)	156 (48.4)
	< 1 %	163 (50.8)	166 (51.6)
Disease stage at initial diagnosis, n (%)	Stage I- III	[redacted]	[redacted]
	Stage IV	[redacted]	[redacted]
	Not reported	[redacted]	[redacted]
Disease status at study entry, n (%)	*De novo* metastatic	184 (57.3)	187 (57.7)
	Recurrent – distant	72 (22.4)	60 (18.5)
	Recurrent – loco-regional	21 (6.5)	25 (7.7)
	Unresectable advanced	44 (13.7)	52 (16.0)
Number of organs with metastases, n (%)	≤ 1	158 (49.2)	158 (48.8)
	≥ 2	163 (50.8)	166 (51.2)
Location at initial diagnosis, n (%)	Upper thoracic	[redacted]	[redacted]
	Middle thoracic	[redacted]	[redacted]
	Lower thoracic	[redacted]	[redacted]
	Gastroesophageal junction	[redacted]	[redacted]
	Not reported	[redacted]	[redacted]

CHEMO: chemotherapy; ECOG PS: Eastern cooperative oncology group performance scale; NIVO-CHEMO: nivolumab with chemotherapy;

* does not include indeterminate patients

B.2.6.3 CheckMate 648: Results

At the database lock ([redacted]), minimum follow-up was 20 months. A summary of the key primary outcomes (OS and PFS for the patients with PD-L1 ≥ 1%) from CheckMate 648 is

provided in Table 10 and the secondary outcomes (OS and PFS for all randomised patients) are provided in Table 11.

Table 10. CheckMate 648: primary outcomes, randomised patients with tumour cell PD-L1 ≥1%

Endpoint		[redacted]		[redacted]	
		NIVO-CHEMO (n=158)	**CHEMO (n=157)**	**NIVO-CHEMO (n=158)**	**CHEMO (n=157)**
OS	Events, n (%)	[redacted]	[redacted]	98 (62)	121 (77.1)
	Median OS (95% CI), months	[redacted]	[redacted]	15.4 (11.9, 19.5)	9.1 (7.7, 10.0)
	12-month OS rate (95% CI), %	[redacted]	[redacted]	0.54 (0.37, 0.8)	NA
	HR (99.5% CI)	[redacted]	[redacted]	0.5 (0.4, 0.71)	NA
	Stratified 2-sided log-rank test p-value	[redacted]	[redacted]	<0.001	NA
PFS per BICR	Events, n (%)	[redacted]	[redacted]	117 (74.1)	100 (63.7)
	HR (95% CI)	[redacted]	[redacted]	0.7 (0.5, 0.9)	NA
	Median (95% CI)	[redacted]	[redacted]	6.9 (5.7, 8.3)	4.4 (2.9, 5.8)
	PFS rate (95% CI) at 12 months	[redacted]	[redacted]	25.41 (18.2,22.2)	10.45 (4.7,18.8)
	PFS rate (95% CI) at 18 months	[redacted]	[redacted]	-	-
	Stratified 2-sided log-rank test p-value	[redacted]	[redacted]	0.002	NA

CI: confidence internal; BICR: blinded independent central review; OS: overall survival; PFS: progression-free survival.
Source: CheckMate 648 [redacted] Summary data,[13] Doki et al. (2022)[63]

Table 11. CheckMate 648: secondary outcomes, all randomised patients

Endpoint		[redacted]		[redacted]	
		NIVO-CHEMO (n=321)	CHEMO (n=324)	NIVO-CHEMO (n=321)	CHEMO (n=324)
OS	Events, n (%)	[redacted]	[redacted]	209 (65.1)	232 (71.6)
	Median OS (95% CI), months	[redacted]	[redacted]	13.2 (11.1, 15.7)	10.7 (9.4, 11.9)
	12-month OS rates (95% CI), %	[redacted]	[redacted]	53.53 (47.8,58.9)	44.32 (38.6,49.9)
	HR (99.5% CI)	[redacted]	[redacted]	0.7 (0.6, 1.0)	NA
	Stratified 2-sided log-rank test p-value	[redacted]	[redacted]	0.0021	NA
PFS per BICR	Events, n (%)	[redacted]	[redacted]	235 (73.2)	210 (64.8)
	HR (95% CI)	[redacted]	[redacted]	0.8 (0.6, 1.0)	NA
	Median (95% CI)	[redacted]	[redacted]	5.8 (5.6, 7.0)	5.6 (4.3, 5.9)
	PFS rate (95% CI) at 12 months	[redacted]	[redacted]	23.62 (18.63, 28.95)	16.02 (11.02, 21.86)
	PFS rate (95% CI) at 18 months	[redacted]	[redacted]	-	-
	Stratified 2-sided log-rank test p-value	[redacted]	[redacted]	0.0355	NA

CI: confidence internal; BICR: blinded independent central review; OS: overall survival; PFS: progression-free survival.
Source: CheckMate 648 [redacted] Summary data[13] Doki et al. (2022)[63]

B.2.6.3.1 Overall survival

In the subgroup of patients with tumour cell PD-L1 ≥1%, treatment with NIVO-CHEMO demonstrated a statistically significant improvement in OS in comparison to CHEMO alone ([redacted]: median OS 15.4 months compared to 9.1 months; HR: 0.54 [0.37, 0.80], p < 0.0001) (Table 10). A similar improvement in OS was also observed at the 20-month minimum follow-up (DBL [redacted]) (median OS [redacted] months compared to [redacted] months; HR: [redacted]) (Table 10).

A similar clinically relevant improvement in OS was also observed in all randomised patients treated with NIVO-CHEMO compared to CHEMO (median OS [redacted] months compared to [redacted] months; HR: [redacted]) (Table 11). The corresponding Kaplan-Meier plots are presented in Figure 9 and Figure 10.

Figure 9. Overall survival in the NIVO-CHEMO and CHEMO arms – patients with tumour cell PD-L1 ≥1%

Statistical model for hazard ratio and p-value: stratified Cox proportional hazard model and stratified log-rank test.
Symbols represent censored observations
Stratification factors are ECOG Performance Status (0 vs 1), number of organs with metastases (≤1 vs ≥2) as recorded in IRT.
Source: CheckMate 648 [redacted] Summary data[13]

Figure 10. Overall survival in the NIVO-CHEMO and CHEMO arms - all randomised patients

Statistical model for hazard ratio and p-value: stratified Cox proportional hazard model and stratified log-rank test.
Symbols represent censored observations
Stratification factors are ECOG Performance Status (0 vs 1), number of organs with metastases (≤1 vs ≥2) as recorded in IRT.
Source: CheckMate 648 [redacted] Summary data[13]

B.2.6.3.2 Progression-free survival

Treatment with NIVO-CHEMO demonstrated a statistically significant improvement in PFS per BICR when considering patients with tumour cell PD-L1 ≥1% (primary PFS definition) compared with the CHEMO arm ([redacted] DBL: HR: 0.65 [95% CI: 0.46, 0.92], P < 0.0001). The PFS benefit was maintained for the [redacted] DBL (HR: [redacted]).

When considering all randomised patients who received NIVO-CHEMO, the prespecified significance boundary for PFS per BICR was not met in the [redacted] or [redacted] DBLs. The corresponding Kaplan-Meier plots are presented in Figure 11 and Figure 12.

Figure 11. Progression-free survival (per BICR) in the NIVO-CHEMO and CHEMO arms - tumour cell PD-L1 ≥1%

[redacted]

Statistical model for hazard ratio and p-value: stratified Cox proportional hazard model and stratified log-rank test.
Symbols represent censored observations
Stratification factors are ECOG Performance Status (0 vs 1), number of organs with metastases (≤1 vs ≥2), PD-L1 status (≥1% vs. <1% or indeterminate) as recorded in IRT.
Source: CheckMate 648 [redacted] Summary data[13]

Figure 12. Progression-free survival (per BICR) in the NIVO-CHEMO and CHEMO arms - all randomised patients

Statistical model for hazard ratio and p-value: stratified Cox proportional hazard model and stratified log-rank test.

Symbols represent censored observations

Stratification factors are ECOG Performance Status (0 vs 1), number of organs with metastases (≤1 vs ≥2), PD-L1 status (≥1% vs. <1% or indeterminate) as recorded in IRT.

Source: CheckMate 648 Summary data[13]

B.2.6.3.3 Objective response rate and duration of response

In all randomised patients with tumour cell PD-L1 expression ≥1%, an improvement in BICR-assessed ORR (95% CI) was observed with NIVO-CHEMO ( ) compared to CHEMO ( ). Complete responses by BICR were observed in %) patients in the NIVO-CHEMO arm, and %) patients in the CHEMO arm (Table 12).

In all randomised patients, an improvement in BICR-assessed ORR (95% CI) was observed with NIVO-CHEMO (,) compared to CHEMO). Complete responses by BICR were observed in (%) patients in the NIVO-CHEMO arm and %) patients in the CHEMO arm (Table 12).

Table 12. ORR by BICR results from the statistical testing hierarchy

Endpoint		NIVO-CHEMO (n=158)[a]	CHEMO (n=157)[a]
Patients with tumour cell PD-L1 ≥1%	ORR, %		
	95% CI		
	Best overall response, %		
	Complete response		
	Partial response		
	Stable disease		
	Progressive disease		
	Not evaluable		
	Median time to response[b] (range), months		
All randomised patients	ORR, %		
	95% CI		
	Best overall response, %		
	Complete response		

	Partial response	[redacted]	[redacted]
	Stable disease	[redacted]	[redacted]
	Progressive disease	[redacted]	[redacted]
	Not evaluable	[redacted]	[redacted]
	Median time to response[b] (range), months	[redacted]	[redacted]

CI: confidence internal; CR: complete response; ORR: objective response rate.
Source: CheckMate 648 [redacted] Summary data[13]
a. Randomised patients who had target lesion measurements at baseline per BICR assessment
b. Time to response was defined as the time from the start of treatment to the first objective tumour response

Among patients with PD-L1 ≥1, median DOR by BICR (95% CI) was [redacted] months for patients in the NIVO-CHEMO arm compared to [redacted]) months in the CHEMO arm (Table 13). Among all responders, median DOR by BICR (95% CI) was [redacted] months for NIVO-CHEMO vs. [redacted]) months for CHEMO (Table 13).

Table 13. DOR by BICR results from the statistical testing hierarchy

Endpoint		NIVO-CHEMO	CHEMO
All randomised patients with tumour cell PD-L1 ≥1%	Median DOR, months	[redacted]	[redacted]
	95% CI	[redacted]	[redacted]
All randomised patients	Median DOR, months	[redacted]	[redacted]
	95% CI	[redacted]	[redacted]

CI: confidence internal; DOR: duration of response; Max: maximum; Min: minimum; .
Source: CheckMate 648 [redacted] Summary data[13]

B.2.6.3.4 Health-related quality of life (HRQoL)

Changes to patients' quality of life (QoL) were recorded during the CheckMate 648 trial using the EQ-5D-3L utility index and visual analogue scale (VAS), and Functional Assessment of Cancer Therapy-Esophageal (FACT-E) instrument.[63]

Among patients with PD-L1 ≥1%, at baseline, mean (SD) EQ-5D-3L Utility Index scores were similar across the NIVO-CHEMO and CHEMO arms. The mean change from baseline increased during the on-treatment period in both the NIVO-CHEMO arm and the CHEMO arm. These improvements in mean EQ-5D-3L utility index scores were sustained longer and surpassed the minimally important difference (MID) threshold more often in the NIVO-CHEMO arm compared to the CHEMO arm.

A longitudinal mixed-model analysis of FACT-E scores through week 49 showed an overall increase in least-squares mean change from baseline with NIVO-CHEMO (4.98 points; 95% CI, 2.68 to 7.27) and CHEMO (1.54 points; 95% CI, −1.26 to 4.33) in the overall population.[63] However, these improvements from baseline were not clinically meaningful, indicating that health-related QoL was maintained during the treatment period.

Except at baseline, the proportion of patients who reported not being bothered by treatment side-effects over time was similar in those with NIVO-CHEMO to those with CHEMO.[63]

EQ-5D-3L utility index

The EQ-5D index is a standardised index instrument to measure self-reported health status and functioning.

Patients with PD-L1≥1%

Among patients with PD-L1 ≥1%, at baseline, mean (SD) EQ-5D-3L Utility Index scores (based on the UK value set) were similar across for the NIVO-CHEMO arm ([redacted]) and CHEMO arm ([redacted]). The mean change from baseline increased during the on-treatment period in the NIVO-CHEMO arm and the CHEMO arm (Figure 13).

Improvements in mean EQ-5D-3L utility index scores were sustained longer and surpassed the minimally important difference (MID) threshold more often in the NIVO-CHEMO arm compared to the CHEMO arm. Except for Weeks 3, 5 and 7, mean changes from baseline increased at all on-treatment assessments from Week 13 through Week 85 for the NIVO-CHEMO arm. Except for Weeks 3, 31, and 37, mean changes from baseline increased during the on-treatment period for subjects in the CHEMO arm. Increases above the MID threshold (0.08) were observed at Weeks 79 and 85 for the NIVO-CHEMO arm. A decrease below the MID threshold (0.08) at Week 31 was seen in the CHEMO arm. Mean decreases from baseline were observed in the NIVO-CHEMO arms at follow-up visits 1 and 2.

All randomised patients

Among all randomised patients, at baseline, mean (SD) EQ- 5D-3L Utility Index scores for the NIVO-CHEMO arm ([redacted]) were similar to those in the CHEMO arm ([redacted]). The mean change from baseline increased in the NIVO-CHEMO arm and the CHEMO arm (Figure 14).

Improvements in mean EQ-5D-3L utility index scores were sustained longer and surpassed the MID threshold more often in the NIVO-CHEMO arm vs the CHEMO arm. Except for Week

3, mean changes from baseline increased at all on-treatment assessments starting at Week 5 through Week 97 for the NIVO-CHEMO arm and Week 5 through Week 49 for the CHEMO arm. The NIVO-CHEMO arm was above the minimally important difference (MID) threshold (0.08) in Weeks 79, 91, and 97. The CHEMO arm was above the MID threshold at Week 49. Mean decreases from baseline were observed in the NIVO-CHEMO and CHEMO arms at follow-up visits 1 and 2.

Figure 13. Mean changes in EQ-5D-3L utility index score from baseline – patients with PD-L1 ≥1

Error bars represent standard error of the mean

Only timepoints where data is available for ≥5 patients in each treatment group are plotted.

Horizontal reference line indicates the minimum important difference (MID) considered a change of ≥0.08 points from baseline

Figure 14. Mean changes in EQ-5D-3L utility index score from baseline – all randomised patients

Error bars represent standard error of the mean

Only timepoints where data is available for ≥5 patients in each treatment group are plotted.

Horizontal reference line indicates the minimum important difference (MID) considered a change of ≥0.08 points from baseline

EQ-5D-3L VAS

Patients with PD-L1 ≥1%

For patients with PD-L1 ≥1, improvements in mean EQ-5D-3L VAS scores were sustained longer and surpassed the MID threshold more often in the NIVO-CHEMO arm vs the CHEMO arm (Figure 15).

At baseline, mean (SD) EQ-5D-3L VAS scores were similar for the NIVO-CHEMO arm ([redacted]) and the CHEMO arm ([redacted]) arms. Except for Week 3, mean changes from baseline increased at all on-treatment assessments with ≥ 10 patients from Week 5 through Week 85 for the NIVO-CHEMO arm. Except for Week 37, mean changes from baseline increased at all on-treatment assessments with ≥ 10 patients from Week 3 through Week 37 in the CHEMO arm.

Increases above the MID threshold (7.0) were demonstrated at Week 79 for the NIVO-CHEMO arm. Mean decreases from baseline were observed in the NIVO-CHEMO arm at follow-up visit 2 and in the CHEMO arm at follow-up visits 1 and 2.

All randomised patients

For all randomised patients, improvements in mean EQ-5D-3L VAS scores were sustained longer and surpassed the MID threshold more often in the NIVO-CHEMO arm vs the CHEMO arm (

Figure 16).

At baseline, mean (SD) EQ-5D-3L VAS scores for the NIVO-CHEMO arm ([redacted]) were similar to those in the CHEMO arm ([redacted]). Mean changes from baseline increased at all on-treatment assessments where ≥10 patients completed surveys, starting at Week 3 for NIVO-CHEMO through Week 97 vs Week 3 for CHEMO subjects through Week 49.

Increases above the MID threshold (7.0) were demonstrated at Weeks 91 and 97 for the NIVO-CHEMO arm. Mean decreases from baseline were observed in the NIVO-CHEMO and CHEMO arms at follow-up visits 1 and 2.

Figure 15. Mean change in overall self-reported health status EQ-VAS from baseline - patients with PD-L1≥1

Error bars represent standard error of the mean

Only timepoints where data is available for ≥5 patients in each treatment group are plotted.

Horizontal reference line indicates the minimum important difference (MID) considered a change of ≥0.08 points from baseline

Figure 16. Mean change in overall self-reported health status EQ-VAS from baseline - all randomised patients

Error bars represent standard error of the mean

Only timepoints where data is available for ≥5 patients in each treatment group are plotted.

Horizontal reference line indicates the minimum important difference (MID) considered a change of ≥0.08 points from baseline

FACT-E

Scores for the FACT-G physical, family, social and emotional well-being subscales were combined to produce a FACT-G total score for each treatment arm, which provides an overall indicant of generic HRQoL. The FACT-G and ECS score were combined to produce a total score for the FACT-E, which provides a composite measure of general and targeted HRQoL.

Patients with PD-L1 ≥1

In patients with PD-L1 ≥1, at baseline, mean (SD) FACT-E scores for the NIVO-CHEMO ([redacted]) and CHEMO ([redacted]) arms were similar.[64] Except for Week 3, mean changes from baseline increased at all other on-treatment assessments (with ≥ 10 patients) from Week 5 through Week 85 for the NIVO-CHEMO arm and from Week 5 through Week 37 for the CHEMO arm.

The NIVO-CHEMO arm demonstrated increases above the MID threshold (9.1) from Weeks 31 through 85. Mean decreases from baseline were observed in the NIVO-CHEMO and CHEMO arms at follow-up visits 1 and 2.

All randomised patients

In all randomised patients, at baseline, mean (SD) FACT-E scores for the NIVO-CHEMO arm ([redacted]) and CHEMO arm ([redacted]) were similar.[64] Except for Week 3, mean changes from baseline increased at all other on-treatment assessments (with ≥ 10 patients) from Week 5 through Week 97 for the NIVO-CHEMO arm and from Week 5 through Week 49 for the CHEMO arm.

The NIVO-CHEMO arm demonstrated increases above the MID threshold (9.1) at Weeks 43 through 97. Mean decreases from baseline were observed in the NIVO-CHEMO and CHEMO arms at follow-up visits 1 and 2.

FACT-E ECS

The 17-item disease-specific FACT-E ECS assesses concerns related to swallowing, vocalization, breathing, dry mouth, eating, disrupted sleep due to coughing, stomach pain, and weight loss.

Patients with PD-L1 ≥1

In patients with PD-L1 ≥1, FACT-E ECS mean (SD) total baseline scores for the NIVO-CHEMO ([redacted]) and CHEMO ([redacted]) arms were similar.[64]

Mean changes from baseline for the NIVO-CHEMO arm increased at all on-treatment assessments (with ≥ 10 patients) through Week 85 with increases greater than the MID threshold (4.0) at Weeks 13 and 25 through 85.

For the CHEMO arm, mean changes from baseline increased at all on-treatment assessments (with ≥ 10 patients) through Week 37 with increases greater than the MID threshold (4.0) at Weeks 13 through Week 37.

At follow up visits 1 and 2, increases in mean changes from baseline were observed in the NIVO-CHEMO arm at both visits whereas the CHEMO arm showed a decrease at both follow-up visits.

During survival follow-up visits, mean changes from baseline for the NIVO-CHEMO arm were increased through visit 4 (with ≥ 10 patients). There was an increase greater than the MID threshold (4.0) at follow-up visit 4. Mean changes from baseline for the CHEMO arm were increased during the survival follow-up through follow-up visit 3 (with ≥ 10 patients). Increases greater than the MID threshold (4.0) were seen at survival follow-up visits 1 and 2.

All randomised patients

In all randomised patients, FACT-E ECS total mean (SD) baseline scores for the NIVO-CHEMO arm ([redacted]) and CHEMO arm ([redacted]) were similar.[64]

Mean changes from baseline for the NIVO-CHEMO arm increased at all on-treatment assessments (with ≥ 10 subjects) through Week 97, a change greater than the MID (4.0) threshold at Weeks 13 through 97.

For the CHEMO arm, mean changes from baseline increased at all on-treatment assessments (with ≥ 10 patients) through Week 49, with a change greater than the MID (4.0) threshold at Weeks 25 through 49.

At follow-up visit 1 and 2, increases in mean changes from baseline were observed in the NIVO-CHEMO group, whereas the CHEMO arm showed a decrease at both follow-up visits.

Mean changes from baseline for the NIVO-CHEMO arm were increased during the survival follow-up through follow-up visit 5 (with ≥ 10 patients). At survival follow-up visit 4, the increase was greater than the MID (4.0). Mean changes from baseline for the CHEMO arm were increased during the survival follow-up through follow-up visit 6 (with ≥ 10 patients). Increases of greater than 4.0 were seen at survival follow-up visits 2 through 6.

FACT-E GP5

The FACT-E GP5 item is a key PRO measure that assesses the overall bother associated with the side-effects of treatment.

Patients with PD-L1 ≥1

In patients with PD-L1 ≥1, at baseline, patients in the NIVO-CHEMO arm selected "not at all" [redacted]% of the time and "a little bit" [redacted]% for a total of [redacted]% patients identifying as bothered "only a little" or "not at all" by treatment side effects.[64] Except for Week 43, the combined score remained above [redacted]% during the on-treatment period (with ≥ 10 patients) and went above [redacted]% multiple times through Week 97.

Patients in the CHEMO arm had better baseline scores with [redacted]% selecting "not at all" and [redacted]% "a little bit" (Total = [redacted]%). However, the combined score was never above [redacted]% during the on-treatment period (with ≥ 10 patients) through Week 49 and dropped under [redacted]% at Week 37.

All randomised patients

In all randomised patients, at baseline, patients in the NIVO-CHEMO arm selected "not at all" [redacted]% of the time and "a little bit" [redacted]% for a total of [redacted]% patients identifying as bothered "only a little" or "not at all" by treatment side effects.[64] The combined score remained above [redacted]% during the on-treatment period (with ≥ 10 patients) and went above [redacted]% multiple times through Week 97.

At baseline, subjects in the CHEMO arm selected "not at all" [redacted]% and [redacted]% "a little bit" (combined total = [redacted]%). However, the combined total score was never above [redacted]% during the on-treatment period (with ≥ 10 patients) through Week 49 and was under [redacted]% at multiple time points.

FACT-G7

The 27-item FACT-General (FACT-G) generic cancer-related core measure assesses symptoms and treatment-related effects impacting physical well-being (PWB; seven items), social/family well-being (SWB; seven items), emotional well-being (EWB; six items), and functional well-being (FWB; seven items). Seven of these items comprise the FACT-G7, an abbreviated version of the FACT-G that provides a rapid assessment of general HRQoL in cancer patients.

Patients with PD-L1 ≥1

In patients with PD-L1 ≥1, at baseline, mean (SD) FACT-G7 scores were similar between the NIVO-CHEMO arm ([redacted]) and CHEMO arm ([redacted]).[64] Except for Week 3, mean changes from baseline increased at all other on-treatment assessments (with ≥10 patients) from Week 5 through Week 85 for the NIVO-CHEMO arm and from Week 5 through Week 37 for the CHEMO arm. Mean decreases from baseline were observed in the NIVO-CHEMO and CHEMO arms at follow-up visits 1 and 2. Except for follow-up visits 1 and 2 in the CHEMO arm, mean change from baseline decreased at all other survival follow-up visits (with ≥10 patients) for both the NIVO-CHEMO and CHEMO arms.

All randomised patients

In all randomised patients, at baseline, mean (SD) FACT-G7 scores for the NIVO-CHEMO arm ([redacted]) were similar to those in the CHEMO arm ([redacted]).[64] Except for Week 3, mean changes from baseline increased at all other on-treatment assessments (with ≥10 subjects) from Week 5 through Week 97 for the NIVO-CHEMO arm and from Week 5 through Week 49 for the CHEMO arm. Mean decreases from baseline were observed in the NIVO-CHEMO and CHEMO arms at follow-up visits 1 and 2. Except for follow-up visit 5 in the CHEMO arm, mean change from baseline decreased at all other survival follow-up visits (with ≥10 patients) for both the NIVO-CHEMO and CHEMO arms.

B.2.7 Subgroup analysis

OS and PFS were analysed for several pre-specified subgroups, summarised in Table 3.

The median OS and HRs for key subgroup analyses for all randomised patients are detailed in Figure 17 and

Figure 18. Overall, subgroup analyses of OS favoured NIVO-CHEMO over CHEMO (point estimate of HR <1) for all randomised patients. As shown in Figure 19 and

Figure 20, no further enrichment of response is observed at higher TC cut-off thresholds, either PD-L1 ≥5 or PD-L1 ≥10.

Figure 17. Forest plot of subgroup analysis, for age, sex, race, region, ECOG, weight and disease stage at initial diagnosis, on overall survival for all randomised patients treated with NIVO-CHEMO or CHEMO

Source: CheckMate 648 [redacted] Summary data[13]

Figure 18. Forest plot for subgroup analysis, for histologic grade, histologic classification, location, disease status, smoking status, alcohol use, number of organs with metastasis, on overall survival in all randomised patients treated with NIVO-CHEMO or CHEMO

HR is not computed for subset category with less than 10 subjects per treatment group

Source: CheckMate 648 [redacted] Summary report[13]

Figure 19. Forest plot of subgroup analysis, for time from initial diagnosis to randomisation, prior surgery, prior radiotherapy and prior systemic therapy, on overall survival in all randomised patients treated with NIVO-CHEMO or CHEMO

Source: CheckMate 648 [redacted] Summary report[13]

Figure 20. Forest plot of treatment effect on OS by tumour cell PD-L1 cut-offs – all randomised patients treated with NIVO-CHEMO or CHEMO

HR is not computed for subset category with less than 10 subjects per treatment group

Source: CheckMate 648 [redacted] Summary report[13]

B.2.8 Meta-analysis

Direct evidence for comparative efficacy of NIVO-CHEMO vs CHEMO may be drawn from the CheckMate-648 study, and so no meta-analysis is required. An indirect treatment comparison (ITC) considering the efficacy of pembrolizumab for inclusion in the cost-effectiveness model (CEM) is presented in Section B.2.9.

B.2.9 Indirect and mixed treatment comparisons

Key points

- A network meta-analysis (NMA) was conducted with the goal of including pembrolizumab with chemotherapy, as assessed in KEYNOTE-590, as a comparator arm within the CEM.
- The NMA considering the PD-L1 CPS ≥10 population demonstrated that trends were similar across all model families whereby pembrolizumab with chemotherapy improved PFS and OS when compared with chemotherapy across all timepoints.
- OS results generated from the NMA indicate that there is no statistically significant difference between the IO-chemo-based regimens used in OSCC, although point estimates tend to marginally favour PEMBRO-CHEMO.
- There are several limitations of the ITC including only one study informing each comparison with no closed loops in the network, as well as uncertainty and heterogeneity in those studies.

B.2.9.1 Indirect treatment comparison

A network meta-analyses (NMAs) was conducted with the goal of including pembrolizumab with chemotherapy, as assessed in KEYNOTE-590, as a comparator arm within the CEM. The NMA considered the PD-L1 ≥10% (CPS) population, in line with the population reported in KEYNOTE-590. As pembrolizumab is only licensed for use in patients with PD-L1 CPS ≥10%, only these patients were included in the KEYNOTE-590 trial. Therefore, only patients with PD-L1 CPS ≥10% from CheckMate 648 were included in the NMA for comparison with pembrolizumab, a subpopulation of the target population for this submission. A summary of the overlapping TC ≥1% and CPS ≥ 10% populations in the CheckMate 648 trial is presented in Table 14.

Table 14. A summary of the overlapping TC ≥1% and CPS ≥ 10% populations in the CheckMate 648 trial

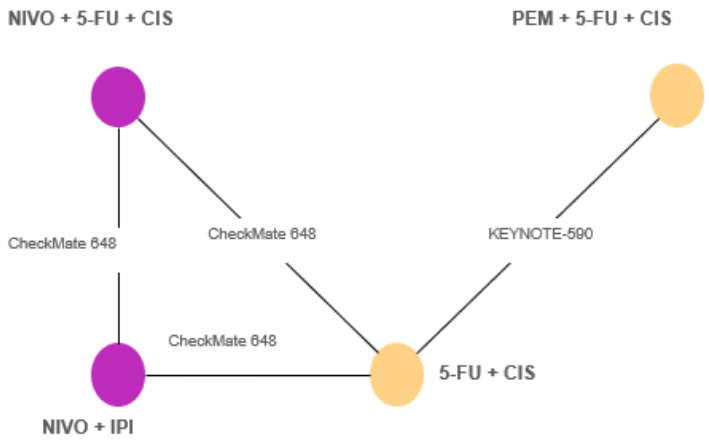
	CPS≥10	CPS < 10 & NA	Total
NIVO-CHEMO			
TC ≥1	96	62	158
TC < 1	39	124	163
Total	135	186	321
CHEMO			
TC ≥1	100	57	157
TC < 1	45	122	167
Total	145	179	324

B.2.9.1.1 Methods

Available data for inclusion in the NMA

The data required to inform the NMA was individual patient-level data (IPD) from both trials. This was available for CheckMate 648, however, for KEYNOTE-590, datasets for the models were sourced from digitized Kaplan-Meier curves and the number of patients at risk over time from which individual patient data (IPD) was recreated using the Guyott algorithm.[86] The network diagram for the included arms of the NMA is presented in Figure 21.

Figure 21. Network diagram



Assessment of comparability

The assessment of comparability was based on data from the all-comers population from KEYNOTE-590, as no baseline characteristics were reported for the OSCC PD-L1 CPS ≥10 population. The assessment found that CheckMate 648 and KEYNOTE-590 were sufficiently similar, in terms of study design and patient baseline characteristics (Appendix L), to conduct an indirect comparison. More detail is provided in Section B.2.9.4.

Under this assumption, survival models were fit to the pembrolizumab with chemotherapy arm of KEYNOTE-590 (Appendix L). For each of the survival models, differences in the survival function parameters between pembrolizumab with chemotherapy and chemotherapy, as assessed in KEYNOTE-590, were estimated and applied to the reference chemotherapy arm from CheckMate 648. The result was the PFS (investigator assessed, IA, as PFS BICR was not assessed during KEYNOTE-590) and OS over time for pembrolizumab with chemotherapy relative to chemotherapy, as assessed in CheckMate 648. This approach preserves randomisation and allows treatment effects to vary over time. This was important as proportional hazards are violated between pembrolizumab with chemotherapy and chemotherapy alone and between nivolumab with chemotherapy and chemotherapy (Appendix L). As the relative treatment effects are estimated via the chemotherapy arms, a robust model is required to model hazards that vary over time.

Model fitting and extrapolation

Both fixed and random-effects models were considered. Given insufficient evidence for estimation of the between-study heterogeneity (characterised by the heterogeneity parameter), fixed-effects models were used.

Traditional univariate NMA models have been extended by Achana et al. to synthesize relative treatment effects related to multiple outcomes.[65] Cope et al. has proposed using these models to synthesise multiple parameters of a survival function.[66] In this method, the alternative survival functions are first fit for the time-to-event outcomes of interest to identify the relevant parameters (and their correlations), which are then used as inputs in a multivariate NMA. The distribution-specific parameter estimates are transformed to a normally distributed scale with accompanying covariance matrix of the transformed parameters. The NMA model in the second step proposed by Cope et al.[66] is based on one specific parametric distribution that is assumed to apply to *all* arms of *all* trials within a network of evidence. It is possible to explore alternative parametric distributions as a series of sensitivity analyses, but alternative distributions cannot be combined within one network of evidence, which would violate the transitivity assumption.

All analyses were performed in a Bayesian framework and involved a model with parameters, data, a likelihood distribution, and prior distributions.[67] These methods employ a generalised linear model framework in which a likelihood is defined for the outcome and a link function is used to transform the outcome to a linear scale. Common distributions used for the analysis of time-to-event data as well as the corresponding survival, hazard functions, link functions, and transformation to linear prediction are presented in Appendix L.

The result of the application of the methods in Cope et al.[66] are differences in each of the survival function parameters between pembrolizumab with chemotherapy and chemotherapy (both from KEYNOTE-590). These differences on the survival function parameters can be applied to chemotherapy as assessed in CheckMate 648 to obtain PFS (IA) and OS over time for pembrolizumab with chemotherapy relative to chemotherapy, as assessed in CheckMate 648.

B.2.9.2 Results

B.2.9.2.1 Pembrolizumab with chemotherapy vs chemotherapy: patients with PD-L1 CPS ≥10

As described in Section B.2.9.1.4, the population of interest within KEYNOTE-590 is the subgroup of OSCC patients with PD-L1 ≥10% (CPS) with the outcomes of interest being PFS (where only PFS [IA] is reported) and OS. However, for PFS (IA), PD-L1 ≥10% (CPS) is only reported for the mixed histology population, which is used for the base case NMA. Fractional polynomial models were not considered for the NMA, due to similar fit and extrapolation to standard parametric and spline models. As the piecewise models are extensions of standard parametric models and less complex versions of the spline models with regards to fit and extrapolation, these were also not included within the NMA. Thus, only standard parametric and spline models were fit to KEYNOTE-590. As 3-, 4-, and 5-knot models were deemed overly complex, only 1- and 2-knot spline models were included in the NMA.

The assessment of proportional hazards assumption and modelling of PFS and OS required to inform the NMA are presented in Appendix L.

Results of the NMA considering PFS are presented in Table 15. Results are also presented as HRs over time as well as averaged HR (similar to constant HR) for standard parametric and spline models in Figure 22 and Figure 23, respectively. Trends were similar across all model families whereby pembrolizumab with chemotherapy improved PFS when compared with chemotherapy across all timepoints. For all models but the log-logistic, log normal, and Weibull, HRs decreased over time. Results for most time points were statistically significant with the exceptions of the spline models at 3 months and after 12 months, and for the standard parametric models after 12 months.

Results of the NMA considering OS are presented in Table 16, Figure 24 and Figure 25. Trends were similar across all model families whereby pembrolizumab with chemotherapy improved OS when compared with chemotherapy across all timepoints. Similar to the NMA of PFS, for all models but the log-logistic, log normal, and Weibull, HRs decreased over time. Results for most time points were statistically significant between 6 and 12 months; however only the gamma and generalized gamma models were statistically significant both at 3 months and at 24 months and thereafter.

Table 15. Tabular results of the indirect treatment comparison of progression-free survival; KEYNOTE-590 patients with PD-L1 ≥10% (combined positive score) and mixed histology

Model family	Model	Hazard ratios (95% credible intervals) for pembrolizumab + chemotherapy versus chemotherapy					
		3 months	6 months	9 months	12 months	24 months	36 months
Standard parametric	Gamma	0.4 (0.56, 0.66)	0.4 (0.49, 0.58)	0.39 (0.47, 0.57)	0.38 (0.45, 0.57)	0.35 (0.43, 0.57)	0.33 (0.42, 0.57)
	Generalized gamma	0.18 (0.57, 0.82)	0.35 (0.59, 0.75)	0.39 (0.57, 0.74)	0.39 (0.56, 0.74)	0.34 (0.53, 0.75)	0.3 (0.51, 0.78)
	Gompertz	0.4 (0.58, 0.83)	0.43 (0.58, 0.78)	0.44 (0.57, 0.74)	0.42 (0.57, 0.75)	0.28 (0.55, 1.05)	0.16 (0.53, 1.63)
	Log-logistic	0.21 (0.5, 0.77)	0.28 (0.51, 0.71)	0.34 (0.54, 0.74)	0.4 (0.58, 0.81)	0.51 (0.71, 1.01)	0.55 (0.79, 1.1)
	Long normal	0.13 (0.45, 0.72)	0.29 (0.54, 0.76)	0.4 (0.58, 0.81)	0.45 (0.62, 0.86)	0.53 (0.69, 0.97)	0.54 (0.72, 1.03)
	Weibull	0.25 (0.54, 0.8)	0.32 (0.55, 0.74)	0.36 (0.55, 0.73)	0.39 (0.56, 0.73)	0.43 (0.56, 0.81)	0.42 (0.57, 0.9)
Spline hazard	1-knot	0.37 (0.59, 0.81)	0.31 (0.53, 0.84)	0.26 (0.48, 0.85)	0.2 (0.44, 0.9)	0.1 (0.38, 1.12)	0.07 (0.35, 1.22)
	2-knot	0.32 (0.58, 1.08)	0.21 (0.43, 0.81)	0.24 (0.48, 0.86)	0.23 (0.53, 1.09)	0.14 (0.66, 2.52)	0.12 (0.71, 3.69)
	3-knot	0.21 (0.61, 1.26)	0.1 (0.4, 0.8)	0.18 (0.43, 1)	0.24 (0.52, 1.07)	0.19 (0.74, 2.97)	0.17 (0.83, 4.2)
Spline odds	1-knot	0.28 (0.59, 1.17)	0.18 (0.41, 0.81)	0.11 (0.34, 0.84)	0.07 (0.31, 0.89)	0.02 (0.31, 1.01)	0.02 (0.33, 1.04)
	2-knot	0.27 (0.57, 1.12)	0.13 (0.4, 0.85)	0.11 (0.42, 0.99)	0.08 (0.45, 1.24)	0.04 (0.55, 1.65)	0.04 (0.6, 1.69)
	3-knot	0.34 (0.73, 1.46)	0.02 (0.21, 0.67)	0.11 (0.37, 0.88)	0.15 (0.56, 1.38)	0.2 (0.92, 2.39)	0.24 (0.98, 2.45)
Spline normal	1-knot	0.3 (0.61, 1.1)	0.21 (0.45, 0.81)	0.12 (0.37, 0.82)	0.08 (0.33, 0.86)	0.03 (0.28, 0.95)	0.02 (0.27, 0.99)
	2-knot	0.25 (0.53, 1)	0.18 (0.44, 0.83)	0.18 (0.47, 0.96)	0.14 (0.49, 1.18)	0.07 (0.55, 1.85)	0.06 (0.56, 2.08)
	3-knot	0.22 (0.67, 1.33)	0.03 (0.27, 0.67)	0.2 (0.38, 0.68)	0.26 (0.52, 0.92)	0.21 (0.78, 2.2)	0.2 (0.85, 2.66)

Figure 22. Results of the indirect treatment comparison of progression-free survival; standard parametric models, KEYNOTE-590 patients with PD-L1 ≥10% (combined positive score) and mixed histology

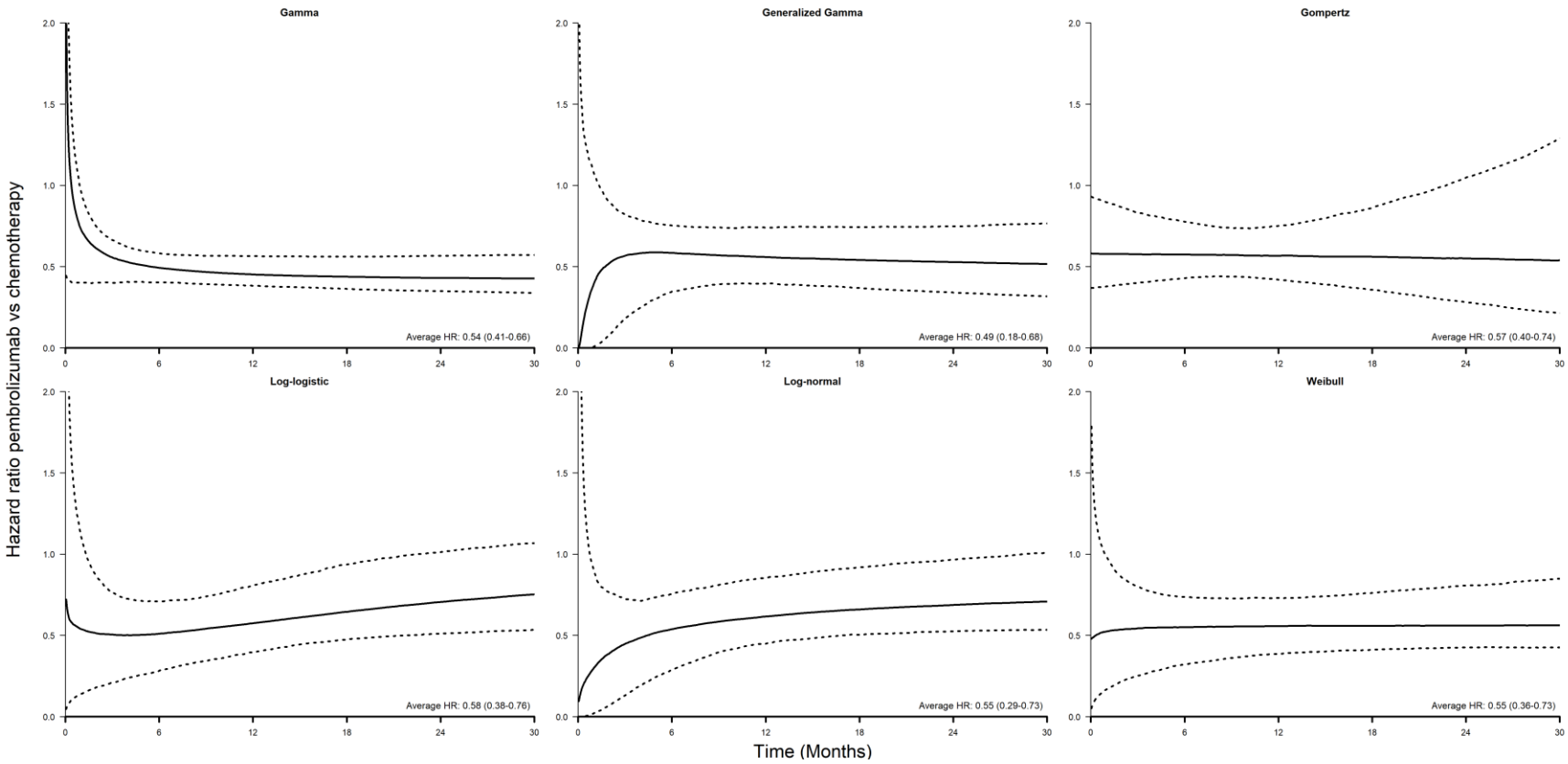


Figure 23. Results of the indirect treatment comparison of progression-free survival; spline models, KEYNOTE-590 patients with PD-L1 ≥10% (combined positive score) and mixed histology

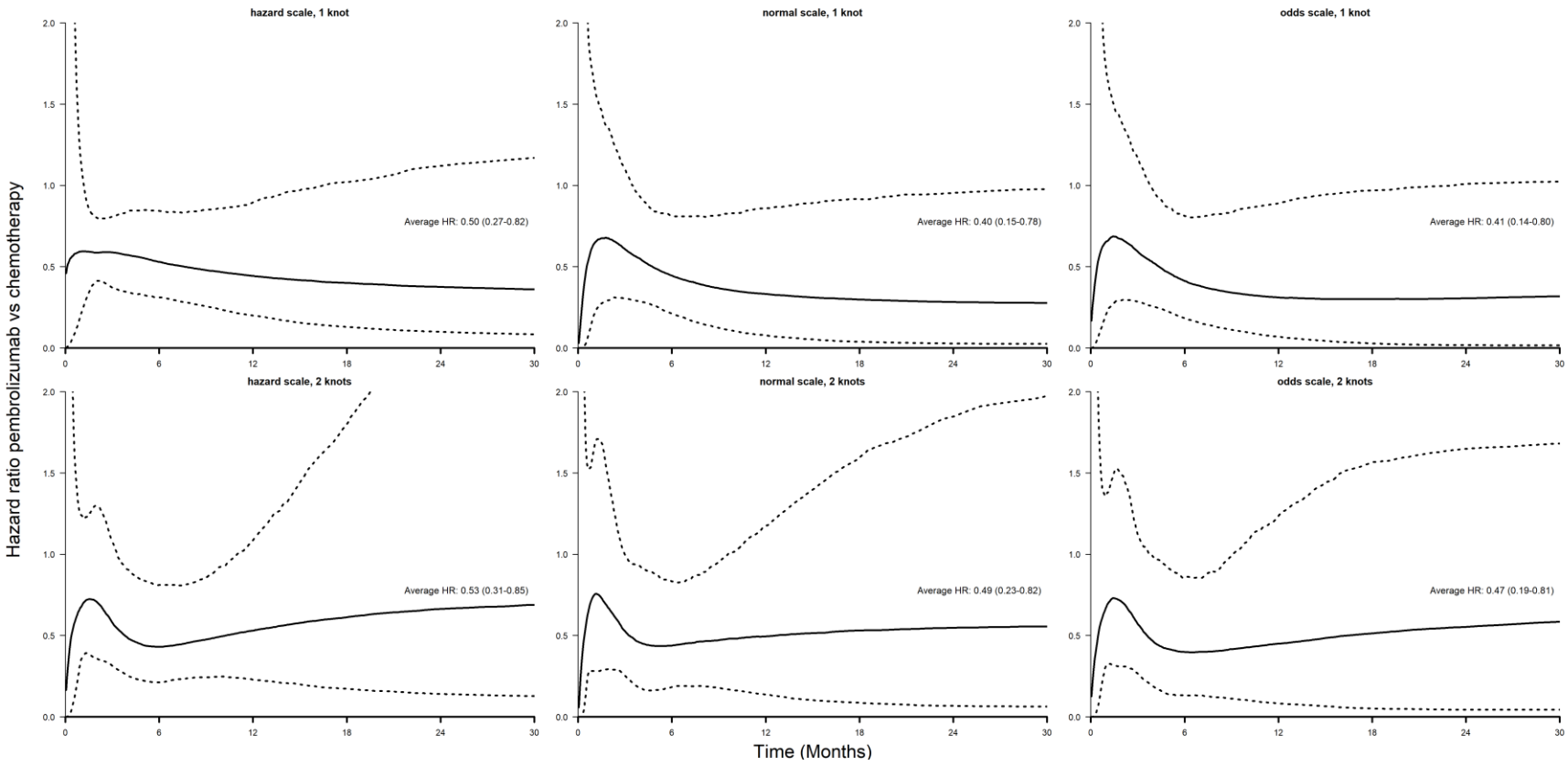


Table 16. Tabular results of the indirect treatment comparison of overall survival; KEYNOTE-590 patients with PD-L1 ≥10% (combined positive score) and squamous cell carcinoma

Model family	Model	Hazard ratios (95% credible intervals) for pembrolizumab + chemotherapy versus chemotherapy					
		3 months	6 months	9 months	12 months	24 months	36 months
Standard parametric	Gamma	0.2 (0.51, 0.79)	0.3 (0.54, 0.74)	0.36 (0.55, 0.73)	0.4 (0.56, 0.74)	0.44 (0.57, 0.82)	0.44 (0.59, 0.85)
	Generalized gamma	0.18 (0.57, 0.82)	0.35 (0.59, 0.75)	0.39 (0.57, 0.74)	0.39 (0.56, 0.74)	0.34 (0.53, 0.75)	0.3 (0.51, 0.78)
	Gompertz	0.4 (0.58, 0.83)	0.43 (0.58, 0.78)	0.44 (0.57, 0.74)	0.42 (0.57, 0.75)	0.28 (0.55, 1.05)	0.16 (0.53, 1.63)
	Log-logistic	0.21 (0.5, 0.77)	0.28 (0.51, 0.71)	0.34 (0.54, 0.74)	0.4 (0.58, 0.81)	0.51 (0.71, 1.01)	0.55 (0.79, 1.1)
	Log normal	0.13 (0.45, 0.72)	0.29 (0.54, 0.76)	0.4 (0.58, 0.81)	0.45 (0.62, 0.86)	0.53 (0.69, 0.97)	0.54 (0.72, 1.03)
	Weibull	0.25 (0.54, 0.8)	0.32 (0.55, 0.74)	0.36 (0.55, 0.73)	0.39 (0.56, 0.73)	0.43 (0.56, 0.81)	0.42 (0.57, 0.9)
Spline hazard	1-knot	0.24 (0.52, 1.13)	0.31 (0.62, 1.19)	0.34 (0.61, 1.08)	0.31 (0.57, 1)	0.08 (0.47, 1.38)	0.05 (0.45, 1.8)
	2-knot	0.22 (0.55, 1.47)	0.3 (0.55, 1)	0.25 (0.54, 1.02)	0.27 (0.55, 1.06)	0.08 (0.66, 2.32)	0.05 (0.69, 3.55)
	3-knot	0.21 (0.52, 1.26)	0.29 (0.61, 1.1)	0.28 (0.56, 0.95)	0.08 (0.48, 1.15)	0.14 (0.75, 2.63)	0.09 (0.84, 4.37)
Spline odds	1-knot	0.24 (0.55, 1.26)	0.28 (0.56, 1.11)	0.24 (0.51, 0.97)	0.18 (0.47, 0.97)	0.05 (0.46, 1.28)	0.05 (0.5, 1.38)
	2-knot	0.18 (0.52, 1.41)	0.24 (0.51, 1.05)	0.2 (0.51, 1.08)	0.22 (0.54, 1.05)	0.05 (0.54, 2.06)	0.03 (0.58, 2.42)
	3-knot	0.29 (0.49, 0.81)	0.35 (0.62, 1.05)	0.26 (0.53, 1)	0.15 (0.42, 0.98)	0.17 (0.67, 1.93)	0.18 (0.75, 2.24)
Spline normal	1-knot	0.2 (0.56, 1.39)	0.28 (0.56, 1.05)	0.27 (0.53, 0.94)	0.2 (0.5, 0.97)	0.1 (0.47, 1.15)	0.08 (0.47, 1.25)
	2-knot	0.15 (0.49, 1.34)	0.27 (0.51, 0.86)	0.25 (0.53, 0.88)	0.35 (0.54, 0.8)	0.07 (0.57, 1.62)	0.04 (0.58, 2.1)
	3-knot	0.27 (0.5, 0.95)	0.23 (0.59, 1.06)	0.25 (0.51, 0.82)	0.08 (0.45, 1.04)	0.13 (0.72, 1.92)	0.09 (0.78, 2.69)

Figure 24. Results of the indirect treatment comparison of overall survival; standard parametric models, KEYNOTE-590 patients with PD-L1 ≥10% (combined positive score) and squamous cell carcinoma

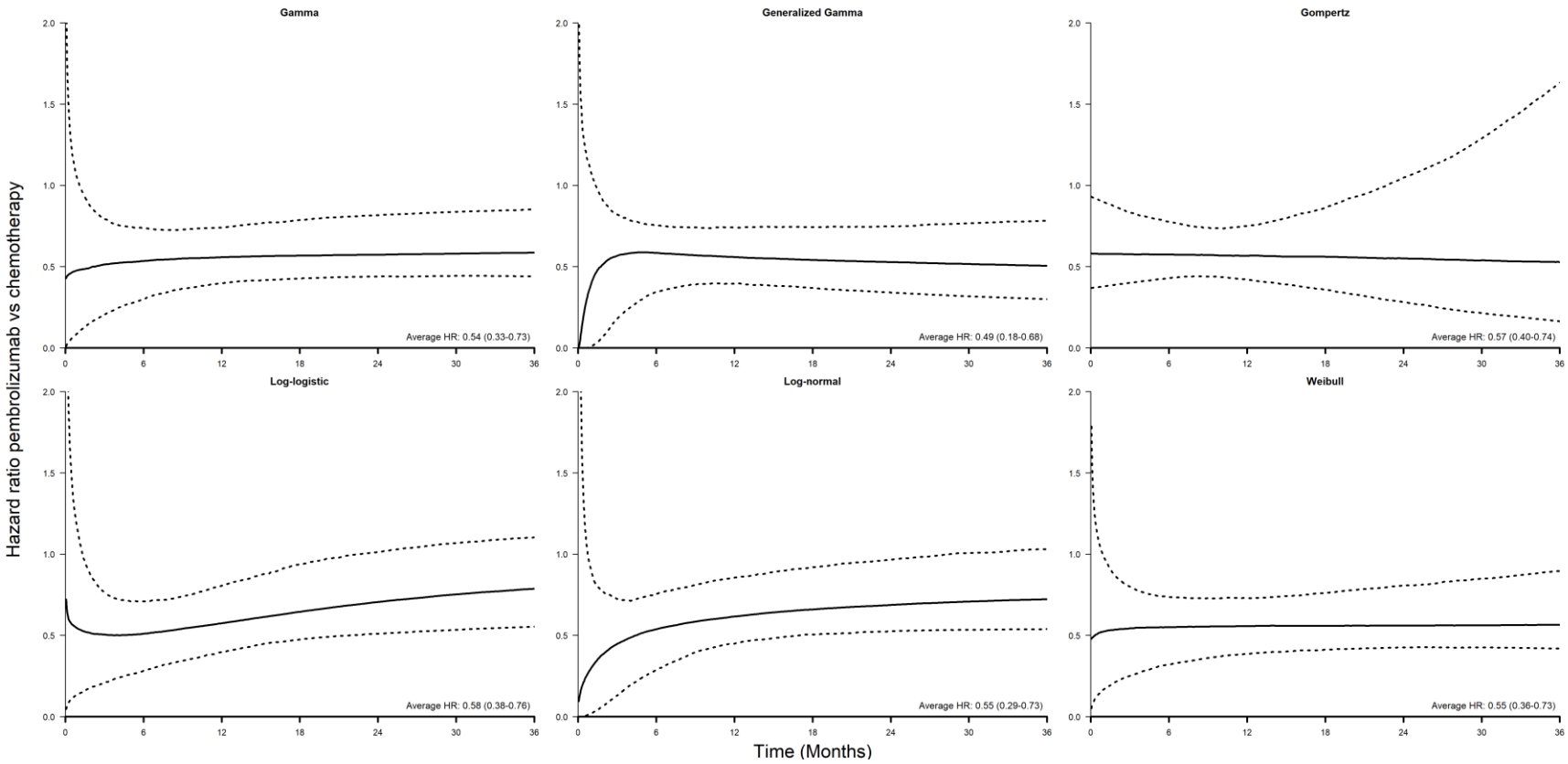
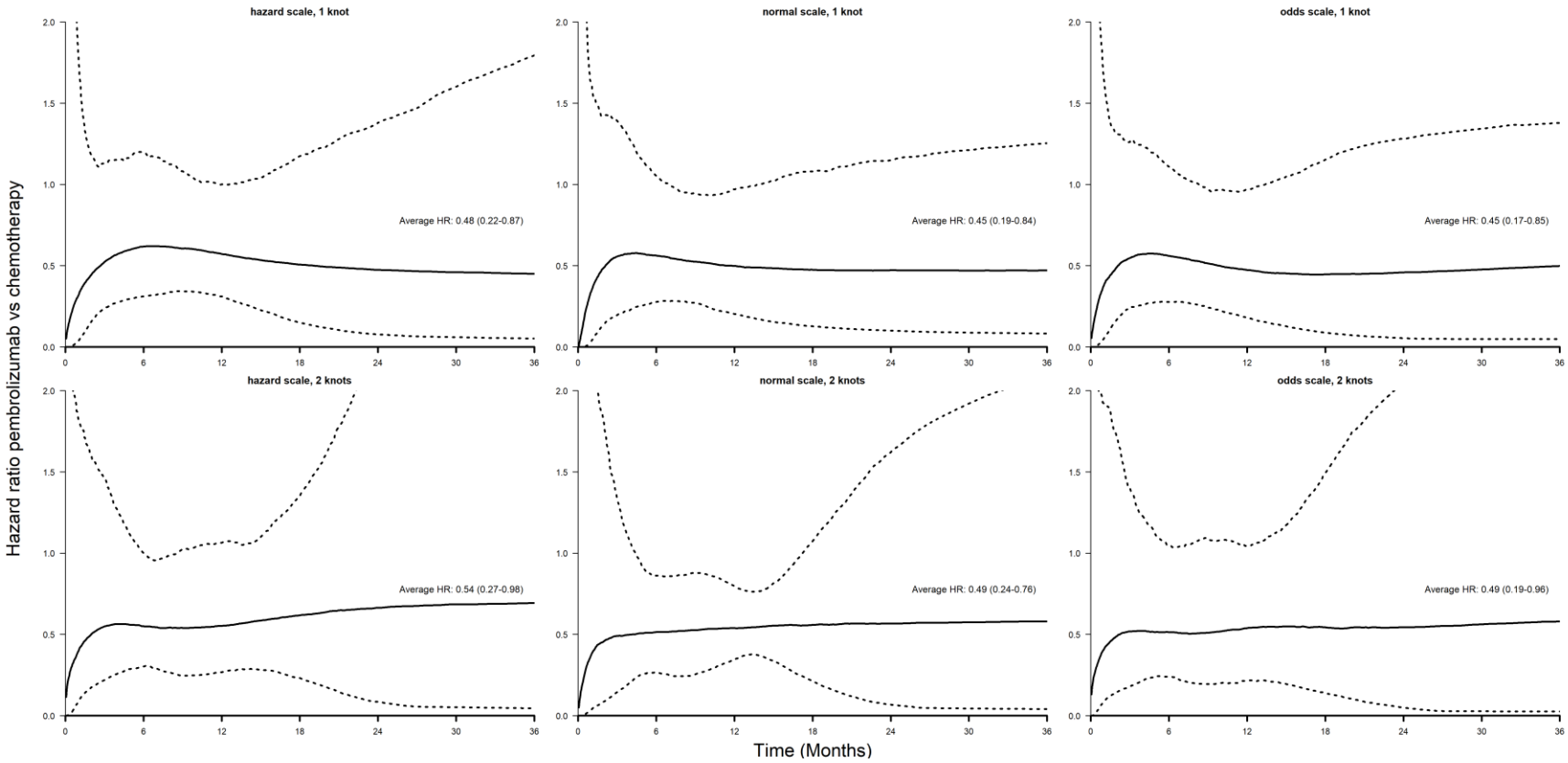


Figure 25. Results of the indirect treatment comparison of overall survival; spline models, KEYNOTE-590 patients with PD-L1 ≥10% (combined positive score) and squamous cell carcinoma



B.2.9.2.2 Pembrolizumab with chemotherapy vs nivolumab with chemotherapy: patients with PD-L1 CPS ≥10

Results of the NMA considering the fixed effects Gamma model for OS of patients treated with pembrolizumab with chemotherapy compared to nivolumab with chemotherapy are presented in Table 17. While the point estimates indicate that pembrolizumab with chemotherapy improves OS compared to nivolumab with chemotherapy, the results are not statistically significant, as the credible interval spans 1, and therefore, the NMA demonstrates that pembrolizumab with chemotherapy has a similar effect on OS as nivolumab with chemotherapy for the treatment of advanced OSCC in patients with PD-L1 CPS ≥10.

Table 17. Results of fixed effects Gamma model for overall survival: HR over time

	Hazard ratios (95% credible intervals) for nivolumab + chemotherapy versus comparators at each timepoint (months)									
	3	**6**	**9**	**12**	**18**	**24**	**30**	**36**	**42**	**48**
Vs chemotherapy	0.69 (0.45, 0.99)	0.70 (0.52, 0.92)	0.70 (0.53, 0.94)	0.71 (0.53, 0.96)	0.71 (0.52, 1.00)	0.72 (0.51, 1.03)	0.72 (0.51, 1.03)	0.72 (0.50, 1.06)	0.72 (0.49, 1.08)	0.73 (0.49, 1.09)
Vs pembrolizumab with chemotherapy	1.34 (0.73, 2.54)	1.30 (0.86, 1.98)	1.27 (0.85, 1.89)	1.26 (0.83, 1.89)	1.24 (0.78, 1.94)	1.23 (0.74, 1.98)	1.22 (0.72, 2.01)	1.21 (0.70, 2.04)	1.21 (0.68, 2.06)	1.20 (0.67, 2.08)
Cells shaded in grey indicate estimates based on model extrapolations										

B.2.9.2.3 Pembrolizumab with chemotherapy vs chemotherapy: patients with PD-L1 TC ≥1 and PD-L1 ≥10

Additionally, an NMA for CheckMate 648 patients that had both PD-L1 ≥10% (CPS) as well as PD-L1 ≥1% (TC) was considered for completeness. However, it was determined that including this population in an NMA would result in an unequal distribution of PD-L1 status, because patients who were both PD-L1 TC ≥1 and PD-L1 CPS ≥10 were removed, therefore enriching the patients with PD-L1 from CheckMate 648. Consequently, it was deemed inappropriate to include this population in any further analysis.

B.2.9.2.3 Summary of results

Overall, OS results generated from the NMA indicate that there is no statistically significant difference between the IO-chemo-based regimens used in the treatment of OSCC, although point estimates tend to marginally favour pembrolizumab in combination with chemotherapy.

B.2.9.3 Results of the assessment of heterogeneity

The assessment for heterogeneity was performed according to the Technical Support Document (TSD) 3 written by the NICE Decision Support Unit (DSU).[68] The key factor identified was the sparsity of the evidence base, as the comparison consists of a single trial informing each direct comparison. However, no significant between trial heterogeneity was identified that would affect the comparability of the trials and prevent their inclusion in the NMA. Therefore, no assessment of heterogeneity in the form of I-square analysis can be estimated.

B.2.9.4 Uncertainties in the indirect treatment comparison

While indirect comparisons provide useful insights in the absence of direct trial-based comparisons, they cannot replace evidence from head-to-head studies, which remain the gold standard. There are several marked limitations of this analysis. Notably, only one study informs each comparison, and with no closed loops in the network, uncertainty and heterogeneity in the included studies will be compounded across the network. In addition, without closed loops in the network, no assessment of consistency can be made. Having only one study to inform a comparison increases uncertainty and relies on the study populations being the same, which is not upheld entirely, particularly with respect to PD-L1 expression. However, limiting the study to the PD-L1 ≥10 population in both studies partly overcomes this. There are differences between the proportions of Asian and non-Asian patients in the trials and the frequency of chemotherapy administration. It is also unknown whether the differences sensitivity of the CPS assays used to detect PD-L1 expression in the trial could introduce any differences between the patient populations. Likewise, the KEYNOTE-590 trial only reports PFS data in the PD-L1 CPS ≥10 population in the mixed histology population. Therefore, it was assumed that this was comparable to the PFS data from the CheckMate 648 trial, which considered an OSCC population. However, this assumption further increases the uncertainty in the analysis.

B.2.10 Adverse reactions

Safety data for NIVO-CHEMO for the treatment of unresectable advanced, recurrent or metastatic, previously untreated OSCC are available from CheckMate 648 for all randomised patients. Nivolumab with chemotherapy was generally well-tolerated, with a similar proportion of patients reporting an AE or treatment-related AE between treatment groups. This is in line with other indications for nivolumab.[69-72] No new safety concerns were identified with NIVO-CHEMO.

B.2.10.1 Overall adverse events

Similar frequencies of all-causality serious adverse events (SAEs) (any grade) occurred in the NIVO-CHEMO (60.0%) and CHEMO (42.8%) arms (Table 18).

Any-grade all-causality AEs leading to discontinuation were reported in [redacted]% of patients in the NIVO-CHEMO arm and [redacted]% of patients in the CHEMO arm. Grade 3-4 all-causality AEs leading to discontinuation were reported in [redacted]% of patients in the NIVO-CHEMO arm and [redacted]% of patients in the CHEMO arm.

Any grade treatment-related AEs (TRAEs) leading to discontinuation were reported in [redacted]% of patients in the NIVO-CHEMO arm and [redacted]% of patients in the CHEMO arm. Grade 3-4 TRAEs leading to discontinuation were reported in [redacted]% of patients in the NIVO-CHEMO arm, and [redacted]% of patients in the CHEMO arm.

A similar number of deaths was also observed between the NIVO-CHEMO and CHEMO arms, [redacted] and [redacted] respectively, with the majority of these attributed to disease progression. Among the [redacted] deaths attributed to study drug toxicity in the NIVO-CHEMO arm, [redacted] were considered related to nivolumab per investigator. Similarly, no death attributed to other in the NIVO-CHEMO arm was assessed as related to nivolumab by the investigator.[64]

Table 18. Overall adverse events: CheckMate 648

Safety parameter	NIVO-CHEMO (n=310)	CHEMO (n=304)
Deaths, n (%)	[redacted]	[redacted]
Primary reason for death		
Disease	[redacted]	[redacted]
Study drug toxicity	[redacted]	[redacted]
Unknown	[redacted]	[redacted]
Other	[redacted]	[redacted]
All-causality AEs		
Any grade	[redacted]	[redacted]
Grade 3-4	[redacted]	[redacted]
All-causality SAEs		
Any grade	[redacted]	[redacted]
Grade 3-4	[redacted]	[redacted]
All-causality AEs leading to discontinuation		
Any grade	[redacted]	[redacted]
Grade 3-4	[redacted]	[redacted]
Treatment-related AEs		
Any grade	[redacted]	[redacted]
Grade 3-4	[redacted]	[redacted]
Treatment-related SAEs		
Any grade	[redacted]	[redacted]
Grade 3-4	[redacted]	[redacted]
Treatment-related AEs leading to discontinuation		
Any grade	[redacted]	[redacted]
Grade 3-4	[redacted]	[redacted]

CI: confidence internal; OS: overall survival; PFS: progression-free survival.
Source: CheckMate 648 [redacted] Summary data[13]

B.2.10.2 Adverse events with potential immunologic aetiology

The most commonly experienced AEs with potential immunologic aetiology (any grade, all cause) were:

- Gastrointestinal, skin and renal ([redacted]%, [redacted]% and [redacted]%, respectively) for patients treated with NIVO-CHEMO (Table 19)
- Renal, gastrointestinal and skin ([redacted]%, [redacted]% and [redacted]%) for patients treated with CHEMO (Table 19)
-

Table 19. Adverse events with potential immunologic aetiology: all causality

Safety parameter	NIVO-CHEMO (n=310)		CHEMO (n=304)	
	Any grade	Grade 3-4	Any grade	Grade 3-4
All causality				
Endocrine	[redacted]	[redacted]	[redacted]	[redacted]
Gastrointestinal	[redacted]	[redacted]	[redacted]	[redacted]
Hepatic	[redacted]	[redacted]	[redacted]	[redacted]
Pulmonary	[redacted]	[redacted]	[redacted]	[redacted]
Renal	[redacted]	[redacted]	[redacted]	[redacted]
Skin	[redacted]	[redacted]	[redacted]	[redacted]

CI: confidence internal; OS: overall survival; PFS: progression-free survival.
Source: CheckMate 648 [redacted] Summary data[13]

The most commonly experienced treatment-related AEs with potential immunologic aetiology of any grade were renal for both NIVO-CHEMO and CHEMO, at [redacted]% and [redacted]% respectively (Table 20).

Table 20. Treatment-related adverse events with potential immunologic aetiology

Safety parameter	NIVO-CHEMO (n=310)		CHEMO (n=304)	
	Any grade	Grade 3-4	Any grade	Grade 3-4
Treatment-related adverse events				
Endocrine	[redacted]	[redacted]	[redacted]	[redacted]
Gastrointestinal	[redacted]	[redacted]	[redacted]	[redacted]
Hepatic	[redacted]	[redacted]	[redacted]	[redacted]
Pulmonary	[redacted]	[redacted]	[redacted]	[redacted]
Renal	[redacted]	[redacted]	[redacted]	[redacted]
Skin	[redacted]	[redacted]	[redacted]	[redacted]

CI: confidence internal; OS: overall survival; PFS: progression-free survival.
Source: CheckMate 648 [redacted] Summary data[13]

The most commonly experienced treatment-related AEs with potential immunologic aetiology leading to discontinuation of any grade were renal for the NIVO-CHEMO and CHEMO arms, [redacted]% and [redacted]%, respectively (Table 21).

Table 21. TRAEs with potential immunologic aetiology leading to discontinuation

Safety parameter	NIVO-CHEMO (n=310)		CHEMO (n=304)	
	Any grade	Grade 3-4	Any grade	Grade 3-4
Endocrine	[redacted]	[redacted]	[redacted]	[redacted]
Gastrointestinal	[redacted]	[redacted]	[redacted]	[redacted]
Hepatic	[redacted]	[redacted]	[redacted]	[redacted]
Pulmonary	[redacted]	[redacted]	[redacted]	[redacted]
Renal	[redacted]	[redacted]	[redacted]	[redacted]
Skin	[redacted]	[redacted]	[redacted]	[redacted]
CI: confidence internal; OS: overall survival; PFS: progression-free survival. Source: CheckMate 648 [redacted] Summary data[13]				

Other events of special interest experienced in the three arms are summarised in Table 22. In the NIVO-CHEMO arm were two cases ([redacted]%) of uveitis and two cases ([redacted]%) of myositis/rhabdomyolysis. There were no events of special interest experienced by patients in the CHEMO group.

Table 22. Other events of special interest summary

Safety parameter	NIVO-CHEMO (n=310)		CHEMO (n=304)	
Myasthenic syndrome	[redacted]	[redacted]	[redacted]	[redacted]
Demyelination event	[redacted]	[redacted]	[redacted]	[redacted]
Guillain-Barre syndrome	[redacted]	[redacted]	[redacted]	[redacted]
Pancreatitis event	[redacted]	[redacted]	[redacted]	[redacted]
Uveitis event	[redacted]	[redacted]	[redacted]	[redacted]
Encephalitis event	[redacted]	[redacted]	[redacted]	[redacted]
Myocarditis event	[redacted]	[redacted]	[redacted]	[redacted]
Myositis/rhabdomyolysis event	[redacted]	[redacted]	[redacted]	[redacted]
Graft versus host disease	[redacted]	[redacted]	[redacted]	[redacted]
Source: CheckMate 648 [redacted] Summary data[13]				

B.2.11 Ongoing studies

Checkmate 648 remains ongoing to further follow-up.

B.2.12 Innovation

Nivolumab is a checkpoint inhibitor immunotherapy agent that utilises the body's immune system to destroy cancer cells (see Section B.1.3.6). The benefits of nivolumab in combination with chemotherapy include:

- **Improved efficacy outcomes vs. standard of care:** In the Checkmate 648 trial, among patients with PD-L1 expression ≥1%:
    - NIVO-CHEMO significantly extended median OS to [redacted] months vs. [redacted] months in patients who received CHEMO, (HR: [redacted]) (Table 10).
    - NIVO-CHEMO demonstrated a statistically significant improvement in median PFS per BICR of [redacted] months compared to [redacted] months in patients treated with CHEMO alone (HR: [redacted] [95% CI: [redacted]]) (Table 11).
    - A higher BICR-assessed ORR (95% CI) was observed with NIVO-CHEMO ([redacted]) vs. CHEMO ([redacted]) (Table 12).
    - Median DOR per BICR (95% CI) was higher in the NIVO-CHEMO arm vs. the CHEMO arm ([redacted] months vs [redacted] months) (Table 13).
- **Maintained quality of life:** As detailed in Section B.2.6.3.4, health-related quality of life was maintained over the course of the treatment period with NIVO-CHEMO.
- **Acceptable safety profile:**
    - Nivolumab has a known safety profile – no new safety signals were identified in CheckMate 648.
    - Rate of treatment-related adverse events (TRAEs) among all treated patients was [redacted]% in the NIVO-CHEMO arm and was [redacted]% in the CHEMO arm.
    - Rate of grade 3-4 TRAEs among all treated patients was [redacted]% in the NIVO-CHEMO group vs. [redacted]% in the CHEMO group.
    - Rate of grade 3-4 TRAEs leading to discontinuation among all treated patients was [redacted]% for the NIVO-CHEMO group and [redacted]% in the CHEMO group.
- **An additional treatment option for patients with high unmet need:**
    - Systemic treatment options at first line for unresectable advanced, recurrent or metastatic OSCC are limited to chemotherapy only and pembrolizumab combined with chemotherapy for patients who are HER2 negative with CPS ≥ 10.

- Outcomes in OC are poor, with only ~15% of people diagnosed with OC surviving for five years or more (2013-2017).[11]

- Addition of nivolumab to chemotherapy, with proven efficacy and tolerability, provides a new systemic treatment option for OSCC patients. This is especially significant for patients with PD-L1 CPS <10%, who do not have an immunotherapy option.

In summary, adoption of nivolumab in combination with chemotherapy by NHS England would represent a significant advancement in the management of this life-threating condition.

B.2.13 Interpretation of clinical effectiveness and safety evidence

B.2.13.1 Principal findings from the clinical evidence

Prognosis for advanced OSCC remains poor: a retrospective analysis has shown that nearly half of patients undergoing systemic treatment for OSCC in the first-line setting do not respond to their treatment and over a third of patients progress to the next line of treatment.[42] There is currently a high unmet need for effective first-line treatments for patients with advanced OC, particularly in patients with PD-L1 CPS <10, where doublet palliative chemotherapy is the only therapy available.

CheckMate 648 demonstrated that the use of nivolumab with fluoropyrimidine- and platinum-based chemotherapy has significant benefits when considered as a first-line for unresectable advanced, recurrent or metastatic OSCC:

- Among patients with PD-L1 ≥1, there was a significant median OS benefit of 6.3 months (HR: 0.54 [0.37, 0.80]) and PFS benefit of 2.5 months (HR: 0.77 [0.64, 0.92]) in the NIVO-CHEMO group vs. CHEMO.
- Among all randomised patients, there was a significant median OS benefit of [redacted] months in the NIVO-CHEMO group vs. CHEMO.
- Health-related quality of life was maintained over the course of the treatment period with NIVO-CHEMO.
- No new safety signals were identified for nivolumab and fewer than half of the TRAEs in the NIVO-CHEMO group were grade 3 or 4.

Overall, NIVO-CHEMO demonstrated a favourable benefit-risk profile in patients with previously untreated, unresectable advanced, recurrent or metastatic OSCC.

B.2.13.2 Strengths and limitations of the clinical evidence base

B.2.13.2.1 Limitations of the study evidence

The clinical efficacy of nivolumab is informed using the phase III, CheckMate 648 trial, which had an open-label study design. However, this minor limitation should not affect the generalisability of the study to the UK population and should be viewed within the context of the study's strengths and the high unmet need in this patient population.

- **Open-label study design:** The open-label study design of CheckMate 648 makes it possible that the knowledge of the treatment could have influenced patient responses with regards to health-related quality of life. However, an open-label design was considered appropriate because of the differences in the dosing regimens and associated toxicities for each treatment group. The primary endpoint of overall survival is an objective measure, which would not be affected by the open-label nature of the study. Furthermore, involvement of an independent data monitoring committee for safety assessments ensured anonymity of the treatment groups during data review.

B.2.13.2.2 Strengths of the study evidence

- **Robust study design:** CheckMate 648 is a well-designed, high-quality phase III randomised controlled trial, which provides direct comparative evidence on the clinical efficacy of nivolumab with chemotherapy compared to chemotherapy alone. The patient cohorts in each arm were large and randomised 1:1:1 (321, 325 and 324 in the nivolumab with chemotherapy, nivolumab with ipilimumab and chemotherapy arms, respectively). There are no other large scale trials where the primary endpoints consider this patient population, patients with PD-L1 CPS ≥1. The CheckMate 648 trial included a large proportion of patients with PD-L1 ≥1, 49% in the NIVO-CHEMO arm and 48% in the CHEMO arm.

- **Maturity of survival data:** CheckMate 648 provides survival data that may be considered relatively mature, placing less reliance on the need for survival extrapolation though parametric curve fitting.

- **Both primary endpoints were met:** Both primary endpoints, OS and PFS in patients with PD-L1 ≥1, were met. Among patients with PD-L1 ≥1, there was a significant median OS benefit of [redacted] months in the NIVO-CHEMO group vs. CHEMO, and a significant PFS benefit of [redacted] months found in the NIVO-CHEMO group vs. CHEMO.

- **Relevant comparator:** The CheckMate 648 trial compared the safety and efficacy of NIVO-CHEMO compared to CHEMO alone. The CHEMO regimen used was cisplatin-5-fluorouracil, which is considered standard of care in the UK, and therefore, was the most appropriate comparator when considering the benefits of introducing NIVO-CHEMO.

- **Inclusion of non-Asian study populations:** OC clinical studies typically enrol most patients in Asian countries, due to the higher incidence rate of OC, and particularly OSCC, in these countries. Aligned with this, most patients recruited in CheckMate 648 were from Asian countries. However, CheckMate 648 included a substantial proportion of non-Asian patients, so that the patient population is more reflective of that observed in UK clinical practice. Further, the results are felt to be applicable to both Asian and non-Asian populations:

    - Studies show little variation in genome-wide mutations, gene expression profiles or gene methylations between Asian and Caucasian cancer patients, reflecting the common characteristics of OSCC tumours from different populations.[73]

    - In NICE TA737, one clinical expert explained that the OSCC biology and aetiology for Asian vs. non-Asian patients are similar.[1] Further, the treatment paradigms for advanced OSCC (and oesophageal adenocarcinoma) are similar in Europe, the US and Asia, with standard treatment being platinum and fluoropyrimidine chemotherapy, as highlighted in the combined international ESMO-JSMO guidelines (2018).[74]

    - A previous NICE submission (TA746) considering nivolumab monotherapy for unresectable, advanced oesophageal cancer after standard chemotherapy has failed, conducted an SLR to determine the differences in patient characteristics and survival outcomes between Asian and Western population.[75] The SLR found that OS was comparable between Asian and Western populations with OSCC (median OS: 7.5 versus 7.4 months; mean one-year survival was 21.1% in Asian and 27.9% in Western patients).[72]

    - Further, during an advisory board held by BMS, UK clinicians felt that there was no biological reason to consider the populations to be different.[2]

- As demonstrated in Section B.2.13.4.1, the baseline characteristics of patients enrolled in CheckMate 648 aligned closely with OSCC patient cohorts from similar studies conducted in the UK (Table 23).

B.2.13.3 Relevance of the evidence base to the decision problem

The submission presents evidence from the CheckMate 648 study, which studied the safety and efficacy of NIVO-CHEMO in patients with untreated, advanced OSCC, in line with the decision problem. The trial demonstrates the clinical efficacy of NIVO-CHEMO and provides evidence for the beneficial impact of nivolumab with chemotherapy in a Western patient population. Further, outcomes considered in the submission closely mirror the decision problem set out by NICE.

Thus, the evidence base presented within this submission represents the best available evidence and is directly relevant to the decision problem.

B.2.13.4 External validity of study results to patients in routine clinical practice

Patients enrolled in the available studies can be considered broadly representative of UK practice, in terms of baseline characteristics, with subgroups provided for analysis where possible (see Section B.2.7).

B.2.13.4.1 Relevance to the UK patient population

As discussed in Section B.2.13.2.2, OC clinical studies typically enrol most patients in Asian countries, due to the higher incidence rate of OC, and particularly OSCC, in these countries. Aligned with this, most patients recruited in CheckMate 648 were from Asian countries. However, CheckMate 648 included a substantial proportion of non-Asian patients, so that the patient population is more reflective of that observed in UK clinical practice. Analysis in Asian and non-Asian patient subgroups showed favourable survival outcomes for nivolumab with chemotherapy in both subgroups (Section B.2.7). During an advisory board held by BMS, UK clinicians did not feel that the study location would effect the applicability of the results to the UK setting.[2] As outlined in Section B.2.13.2.2, this is aligned with broader evidence supporting the applicability of evidence from CheckMate 648 to the UK population.

Similarly, CheckMate 648 also considered a highly relevant comparator during the trial (cisplatin-5-fluorouracil), which is considered standard of care in the UK, and therefore, the benefits demonstrated with NIVO-CHEMO compared to CHEMO alone are directly applicable to current UK clinical practice. Section B.2.13.4.2 further explores alternative comparators used in UK clinical practice.

The baseline characteristics of patients enrolled in CheckMate 648 are similar to those enrolled in other UK studies of OSCC. The median age of patients enrolled in CheckMate-648 was 64 years, which aligns to the median age of patients enrolled in similar UK studies of OSCC. For instance, Shyamalee et al. (2021), a real-world evidence study of the outcomes of UK OSCC patients treated with best supportive care, observed a median age of 63 years for patients treated with palliative chemotherapy.[76] Likewise, the western cohort of OSCC patients who initiated first-line therapy included in the non-observational study conducted by Jaffe et al. (2022) had a mean age of 62.9 years.[36]

Slightly few patients with ECOG status of 0 were enrolled in CheckMate 648 compared to the Shyamalee study. Clinical trials commonly specify performance scores as an inclusion criterion, typically based on either ECOG or Karnofsky scale. This leads to limited evidence of net clinical benefit for patients with certain performance scores, typically those with worse scores. This absence of evidence contributes to a reluctance to provide certain treatments to patients of reduced performance score. However, this is limited evidence to suggest different outcomes between patients with different performance score.

A 2017 SLR and meta-analysis of RCTs assessed clinical benefit by performance score subgroups. This identified 110 RCTs, with 66 (60%) reporting performance score subgroups for efficacy and none reporting subgroups for toxicity. For these 66 RCTs, pooled HRs for good performance score and reduced performance score subgroups were 0.65 (95% CI: 0.61-0.70) and 0.67 (95% CI: 0.62- 0.72), respectively, with no difference between the two groups (p=0.68). Sensitivity analyses based on drug or cancer type and type of endpoints (OS or PFS) demonstrated similar results.[77]

A comparison of the CheckMate 648 trial population with the patient populations included in the Shyamalee and Jaffe studies is presented in Table 23. This demonstrates that the baseline characteristics of trial population of CheckMate 648 are comparable to those of other UK OSCC cohorts and so can be considered broadly representative of those patients seen in UK clinical practice.

Table 23. A comparison of the baseline characteristics of patients in the CheckMate 648 trial with those in the Shyamalee and Jaffe UK studies

Baseline characteristic		CheckMate 648		Shyamalee[76]	Jaffe[36]
		NIVO-CHEMO	CHEMO		
Cohort size		321	324	219	1,049
Age	Median (range), years	64 (40-90)	64 (26-81)	63	-
	Mean (SD), years	-	-	-	62.9 (10.6)
Sex	Male (%)	78.8	84.9	48	82.7
ECOG PS, n (%)	0	150 (46.7)	154 (47.5)	6 (27)	-
	1	171 (53.3)	170 (52.5)	9 (41)	-
CHEMO: chemotherapy; ECOG PS: Eastern cooperative oncology group performance scale; NIVO-CHEMO: nivolumab with chemotherapy;					

B.2.13.4.2 Comparison of CheckMate 648 with published evidence

B.2.13.4.2.1 Comparison with UK studies

CheckMate 648 can be considered highly relevant to UK clinical outcomes. Although no studies were identified to assess UK outcomes or baseline characteristics in advanced OSCC patients, studies are available for gastro-oesophageal adenocarcinoma. Table 24 shows a comparison of baseline characteristics versus UK specific studies. CheckMate 648 enrolled a similar proportion of male patients to other UK studies. Similarly, the median age was comparable to UK studies. However, patients in CheckMate 648 typically had a lower ECOG status and were more likely to have locally advanced or recurrent disease.

Table 24. Comparison of CheckMate 648 baseline characteristics versus those from UK-specific studies

		CheckMate 648[63,78]		Cougar-2[79]		Royal Marsden retrospective review[80]
		NIVO+ CHEMO	CHEMO	Docetaxel	Active symptom control	
N		**158**	**157**	**84**	**84**	**511**
Sex, male (%)		125 (79%)	131 (83%)	69 (82%)	67 (80%)	384 (75%)
Median age (range), years		64 (40–85)	62 (28–81)	65 (28–84)	66 (36–84)	66 (24-90)**
ECOG status	0	71 (45%)	70 (45%)	24 (28%)	22 (26%)	64 (13%)
	1	87 (55%)	86 (55%)	46 (55%)	50 (60%)	276 (54%)
	2	0	0	14 (17%)	12 (14%)	87 (17%)
Disease status	Locally advanced or recurrent	[redacted]	[redacted]	11 (13%)	10 (12%)	68 (13%)*
	Metastatic disease	[redacted]	[redacted]	73 (87%)	74 (88%)	335 (66)*
Histology	Adenocarcinoma	-	-	84 (100%)	84 (100%)	511 (100%)
	Squamous cell carcinoma	156 (99)	155 (99)	-	-	-

* 21% of patients had relapsed metastatic disease after radical treatment.
** Age at diagnosis, not study baseline
Baseline characteristics and demographics presented for CheckMate 648 patients with tumour cell PD-L1 ≥1%

Prognosis is notably poor for patients with locally advanced or metastatic gastro-oesophageal cancer. Although a small proportion of patients demonstrate improved outcomes versus the overall cohort, this proportion is very small. For this reason, it may be implausible for approximately 10% of patients receiving chemotherapy to survive at 36 months, as observed during CheckMate 648.

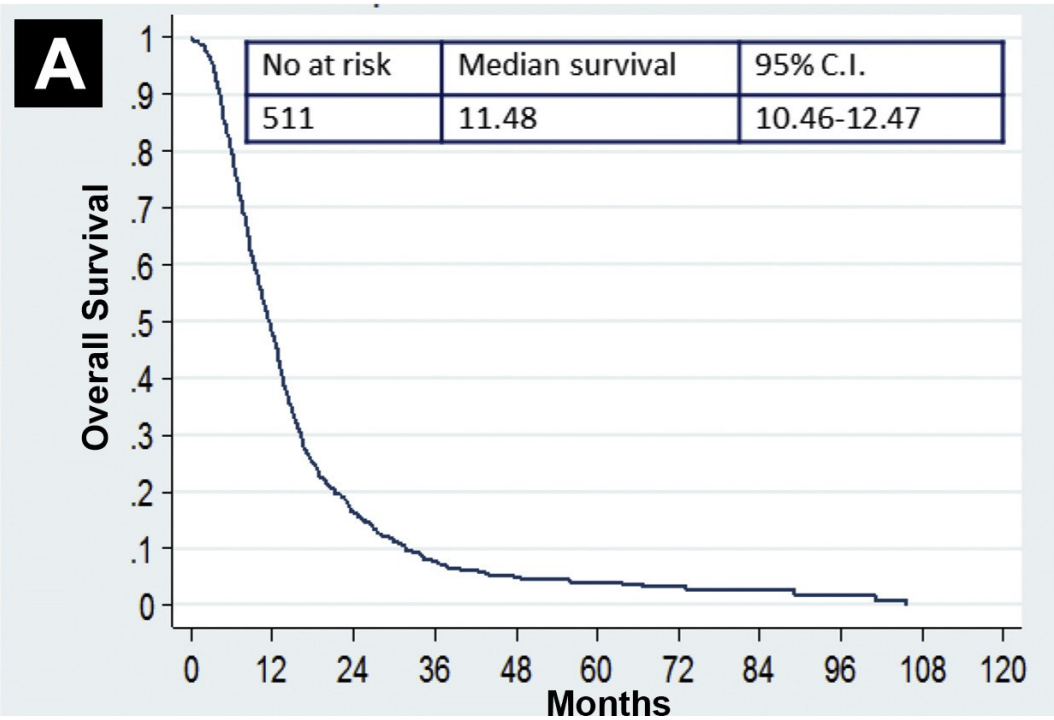
A comparison between CheckMate 648 and previously published studies is provided below.

B.2.13.4.2.1.1 Royal Marsden retrospective review

A retrospective analysis was undertaken of patients who had received at least one cycle of chemotherapy for gastro-oesophageal adenocarcinoma in the advanced disease setting at the Royal Marsden Hospital from April 2009 to November 2015.[81] Baseline characteristics are described in Table 24.

Median survival was slightly longer than observed during CheckMate 648 (11.5 months versus 9.07 months). However, survival at 24 months was generally comparable between studies, as observed in Figure 26. Although this is followed by a low hazard, there are less than 10% of patients surviving at 36 months and this continues to decrease over the long-term follow up.

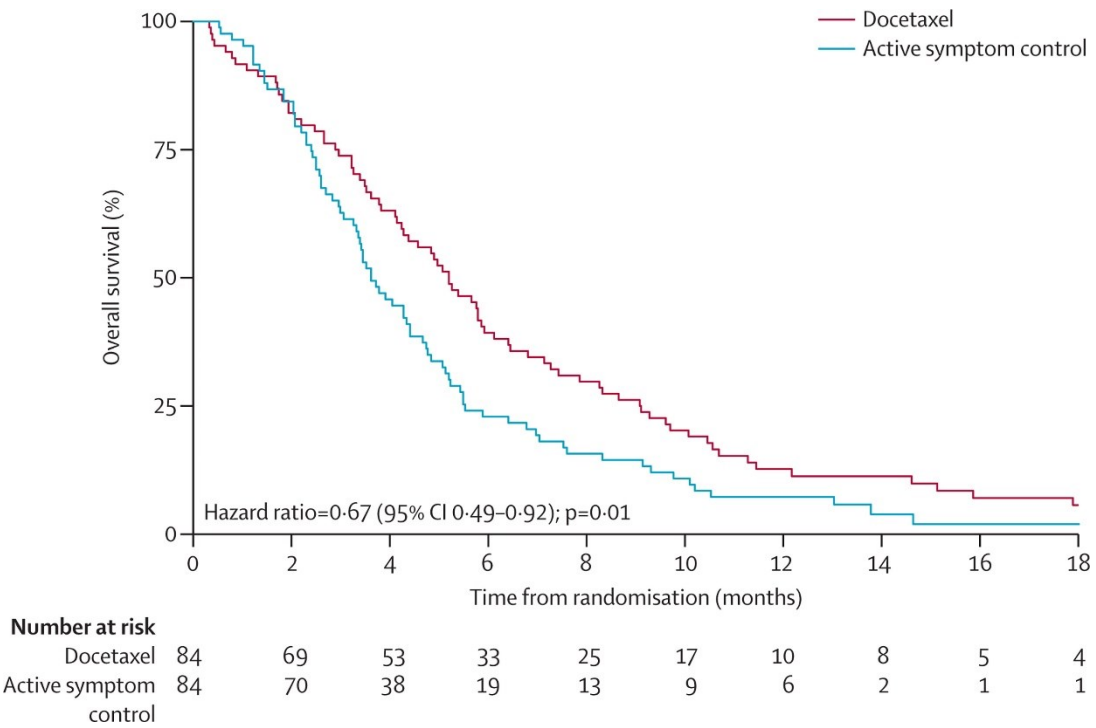
Figure 26. Overall survival for patients receiving chemotherapy for gastro-oesophageal adenocarcinoma at the Royal Marsden Hospital[81]



B.2.13.4.2.1.2 COUGAR-2

COUGAR-2 was a randomised, controlled trial assessed docetaxel versus active symptom control in previously treated UK patients with advanced gastro-oesophageal adenocarcinoma.[79] Median OS was 5.2 months in patients receiving docetaxel and 3.6 months in patients receiving active symptom control. However, patients continue to experience increased hazard over time, as illustrated in Figure 27, although this is limited by lack of follow-up.

Figure 27. Overall survival during COUGAR-2[79]



B.2.13.4.2.2 Published evidence describing outcomes for immunotherapy treatment

Immunotherapies are extensively studies in gastro-oesophageal cancers. One systematic literature review (SLR) of immune checkpoint inhibitors for gastro-oesophageal cancers identified six studies in the first-line setting, eight studies in the second-line setting and three studies assessing maintenance treatment.[82] This analysis demonstrated statistically significant overall survival benefit in the first-line setting in gastric and gastro-oesophageal adenocarcinomas (hazard ratio [HR] = 0.83, 95% confidence interval [CI] = 0.76 to 0.90, P <0.001; based on 4 studies) and OSCC (HR = 0.72, 95% CI = 0.64 to 0.81, P <0.001; based on 3 studies). Additionally, patients in the second-line setting with OSCC derive survival benefit from immunotherapies in the second-line setting (HR = 0.74, 95% CI = 0.68 to 0.82, P <0.001).[82]

Studies identified during this SLR that assessed previously untreated gastro-oesophageal cancer in a UK-relevant population are described below.

B.2.13.4.2.1 KEYNOTE-590

KEYNOTE-590 was a randomised, placebo-controlled, double-blind, phase 3 study assessing pembrolizumab plus chemotherapy or placebo plus chemotherapy in patients with previously

untreated, locally advanced, unresectable or metastatic oesophageal cancer or Siewert type 1 gastro-oesophageal junction cancer.[83]

Baseline characteristics were similar between KEYNOTE-590 and CheckMate 648 (Table 25). Age was similar between study populations, as was the proportion of males enrolled, while ECOG status was slightly lower in patients enrolled onto CheckMate 648. However, 27% of patients in KEYNOTE-590 had adenocarcinoma, whereas these patients were not eligible for CheckMate 648. Additionally, CheckMate 648 enrolled more Asian patients than KEYNOTE-590 and fewer patients with metastatic disease. These key differences are all considered to be influential in long-term outcomes.[83]

As can be seen in Figure 28, outcomes in KEYNOTE-590 are broadly aligned with CheckMate 648 (median OS for OSCC subgroup: 12.6 months for pembrolizumab versus 9.8 months for chemotherapy; OSCC and PD-1 combined positive score [CPS] ≥ 10: 13.9 months versus 8.8 months).[83] Similar to CheckMate 648, survival for chemotherapy decreased below 20% by 18 months, but then hazard decreased and few events are observed after this point. Approximately 30% of patients in the pembrolizumab arm remain alive at 24 months, which is broadly comparable to the nivolumab arm of CheckMate 648; however, in KEYNOTE-590, this hazard continues to decline, which is not observed in CheckMate 648.

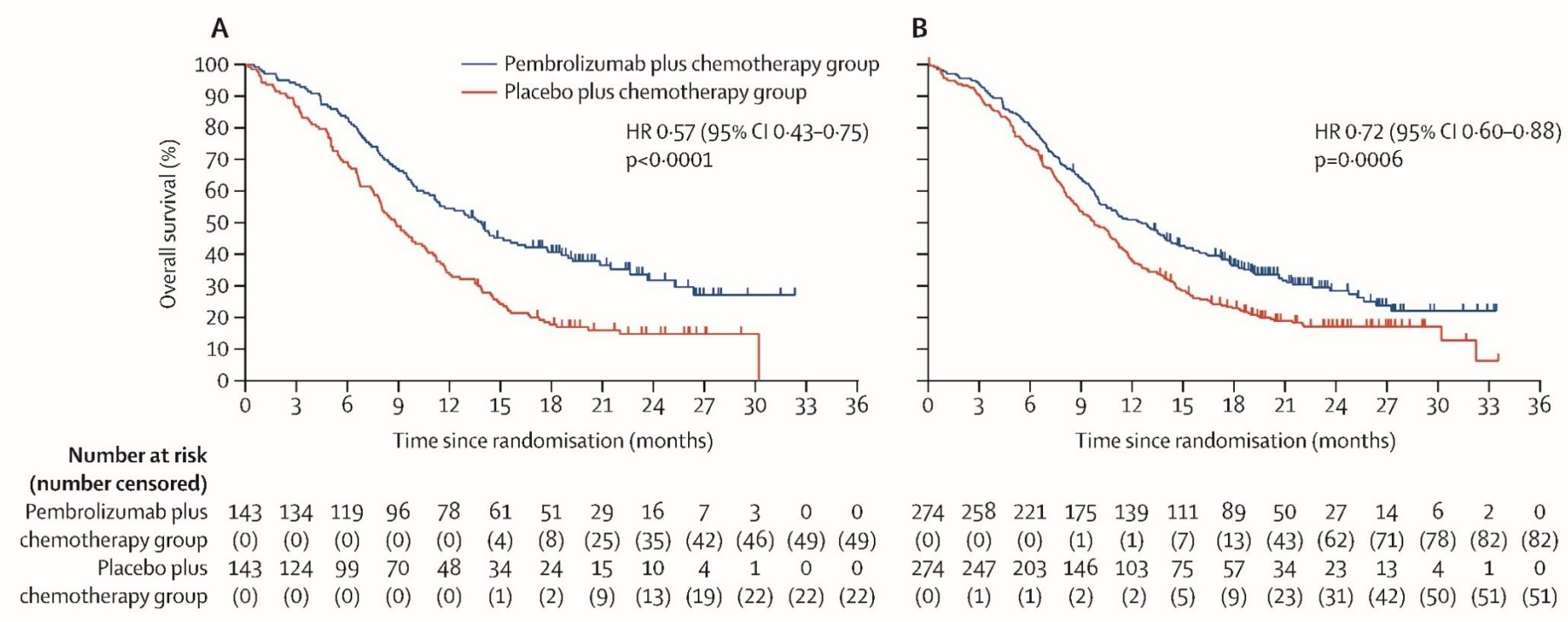
Of note, response durations of 24 months or longer occurred in 18% of patients in the pembrolizumab plus chemotherapy group and 6% of patients in the placebo plus chemotherapy group, which is broadly aligned with CheckMate 648.[83]

Similar to CheckMate 648, 161 (43%) patients in the pembrolizumab plus chemotherapy group versus 177 (47%) in the placebo plus chemotherapy group received subsequent anticancer therapy.[83] However, 22 (6%) in the pembrolizumab plus chemotherapy group versus 35 (9%) in the placebo plus chemotherapy group received subsequent immunotherapy.[83]

Table 25. Comparison of KEYNOTE-590 and CheckMate 648

	KEYNOTE-590[83]				CHECKMATE 648 (PD-L1 ≥1%[63]			
	Pembrolizumab +CHEMO (n=373)		CHEMO (n=376)		NIVO+CHEMO (n=158)		CHEMO (n=157)	
Age, (years)								
Median (range)	64 (28–94)		62 (27–89)		64 (40-85)		62 (28-81)	
Sex								
Male	306	82%	319	85%	125	79%	131	83%
Asia region	196	53%	197	52%	114	72%	113	72%
Race								
Asian	201	54%	199	53%	116	73%	113	72%
White	139	37%	139	37%	38	24%	38	24%
ECOG performance status								
0	149	40%	150	40%	71	45%	70	45%
1	223	60%	225	60%	87	55%	86	55%
2	1	<1%	1	<1%	-	-	-	-
Oesophageal squamous cell carcinoma	274	73%	274	73%	156	99%	155	99%
Adenocarcinoma	99	27%	102	27%	-	-	-	-
Oesophageal adenocarcinoma	58	16%	52	14%	-	-	-	-
Siewert type 1 gastro-oesophageal junction adenocarcinoma	41	11%	50	13%	-	-	-	-
Disease status								
Metastatic	344	92%	339	90%	85	54%	89	57%
Unresectable locally advanced	29	8%	37	10%	73	46%	68	43%
PD-L1 CPS ≥10	186	50%	197	52%	-	-	-	-

Figure 28. KEYNOTE-590 overall survival outcomes; A) Patients with oesophageal squamous cell carcinoma and PD-L1 CPS of 10 or more; B) Patients with oesophageal squamous cell carcinoma[83]



B.2.13.4.2.2 CheckMate 649

CheckMate 649 (NCT02872116) is a Phase III, open-label, randomised, multi-centre trial assessing NIVO+CHEMO, NIVO+IPI or CHEMO in previously untreated advanced gastric, GOJ or oesophageal adenocarcinoma.[84]

Baseline characteristics were broadly comparable between studies, as shown in Table 26, with the exception of race (CheckMate 649 enrolled fewer Asian patients) and tumour site and histology, due to the differing eligibility criteria.

CheckMate 649 reported short median OS (11.6 months) for patients receiving chemotherapy[84] (Figure 29). Aligned with CheckMate 648, approximately 20% of the patients in the chemotherapy arm are alive at 24 months; although the hazard decreases after this point, survival is just above 10% by 36 months. However, unlike CheckMate 648, outcomes for NIVO+CHEMO remain at slightly below 20% at 36 months, with no events occurring in the long-term follow up.[84]

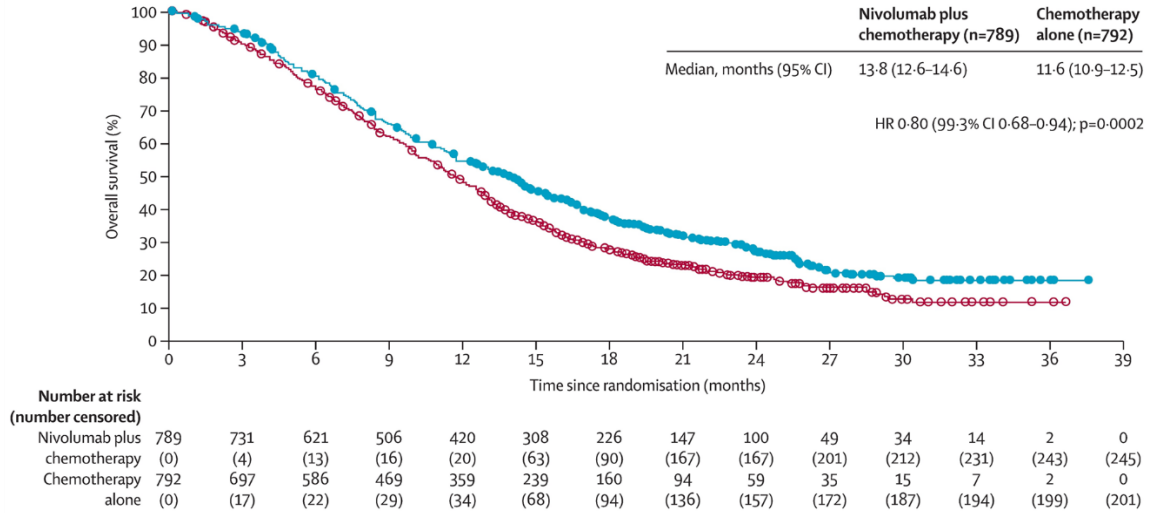
Similar to CheckMate 648, more patients during CheckMate 649 in the CHEMO arm went on to receive subsequent treatment (41% versus 38% for NIVO+CHEMO), systemic treatments (39% versus 34% for NIVO+CHEMO) and specifically an anti-PD-1/PD-L1 therapy (8% versus 2% for NIVO+CHEMO).[84]

Table 26. Comparison of baseline characteristics for CheckMate 649 and CheckMate 648

	CheckMate 649[84]		CheckMate 648[63]	
	NIVO+CHEMO N=789	CHEMO N=792	NIVO+CHEMO N=789	CHEMO N=792
Median age, years (range)	62 (54-69)	61 (53-68)	64 (40–85)	62 (28–81)
Sex, male (%)	540 (68)	560 (71)	125 (79)	131 (83)
Race, n (%)				
White	556 (70)	541 (68)	38 (24)	38 (24)
Asian	186 (24)	189 (24)	116 (73)	113 (72)
Region, n (%)				
Asia	178 (23)	178 (22)	114 (72)	113 (72)
Initial diagnosis, n (%)				
Gastroesophageal junction cancer	132 (17)	128 (16)	-	-
Gastric cancer	554 (70)	556 (70)	-	-
Oesophageal adenocarcinoma	103 (13)	108 (14)	-	-
OSCC	-	-	156 (99)	155 (99)
Disease status classification, n (%)				
Locally recurrent	5 (1)	2 (<1)	53 (34)	41 (26)
Metastatic	757 (96)	756 (95)	85 (54)	89 (57)
Locally advanced	27 (3)	34 (4)	20 (13)	18 (11)
ECOG PS				
0	326 (41)	336 (42)	71 (45)	70 (45)
1	462 (59)	452 (57)	87 (55)	86 (55)

CHEMO: chemotherapy; ECOG: Eastern Cooperative Oncology Group; PS: performance status.
Source: Janjigian et al 2021[84]

Figure 29. CheckMate 649 overall survival[84]



B.2.13.4.2.3 KEYNOTE 062

KEYNOTE-062 was a randomised, controlled, partially-blinded, phase 3 study assessing pembrolizumab (with or without chemotherapy) or chemotherapy in patients with previously untreated, locally advanced, unresectable or metastatic gastric cancer or gastro-oesophageal junction cancer with PD-L1 CPS of 1 or greater.[85] The pembrolizumab monotherapy arm is not further discussed, as the pembrolizumab plus chemotherapy arm provides a more relevant comparison to CheckMate 648.

Baseline characteristics were comparable between studies, with the exception that KEYNOTE-062 enrolled fewer male patients and CheckMate 648 enrolled fewer patients with metastatic disease. Additionally, tumour site and histology differed due to eligibility criteria differences.

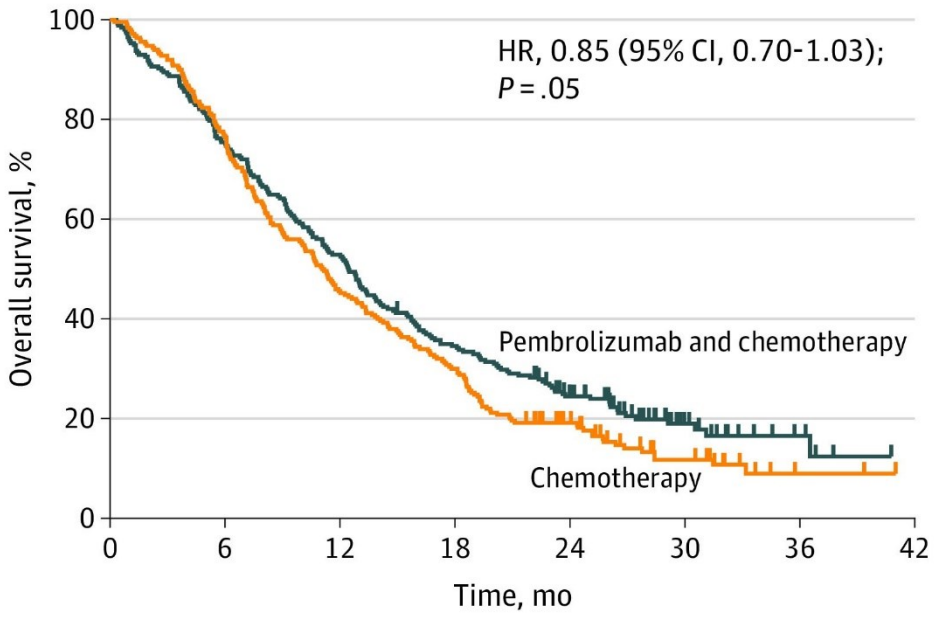
As can be seen in Figure 30, outcomes in KEYNOTE-062 are aligned with CheckMate 648 (median OS: 12.5 months for pembrolizumab plus chemotherapy versus 11.1 months for chemotherapy).[85] Similar to CheckMate 648, survival for chemotherapy was approximately 20% by 24 months, reaching approximately 10% by 30months. [85]

Similar to CheckMate 648, more patients during KEYNOTE-062 in the CHEMO arm went on to receive subsequent treatment (54% versus 47% for pembrolizumab plus chemotherapy) and specifically an anti-PD-1/PD-L1 therapy (13% versus 4% for pembrolizumab plus chemotherapy).[85]

Table 27. Comparison of KEYNOTE-062 and CheckMate 648 baseline characteristics

	KEYNOTE-062[85]				CHECKMATE 648 (PD-L1 ≥1%[63]			
	Pembrolizumab +CHEMO (n=257)		CHEMO (n=250)		NIVO+CHEMO (n=158)		CHEMO (n=157)	
Age, (years)								
Median (range)	62 (22–83)		62.5 (23–87)		64 (40-85)		62 (28-81)	
Sex								
Male	195	76%	179	72%	125	79%	131	83%
Asia region	64	25%	61	24%	114	72%	113	72%
ECOG performance status								
1	138	54%	135	54%	87	55%	86	55%
Oesophageal squamous cell carcinoma	-	-	-	-	156	99%	155	99%
Adenocarcinoma	257	100%	250	100%	-	-	-	-
Gastric adenocarcinoma	170	61%	181	72%	-	-	-	-
Siewert type 1 gastro-oesophageal junction adenocarcinoma	85	33%	67	27%	-	-	-	-
Disease status								
Metastatic	243	95%	235	94%	85	54%	89	57%
PD-L1 CPS ≥10	99	39%	90	36%	-	-	-	-

Figure 30. KEYNOTE-062 overall survival



No. at risk (No. censored)

Pembrolizumab and chemotherapy	257 (0)	194 (0)	136 (0)	88 (0)	52 (0)	17 (23)	5 (44)	0 (50)
Chemotherapy	250 (0)	192 (0)	114 (0)	75 (0)	38 (0)	15 (18)	2 (29)	0 (32)

B.2.13.4.2.3 Conclusions on comparison with published evidence

CheckMate 648 baseline characteristics and outcomes are well aligned to the published evidence base, and so can be considered highly relevant to UK clinical practice.

However, it should be noted that long-term outcomes for the CheckMate 648 CHEMO arm are more optimistic than observed during previous studies, while the NIVO-CHEMO has higher long-term hazard than observed in other immuno-oncology therapies for gastro-oesophageal cancer. This may be driven by the high rate of subsequent treatment use in the CHEMO arm. Alternatively, this be confounded by the limited patient numbers informing long-term follow up, so that the responder population has an outsized contribution to the long-term follow up. However, these data challenges impact on interpretation of the study and long-term extrapolation of outcomes.

B.2.13.4.2 UK standard of care

As outlined in Section B.1.3.5, UK guidelines recommend chemotherapy for patients who have previously untreated, unresectable advanced, recurrent or metastatic OSCC, including doublet treatment with 5-fluorouracil or capecitabine in combination with cisplatin or oxaliplatin or triplet treatment including epirubicin with a fluoropyrimidine (5-fluorouracil or capecitabine) and a platinum agent (cisplatin or oxaliplatin).[12] However, as outlined in Section B.1.1, doublet chemotherapies are more commonly used for treatment of OSCC in the UK. Use of epirubicin-based triplet therapies for OSCC is declining in the UK, as there is limited evidence to support clinical benefit in the context of increased adverse events. This is confirmed by TA737, where clinical experts contacted by NICE explained that dual therapy regimens are preferred, while the Cancer Drugs Fund clinical lead confirmed that the use of triple regimens is rapidly diminishing.[1] In this context, the committee concluded that a dual chemotherapy regimen would be the appropriate comparator for TA737.[1] During an advisory board conducted by BMS, NHS clinicians confirmed that triplet therapy is rarely used in the UK and would not be considered standard of care.[2]

Clinical advisors to BMS have also confirmed that doublet chemotherapies have similar outcomes in cases of advanced OSCC.[2] This is aligned with clinical advice obtained during TA737, where clinicians advised the ERG that doublet regimens are of exchangeable effectiveness (i.e. exhibit a class effect).[86] However, clinical advice to the ERG also noted that regimens with fluorouracil are rarely given due to the lengthy infusion time, with use only where patients are unable to swallow capecitabine tablets.[86] Further, clinical experts stated during TA737 that oxaliplatin is more commonly used than cisplatin, as it is better tolerated and has a shorter infusion time.[1] As a result, clinical experts felt that capecitabine plus oxaliplatin

(XELOX) would be the primary comparator in this patient population.[1] This aligns with clinical advice during an ongoing gastro-oesophageal adenocarcinoma NICE appraisal.[87]

Despite this, the CheckMate 648 study included a chemotherapy comparator arm, comprising of 5-fluorouracil and cisplatin, which is a relevant comparator to the UK setting for treatment of untreated, unresectable OSCC, as confirmed by the NHS clinicians during the advisory board and in line with the NICE scope for this indication.[2]

Additionally, the decision problem includes pembrolizumab plus chemotherapy. However, it should also be noted that pembrolizumab was only recently recommended by NICE (October 2021)[1] and is hence not yet standard of care. Further, pembrolizumab plus chemotherapy is only used in a subgroup of the relevant patient population (patients with PD-L1 CPS ≥10) and hence is only relevant to part of the decision problem. This was not included in the CheckMate 648 trial as the trial was initiated prior to the approval of pembrolizumab. Despite this, an ITC comparing NIVO-CHEMO with pembrolizumab plus chemotherapy is presented in Section B.2.9.

B.2.13.4.3 Measurement of PD-L1 in UK clinical practice

Assessment of PD-L1 status is not yet clinical practice in the UK for patients with OSCC, although this will be changing following availability of pembrolizumab for patients with PD-L1 CPS ≥10 (October 2021)[1] PD-L1 CPS is a scoring method that evaluates the number of PD-L1 positive cells (tumour, lymphocytes and macrophages) divided by the total number of tumour cells, multiplied by 100[88]. Hence, it is a composite score that allows the capture of PD-L1 positive tumour and immune cells in a single reading.[89] For the purposes of clinical trials, this is preferred over tumour PD-L1 score, which only reflects the percentage of tumour cells that are positive for PD-L1 expression. However, it is not yet confirmed which approach will be used in clinical practice for patients with OSCC.

However, assessment of PD-L1 has been rapidly evolving, so that clinical trials often use tumour cell PD-L1 score, as per CheckMate 648, which applied the tumour cell PD-L1 measure to define the patient subgroup for the primary endpoint. These measures do have significant overlap, as shown in Table 28. Of the 158 patients in the NIVO-CHEMO arm with tumour cell PD-L1 ≥1%, [redacted] also had CPS ≥10. Similarly, in the CHEMO arm, 157 patients had tumour cell PD-L1 ≥1% and of these [redacted] had CPS ≥10. Further, although outcomes are slightly better in patients with CPS ≥10 (Table 29), there is significant benefit in patients with tumour cell PD-L1 ≥1%.

As such, further subgrouping of the OSCC patient population may improve average patient outcomes but would exclude patients who would derive significant benefit from immuno-oncology therapies and are currently limited to standard chemotherapy options.

Table 28. CheckMate 648 frequency of PD-L1 by SPC status

	NIVO-CHEMO		CHEMO	
	ITT	Tumour cell PD-L1 ≥1%	ITT	Tumour cell PD-L1 ≥1%
ITT	[redacted]	[redacted]	[redacted]	[redacted]
ITT with CPS score	[redacted]	[redacted]	[redacted]	[redacted]
CPS≥5	[redacted]	[redacted]	[redacted]	[redacted]
CPS≥10	[redacted]	[redacted]	[redacted]	[redacted]

Table 29. CheckMate 648: impact of alternative PD-L1 measurement scores on outcomes

		Tumour cell PD-L1 ≥1%		CPS ≥10	
		NIVO-CHEMO	CHEMO	NIVO-CHEMO	CHEMO
N		[redacted]	[redacted]	[redacted]	[redacted]
OS (months)	Median	[redacted]	[redacted]	[redacted]	[redacted]
	Restricted mean	[redacted]	[redacted]	[redacted]	[redacted]
PFS (months)	Median	[redacted]	[redacted]	[redacted]	[redacted]
	Restricted mean	[redacted]	[redacted]	[redacted]	[redacted]

B.2.13.4.3 Application of NICE end-of-life criteria to nivolumab with chemotherapy use in oesophageal cancer

Outcomes are known to be poor in OSCC patients with untreated, unresectable advanced, recurrent or metastatic disease, although there is a paucity of evidence describing this patient population. These patients have limited treatment options and estimates of OS at 1 year are around 44%, as reported in patients in the chemotherapy arm from CheckMate 648.[60] Therefore, there is a high degree of unmet clinical need in this patient population, which would be addressed by the availability of nivolumab with chemotherapy.

The case for application of NICE end-of-life criteria for nivolumab with chemotherapy for the treatment of OSCC is set out in Table 30, and based on this evidence, nivolumab is considered to meet both criteria for end-of-life.

Table 30. End-of-life criteria

Criterion	Data available	Reference in submission (section and page number)
The treatment is indicated for patients with a short life expectancy, normally less than 24 months	Available therapies in patients with untreated, unresectable, advanced, recurrent or metastatic OSCC are associated with poor outcomes, although data describing this patient population are limited. Based on available data, median OS for platinum-based chemotherapy, 5-fluorouracil and cisplatin, as observed during CheckMate 648, was 10.7 months.	B.2.6.3.1
There is sufficient evidence to indicate that the treatment offers an extension to life, normally of at least an additional 3 months, compared with current NHS treatment	The mean OS is more representative of the survival benefit associated with nivolumab with chemotherapy. However, it is acknowledged that extrapolated outputs are subject to uncertainty due to the potential variation in extrapolations. However, when data are restricted to the observed period, restricted mean OS is [redacted] months in the nivolumab with chemotherapy arm and [redacted] months in the chemotherapy arm, providing [redacted] months of survival benefit. Based on model output, mean OS extrapolated over a life-time horizon was [redacted] years in the nivolumab with chemotherapy arm and 1.4 years in the chemotherapy arm (an improvement of [redacted] years). Based on this evidence, it can be concluded that end-of-life criteria are met.	B.2.6.3.1

B.3 Cost effectiveness

Summary of cost-effectiveness

- A *de novo* partitioned survival model was developed to assess the cost-effectiveness of nivolumab and chemotherapy compared to chemotherapy and pembrolizumab and chemotherapy for adults with untreated unresectable metastatic OSCC (consistent with the population in the Checkmate 648 trial).
- Use of NIVO+CHEMO will result in additional discounted QALYs and life years of [redacted] and [redacted], respectively, compared to CHEMO.
- Discounted incremental costs with NIVO-CHEMO were estimated to be [redacted] versus CHEMO under base case assumptions and the resultant ICER was £34,366 per QALY, which is considered to be cost-effective at a willingness-to-pay threshold of £50,000 per QALY.
- Extensive sensitivity analyses were undertaken, reflecting the assumptions required to undertake plausible, robust and transparent base case analysis.
- In the probabilistic sensitivity analysis, NIVO-CHEMO was cost-effective in 88.7% of scenarios at a willingness-to-pay threshold of £50,000 per QALY.
- In the deterministic sensitivity analysis, NIVO-CHEMO was cost-effective in all scenarios at a willingness-to-pay threshold of £50,000 per QALY.

B.3.1 Published cost-effectiveness studies

In line with the NICE Guide to the methods of technology appraisal 2013,[90] an SLR was conducted to identify cost-effectiveness studies for the treatment of advanced OSCC. In brief, electronic database searches (Medical Literature Analysis and Retrieval System Online [MEDLINE, via Ovid], Excerpta Medica dataBASE [Embase, via Ovid], Database of Abstracts of Reviews and Effects [DARE], Health Technology Assessment [HTA], and National Health Service Economic Evaluations Database [NHS EED]) were conducted on April 28, 2021. A total of 23 unique studies describing full economic evaluations of interventions aimed at managing previously untreated advanced or metastatic OSCC were included. Of these, nine studies were prioritised for extraction as they evaluated pharmacological interventions, whereas the remaining 14 studies evaluated non-pharmacological interventions, including esophagectomy, stents and brachytherapy, which were not deemed relevant to the objective of this SLR. Full details of the process and methods to identify and select the relevant cost-effectiveness evidence, including PRISMA diagrams, are provided in Appendix H.

B.3.2 Economic analysis

The economic case presented in this submission is based on conventional cost-utility analysis, assessing the use of NIVO-CHEMO versus relevant comparators for first-line treatment of unresectable advanced, recurrent or metastatic OSCC with tumour cell PD-L1 ≥1%.

B.3.2.1 Model structure

A partitioned survival model (PSM) approach has been utilised. It is acknowledged that modelling of subsequent treatment with immunotherapies (particularly nivolumab in the second-line setting) may indicate that use of a Markov approach would be appropriate. However, this aligns with previous UK HTAs for oesophageal cancer[1,72] and advice provided by ERGs for ongoing an HTA in gastro-oesophageal cancer.[87] Further, a PSM adheres to the NICE DSU guidelines[91] and is fully flexible, allowing extensive exploration of survival parameterisations and other input parameters. Moreover, a PSM may replicate survival outcomes with a higher degree of accuracy compared with a Markov model, although differences in outcomes should be minimal, particularly where appropriate transition rates have been derived.[92] Lastly, the structure of the PSM accommodates several treatment discontinuation options, which is of importance in the appraisal of nivolumab in combination with chemotherapy, where therapies may be continued beyond progression.

Aligned with the PSM approach, the economic model includes three mutually exclusive health states representing progression-free disease, post-progression and death, stratified by on-treatment versus discontinued (Figure 31). Further details regarding the modelling approach and inputs are detailed in Appendix M.

In a three-state PSM, health state occupancy is determined by survival curves, namely overall survival (OS) and progression-free survival (PFS). Figure 32 shows the health states of a PSM. The area under each curve shows the health state occupancy, with the area under the OS curve showing the proportion of patients alive at a given time and the area under the PFS curve showing the proportion of patients who are progression-free at a given time. The proportion of patients alive with progressed disease is the difference between OS and PFS curves.

These health states reflect disease severity and determine use of healthcare resources, health-related quality of life and mortality rates. The economic model structure has been chosen to reflect the most important treatment outcomes for OSCC patients: survival (progression free and overall), side effects, symptom control and quality of life. Survival curves have been applied to estimate PFS and OS in each treatment arm, while health state utilities

and costs have been applied to reflect the symptom control and quality of life experienced by patients receiving NIVO-CHEMO or comparators. Treatment-specific AE probabilities, alongside AE event-specific costs, are used to estimate the incidence and economic consequences associated with treatment-related AEs (Section B.3.3.6).

Figure 31. Three-state model structure

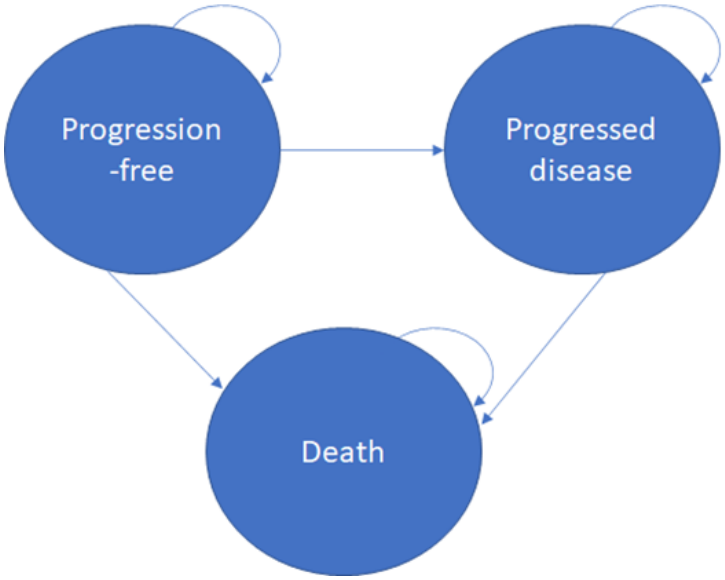
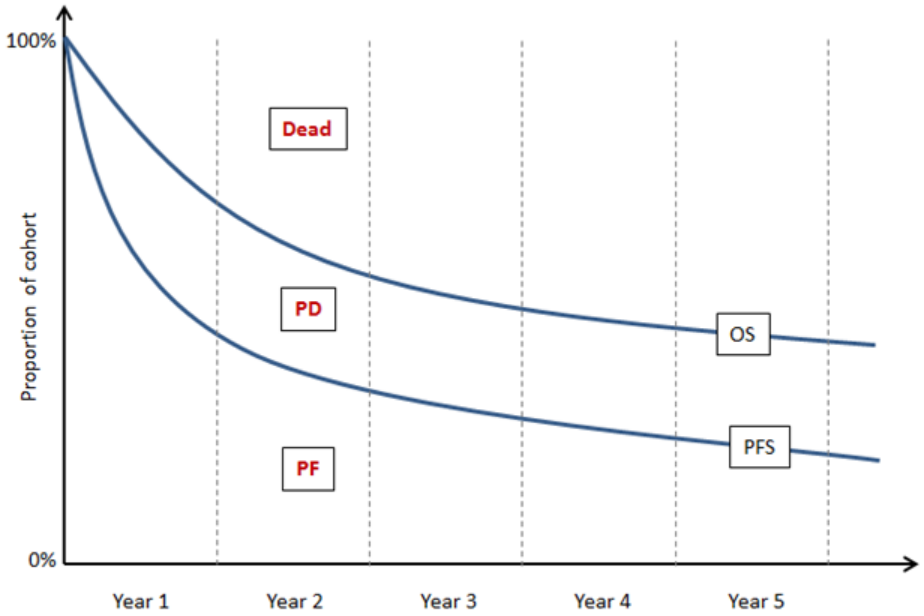


Figure 32. Overview of PSM method



OS: Overall survival; PD: Progressed disease; PF: Progression-free; PFS: Progression-free survival

Each first-line treatment has unique survival curves, OS and PFS, which determines the time spent in each health state (pre-progression, post-progression and death). This represents the treatment efficacy. Each first-line treatment also has a unique time on treatment (ToT) curve;

this determines how patients move through lines of treatment. In each health state, patients accrue treatment costs based on drug acquisition and administration, and health care resources use costs while in that health state, based on disease monitoring and management. Utilities are applied per health state, and a disutility is applied as a one-off utility decrement in the four cycles before death. Further details on clinical efficacy, costs and utility inputs can be found in Sections B.3.3, B.3.4 and B.3.5.

To reflect the nature of OSCC and available evidence, the model assumes that progression phases are consecutive, which means patients are not able to revert to pre-progression from more advanced phases of the disease. Although patients may be able to respond to therapy following progression, patients are still considered to have a higher hazard and an increased resource use. As evidence for this, patients enrolled in ATTRACTION-3 were still able to achieve a complete or partial response, but OS remained low.[29] Hence, this assumption can be considered appropriate.

Using a weekly cycle length, the model predicts the proportion of the population who experience a progression or death event. Weekly cycles were considered appropriate for this evaluation because it enables the model to reflect the timings of drug administrations associated with both NIVO-CHEMO and comparator therapies. Weekly cycles further capture a realistic minimum time during which the symptoms or responses can change in UK clinical practice.

A summary of the features of the PSM in presented in Table 31 and Table 32.

Table 31. Features of the economic analysis (per NICE template)

Feature	Model functionality	TA737[1] approach	Rationale
Time horizon	Lifetime horizon of 40 years	Lifetime horizon of 30 years	NICE reference case[93]
Cycle length	1 week (no half-cycle correction)	1 week with half-cycle correction	This is aligned to an economic model in a similar indication, a 1-week cycle length has been chosen as it is sufficient to capture treatment benefit and disease progression. No half-cycle correction is applied, in line with ERG comments during TA737.[86]
Source of utilities	Checkmate 648 EQ-5D-3L	KEYNOTE-590 EQ-5D-5L	NICE reference case[93]
Source of costs	Drug acquisition costs from BNF and eMIT (as appropriate). Drug administration costs aligned with TA737 ERG preferred approach. Disease management costs aligned with TA737 ERG preferred approach. Adverse event costs aligned with NHS reference costs and TA737 (where possible). Cost of end of life aligned with previously oncology HTAs.	Not applicable	NICE reference case[93]
Duration of treatment effect	No treatment waning effect applied.	ERG preferred a treatment waning scenario applied between 5 and 7 years. MSD did include a treatment waning scenario, based on the observed evidence. The committee concluded that all scenarios provided plausible estimates of overall survival and the preferred scenarios were not greatly different.[1]	There is now long-term evidence of a robust and durable treatment effect lasting beyond discontinuation for immunotherapies.[94]

Table 32. Additional features of the economic analysis

Feature	Model functionality	TA737[1] approach	Rationale
Type of economic evaluation	Cost-utility analysis using partitioned survival model	Cost-utility analysis using partitioned survival model	This approach aims to capture the impact on costs, life years and quality of life of introducing nivolumab as a first-line add on treatment
Setting and perspective on costs and outcomes	National Health Service and Personal Social Services (NHS and PSS)	NHS and PSS	NICE reference case[93]
Population	Patients with advanced unresectable, recurrent or metastatic previously untreated oesophageal cancer with PD-L1 ≥1%	Untreated, unresectable locally advanced or metastatic oesophageal cancer or HER-2 negative gastroesophageal junction adenocarcinoma (subgroup with PD-L1 CPS ≥10)	This is aligned to CheckMate 648 trial[95]
Intervention	Nivolumab in combination with chemotherapy (fluorouracil plus cisplatin)	Pembrolizumab plus chemotherapy (fluorouracil plus cisplatin)	This is aligned to CheckMate 648 trial[95]
Comparator	Chemotherapy (fluorouracil plus cisplatin) for all patients or pembrolizumab with chemotherapy (fluorouracil plus cisplatin) for patients with PD-L1 ≥10 (as a scenario)	Chemotherapy: fluorouracil plus cisplatin as base case analysis; scenarios assessed alternatives	This is aligned to CheckMate 648 trial[95] and current UK practice
Subsequent treatments	Intervention: Taxane monotherapy (docetaxel or paclitaxel) Comparator: Taxane monotherapy (docetaxel or paclitaxel) or nivolumab	KEYNOTE-590 subsequent treatments; updated to include nivolumab in the comparator arm.	This is aligned to current UK practice as confirmed during an advisory board held by BMS[2]. Additionally, this is aligned
Discounting	3.5% costs and health outcomes	3.5% costs and health outcomes	NICE reference case[96]

B.3.2.1.1 Derivation of health state occupancy estimates

Health state occupancy is defined by treatment specific PFS and OS extrapolations, derived from available data. It is assumed that these PFS and OS data implicitly include the effects of any subsequent treatment that may have been administered; hence, the benefits of subsequent treatment are captured.

B.3.2.1.2 Derivation of treatment line occupancy

Patients enter the model and can receive NIVO-CHEMO or a comparator treatment. Following treatment cessation, patients receive a subsequent line of therapy. As a simplifying assumption, it is assumed that patients may not discontinue this final line of therapy.

In the base case analysis, the proportion of patients on initial or subsequent treatment lines is based on the following criteria:

- Observed time on treatment data
- Treatment cessation (where treatment duration is specified, for example in set treatment durations or stopping rules)

B.3.2.1.3 Outcome measures

The primary model output is the incremental cost-effectiveness ratio (ICER) expressed as incremental costs per QALY gained. Additionally, the model provides an overview of other outcomes, such as LYs gained, and clinically relevant outcomes, such as predicted median OS and PFS.

B.3.2.2 Patient population

The economic evaluation considered the use of [redacted]

In the base case analysis, baseline patient parameters are derived from the baseline characteristics of patients enrolled in CheckMate 648, as detailed in Table 33.

Table 33. Baseline patient characteristics

Characteristic	Mean value (Standard error)	Source
Proportion male	[redacted]	CheckMate 648 PD-L1 ≥1% population (including NIVO-IPI arm)
Baseline age	[redacted]	

B.3.2.3 Intervention technology and comparators

As outlined in Section B.1.1, doublet chemotherapies are more commonly used for treatment of OSCC in the UK. Use of epirubicin-based triplet therapies for OSCC is declining in the UK, as there is limited evidence to support clinical benefit in the context of increased adverse events. This is confirmed by TA737, where clinical experts contacted by NICE explained that dual therapy regimens are preferred, while the Cancer Drugs Fund clinical lead confirmed that the use of triple regimens is rapidly diminishing.[1] In this context, the committee concluded that a dual chemotherapy regimen would be the appropriate comparator for TA737.[1] This aligns with expert advice provided to BMS.[2] Hence, assessment of epirubicin-based triplet therapy is not considered within the economic model.

Clinical advisors to BMS have also confirmed that doublet chemotherapies have similar outcomes in cases of advanced OSCC.[2] This is confirmed by clinical advice obtained during TA737, where clinicians advised the ERG that doublet regimens are of exchangeable effectiveness (i.e. exhibit a class effect).[86] However, clinical advice to the ERG also noted that regimens with fluorouracil are rarely given due to the lengthy infusion time, with use only where patients are unable to swallow capecitabine tablets.[86] Further, clinical experts stated during TA737 that oxaliplatin is more commonly used than cisplatin, as it is better tolerated and has a shorter infusion time.[1] In conclusion, clinical experts felt that capecitabine plus oxaliplatin (XELOX) would be the primary comparator in this patient population.[1] This aligns with clinical advice during an ongoing gastro-oesophageal adenocarcinoma NICE appraisal.[87] Further, clinical advisors to BMS have confirmed that XELOX should be considered the primary comparator.[2]

As CheckMate 648 provides direct comparative evidence for NIVO-CHEMO versus CHEMO, this evidence is considered as the base case analysis, and the economic model uses the costs of cisplatin plus fluorouracil is the chemotherapy of interest. However, scenario analyses are presented to address additional doublet chemotherapy regimens, which are assumed to have equal effectiveness but different cost and administration profiles (Q2W/Q3W vs Q4W). It is suggested the choice of therapy would not be impacted by addition of nivolumab (i.e., a patient who would have received XELOX would receive NIVO+XELOX as opposed to NIVO+FOLFOX).[13] Hence, scenario analysis compares nivolumab plus one doublet (e.g. XELOX) versus that same doublet (XELOX).

Additionally, the decision problem includes pembrolizumab plus chemotherapy. However, it should also be noted that pembrolizumab was only recently recommended by NICE (October 2021)[1] and is hence not yet considered as a standard of care. Further, pembrolizumab plus

chemotherapy is only used in a subgroup of the relevant patient population (patients with PD-L1 CPS ≥10) and hence is only relevant to part of the decision problem. As a result, a scenario analysis is presented versus pembrolizumab plus chemotherapy.

Table 34 provides an overview of the intervention and comparator applied in the base case analysis and scenario analyses.

Table 34. Definition of intervention and comparators

	Intervention	Comparators
Base case analysis	NIVO-CHEMO: nivolumab plus fluorouracil and cisplatin	CHEMO: fluorouracil and cisplatin
Scenario analysis	NIVO-CHEMO: nivolumab plus fluorouracil and oxaliplatin (FOLFOX)	CHEMO: fluorouracil and oxaliplatin (FOLFOX)
	NIVO-CHEMO: nivolumab plus capecitabine and oxaliplatin (XELOX)	CHEMO: capecitabine and oxaliplatin (XELOX)
	NIVO-CHEMO: nivolumab plus capecitabine and cisplatin	CHEMO: capecitabine and cisplatin
	NIVO-CHEMO: nivolumab plus fluorouracil and cisplatin	Pembrolizumab plus fluorouracil and cisplatin (in patients with PD-L1 CPS ≥10)

B.3.3 Clinical parameters and variables

B.3.3.1 Survival analysis approach

Clinical data to inform PFS and OS for NIVO-CHEMO and CHEMO are derived from CheckMate 648. However, follow-up was less than the maximum time horizon of the model. Therefore, extrapolation of survival data from the study was required to inform long-term outcomes, undertaken with reference to the guidance from the NICE Decision Support Unit (DSU) and Bagust an Beale (2014).[97,98] The model selection algorithm was used to select a suitable model (Figure 33).

A full description of the methods used is provided in Appendix N. In brief, several approaches were considered for the survival analysis. Progression events were based on BICR-assessed outcomes from CheckMate 648 and were defined as in this study. Death events from CheckMate 648 were used to inform OS modelling. Parametric survival functions were fitted to the extracted data using the R statistics environment, including exponential, Weibull, log-logistic, lognormal, Gompertz and generalised gamma survival distributions. Additionally, semi-parametric models were considered assessing the impact of different split points and subsequent parametric functions, in line with the approach taken in recent appraisals of immuno-oncology agents.[99,100]

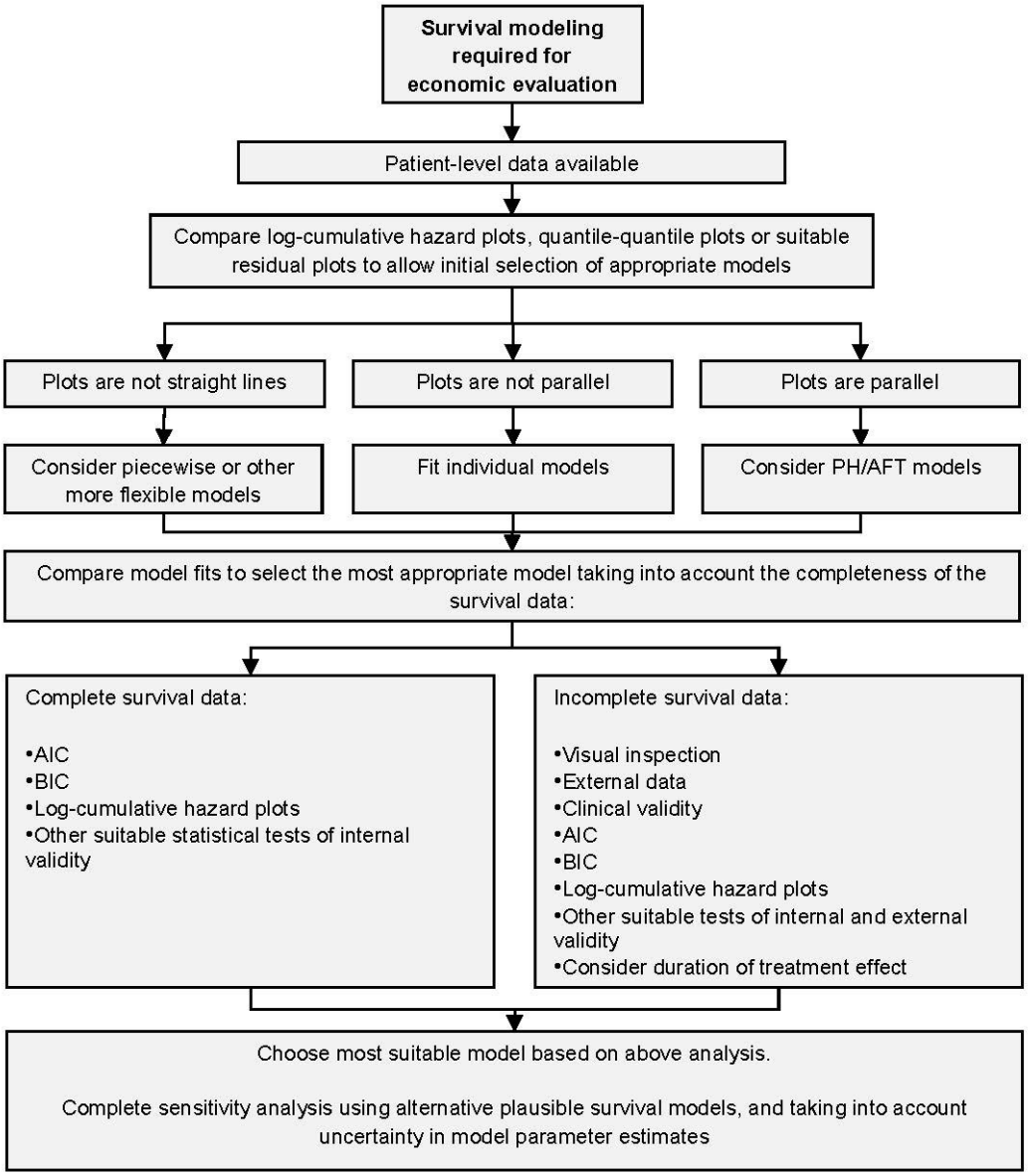


Figure 33. Survival model selection process algorithm

Source: NICE Decision Support Unit Document 14: Survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data.[98]
AFT: accelerated failure time; AIC: Akaike Information Criterion; BIC: Bayesian Information Criterion; PH: proportional hazards.

Goodness-of-fit was evaluated using the Akaike and Bayesian Information Criteria (AIC and BIC, respectively); minimisation of these measures is used to indicate goodness-of-fit whilst penalising overfitting, so that a smaller value demonstrates a more appropriate fit. In addition to assessment of goodness-of-fit statistics, the appropriateness of the parametric extrapolation was by visual inspection of the fit over the observed period and consideration of the log cumulative hazard plots.

It is worth noting that while the above methods for validating the extrapolation of progression and death events are appropriate, they are also necessarily constrained by derivation from observed data, which is, as previously indicated, limited by the uncertainty in the tail of the data. Therefore, the plausibility of the extrapolation was assessed through consideration of the long-term hazard profile and the extrapolated mean survival estimates. Additionally, clinical expert opinion was sought to ensure that the survival extrapolation approach can be considered appropriate.

B.3.3.1.1. Overall survival

As discussed in Appendix N, the proportional hazard assumption is violated for OS due to non-parallelism (See Figure 34). As a result, independent models were considered.

Figure 34. CheckMate 648 overall survival patients with tumour cell PD-L1 ≥1% from October 2021 DBL – Complementary log-log plot: CHEMO vs NIVO+CHEMO.

OS demonstrates a clear difference in hazard profile between the arms. During the first 6 months, the NIVO + CHEMO and the CHEMO KM curves are very similar. After 6 months, patients in the CHEMO arm have a higher hazard than the NIVO+CHEMO arm. From month 10, NIVO+CHEMO presents greater OS than CHEMO up to the end of the observed period. As a result, the NIVO+CHEMO (Figure 35) hazard rate peaks in the first 2–3 months, whereas in the CHEMO (Figure 36) arm, the hazard continually increases until reaching a peak at approximately 8–9 months.

Additional follow up has confirmed trends observed in the early part of the trial. However, this has not addressed the data challenges. As shown in Figure 36 and Appendix N, the hazard observed after 24 months in the CHEMO arm is decreasing. While this is plausible based on the March 2020 data (Appendix N), the CHEMO arm hazard is approaching general population mortality in the [redacted] (Figure 36), which is implausible.

Parametric models were explored but did not adequately reflect this change in observed hazard (as outlined in Appendix N). In particular, the parametric models were unable to reflect the CHEMO arm observed data after 20 months. Hence, a semi-parametric approach was considered appropriate as it reflected the high initial hazard but applied the maximum amount of data to inform the long-term extrapolation.

Applying Kaplan-Meier data until 6.9 months followed by parametric extrapolation enabled the initial hazard to be modelled appropriately and captured the high rate of events between study entry and six months. Further, there is significant overlap between the NIVO-CHEMO and

CHEMO arms in the first six months, with divergence after this point. Switching to parametric extrapolation from 6.9 months uses the maximum number of events to inform long-term extrapolation and describe the lower long-term hazard.

Several models were inappropriate for use for the CHEMO arm (Figure 33), including exponential and Weibull, which produced a poor fit to the observed data. Based on goodness of fit statistics, the best fit could be considered to be Gompertz, log-logistic, lognormal or generalised gamma. However, the survival extrapolations described for Gompertz and log-logistic could not considered plausible. Outcomes using the lognormal approach may also be considered optimistic, particularly in the context of the observed restricted mean OS ([redacted] months) but align with the observed hazard profile. It should be noted that this finding is replicated across the semi-parametric cut-points: improving the fit to the observed CHEMO arm data provides less clinically plausible outcomes. As such, the lognormal extrapolation may provide a balance between optimal fit to the observed data and the plausibility of the long-term predicted survival outcomes.

Several extrapolations are plausible in the NIVO-CHEMO arm. However, it should be noted that several predict implausibly short mean survival outcomes. Based on the observed data, restricted mean OS is [redacted] months in the NIVO-CHEMO arm, with [redacted] of patients without a death event at end of follow up (based on [redacted] events observed in 158 patients during the trial period). As a result, the mean OS predicted by the generalised gamma, exponential and Weibull functions are implausibly short. As the long-term follow-up is likely to reflect the responder population (See section 4.2.4 in Appendix N), as observed in the CHEMO arm, the lognormal function is likely to provide the optimal choice for the economic model. Further, this reflects the hazard profile observed for immunotherapies in general.

Overall, despite an initially higher hazard for NIVO+CHEMO, the magnitude of hazard is much greater for CHEMO overall. None of the hazard functions are monotonic, CHEMO is unimodal in shape whereas NIVO+CHEMO has a changing hazard.

Figure 35. OS, Smoothed hazard function estimates: NIVO+CHEMO

R-P: Royston-Parmar. Confidence interval is shown around b-spline estimator.

Figure 36. OS, Smoothed hazard function estimates: CHEMO

R-P: Royston-Parmar. Confidence interval is shown around b-spline estimator.

Figure 37. CheckMate 648 in patients with tumour cell PD-L1 ≥1% from [redacted], NIVO+CHEMO: Semi-parametric OS models overlaid upon Kaplan-Meier – 6.9 months cut point

Figure 38. CheckMate 648 in patients with tumour cell PD-L1 ≥1% from [redacted], CHEMO: Semi-parametric OS models overlaid upon Kaplan-Meier – 6.9 months cut point

B.3.3.1.2 BICR-assessed PFS

Progression events were based on BICR-assessed outcomes from CheckMate 648 and were defined as in this study (See section 3.2 of Appendix N for more detail).

The PFS (BICR) Kaplan-Meier has a highly stepped appearance, caused by regular tumour assessment times (every 6 weeks). While progression events can occur at any time, progression is actively monitored at these timepoints, causing an increase in progression events identified during these periods.

As in the OS curve, CHEMO and NIVO+CHEMO Kaplan-Meier data is aligned at the start of the trial data, with diverge after approximately 2–3 months. After this point KMs diverge with NIVO+CHEMO KM lying clearly above CHEMO. Overall, the NIVO+CHEMO arm presents better PFS (BICR) across all of the observed period. Reflecting the Kaplan-Meier data, the CHEMO and NIVO-CHEMO BICR-assessed PFS demonstrates an increase in hazard during the first 2-3 months (Royston-Palmer spline). Overall, the hazard for CHEMO is higher, it increases steadily and has a sustained higher hazard after initial peak.

Figure 39. BICR-assessed PFS: Smoothed hazard function estimates: NIVO+CHEMO arm

Figure 40. BICR-assessed PFS: Smoothed hazard function estimates: CHEMO arm

As with OS, parametric models were explored but did not adequately reflect this change in observed hazard (as outlined in Appendix N). In particular, the parametric models were unable to reflect the CHEMO arm observed data after 10 months and the NIVO-CHEMO arm data after 20 months. Hence, a semi-parametric approach was considered appropriate as it reflected the high initial hazard but applied the maximum amount of data to inform the long-term extrapolation.

It was determined that the 6.9 month cut point was appropriate for PFS data. Applying Kaplan-Meier data until 6.9 months followed by parametric extrapolation enabled the initial hazard to be modelled appropriately and captured the high rate of events between study entry and six months. Further, there is significant overlap between the NIVO-CHEMO and CHEMO arms in the first three months, with divergence after this point. Switching to parametric extrapolation from 6.9 months uses the maximum number of events to inform long-term extrapolation and describe the lower long-term hazard.

For the NIVO+CHEMO arm (Figure 41), the exponential and Weibull provided poor visual fit to the observed data. Additionally, Gompertz and log-logistic predicted mean PFS outcomes that could not be considered plausible as they do not converge or do not converge in the long-term suggesting long-term survival, which is clinically implausible. As such, generalised gamma was considered to provide the best fit, in line with goodness of fit statistics.

Several models predicted implausibly long mean PFS for the CHEMO arm (Figure 42), including Gompertz, log-logistic and lognormal. After exclusion of these models, based on goodness of fit statistics, the best fit could be considered to be Weibull.

Figure 41. CheckMate 648 PD-L1 ≥1% [redacted], NIVO+CHEMO: Semi-parametric OS models overlaid upon Kaplan-Meier – 6.9 months cut point

Figure 42. CheckMate 648 PD-L1 ≥1% [redacted], CHEMO: Semi-parametric PFS (BICR) models overlaid upon Kaplan-Meier – 6.9 months cut point

B.3.3.1.3 Clinical rationale and validation of survival extrapolation

Clinicians were consulted regarding their opinion upon the long-term overall survival and progression-free survival of patients in the NIVO+CHEMO treatment arm. The advisory board

report[2] states that “when considering extrapolation approaches to model long-term survival, it was considered important that these are based on clinical plausibility.”

The key concern was that there is a lot of uncertainty in the tails at the end of the survival curves, due to a small number of patients surviving to the end of the trial. As a result, several scenario analyses have been undertaken, assessing different survival modelling approaches (Section B.3.8.3.1).

There are no other studies with which to validate the results for extrapolation of the NIVO+CHEMO arm other than the informing trial, CheckMate 648. While KEYNOTE-590 enrolled a similar patient population, the overlap was not complete and follow up time was less than CheckMate 648, so that conclusions could not be drawn. For this reason, the extrapolated curves and approaches were compared to the observed CheckMate 648 data visually and statistically (using AIC and BIC goodness of fits statistics) as much as possible. This method informed selection of the most appropriate modelling approach and fit as a form of validation.

Table 35. Survival extrapolations applied in the economic model

	NIVO-CHEMO	CHEMO
Overall survival	Semi-parametric (6.9 month cut point); Lognormal	Semi-parametric (6.9 month cut point); Lognormal
Progression-free survival	Semi-parametric (6.9 month cut point); Generalised gamma	Semi-parametric (6.9 month cut point); Weibull

B.3.3.1.4 All-cause mortality

Individuals randomised into clinical trials are likely to be slightly younger and healthier than the overall oesophageal cancer patient population in the UK. The mean age of patients in CheckMate 648 is [redacted] years, increasing the likelihood that most deaths observed over the trial period were cancer-related.

Therefore, the model includes age and gender-adjusted mortality based on information from UK life tables. These values (based on UK lifetables)[101] are included in every cycle in addition to the disease-related mortality values and are applied multiplicatively. As some deaths of the individuals randomised into a clinical trial are likely to be non-cancer related, some form of double-counting will occur. However, as the effect applies equally to all comparators, it is likely to have a negligible impact on predicted survival (and hence cost-effectiveness).

B.3.3.2 Treatment discontinuation

The economic model applies treatment discontinuation using time on treatment (ToT) data during CheckMate 648, defined as time from randomisation to last dose of treatment. The timing of discontinuations was assumed to impact on treatment costs and resource use.

B.3.3.2.1 Nivolumab plus chemotherapy

Patient-level data from CheckMate 648 were obtained describing discontinuation due to progression, study drug toxicity, AEs unrelated to study therapy and withdrawal of patient consent. Kaplan-Meier estimates of ToT were complete at the end of the trial follow-up period, in that the number of patients at risk of discontinuation at the end of follow-up was 0. As such the Kaplan-Meier curves themselves were used in the model to estimate ToT, ensuring complete consistency with the clinical trial data.

Kaplan-Meier data for ToT for NIVO-CHEMO are summarised in Figure 43.

Figure 43. Time on treatment: CheckMate 648 Kaplan-Meier – NIVO-CHEMO (PD-L1 ≥1% subgroup)

B.3.3.2.2 Chemotherapy

CheckMate 648 is a randomised controlled phase 3 study that includes cisplatin plus fluorouracil as a chemotherapy arm. As described for NIVO-CHEMO, patient-level data from CheckMate 648 were obtained describing discontinuation due to progression, study drug toxicity, AEs unrelated to study therapy and withdrawal of patient consent. Kaplan-Meier estimates of ToT were complete at the end of the trial follow-up period, in that the number of patients at risk of discontinuation at the end of follow-up was 0. As such the Kaplan-Meier curves themselves were used in the model to estimate ToT, ensuring complete consistency with the clinical trial data. Kaplan-Meier data for ToT for CHEMO are summarised in Figure 44.

The ToT for additional comparators in scenario analysis was assumed to be the same as the chemotherapy arm of the CheckMate 648 study.

Figure 44. Time on treatment: CheckMate 648 Kaplan-Meier – CHEMO (PD-L1 ≥1% subgroup)

B.3.3.2.3 Subsequent therapies

Second-line palliative chemotherapy is recommended for patients who have progressed on the first-line therapy; however, there is uncertainty around composition of therapy. Specific chemotherapy regimens are not defined in the NICE clinical guidelines in the second line setting.[102-104] Similar to UK guidance, guidelines from the European Society for Medical Oncology (ESMO) recommend palliative chemotherapy in the management of advanced or metastatic OC.[18] Second-line chemotherapy using taxane monotherapy (docetaxel, paclitaxel) is recommended for patients with OSCC.[18] However, patients in the UK may also receive second-line nivolumab treatment following previous chemotherapy.[72]

In the economic model, patients receive a subsequent therapy following discontinuation, as outlined in Table 36. As a simplifying assumption, it is assumed that all patients receiving an immunotherapy (i.e. NIVO-CHEMO or PEMBRO+CHEMO) in the first line setting receive single agent taxane as subsequent therapy; this is aligned with clinical expert opinion.[2] Clinical advisors to BMS advise that patients would not receive subsequent PD-L1 inhibitors following previous PD-L1 inhibitor use.[2] Second line standard of care is defined as equal proportions of patients receiving paclitaxel and docetaxel. This aligns to a previously published study of UK clinical practice, which identified that more than half (54%) of patients receiving second-line therapy receive single agent treatment and the most common second-line treatment is paclitaxel (35% of use).[80]

Patients in the CHEMO arm receive nivolumab monotherapy, in line with budget impact modelling assumptions during TA707, where nivolumab monotherapy displaced the majority of taxane use.[72]

All patients discontinue treatment during CheckMate 648, most commonly due to disease progression or study drug toxicity. However, not all patients received subsequent treatment. This may be related to patient comorbidities or fitness. Additionally, CheckMate 648 applied a stopping rule at 24 months and patients with a complete response may not receive subsequent treatment. To reflect this outcome, only a proportion of patients receive subsequent treatment costs, aligned with CheckMate 648 subsequent treatment usage.

Table 36. Subsequent therapy applied in model

Treatment arm	Subsequent treatment	Proportion of patients
NIVO-CHEMO	Single agent taxane; assumed equal use of docetaxel and paclitaxel	49.4%
CHEMO	Nivolumab monotherapy	56.7%
PEMBRO+CHEMO	Single agent taxane; assumed equal use of docetaxel and paclitaxel	Aligned with NIVO-CHEMO
CHEMO: chemotherapy; NIVO: nivolumab; PEMBRO: pembrolizumab		

B.3.3.2.3.1 Impact of subsequent therapies in CheckMate 649

Among CheckMate 648 patients with PD-L1 ≥1%, subsequent cancer therapy was received by a lower proportion (53%) of patients in the NIVO-CHEMO treatment arm compared to the CHEMO arm (66%).[63] Further, fewer patients received subsequent systemic therapy in the NIVO-CHEMO treatment arm (49%) than in the CHEMO arm (57%). Additionally, as shown in Table 37, more patients in the CHEMO arm received a PD-1 or PD-L1 inhibitor (15% versus 6%) and this was most commonly nivolumab (10% versus 5%) or camrelizumab (10% versus 5%).[63]

Taxane usage, and particularly paclitaxel use was relatively high during CheckMate 648 (paclitaxel: 26% in the NIVO-CHEMO arm and 24% in the CHEMO arm), reflecting around 70% of subsequent systemic therapy use, which is aligned to UK clinical practice. Similarly, use of nivolumab after previous chemotherapy is aligned with UK standard of care. However, use of camrelizumab does not reflect the UK patient pathway.

As outlined in Appendix N, subsequent treatment is highly influential in long-term outcomes, particularly in the CHEMO arm.

Table 37. CheckMate 648 subsequent treatment (PD-L1 ≥1%)[63]

	NIVO-CHEMO (N=158)	CHEMO (N=157)
Any subsequent treatment	**83 (53)**	**104 (66)**
Subsequent radiotherapy	**35 (22)**	**52 (33)**
Curative	4 (3)	3 (2)
Palliative	32 (20)	49 (31)
Subsequent surgery	**4 (3)**	**2 (1)**
Curative	1 (<1)	0
Palliative	3 (2)	1 (<1)
Other	0	1 (<1)

Subsequent systemic therapy	78 (49)	89 (57)
Anti-PD-1/PD-L1	9 (6)	23 (15)
Nivolumab	8 (5)	16 (10)
Pembrolizumab	0	2 (1)
Camrelizumab	8 (5)	16 (10)
Other systemic therapies	77 (49)	82 (52)
Fluorouracil	22 (14)	33 (21)
Cisplatin	20 (13)	22 (14)
Paclitaxel	41 (26)	37 (24)
Docetaxel	23 (15)	20 (13)
Oxaliplatin	6 (4)	4 (3)
Carboplatin	7 (4)	6 (4)
Nedaplatin	13 (8)	9 (6)
Gimeracil; oteracil potassium; tegafur	8 (5)	5 (3)
Irinotecan	1 (<1)	5 (3)

B.3.3.3 Adverse events

Treatment-related AEs are an inevitable consequence of any intervention, and these events are applied in the economic model, affecting the costs and disutilities accrued by patients on each intervention.

AEs were selected on the basis of relevance to NIVO-CHEMO treatment. The ten most frequently occurring treatment-related grade 3–4 serious AEs were included in the economic model. Each treatment has a unique AE profile, with each AE requiring an AE-specific cost of management in the cycle in which the AE occurs. Each AE also has an AE specific utility decrement, applied additively to the health state utility values in the cycle in which the AE occurs.

These AEs were applied in the model as a one-off cost in the first cycle only. This is in line with TA737[1], where AEs are only modelled upon treatment initiation. Therefore, the proportion of the cohort demonstrated in Table 38 receives the costs and utility decrements associated with that AE.

Table 38. CheckMate 648 grade 3–4 treatment-related serious adverse events rates

Adverse event	Nivo -CHEMO (n = 310)			CHEMO (n = 304)			PEMBRO+CHEMO (n = 370)		
	n	%	SE	n	%	SE	n	%	SE
Source	CheckMate 648[51]						KEYNOTE-590[105]		
Total patients with an event	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	47	12.70%	1.92%
Vomiting	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	9	2.43%	0.80%
Hyponatraemia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	0	0.00%	0.00%
Pneumonitis	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	12	3.24%	0.92%
Hepatic function abnormal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	0	0.00%	0.00%
Adrenal insufficiency	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	0	0.00%	0.00%
Acute kidney injury	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	11	2.97%	0.88%
Colitis	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	0	0.00%	0.00%
Nausea	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	0	0.00%	0.00%
Dehydration	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	6	1.62%	0.66%
Febrile neutropenia	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	9	2.43%	0.80%

B.3.4 Measurement and valuation of health effects

B.3.4.1 Health-related quality of life studies

In line with the NICE guidelines to the methods of technology appraisal 2013,[90] studies describing health-related quality-of-life for patients with OSCC were identified systematically. Relevant studies were identified by searching the following databases: Medical Literature Analysis and Retrieval System Online (MEDLINE, via Ovid), Excerpta Medica dataBASE (Embase, via Ovid), Cochrane Central Register of Controlled Trials (via Cochrane Library), and the University of Sheffield School of Health and Related Research Health Utilities Database (ScHARRHUD). The database searches were executed on April 27, 2021 and identified 5,439 abstracts. Of the 39 publications moving to full-text screening, 10 were eligible for inclusion in the SLR. Following the addition of one record from a hand search, 11 unique publications representing seven studies were included in the SLR. The methods and results of the SLR are fully described in Appendix I.

B.3.4.2 CheckMate 648 health-related quality of life data

CheckMate 648 included assessment of health-related quality of life during the study, which can be used to derive utilities for modelling analysis. Assessments of EQ-5D-3L status in CheckMate 648 were carried out every 6 weeks during the treatment phase and every 12 weeks in the follow-up phase.

In the NIVO-CHEMO arm, 306 of the 321 patients (95.3%) provided a baseline questionnaire and hence were able to inform outcomes. Similarly, in the chemotherapy arm, 298 of the 324 patients (92.0%) provided a baseline questionnaire. Completed questionnaires were sourced from the [redacted] DBL for the overall population of CheckMate 648; as quality of life is not anticipated to vary by PD-L1 status, use of the overall population increased the data informing this analysis. Among all randomised patients, at baseline, mean (SD) EQ- 5D-3L Utility Index scores for the NIVO-CHEMO arm ([redacted]) were similar to those in the CHEMO arm ([redacted]). The mean change from baseline increased in the NIVO-CHEMO arm and the CHEMO arm (Figure 14). Improvements in mean EQ-5D-3L utility index scores were sustained longer and surpassed the MID threshold more often in the NIVO-CHEMO arm vs the CHEMO arm.

Two health state models were assessed: progression-based health state model and a time-to-death health state model. To estimate the mean values of EQ-5D-3L for each health state, a mixed model approach was used to account for repeated EQ-5D-3L measurements per patient within a health state (mixed model for repeated measures [MMRM]). For each health

state model, two statistical models were fit: one with and one without treatment. The variable(s) defining health states, treatment, and their interaction, if any, were included in the model as fixed effects. The model with treatment included interactions of treatment by health state variable in the model. A random intercept was used to account for repeated measurements within each patient. An unstructured covariance structure was used.

The Akaike information criterion (AIC) and Bayesian information criterion (BIC) based on the maximum likelihood approach were used to examine the extent of improvement in model fit after including treatment, where lower AIC and BIC values indicate better fit. The -2*log-likelihood (2*logL) statistics were also presented, as well as results from chi-square tests of the statistical significance between nested models with and without treatment.

The utility estimates using the United Kingdom (UK)-weight index are initially presented,[106] which are based on the time trade-off valuation technique methodology (Dolan et al, 1997[106]).

Outcomes from CheckMate 648 are presented in Table 39 and Table 40.

Table 39. CheckMate 648 EQ-5D-3L utility index pre- and post-progression: Number of Patients, Observations, and least square mean estimates (all randomised patients; [redacted]DBL)

	Health State	Overall	Nivolumab + Chemotherapy	Chemotherapy
All randomised patients				
Patient numbers/ observation numbers	Pre-progression	[redacted]	[redacted]	[redacted]
	Post-progression	[redacted]	[redacted]	[redacted]
Least squares means (SE) (95% CI)	Pre-progression	[redacted]	[redacted]	[redacted]
	Post-progression	[redacted]	[redacted]	[redacted]
PD-L1 ≥1%				
Patient numbers/ observation numbers	Pre-progression	[redacted]	[redacted]	[redacted]
	Post-progression	[redacted]	[redacted]	[redacted]
Least squares means (SE) (95% CI)	Pre-progression	[redacted]	[redacted]	[redacted]
	Post-progression	[redacted]	[redacted]	[redacted]

Table 40. CheckMate 648 EQ-5D-3L utility index Time-to-death LS Mean (95% CI) – all randomised patients, [redacted]DBL

Time-to-death Category	Overall	Nivolumab + Chemotherapy	Chemotherapy
Overall	[redacted]	[redacted]	[redacted]
>180 Days	[redacted]	[redacted]	[redacted]
91–180 days	[redacted]	[redacted]	[redacted]
31–90 days	[redacted]	[redacted]	[redacted]
0–30 days	[redacted]	[redacted]	[redacted]

B.3.4.2.1 Mapping

EQ-5D data were collected in CheckMate 648 in line with the NICE reference case. Utility values for health states and AEs for which CheckMate 648 data could not be used were obtained from the literature. Therefore, there was no need to use mapping techniques.

B.3.4.2.2 Economic model health state utility values

Utility values determined by health state occupancy, stratified by progression status, are employed in the model in line with the preferred approach during TA737.[1,86] The health state utility values from the BMS utility analysis (Appendix O) are displayed in Table 41.

Table 41. Health state utility values

Health state	Mean value (SE)	Source
Pre-Progression	[redacted]	CheckMate 648[95]
Post-Progression	[redacted]	

B.3.4.2.3 End of life utility decrement

End of life utility decrement represents the deterioration of the condition, and thus the reduction in quality of life, in the time prior to death for a patient with OC. The value (SE) used for this decrement is [redacted] This value was derived from values sourced from the BMS utility analysis[107], specifically, the difference between the overall utility of a patient on treatment in the post-progression disease state ([redacted]) and the overall utility of the patient on treatment in the 30 days before death ([redacted]). The post-progression health state utility value was chosen to derive this decrement as it provides the most conservative estimate and because patients mostly die in the post-progression state. As the estimate of the utility prior to death was in the 30 days before death, the estimated utility decrement is applied in the four cycles (four weeks) before death.

B.3.4.3 Adverse reactions

AEs were selected on the basis of relevance to NIVO-CHEMO treatment. The ten most frequently occurring treatment-related grade 3–4 serious AEs were included in the economic model. AEs have a negative impact on quality of life each time a patient experiences an AE and, therefore, results in a reduction in total utility. The utility decrements associated with each AE are applied additively to the health state utility in the cycle in which the AE occurs. Each adverse events' utility decrement is displayed in Table 42.

Table 42. Adverse event utility decrements applied in economic model

Adverse event	Utility decrement	SE	Source
Vomiting	0.048	0.016	Nafees 2008[108]
Hyponatraemia	0.000	0.000	TA484[109]
Pneumonitis	0.037	0.004	TA578[110]
Hepatic function abnormal	0.119	0.012	Assumption
Adrenal insufficiency	0.119	0.012	Assumption
Acute kidney injury	0.048	0.016	Assumption
Colitis	0.047	0.005	Assumption
Nausea	0.048	0.016	Nafees 2008[108]
Dehydration	0.119	0.012	Assumption
Febrile neutropenia	0.090	0.016	Nafees 2008[108]

B.3.4.4 Health-related quality-of-life data used in the cost-effectiveness analysis

Table 43 provides an overview of the utility values applied in the cost-effectiveness analysis.

Table 43. Summary of utility values for cost-effectiveness analysis

State	Utility value: mean (standard error)	Reference in submission	Justification
Health state utilities			
Pre-Progression	[redacted]	Table 41	Progression-stratified utility values sourced from CheckMate 648
Post-Progression	[redacted]	Table 41	
Utility decrements			
End of life	[redacted]	B.3.4.2.3	Reflects the impact of end of life at a patient level, based on CheckMate 648 data
Vomiting	0.048 (0.016)	Table 42	Nafees 2008[108]
Hyponatraemia	0.000	Table 42	TA484[109]
Pneumonitis	0.037 (0.004)	Table 42	TA578[110]
Hepatic function abnormal	0.119 (0.012)	Table 42	Assumption
Adrenal insufficiency	0.119 (0.012)	Table 42	Assumption
Acute kidney injury	0.048 (0.016)	Table 42	Assumption
Colitis	0.047 (0.005)	Table 42	Assumption
Nausea	0.048 (0.016)	Table 42	Nafees 2008[108]
Dehydration	0.119 (0.012)	Table 42	Assumption
Febrile neutropenia	0.090 (0.016)	Table 42	Nafees 2008[108]

B.3.5 Cost and healthcare resource use identification

In line with the NICE guidelines to the methods of technology appraisal 2013,[90] studies describing costs and healthcare resource use for patients with OSCC were identified systematically, during the cost-effectiveness SLR. Relevant studies were identified by searching the following databases: Medical Literature Analysis and Retrieval System Online (MEDLINE, via Ovid), Excerpta Medica dataBASE (Embase, via Ovid), Cochrane Central Register of Controlled Trials (via Cochrane Library), and the University of Sheffield School of Health and Related Research Health Utilities Database (ScHARRHUD). The searches were executed on April 27, 2021, and are fully described in Appendix G.

Costs have been categorised as relating to the intervention/comparator, subsequent therapies, monitoring and management of the disease, management of AEs, and terminal care. Costs have been sourced from the relevant UK literature and NHS reference costs.[111-114] Where values for standard errors are not available, a default value of 20% of the mean has been used.

B.3.5.1 Intervention and comparator cost and resource use

B.3.5.1.1 Cost of initial treatment-related costs

The costs of each therapy are applied each cycle where treatment is continued and include drug procurement and administration costs. Treatment modifiers were applied to the acquisition and administration costs, accounting for missed or delayed doses during CheckMate 648 (described in Section B.3.5.1.1.1). Costs for initial treatment are aligned with the time on treatment curves described in Section B.3.3.2.

Costs of the interventions and comparator comprise the unit costs of the treatment, costs according to the dose and frequency administered to patients and the administration of treatment. An overview of drug acquisition costs and administration costs is provided in Table 44 and Table 45, respectively. A breakdown of the costs for the intervention, nivolumab in combination with chemotherapy, is displayed in Table 46. A breakdown of the costs for the comparators, chemotherapy (fluorouracil plus cisplatin) and pembrolizumab in combination with chemotherapy (fluorouracil plus cisplatin), is displayed in Table 47 and Table 48, respectively.

Additionally, patients in the nivolumab treatment arm are assumed to receive a cost for PD-L1 testing, which is £42.61, in line with the cost applied during TA737.[1]

Table 44. Administration costs

Details	Mean value	Source
Oral tablets	£0.00	-
Deliver Simple Parenteral Chemotherapy at First Attendance	£284.05	NHS reference costs: weighted average of SB12Z[115]
Deliver Complex Chemotherapy, including Prolonged Infusion Treatment, at First Attendance	£431.72	NHS reference costs: SB14Z day case and reg day/night[115]

Table 45. Drug acquisition costs

Drug	Formulation	Acquisition cost	Source
Capecitabine	150mg tablets pack size 60	£4.43	eMIT database[116]
	300mg tablets pack size 60	£7.77	
	500mg tablets pack size 120	£26.30	
Cisplatin	100mg/100ml solution for infusion vials	£8.73	eMIT database[116]
	50mg/50ml solution for infusion vials	£5.38	
Fluorouracil	1g/20ml (5%) solution for infusion vial	£2.35	eMIT database[116]
	2.5g/100ml (2.5%) solution for infusion vial	£3.79	
	2.5g/50ml (5%) solution for infusion vial	£4.01	
	500mg/10ml (5%) solution for infusion vial	£1.77	
	5g/100ml (5%) solution for infusion vials	£8.58	
Nivolumab*	240mg/24ml concentrate for solution for infusion vial	£2,633	BNF[117]
Pembrolizumab*	100mg/4ml concentrate for solution for infusion vial	£2,630	BNF[117]
* Patient access schemes available			

Table 46. Drug acquisition and administration unit costs for nivolumab in combination with chemotherapy (fluorouracil plus cisplatin)

	Nivolumab	Fluorouracil	Cisplatin	Source
Dosing regimen	240 mg, on day 1 every 2 weeks	800 mg/m^2, on day 1 through day 5 every 4 weeks	80 mg/m^2, on day 1 every 4 weeks	CheckMate 648 trial[95]
Dose received	240 mg	1331 mg (6656 mg over 5 days)	133 mg	Assuming body surface area of 1.66m^2, calculated using CheckMate 648 data[95]
Unit cost	£2,633.00 (PAS cost: [redacted])	£15.64	£14.11	Table 45
Admin method	Intravenous as a 30 minute infusion on day 1 and day 15 of each 28 day cycle	Intravenous continuous infusion on days 1–5 of 28 day cycle	Intravenous as a 30–120 minute infusion on day 1 of 28 day cycle	CheckMate 648 trial[95]
Day 1 administration cost	£431.72			Table 44
Day 15 administration cost	£284.05			Table 44
PD-L1 test cost	£42.61			TA737[1]
All therapies assume wastage.				

Table 47. Drug acquisition and administration unit costs for chemotherapy (fluorouracil plus cisplatin)

	Fluorouracil	Cisplatin	Source
Dosing regimen	800 mg/m2, on day 1 through day 5 every 4 weeks	80 mg/m2, on day 1 every 4 weeks	CheckMate 648 trial[95]
Dose received	1331 mg (6,656 mg over 5 days)	133 mg	Assuming body surface area of 1.66m^2, calculated using CheckMate 648 data[95]
Unit cost	£15.64	£14.11	Table 45
Admin method	Intravenous continuous infusion on days 1–5 of 28 day cycle	Intravenous as a 30–120 minute infusion on day 1 of 28 day cycle	CheckMate 648 trial[95]
Day 1 administration cost	£431.72		Table 44
All therapies assume wastage.			

Table 48. Drug acquisition and administration unit costs for pembrolizumab in combination with chemotherapy (fluorouracil plus cisplatin)

	Pembrolizumab	Fluorouracil	Cisplatin	Source
Dosing regimen	200 mg, on day 1 every 3 weeks	800 mg/m2, on day 1 through day 5 every 3 weeks	80 mg/m^2 ,on day 1 every 3 weeks	KEYNOTE-590[105]
Dose received	200 mg	1,331 mg (6,656 mg over 5 days)	133 mg	KEYNOTE-590[105]
Unit cost	£5,260.00	£15.64	£14.11	Table 45
Admin method	Intravenous as a 30 minute infusion on day 1 each 21 day cycle	Intravenous continuous infusion on days 1–5 of 21 day cycle	Intravenous as infusion on day 1 of 21 day cycle	KEYNOTE-590[105]
Day 1 administration cost	£431.72			Table 44
PD-L1 test cost	£42.61			TA737[1]
All therapies assume wastage.				

B.3.5.1.1.1 Dose intensity

In order to account for the observance that not all patients will follow the dosing regimen prescribed, leading to them missing or delaying doses, a treatment modifier is applied to the cost of each component of each intervention in the model. This reflects the proportion of doses delayed versus those administered during CheckMate 648. The treatment modifier of each intervention component is presented in Table 49.

Table 49. CheckMate 648 proportion of patients receiving a dose

Treatment		Treatment modifier	Source
Nivolumab in combination with chemotherapy (fluorouracil plus cisplatin)	Nivolumab	[redacted]	CheckMate 648 trial,[95]
	Fluorouracil	[redacted]	
	Cisplatin	[redacted]	
Chemotherapy (fluorouracil plus cisplatin)	Fluorouracil	[redacted]	
	Cisplatin	[redacted]	
Pembrolizumab in combination with chemotherapy (fluorouracil plus cisplatin)	Pembrolizumab	[redacted]	Assumed equivalent to NIVO-CHEMO due to lack of data
	Fluorouracil	[redacted]	
	Cisplatin	[redacted]	

B.3.5.1.2 Subsequent treatment

In clinical practice, OC patients who discontinue their first-line therapy are likely to receive a subsequent therapy, with the possible subsequent therapies determined by the treatment they received in the first-line. Reflecting this, the economic model assumes that patients discontinuing initial treatment receive a subsequent therapy. The composition of subsequent treatment and underpinning assumptions are described in Section B.3.3.2.3.

As a PSM is unable to track individual patients through lines of therapy, cyclical second-line average costs are calculated; these are displayed in Table 50. The frequency of each second-line treatment dependent on the first-line treatment and the resultant average weighted costs applied in the model are displayed in Table 51.

Table 50. Subsequent treatment costs

	Nivolumab	Taxane: docetaxel	Taxane: paclitaxel	Source
Dosing regimen	240 mg, on day 1 every 2 weeks	75 mg/m2, on day 1 every 2 weeks	100 mg/m2, on day 1 every week for 6 weeks, followed by a 2 week break	ATTRACTION-3[118]
Dose received	240 mg	125 mg	166 mg	ATTRACTION-3[118]
Unit cost	£2,633.00	£17.95	£14.44	Table 45
Admin method	Intravenous	Intravenous	Intravenous	ATTRACTION-3[118]
Admin cost	£284.05	£284.05	£284.05	Table 44
Average cyclical cost	£1,458.52	£129.50	£191.62	Derived
All therapies assume wastage.				

Table 51. Weighted average subsequent treatment costs

	Second-line treatment frequency			Second-line weighted average cyclical cost
	Nivolumab	Taxane: docetaxel	Taxane: paclitaxel	
NIVO-CHEMO*	0%	24.7%	24.7%	£79.26
CHEMO*	56.7%	0%	0%	£826.80
PEMBRO+CHEMO*	0%	24.7%	24.7%	£79.26
Assumptions information composition of subsequent treatment are presented in Table 36. * Please note that not all patients will receive subsequent treatment, e.g. due to comorbidities or insufficient fitness, and that patients that received I-O therapy as first-line treatment will not receive I-O therapy as second-line treatment.				

Duration of subsequent treatment

Second-line treatment is only given to patients for a finite time period. In order to prevent implausible accrual of second-line treatment costs, functionality is included in the model, which moves patients from second-line treatment to no treatment. This functionality uses the median ToT data for second-line treatments to derive a cyclical second-line treatment discontinuation rate for the available second-line treatments. The second line treatment discontinuation rates are weighted based on the frequency of use of treatment in the second line and combined to form an average second line cyclical discontinuation rate, both for the treatment and control arms.

Table 52. Second-line cyclical discontinuation

	Nivolumab	*Source*	Taxane: docetaxel or paclitaxel	*Source*
Median time on treatment (weeks)	12.00	TA707[119]	11.00	TA707[119]
Cyclical discontinuation rate	0.056	Derived	0.061	Derived

B.3.5.1.2 Health state-specific disease management costs

Monitoring and disease management costs vary by health state. These costs are associated with healthcare resource use. The frequency of resource use in each health state has been sourced through the literature using TA737.[1] The cost for each resource use is sourced from the NHS reference costs 2019-2020.[111] The calculations and total cyclical (1 week) health state costs, which are used as model inputs, are displayed in Table 53.

Table 53. Health state costs

Resource Use	Cost[111]	Weekly frequency pre-progression[1,111]	Weekly frequency post-progression[1,111]
CT scan	£103.31	0.08	0.08
Blood test	£2.53	0.33	1.00
Kidney	£33.80	0.33	1.00
Hepatic	£33.80	0.33	1.00
Consultant	£203.14	0.25	0.25
Total cost (SE)		£82.77 (£16.55)	£129.52 (£25.90)
CT: computed tomography; SE: standard error			

B.3.5.1.3 Terminal care costs

Terminal care costs represent the management, monitoring and resource use for patients with OC in the months prior to death and are applied to patients who enter the death state as a one-off cost. The terminal care cost used in the model, sourced from Georghiou et al. (2014),[120] is £9,171.92, with a SE of £1,834.38, adjusted to account for inflation.

This terminal care cost remains higher than that applied during TA737,[1] which used an earlier derived cost. However, the ERG noted several concerns with this cost,[1] so this externally sourced and published cost has been applied.

B.3.5.2 Adverse events

As outlined in Section B.3.3.6 and B.3.4.3, the economic model includes the most common grade 3–4 drug-related serious adverse events (AEs) [grade3-4] rates reported during CheckMate 648. Each treatment has a unique AE profile, with each AE requiring an AE

specific cost of management in the cycle in which the AE occurs. Each AE also has an AE specific utility decrement, applied additively to the health state utility values in the cycle in which the AE occurs.

As first-line treatment is the key feature of the model and AEs have a negligible impact on the modelled results, AEs are assumed to have a zero incidence in subsequent treatments. This is in line TA737[1], where there is no modelling of AEs in subsequent treatments.

The cost and utility decrement associated with each AE are summarised in Table 54, as well as the incidence for each first-line treatment.

Table 54. Adverse event costs

Adverse event (AE)	AE cost (SE)	Source
Vomiting	£471.95 (£94.39)	NHS reference costs 2019-2020 (FD10M NES)[111]
Hyponatraemia	£1,164.14 (£232.83)	NHS reference costs 2019-2020 (KC05H - NES)[111]
Pneumonitis	£1,909.33 (£381.87)	NHS reference costs 2019-2020 (weighted average DZ111K,L,M,N,P,Q,R,S,T,U,V – total HRGs)[111]
Hepatic function abnormal	£2,461.04 (£492.21)	NHS reference costs 2019-2020 (weighted average GC01C,D,E,F – total HRGs)[111]
Adrenal insufficiency	£2,079.75 (£415.95)	Chauhan 2013[121]
Acute kidney injury	£1,961.20 (£392.24)	NHS reference costs 2019-2020 (weighted average LA07H,J,K,L,M,N,P – total HRGs)[111]
Colitis	£2,426.57 (£485.31)	Copley-Merriman 2018[122]
Nausea	£471.95 (£94.39)	NHS reference costs 2019-2020 (FD10M -NES)[111]
Dehydration	£1,329.93 (£265.99)	NHS reference costs 2019-2020 (weighted average KC05G,H,J,K,L,M,N – total HRGs)[111]
Febrile neutropenia	£4,755.76 (£951.15)	Copley-Merriman 2018[122]
HRGs: Healthcare Resource Groups; NES: Non-Elective Short Stay		

B.3.6 Summary of base case analysis inputs and assumptions

A summary of the base case analysis inputs and assumptions are provided in Table 55 and Table 56, respectively.

Table 55. Summary of variables applied in the economic model

Variable	Value	Measurement of uncertainty and distribution	Section
Baseline parameters			
Baseline parameters	Table 33	SE (age: normal; sex: beta)	B.3.2.2
Survival and progression functions			
Overall survival	Table 35	Described in Section B.3.3.1	B.3.3.1
Progression-free survival			
All-cause mortality	Not applicable	None	B.3.3.1.4
Clinical parameters			
Time on treatment	Figure 43, Figure 44	Described in Section B.3.3.2	B.3.3.2
AE prevalence	Table 38	SE (beta)	B.3.3.3
Utilities			
Health state utilities	Table 43	SE (beta)	B.3.4.4
Costs			
Medication costs	Table 46, Table 47, Table 48, Table 49	Not applicable	B.3.5.1.1
Health state costs	Table 53	SE (gamma)	B.3.5.1.4
Terminal care costs	B.3.5.1.3	SE (gamma)	B.3.5.1.3
AE costs	Table 54	SE (gamma)	B.3.5.2
Subsequent therapy costs	Table 50, Table 51, Table 52	Not applicable	B.3.5.1.2
AE: adverse events; SE: standard error.			

Table 56. Assumptions applied in the economic model

Assumption	Rationale
Baseline parameters are derived from CheckMate 648 cohort, which is assumed to be reflective of patients seen in UK clinical practice for the anticipated MA.	Although there may be differences between characteristics in CheckMate 648 and OSCC patients in UK clinical practice, these can be considered small, as shown in Table 23. Sensitivity analyses (probabilistic and deterministic) have been conducted to assess the impact of variability in these parameters.
The model applies a weekly cycle length, which is assumed to be sufficiently granular to accurately reflect costs and benefits when modelling OC.	Previous OC evaluations assessed by NICE had applied weekly cycle lengths, which was considered appropriate by ERG.[1,72,75,87] This cycle length is short enough to reflect the treatment cycles for patients and reflects the frequency of follow-up for patients and reflects the frequency of follow-up for patients and a realistic minimum time during which symptoms or response can change.
To reflect the nature of OC and available evidence, the model assumes that OC phases are consecutive, and patients cannot revert to pre-progression from more advanced phases of the disease.	This assumption has been validated by clinicians and is in line with other HTAs and economic analyses assessing the OC population.
Identification of most appropriate survival curves describing PFS and OS inform extrapolation	Extensive analyses have been undertaken to identify appropriate and conservative survival curves describing NIVO+CHEMO efficacy, with reference to the guidance from the NICE Decision Support Unit (DSU)[91] and Bagust and Beale (2014)[97]. The approach and identified survival extrapolations have been validated by clinical and health economic experts. However, to address the uncertainty around this parameter, scenario analyses have been conducted by applying alternative assumptions around extrapolations, as presented in Section B.3.3.1.
Efficacy has been based on BICR-assessed data, rather than investigator-assessed data	During CheckMate 648, the two measures of response of PFS were comparable. However, BICR was designated as the primary endpoint and may be considered slightly more conservative.
As a simplification, it is assumed that all adverse events occur in the first cycle of treatment.	The majority of patients during CheckMate 648 have discontinued treatment within the current database lock, so that the data can be considered an accurate reflection of the safety profile. AEs are often only observed to occur soon after treatment initiation, so that this may not be well reflected by assuming a constant rate per cycle.
It was assumed that health state utilities, pre-progression, post-progression and the disutility of death, are the same for the treatment and control arm.	This is based on evidence observed during CheckMate 648, described in Section B.3.4.2.

Assumption	Rationale
It was assumed that patients receiving pembrolizumab in combination with chemotherapy experience missing or delayed doses in line with nivolumab during CheckMate 648.	Currently, there is no published data available to inform proportion of received doses of pembrolizumab. As the mechanism of action is similar, this seems an appropriate assumption.
The health state resource use is derived from evidence presented in TA737.	Robust estimates of health state resource use for patients in this setting are not publicly available, given the limited alternative treatment available for which evidence may have previously gathered. In order to provide relevant economic evaluations and facilitate comparison between these appraisals, health state resource use from TA737 is applied.
Subsequent treatment for NIVO-CHEMO and PEMBRO+CHEMO is assumed to be single agent taxane (equal use of paclitaxel and docetaxel).	During CheckMate 648, taxane use reflected around 70% of subsequent systemic therapy use, indicating the plausibility of this assumption. Docetaxel and paclitaxel have similar efficacy and cost.
Subsequent treatment for CHEMO is assumed to nivolumab monotherapy.	This aligns with the current UK treatment pathway and is aligned with budget impact assumptions applied during TA707.[72]
AE utility decrement values were assumed for certain AEs.	Values were assumed for those AEs where published data was not available. However, deterministic sensitivity analysis has been presented to show the impact of AE utility decrements.
No treatment waning has been assumed.	Evidence supports a robust and durable treatment effect lasting beyond discontinuation for immunotherapies.[94] Further, during TA737, the committee concluded that all scenarios provided plausible estimates of overall survival and the treatment waning scenarios were not greatly different from those without treatment waning.[1] This is of particular relevance given the low long-term hazard in the CHEMO arm of CheckMate 648.

B.3.7 Base-case results

The results of the base case analysis are summarised in Table 57.

For patients treated with chemotherapy, the model predicted [redacted] discounted life years with an accrual of [redacted] discounted QALYs. Nivolumab use with chemotherapy was estimated to result in an additional [redacted] discounted QALYs (total [redacted] QALYs) and an additional [redacted] discounted life years (total [redacted] life years).

Total discounted costs associated with nivolumab and chemotherapy were predicted to be £[redacted]. Incremental costs were predicted to be £[redacted] compared to chemotherapy alone, under base-case assumptions. The resulting ICER estimate for nivolumab with chemotherapy

versus chemotherapy alone was £33,272 per QALY gained. Therefore, the base-case ICER is below the £50,000 per QALY willingness-to-pay threshold.

Table 57. Overview of base case analysis results (with PAS; discounted)

	NIVO-CHEMO	CHEMO	Incremental
Life years	[redacted]	[redacted]	[redacted]
QALYs	[redacted]	[redacted]	[redacted]
Total costs (£)	[redacted]	[redacted]	[redacted]
ICER (£/QALY)			£33,272
ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year			

Table 58. Detailed base case analysis results

	NIVO-CHEMO	CHEMO
Patient level survival (undiscounted)		
Median PFS (years)	[redacted]	0.383
Mean PFS (years)	[redacted]	0.576
Median OS (years)	[redacted]	0.747
Mean OS (years)	[redacted]	1.382
Patient-level progression		
Time in pre-progression (years)	[redacted]	0.576
Time in post-progression (years)	[redacted]	0.807
Costs (with PAS)		
Health state costs	[redacted]	£7,290
Treatment costs	[redacted]	£11,355
AE costs for initial therapy	[redacted]	£82
Terminal care costs	[redacted]	£8,768
Total costs	[redacted]	£27,494
Health benefits		
HS QALYs	[redacted]	0.931
Adverse event utility	[redacted]	-0.0001
Time-to-death utility	[redacted]	-0.0142
Total QALYs	[redacted]	0.917
Total LYs (undiscounted)	[redacted]	1.382
Incremental results		
Incremental total costs	-	[redacted]
Incremental QALYs	-	[redacted]
Incremental LYs (undiscounted)	-	[redacted]
Cost/QALY	**-**	**£33,272**
AE: adverse event; HS: health state; LY: life year; OS: overall survival; PFS: progression free survival; QALY: quality-adjusted life year; ToT: Time on Treatment.		

B.3.8 Sensitivity analysis

In order to assess the impact of parameters on the model outcomes, deterministic sensitivity analyses have been used to vary the data inputs by a set amount. Uncertainty around the input data has been assessed using probabilistic analyses, while alternative assumptions have been examined in scenario analyses.

B.3.8.1 Probabilistic sensitivity analysis

In the probabilistic sensitivity analysis (PSA), a non-parametric bootstrapping approach will be taken, sampling values from distributions around the means of input parameters in the model. Sampling utilises information of the mean and standard error of parameters to derive an estimated value using an appropriate distribution (costs: gamma, age and survival parameters: normal, utilities, probabilities and proportions: beta). These analyses are used to estimate the overall uncertainty that exists in the model results due to uncertainty in the chosen input parameters.

The majority of parameters included in the PSA are sampled independently, with the exception of semi-parametric survival estimates, where parameters associated with individual survival function are sampled using a common random number.

Several inputs are derived from sources where it has not been possible to ascertain standard errors. To assess uncertainty surrounding these inputs, the standard error has been assumed to be 20% of the mean value for the purposes of the PSA.

In order to enable the model results to converge to a sufficient degree of accuracy, 1000 simulations of the model were required.

Results from 1,000 iterations of the model using probabilistic values can be seen in Table 59 and show that results are in line with the deterministic analysis. The scatterplot shows that there is limited spread in the values from each iteration and these are predominantly contained in the north east quadrant under the willingness-to-pay threshold, demonstrating cost-effectiveness (Figure 45). Out of the 1,000 iterations, approximately 88.7% estimated nivolumab to be cost effective (Figure 46) demonstrating a high certainty in the base case results.

Table 59. Probabilistic sensitivity analysis results (combined CR/PR)

	NIVO-CHEMO	CHEMO	Incremental
Life years	2.310	1.304	1.007
QALYs	1.669	0.940	0.730
Total costs (£)	£51,416	£27,533	£23,883
ICER (£/QALY)			£32,736
ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year			

Figure 45. Scatterplot of probabilistic results

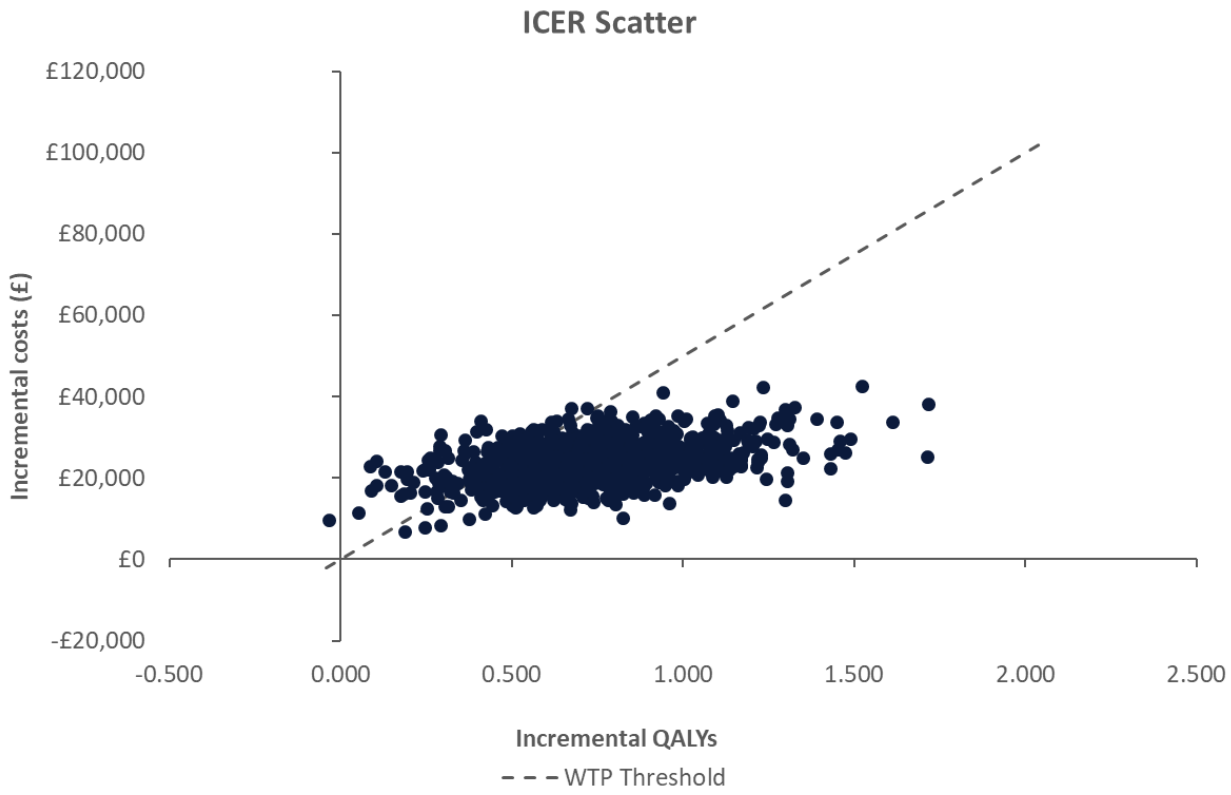
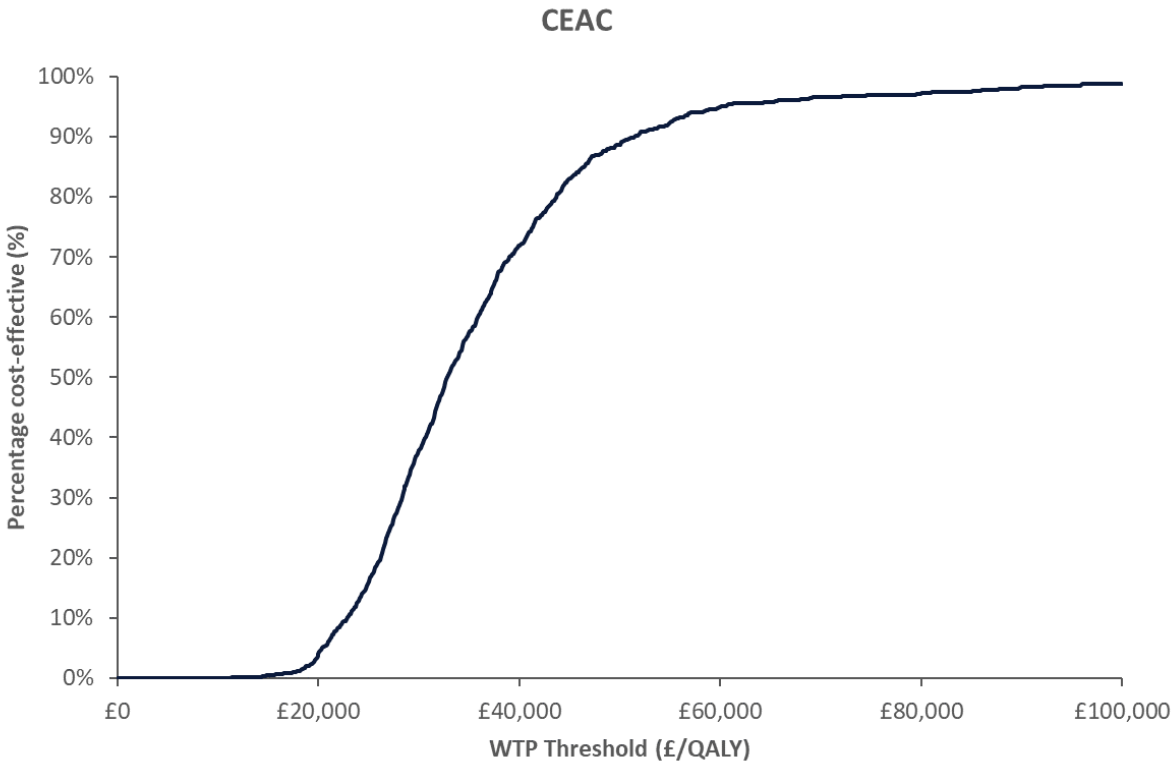


Figure 46. Cost-effectiveness acceptability curve



B.3.8.2 Deterministic sensitivity analysis

A range of one-way (deterministic) sensitivity analyses have been conducted, regarding the following assumption and parameters:

- Time horizon (260 weeks [5 year] and 520 weeks [10 years])
- Discounting: costs (0% and 6%)
- Discounting: benefits (0% and 6%)
- Baseline characteristics: age (± 20%, impacting on all-cause mortality)
- Baseline characteristics: sex (0% and 100% male, impacting on all-cause mortality)
- Health state costs: pre-progression and post-progression (± 20%)
- Health state costs: terminal care costs (± 20%)
- Initial treatment costs (± 20%)
- Subsequent treatment costs (± 20%)
- Adverse event costs (± 20%)
- Health state utility: pre-progression and post-progression (± 20%)
- End of life utility (± 20%)
- Adverse event disutility (± 20%)
- 2nd line time on treatment (± 20%)
- Treatment modifier: proportion receiving dose (± 20%)

- Adverse event probability (± 20%)
- Subsequent treatment ToT (± 20%)

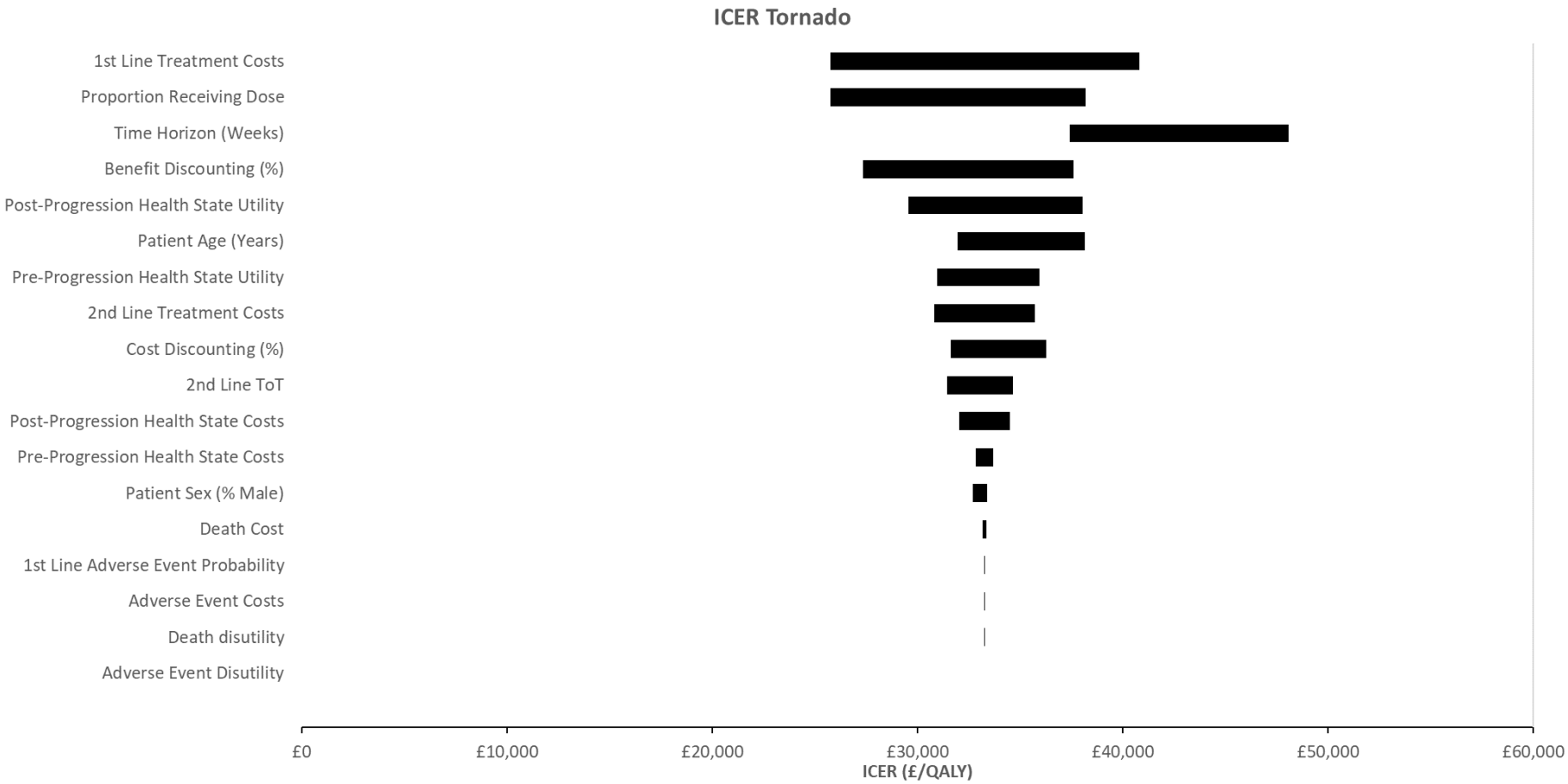
Deterministic sensitivity analysis (DSA) results indicate the parameters that influence the results and conclusions of the decision problem to the greatest degree (Table 60, Figure 47). Parameters with the greatest impact are first-line treatment costs, proportion of patients receiving a dose and the post-progression health state utility. NIVO+CHEMO was cost-effective in the majority of scenarios at a WTP threshold of £50,000/QALY, which indicates that the ICER is relatively stable across analyses.

Table 60. Deterministic sensitivity results

Scenario	Parameter variation	Incremental		ICER (£/QALY)	Base case ICER (£/QALY)
		Costs	QALY		
Time horizon (weeks	260	[redacted]	[redacted]	£48,094	£33,272
	520	[redacted]	[redacted]	£37,426	£33,272
Cost discount rate (%)	0%	[redacted]	[redacted]	£36,261	£33,272
	6%	[redacted]	[redacted]	£31,637	£33,272
Benefit discount rate (%)	0%	[redacted]	[redacted]	£27,334	£33,272
	6%	[redacted]	[redacted]	£37,589	£33,272
Patient age (years)	80%	[redacted]	[redacted]	£31,945	£33,272
	120%	[redacted]	[redacted]	£38,147	£33,272
Patient sex (% male)	0%	[redacted]	[redacted]	£32,709	£33,272
	100%	[redacted]	[redacted]	£33,396	£33,272
Pre-progression health state cost	80%	[redacted]	[redacted]	£32,863	£33,272
	120%	[redacted]	[redacted]	£33,680	£33,272
Post-progression health state cost	80%	[redacted]	[redacted]	£32,027	£33,272
	120%	[redacted]	[redacted]	£34,516	£33,272
Terminal care cost	80%	[redacted]	[redacted]	£33,360	£33,272
	120%	[redacted]	[redacted]	£33,184	£33,272
1st Line treatment costs	80%	[redacted]	[redacted]	£25,745	£33,272
	120%	[redacted]	[redacted]	£40,798	£33,272
2nd Line treatment costs	80%	[redacted]	[redacted]	£35,734	£33,272
	120%	[redacted]	[redacted]	£30,810	£33,272
Adverse event costs	80%	[redacted]	[redacted]	£33,247	£33,272
	120%	[redacted]	[redacted]	£33,297	£33,272
Pre-progression health state utility	80%	[redacted]	[redacted]	£35,927	£33,272
	120%	[redacted]	[redacted]	£30,981	£33,272
Post-progression health state utility	80%	[redacted]	[redacted]	£38,066	£33,272
	120%	[redacted]	[redacted]	£29,550	£33,272
End of life disutility	80%	[redacted]	[redacted]	£33,276	£33,272
	120%	[redacted]	[redacted]	£33,267	£33,272
Adverse event disutility	80%	[redacted]	[redacted]	£33,271	£33,272
	120%	[redacted]	[redacted]	£33,272	£33,272

Proportion receiving dose	80%	[redacted]	[redacted]	£25,745	£33,272
	120%	[redacted]	[redacted]	£38,178	£33,272
1st Line adverse event probability	80%	[redacted]	[redacted]	£33,246	£33,272
	120%	[redacted]	[redacted]	£33,297	£33,272
2nd Line Time on treatment	80%	[redacted]	[redacted]	£31,447	£33,272
	120%	[redacted]	[redacted]	£34,654	£33,272

Figure 47. Tornado diagram



B.3.8.3 Scenario analysis

B.3.8.3.1 Impact of alternative survival assumptions

Survival modelling using long-term extrapolation of parametric functions is subject to considerable uncertainty despite efforts to robustly and transparently provide survival curves that best represent patients in clinical practice. In order to assess the impact of alternative parametric fittings on the cost-effectiveness of NIVO-CHEMO, survival curves described in the survival analysis report (Appendix N) have been applied within the model as scenario analyses (Table 61). Additionally, a response-stratified approach was considered as scenario.

This analysis should be viewed within the context of identifying the most appropriate survival extrapolation, as detailed in Section B.3.3.1. Parametric extrapolation of survival data from CheckMate 648 was undertaken with reference to the guidance from the NICE Decision Support Unit (DSU)[123] and Bagust and Beale (2014).[97] Plausible extrapolations have been assessed based on the criteria provided in Appendix N.

Table 61. Alternative extrapolations applied during scenario analysis

	CHEMO		NIVO+CHEMO	
	OS	**PFS**	**OS**	**PFS**
Parametric	Log-logistic	Lognormal	Log-normal	Lognormal
Semi-parametric ([redacted] DBL): 6.9 month cut-point (alternative Oct 2021 database lock approach)	Generalised Gamma	Weibull	Exponential	Generalised Gamma
Semi-parametric (March 2020 DBL): 6.9 month cut-point	Weibull	Weibull	Weibull	Weibull
Response-stratified approach				
Complete Response/Partial Response	Generalised Gamma	Generalised Gamma	Lognormal	Generalised Gamma
Stable Disease	Generalised Gamma	Generalised Gamma	Weibull	Lognormal
Progressive Disease/Unable to determine	Lognormal	Log-logistic	Lognormal	Log-logistic

The impact of applying alternative survival extrapolations for the NIVO-CHEMO arm is shown in Table 62. As can be seen, all scenarios increase the ICER compared with the base case analysis, but almost all remain below the £50,000 willingness-to-pay threshold. However, discounted incremental QALYs and costs remain broadly consistent ([redacted] for QALYs; [redacted] for costs), indicating the consistency in benefit associated with long-term outcomes.

Table 62. Scenario analysis: impact of alternative survival approaches

Scenarios	Incremental costs (£)	Incremental LYs	Incremental QALYs	ICER (£/QALY)
Base case analysis	[redacted]	[redacted]	[redacted]	**£33,272**
Scenario: response-stratified approach	[redacted]	[redacted]	[redacted]	**£44,958**
Scenario: parametric survival approach	[redacted]	[redacted]	[redacted]	**£38,072**
Scenario: alternative [redacted] database lock approach	[redacted]	[redacted]	[redacted]	**£62,594**
Scenario: March 2020 database lock	[redacted]	[redacted]	[redacted]	**£39,054**
CHEMO: chemotherapy; ICER, incremental cost-effectiveness ratio; LYs: life years; NIVO: nivolumab; QALYs, quality-adjusted life years				

B.3.8.3.2 Alternative comparators

The base case analysis informed by CheckMate 648 compares NIVO-CHEMO versus CHEMO, where chemotherapy is assumed to be cisplatin plus fluorouracil. As outlined in Section B.3.2.3, this can be considered clinically appropriate based on current guidelines, clinical evidence and expert opinion.

However, in order to inform decision-making, a comparison of NIVO-CHEMO against other potential comparators has been provided as a scenario analysis, specifically FOLFOX, XELOX and cisplatin plus capecitabine. Efficacy is assumed to be equivalent between doublet therapies, as per clinical expert opinion.

Additionally, scenario analyses were undertaken versus pembrolizumab plus chemotherapy. Efficacy inputs were derived from the ITC, described in Section B.2.9, were applied to the CHEMO survival curves to generate the PEMBRO-CHEMO survival curves. The ToT curves were derived from the PEMBRO-CHEMO arm of KEYNOTE-590 with a mean ToT of 33.67 weeks.[1]

As described in Table 63, the efficacy and cost profiles of PEMBRO-CHEMO and NIVO-CHEMO are comparable, resulting in marginal differences in total costs, LYs and QALYs between the two treatment options (£[redacted], respectively), leading to a not cost-effective ICER. When the additional alternative comparators, FOLFOX, XELOX and cisplatin plus capecitabine, are run as scenarios, the ICERs decrease compared to the base case (£30,068 for FOLFOX scenario, £32,975 for XELOX scenario and £33,162 for cisplatin plus capecitabine, compared to £33,272 in the base case). As the CHEMO efficacy profile is used for the additional alternative comparators, these reductions in the ICER result from the decrease in the incremental costs due to the increased cost of the comparators.

Table 63. Scenario analysis: impact of alternative comparators

Technologies	Incremental costs (£)	Incremental LYs	Incremental QALYs	ICER (£/QALY)
Base case analysis	[redacted]	[redacted]	[redacted]	**£33,272**
Scenario: Pembrolizumab (HR approach)	[redacted]	[redacted]	[redacted]	**-£5,582**
Scenario: FOLFOX	[redacted]	[redacted]	[redacted]	**£30,068**
Scenario: XELOX	[redacted]	[redacted]	[redacted]	**£32,975**
Scenario: cisplatin plus capecitabine	[redacted]	[redacted]	[redacted]	**£33,162**
CHEMO: chemotherapy; ICER, incremental cost-effectiveness ratio; LYs: life years; NIVO: nivolumab; QALYs, quality-adjusted life years				

As can be seen in Table 63, the efficacy and cost profiles of PEMBRO-CHEMO and NIVO-CHEMO are comparable, resulting in marginal differences in total costs, LYs and QALYs between the two treatment options (£478, -0.001, and -0.086, respectively), leading to a not cost-effective ICER. When the additional alternative comparators, FOLFOX, XELOX and cisplatin plus capecitabine, are run as scenarios, the ICERs decrease compared to the base case (£30,068 for FOLFOX scenario, £32,975 for XELOX scenario and £33,162 for cisplatin plus capecitabine, compared to £33,272 in the base case). As the CHEMO efficacy profile is used for the additional alternative comparators, these reductions in the ICER result from the decrease in the incremental costs due to the increased cost of the comparators.

B.3.8.3.3 Removal of the treatment modifier

A treatment modifier was used in the base case analysis to reflect doses that were missed or delayed during CheckMate 648. To explore the impact of this on the ICER, a scenario was run without the treatment modifier and results are displayed in Table 64. The removal of the treatment modifier increased the ICER to £39,598 per QALY. However, this remained below the £50,000 per QALY willingness-to-pay threshold.

Table 64. Scenario analysis: impact of removing treatment modifier

Technologies	Incremental costs (£)	Incremental LYs	Incremental QALYs	ICER (£/QALY)
Base case analysis	[redacted]	[redacted]	[redacted]	**£33,272**
Scenario: removing treatment modifier	[redacted]	[redacted]	[redacted]	**£38,512**
CHEMO: chemotherapy; ICER, incremental cost-effectiveness ratio; LYs: life years; NIVO: nivolumab; QALYs, quality-adjusted life years				

B.3.8.3.4 Impact of alternative utility assumptions

In the base case analysis, time to death utilities were implemented in the month prior to death. A scenario exploring the impact of not using time to death utilities was conducted. Results are displayed in Table 65, where the ICER increased slightly but remained below the £50,000 per QALY willingness-to-pay threshold.

Table 65. Scenario analysis: impact of removing time to death utilities

Technologies	Incremental costs (£)	Incremental LYs	Incremental QALYs	ICER (£/QALY)
Base case analysis	[redacted]	[redacted]	[redacted]	**£33,272**
Scenario: removing time to death utilities	[redacted]	[redacted]	[redacted]	**£33,295**

CHEMO: chemotherapy; ICER, incremental cost-effectiveness ratio; LYs: life years; NIVO: nivolumab; QALYs, quality-adjusted life years

B.3.8.4 Summary of the sensitivity analysis results

The sensitivity analyses show that the base case analysis is robust to the natural variation that may be seen in clinical practice. The PSA shows that in 71.6% of instances, nivolumab with chemotherapy would be considered cost-effective, which is within normal bounds. The most influential parameters on cost-effectiveness are the first-line treatment costs, proportion receiving a dose and the post-progression health state utility.

B.3.10 Validation

B.3.10.1 Validation of cost-effectiveness analysis

In general, where no evidence was identified to validate the results of the cost-effectiveness analysis, simple assumptions have been made based on independent sources, such as published literature, OC guidelines or previous NICE appraisals in the field of OC. These assumptions were assessed for clinical plausibility; uncertainty will be characterised through the use of sensitivity analyses. Extensive sensitivity analyses were then undertaken, and all ICERs remain below a £50,000/QALY threshold.

A technical review of the cost-effectiveness model was conducted by an independent economist. Further, the relevance of the model structure and assumptions were validated through consultation with UK clinicians. This allowed the model approach to be validated and permitted areas of disagreement to be resolved prior to generation of model results. In addition, quality control was undertaken, whereby a cell-by-cell verification process was conducted to allow checking of all input calculation, formulae and visual basic code.

B.3.10.2 Exploration of survival extrapolation techniques for cancer immunotherapy

Limited clinical trial follow-up and low event rates for survival endpoints introduce uncertainty in survival extrapolation for immunotherapies. Traditional conservative approaches can adversely impact estimates of cost-effectiveness, impacting on HTA outcomes as well as restricting patients' access to these medicines. There is also growing evidence that parametric survival models may be unable to capture the characteristic plateau observed in the latter period of immunotherapy survival curves, as well as to model hazard functions with multiple inflection points.[124-127]

To examine this problem, two case studies have been undertaken to assess a range of extrapolation methods to patient-level survival data to assess their predictive accuracy over time. Multiple extrapolation methods were examined: standard parametric models, natural cubic splines, piecewise models combining Kaplan-Meier data with an exponential or non-exponential distribution, response-based landmark models, mixture cure models and parametric mixture models. Data from two separate studies were assessed:

- CheckMate 067: Phase III randomised controlled study that compared PFS and OS of nivolumab monotherapy and NIVO+IPI versus ipilimumab monotherapy in patients with previously untreated, unresectable or metastatic melanoma, using data cuts at 28 months, 40 months, 52 months and 60 months.[128]

- CheckMate 025: Phase III randomised controlled trial comparing nivolumab with everolimus for previously treated advanced renal cell carcinoma, using data cuts at 15 months, 27 months, 39 months and 64 months.[129]

The extrapolation models were fitted to the earlier database locks and NICE DSU 14 was used to inform model selection. The extrapolations for each model were compared with the observed data from the latest database locks.

In the CheckMate 025 case study, all extrapolation methods, with the exception of mixture models, underestimated landmark and mean OS for nivolumab compared to long-term follow-up data. OS estimates for everolimus tended to be more accurate, with four of the six methods providing landmark OS estimates within the 95% confidence interval of observed OS as per the latest dataset. The predictive accuracy of survival extrapolation methods fitted to nivolumab also showed greater variation than for everolimus.[129]

In the CheckMate 067 case study, the parametric models, spline models and piecewise models consistently underestimated survival at 60 months. The methods that explicitly model heterogeneity in the patient population (mixture cure models, parametric mixture models and response-based landmark models) generally aligned with each other and provided accurate and consistent estimates of OS across a range of follow-up periods, including landmark survival at 60 months.[128]

In a similar study, survival modelling from TA319 (ipilimumab in melanoma) was revisited to assess the accuracy of extrapolation methods.[127] In addition to the piecewise survival model used in TA319, alternative models were assessed (fit to trial data with minimum follow-up of 3 years), including parametric, spline-based, mixture, and mixture-cure models. These were compared against a longer-term data cut (5-year follow-up). Only the survival model used in TA319 and a mixture-cure model provided 5-year survival predictions close to those observed in the 5-year follow-up data set. Standard parametric, spline, and non–curative-mixture models substantially underestimated 5-year survival.[127]

Based on this evidence, it can be concluded that estimating long-term survival for NIVO-CHEMO through piecewise Kaplan-Meier and extrapolation may underestimate the long-term survival benefit of this therapy, and the current estimates may be considered conservative.

B.3.10.3 Comparison of outputs with TA737

TA737 provides outputs for PEMBRO+CHEMO and CHEMO, although these outputs apply a PAS discount to costs.[130] A comparison of outputs for the current submission versus TA737 is provided in Table 66. As can be seen, predicted LYs are broadly comparable with values output from TA737. Predicted costs are also comparable. QALY outcomes have slightly more variation than those produced during TA737. However, this may be due to slight differences in PFS and OS outcomes.

Table 66. Comparison of outcomes for trifluridine-tipiracil

		Current appraisal	TA737[130]	
			Company	ERG
Total LYs	CHEMO	1.28	1.37	NR
	Intervention	2.28	2.13	NR
Incremental costs (£)*		23,999	27,165	28,007
Incremental QALYs		0.70	0.63	0.54
* applies PAS for intervention arms				

B.3.10.4 Comparison of economic model output with CheckMate 648 data

A comparison between the economic model output and the CheckMate 648 data was carried out as an additional validation exercise. The output of this validation exercise is displayed in Table 67. As can be seen, there is only a small variation between the CheckMate 648 data and the model output, confirming the model results provide a good representation of the available data.

Table 67. Comparison of economic model output with CheckMate 648 data

		NIVO-CHEMO			CHEMO		
		PLD	Preferred Survival Curves	Model Output	PLD	Preferred Survival Curves	Model Output
OS	1 year	57.62%	53.92%	53.92%	37.26%	34.10%	34.33%
	2 years	31.93%	31.93%	31.93%	11.95%	15.15%	15.26%
	3 years	17.93%	22.85%	22.73%	10.14%	9.24%	9.23%
	5 years	NR	14.35%	14.30%	NR	4.72%	4.73%
	10 years	NR	6.94%	6.93%	NR	1.70%	1.71%
PFS	1 year	25.39%	26.50%	26.31%	10.30%	13.24%	12.70%
	2 years	11.79%	10.49%	10.44%	2.75%	3.11%	3.16%
	3 years	5.90%	5.21%	5.12%	2.75%	0.96%	1.32%
	5 years	NR	1.70%	1.91%	NR	0.13%	0.41%
	10 years	NR	0.21%	0.48%	NR	0.00%	0.07%

CHEMO: chemotherapy; NIVO: nivolumab; OS: Overall survival; PFS: Progression-free survival; PLD: Patient-level data

B.3.11 Interpretation and conclusion of economic analysis

Base case analysis

- Use of NIVO-CHEMO results in an increased mean OS ([redacted] years versus [redacted] years, undiscounted), as well as additional discounted QALYs and life years of up to [redacted] and [redacted], respectively.
- Discounted incremental costs were estimated to be [redacted] under base case assumptions and the resultant ICERs were £33,272 per QALY, which is considered to be cost-effective at a willingness-to-pay threshold of £50,000 per QALY.

Sensitivity analysis

- In the probabilistic sensitivity analysis, NIVO-CHEMO was cost-effective in 88.7% of scenarios at a willingness-to-pay threshold of £50,000 per QALY.
- In the deterministic sensitivity analysis, NIVO-CHEMO was cost-effective in all scenarios at a willingness-to-pay threshold of £50,000 per QALY.
- Extensive scenario analyses were undertaken, reflecting the assumptions required to undertake plausible, robust and transparent base case analysis. Within these scenario analyses, almost all ICERs remain below the £50,000 per QALY threshold
- Therefore, NIVO-CHEMO can be considered a cost-effective use of NHS resources.

Prognosis for advanced OSCC remains poor: a retrospective analysis has shown that nearly half of patients undergoing systemic treatment for OSCC in the first-line setting do not respond to their treatment and over a third of patients progress to the next line of treatment.[42] There is currently a high unmet need for effective first-line treatments for patients with advanced OC, particularly in patients with PD-L1 CPS <10, where doublet palliative chemotherapy is the only therapy available.

In the base case analysis, it was estimated that NIVO-CHEMO use would result in [redacted] discounted QALYs and [redacted] discounted LYs. Discounted incremental costs were estimated to be [redacted] compared to chemotherapy alone under base case assumptions and the resultant ICER was £33,272 per QALY, which can be considered cost-effective at a willingness-to-pay threshold of £50,000 per QALY.

A large number of sensitivity analyses have been undertaken, assessing the impact of variation in all variables and assumptions applied within the model. In the DSA and PSA, NIVO+CHEMO was cost-effective in the majority of scenarios at a WTP threshold of £50,000/QALY. This indicates that the ICER is relatively stable across analyses.

This analysis has been designed to be aligned with TA737, facilitating review and transparency. For this reason, a comparison of published outcomes from TA737 have been provided within Section B.3.10.3. Although there remains a number of evidence gaps, aligning with a previously undertaken NICE HTA supports a robust approach to analysis.

The addition of nivolumab to standard chemotherapy for adults with unresectable, advanced, recurrent or metastatic, previously untreated OSCC would provide an

opportunity to make a significant and substantial impact on health-related benefits, address a current unmet need, and would represent a further, significant advance in the management of this end-of-life condition.

References

1. National Institute for Health and Care Excellence. Pembrolizumab with platinum- and fluoropyrimidine-based chemotherapy for untreated advanced oesophageal and gastro-oesophageal junction cancer [TA737]. 2021. Available at: https://www.nice.org.uk/guidance/ta737/chapter/1-Recommendations [Accessed 09/12/21].
2. Bristol-Myers Squibb. Advisory Board Report from 16 July 2021 [data on file]. 2021.
3. European Medicines Agency. OPDIVO Summary of Product Characteristics. 2022. Available at: https://www.ema.europa.eu/en/medicines/human/EPAR/opdivo#product-information-section [Accessed 04/01/2022].
4. Bristol-Myers Squibb. Clinical Protocol CA209648 [Data on file]. 2020.
5. Cancer Research UK. What is oesophageal cancer? 2019. Available at: https://www.cancerresearchuk.org/about-cancer/oesophageal-cancer/about [Accessed 15/06/2021].
6. Smyth EC, Lagergren J, Fitzgerald RC, et al. Oesophageal cancer. Nat Rev Dis Primers. 2017;3:17048.
7. Cancer Research UK. Oesophageal cancer incidence statistics. 2019. Available at: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/oesophageal-cancer/incidence [Accessed 15/06/2021].
8. BMJ Best Practice. Oesophageal cancer. 2018. Available at: https://bestpractice.bmj.com/topics/en-gb/1029/pdf/1029.pdf [Accessed 12/12/21].
9. World Cancer Research Fund. Oesophageal cancer. 2022. Available at: https://www.wcrf-uk.org/cancer-types/oesophageal-cancer/ [Accessed 12/01/2022].
10. National Cancer Institute. SEER Cancer Statistics Review (CSR) 1975-2016. 2019. Available at: https://seer.cancer.gov/csr/1975_2016/results_merged/sect_08_esophagus.pdf [Accessed 15/06/2021].
11. Cancer Research UK. Oesophageal cancer survival. 2016. Available at: https://www.cancerresearchuk.org/about-cancer/oesophageal-cancer/survival [Accessed 15/06/2021].
12. National Institute for Health and Care Excellence. Oesophago-gastric cancer: assessment and management in adults. NICE guideline [NG83]. 2018. Available at: https://www.nice.org.uk/guidance/ng83 [Accessed 15/06/2021].
13. Bristol-Myers Squibb. CheckMate 648 Oct 2021 Summary data [Data on file]. 2021.
14. Cancer Research UK. Oesophageal cancer statistics. 2017. Available at: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/oesophageal-cancer [Accessed 15/06/2021].
15. Mayo Clinic. Esophageal cancer—Symptoms and causes. . 2021. Available at: https://www.mayoclinic.org/diseases-conditions/esophageal-cancer/symptoms-causes/syc-20356084 [Accessed 24/01/2022].
16. Short MW, Burgers KG, Fry VT. Esophageal Cancer. Am Fam Physician. 2017;95(1):22-8.
17. Kleinberg L, Kelly R, Yang S, et al. 74 - Cancer of the Esophagus. In: Niederhuber JE, Armitage JO, Doroshow JH, Kastan MB, Tepper JE, editors. Abeloff's Clinical Oncology (Fifth Edition). Philadelphia: Churchill Livingstone; 2014. p. 1207-39.e7.
18. Lordick F, Mariette C, Haustermans K, et al. Oesophageal cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Annals of Oncology. 2016;27:v50-v7.
19. Wu S-G, Zhang W-W, He Z-Y, et al. Sites of metastasis and overall survival in esophageal cancer: a population-based study. Cancer management and research. 2017;9:781-8.

20. Wu SG, Zhang WW, Sun JY, et al. Patterns of Distant Metastasis Between Histological Types in Esophageal Cancer. Front Oncol. 2018;8:302.
21. Ai D, Zhu H, Ren W, et al. Patterns of distant organ metastases in esophageal cancer: a population-based study. J Thorac Dis. 2017;9(9):3023-30.
22. Cancer Research UK. Oesophageal cancer incidence statistics. 2022. Available at: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/oesophageal-cancer/incidence#heading-Zero [Accessed 12/01/2022].
23. Cancer Research UK. Oesophageal cancer mortality statistics. 2021. Available at: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/oesophageal-cancer/mortality [Accessed 15/06/2021].
24. GLOBOCAN. Oesophagus Fact Sheet. . 2020. Available at: https://gco.iarc.fr/today/data/factsheets/cancers/6-Oesophagus-fact-sheet.pdf [Accessed 12/01/2022].
25. Office for National Statistics. Cancer survival in England - adults diagnosed. 2019. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/datasets/cancersurvivalratescancersurvivalinenglandadultsdiagnosed [Accessed 15/06/2021].
26. Wu S-G, Xie W-H, Zhang Z-Q, et al. Surgery Combined with Radiotherapy Improved Survival in Metastatic Esophageal Cancer in a Surveillance Epidemiology and End Results Population-based Study. Scientific reports. 2016;6:28280-.
27. Cao W, Xu C, Lou G, et al. A phase II study of paclitaxel and nedaplatin as first-line chemotherapy in patients with advanced esophageal cancer. Jpn J Clin Oncol. 2009;39(9):582-7.
28. Hiramoto S, Kato K, Shoji H, et al. A retrospective analysis of 5-fluorouracil plus cisplatin as first-line chemotherapy in the recent treatment strategy for patients with metastatic or recurrent esophageal squamous cell carcinoma. Int J Clin Oncol. 2018;23(3):466-72.
29. Kato K, Cho BC, Takahashi M, et al. Nivolumab versus chemotherapy in patients with advanced oesophageal squamous cell carcinoma refractory or intolerant to previous chemotherapy (ATTRACTION-3): a multicentre, randomised, open-label, phase 3 trial. Lancet Oncol. 2019;20(11):1506-17.
30. Liu Y, Ren Z, Yuan L, et al. Paclitaxel plus cisplatin vs. 5-fluorouracil plus cisplatin as first-line treatment for patients with advanced squamous cell esophageal cancer. 2016. Available at: http://europepmc.org/abstract/MED/27822423
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5088297/pdf/ajcr0006-2345.pdf [Accessed 2016].
31. Ross P, Nicolson M, Cunningham D, et al. Prospective Randomized Trial Comparing Mitomycin, Cisplatin, and Protracted Venous-Infusion Fluorouracil (PVI 5-FU) With Epirubicin, Cisplatin, and PVI 5-FU in Advanced Esophagogastric Cancer. Journal of Clinical Oncology. 2002;20(8):1996-2004.
32. Wang J, Chang J, Yu H, et al. A phase II study of oxaliplatin in combination with leucovorin and fluorouracil as first-line chemotherapy in patients with metastatic squamous cell carcinoma of esophagus. Cancer Chemother Pharmacol. 2013;71(4):905-11.
33. Arnold M, Morgan E, Bardot A, et al. International variation in oesophageal and gastric cancer survival 2012–2014: differences by histological subtype and stage at diagnosis (an ICBP SURVMARK-2 population-based study). Gut. 2021:gutjnl-2021-325266.
34. Shi J-F, Huang H-Y, Guo L-W, et al. Quality-of-life and health utility scores for common cancers in China: a multicentre cross-sectional survey. The Lancet. 2016;388:S29.
35. Doherty MK, Leung Y, Su J, et al. Health utility scores from EQ-5D and health-related quality of life in patients with esophageal cancer: a real-world cross-sectional study. Dis Esophagus. 2018;31(12).

36. Jaffe D, DeCongelio M, Gricar J. HSR21-055: Treatment Patterns and Healthcare Resource Utilization in First-Line Therapy for Advanced Esophageal Squamous Cell Carcinoma in Asian and Western Countries. Journal of the National Comprehensive Cancer Network J Natl Compr Canc Netw. 2021;19(3.5):HSR21-055-HSR21-.
37. Davidson M, Chau I, Cunningham D, et al. Impact of tumour histological subtype on chemotherapy outcome in advanced oesophageal cancer. World J Gastrointest Oncol. 2017;9(8):333-40.
38. van Rossum PSN, Mohammad NH, Vleggaar FP, et al. Treatment for unresectable or metastatic oesophageal cancer: current evidence and trends. Nat Rev Gastroenterol Hepatol. 2018;15(4):235-49.
39. GBD. Esophageal cancer—Level 3 cause. Institute for Health Metrics and Evaluation. 2020. Available at: http://www.healthdata.org/results/gbd_summaries/2019/esophageal-cancer-level-3-cause [Accessed 24/01/2022].
40. Bedi PS, Rai MP, Kandola SK, et al. In-Hospital Resource Utilization and Treatment Outcomes Among Patients With Esophageal Cancer: A Nationwide Analysis: 337. Official journal of the American College of Gastroenterology | ACG. 2018;113.
41. Hofmarcher T, & Lindgren, P. . The cost of cancers of the digestive system in Europe. 2020. Available at: https://ihe.se/en/publicering/cost-of-cancers-of-the-digestive-system/ [Accessed 24/01/2022].
42. Jaffe D, DeCongelio M, Dubell A, et al. PCN31 TREATMENT PATTERNS AND HEALTHCARE RESOURCE UTILIZATION (HCRU) IN SECOND-LINE (2L) ADVANCED ESOPHAGEAL SQUAMOUS CELL CARCINOMA IN ASIAN VERSUS WESTERN COUNTRIES. Value in Health. 2019;22:S441.
43. National Institute for Health and Care Excellence. Oesophageal and gastric cancer overview. NICE pathways. 2018. Available at: https://pathways.nice.org.uk/pathways/oesophageal-and-gastric-cancer#path=view%3A/pathways/oesophageal-and-gastric-cancer/oesophageal-and-gastric-cancer-overview.xml&content=view-index [Accessed 15/06/2021].
44. National Institute for Health and Care Excellence. Oesophago-gastric cancer. Quality standard [QS176] 2018. Available at: https://www.nice.org.uk/guidance/QS176 [Accessed 15/06/2021].
45. Brahmer JR, Drake CG, Wollner I, et al. Phase I study of single-agent anti-programmed death-1 (MDX-1106) in refractory solid tumors: safety, clinical activity, pharmacodynamics, and immunologic correlates. J Clin Oncol. 2010;28(19):3167-75.
46. Freeman GJ, Long AJ, Iwai Y, et al. Engagement of the PD-1 immunoinhibitory receptor by a novel B7 family member leads to negative regulation of lymphocyte activation. J Exp Med. 2000;192(7):1027-34.
47. Blank C, Gajewski TF, Mackensen A. Interaction of PD-L1 on tumor cells with PD-1 on tumor-specific T cells as a mechanism of immune evasion: implications for tumor immunotherapy. Cancer Immunol Immunother. 2005;54(4):307-14.
48. Latchman Y, Wood CR, Chernova T, et al. PD-L2 is a second ligand for PD-1 and inhibits T cell activation. Nat Immunol. 2001;2(3):261-8.
49. Ansell SM. Targeting immune checkpoints in lymphoma. Current Opinion in Hematology. 2015;22(4).
50. Ohigashi Y, Sho M, Yamada Y, et al. Clinical significance of programmed death-1 ligand-1 and programmed death-1 ligand-2 expression in human esophageal cancer. Clinical cancer research. 2005;11(8):2947-53.
51. Bristol-Myers Squibb. Nivolumab+chemotherapy or +ipilimumab - Primary Clinical Study Report for Study CA209648. [Data on file]. 2021.
52. Berman D, Korman A, Peck R, et al. The development of immunomodulatory monoclonal antibodies as a new therapeutic modality for cancer: The Bristol-Myers Squibb experience. Pharmacology & Therapeutics. 2015;148:132-53.

53. Frelaut M, du Rusquec P, de Moura A, et al. Pseudoprogression and Hyperprogression as New Forms of Response to Immunotherapy. BioDrugs. 2020;34(4):463-76.
54. West H. Immune Checkpoint Inhibitors. JAMA Oncology. 2015;1(1):115-.
55. Chiou VL, Burotto M. Pseudoprogression and Immune-Related Response in Solid Tumors. Journal of Clinical Oncology. 2015;33(31):3541-3.
56. Borcoman E, Nandikolla A, Long G, et al. Patterns of Response and Progression to Immunotherapy. Am Soc Clin Oncol Educ Book. 2018;38:169-78.
57. Park R, Williamson S, Kasi A, et al. Immune Therapeutics in the Treatment of Advanced Gastric and Esophageal Cancer. Anticancer Research. 2018;38(10):5569.
58. Kitagawa Y, Uno T, Oyama T, et al. Esophageal cancer practice guidelines 2017 edited by the Japan esophageal society: part 2. Esophagus. 2019;16(1):25-43.
59. Clinicaltrials.gov. A Study to Evaluate Efficacy in Subjects With Esophageal Cancer Treated With Nivolumab and Ipilimumab or Nivolumab Combined With Fluorouracil Plus Cisplatin Versus Fluorouracil Plus Cisplatin (CheckMate 648). 2021. Available at: https://clinicaltrials.gov/ct2/show/NCT03143153 [Accessed 09/06/2021].
60. Chau I, Doki Y, Ajani JA, et al. Nivolumab (NIVO) plus ipilimumab (IPI) or NIVO plus chemotherapy (chemo) versus chemo as first-line (1L) treatment for advanced esophageal squamous cell carcinoma (ESCC): First results of the CheckMate 648 study. Journal of Clinical Oncology. 2021;39(18_suppl):LBA4001-LBA.
61. Maurer W, Bretz F. Multiple Testing in Group Sequential Trials Using Graphical Approaches. Statistics in Biopharmaceutical Research. 2013;5(4):311-20.
62. Centre for Reviews and Dissemination. Systematic reviews: CRD's guidance for undertaking reviews in health care (University of York Centre for Reviews and Dissemination 2009. Available at: https://www.york.ac.uk/media/crd/Systematic_Reviews.pdf [Accessed 01/04/2021].
63. Doki Y, Ajani JA, Kato K, et al. Nivolumab Combination Therapy in Advanced Esophageal Squamous-Cell Carcinoma. N Engl J Med. 2022;386(5):449-62.
64. Bristol-Myers Squibb. CA209648 BMS-936558 Clinical study report supplemental tables [data on file]. 2021.
65. Achana FA, Cooper NJ, Bujkiewicz S, et al. Network meta-analysis of multiple outcome measures accounting for borrowing of information across outcomes. BMC Med Res Methodol. 2014;14:92.
66. Cope S, Chan K, Jansen J. Multivariate Network Meta-Analysis Of Survival Function Parameters. Research synthesis methods. 2020.
67. Dias S, Ades AE, Welton NJ, et al. Generalised linear models. In: Network meta-analysis for decision-making.: John Wiley & Sons; 2018.
68. National Institute for Health and Care Excellence Decision Support Unit. NICE DSU TECHNICAL SUPPORT DOCUMENT 3: HETEROGENEITY: SUBGROUPS, META-REGRESSION, BIAS AND BIAS-ADJUSTMENT. . 2012. Available from: http://nicedsu.org.uk/wp-content/uploads/2016/03/TSD3-Heterogeneity.final-report.08.05.12.pdf.
69. National Institute for Health and Care Excellence. Nivolumab in combination with ipilimumab for treating advanced melanoma [TA400]. 2016. Available at: [Accessed 25/06/2021].
70. National Institute for Health and Care Excellence. Nivolumab for previously treated advanced renal cell carcinoma [TA417]. 2016. Available at: [Accessed 25/06/2021].
71. National Institute for Health and Care Excellence. Nivolumab for treating advanced (unresectable or metastatic) melanoma [TA384]. 2016. Available at: [Accessed 25/06/2021].
72. National Institute for Health and Care Excellence. Nivolumab for previously treated unresectable advanced or recurrent oesophageal cancer [TA707]. 2021. Available at: [Accessed 25/06/2021].

73. Chen S, Zhou K, Yang L, et al. Racial Differences in Esophageal Squamous Cell Carcinoma: Incidence and Molecular Features. BioMed Research International. 2017;2017:1204082.
74. Muro K, Lordick F, Tsushima T, et al. Pan-Asian adapted ESMO Clinical Practice Guidelines for the management of patients with metastatic oesophageal cancer: a JSMO-ESMO initiative endorsed by CSCO, KSMO, MOS, SSO and TOS. Ann Oncol. 2019;30(1):34-43.
75. National Institute for Health and Care Excellence. Nivolumab for adjuvant treatment of resected oesophageal or gastro-oesophageal junction cancer [TA746]. 2021. Available at: https://www.nice.org.uk/guidance/TA746 [Accessed 07/03/2022].
76. Shyamalee J, Baxter MA, Walsh S, et al. 1720P Real-world outcomes for advanced oesophageal squamous cell carcinoma with supportive care alone. Annals of Oncology. 2021;32:S1194.
77. Cheng S, Qureshi M, Pullenayegum E, et al. Do patients with reduced or excellent performance status derive the same clinical benefit from novel systemic cancer therapies? A systematic review and meta-analysis. ESMO Open. 2017;2(4):e000225.
78. Bristol-Myers Squibb. Primary Clinical Study Report for Study CA209648: A Randomized Phase 3 Study of Nivolumab plus Ipilimumab or Nivolumab Combined with Fluorouracil plus Cisplatin versus Fluorouracil plus Cisplatin in Subjects with Unresectable Advanced, Recurrent or Metastatic Previously Untreated Esophageal Squamous Cell Carcinoma. 2021.
79. Ford HER, Marshall A, Bridgewater JA, et al. Docetaxel versus active symptom control for refractory oesophagogastric adenocarcinoma (COUGAR-02): an open-label, phase 3 randomised controlled trial. The Lancet Oncology. 2014;15(1):78-86.
80. Davidson M, Cafferkey C, Goode EF, et al. Survival in Advanced Esophagogastric Adenocarcinoma Improves With Use of Multiple Lines of Therapy: Results From an Analysis of More Than 500 Patients. Clinical Colorectal Cancer. 2018;17(3):223-30.
81. Davidson M, Cafferkey C, Goode EF, et al. Survival in Advanced Esophagogastric Adenocarcinoma Improves With Use of Multiple Lines of Therapy: Results From an Analysis of More Than 500 Patients. Clin Colorectal Cancer. 2018;17(3):223-30.
82. Kamposioras K, Ntellas P, Nikolaou M, et al. Immunotherapy Efficacy in the Initial Lines of Treatment in Advanced Upper Gastrointestinal Malignancies: A Systematic Review of the Literature. JNCI cancer spectrum. 2021;5(6):pkab088.
83. Sun JM, Shen L, Shah MA, et al. Pembrolizumab plus chemotherapy versus chemotherapy alone for first-line treatment of advanced oesophageal cancer (KEYNOTE-590): a randomised, placebo-controlled, phase 3 study. Lancet. 2021;398(10302):759-71.
84. Janjigian YY, Shitara K, Moehler M, et al. First-line nivolumab plus chemotherapy versus chemotherapy alone for advanced gastric, gastro-oesophageal junction, and oesophageal adenocarcinoma (CheckMate 649): a randomised, open-label, phase 3 trial. The Lancet. 2021;398(10294):27-40.
85. Shitara K, Van Cutsem E, Bang Y-J, et al. Efficacy and Safety of Pembrolizumab or Pembrolizumab Plus Chemotherapy vs Chemotherapy Alone for Patients With First-line, Advanced Gastric Cancer: The KEYNOTE-062 Phase 3 Randomized Clinical Trial. JAMA Oncology. 2020;6(10):1571-80.
86. National Institute for Health and Care Excellence. Pembrolizumab with platinum-based chemotherapy for untreated advanced oesophageal cancer [ID3741]. Committee Papers. 2021. Available from: https://www.nice.org.uk/guidance/ta737/documents/committee-papers [accessed 07/03/2022].
87. National Institute for Health and Care Excellence. Nivolumab with platinum- and fluoropyrimidine-based chemotherapy for untreated HER2-negative advanced gastric, gastro-oesophageal junction or oesophageal adenocarcinoma [ID1465]. 2022.

Available from: https://www.nice.org.uk/guidance/indevelopment/gid-ta10352 [accessed 07/03/2022].
88. Agilent. Interpretation Manual - Gastric or Gastroesophageal Junction Adenocarcinoma. PD-L1 IHC 22C3 pharmDx is FDA-approved for in vitro diagnostic use. 2019. Available at: https://www.agilent.com/cs/library/usermanuals/public/29219_pd-l1-ihc-22C3-pharmdx-gastric-interpretation-manual_us.pdf [Accessed 06/01/2021].
89. Kulangara K, Hanks DA, Waldroup S, et al. Development of the combined positive score (CPS) for the evaluation of PD-L1 in solid tumors with the immunohistochemistry assay PD-L1 IHC 22C3 pharmDx. Journal of Clinical Oncology. 2017;35(15_suppl):e14589-e.
90. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal 2013 - Process and methods [PMG9]. 2013. Available from: https://www.nice.org.uk/process/pmg9/resources/guide-to-the-methods-of-technology-appraisal-2013-pdf-2007975843781 [accessed 20 Aug 2020].
91. National Institute for Health and Care Excellence Decision Support Unit. Technical Support Document 14: Survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data. 2016. Available at: http://www.nicedsu.org.uk/NICE%20DSU%20TSD%20Survival%20analysis.updated%20March%202013.v2.pdf [Accessed 23/04/2021].
92. National Institute for Health and Care Excellence Decision Support Unit. Technical Support Document 18: Methods for Population-Adjusted Indirect Comparisons in Submissions to NICE. 2016. Available at: http://www.nicedsu.org.uk/Populationadjusted-ICs-TSD(3026862).htm. [Accessed January 2021].
93. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal 2013: The reference case. 2013. Available from: https://www.nice.org.uk/process/pmg9/chapter/the-reference-case [accessed 10/08/2021].
94. Regan MM, Mantia CM, Werner L, et al. Treatment-free survival over extended follow-up of patients with advanced melanoma treated with immune checkpoint inhibitors in CheckMate 067. Journal for ImmunoTherapy of Cancer. 2021;9(11):e003743.
95. Chau Iea. Nivolumab plus ipilimumab or nivolumab plus chemotherapy versus chemotherapy as first-line treatment for advanced esophageal squamous cell carcinoma: first results of the CheckMate 648 study. ASCO annual meeting, Abstract number LBA4001. 2021.
96. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal (PMG9) Section 5.6 Discounting. 2013. Available from: https://www.nice.org.uk/process/pmg9/chapter/the-reference-case#discounting [accessed 23/04/2021].
97. Bagust A, Beale S. Survival analysis and extrapolation modeling of time-to-event clinical trial data for economic evaluation: an alternative approach. Med Decis Making. 2014;34(3):343-51.
98. National Institute for Health and Care Excellence Decision Support Unit. Technical Support Document 14: Survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data. 2013. Available at: http://nicedsu.org.uk/wp-content/uploads/2016/03/NICE-DSU-TSD-Survival-analysis.updated-March-2013.v2.pdf [Accessed September 2020].
99. Excellence NIfHaC. Technology appraisal guidance [TA428]. Pembrolizumab for treating PD-L1-positive non-small-cell lung cancer after chemotherapy. 2017. Available from: https://www.nice.org.uk/guidance/ta428 [accessed 27/07/2017].
100. National Institute for Health and Care Excellence. Technology appraisal guidance [TA531]. Pembrolizumab for untreated PD-L1-positive metastatic non-small-cell lung cancer. 2018. Available from: https://www.nice.org.uk/guidance/TA531.

101. Office for National Statistics. National life tables: UK 2018-2020. 2021. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/bulletins/nationallifetablesunitedkingdom/2018to2020 [Accessed
102. National Institute for Health and Care Excellence. Oesophago-gastric cancer: assessment and management in adults. NICE guideline [NG83]. 2018. Available at: https://www.nice.org.uk/guidance/ng83 [Accessed 11 September 2018].
103. National Institute for Health and Care Excellence. Oesophageal and gastric cancer overview. NICE pathways. 2018. Available at: https://pathways.nice.org.uk/pathways/oesophageal-and-gastric-cancer#path=view%3A/pathways/oesophageal-and-gastric-cancer/oesophageal-and-gastric-cancer-overview.xml&content=view-index [Accessed 12 September 2018].
104. National Institute for Health and Care Excellence. Oesophago-gastric cancer. Quality standard [QS176]. 2018; Updated 28 May 2019. Available from: https://www.nice.org.uk/guidance/QS176.
105. clinicaltrials.gov. First-line Esophageal Carcinoma Study With Chemo vs. Chemo Plus Pembrolizumab (MK-3475-590/KEYNOTE-590) (NCT03189719). 2022. Available at: https://clinicaltrials.gov/ct2/show/NCT03189719 [Accessed 21/01/2022].
106. Dolan P. Modeling valuations for EuroQol health states. Med Care. 1997;35(11):1095-108.
107. Bristol Myers Squibb. Utility analysis from CheckMate 648 [data on file]. 2021.
108. Nafees B, Stafford M, Gavriel S, et al. Health state utilities for non small cell lung cancer. Health Qual Life Outcomes. 2008;6:84.
109. National Institute for Health and Care Excellence. Nivolumab for previously treated non-squamous non-small-cell lung cancer. 2017. Available at: https://www.nice.org.uk/guidance/ta484 [Accessed 21/01/2022].
110. National Institute for Health and Care Excellence. Durvalumab for treating locally advanced unresectable non-small-cell lung cancer after platinum-based chemoradiation [TA578]. 2019. Available at: https://www.nice.org.uk/guidance/ta578/history [Accessed 21/01/2022].
111. NHS England. NHS reference costs 2019-2020. 2021. Available at: https://www.england.nhs.uk/national-cost-collection [Accessed 17/01/2022].
112. Mosteller RD. Simplified calculation of body-surface area. N Engl J Med. 1987;317(17):1098.
113. British National Formulary. Nivolumab. 2022. Available at: https://bnf.nice.org.uk/medicinal-forms/nivolumab.html [Accessed 21/01/2022].
114. GOV.UK. Drugs and pharmaceutical electronic market information tool (eMIT). 2021. Available at: https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit [Accessed 21/01/2022].
115. Department of Health. NHS reference costs 2015 to 2016. 2016. Available from: https://www.gov.uk/government/publications/nhs-reference-costs-2015-to-2016.".
116. Department of Health and Social Care. Drugs and pharmaceutical electronic market information tool (eMIT). 2020. Available from: https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit.
117. British National Formulary. Nivolumab. 2020. Available from: https://bnf.nice.org.uk/drug/nivolumab.html#indicationsAndDoses.
118. clinicaltrials.gov. Study of Nivolumab in Unresectable Advanced or Recurrent Esophageal Cancer (NCT02569242). 2022. Available at: https://www.clinicaltrials.gov/ct2/show/NCT02569242 [Accessed 21/01/2022].
119. National Institute for Health and Care Excellence. Nivolumab for previously treated unresectable advanced or recurrent oesophageal cancer [TA707]. 2021. Available at: https://www.nice.org.uk/guidance/ta707 [Accessed 15/02/2022].

120. Georghiou T, Bardsley M. Exploring the cost of care at the end of life. 2014. Available at: https://www.nuffieldtrust.org.uk/research/exploring-the-cost-of-care-at-the-end-of-life [Accessed 21/01/2022].
121. Chauhan R, Lee D. Adrenal Insufficiency: Burden of Disease and Cost of Illness. Value in Health. 2013;16(7):A436.
122. Copley-Merriman C, Stevinson K, Liu FX, et al. Direct costs associated with adverse events of systemic therapies for advanced melanoma: Systematic literature review. Medicine (Baltimore). 2018;97(31):e11736.
123. National Institute for Health and Care Excellence Decision Support Unit. Technical Support Document 14: Survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data. 2013. Available from: http://www.nicedsu.org.uk/NICE%20DSU%20TSD%20Survival%20analysis.updated%20March%202013.v2.pdf [accessed 07/06/2016].
124. Gibson E, Koblbauer I, Begum N, et al. Modelling the Survival Outcomes of Immuno-Oncology Drugs in Economic Evaluations: A Systematic Approach to Data Analysis and Extrapolation. Pharmacoeconomics. 2017;35(12):1257-70.
125. Ouwens M, Mukhopadhyay P, Zhang Y, et al. Estimating Lifetime Benefits Associated with Immuno-Oncology Therapies: Challenges and Approaches for Overall Survival Extrapolations. Pharmacoeconomics. 2019;37(9):1129-38.
126. Quinn C, Garrison LP, Pownell AK, et al. Current challenges for assessing the long-term clinical benefit of cancer immunotherapy: a multi-stakeholder perspective. Journal for ImmunoTherapy of Cancer. 2020;8(2):e000648.
127. Bullement A, Latimer NR, Bell Gorrod H. Survival Extrapolation in Cancer Immunotherapy: A Validation-Based Case Study. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2019;22(3):276-83.
128. Paly VF, Kurt M, Zhang L, et al. Heterogeneity in Survival with Immuno Oncologic Therapy and Its Implications for Survival Extrapolations: A Case Study in Untreated Metastatic Melanoma [data on file]. Virtual ISPOR Europe 2020. 16-19 November 2020.
129. Klijn SL, Fenwick E, Kroep S, et al. What Did Time Tell Us? A Comparison and Retrospective Validation of Different Survival Extrapolation Methods for Immuno-Oncologic Therapy in Advanced or Metastatic Renal Cell Carcinoma [data on file]. submitted.
130. National Institute for Health and Care Excellence. Technology appraisal guidance [TA737]. Pembrolizumab with platinum- and fluoropyrimidine-based chemotherapy for untreated advanced oesophageal and gastro-oesophageal junction cancer. 2021. Available from: https://www.nice.org.uk/guidance/ta737 [accessed 05/01/2022].

Appendices

In line with the user guide for company evidence submission template, appendices start at C, because document A is the submission summary and document B is the main submission.

Appendix number	Appendix Title	Location
C	Nivolumab draft SmPC NB: A version of the European public assessment report or scientific discussion is not yet available	Provided as a separate document
D	Checklist of confidential information	Provided as a separate document
E	Identification, selection and synthesis of clinical evidence: systematic literature review report	Provided as a separate document
F	Subgroup analysis	Provided in the main body of the report
	F.1: CheckMate 648 Clinical Study Report	Provided as a separate document
	F.2: CheckMate 648 Clinical Study protocol	Provided as a separate document
G	Adverse reactions	Provided in the main body of the report
H	Published cost-effectiveness studies: systematic literature review	Provided as a separate document
I	Health-related quality-of-life studies: systematic literature review	Provided as a separate document
J	Cost and healthcare resource identification:	Provided within Appendix H
K	Clinical outcomes and disaggregated results from the model	Provided in the main body of the report
L	Indirect treatment comparison report	Provided as a separate document
M	M.1: Economic model user guide	Provided as a separate document
	M.2: Economic model technical report	Provided as a separate document
N	Survival analysis report	Provided as a separate document
O	Utility analysis report	Provided as a separate document

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Nivolumab in combination with fluoropyrimidine- and platinum-based combination chemotherapy for the first-line treatment of adult patients with unresectable advanced, recurrent, or metastatic oesophageal squamous cell carcinoma with tumour cell PD-L1 expression ≥1% [ID2712]

Clarification questions

May 2022

File name	Version	Contains confidential information	Date
		Yes/no	

Notes for company

Highlighting in the template

Square brackets and grey highlighting are used in this template to indicate text that should be replaced with your own text or deleted. These are set up as form fields, so to replace the prompt text in [grey highlighting] with your own text, click anywhere within the highlighted text and type. Your text will overwrite the highlighted section.

To delete grey highlighted text, click anywhere within the text and press DELETE.

Section A: Clarification on effectiveness data

Literature searches

A1. Please clarify which conference resources were searched in the following sections of the company submission (CS):

- **Appendix E (p.8) states that Northern Light was searched, however the searches presented on pp.52-54 are from Embase.com**
- **Appendix H (p.8) states that Embase.com was searched, however the search presented on p.29 is from Northern Light.**
- **Appendix I (p.9) states that Embase.com was searched, however the search presented on pp.24-25 is from Northern Light.**

Response: Appendix E (clinical SLR report) incorrectly stated that Northern Lights was searched for conference abstracts; EMBASE was used for the search of conference abstracts in the clinical SLR.

Both Appendix H (economic SLR report) and Appendix I (HRQoL SLR report) incorrectly state that EMBASE was used for the search of conference abstracts in the economic and HRQoL SLRs; the Northern Lights database was used for the search of conference abstracts for both of these SLRs.

In summary, the search strategies as presented were correct; however, the text of the SLR reports was incorrect.

A2. Please provide the strategies used for the ClinicalTrials.gov search in Appendix E.

Response: The text regarding the search strategy used for the clinicaltrials.gov search was amended to:

> *Searches for clinicaltrials.gov were conducted by screening all trial entries identified when searching 'esophageal cancer' as the condition or disease and limiting to entries with results. Note that when the term 'esophageal cancer' is searched in clinicaltrials.gov, this includes oesophageal neoplasm and oesophageal carcinoma. Identified entries were screened according to the eligibility criteria presented in Table 1.*

Table 1. Eligibility criteria for the global systematic literature review and focused OSCC-only population

Criteria	Global inclusion criteria	OSCC-specific inclusion criteria
Population	Adult patients with previously untreated unresectable advanced or metastatic oesophageal cancer	Adult patients with previously untreated advanced or metastatic oesophageal squamous cell carcinoma
Interventions	Eligible interventions include any of the following treatments as monotherapy or in combination with one or more of the other treatments: • Nivolumab • Anthracycline • Capecitabine • Carboplatin • Cetuximab • Cisplatin • Docetaxel • Epirubicin • Exaliplatin • Fluorouracil • Ipilimumab • Irinotecan • Leucovorin • Oxaliplatin • Paclitaxel • Pembrolizumab	

Comparators	Eligible comparators include the following: • Placebo, observation, physician's choice, or best supportive care • Any intervention of interest • Any treatment that facilitates an indirect comparison	Comparators were ineligible if they evaluated radiotherapy as monotherapy or in combination with other eligible interventions
Outcomes	Studies must report at least one of the following outcomes*: • Overall survival • Disease-free survival, progression-free survival, or time-to-progression • Distant metastatic-free survival In addition, the following outcomes will also be extracted, where reported: • Any adverse events • All cause grade 3/4 adverse events Overall discontinuations	
Study design	Randomized controlled trials only	
Language	Only studies published in English	

A3. The numbers provided in the PRISMA flow diagram on p.28 of the CS (Document B, Section B.2.1)/p.12 of Appendix E do not match the numbers of records retrieved as documented in Appendix A of the clinical systematic literature review (SLR) report in Appendix E. For example, the flow diagram shows Embase as 398, however 4,001 records were found by the original Embase search, and 484 by the update. The other database results are similarly incorrectly documented.

Please provide a full PRISMA flow diagram showing all results for both original and update searches before and after deduplication.

Response: The reason for this discrepancy is that for each of the SLR updates (of which all three are updates), duplicates were removed in two stages in the first instance: when identified studies were combined across databases as well as manually during the update. The error in the PRISMA flow diagram is that the total number of 'Records identified through database searching' only reflects results after the first de-duplication. The original clinical SLR and economic SLR PRISMA flow diagrams have now been amended to accurately reflect all duplicates that were removed prior to abstract screening (Figure 1 and Figure 3). In addition, the PRISMA

flow diagrams for the updated SLRs have also been provided below (Figure 2 and Figure 4). Note that the same method was used for the HRQoL SLR: however, this did not result in a discrepancy in the PRISMA diagram.

Figure 1. PRISMA flow diagram for original clinical systematic literature review (original search executed January 2021)

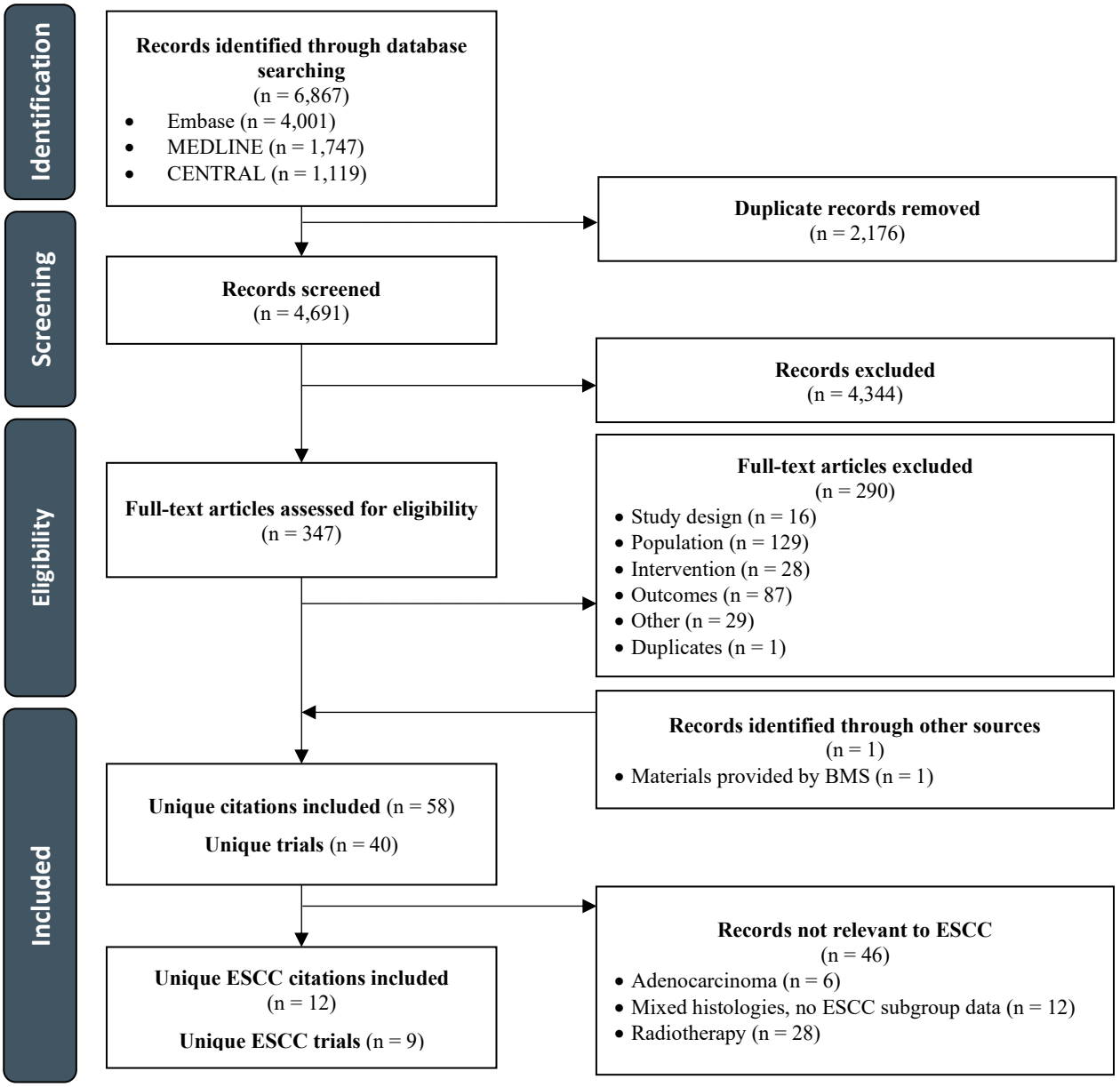
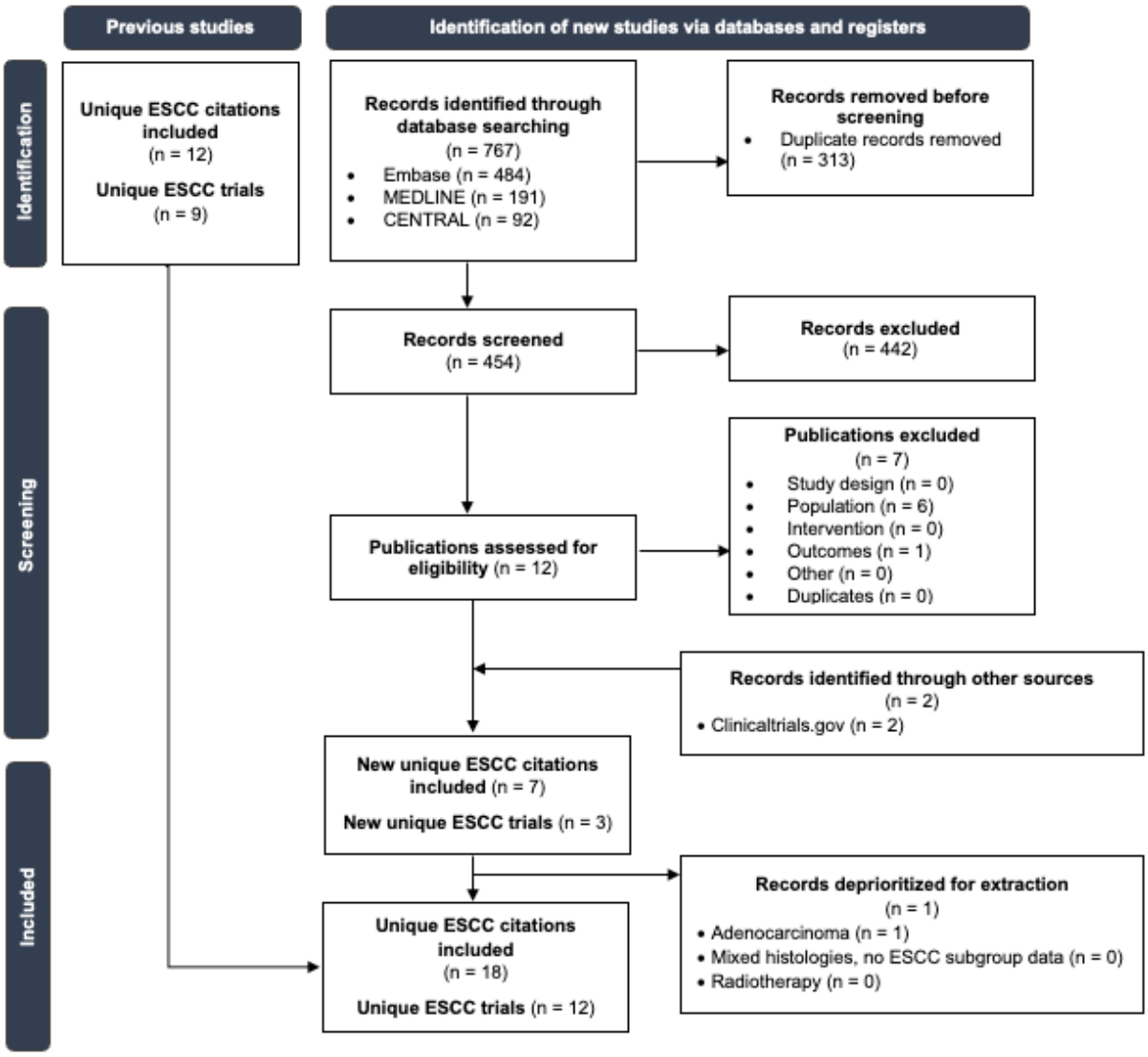


Figure 2. PRISMA flow diagram for updated clinical systematic literature review (search executed October 2021)



SCC – squamous cell carcinoma

Figure 3. PRISMA flow diagram for original economic systematic literature review (original search executed April 2021)

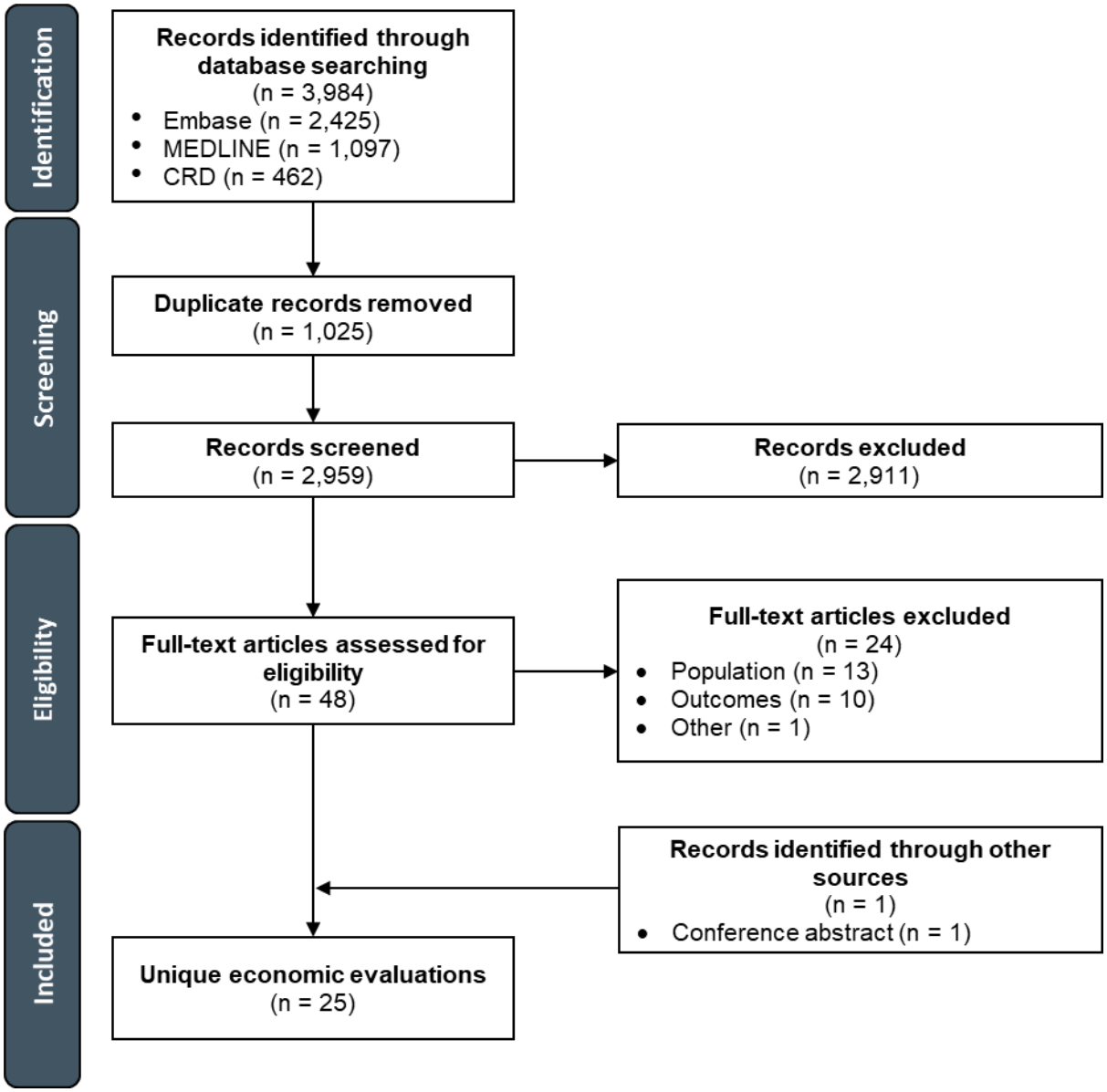
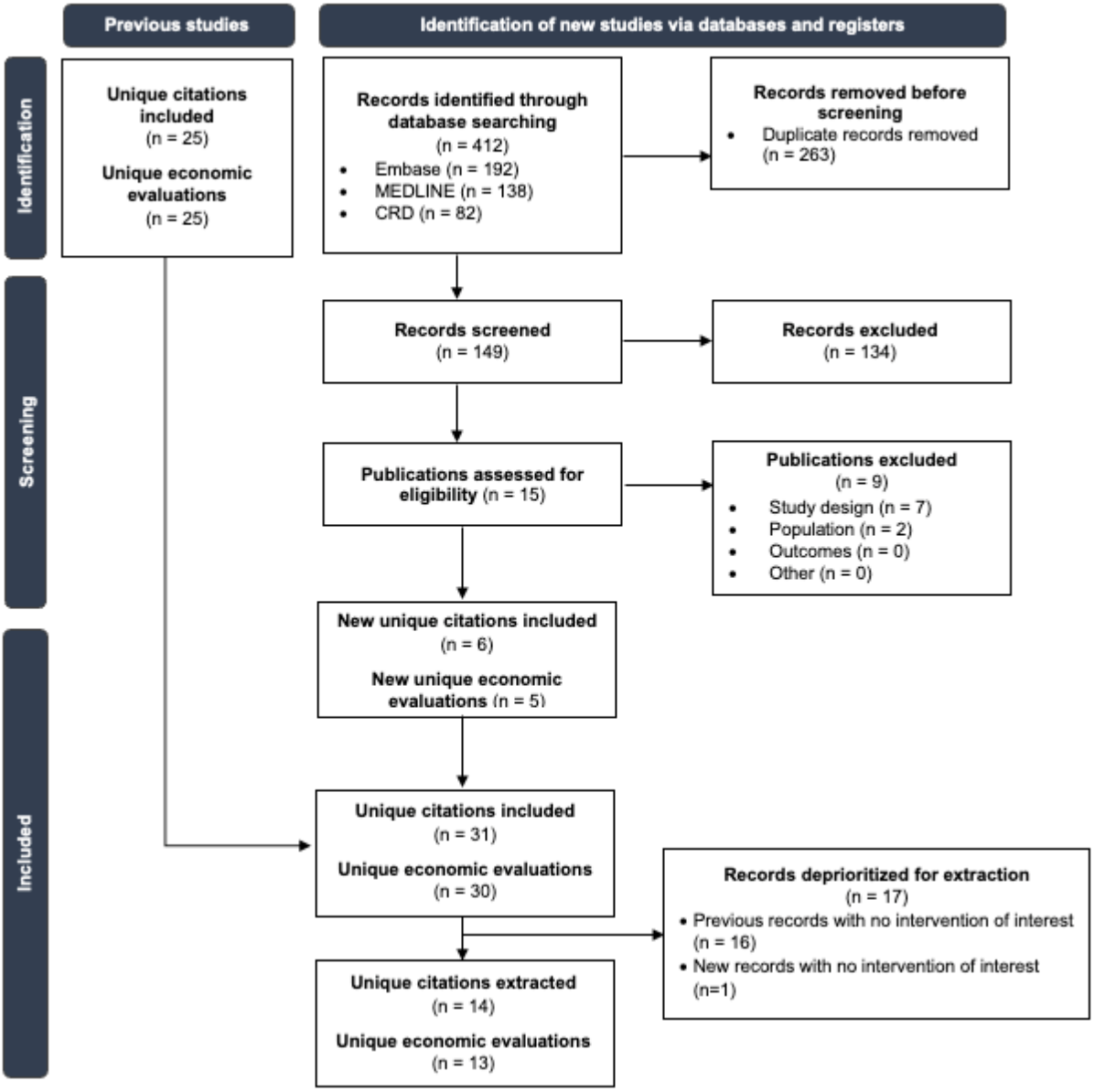


Figure 4. PRISMA flow diagram for updated economic systematic literature review (search executed October 2021)



A4. Please confirm if the CENTRAL searches were conducted via the Cochrane Library (as stated on p.26 of the CS [Document B, Section B.2.1]), or via EBM Reviews (as stated on p.49-50 of Appendix E). If the latter, please state the host (e.g., Ovid/EBSCO, etc.) used.

Response: CENTRAL searches were conducted via EBM reviews through Ovid for the clinical SLR and via Cochrane Library for the HRQoL SLR.

A5. Please confirm the date that the cost-effectiveness searches were conducted on. Appendix H (p.7) states that they were conducted on 14 January

2021, however Appendix A of the economic SLR report in Appendix H (p.26-29) provides a search date of 28 April 2021.

Response: For the economic SLR, the original searches were conducted on 28 April 2021 for all databases while the updated searches were conducted on 22 October 2021 for all databases. One exception is the updated conference searches, which were conducted after completion of the relevant conferences (September 30, 2021).

Decision problem

A6. Priority question: Chemotherapy is a comparator for the programmed death ligand 1 (PD-L1) ≥1% tumour cells (TC) population. However, pembrolizumab plus chemotherapy is a comparator for the ≥10 combined positive score (CPS) population. Therefore, the population of PD-L1 ≥1% TC and ≥10 CPS population is relevant to both comparators, but the populations of PD-L1 ≥1% TC and less than 10 CPS is relevant to only chemotherapy.

a) **Please clarify precisely which population according to PD-L1 TC and CPS status are relevant to which comparator treatments.**

Response: Table 2 clarifies which population is relevant to which comparator. Since KEYNOTE-590 does not provide data for patients with CPS <10%, this subpopulation is not a comparator for NIVO+CHEMO. More importantly, the label for nivolumab in this indication includes patients with tumour cell PD-L1 expression ≥1% and not CPS ≥ 10%.

Table 2. Population according to PD-L1 TC and CPS status

Population	CHEMO	PEMBRO+CHEMO
PD-L1 ≥1% and CPS ≥10%	Yes	Yes
PD-L1 ≥1% and CPS <10%	Yes	No

b) **Please conduct separate effectiveness analyses of nivolumab plus chemotherapy versus each comparator using data from the relevant population, including:**

   i. **versus chemotherapy and pembrolizumab plus chemotherapy in the PD-L1 ≥1% TC and ≥10% CPS population.**

Response: BMS would like to confirm that the indication, as adopted by the CHMP and aligned with its licence,[1] includes patients with tumour cell PD-L1 expression ≥1% and not CPS ≥ 10%. As noted by the EMA, nivolumab plus chemotherapy demonstrated superiority over chemotherapy alone in terms of OS, PFS and ORR in

the first-line treatment of patients with unresectable advanced, recurrent or metastatic OSCC with tumour cells expressing PD-L1 ≥1%.[1]

By contrast, pembrolizumab is indicated in combination with platinum and fluoropyrimidine-based chemotherapy for the first-line treatment of locally advanced unresectable or metastatic carcinoma of the oesophagus or HER-2 negative gastroesophageal junction adenocarcinoma in adults whose tumours express PD-L1 with a CPS ≥10%. Aligned with this, KEYNOTE-590 provided evidence for patients with PD-L1 CPS ≥10%, stratified by tumour location and histology for some outcomes. KEYNOTE-590 did not provide evidence for OSCC patients with PD-L1 TC ≥1%, limiting comparisons that can be drawn in the population of primary interest for nivolumab. Further, limited data from KEYNOTE-590 were available for OSCC patients with PD-L1 CPS ≥10%.

For this reason, the ITC was conducted in the CPS ≥ 10% population, facilitating a comparison between nivolumab and pembrolizumab in the population of interest for pembrolizumab, despite the fact that the label populations for pembrolizumab and nivolumab differ.

Data from the sub-population of patients with both PD-L1 ≥ 1% and PD-L1 CPS ≥10% is available from CheckMate 648. However, this data is not provided in KEYNOTE-590, precluding a comparison between nivolumab and pembrolizumab in this specific subgroup.

It is acknowledged that there is significant overlap between the PD-L1 TC ≥ 1% and PD-L1 CPS ≥10% populations. Of the [redacted] patients in the NIVO-CHEMO arm with tumour cell PD-L1 ≥1% and available CPS data, [redacted] also had PD-L1 CPS ≥10%. However, [redacted] patients who had PD-L1 CPS ≥10 in the ITT population did not have PD-L1 TC ≥ 1%, as outlined in Table 3, demonstrating that not all patients with PD-L1 TC ≥ 1% have PD-L1 CPS ≥10%. As a result, a subgroup of patients from CheckMate 648 with PD-L1 TC ≥1% and PD-L1 CPS ≥10% would enrich the population with those patients likely to have best response by both PD-L1 assessment criteria. As this enriched subgroup would be compared against the published KEYNOTE-590 PD-L1 CPS ≥10%, which would include patients with PD-L1 TC < 1%, this would be a biased comparison.

Table 3. CheckMate 648 frequency of PD-L1 by CPS status

	NIVO-CHEMO		CHEMO	
	ITT	Tumour cell PD-L1 ≥1%	ITT	Tumour cell PD-L1 ≥1%
ITT	[redacted]	[redacted]	[redacted]	[redacted]
ITT with CPS score	[redacted]	[redacted]	[redacted]	[redacted]
CPS ≥5%	[redacted]	[redacted]	[redacted]	[redacted]
CPS ≥10%	[redacted]	[redacted]	[redacted]	[redacted]

Clinically, this subpopulation does not exist as medical decisions which drug to use would be based on CPS or TC.

Furthermore, CheckMate 648 is powered to detect significant differences based on the subgroup of patients with TC PD-L1 expression ≥1% (i.e. the primary trial endpoint) and not the subgroup with CPS ≥10%. Reducing the sample size further using subgroup analysis limits the conclusions that can be drawn from the analysis.

Results for the analysis in the population of patients with PD-L1 TC ≥1% and PD-L1 CPS ≥10% are presented in response to Question A19b, but should be considered with caution since only an overlap analysis could be conducted as PD-L1 TC ≥1% data was not available from KEYNOTE-590 (see Table 6 of updated Appendix L - NMA report).

Of additional note, the survival analysis presented in the company submission is based on the data of OSCC patients with TC PD-L1 ≥1%. While an analysis is presented for OSCC patients with PD-L1 CPS ≥10%, there remains some uncertainty, as data was not available for this subgroup from KEYNOTE-590. Data for this subpopulation from KEYNOTE-590 was only available for the mixed histology population, limiting the conclusions that can be drawn.

ii. versus chemotherapy in the PD-L1 ≥1% TC and less than 10 CPS population.

Response: As stated in i) the sample size calculation of CheckMate 648 is based on PD-L1 TC ≥1% and not CPS. The study is not powered for an analysis that would include fewer patients as suggested when restricting further to patients with PD-L1 ≥ 1% TC and CPS <10%. Additionally, the HR for OS for the patient population

with CPS <10%, as presented in Table 7.4.2.2-1 of the CSR, is [redacted] and taking the confidence interval into account comparable to PD-L1 <1% TC, [redacted] as stated in Table 7.4.1.1-1, which suggests that an analysis as requested probably will not have a huge impact on the results.

Furthermore, chemotherapy is still standard of care (SOC) regardless of PD-L1 status and not all patients who have a CPS >10% will receive pembrolizumab as it is reimbursed in this indication in October 2021 and is not yet widely used in clinical practice.

A7. Priority question: In Document B, Section B.1.3.5, the CS states the following as the basis for the unmet need that nivolumab is intended to address: "*Pembrolizumab in combination with platinum and fluoropyrimidine-based chemotherapy has recently been recommended for the treatment of untreated, locally advanced unresectable or metastatic oesophageal carcinoma in adults whose tumours express PD-L1 with a combined positive score (CPS) ≥10. However, this indication does not include patients with CPS <10, who are covered by the indication for nivolumab with chemotherapy, which includes patients with PD-L1 CPS ≥1, and therefore, there remains an unmet need for therapeutic options covering this patient population*". Table 28 of the CS (Document B, Section B.2.13.4.3) shows that there is an observed overlap between PD-L1 CPS ≥10 and TC≥1% which is [redacted] in the nivolumab plus chemotherapy arm and [redacted] in the chemotherapy arm (for the intention-to-treat population with available CPS). The remaining proportion is of the patient population for which there is an unmet need for therapeutic options i.e., PD-L1 CPS <10 and TC≥1%.

a) **Please clarify the extent to which each of these methods of determining PD-L1 status will be used in NHS clinical practice**.

Response: PD-L1 testing of tumour tissue has been used as a biomarker of response to antagonist medications. The NHS has a growing network of pathologists who are trained and capable of confirming PD-L1 status to ensure oncologists can target these precision medicines to the patients who will benefit the most. PD-L1 testing can be divided into two method types: tumour cell/tumour proportion scoring method (TC/TPS) and the combined positive score method (CPS). TC and TPS are

very similar. To obtain a TC score you divide the number of PD-L1 stained tumour cells by the total number of viable tumour cells and multiple by 100. For TPS the numerator is the number of PD-L1 positive tumour cells. It is important to stress that the output from both scoring methods in the same. The TC/TPS method was used in CheckMate 648 and is a common method used in other cancer types. The CPS method was developed as a new scoring method and compared with the TPS in patients treated with pembrolizumab in GC/GEJC.[2] CPS is the ratio of the number of all PD-L1 expressing cells (tumour and non-tumour) to the number of all tumour cells.[3]

In our engagement process with pathology services within the NHS, we understand that even though PD-L1 testing is new to the gastrointestinal tumour therapy area, it is increasingly common in other types of cancer. The CPS method is deployed in both adeno and squamous cell carcinoma of the oesophagus where the TPS method is well established in lung, head & neck and bladder cancer therapy areas.

Assessment of PD-L1 status is becoming routine in clinical practice in the UK for patients with OSCC, following the introduction of pembrolizumab for patients with PD-L1 CPS ≥10% ([redacted] DBL) with locally advanced unresectable or metastatic oesophageal adenocarcinoma (OAC), oesophageal squamous cell carcinoma (OSCC) or gastro-oesophageal junction cancer (GOJC). The summary of product characteristics for pembrolizumab states that PD-L1 status should be assessed via a well-validated and robust methodology to minimise false negative or false positive determination.

In the UK, the Dako PD-L1 IHC 22C3 pharmDx test is currently used in clinical practice for lung cancer[4] and was used during the KEYNOTE-590 trial to determine PD-L1 status.[5] The FDA does not require PD-L1 testing, whereas EMA requires a PD-L1 assay but that is primarily based on its interpretation of the data from KEYNOTE-590 and CheckMate 648 and the scoring method used within those trials.

The Dako PD-L1 IHC 22C3 pharmDx assay can assess PD-L1 expression using both CPS or TPS and so there is no additional burden to patients as only one sample is required. Additionally, the costs difference of these two methods is negligible and would have a marginal impact on the ICER or BIM results. It is likely

that it will become routine practice to assess both TPS and CPS during the same test to determine which OSCC patients are suitable for either pembrolizumab or nivolumab treatment. TPS is generally considered to be an easier test to conduct from a pathology perspective.

As part of BMS' commitment to the NHS, we have scheduled trainings being offered to pathologists across the country to ensure they are confident and comfortable in conducting PD-L1 using the TC/TPS method.

b) Please provide evidence as to the proportion of patients in each of these categories, PD-L1 CPS <10 and TC≥1%, and PD-L1 CPS ≥10 and TC≥1%, in NHS clinical practice.

Response: As pembrolizumab was only recommended for the treatment of OSCC in October 2021,[6] testing for PD-L1 is only just starting to be implemented in routine clinical practice in the UK. Therefore, there is limited evidence available in a UK or European population to determine the relative proportions of patients in each of the categories (PD-L1 CPS <10%, TC ≥1% and CPS ≥10% and TC ≥1%. A calculation from CheckMate 648 cannot be performed as the number of UK patients included in the trial is limited.

Table 4Table 4 presents the proportions of patients with either PD-L1 CPS <10% or TC ≥1% reported in the KEYNOTE-590 and CheckMate 648 trials, respectively.

Table 4. The proportion of PD-L1 expression reported in KEYNOTE-590 and CheckMate 648

	KEYNOTE-590[5]	**CheckMate 648[7]**
PD-L1 CPS <10%	46.3%	Can be calculated from PLD
PD-L1 CPS ≥10%	51.1%	Can be calculated from PLD
TC ≥1%	Not reported	49.2%
PD-L1 CPS ≥10 and TC ≥1%	Not reported	Can be calculated from PLD

Furthermore, the relative proportions of patients with different PD-L1 expression would only influence the results of the budget impact analyses in this submission, where it is used to determine how many patients would receive nivolumab with chemotherapy.

Systematic literature review (SLR)

A8. The clinical effectiveness SLR limits inclusion to studies published in English (Table 1 in Appendix E). The same language restriction is applied to the SLRs of cost-effectiveness studies (Table 1 in Appendix H) and health-related quality of life (HRQoL) studies (Table 1 in Appendix I).

For each SLR, please describe the volume of relevant literature omitted because of this restriction and discuss the impact on study retrieval and estimates of clinical effectiveness, cost-effectiveness and HRQoL estimates.

Response: Table 5 below shows the differences that can be attributed to restricting to English publications for the clinical SLR (taken from the EMBASE and MEDLINE searches executed January 2021). Note that numbers could not be compared for the update as additional restrictions were applied after the language restriction. The table does not screen these additional studies; however, given the low number of hits and that pembrolizumab was key comparator of interest for which data was identified through the KEYNOTE-590 study, it is unlikely that relevant studies were missed. For both the economic and HRQoL SLRs, as the focus of the studies is for submission to the UK and thus, the UK population being the most relevant, it is unlikely that relevant studies were missed in the restriction to English publications.

Table 5. Comparison of studies identified with search strategies restricting and not restricting to English studies for the clinical systematic literature review

Database	Search; restricting to English publications	Search prior to restriction to English publications	Difference in studies identified
Clinical systematic literature review			
EMBASE	4,001	4,236	235
MEDLINE	1,747	2,021	274

A9. Appendix C within the clinical SLR report in Appendix E lists 46 publications excluded at the full-text screening stage. This number is

discrepant with both the PRISMA flow diagram (Figure 1 of Appendix E) and the narrative summary of study retrieval in Appendix E (Study selection, p11).

Please clarify the number of studies (and publications) excluded at the full-text screening stage during the January 2021 original search and the October 2021 update.

Response: In total, 297 studies were excluded during full-text screening, 290 in the original SLR conducted in January 2021 and 7 in the updated SLR performed in October 2021.

A10. Figure 1 (PRISMA flow diagram) of Appendix H (cost-effectiveness SLR) indicates that 26 records were excluded at the full-text screening stage. Similarly, Figure 1 of Appendix I (HRQoL SLR) suggests that 29 records were excluded at the full-text screening stage. Neither appendix shows further details of these excluded studies. Please provide tabulation of these studies, including full bibliographic details against the reason for exclusion per individual study.

Response: Full details of the studies referred to, as well as the reasons for exclusion, are provided in Table 6 and Table 7 for the economic and HRQoL SLRs, respectively. Note that the 26 publications excluded at the full text screening stage are a combination of the 9 studies excluded according to the pre-specified eligibility criteria and an additional 17 studies that were deprioritized as they did not evaluate interventions of interest (detailed in Table 7).

Table 6. Full text publications excluded and deprioritized in the economic systematic literature review

Author	Year	Title	Exclusion reason
Full text publications excluded in original systematic literature review (n=24)			
Birch	1998	A cost-benefit comparison of self-expanding metal stents and atkinson tubes for the palliation of obstructing esophageal tumors	Outcomes: No outcomes of interest
Boshier	2018	Endoscopic therapy and surveillance versus esophagectomy for early esophageal adenocarci-noma: A review of early outcomes and cost analysis	Outcomes: No modelled direct/indirect costs
Broe	2013	Evaluating the clinical efficacy and cost effectiveness of direct access endoscopy	Population: Not EC population
Buyukkaramikli	2017	Ramucirumab for treating advanced gastric cancer or gastro-oesophageal junction adenocarcinoma previously treated with chemotherapy: An evidence review group perspective of a nice single technology appraisal	Population: Not EC population

Chevrou-Severac	2012	Cost-effectiveness analysis of immunonutrition for upper gastrointestinal cancer patients undergoing surgery in British hospitals	Population: Not EC population
Chu	2017	Surgical versus endoscopic management of t1 esophageal adenocarcinoma: A modelling decision analysis incorporating age and comorbidity	Population: Not EC population
Dimofte	2004	Cost-effectiveness of endoscopically placed stents in the palliation of locally advanced esophageal carcinoma	Outcomes: No outcomes of interest
Goenka	2011	The role of surgical resection following primary chemoradiation therapy in esophageal squamous cell carcinoma: A decision analysis	Outcomes: No modelled direct/indirect costs
Harewood	2001	Cost minimization analysis of alternative strategies for initial staging of esophageal cancer	Other: Editorial
Heise	2001	Expense and benefit of neoadjuvant treatment in squamous cell carcinoma of the esophagus	Outcomes: No modelled costs
Horgan	2011	Capecitabine or infusional 5-fluorouracil for gastroesophageal cancer: A cost-consequence analysis	Population: Not EC population
Hulscher	2002	Extended transthoracic resection compared with limited transhiatal resection for adenocarcinoma of the esophagus	Population: Not EC population
Hung	2014	Generalized cost-effectiveness analysis for care of major cancers and other major illnesses in Taiwan	Population: Not EC population
Luo	2018	Economic evaluation of ramucirumab as second-line chemotherapy for previously treated advanced gastric or gastro-oesophageal junction adenocarcinoma in china	Population: Not EC population
Meads	2016	The cost effectiveness of docetaxel and active symptom control versus active symptom control alone for refractory oesophagogastric adenocarcinoma: Economic analysis of the cougar-02 trial	Population: Not EC population
Nicholson	1999	The cost effectiveness of metal oesophageal stenting in malignant disease compared with conventional therapy	Outcomes: No modelled direct/indirect costs
Ostvar	2018	Cost-effectiveness of immune checkpoint inhibition in metastatic gastric and esophageal tumors	Population: Not EC population
Ruhstaller	2017	Intergroup phase iii trial of neo-adjuvant chemotherapy, followed by chemoradiation and surgery with and without cetuximab in locally advanced esophageal carcinoma: First results from the health economic analysis of sakk 75/08 trial	Outcomes: No modelled direct/indirect costs
Sihvo	2002	Inoperable adenocarcinoma of the oesophagogastric junction: A comparative clinical study of laser coagulation versus self-expanding metallic stents with special reference to cost analysis	Outcomes: No outcomes of interest
Virik	2020	Economic evaluation of trifluridine/tipiracil (tt) versus nivolumab (n) in patients with advanced/metastatic gastric cancer (gc) or gastro-esophageal junction cancer (gejc) in Canada	Population: Not EC population
Ward	2021	Global costs, health benefits, and economic benefits of scaling up treatment and imaging modalities for survival of 11 cancers: A simulation-based analysis	Population: Not EC population
Wheat	2015	Relative cost per life-year gained of treatment with curative intent for t3nxm0 upper gastrointestinal cancer	Population: Not EC population
Xinopoulos	2004	Natural course of inoperable esophageal cancer treated with metallic expandable stents: Quality of life and cost-effectiveness analysis	Outcomes: No modelled direct/indirect costs
Xinopoulos	2005	Palliative treatment of advanced esophageal cancer with metal-covered expandable stents. A cost-effectiveness and quality of life study	Outcomes: No modelled direct/indirect costs
Full text publications excluded in original systematic literature review (n=9)			
Dunn	2021	Transition from esophagectomy to endoscopic therapy for early esophageal cancer	Outcomes: Not modeled direct/indirect costs

Gheorghe	2021	Economic impact of avoidable cancer deaths caused by diagnostic delay during the COVID-19 pandemic: A national population-based modelling study in England, UK	Population: Not EC population
Gregory	2021	Cost Effectiveness Analysis of Partially Covered, Fully Covered, and Sutured Fully Covered Self Expanding Metal Stents for Palliation of Malignant Esophageal Dysphagia	Population: Not EC population
Huang	2021	Therapeutic Effect and Cost-Benefit Analysis of Three Different Nutritional Schemes for Esophageal Cancer Patients in the Early Post-operative Period	Outcomes: Not modeled direct/indirect costs
Ichimura	2021	Cost-effectiveness of primary prophylaxis of febrile neutropenia with pegfilgrastim in docetaxel, cisplatin and 5-fluorouracil therapy for esophageal cancer	Population: Not EC population
Lai	2021	QALYs and medical costs saved from prevention of a cancer: Analysis of nation-wide real-world data of Taiwan with lifetime horizon	Outcomes: Not modeled direct/indirect costs
Navaratnam	2021	Real-World Assessment of the Treatment Patterns, Healthcare Resource Use, and Survival Outcomes Associated with Non-Resectable Advanced Esophageal Cancers in South Korea	Outcomes: Not modeled direct/indirect costs
Yang	2021	Cancer death and potential years of life lost in Feicheng City, China: Trends from 2013 to 2018	Outcomes: Not modeled direct/indirect costs
Ontario Health	2021	Proton beam therapy for cancer in children and adults: A health technology assessment	Population: Not EC population
Full text publications deprioritized for extraction in original systematic literature review (n=16)			
Adamson	2021	D4: Can we afford pet-ct for oesophageal cancer management?	Intervention assessed not of interest
Azmi	2017	Cost-effectiveness analysis of thoracoscopic versus open esophagectomy for esophageal cancer: A population-based study	Intervention assessed not of interest
Chao	2020	Surgical vs endoscopic management of t1 esophageal adenocarcinoma: A modelling decision analysis	Intervention assessed not of interest
Chu	2018	Cost-effectiveness of palliation of unresectable esophageal cancer	Intervention assessed not of interest
da Silveira	2008	Cost-effectiveness in the management of patients with oesophageal cancer	Intervention assessed not of interest
Farndon	1998	Adjuvant statin therapy for esophageal adenocarcinoma: A cost-utility analysis	Intervention assessed not of interest
Fong Soe Khoie	2018	Modelling the cost-effectiveness of strategies for treating esophageal adenocarcinoma and high-grade dysplasia	Intervention assessed not of interest
Gordon	2012	Ct or eus for the initial staging of esophageal cancer? A cost minimization analysis	Intervention assessed not of interest
Hadzijahic	2000	A cost analysis of endoscopic ultrasound in the evaluation of esophageal cancer	Intervention assessed not of interest
Harewood	2002	Cost-effectiveness of minimally invasive versus open esophagectomy for esophageal cancer	Intervention assessed not of interest
Lee	2013	Cost-effectiveness of minimally invasive esophagectomy for esophageal squamous cell carcinoma	Intervention assessed not of interest
Liu	2018	Economic analysis of esophageal stenting for management of malignant dysphagia	Intervention assessed not of interest
Rao	2009	A pragmatic randomised controlled trial of the cost-effectiveness of palliative therapies for patients with inoperable oesophageal cancer	Intervention assessed not of interest
Shenfine	2005	Photodynamic therapy for the treatment of early esophageal cancer: A systematic review and economic evaluation	Intervention assessed not of interest
Health Technology & Policy Unit	2009	An analysis of multiple staging management strategies for carcinoma of the esophagus: Computed tomography, endoscopic ultrasound, positron emission tomography, and thoracoscopy/laparoscopy	Intervention assessed not of interest

Wallace	2002	Cost-effectiveness of proton therapy for esophageal cancer	Intervention assessed not of interest
Studies deprioritized for extraction in original systematic literature review (n=1)			
Adamson	2021	Palliative radiotherapy combined with stent insertion to reduce recurrent dysphagia in oesophageal cancer patients: the ROCS RCT	Intervention assessed not of interest

Table 7. Full text publications excluded in the health-related quality of life systematic review

Author	Year	Title	Exclusion reason
Full text publications excluded in original systematic literature review (n=29)			
Not specified	2018	Effects of different radiotherapy regimens combined with tp regimen on the survival and quality of life of patients with middle and advanced esophageal cancer	Other: Chinese paper
Amdal	2017	Improved treatment decisions in patients with esophageal cancer	Outcomes: EORTC QLQ-C30 and QLQ-OG25 reported
Amdal	2013	Palliative brachytherapy with or without primary stent placement in patients with oesophageal cancer, a randomised phase iii trial	Outcomes: EORTC QLQ-C30 and QLQ-OG25 reported
Bergquist	2012	Combined stent insertion and single high-dose brachytherapy in patients with advanced esophageal cancer--results of a prospective safety study	Outcomes: EORTC QLQ-C30 and QLQ-QES18 reported
Boshier	2020	Assessment of health related quality of life and digestive symptoms in long-term, disease free survivors after esophagectomy	Population: Not unresectable
Chang	2016	Quality-of-life measures as predictors of post-esophagectomy survival of patients with esophageal cancer	Outcomes: EORTC QLQ-C30 and QLQ-QES18 reported
Chen	2021	Effects of rehabilitation program on quality of life, sleep, rest-activity rhythms, anxiety, and depression of patients with esophageal cancer: A pilot randomized controlled trial	Population: Mixed stage I-IV population; no subgroup data for advanced EC patients
Chu	2018	Surgical vs endoscopic management of t1 esophageal adenocarcinoma: A modelling decision analysis	Population: Not unresectable
Didden	2018	Fully vs. Partially covered selfexpandable metal stent for palliation of malignant esophageal strictures: A randomized trial (the copac study)	Outcomes: EORTC QLQ-C30 and QLQ-QES18 reported
Ding	2019	The health utility level and quality of life of patients with precancerous lesion or cancer of the digestive tract in Beijing: A cross-sectional survey	Population: Unable to identify cancer stage due to insufficient data
Ding	2021	Using a Chinese time trade-off approach to explore the health utility level and quality of life of cancer patients in urban China: A multicentre cross-sectional study	Population: Mixed stage I-IV population; no subgroup data for advanced EC patients
Dutton	2014	Gefitinib for oesophageal cancer progressing after chemotherapy (cog): A phase 3, multicentre, double-blind, placebo-controlled randomised trial	Outcomes: EORTC QLQ-C30 and QLQ-OG25 reported
EUCTR	2013	Intravenous iron in the management of anaemia in palliative oesophageal and gastric cancer	Population: Mixed cancer populations; no subgroup data for advanced EC patients
Forootan	2019	Efficacy of chemoradiotherapy on health-related quality of life in esophageal cancer patients with dysphagia	Outcomes: RTOG and EORTC QoL used
Kuo	2018	Cancer impact, complementary/alternative medicine beliefs, and quality of life in cancer patients	Population: Mixed cancer populations; no subgroup data for advanced EC patients
Lee	2015	Capecitabine in combination with either cisplatin or weekly paclitaxel as a first-line treatment for	Outcomes: EORTC QLQ-QES18

		metastatic esophageal squamous cell carcinoma: A randomized phase ii study	
Li and Liu	2016	Clinical efficacy and safety of nedaplatin combined with paclitaxel liposome in treatment of advanced esophageal cancer	Other: Chinese paper
Liu	2012	Long-term outcome of irradiation with or without chemotherapy for esophageal squamous cell carcinoma: A final report on a prospective trial	Outcomes: QoL assessed by KPS, Diet, Cough and Hemoptysis
Maishman	2021	A phase ii study of biodegradable stents plus palliative radiotherapy in oesophageal cancer	Outcomes: EORTC QLQ-C30 and QLQ-OG25 reported
NCT	2013	Rocs (radiotherapy after oesophageal cancer stenting) study	Population: Mixed stage I-IV population; no subgroup data for advanced EC patients
Penniment	2020	An economic evaluation of palliation of dysphagia in esophageal cancer: Analysis of the trog 03.01/ncic es.2 phase iii study in advanced esophageal cancer in patients treated with radiotherapy versus chemoradiotherapy	Outcomes: No QoL outcomes reported
Qiu	2020	Effect of whole-course nutrition management on patients with esophageal cancer undergoing concurrent chemoradiotherapy: A randomized control trial	Outcomes: EORTC QLQ-C30 reported
Rahouma	2017	New chemotherapy regimen; does it really work for esophageal cancer adenocarcinoma?	Population: Not advanced or metastatic EC
Russell	2013	Cancer of oesophagus or gastricus - new assessment of technology of endosonography (cognate): Report of pragmatic randomised trial	Population: Not unresectable advanced or metastatic esophageal cancer
Seike	2011	A new candidate supporting drug, rikkunshito, for the qol in advanced esophageal cancer patients with chemotherapy using docetaxel/5-fu/cddp	Outcomes: QoL was assessed by customised form, evaluating sleep, mood, volition, activity of daily living
Su	2019	Health-related quality of life among cancer survivors in rural China	Population: Patients were cancer survivors, stage unclear
Trudel	2016	Longitudinal evaluation of trial outcome index scores in patients with esophageal cancer	Population: Not advanced or metastatic EC
Tsuda	2011	Prospective study of definitive chemoradiotherapy with s-1 and nedaplatin in patients with stage ii/iii (non-t4) esophageal cancer	Outcomes: No QoL outcomes reported
Yamashita	2014	Longitudinal assessments of quality of life and late toxicities before and after definitive chemoradiation for esophageal cancer	Population: Mixed stage I-IV population; no subgroup data for advanced EC patients
Full text publications excluded in original systematic literature review (n=27)			
Adamson	2021	Palliative radiotherapy combined with stent insertion to reduce recurrent dysphagia in oesophageal cancer patients: The rocs RCT	Outcomes: EORTC QLQ-C30 reported
Ajani	2021	O-15 Randomized, phase 3 study of second-line tislelizumab vs chemotherapy in advanced or metastatic esophageal squamous cell carcinoma (RATIONALE 302) in the overall population and Europe/North America subgroup	Outcomes: No QoL nor utilities reported
Chen	2021	Application effect of continuous nursing in patients with advanced esophageal cancer after esophageal stent implantation	Outcomes: EORTC QLQ-C30 reported
Elimova	2021	Health-related quality of life (HRQOL) in patients (pts) with advanced gastric cancer/gastroesophageal junction cancer (GC/GEJC) or esophageal adenocarcinoma (EAC): Results of nivolumab plus chemotherapy (NIVO+chemo) versus chemo from CheckMate 649	Population: Low EC population

Elliott	2021	PL11.02 ensure: An international multicentre study exploring whether surveillance after esophageal cancer surgery impacts oncological and quality of life outcomes	Population: Not unresectable advanced or metastatic esophageal cancer
Elliott	2021	Intensive surveillance after curative intent surgery for esophageal cancer: Initial results of the ensure study	Population: Not unresectable advanced or metastatic esophageal cancer
Gheorghe	2021	Economic impact of avoidable cancer deaths caused by diagnostic delay during the COVID-19 pandemic: A national population-based modelling study in England, UK	Study design: No QoL nor utilities
Gourzoulidis	2021	Cost-effectiveness of trifluridine/tipiracil as a third-line treatment of metastatic gastric cancer, including adenocarcinoma of the gastrohesophageal junction, among patients previously treated in Greece	Outcomes: No result data available on QoL
Hall	2021	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	Population: Low EC population (~30-50%)
Hauser	2021	PCN9 A Real-World Assessment of the Treatment Patterns, Health Resource Utilization and Outcomes Associated with Esophageal Cancer in the United States	Outcomes: No result data available on QoL
Ho	2021	PCN46 Economic Evaluation of Esophageal Cancer Screening Among Patients with Oral Cavity Cancer in Taiwan	Outcomes: No result data available on QoL
Ichimura	2021	Cost-effectiveness of primary prophylaxis of febrile neutropenia with pegfilgrastim in docetaxel, cisplatin and 5-fluorouracil therapy for esophageal cancer	Outcomes: No result data available on QoL; utilities derived from Sugimoto 2018, which was not EC specific
Jezerskyte	2021	Postoperative Complications and Long-Term Quality of Life After Multimodality Treatment for Esophageal Cancer: An Analysis of the Prospective Observational Cohort Study of Esophageal-Gastric Cancer Patients (POCOP)	Outcomes: EORTC QLQ-C30 reported
Klevebro	2021	Severe Dumping Symptoms Are Uncommon Following Transthoracic Esophagectomy But Significantly Decrease Health-Related Quality of Life in Long-Term, Disease-Free Survivors	Population: Not unresectable advanced or metastatic esophageal cancer
Lai	2021	QALYs and medical costs saved from prevention of a cancer: Analysis of nation-wide real-world data of Taiwan with lifetime horizon	Outcomes: No result data available on QoL
Lin	2021	Cost-Effectiveness Analysis of Camrelizumab Immunotherapy versus Docetaxel or Irinotecan Chemotherapy as Second-Line Therapy for Advanced or Metastatic Esophageal Squamous Cell Carcinoma	Outcomes: EORTC QLQ-C30 reported
Orsini	2021	PCN43 Cost-Effectiveness of Nivolumab for the Treatment of Advanced Esophageal Squamous Cell Carcinoma Refractory or Intolerant to Previous Chemotherapy: A United-States Payer Perspective	Outcomes: No result data available on QoL
Qu	2021	1409P Cost-effectiveness of pembrolizumab plus platinum and fluoropyrimidine-based chemotherapy as first-line treatment of advanced esophageal cancer in the United States	Outcomes: No result data available on QoL
Silvers	2021	PCN67 Estimating Cost-Effectiveness of Pembrolizumab in Advanced Esophageal Cancer Based on the Keynote-181 Trial	Outcomes: No result data available on QoL
Soni	2021	Nivolumab in gastric/gastroesophageal junction cancer: Real-world data from UK Early Access to Medicines Scheme	Population: >30% gastric population and no subgroup QoL data
Thomas	2021	Stronger therapeutic alliance is associated with better quality of life among patients with advanced cancer	Outcomes: FACIT-Pal reported

Van Cutsem	2021	Health-related quality of life in advanced gastric/gastroesophageal junction cancer with second-line pembrolizumab in KEYNOTE-061	Population: > 65% gastric population and no subgroup QoL data
Van Cutsem	2021	Quality of life with first-line pembrolizumab for PD-L1-positive advanced gastric/gastroesophageal junction adenocarcinoma: results from the randomised phase III KEYNOTE-062 study	Population: > 65% gastric population and no subgroup QoL data
Vimolratana	2021	Two-Year Quality of Life Outcomes After Robotic-Assisted Minimally Invasive and Open Esophagectomy	Population: Not unresectable advanced or metastatic esophageal cancer
Yang	2021	Cost-effectiveness analysis of camrelizumab in the second-line treatment for advanced or metastatic esophageal squamous cell carcinoma in China	Outcomes: No result data available on QoL; utilities derived from Al-Batran 2016, which was not EC specific
Wang	2021	PCN73 Cost-Effectiveness of Pembrolizumab Versus Chemotherapy for Advanced Esophageal Squamous Cell Carcinoma (ESCC) As the Second-LINE Treatment Among the Chinese Population	Outcomes: No result data available on QoL
Zeng	2021	Rehabilitation Nursing Intervention Can Improve Dysphagia and Quality of Life of Patients Undergoing Radiotherapy for Esophageal Cancer	Outcomes: EORTC QLQ-C30 reported

A11. Document B, Section B.2.13.4.2.1 reports a comparison of the CheckMate 648 trial with UK studies, specifically to Cougar-2 and the Royal Marsden retrospective review.

Please provide details on how these two studies were identified and included in the CS since they were not part of the SLR results executed by the company.

Response: The objective of the SLR was to identify studies in OSCC. Neither Cougar-2 nor the Royal Marsden retrospective review included OSCC patients. Since these studies were conducted in the UK and included patients with gastro-oesophageal adenocarcinoma, which is a patient population similar to OSCC, they were used to validate the CheckMate 648 study results. Both studies were identified through a targeted search.

A12. Document B, Section B.2.13.4.2.3 reports on the KEYNOTE-062 trial. This section is describing studies that were identified during the company's SLR, but KEYNOTE-062 was not included in the results as presented in Appendix E.

Please clarify why this study is included in this part of the CS and provide details on how it was identified and included in the CS.

Response: Similar to the Cougar-2 and the Royal Marsden retrospective review, KEYNOTE-062 was identified in a targeted search and used to validate the baseline

characteristics and results from CheckMate 648 since patients with gastric or gastro-oesophageal junction cancer were included in that study. The CS only stated that KEYNOTE-062 was identified in a published SLR,[8] but not within the SLR that is presented in Appendix E.

Trial and data analysis

A13. The company states in Document B, Section B.2.13.4 that "*CheckMate 648 provides survival data that may be considered relatively mature*".

Please discuss how the maturity of the survival data was assessed.

Response: Survival data is typically considered mature where the median point has been reached. Per Appendix N Section 4.2.3, as 25.3% of patients remain alive at the end of follow-up, the CheckMate 648 data can be considered relatively mature. The maturity of the survival follow-up may be compared to that of KEYNOTE-062 trial at the time of appraisal in NICE TA737. In the CPS ≥10% subgroup, OS was 15% and 26% in the standard of care and pembrolizumab + standard of care respectively at 27 months, the last point at which at least 10 surviving patients had survival follow-up beyond (Committee papers, company submission, figure 5[6] The PD-L1 ≥1% subgroup of CheckMate 648, survival at 27 months was 11% and 29% in the chemotherapy and nivolumab + chemotherapy arms respectively, and at month 33 (where 15 patients remained within follow-up on the nivolumab + chemotherapy arm) was 10% and 21% respectively [BMS data on file]. The survival data were thus considered mature relative to contemporary technology assessments.

A14. Document B, Section B.2.13.4 of the CS states that, "*Patients enrolled in the available studies can be considered broadly representative of UK practice, in terms of baseline characteristics…*"

Please discuss the representativeness of the trial population to UK clinical practice and provide supporting documents if needed.

Response: During an advisory board conducted by BMS, UK clinicians believed that the baseline characteristics of patients enrolled in CheckMate 648 were representative of the patients they treat in UK clinical practice.[9] They felt that the eligibility criteria reflected clinical practice in the UK and was aligned with the decision problem. When asked if the baseline characteristics observed in patients randomised in CheckMate 648 were representative of those seen in UK clinical practice, the clinicians agreed that the trial patients were broadly younger than they would expect, but otherwise, the patients were representative of UK OSCC patients. Similarly, the discontinuation data observed during CheckMate 648 was considered

to be aligned to the clinicians' expectations and the data observed during the REAL2 trial.[10]

As presented in Table 23 of the CS, the age, sex and ECOG PS at baseline for patients enrolled in CheckMate 648 is similar to that observed in the Shyamalee et al.[11] and Jaffe et al.[12] studies which were conducted in UK OSCC patient populations.

The clinicians agreed during the advisory board that the high proportion of Asian patients in the CheckMate 648 trial (70%) was not an issue when applying the trial data to a UK population. It was explained that in oesophageal adenocarcinoma the imbalance between Asian and non-Asian patients would be an issue as patients are treated over several different lines of therapy. However, this is not the same in OSCC and so should not be considered an issue. It was confirmed that there was no biological reason to consider the populations to be different.

This is further supported by the data presented in Section B.2.7 where subgroup analysis demonstrated favourable OS for nivolumab with chemotherapy in both Asian and non-Asian populations.

Additionally, the imbalance between Asian and non-Asian patients was not considered to be an issue in the pembrolizumab assessment in OSCC.[6]

A15. Document B, Section B.2.13.4.2.1 of the CS states that "*no studies were identified to assess UK outcomes or baseline characteristics in advanced OSCC patients*". On the other hand, in Document B, Section B.2.13.4.2.2 the CS presents "*Studies identified during this SLR that assessed previously untreated gastro-oesophageal cancer in a UK-relevant population*". In Appendix E where the SLR results are presented in detail, there are two studies (CheckMate 648 and KEYNOTE-590) that include populations from the UK.

Please confirm whether UK populations where available from other sources.

Response: As stated in the company's submission, no studies were identified that assessed UK outcomes or baseline characteristics in advanced OSCC patients since neither KEYNOTE-590 nor CheckMate 648 presented outcomes and baseline

characteristics separately for the UK patients. Nevertheless, the study populations itself can be deemed relevant to the UK population as confirmed for KEYNOTE-590 in TA737, although it needs to be mentioned that KEYNOTE-590 contained a mixed population of OAC, OSCC and GOJC patients.[6]

No other studies were identified that contained information on UK patients with OSCC.

A16. Document B, Section B.2.13.4.2 of the CS compares the results of CHECKMATE 648 with five other studies.

Please provide a tabular presentation of CHECKMATE 648 plus all five comparator studies to facilitate an overall comparison.

Response: The baseline characteristics of patients enrolled in CheckMate 648 and the five comparator studies identified are presented in Table 8.

The median overall survival of patients enrolled in CheckMate 648 and the five comparator studies are presented in Table 9.

Table 8. Baseline characteristics of patients enrolled in CheckMate 648 and the five comparator studies identified

		Oesophageal cancers					Gastro-oesophageal cancers					
		CheckMate 648[13,14] (PD-L1 ≥1%)		Cougar-2[15]		Royal Marsden retrospective review[16]	KEYNOTE-590[5]		CHECKMATE 649[7]		KEYNOTE-062[17]	
		NIVO+ CHEMO	CHEMO	Docetaxel	Active symptom control		PEMBRO+ CHEMO	CHEMO	NIVO+C HEMO	CHEMO	PEMBRO +CHEMO	CHEMO
N		**158**	**157**	**84**	**84**	**511**	**373**	**376**	**789**	**792**	**257**	**250**
Sex, male (%)		125 (79%)	131 (83%)	69 (82%)	67 (80%)	384 (75%)	306 (82%)	319 (85%)	540 (68)	560 (71)	195 (76%)	179 (72%)
Median age (range), years		64 (40–85)	62 (28–81)	65 (28–84)	66 (36–84)	66 (24-90)**	64 (28-94)	62 (27-89)	62 (54-69)	61 (53-68)	64 (25%)	61 (24%)
ECOG status	0	71 (45%)	70 (45%)	24 (28%)	22 (26%)	64 (13%)	149 (40%)	150 (40%)	326 (41)	336 (42)	-	-
	1	87 (55%)	86 (55%)	46 (55%)	50 (60%)	276 (54%)	223 (60%)	225 (60%)	462 (59)	452 (57)	138 (54%)	135 (54%)
	2	0	0	14 (17%)	12 (14%)	87 (17%)	1 (<1%)	1 (<1%)	-	-	-	-
Disease status	Locally advanced or recurrent	[redacted]	[redacted]	11 (13%)	10 (12%)	68 (13%)*	29 (8%)	37 (10%)	32 (4%)	36 (5%)	-	-
	Metastatic disease	[redacted]	[redacted]	73 (87%)	74 (88%)	335 (66)*	344 (92%)	339 (90%)	757 (96)	756 (95)	243 (95%)	235 (94%)
Histology	Adenocarcinoma	-	-	84 (100%)	84 (100%)	511 (100%)	99 (27%)	102 (27%)	-	-	257 (100%)	250 (100%)
	Squamous cell carcinoma	156 (99)	155 (99)	-	-	-	-	-	-	-	-	-

* 21% of patients had relapsed metastatic disease after radical treatment.
** Age at diagnosis, not study baseline
Baseline characteristics and demographics presented for CheckMate 648 patients with tumour cell PD-L1 ≥1%

Table 9. Overall survival of patients enrolled in CheckMate 648 and the five comparator studies identified

	Oesophageal cancers					Gastro-oesophageal cancers					
	CheckMate 648[13,14] (PD-L1 ≥1%)		Cougar-2[15]		Royal Marsden retrospective review[16]	KEYNOTE-590[5]		CHECKMATE 649[7]		KEYNOTE-062[17] (PD-L1 CPS ≥1%)	
	NIVO+ CHEMO	CHEMO	Docetaxel	Active symptom control		PEMBRO+ CHEMO*	CHEMO*	NIVO+ CHEMO	CHEMO	PEMBRO +CHEMO	CHEMO
N	**158**	**157**	**84**	**84**	**511**	**373**	**376**	**789**	**792**	**257**	**250**
Median survival	15.1	9.1	5.2	3.6	11.5	12.6	9.8	14.4	11.1	10.6	11.1
95% CI	11.9-18.6	7.7-10.0	4.1-5.9	3.3-4.4	10.5-12.5	10.2-14.3	8.6-11.1	13.1-16.2	10.0-12.1	7.7-13.8	9.2-12.8

* Results presented for the OSCC subgroup
OS outcomes presented for CheckMate 648 and KEYNOTE-062 patients with tumour cell PD-L1 ≥1%

A17. In Document B, Section B.2.13.4.2 the CS compares the results of CHECKMATE 648 with five other studies. The CS provides data on the patients' PD-L1 CPS ≥10 status for the KEYNOTE-062 and the CHECKMATE 649 trials but not for the other three comparator studies. Please confirm whether PD-L1 status is available for the rest of the studies. If the PD-L1 status is not available, please discuss the comparability to the CHECKMATE 648 trial.

Response: A summary of the data available for the different PD-L1 subgroups in CheckMate 648 and the comparator trials is presented in Table 10. While CheckMate 648, KEYNOTE-590, CheckMate 649 and KEYNOTE-062 presented data for specific PD-L1 patient populations, the Cougar-2 trial and the Royal Marsden review did not.

The Cougar-2 trial and the Royal Marsden review do not consider the use of immunotherapies and therefore, the outcomes that were reported for chemotherapy treatment are comparable to the outcomes reported in the chemotherapy arm of CheckMate 648, as the treatment effect of chemotherapy should not be affected by PD-L1 expression.

Table 10. Data availability for PD-L1 subgroups in CheckMate 648 and the comparator trials

	Oesophageal cancers			Gastro-oesophageal cancers		
	CheckMate 648[13,14]	**Cougar-2**[15]	**Royal Marsden retrospective review**[16]	**KEYNOTE-590**[5]	**CheckMate 649**[7]	**KEYNOTE-062**[17]
PD-L1 ≥1%TC	Yes	No	No	No	No	Yes
PD-L1 ≥5% CPS	No	No	No	No	Yes	No
PD-L1 ≥10% CPS	No	No	No	Yes	No	Yes
PD-L1 <10% CPS	No	No	No	Yes	No	No

* Results presented for the OSCC subgroup
OS outcomes presented for CheckMate 648 and KEYNOTE-062 patients with tumour cell PD-L1 ≥1%

A18. Priority question: The majority of patients included in the CHECKMATE 648 trial were Asian: n = 227 (71% NIVO-CHEMO arm) and n = 227 (70% CHEMO arm) living in Asian countries n = 225 (70% NIVO-CHEMO arm), n = 226 (70% CHEMO arm). The CS states that "*CheckMate 648 included a substantial proportion of non-Asian patients, so that the patient population is more*

***reflective of that observed in UK clinical practice*.” Also, the results of subgroup analyses on median overall survival, as illustrated in Figure 17, are [redacted] for Asian participants and [redacted] for non-Asian participants, with the respective hazard ratios (HRs) being [redacted] and [redacted].**

a) Can the company confirm how many patients, if any, come from the UK?

Response: In CheckMate 648, there were 5 centres in the UK which enrolled [redacted] patients in total. Of these [redacted] patients, [redacted] had PD-L1 >1% expression, and 9 were in the NIVO+CHEMO arm and [redacted] were in the chemotherapy treatment arm.

b) Please provide evidence as to the comparability with NHS clinical practice.

Response: During an advisory board conducted by BMS, clinicians specialising in the treatment of oesophageal cancer in the UK stated that the patients enrolled in CheckMate 648 were representative of the patients they treat in the UK.[9] They stated that the eligibility criteria of the trial reflected UK clinical practice. When shown the baseline characteristics of patients enrolled in the trial, UK clinicians believed that the patients randomised in CheckMate 648 were representative of those seen in UK clinical practice.

It was agreed between the clinicians that the high proportion of Asian patients in the CheckMate 648 trial (70%) was not an issue and would not affect how the trial data aligns with patients from the UK. It was explained that in oesophageal adenocarcinoma, the imbalance between Asian and non-Asian patients would be an issue as these patients are treated over several lines of therapy. However, this is not the same for OSCC, where the treatment paradigms for advanced OSCC are similar in Europe, the US and Asia,[18] and therefore, this should not be considered an issue. It was agreed that there was no biological reason to consider Asian and non-Asian OSCC populations to be different, which is also supported by studies comparing the genetic profiles between Asian and non-Asian OSCC patients and previous NICE appraisals for OSCC.[6,19]

Despite the large proportion of Asian patients, the baseline characteristics of patients enrolled in CheckMate 648 are similar to those enrolled in UK studies of OSCC. The median age of patients enrolled in CheckMate-648 was 64 years, which aligns to the

median age of patients enrolled in similar UK studies of OSCC. For instance, Shyamalee et al. (2021), a real-world evidence study of the outcomes of UK OSCC patients treated with best supportive care, observed a median age of 63 years for patients treated with palliative chemotherapy.[11] Likewise, the western cohort of OSCC patients who initiated first-line therapy included in the non-observational study conducted by Jaffe et al. (2022) had a mean age of 62.9 years.[12]

A comparison of the CheckMate 648 trial population with the patient populations included in the Shyamalee et al.[11] and Jaffe et al.[12] studies is presented in Table 11 (Table 23 of the CS). This demonstrates that the baseline characteristics of trial population of CheckMate 648 are comparable to those of other UK OSCC cohorts and so can be considered broadly representative of those patients seen in UK clinical practice.

Table 11. A comparison of the baseline characteristics of patients in the CheckMate 648 trial with those in the Shyamalee and Jaffe UK studies

Baseline characteristic		CheckMate 648[7]		Shyamalee[11]	Jaffe[12]
		NIVO-CHEMO	CHEMO		
Cohort size		321	324	219	1,049
Age	Median (range), years	64 (40-90)	64 (26-81)	63	-
	Mean (SD), years	-	-	-	62.9 (10.6)
Sex	Male	78.8%	84.9%	48%	82.7%
ECOG PS, n (%)	0	150 (46.7)	154 (47.5)	6 (27)	-
	1	171 (53.3)	170 (52.5)	9 (41)	-
CHEMO: chemotherapy; ECOG PS: Eastern cooperative oncology group performance scale; NIVO-CHEMO: nivolumab with chemotherapy;					

c) Please discuss the implications in terms of effectiveness of any discrepancy between CheckMate 648 and NHS clinical practice.

Response: Despite a large proportion of the patients in CheckMate 648 being Asian (70%), this should not affect the comparability of the effectiveness results to UK clinical practice. This was confirmed during an advisory board meeting conducted by BMS by clinicians specialising in the treatment of OSCC in the UK, who explained that there was no biological reason to consider Asian and non-Asian OSCC patients to be different and to respond differently to treatment.[9]

When comparing the baseline characteristics of patients enrolled in CheckMate 648 with other studies conducted in OSCC patients from the UK (Shyamalee et al.[11] and Jaffe et al.[12] Table 11), it was noted that slightly fewer patients with ECOG status of 0 were enrolled in CheckMate 648 compared to the Shyamalee study. Clinical trials commonly specify performance scores as an inclusion criterion, typically based on either ECOG or Karnofsky scale. This leads to limited evidence of net clinical benefit for patients with certain performance scores, typically those with worse scores. This absence of evidence contributes to a reluctance to provide certain treatments to patients of reduced performance score. However, there is limited evidence to suggest different outcomes between patients with different performance score. Therefore, this difference in performance scores should not impact the comparability of the effectiveness of nivolumab with chemotherapy observed in CheckMate 648 with UK clinical practice.

This is further supported by subgroup analysis from the CheckMate 648 trial, which demonstrated a survival benefit with nivolumab and chemotherapy in both the Asian and non-Asian subgroups (Section B.2.7, reproduced in Figure 5).

Figure 5. Forest plot of subgroup analysis, for age, sex, race, region, ECOG, weight and disease stage at initial diagnosis, on overall survival for all randomised patients treated with NIVO-CHEMO or CHEMO

Source: CheckMate 648 [redacted] Summary data[20]

Response: A previous NICE submission (TA746)[21] considering nivolumab monotherapy for unresectable, advanced oesophageal cancer after standard chemotherapy has failed, conducted an SLR to determine the differences in patient characteristics and survival outcomes between Asian and Western population.[21] The SLR found that OS was comparable between Asian and Western populations with OSCC (median OS: 7.5 versus 7.4 months; mean one-year survival was 21.1% in Asian and 27.9% in Western patients).[22]

When considering the common chemotherapy arms from CheckMate 648 and studies conducted in the UK, in both oesophageal and gastro-oesophageal cancers, the median OS of patients from CheckMate 648 treated with chemotherapy was similar to that observed in similar patients treated in with chemotherapy in the UK (Table 9, 9.1 months compared to 5.2, 11.5, 9.8, 11.1 and 11.1, respectively).[5,7,11,12,15,17] This demonstrates that the survival outcomes of the CheckMate 648 trial are comparable to studies conducted in the UK, and therefore, the inclusion of a large proportion of Asian patients does not affect the generalisability of the trial results to UK clinical practice in terms of effectiveness.

It needs to be noted that, although the median OS for patients receiving chemotherapy in CheckMate 648 is comparable to other UK studies[7,11,12,15,17], the long-term results of the chemotherapy arm in CheckMate differ from medical expectations. It is clinically implausible that approximately 10 % of patients receiving chemotherapy are alive at 36 months as observed in the clinical study.

Network meta-analysis (NMA)

A19. Priority question: In Document B, Section B.2.9.2.3 the CS states that a network meta-analysis (NMA) for CheckMate 648 patients that had both PD-L1 ≥10% (CPS) as well as PD-L1 ≥1% (TC) was deemed inappropriate "*…because patients who were both PD-L1 TC ≥1 and PD-L1 CPS ≥10 were removed, therefore enriching the patients with PD-L1 from CheckMate 648.*" However, in Appendix L it is stated: "*Additionally, an ITC for CheckMate 648 patients that had both PD-L1 ≥10% (CPS) as well as PD-L1 ≥1% (TC) was also conducted; survival models and ITC results are presented in Appendix J.*" (p.20)

a) **Please explain how patients with both PD-L1 TC ≥1 and PD-L1 CPS ≥10 have been removed, what is meant by 'enriching the patients with PD-L1' and how this diminishes the validity of an NMA in this population.**

Response: The text in question contains a typing error. The sentence should read as follows: "… because patients who were both PD-L1 TC ≥ 1% and PD-L1 CPS <10% were removed, therefore enriching the patients with PD-L1 from CheckMate 648."

This population is 'enriched' because in order to isolate this group you have to remove patients who are CPS ≥10% and TC PD-L1 <1% in the KEYNOTE-590 trial and data is not available for the TC PD-L1<1% population (see response to A6). BMS knows that results are less efficacious TC PD-L1 < % patients, therefore this can be viewed as causing some potential bias. This comparison differs from the base case which (although imperfect) compares on CPS ≥10% patients in both trials. However, this does not align with the license for nivolumab.

b) **Please present the results of an indirect treatment comparison (ITC) for CheckMate 648 patients that had both PD-L1 ≥10% (CPS) as well as PD-L1 ≥1% (TC).**

Response: Results for the population with PD-L1 ≥1% and CPS ≥10% are presented in Appendix I of the Appendix L NMA report and in Table 12 and Table 13 below.

Table 12. Tabular results of the ITC of PFS; CPS ≥10% and TC ≥1%

Model family	Model	Hazard ratios (95% credible intervals) for pembrolizumab + chemotherapy versus chemotherapy					
		3 months	*6 months*	*9 months*	*12 months*	*24 months*	*36 months*
Standard parametric	Gamma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Generalized gamma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Log normal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Weibull	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Spline hazard	1-knot	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	2-knot	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Spline odds	1-knot	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	2-knot	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Spline normal	1-knot	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	2-knot	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table 13. Tabular results of the ITC of OS; CPS ≥10% and TC ≥1%

Model family	Model	Hazard ratios (95% credible intervals) for pembrolizumab + chemotherapy versus chemotherapy					
		3 months	*6 months*	*9 months*	*12 months*	*24 months*	*36 months*
Standard parametric	Gamma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Generalized gamma						
	Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Log-logistic						
	Log normal						
	Weibull						
Spline hazard	1-knot						
	2-knot						
Spline odds	1-knot						
	2-knot						
Spline normal	1-knot						
	2-knot						

A20. Priority question: In Document B, Section B.2.9.1.1, the company states that the "*assessment of comparability was based on data from the all-comers population from KEYNOTE-590, as no baseline characteristics were reported for the OSCC PD-L1 CPS ≥10 population. The assessment found that CheckMate 648 and KEYNOTE-590 were sufficiently similar, in terms of study design and patient baseline characteristics (Appendix L), to conduct an indirect comparison. More detail is provided in Section B.2.9.4.*"

The Evidence Review Group (ERG) was unable to locate the assessment of study design and baseline characteristics in Appendix L. Please provide further information. Detailed comparison of the datasets is required to support their similarity which underpins the validity of the executed network meta-analysis.

Response: The NMA report (Appendix L) was updated since the initial submission and is provided within this response.

A feasibility assessment was conducted (Section 5.1 of Appendix L), trial, treatment and patient characteristics compared. It was concluded that under the following assumptions the NMA could be conducted:

- KEYNOTE-590 did not allow patients with prior treatment experience while CheckMate 648 allowed patients with prior treatment provided it was completed more than six months prior to trial enrolment resulting in nearly 80% of patients with prior treatment experience in CheckMate 648. It is assumed that these differences do not act as treatment effect modifiers.
- Cycle lengths differed between CheckMate 648 and KEYNOTE-590 resulting in longer planned treatment duration for CheckMate 648 (though median treatment durations were 3.4 and 5.8 months for the chemotherapy arms, respectively). It is assumed that these differences do not act as treatment effect modifiers.
- Though no differences in treatment effect modifiers could be identified through analysis of baseline patient characteristics, it is assumed this is also true for the OSCC patients with PD-L1 ≥10% (CPS) within each trial.

A21. Priority question: Appendix L of the CS states that both JAGS and R were used for the ITC analyses. Please report all the packages used and provide the code that was used to run the ITCs as well as the datasets that the analyses were based on, derived from both studies (CheckMate 648 and KEYNOTE-590).

Response: The packages and code are attached to this response.

A22. Priority question: In Document B, Section B.2.9.1.1 the company states that "*This approach preserves randomisation and allows treatment effects to vary over time.*" (p. 70) Please explain how this is the case.

Response: The method used to include pembrolizumab + chemotherapy was estimation of the relative treatment effect between pembrolizumab + chemotherapy and chemotherapy (as assessed in KEYNOTE-590) and then this relative treatment effect was applied to the chemotherapy arm from CheckMate 648. The assumptions of this analysis are outlined in the response to A27, below. Under the assumptions listed there, we conclude that there are no differences in treatment effect modifiers between the chemotherapy arms in the CheckMate 648 trial and KEYNOTE-590 and with that, the outcome would be similar to that of a complete network meta-analysis, where randomization is preserved using the principle of transitivity. However, it must be acknowledged that this analysis was only conducted as estimation of the RTE within the KEYNOTE-590 trial and thus, does not require consideration of preserving randomisation for this analysis.

A23. Priority question: In Document B, Section B.2.9.2.1 the company states that "*Fractional polynomial models were not considered for the NMA, due to similar fit and extrapolation to standard parametric and spline models.*" (p. 72) Please provide further explanation as to why such models were deemed inappropriate.

Response: One point of clarification is that although a multivariate normal network meta-analysis framework was utilized for the estimation of relative treatment effects (RTEs), this analysis does not incorporate the entire evidence base and thus, should be referred to as the estimation of RTEs for pembrolizumab + chemotherapy versus chemotherapy (as assessed in KEYNOTE-590), not a network meta-analysis.

With regards to the fractional polynomial models, though these models were considered for survival modelling of independent treatment arms, for the comparison of pembrolizumab + chemotherapy versus chemotherapy in patients with PD-L1 CPS ≥10, these models were not used to estimate RTEs as fractional polynomials were not recommended to be considered for the base case of the cost-effectiveness model. The reason for this is, as illustrated in Figure 6 and Figure 7 below for both OS and PFS (IA), the fractional polynomial models evaluated either provided nearly identical extrapolations to the standard parametric and spline-based models evaluated or generated models with long tails that were considered clinically implausible for chemotherapy survival. Furthermore, this rational was presented to clinical and health economic experts at an advisory board where the health economic experts confirmed the similarity of the spline and standard parametric models to the fractional polynomial models and agreed incorporating fractional polynomial models into the cost-effectiveness model by means of estimation of RTEs would likely not lead to different estimates.

Figure 6. Models fit to CheckMate 648 chemotherapy OS in patients with PD-L1≥10% (combined positive score): fractional polynomial (A), standard parametric (B), spline hazard (C), spline normal (D), spline odds (E)

A)

B)

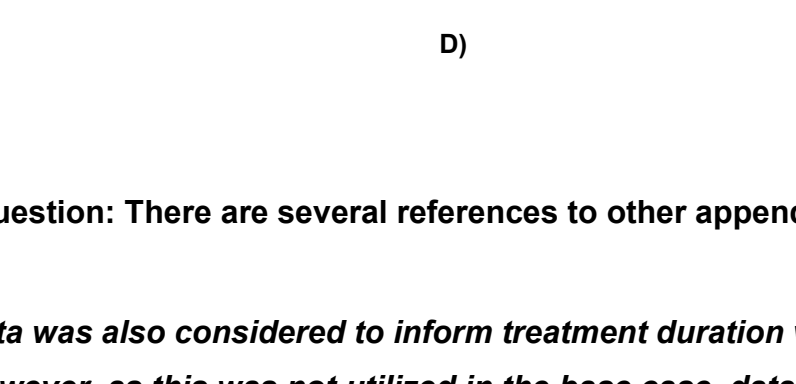
C)

D)

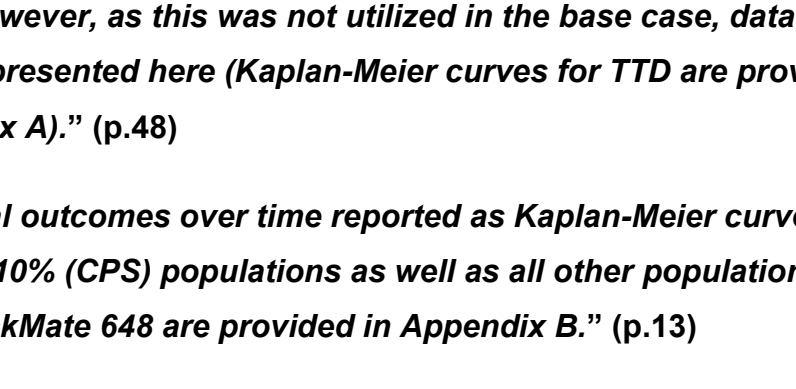
E)

Figure 7. Models fit to CheckMate 648 chemotherapy PFS (investigator-assessed) in patients with PD-L1 ≥10% (combined positive score): fractional polynomial (A), standard parametric (B), spline hazard (C), spline normal (D), spline odds (E)

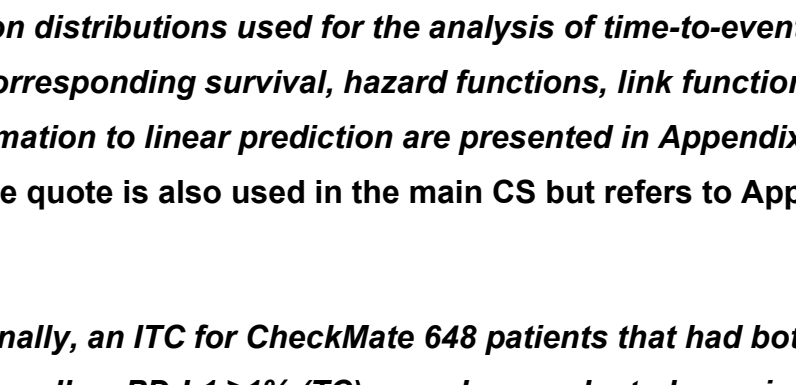
A)



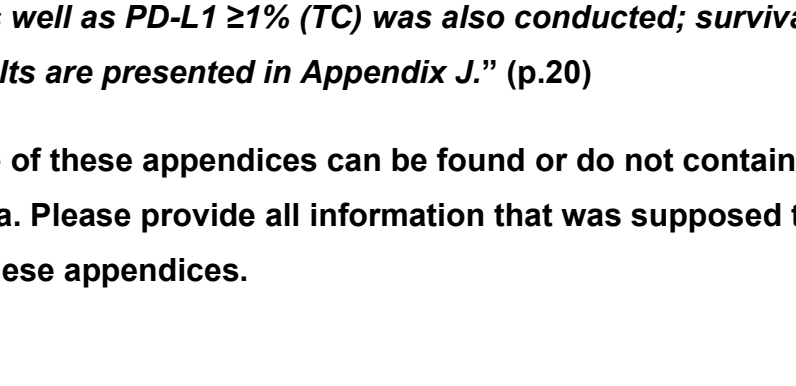
B)



C)



D)



E)

A24. Priority question: There are several references to other appendices in Appendix L:

- **"*TTD data was also considered to inform treatment duration within the CEM, however, as this was not utilized in the base case, data and analyses are not presented here (Kaplan-Meier curves for TTD are provided in Appendix A).*" (p.48)**
- **"*Survival outcomes over time reported as Kaplan-Meier curves for the ITT and PD-10% (CPS) populations as well as all other populations of interest for CheckMate 648 are provided in Appendix B.*" (p.13)**
- **"*Common distributions used for the analysis of time-to-event data as well as the corresponding survival, hazard functions, link functions, and transformation to linear prediction are presented in Appendix C*". (p.48). The same quote is also used in the main CS but refers to Appendix L (p. 71).**
- **"*Additionally, an ITC for CheckMate 648 patients that had both PD-L1 ≥10% (CPS) as well as PD-L1 ≥1% (TC) was also conducted; survival models and ITC results are presented in Appendix J.*" (p.20)**

However, none of these appendices can be found or do not contain the referenced data. Please provide all information that was supposed to be contained in these appendices.

Response: The appendices are provided as sub-appendix A, B, C and I (in the company submission referred to as Appendix J) to the updated Appendix L with this response.

A25. Document B, Section B.2.9.2.1 of the CS includes the statement: "…*only standard parametric and spline models were fit to KEYNOTE-590. As 3-, 4-, and 5-knot models were deemed overly complex, only 1- and 2-knot spline models were included in the NMA.*" Please provide further explanation on why the above models were deemed overly complex.

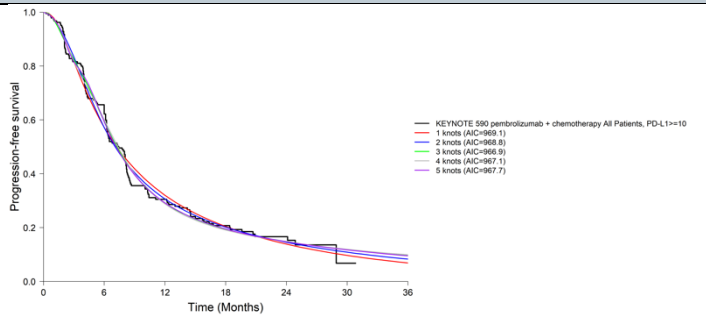
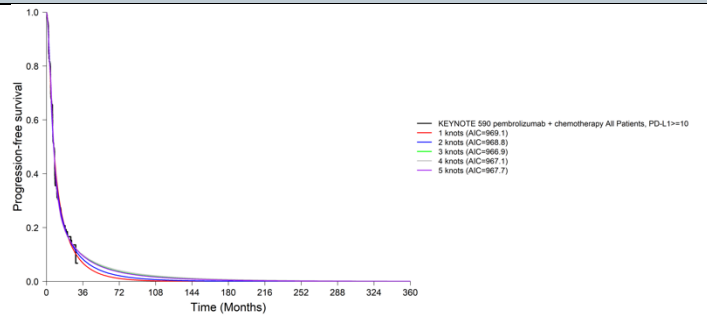
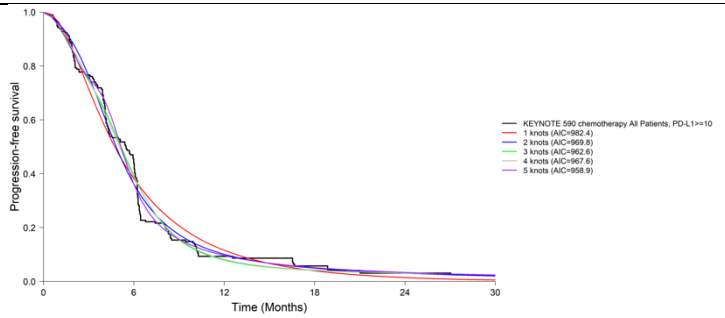
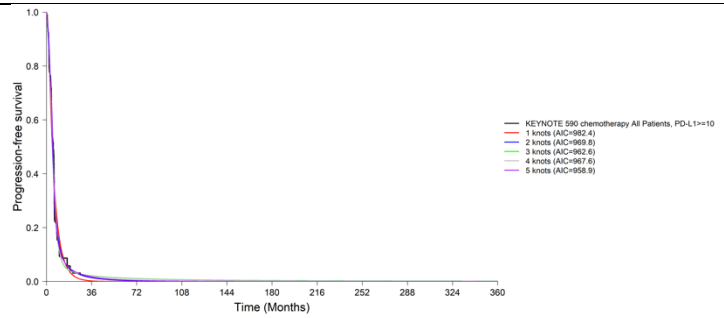
Response: It was incorrectly reported that 3-knot models were deemed overly complex; this only applied to 4- and 5-knot models (1-, 2-, and 3-knot models were all included for the estimation of RTEs). The decision to not use the 4- and 5-knot models was based on two factors: 1) they were deemed overly complex when compared to the observed hazards, and 2) the observed fit and extrapolation to the data was similar. Further justification is provided below.

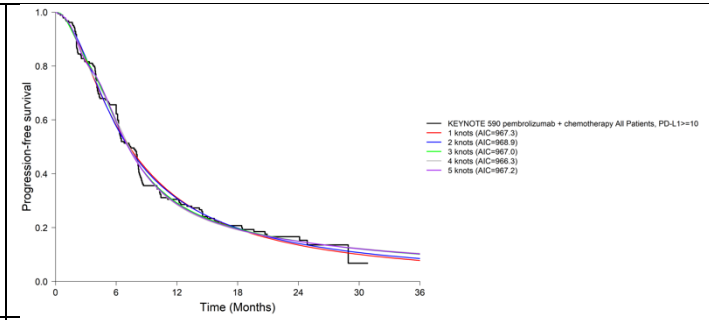
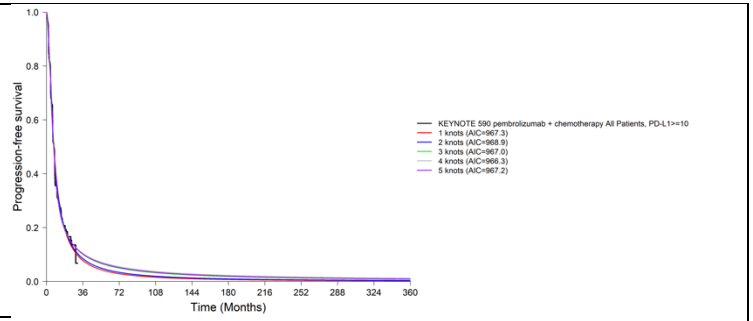
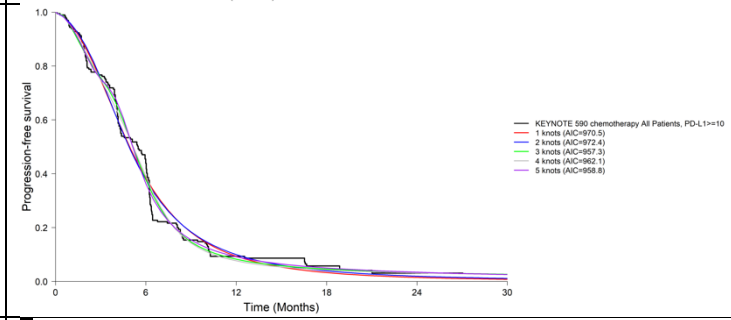
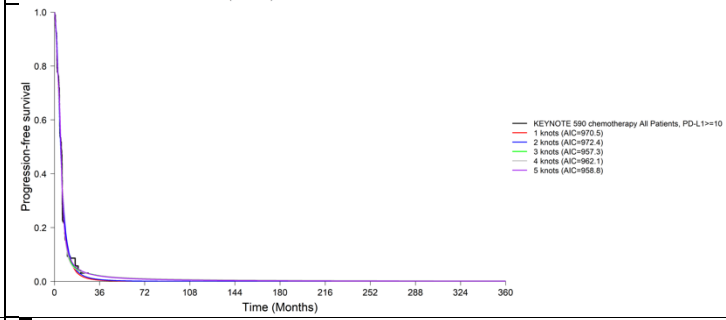
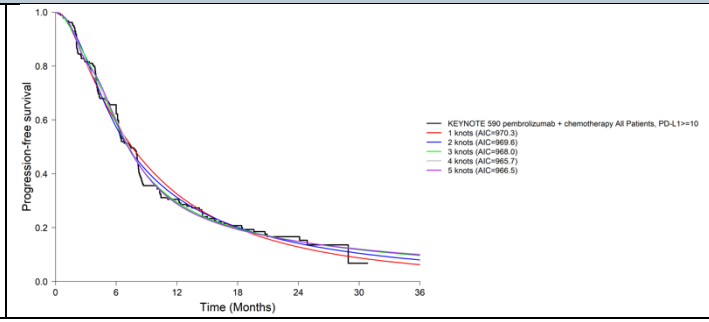
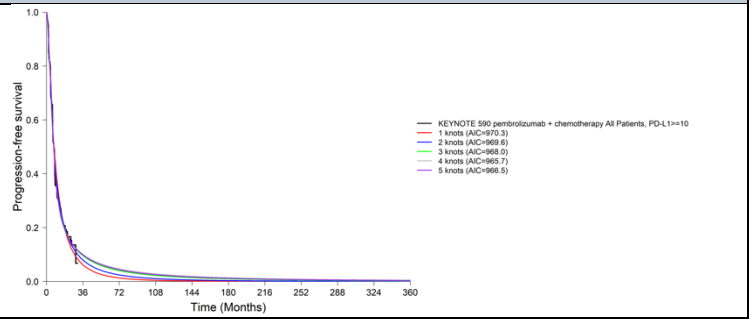
Figure 8 below shows the observed hazards over time of the three treatment arms relevant for the estimation of RTEs. As the goal of the spline models is to capture the complexity in the hazards with a sufficient number of knots, inflections points provide an important indicator of complexity. In the current case, Figure 8 demonstrates that there are no treatment arms for which there is more than one inflection point, indicating that the flexibility in models with more than 2 knots may be unnecessary.

Further, Figure 9 and Figure 10 show the comparison of the spline models over time for the treatment arms relevant to the estimation of RTEs (results taken from Appendix H, Appendix K, and Appendix L of the survival analysis report). As is shown in these figures, the fit and long-term extrapolation of the 4-knot and 5-knot spline models are comparable to the 1-knot, 2-knot, and 3-knot models.

Figure 8. Hazards over time for progression-free survival (investigator-assessed; left) and overall survival (right) in patients with PD-L1 ≥10% (combined positive score) for treatment arms relevant to the estimation of relative treatment effects

Figure 9. Modelled survival over time for progression-free survival (investigator-assessed) in patients with PD-L1 (CPS) ≥10% (mixed histology for KEYNOTE-590) for treatment arms relevant to the estimation of relative treatment effects

Population	Observed period	30-year extrapolation
Spline (hazard)		
Pembrolizumab + chemotherapy		
Chemotherapy (KEYNOTE-590)		
Chemotherapy (CheckMate 648)	[redacted]	[redacted]
Spline (odds)		

Pembrolizumab + chemotherapy		
Chemotherapy (KEYNOTE-590)		
Chemotherapy (CheckMate 648)	[redacted]	[redacted]
Spline (normal)		
Pembrolizumab + chemotherapy		

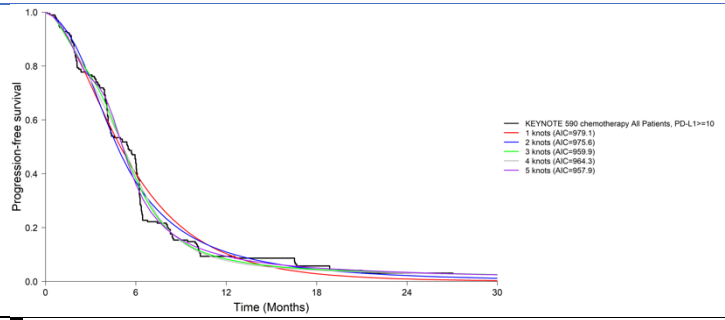
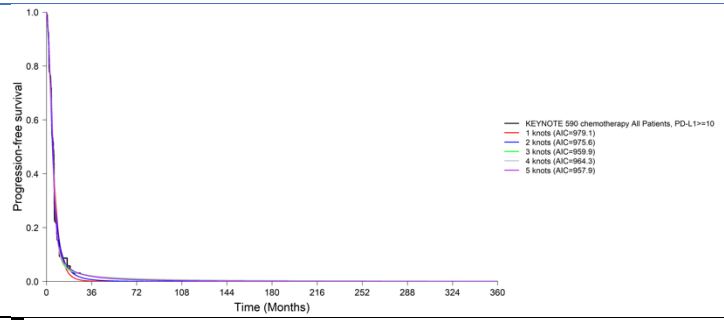
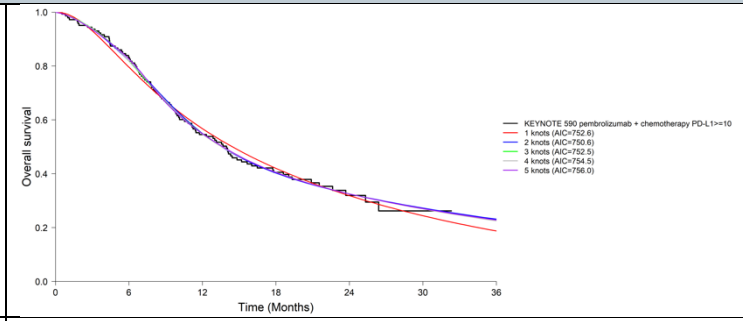
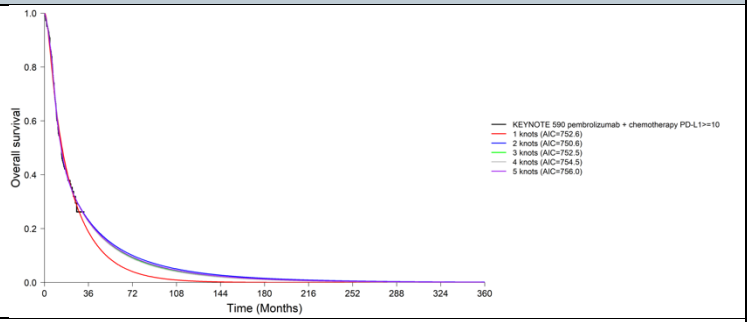
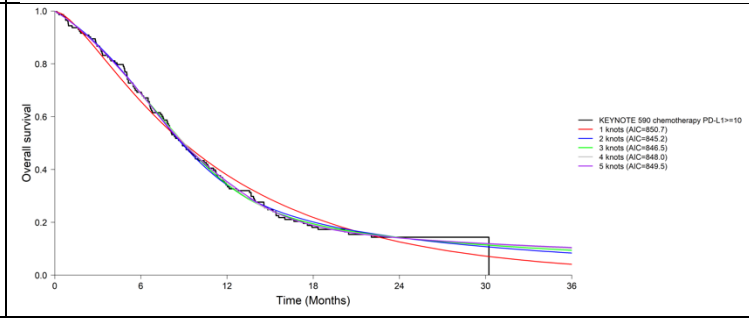
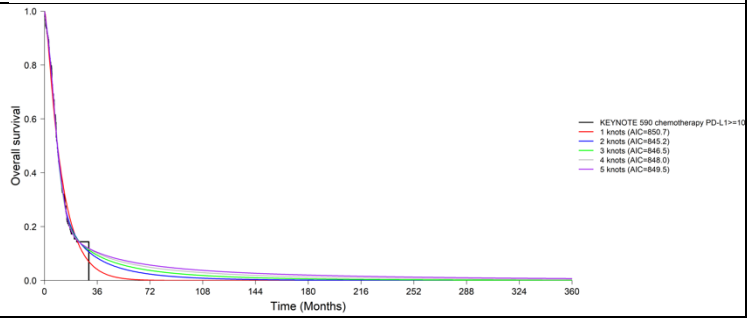
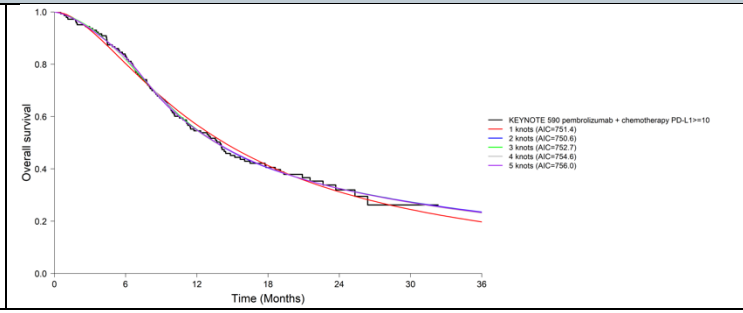
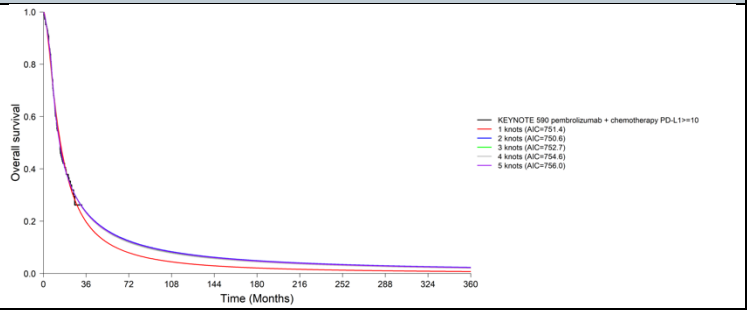
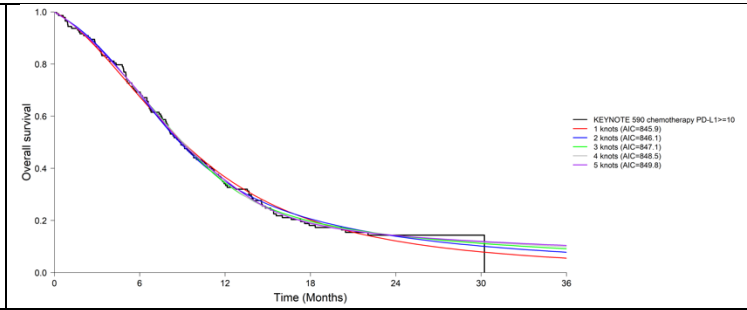
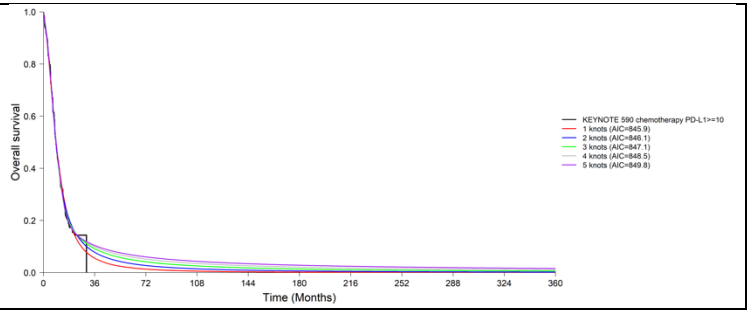
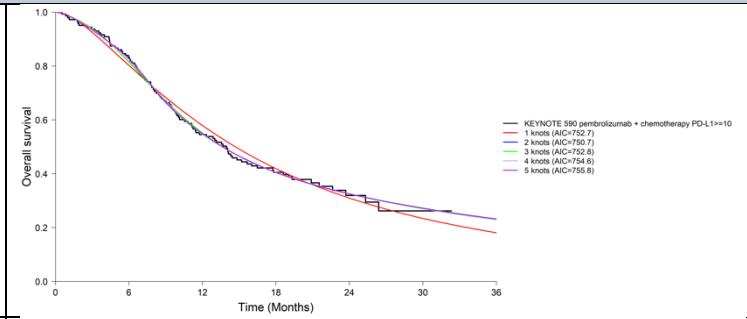
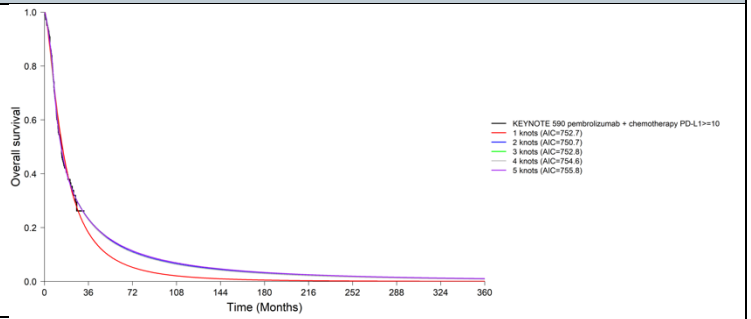
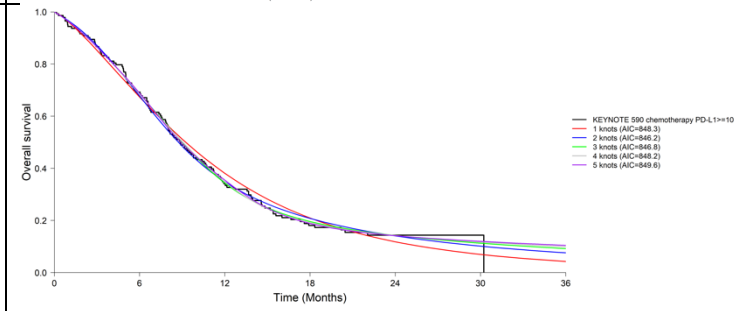
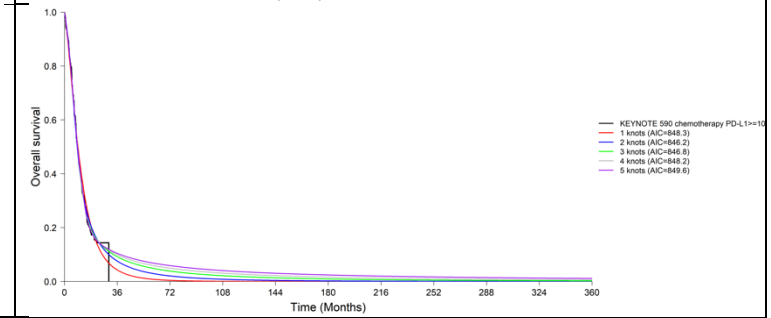
Chemotherapy (KEYNOTE-590)		
Chemotherapy (CheckMate 648)	[redacted]	[redacted]

Figure 10. Modelled survival over time for overall survival in patients with PD-L1 (CPS) ≥10% (OSCC only data) for treatment arms relevant to the estimation of relative treatment effects

Population	Observed period	30-year extrapolation
Spline (hazard)		
Pembrolizumab + chemotherapy		
Chemotherapy (KEYNOTE-590)		
Chemotherapy (CheckMate 648)	[redacted]	[redacted]
Spline (odds)		
Pembrolizumab + chemotherapy		

Chemotherapy (KEYNOTE-590)		
Chemotherapy (CheckMate 648)	[redacted]	[redacted]
Spline (normal)		
Pembrolizumab + chemotherapy		
Chemotherapy (KEYNOTE-590)		
Chemotherapy (CheckMate 648)	[redacted]	[redacted]

A26. Priority question: As reported in Appendix L of the CS, fitness of all tested models in the ITC analyses are assessed according to the Akaike information criterion (AIC) value. Spline hazard, odds and normal models, using 1-knot to 5-knots were fitted. In the progression free survival (PFS) ITCs, the choice of the best fitness was based only on the AIC of the pembrolizumab + chemotherapy arm. The chemotherapy arm, which had contradicting AIC results, was not taken into consideration. For example, the 5-knots spline hazard model in the chemotherapy arm has a lower AIC of 958.9 than the 1-knot model (982.4) but the latter is preferred. Please provide justification on why the chemotherapy arm AIC values were not considered.

Response: One key clarification point is that the model choice for the estimation of RTEs *did* consider AIC for both chemotherapy from CheckMate 648 and pembrolizumab + chemotherapy for KEYNOTE-590. To illustrate this process of model selection, the AICs for models relevant to the estimation of RTEs are presented in Table 14 and Table 15 in addition the sum of the AICs across treatment arm for each model.

Though 5-knot and 4-knot models provide the best fit within specific spline approaches, these models have a total AIC of < 5 points of spline models with fewer knots. Furthermore, as outlined above in Figure 8, the observed hazards over time of the three treatment arms show that no treatment within the estimation of RTEs had more than one inflection point indicating that the flexibility of 2-knot, 3-knot, 4-knot, and 5-knot models may be unnecessary. In addition, Figure 9 and Figure 10 show the fit and long-term extrapolation of the 4-knot and 5-knot spline models are comparable to the 1-knot, 2-knot, and 3-knot models. Therefore, inclusion of the 4-knot and 5-knot spline models would have a marginal impact on model selection and results.

Table 14. Summary of goodness of fit statistics for progression-free survival (investigator-assessed) for spline models in the PD-L1 ≥10% (combined positive score) population

Approach	Pembrolizumab + chemotherapy			Chemotherapy (KEYNOTE-590)			Chemotherapy (CheckMate 648)			Total	
	Distribution	*AIC*	*BIC*	*Distribution*	*AIC*	*BIC*	*Distribution*	*AIC*	*BIC*	***AIC***	***BIC***
Spline (hazards)	1 knot	969.14	978.81	1 knot	982.41	981.85	1 knot	814.93	823.86	2766.48	2784.52
	2 knots	968.75	981.66	2 knots	969.79	979.07	2 knots	808.68	820.59	2747.22	2781.32
	3 knots	966.87	983	3 knots	962.65	987.3	3 knots	809.7	824.58	2739.22	2794.88
	4 knots	967.08	986.43	4 knots	967.6	982.92	4 knots	806.79	824.65	2741.47	2794
	5 knots	967.7	990.28	5 knots	958.87	992.26	5 knots	808.31	829.15	2734.88	2811.69
Spline (normal)	1 knot	970.27	979.95	1 knot	979.09	980.93	1 knot	813	821.93	2762.36	2782.81
	2 knots	969.64	982.54	2 knots	975.56	976.29	2 knots	808.38	820.29	2753.58	2779.12
	3 knots	967.97	984.1	3 knots	959.88	983.99	3 knots	809.87	824.75	2737.72	2792.84
	4 knots	965.69	985.04	4 knots	964.29	988.7	4 knots	808.9	826.76	2738.88	2800.5
	5 knots	966.46	989.04	5 knots	957.94	988.94	5 knots	809.62	830.46	2734.02	2808.44
Spline (odds)	1 knot	967.34	977.02	1 knot	970.49	973.71	1 knot	809.65	818.58	2747.48	2769.31
	2 knots	968.92	981.82	2 knots	972.4	981.76	2 knots	807.82	819.72	2749.14	2783.3
	3 knots	967	983.13	3 knots	957.29	981.84	3 knots	809.17	824.06	2733.46	2789.03
	4 knots	966.29	985.64	4 knots	962.15	980.34	4 knots	807.47	825.34	2735.91	2791.32
	5 knots	967.17	989.75	5 knots	958.78	985.54	5 knots	808.63	829.47	2734.58	2804.76

Table 15. Summary of goodness of fit statistics for overall survival for spline models in the PD-L1 ≥10% (combined positive score) population

Approach	Pembrolizumab + chemotherapy			Chemotherapy (KEYNOTE-590)			Chemotherapy (CheckMate 648)			Total	
	Distribution	*AIC*	*BIC*	*Distribution*	*AIC*	*BIC*	*Distribution*	*AIC*	*BIC*	***AIC***	***BIC***
Spline (hazards)	1 knot	752.61	761.49	1 knot	850.71	857.04	1 knot	670.8	679.73	2274.12	2298.26
	2 knots	750.62	762.47	2 knots	845.19	861.35	2 knots	672.59	684.49	2268.4	2308.31
	3 knots	752.45	767.27	3 knots	846.54	865.77	3 knots	674.07	688.95	2273.06	2321.99
	4 knots	754.47	772.25	4 knots	847.99	870.26	4 knots	664.91	682.77	2267.37	2325.28
	5 knots	755.96	776.7	5 knots	849.52	859.6	5 knots	673.88	694.72	2279.36	2331.02
Spline (normal)	1 knot	752.73	761.62	1 knot	848.25	858.08	1 knot	669.1	678.03	2270.08	2297.73
	2 knots	750.71	762.56	2 knots	846.23	861.61	2 knots	671.05	682.96	2267.99	2307.13
	3 knots	752.81	767.63	3 knots	846.79	865.96	3 knots	672.99	687.87	2272.59	2321.46
	4 knots	754.55	772.33	4 knots	848.18	857.14	4 knots	665.02	682.88	2267.75	2312.35
	5 knots	755.76	776.5	5 knots	849.55	870.29	5 knots	673.59	694.43	2278.9	2341.22
Spline (odds)	1 knot	751.4	760.29	1 knot	845.85	854.74	1 knot	669.28	678.21	2266.53	2293.24
	2 knots	750.63	762.48	2 knots	846.08	857.93	2 knots	669.72	681.63	2266.43	2302.04
	3 knots	752.65	767.47	3 knots	847.07	861.89	3 knots	671.49	686.37	2271.21	2315.73
	4 knots	754.57	772.35	4 knots	848.47	866.25	4 knots	664.73	682.59	2267.77	2321.19
	5 knots	755.98	776.72	5 knots	849.84	870.58	5 knots	673.72	694.56	2279.54	2341.86

A27. Priority question: Document B, Section B.2.9.1.1 of the CS includes the statement: "*Both fixed and random-effects models were considered. Given insufficient evidence for estimation of the between-study heterogeneity (characterised by the heterogeneity parameter), fixed-effects models were used.*"

a) Please provide the results of both fixed-effect and random-effects models.

Response: As previously mentioned (see response to A23), only RTEs are being estimated. Though random effects models were considered (consistent with recommendations from the NICE Decision Support Unit and Technical Support Documents 14, specifically), this is relevant for cases where, in a given network of evidence, there are multiple trials informing one or more direct comparisons where between-study heterogeneity can be estimated and assessed. However, the current estimation of RTEs only utilized one trial and thus, there does not exist any between-study heterogeneity to estimate. For this reason, the random effects model was not employed.

b) Please provide details on the assessment of between-study heterogeneity.

Response: The feasibility assessment is presented in Section 5.1 of the updated Appendix L NMA report. The key findings resulting from the feasibility assessment are:

- KEYNOTE-590 did not allow patients with prior treatment experience while CheckMate 648 allowed patients with prior treatment provided it was completed more than six months prior to trial enrolment resulting in nearly 80% of patients with prior treatment experience in CheckMate 648. It is assumed that these differences do not act as treatment effect modifier

- Cycle lengths differed between CheckMate 648 and KEYNOTE-590 resulting in longer planned treatment duration for CheckMate 648 (though median treatment durations were 3.4 and 5.8 months for the chemotherapy arms, respectively). It is assumed that these differences do not act as treatment effect modifiers

- Though no differences in treatment effect modifiers could be identified through analysis of baseline patient characteristics (only available for the mixed histology population in KEYNOTE-590), it is assumed this is also true for the ESCC patients with PD-L1 ≥10% (CPS) within each trial.

c) Was the deviance information criterion (DIC) explored as a method to compare the fitness of the models? If so, please provide the results of this assessment.

Response: For the reasons provided in response to A27a above, random effects models were not employed and thus, the Deviance Information Criterion was not estimated.

A28. Priority question: The company states in Document B, Section B.2.9.3 that "*…no significant between trial heterogeneity was identified that would affect the comparability of the trials and prevent their inclusion in the NMA. Therefore, no assessment of heterogeneity in the form of I-square analysis can be estimated*". Please provide detailed information on how heterogeneity was assessed and why statistical exploration of heterogeneity was not attempted.

Response: The complete feasibility assessment is now presented in Section 5.1 of the updated Appendix L NMA report. Statistical exploration of heterogeneity was not attempted as only one trial informed the estimation of RTEs.

A29. Priority question: Document B, Section B.2.9.1.1 of the CS and Appendix L provide an overview of the methods that were used in the NMA, which entail the estimation of the "*…differences in each of the survival function parameters between pembrolizumab with chemotherapy and chemotherapy (both from KEYNOTE-590)*" and applying them "*… to chemotherapy as assessed in CheckMate 648 to obtain PFS (IA) and OS over time for pembrolizumab with chemotherapy relative to chemotherapy, as assessed in CheckMate 648*". The results of this process are reported but not the survival function parameters.

a) Please provide a full explanation as to the choice of parametric survival distributions that were tested in the analysis.

Response: The range of standard parametric models tested were consistent with those recommended by TSD 14 and comprised exponential, Weibull, Gompertz, log-logistic, log normal, gamma, and generalized gamma.

b) Please provide the results of the differences in the survival function parameters estimates that were used as the NMA model parameters.

Response: As described in Appendix L of the company submission, the parameters and uncertainty matrices of the parametric survival models fitted independently to the CPS ≥10% subgroups of the pembrolizumab + chemotherapy and chemotherapy arms of KEYNOTE-590 were synthesized using the method of Cope et al. (2020).[23] This method is applied in a Bayesian framework, and results in a posterior distribution of survival time distribution parameters and differences between these parameters. These parameter differences are assumed to represent the relative treatment effect of each pair of treatments, and that by applying the parameter difference to the parameters of a model representing outcomes upon the reference treatment, the parametric model predicting outcomes upon the investigational treatment may be formed. The instantaneous hazard at any point in time upon the reference treatment model and upon the scaled investigational treatment model may be calculated, and the ratio between these expressed as a point hazard ratio. Due to structural differences between the survival models used in the indirect treatment comparison (which were fully parametric, single-piece models) and the survival models used in the economic model (which were piecewise non-parametric/parametric models) the parameter differences could not be used directly in the economic model. Instead, the time-varying hazard ratios predicted by applying the parameter deltas to the model of the CPS ≥ 10% subgroup of the chemotherapy arm of KEYNOTE-590 were used to scale the chemotherapy survival models used in the economic model. This is described further in question A32.

The deltas between pembrolizumab + chemotherapy and chemotherapy of the parameters upon the (-Inf, Inf) transformed scale (posterior median and covariance matrix) are reproduced here as requested in Table 16 (PFS) and Table 17 (OS). However, we note that they have no direct use in the economic model as specified, as they are only assumed transitive to other fully parametric models of the same

distribution family, and therefore cannot be directly applied to the chemotherapy (reference) piecewise models used.

Table 16. Posterior distribution of parameter differences - PFS - pembrolizumab + chemotherapy vs chemotherapy

Distribution	Parameter	Trans.	Median Δ	Variance-covariance		
Log-logistic	shape	ln	-2.60E-01	9.15E-03	-6.51E-04	
	scale	ln	4.36E-01	-6.51E-04	9.46E-03	
Weibull	shape	ln	-1.65E-01	7.54E-03	1.19E-03	
	scale	ln	5.44E-01	1.19E-03	8.63E-03	
Gompertz	shape	I()	-3.25E-02	3.72E-04	-2.27E-03	
	rate	ln	-4.15E-01	-2.27E-03	2.66E-02	
Lognormal	mean(log)	I()	4.86E-01	9.96E-03	9.93E-04	
	sd(log)	ln	1.93E-01	9.93E-04	6.64E-03	
Gamma	shape	ln	-3.09E-01	2.14E-02	2.35E-02	
	rate	ln	-8.90E-01	2.35E-02	3.42E-02	
Generalised gamma	mu	I()	3.18E-01	2.33E-02	-4.00E-03	3.18E-02
	sigma	ln	2.70E-01	-4.00E-03	8.15E-03	-9.67E-03
	Q	I()	-4.28E-01	3.18E-02	-9.67E-03	7.42E-02

Table 17. Posterior distribution of parameter differences - OS - pembrolizumab + chemotherapy vs chemotherapy

Distribution	Parameter	Trans.	Median Δ	Variance-covariance		
Log-logistic	shape	ln	-6.27E-02	1.41E-02	-1.49E-03	
	scale	ln	4.97E-01	-1.49E-03	1.56E-02	
Weibull	shape	ln	1.99E-02	1.39E-02	-3.10E-04	
	scale	ln	4.80E-01	-3.10E-04	1.29E-02	
Gompertz	shape	I()	-2.62E-03	4.28E-04	-3.99E-03	
	rate	ln	-5.39E-01	-3.99E-03	5.59E-02	
Lognormal	mean(log)	I()	5.44E-01	1.85E-02	2.90E-03	
	sd(log)	ln	1.63E-02	2.90E-03	1.12E-02	
Gamma	shape	ln	2.23E-02	2.98E-02	3.46E-02	
	rate	ln	-4.63E-01	3.46E-02	5.33E-02	
Generalised gamma	mu	I()	3.55E-01	3.96E-02	-1.57E-02	6.34E-02
	sigma	ln	1.31E-01	-1.57E-02	2.35E-02	-4.42E-02
	Q	I()	-4.24E-01	6.34E-02	-4.42E-02	1.65E-01

c) Please provide a landmark analysis of PFS and overall survival (OS) for pembrolizumab plus chemotherapy relative to chemotherapy (as assessed in KEYNOTE-590) in a tabular form so that comparison can be made

between the PFS (investigator-assessed, IA) and OS over time for pembrolizumab plus chemotherapy relative to chemotherapy (as assessed in CheckMate 648).

Response: The landmark analyses for pembrolizumab plus chemotherapy relative to chemotherapy in KEYNOTE-590 and in CheckMate 648 are displayed in Table 18 (PFS) and Table 19 (OS).

Table 18. Progression free survival for landmark analysis for pembrolizumab plus chemotherapy relative to chemotherapy

PFS	1 year	2 years	3 years	3.5 years
CHEMO KEYNOTE-590	8.3%	3.4%	-	-
PEMBRO +CHEMO KEYNOTE-590	29.9%	15.3%	-	-
CHEMO CheckMate 648	[redacted]	[redacted]	[redacted]	[redacted]
PEMBRO + CHEMO CheckMate 648	[redacted]	[redacted]	[redacted]	[redacted]
CHEMO: chemotherapy; PFS: Progression free survival				

Table 19. Overall survival for landmark analysis for pembrolizumab plus chemotherapy relative to chemotherapy

OS	1 year	2 years	3 years	3.5 years
CHEMO KEYNOTE-590	37.7%	17.0%	-	-
PEMBRO +CHEMO KEYNOTE-590	54.5%	31.9%	-	-
CHEMO CheckMate 648	[redacted]	[redacted]	[redacted]	[redacted]
PEMBRO + CHEMO CheckMate 648	[redacted]	[redacted]	[redacted]	[redacted]
CHEMO: chemotherapy; OS: Overall survival				

A30. Priority question: Appendix N presents landmark analyses only of overall survival for PD-L1 ≥1% (TC) from CheckMate 648.

a) Please provide similar landmark analyses for both PD-L1≥1% (TC) and PD-L1 ≥10% (CPS):

i. Kaplan-Meier data for pembrolizumab plus chemotherapy and chemotherapy from CheckMate 648

Response: The following figure presents outcomes of overall survival of patients from CheckMate 648 PD-L1 ≥10% (CPS) DBL.

Landmark survival analysis results for CheckMate 648 and KEYNOTE-590 are also provided in Table 20.

Table 20. Landmark overall survival for CheckMate 648 and KEYNOTE-590

	Overall Survival (95% CI) (%)					
	CheckMate 648*				KEYNOTE-590†	
	PD-L1 ≥1% (TC)		PD-L1 ≥10% (CPS)		PD-L1 ≥10% (CPS)	
Time (Months)	**CHEMO**	**NIVO+ CHEMO**	**CHEMO**	**NIVO+ CHEMO**	**CHEMO**	**PEMBRO + CHEMO**
6	[redacted]	[redacted]	[redacted]	[redacted]		
12	[redacted]	[redacted]	[redacted]	[redacted]	37.7%	54.5%
18	[redacted]	[redacted]	[redacted]	[redacted]		
24	[redacted]	[redacted]	[redacted]	[redacted]	17.0%	31.9%
30	[redacted]	[redacted]	[redacted]	[redacted]		
36	[redacted]	[redacted]	[redacted]	[redacted]		

**BMS data on file*

†Reconstructed PLD from digitisation of Figure 1a, Sun et al (2021)[5]

ii. Each parametric or semi-parametric extrapolation model for nivolumab plus chemotherapy, pembrolizumab plus chemotherapy and chemotherapy.

Response: Figure 11 shows the fits of the standard parametric OS models in the CHEMO arm for the PD- L1 ≥10% (CPS).

Even though log-logistic provides the best statistical fit per goodness-of-fit statistics, namely AIC and BIC statistics, it is not overly clinically plausible and so the generalised gamma is deemed to be the best choice of extrapolation.

Figure 11. CheckMate 648 PD-L1 ≥10% (CPS) [redacted] DBL standard parametric models for OS overlaid upon Kaplan-Meier: CHEMO

[redacted]

Figure 12 shows how standard parametric models fit the OS trial data of CheckMate 648 in the NIVO+CHEMO arm.

Lognormal presents the most statistically and clinically plausible model selection to represent this data.

Figure 12. CheckMate 648 PD-L1 ≥10% (CPS) [redacted] DBL Standard parametric models for OS overlaid upon Kaplan-Meier: NIVO+CHEMO

[redacted]

Figure 13 shows how standard parametric models fit the PFS (BICR assessed) trial data of CheckMate 648 in the NIVO+CHEMO arm.

Figure 13. CheckMate 648 PD-L1 ≥10% (CPS) [redacted] DBL Standard parametric models for PFS (BICR assessed) overlaid upon Kaplan-Meier: CHEMO

[redacted]

Figure 14 shows the fits of the standard parametric PFS (BICR assessed) models in the NIVO+CHEMO arm for the PD-L1 ≥10% (CPS).

Figure 14. CheckMate 648 PD-L1 ≥10% (CPS) [redacted] DBL Standard parametric models for PFS (BICR assessed) overlaid upon Kaplan-Meier: NIVO+CHEMO

[redacted]

Semi-parametric extrapolations were also considered, the results of which are presented below.

Figure 15. CheckMate 648 PD-L1 ≥10% (CPS) [redacted] DBL Semi-parametric (6.9 month cut-point) models for OS overlaid upon Kaplan-Meier: CHEMO

[redacted]

Figure 16. CheckMate 648 PD-L1 ≥10% (CPS) [redacted] DBL Semi-parametric (6.9 month cut-point) models for OS overlaid upon Kaplan-Meier: NIVO+CHEMO

[redacted]

Figure 17. CheckMate 648 PD-L1 ≥10% (CPS) [redacted] DBL Semi-parametric (6.9 month cut-point) models for PFS (BICR) overlaid upon Kaplan-Meier: CHEMO

[redacted]

Figure 18. CheckMate 648 PD-L1 ≥10% (CPS) [redacted] DBL Semi-parametric (6.9 month cut-point) models for PFS (BICR) overlaid upon Kaplan-Meier: NIVO+CHEMO

Figure 19. CheckMate 648 PD-L1 ≥10% (CPS) [redacted] DBL Semi-parametric (6.9 month cut-point) models for PFS (BICR) overlaid upon Kaplan-Meier: NIVO+CHEMO

[redacted]

iii. Any external validation data including for chemotherapy.

Response: No external validation of these data was conducted as this was not feasible within the time frame of this response.

b) Please provide an analysis of the choice of extrapolation model based on the clinical plausibility of the results of the landmark analysis and in line with NICE Technical Support Document (TSD) 14.

Response: In deciding a choice of extrapolation, the six standard parametric models per NICE TSD 14[24] were considered.

For each grouping, the best fitting standard parametric models per the goodness of fit statistics (per AIC and BIC fits) that are considered clinically plausible were selected as the choices of extrapolation.

For OS, in the CHEMO arm, despite log-logistic presenting the best statistical fit per AIC and BIC, it cannot be deemed clinically plausible. Further, of the clinically plausible curves, the generalised gamma is the best fitting curve. For the NIVO+CHEMO arm, the choice of survival model is the lognormal as it has the greatest statistical fit to the data, as demonstrated by its goodness of fit statistic.

For the PFS (BICR) outcome, in the CHEMO arm, despite both presenting good statistical fits to the data, the log-logistic and lognormal AIC and BIC score does not differ enough to suggest that the lognormal has significantly greater statistical fit than the log-logistic. Of these two distributions, the log-logistic fits the tail of the data better and so is the best choice of model. For the NIVO+CHEMO arm, the lognormal is the best fitting model that is also clinically plausible (generalised gamma presents good fit but is not clinically plausible as it implies long-term survival for these patients).

Semi-parametric models were also considered in order to provide the most appropriate fits to the observed data. The 6.9 month cut-point was chosen as to

avoid the sharp change in the hazard observed in the first six months for NIVO+CHEMO and CHEMO. Similar to the selection of standard parametric models, clinical plausibility was considered as well as goodness of fit and also the fit to tail of the data was considered.

As observed in Table 21, the observed survival data supports use of the base case survival approach.

Table 21. Survival Distribution selection for CheckMate 648 PD-L1 ≥10% (CPS) [redacted] DBL

Outcome	CHEMO	NIVO+CHEMO
OS		
Semi-parametric (6.9 month cut-point)	Weibull	Generalised Gamma
PFS (BICR assessed)		
Semi-parametric (6.9 month cut-point)	Log-logistic	Generalised Gamma
BICR: blinded independent central review; CHEMO: chemotherapy; NIVO: nivolumab; OS: overall survival; PFS: progression-free survival.		

A31. Priority question: Section 3.2 of Appendix L states that: "…*it was assumed that there were no differences in treatment effect modifiers between the chemotherapy arms of KEYNOTE-590 and CheckMate 648*". Please provide detailed evidence to support this statement.

Response: As part of the assessment of heterogeneity undertaken during the NMA, treatment effect modifiers were identified and compared between the populations in the CheckMate 648 and KEYNOTE-590 trial to identify potential sources of heterogeneity. Baseline characteristics were assessed; however, these were only available for the ITT population from KEYNOTE-590, and therefore, these comparisons should be interpreted with caution.

Patient age was similar across studies, with mean or median age ranging from 63 to 64 years. The proportion of Asian patients were: KEYNOTE-590, (52%) and CheckMate 648 (East Asian 57% and rest of Asia 13%).[25] Both trials reported ECOG PS at baseline with similar proportions of patients with ECOG PS 0 (40% to 47%) or 1 (53% to 60%). CheckMate 648 reported the number of organ metastases; 49% with one or less organ metastases and 51% with two or more organ metastases.

KEYNOTE-590 did not report number of organ metastases. Liver metastases was not reported by either study.

A32. Priority question: In Document B, Section B.2.9.2.2 of the CS the company reports the ITC results of the NMA analysis for "*pembrolizumab with chemotherapy vs nivolumab with chemotherapy: patients with PD-L1 CPS ≥10*" using a fixed effects Gamma model.

a) **Please provide the justification for using the Gamma model.**

Response: Section B.2.9.2.2 presents ITC results to illustrate the non-significant results of the comparison for OS. The fixed effects gamma model was used as an example but is not used in our analysis. This is aligned with other models assessed within the Appendix L (ITC report) of the company submission.

Within the economic model, the results of the estimation of relative treatment effects between PEMBRO+CHEMO and CHEMO were applied by modifying the chemotherapy PFS and OS curves by the output time-varying hazard ratios (HRs). The models used in this case were matched in family to the extrapolative portions of the chemotherapy models, i.e. for PFS, the chemotherapy PFS model (piecewise Kaplan-Meier / Weibull) was scaled by the time-varying HRs from the NMA using fully parametric Weibull models to represent PFS; for OS, the chemotherapy OS model (piecewise Kaplan-Meier / lognormal) was scaled by the time-varying HRs from the NMA using fully parametric lognormal models to represent OS. Due to the structural constraints upon the hazard function implied by the different distribution families, it was inappropriate to use time-varying hazard ratios that were derived from models from a different family; therefore, although the gamma model was used for demonstrating clinical difference, for consistency with the economic model, the results of these analyses using consistent distribution families were used.

In practice, the adjustment of the survival curves by these time-varying hazard ratios was undertaken offline. To do so, the following process was undertaken (t in unit "timestep" unless otherwise specified):

1. Evaluate accumulated hazard to each model timestep using the relationship $H(t) = -\ln\big(S(t)\big)$

2. Calculate mean hazard experienced during timestep as $h(t) = H(t) - \big(H(t-1)\big)$

3. Multiply this hazard by the reported point hazard ratio at nearest time ≤ t for t ≤ 36 months, else hold at hazard ratio for t = 36 months

4. Accumulate these hazards to form a scaled cumulative hazard $H_{scaled}(t) = \sum_{i=0}^{t} h(i)$

5. Convert scaled cumulative hazard to scaled survival using reverse of relation (1)

The time-varying hazard ratios used in this process are given in Table 22 (PFS) and Table 23 (OS).

Table 22. Time-varying hazard ratio of PFS for parametric models; PEMBRO+CHEMO vs CHEMO

Model	Hazard ratio at month					
	3	6	9	12	24	36
Gamma	0.52	0.47	0.45	0.43	0.42	0.42
Gen. gamma	0.56	0.49	0.47	0.45	0.40	0.38
Gompertz	0.60	0.54	0.49	0.45	0.30	0.21
Log-logistic	0.53	0.52	0.57	0.61	0.70	0.73
Lognormal	0.54	0.54	0.55	0.56	0.58	0.58
Weibull	0.56	0.49	0.45	0.43	0.37	0.35

Table 23. Time-varying hazard ratio of OS for parametric models; PEMBRO+CHEMO vs CHEMO

Model	Hazard ratio at month					
	3	6	9	12	24	36
Gamma	0.49	0.52	0.54	0.55	0.57	0.58
Gen. gamma	0.53	0.56	0.55	0.54	0.52	0.5
Gompertz	0.58	0.58	0.57	0.57	0.55	0.53
Log-logistic	0.46	0.48	0.52	0.57	0.73	0.82
Lognormal	0.40	0.51	0.56	0.60	0.68	0.72
Weibull	0.53	0.54	0.55	0.55	0.55	0.56

b) **Please provide NMA analysis results using the rest of the recommended models according to according to NICE TSD 14.**

Response: The NMA results for the different models are presented in sub-appendix C of the updated Appendix L.

A33. The company states in Section 3.2 of Appendix L that "*All analyses were performed in a Bayesian framework and involved a model with parameters, data and a likelihood distribution, and prior distributions*". Please provide the details of the Bayesian framework that was applied.

Response: The RTEs of pembrolizumab + chemotherapy versus chemotherapy were synthesized in a Bayesian framework. For a given parametric survival distribution, the data was the arm-level scale and shape parameters, the likelihood was a multivariate Normal distribution, and the parameters of interest are the relative treatment effects, i.e. scale and shape parameter d's. Parameters have been provided as part of the response to A29b. Normal non-informative prior distributions with a mean of 0 and a variance of 10,000 were used for the relative treatment effect parameters estimated**.**

A34. The network diagram for the included arms of the NMA presented in Document B, Section B.2.9.1.1 of the CS (Figure 21), includes intervention 2 of CheckMate 648 (NIVO+IPI: nivolumab 3 mg/kg every 2 weeks (Q2W) intravenously (IV) + ipilimumab 1 mg/kg every 6 weeks (Q6W) IV). Please confirm whether this arm was included in the NMA.

Response: The NIVO+IPI arm was included in the NMA as presented in Figure 5 of the NMA report (updated Appendix L). This treatment arm was included as it was part of the CheckMate 648 trial. The inclusion of the NIVO+IPI treatment arm does not affect the results of the NMA when considering the comparison of nivolumab with chemotherapy and pembrolizumab with chemotherapy using the common chemotherapy alone arm.

A35. Document B, Section B.2.9.1.1 of the CS reports that individual patient data (IPD) for the KEYNOTE-590 trial were not available and that "…*datasets for the models were sourced from digitized Kaplan-Meier curves and the*

***number of patients at risk over time from which IPD was recreated using the Guyott algorithm*".**

a) Please provide further details of the efficiency of these methods.

Response: Full details of the methods can be found in the attached publication, Guyot et al 2012.[26] In short, the publication states a mean error of -0.103% (95% confidence interval [CI] -0.260, 0.055). As an example, the article states that "…if the original survival probability estimate was 50%, we would expect survival probability based on reconstructed data to be 49.897% (95% CI 49.740, 50.055). There is therefore no significant systematic error." The article also states that "As the level of information available is decreased by successively removing data on numbers at risk and number of events, the mean error and the reproducibility standard deviation remain unaltered. There is, however, a slight fall in accuracy as assessed by mean absolute error and exemplar variance. In addition, this method has been commonly used in NICE submissions with three in the last six months utilizing the Guyot algorithm to reproduce individual patient-level data (TA528, ID3802, ID1557).

b) What is the margin of error expected when using the above methods and have they been considered during analyses?

Response: The margin of error has been listed in the answer above and are further detailed in the source publication, Guyot et al 2012.[26] The margin of error was considered in the analysis, however, given the number of patients at risk is provided at three-month intervals (high level information according to Guyot et al. 2012), this was considered sufficiently accurate to not require additional uncertainty to be included within the estimation of RTEs.

c) Please provide the digitized survival data extracted from the Kaplan-Meier curves.

Response: Digitised KEYNOTE-590 data informing analyses have been provided separately.

A36. According to Section 3.3 of Appendix L "*…the parameters of the different models were estimated using a Markov Chain Monte Carlo (MCMC) method implemented in the JAGS software package.[19] A first series of iterations from*

***the JAGS sampler were discarded as 'burn-in', and the inferences were based on additional iterations using two chains.*"**

a) **Please report how many iterations were used.**

Response: Two chains were used, each with 20,000 burn-in and 20,000 iterations.

b) **Was the convergence of the two chains tested?**

Response: Convergence was assessed using trace plots, density plots, Gelman-Rubin-Brooks plots, and auto-correlation plots. These have now been provided separately. Note that these have only been provided for the standard parametric models as spline models, which required consistent knot locations across arms, were fit using trial-level data and the *flexsurv* package.

Section B: Clarification on cost-effectiveness data

Model structure and implementation

B1. Priority question: Duration of subsequent treatment is not modelled via a separate health state, and it appears that there is only one line of subsequent treatment. Please provide justification for this aspect of the model structure by reference to previous technology appraisals and NHS clinical practice.

Response: Patient discontinuation of first line treatment is based on treatment specific time on treatment (ToT) curves. Patients then discontinue from second line treatment to no treatment based on a treatment specific cyclical discontinuation rate derived from the average ToT for the corresponding treatment in TA707.[22] The relevant costs are applied to each patient receiving each treatment per cycle, with a zero cost assumed for no treatment.

One of the limitations of a partitioned survival model (PSM) is its inability to explicitly track individual patients over time through subsequent lines of treatment. The PSM approach only captures patient status on a cohort-level between progression-free (denoted by the PFS curve), progressed disease (denoted by the difference between OS and PFS curves), and death (1 minus OS curve). It does not track individual patients (i.e. does not explicitly capture how long specific patients have spent within progressed disease, nor how long they have spent on subsequent treatments). Due to this limitation, discontinuation from second line treatment cannot be tracked or explicitly captured within the model, and so cannot be health state specific. . Best supportive care (BSC) would be the relevant a third-line treatment, with patients who discontinue second line treatment receiving BSC until death. The recent NICE appraisal for nivolumab in previously treated oesophageal cancer (TA707) utilised BSC for subsequent therapies in each arm,[22] reflecting clinical practice. However, the inclusion of BSC biases against treatments that increase survival through an additional cost. Therefore, the inclusion of BSC would create a bias towards the control arm, the arm which provides lower survival. To illustrate this, modelled patients, on average, remain alive longer after discontinuing second-line treatment in the treatment arm as compared to the control arm in the base case (1.739 and 0.845 years, respectively). Therefore, the inclusion of BSC would result in 0.891 years of

additional ‘third line’ treatment costs for the treatment arm compared to the control arm as a result of the increased survival that the treatment arm provides.

Additionally, BSC components are palliative as opposed to curative, and therefore are implicitly encompassed by the cost of terminal care as opposed to a subsequent line of treatment. Accordingly, in the model, the inclusion of BSC as a third line treatment was concluded to be inappropriate, and so no treatment was included in the third line.

These assumptions are in keeping with TA737 which only explicitly incorporated one line of subsequent treatments and did not explicitly incorporate further discontinuation to BSC.[6] Additionally, the options of subsequent treatment (further discussed in answer to B2 below) align with UK clinical practice and have been validated by UK clinicians.

B2. Priority question: It appears that the only subsequent treatment in the economic model was systemic therapy and only either nivolumab monotherapy or single agent taxanes.

a) Please explain why radiotherapy and surgery were not included.

Response: Radiotherapy and surgery were not included as subsequent treatments as these are considered palliative and not curative within the UK and therefore are encompassed implicitly within the cost of terminal care as opposed to a subsequent line of treatment. This approach is in keeping with TA737, where neither radiotherapy nor surgery were incorporated as subsequent treatments.[6]

b) Please explain why systemic therapies other than taxanes were administered in CheckMate 648 if recommended 2nd line chemotherapy is taxane monotherapy.

Response: CheckMate 648 was an international clinical trial, whereas the economic model has been tailored to a UK population. As such, subsequent therapies from CheckMate 648 incorporate treatments applicable to countries other than the UK. Conversely, only taxanes and nivolumab are relevant subsequent therapies to a UK population, as confirmed by clinicians during an advisory board meeting conducted by BMS and thus only these options are incorporated into the economic model.

c) **Please expand on the underlying budget impact modelling assumptions during NICE TA707 on subsequent therapies stated in the CS that dictated for patients on the chemotherapy arm to receive nivolumab monotherapy.**

Response: Budget impact assumptions from TA707 are not publicly available. However, within this submission for second line nivolumab in OSCC, nivolumab displaced the majority of taxane use. This indicates that, where nivolumab is applicable at second line, nivolumab would replace the use of taxanes. Therefore, within the chemotherapy arm of the company submission, nivolumab is used as the subsequent treatment (as opposed to single agent taxanes which are used in the NIVO+CHEMO arm).

d) **Document B, Section B.3.3.2.3 of the CS states that, *"clinical advisors to BMS advise that patients would not receive subsequent PD-L1 inhibitors following previous PD-L1 inhibitor use."* Please discuss (with supporting evidence) the subsequent therapies expected to be given to patients who have progressed on the study treatments in UK clinical practice as per clinical expert opinion.**

Response: During an advisory board meeting conducted by BMS,[9] clinicians specialising in the treatment of OSCC in the UK stated that if nivolumab combination therapy was approved as a first-line treatment, then they would not offer an immunotherapy-containing second-line therapy. It was generally believed that a docetaxel or paclitaxel-containing regimen would be offered in the second-line after a nivolumab-containing first-line regimen. This is in-line with current ESMO guidance, which recommends taxanes as monotherapy in second-line therapy for advanced or metastatic OSCC.[27]

During the NICE appraisal of pembrolizumab with chemotherapy for untreated oesophageal and gastro-oesophageal cancer (TA737),[6] it was deemed preferable to give treatment with a PD-L1 inhibitor early in the treatment pathway. During the appraisal, clinical experts explained that because pembrolizumab and nivolumab were both PD-L1 inhibitors, it would not be suitable to give nivolumab as a second-line treatment after pembrolizumab with chemotherapy and stated that it was likely that immunotherapy is more effective when used earlier.

During the NICE appraisal for nivolumab for previously treated unresectable advanced or recurrent oesophageal cancer (TA707),[22] clinical experts explained that people with unresectable advanced, recurrent or metastatic OSCC, whose disease has progressed after fluoropyrimidine and platinum-based combination therapy, receive the taxanes, paclitaxel and docetaxel as second-line therapy. The NHS England clinical lead noted that taxanes have limited efficacy and patients are often not well enough to have third-line treatment if taxanes do not control the disease. Patients who are unable to tolerate taxane chemotherapy receive best supportive care, which does not affect disease progression.

Therefore, according to current NICE guidelines and clinical expert feedback from an advisory board conducted by BMS and from previous NICE appraisals,[6,9,21,22] the second-line therapy for patients with advanced OSCC who have progressed on current first-line treatment, would be nivolumab or taxanes for patients who have received fluoropyrimidine and platinum-based combination therapy first-line. In patients who receive a PD-1 inhibitor first-line, only taxanes would be offered as second-line therapy.

Subsequent therapies within the economic model are within Table 36 of the company submission (reproduced below in Table 24). These are in line with clinical expert opinion: patients who have been treated with PD-L1 inhibitors (nivolumab or pembrolizumab), would not go on to receive PD-L1 inhibitors at subsequent lines. As such, first line NIVO+CHEMO and PEMBRO+CHEMO patients receive single use taxanes. Conversely, first line chemotherapy patients may receive nivolumab at second line.

Table 24. Subsequent therapy applied in model

Treatment arm	Subsequent treatment	Proportion of patients
NIVO+CHEMO	Single agent taxane; assumed equal use of docetaxel and paclitaxel	[redacted]%
CHEMO	Nivolumab monotherapy	[redacted]%
PEMBRO+CHEMO	Single agent taxane; assumed equal use of docetaxel and paclitaxel	Aligned with NIVO+CHEMO

CHEMO: chemotherapy; NIVO: nivolumab; PEMBRO: pembrolizumab
Source: CheckMate 648 October 2021

e) **In the CheckMate 648 trial, 10% of patients with PD-L1 ≥1% on the chemotherapy arm go on to receive nivolumab monotherapy. Please provide supporting evidence for the 56.7% of patients on the chemotherapy arm that go on to receive nivolumab monotherapy in the economic model.**

Response: The proportions of patients who go on to subsequent therapy ([redacted]% in the NIVO+CHEMO arm, [redacted] % in the CHEMO arm) are sourced from CheckMate 648 (sourced from the latest DBL, [redacted]).

f) **Please discuss the implications on effectiveness of patients in the nivolumab or pembrolizumab plus chemotherapy arms subsequently receiving only taxanes and those in the chemotherapy arm receiving only nivolumab monotherapy as opposed to what was actually received in the CheckMate 648 trial.**

Response: According to the latest DBL, [redacted]% of the NIVO+CHEMO patients received subsequent therapy, of which, [redacted]% received anti-PD(-L)1 and [redacted]% received other systemic anticancer therapy with some patients receiving a combination of anti-PD(-L)1 and other systemic therapy. In contrast, [redacted]% of the patients in the CHEMO arm received subsequent therapy with [redacted]% receiving anti-PD(-L)1 and [redacted]% receiving other systemic therapy. In the company's economic model, the NIVO+CHEMO and PEMBRO+CHEMO patients would receive only taxanes as subsequent therapy, whereas the CHEMO patients would receive nivolumab. None of those patients would receive a combination therapy. The approach in the economic model is more conservative as patients in the CHEMO arm would highly benefit from a subsequent treatment with nivolumab. In contrast, there would be a slight underestimation of the subsequent treatment effectiveness in the NIVO+CHEMO and PEMBRO+CHEMO arms if all patients would receive taxanes subsequently and none a PD(-L)1 treatment. It should be noted that all but one patient in the NIVO+CHEMO arm that received subsequent systemic therapy received other systemic anticancer therapy so the implications for this treatment arm should be marginal.

The conservative approach chosen overestimates the effectiveness in the CHEMO arm and slightly underestimates the effectiveness in the NIVO+CHEMO and potentially PEMBRO+CHEMO arm leading to a higher ICER.

g) Please conduct an analysis of OS and PFS in both arms of CheckMate 648 adjusting for switching to anti-PD-1/PD-L1 therapies by reference to TSD 16.

Response: Table 25 contains the number of patients whose actual treatment differs from their planned treatment in the overall intention-to-treat (ITT) and the PD-L1 ≥1% population.

From the table, we can see that the extent of treatment switching throughout the CheckMate 648 data is small, with [redacted]% of patients of the ITT population whose actual treatment is different from the originally planned treatment, and similarly [redacted]% of patients from the PD-L1 ≥1% experience this. Therefore, we would suggest that this is unlikely to distort the results to any great degree. Despite there being a larger proportion of treatment switching in the control arm (CHEMO) than in the NIVO+CHEMO arm, the difference is not enough to suggest a high degree of selection bias mentioned in TSD 16. Consequently, we do not believe it is necessary to undertake the additional analysis requested.

Table 25. Number and proportion of patients from each subpopulation whose treatment have switched

Treatment	ITT (n=970)	PD-L1 ≥1% (n=473)
	Proportion of patients who switched treatment	Proportion of patients who switched treatment
CHEMO	[redacted]% (n=20)	[redacted]% (n=12)
NIVO+CHEMO	[redacted]% (n=11)	[redacted]% (n =3)
NIVO+IPI	[redacted]% (n=3)	[redacted]% (n=0)
Any Treatment	[redacted]% (n=34)	[redacted]% (n=15)
CHEMO: chemotherapy; IPI: ipilimumab; NIVO: nivolumab.		

h) Please conduct scenario analyses using adjusted data in the economic model, including variation in the proportion of patients who experience the treatment effect of anti-PD-1/PD-L1 therapies to better reflect NHS clinical practice (see Ouwens M, Darilay A, Zhang Y, Mukhopadhyay P, Mann H,

Ryan J, et al. Assessing the influence of subsequent immunotherapy on overall survival in patients with unresectable stage III non-small cell lung cancer from the PACIFIC study. *Current Therapeutic Research, Clinical and Experimental* 2021;95:100640.)

Response: Since we believe that it is not necessary to undertake analyses to adjust for treatment switching, scenario analyses will not be conducted either.

B3. Document B, Section B.3.2.1.2 of the CS states that: "*Following treatment cessation, patients receive a subsequent line of therapy. As a simplifying assumption, it is assumed that patients may not discontinue this final line of therapy.*" Please justify why the approach of specifying a maximum number of treatment cycles has not been taken, for each subsequent treatment.

Response: This sentence within Document B of the Company Submission is erroneous, patients do discontinue second line therapy. As previously described, due to the PSM approach, time in health state cannot be tracked for subsequent health states. Therefore, a treatment cycle-based approach cannot be used. Instead, second line time on treatment is incorporated for subsequent therapies, reflecting the second line nivolumab OSCC submission.[22] Mean time on treatment is used for subsequent therapies. This data is used to calculate and adjust weekly acquisition and administration subsequent treatment costs accordingly. Hence, although number of treatment cycles cannot be incorporated for subsequent therapies, time on treatment for subsequent therapies is captured.

Patient population

B4. Priority question: Please confirm whether the demographic parameters used in the model (age, proportion of males) are representative for the UK clinical practice.

Response: The median age and proportion male in CheckMate 648[14] and other oesophageal clinical trials are displayed in Table 26. The Cougar-2[15] trial, a UK specific clinical trial, has baseline characteristics that lie close to the baseline characteristics of CheckMate 648[14]. Accordingly, the demographic parameters used in the model, taken from CheckMate 648[14], are representative of the UK clinical practice. Additionally, TA737 utilised data from KEYNOTE-590, with a median age of 62.4 years old, and 83.4% male; both of which are closely aligned to CheckMate 648 data (62.6 years old, 81.8% male).[6] Within the TA737 submission, the ERG agreed that age and proportion male were representative of the target population.[6] This further highlights the generalisability of the demographics used within the company submission herein.

Table 26. Comparison between CheckMate 648 and other OC clinical trials

Trial	Treatment	Age (years)		Proportion male (%)
		Median	Range	
CheckMate 648[14]	**NIVO + CHEMO**	**64**	**40–85**	**79%**
	CHEMO	**62**	**28–81**	**83%**
Cougar-2[15]	Docetaxel	65	28–84	82%
	Active symptom control	66	36–84	80%
KEYNOTE-590[5]	Pembrolizumab + CHEMO	64	28–94	82%
	CHEMO	62	27–89	85%
CheckMate 649[7]	NIVO + CHEMO	62	54-69	68%
	CHEMO	61	53-68	71%
KEYNOTE-062[17]	Pembrolizumab + CHEMO	62	22–83	76%
	CHEMO	62.5	23–87	72%

Interventions and comparators

B5. Priority question: With reference to question A6, please conduct fully incremental cost effectiveness analyses including all relevant comparators for each relevant population.

Response: Cost-effectiveness analyses considering the PD-L1 TC ≥1% and PD-L1 CPS ≥10% patient populations have been conducted, which align to the licenced populations for nivolumab and pembrolizumab, respectively.

The cost effectiveness results for the PD-L1 TC ≥1% population for NIVO-CHEMO versus CHEMO and NIVO-CHEMO versus PEMBRO-CHEMO are presented in Table 27 and Table 28, respectively. Fully incremental analyses are presented in Table 29 and Figure 20.

Please note, these results are calculated using the most recent subsequent treatment costs based on the updated proportion of patients receiving a subsequent treatment as outlined in response to B2.

Table 27. NIVO + CHEMO versus CHEMO, [redacted], 6.9 month cut-point, PD-L1 ≥1%

	NIVO-CHEMO	CHEMO	Incremental
Life years	[redacted]	[redacted]	[redacted]
QALYs	[redacted]	[redacted]	[redacted]
Total costs (£)	[redacted]	[redacted]	[redacted]
ICER (£/QALY)			£33,357
ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year			

Table 28. NIVO + CHEMO versus PEMBRO + CHEMO, [redacted] DBL, 6.9 month cut-point, PD-L1 ≥1%

	NIVO-CHEMO	PEMBRO + CHEMO	Incremental
Life years	[redacted]	[redacted]	[redacted]
QALYs	[redacted]	[redacted]	[redacted]
Total costs (£)	[redacted]	[redacted]	[redacted]
ICER (£/QALY)			-£5,594
ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year			

Table 29. Fully incremental analysis, base case population (NIVO+CHEMO vs. PEMBRO+CHEMO vs CHEMO, [redacted] DBL

Treatment	Total costs (discounted, £)	Total QALYs (discounted)	ICER (£/QALY)/result
NIVO-CHEMO	[redacted]	[redacted]	-
PEMBRO + CHEMO	[redacted]	[redacted]	£29,204*
CHEMO	[redacted]	[redacted]	Dominated
*ICER versus CHEMO			

[redacted]

Figure 20. Cost-efficiency frontier, base case population (NIVO+CHEMO vs. PEMBRO+CHEMO vs CHEMO, [redacted] DBL

Treatment effectiveness and extrapolation

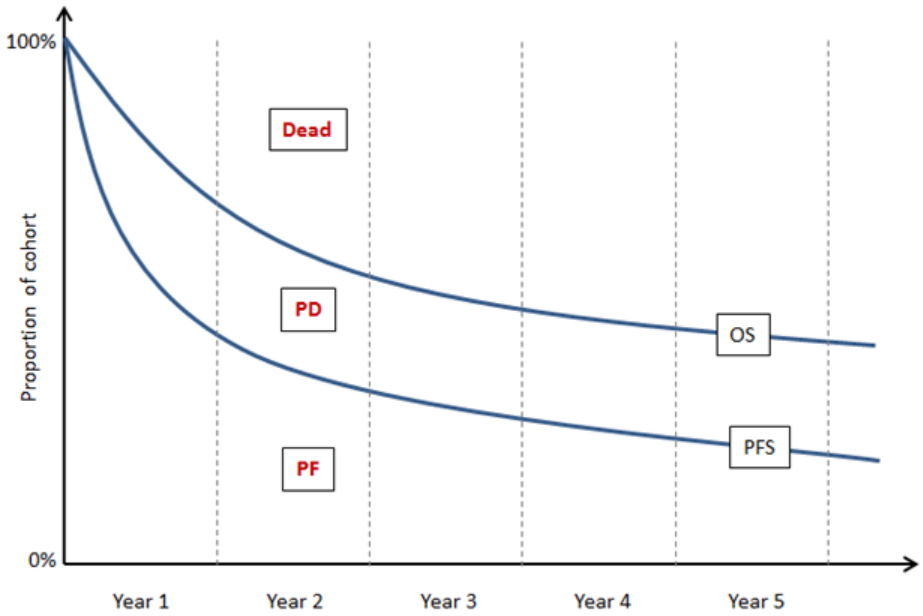
B6. Priority question: Document B, Section B.3.3 of the CS entitled "*Clinical parameters and variables*" does not provide an overview of the clinical parameters and variables used in the economic model.

a) Please provide an overview of all transition probabilities used in the model with sources.

b) Please justify the sources and calculations used to inform the transition probabilities in the model.

Response for a) and b): The economic model is a partitioned survival model. As such, health state occupancy for progression-free, progressed disease, and death is determined solely via PFS and OS curves (see Figure 21). There are no explicit transition probabilities between states. The choice of PFS and OS curves within the base case are described within section B.3.3 of the Company Submission.

Figure 21. Overview of PSM method



OS: Overall survival; PD: Progressed disease; PF: Progression-free; PFS: Progression-free survival

B7. Priority question: Document B, Section B.3.3.1.1 of the CS states that a semi-parametric approach with Kaplan-Meier data until 6.9 months was used for extrapolating OS. It also appears to be the case that the choice survival extrapolation approach within the economic model does not include the full range of models considered for both OS and PFS.

a) Please provide a justification for why precisely 6.9 months was chosen.

Response: As discussed in the survival appendix (Appendix N, Section 4.2.3): "A number of potential cut points were considered, avoiding assessment windows due to the rapid change in hazard near the model start time implied by these periods. As a compromise the between maximation of data for use in extrapolation and removal of the largest hazard discontinuities, a time of 6.9 months was chosen. This timepoint avoids the sharp change in hazard observed in the first six months for NIVO+CHEMO and CHEMO."

b) Given the apparent inflexion point of about 6 months in the smoothed hazard plot for chemotherapy, please provide scenario analyses for later cut-points, including 12 months and 20 months (minimum follow-up in the trial).

Response: Figure 22 to Figure 25 present the outcomes at a 12 month cut-point. It is worth noting that the company is hesitant to suggest that PFS outcome is significantly meaningful as there are next to no events after the 12 month cut-point; this is especially apparent in the CHEMO arm. Consequently, the 20 month cut-point is deemed unmeaningful and analysis is not presented.

The clinical plausibility and goodness of fit statistics (AIC and BIC) were considered in deciding on which semi-parametric extrapolation is the most appropriate.

For overall survival, the lognormal presents the best statistical-fit of clinically plausible models in the CHEMO arm and Gompertz presented the best statistical-fit of clinically plausible models in the NIVO+CHEMO arm.

For BICR-assessed progression-free survival, in both CHEMO and NIVO+CHEMO arms, the best fitting model per statistical fit is the lognormal.

Figure 22. CheckMate 648 for patients with tumour cell PD-L1 ≥1% from [redacted] DBL, CHEMO: Semi-parametric OS models overlaid upon Kaplan-Meier – 12 month cut point

Figure 23. CheckMate 648 for patients with tumour cell PD-L1 ≥1% from [redacted] DBL, NIVO+CHEMO: Semi-parametric OS models overlaid upon Kaplan-Meier – 12 month cut point

Figure 24. CheckMate 648 PD-L1 ≥1% [redacted] DBL, CHEMO: Semi-parametric PFS (BICR) models overlaid upon Kaplan-Meier – 12 months cut point

Figure 25. CheckMate 648 PD-L1 ≥1% [redacted] DBL, NIVO+CHEMO: Semi-parametric PFS (BICR) models overlaid upon Kaplan-Meier – 12 months cut point

c) **Please provide a cost-effectiveness model that permits choice of cut-point for OS as well as all fully parametric and semi-parametric models for both OS and PFS.**

Response: An updated cost-effectiveness model will be provided with this response.

B8. Priority question: The ERG notes that there are no plots of hazard ratio over time between nivolumab plus chemotherapy and any comparator.

a) **Please provide plots of hazard ratios over time from the smoothed hazards from the Kaplan-Meier data for nivolumab plus chemotherapy versus all comparators, including chemotherapy and pembrolizumab plus chemotherapy.**

Response: The following figures show the hazard ratio plots over time from the smoothed hazards from the Kaplan-Meier data of the CheckMate 648 trial data.

For overall survival (OS), the hazard ratio increases over time implying that the likelihood of patients on the CHEMO arm as opposed to those on the NIVO+CHEMO arm having an OS event decreases steadily over time but always remains the more likely arm for an event to occur on.

The opposite is true for the PFS outcome, where the hazard ratio decreases over time implying that over the course of the study, the likelihood that a PFS event occurs in the CHEMO arm rather than the NIVO+CHEMO arm increases.

Figure 26. CheckMate 648 PD-L1 ≥1% [redacted] DBL hazard ratio over time plot of Overall Survival: CHEMO vs NIVO+CHEMO.

Figure 27. CheckMate 648 PD-L1 ≥1% [redacted] DBL, hazard ratio over time plot of Progression-free survival (BICR-assessed): CHEMO vs NIVO+CHEMO

b) Please provide HR plots over time for all extrapolations (parametric and semi-parametric).

Response: Plots that demonstrate an initially high hazard ratio highlight that the hazard of the control arm (CHEMO arm) is near to zero at the beginning of the trial data.

Figure 28. CheckMate 648 PD-L1 ≥1% [redacted] DBL, hazard ratio of exponential distribution over time plot of OS: CHEMO vs NIVO+CHEMO[redacted]

Figure 29. CheckMate 648 PD-L1 ≥1% [redacted] DBL, hazard ratio of Weibull distribution over time plot of OS: CHEMO vs NIVO+CHEMO[redacted]

Figure 30. CheckMate 648 PD-L1 ≥1% [redacted] DBL, hazard ratio of log-logistic distribution over time plot of OS: CHEMO vs NIVO+CHEMO[redacted]

Figure 31. CheckMate 648 PD-L1 ≥1% [redacted] DBL, hazard ratio of lognormal distribution over time plot of OS: CHEMO vs NIVO+CHEMO[redacted]

Figure 32. CheckMate 648 PD-L1 ≥1% [redacted] DBL, hazard ratio of Gompertz distribution over time plot of OS: CHEMO vs NIVO+CHEMO[redacted]

Figure 33. CheckMate 648 PD-L1 ≥1% [redacted] DBL, hazard ratio of Generalised Gamma distribution over time plot of OS: CHEMO vs NIVO+CHEMO[redacted]

Figure 34. CheckMate 648 PD-L1 ≥1% [redacted] DBL, hazard ratio of exponential distribution over time plot of PFS (BICR assessed): CHEMO vs NIVO+CHEMO [redacted]

Figure 35. CheckMate 648 PD-L1 ≥1% [redacted] DBL, hazard ratio of Weibull distribution over time plot of PFS (BICR assessed): CHEMO vs NIVO+CHEMO[redacted]

Figure 36. CheckMate 648 PD-L1 ≥1% [redacted] DBL, hazard ratio of log-logistic distribution over time plot of PFS (BICR assessed): CHEMO vs NIVO+CHEMO[redacted]

Figure 37. CheckMate 648 PD-L1 ≥1% [redacted] DBL hazard ratio of lognormal distribution over time plot of PFS (BICR assessed): CHEMO vs NIVO+CHEMO[redacted]

Figure 38. CheckMate 648 PD-L1 ≥1% [redacted] DBL, hazard ratio of Gompertz distribution over time plot of PFS (BICR assessed): CHEMO vs NIVO+CHEMO [redacted]

Figure 39. CheckMate 648 PD-L1 ≥1% [redacted] DBL, hazard ratio of generalised gamma distribution over time plot of PFS (BICR assessed): CHEMO vs NIVO+CHEMO
[redacted]

For the semi-parametric extrapolations, a cut-point of 6.9 months was used to which the survival models switched from the trial data to a parametric extrapolation.

Figure 40. CheckMate 648 PD-L1 ≥1% [redacted] DBL hazard ratio of semi-parametric (6.9 month cut-point) exponential distribution for OS: CHEMO vs NIVO+CHEMO

Figure 41. CheckMate 648 PD-L1 ≥1% [redacted] DBL hazard ratio of semi-parametric (6.9 month cut-point) Weibull distribution for OS: CHEMO vs NIVO+CHEMO

Figure 42. CheckMate 648 PD-L1 ≥1% [redacted] DBL hazard ratio of semi-parametric (6.9 month cut-point) log-logistic distribution OS: CHEMO vs NIVO+CHEMO

[redacted]

Figure 43. CheckMate 648 PD-L1 ≥1% [redacted] DBL hazard ratio of semi-parametric (6.9 month cut-point) lognormal distribution for OS: CHEMO vs NIVO+CHEMO

Figure 44. CheckMate 648 PD-L1 ≥1% [redacted] DBL hazard ratio of semi-parametric (6.9 month cut-point) Gompertz distribution for OS: CHEMO vs NIVO+CHEMO.

Figure 45. CheckMate 648 PD-L1 ≥1% [redacted] DBL hazard ratio of semi-parametric (6.9 month cut-point) generalised gamma distribution for OS: CHEMO vs NIVO+CHEMO.

[redacted]

Figure 46. CheckMate 648 PD-L1 ≥1% [redacted] DBL hazard ratio of semi-parametric (6.9 month cut-point) exponential distribution for PFS (BICR-assessed): CHEMO vs NIVO+CHEMO.

Figure 47. CheckMate 648 PD-L1 ≥1% [redacted] DBL hazard ratio of semi-parametric (6.9 month cut-point) Weibull distribution for PFS (BICR-assessed): CHEMO vs NIVO+CHEMO.

Figure 48. CheckMate 648 PD-L1 ≥1% [redacted] DBL hazard ratio of semi-parametric (6.9 month cut-point) log-logistic distribution for PFS (BICR-assessed): CHEMO vs NIVO+CHEMO.

Figure 49. CheckMate 648 PD-L1 ≥1% [redacted] DBL hazard ratio of semi-parametric (6.9 month cut-point) lognormal distribution for PFS (BICR-assessed): CHEMO vs NIVO+CHEMO.

[redacted]

Figure 50. CheckMate 648 PD-L1 ≥1% [redacted] DBL hazard ratio of semi-parametric (6.9 month cut-point) Gompertz distribution for PFS (BICR-assessed): CHEMO vs NIVO+CHEMO.

[redacted]

Figure 51. CheckMate 648 PD-L1 ≥1% [redacted] DBL hazard ratio of semi-parametric (6.9 month cut-point) generalised gamma distribution for PFS (BICR-assessed): CHEMO vs NIVO+CHEMO.

[redacted]

c) Please discuss the validity of the choice of most appropriate extrapolation in the context of the comparison with the hazard ratios from the smoothed hazards.

Response: Comparing the plots of the hazard ratio over time of the parametric extrapolations to the hazard ratios from the smoothed hazards, the Gompertz, the generalised gamma and the log-logistic distributions give the most similar hazard ratio profile. This being relatively a monotonically increasing hazard ratio. The exponential produces a constant hazard ratio and so does not follow the trend of the hazard ratio as well as the aforementioned distributions. The Weibull demonstrates a decreasing hazard ratio and is not similar at all to the hazard ratio derived from the smoothed hazard.

In comparing the semi-parametric plots of the hazard ratio over time, for OS we see that the general trend of the hazard ratios of the Weibull, log-logistic and lognormal follow the trend set by the hazard ratio over time plot from the smoothed hazards the closest.

Cost and resource use

B9. Document B, Section B.3.5.1.2 of the CS sates that, "*The frequency of resource use in each health state has been sourced through the literature using TA737*."

a) Please elaborate on this statement.

Response: Healthcare resource use frequency has been sourced from TA737[6], whose resource use was sourced via expert opinion in TA378,[28] accounting for the ERG's comments in TA737 that the post-progression resource use should also be aligned to that of TA378. Note that no treatment-specific healthcare resource use is used.

b) Please provide a table of the frequency and cost of administration for all interventions.

Response: The frequency and cost of administration are provided within the company submission Tables 46, 47, and 48. These tables are reproduced below (Table 30 to Table 32).

Table 30. Drug acquisition and administration unit costs for nivolumab in combination with chemotherapy (fluorouracil plus cisplatin)

	Nivolumab	Fluorouracil	Cisplatin	Source
Dosing regimen	240 mg, on day 1 every 2 weeks	800 mg/m^2, on day 1 through day 5 every 4 weeks	80 mg/m^2, on day 1 every 4 weeks	CheckMate 648 trial[29]
Dose received	240 mg	1,331 mg (6,656 mg over 5 days)	133 mg	Assuming body surface area of 1.66m^2, calculated using CheckMate 648 data[29]
Unit cost	£2,633.00 (PAS cost: [redacted])	£15.64	£14.11	Table 44 of CS
Admin method	Intravenous as a 30 minute infusion on day 1 and day 15 of each 28 day cycle	Intravenous continuous infusion on days 1–5 of 28 day cycle	Intravenous as a 30–120 minute infusion on day 1 of 28 day cycle	CheckMate 648 trial[29]
Day 1 administration cost	£431.72			Table 44 of CS
Day 15 administration cost	£284.05			Table 44 of CS
PD-L1 test cost	£42.61			TA737[6]
All therapies assume wastage.				

Table 31. Drug acquisition and administration unit costs for chemotherapy (fluorouracil plus cisplatin)

	Fluorouracil	Cisplatin	Source
Dosing regimen	800 mg/m^2, on day 1 through day 5 every 4 weeks	80 mg/m^2, on day 1 every 4 weeks	CheckMate 648 trial[29]
Dose received	1,331 mg (6,656 mg over 5 days)	133 mg	Assuming body surface area of 1.66m^2, calculated using CheckMate 648 data[29]
Unit cost	£15.64	£14.11	Table 45 of CS
Admin method	Intravenous continuous infusion on days 1–5 of 28 day cycle	Intravenous as a 30–120 minute infusion on day 1 of 28 day cycle	CheckMate 648 trial[29]
Day 1 administration cost	£431.72		Table 44 of CS
All therapies assume wastage.			

Table 32. Drug acquisition and administration unit costs for pembrolizumab in combination with chemotherapy (fluorouracil plus cisplatin)

	Pembrolizumab	Fluorouracil	Cisplatin	Source
Dosing regimen	200 mg, on day 1 every 3 weeks	800 mg/m^2, on day 1 through day 5 every 3 weeks	80 mg/m^2, on day 1 every 3 weeks	KEYNOTE-590[30]
Dose received	200 mg	1,331 mg (6,656 mg over 5 days)	133 mg	KEYNOTE-590[30]
Unit cost	£5,260.00	£15.64	£14.11	Table 45 of CS
Admin method	Intravenous as a 30 minute infusion on day 1 each 21 day cycle	Intravenous continuous infusion on days 1–5 of 21 day cycle	Intravenous as infusion on day 1 of 21 day cycle	KEYNOTE-590[30]
Day 1 administration cost	£431.72			Table 44 of CS
PD-L1 test cost	£42.61			TA737[6]
All therapies assume wastage.				

B10. Priority question: Please provide the currency codes, descriptions, and settings for all unit costs sourced from NHS reference costs.

Response: The tables below outline the currency code, description, and settings for administration costs (Table 33), healthcare resource use costs (Table 34), and adverse event costs (Table 35).

Table 33. Administration costs

Details	Mean value	Currency code	Description	Setting
Deliver Simple Parenteral Chemotherapy at First Attendance	£284.05	SB12Z[31]	Deliver Simple Parenteral Chemotherapy at First Attendance	Chemotherapy: Weighted average of day case and reg day/night, outpatient and other
Deliver Complex Chemotherapy, including Prolonged Infusion Treatment, at First Attendance	£431.72	SB14Z[31]	Deliver Complex Chemotherapy, including Prolonged Infusion Treatment, at First Attendance	Chemotherapy: Day case and reg day/night

Table 34. Resource use costs

Resource Use	Cost[32]	Currency code	Description	Setting
CT scan	£103.31	RD25Z[31]	Computerised Tomography Scan of Three Areas, without Contrast	Diagnostic imaging: Weighted average of Imaging: Direct Access, Imaging: Outpatient and Imaging: Other
Blood test	£2.53	DAPS05[31]	Haematology	Directly accessed pathology services
Kidney	£33.80	WH15Z[31]	Special Screening, Examinations or Other Genetic Disorders	Directly accessed diagnostic services
Hepatic	£33.80	WH15Z[31]	Special Screening, Examinations or Other Genetic Disorders	Directly accessed diagnostic services
Consultant	£203.14	WF01A[31]	Non-Admitted Face-to-Face Attendance, Follow-up	Consultant led: Medical oncology
CT: computed tomography				

Table 35. Adverse event costs

Adverse event (AE)	AE cost (SE)	Currency code	Description	Setting
Vomiting	£471.95 (£94.39)	FD10M[32]	Non-Malignant Gastrointestinal Tract Disorders without Interventions, with CC Score 0-2	Non-Elective Short Stay
Hyponatraemia	£1,164.14 (£232.83)	KC05H[32]	Fluid or Electrolyte Disorders, with Interventions, with CC Score 0-4	Non-Elective Short Stay
Pneumonitis	£1,909.33 (£381.87)	Weighted average DZ111K,L,M,N,P, Q,R,S,T,U,V[32]	Lobar, Atypical or Viral Pneumonia, with Multiple Interventions, with CC Score 0-8, 9-13 and 14+. Lobar, Atypical or Viral Pneumonia, with Single Intervention, with CC Score 0-7, 8-12 and 13+. Lobar, Atypical or Viral Pneumonia, without Interventions, with CC Score 0-3, 4-6, 7-9, 10-13 and 14+.	Total Healthcare Resource Groups
Hepatic function abnormal	£2,461.04 (£492.21)	Weighted average GC01C,D,E,F[32]	Liver Failure Disorders with Multiple Interventions. Liver Failure Disorders with Single Intervention. Liver Failure Disorders without Interventions, with CC Score 0-4 and 5+.	Total Healthcare Resource Groups
Acute kidney injury	£1,961.20 (£392.24)	Weighted average LA07H,J,K,L,M,N, P[32]	Acute Kidney Injury with Interventions, with CC Score 0-5, 6-10 and 11+. Acute Kidney Injury without Interventions, with CC Score 0-3, 4-7, 8-11 and 12+.	Total Healthcare Resource Groups
Nausea	£471.95 (£94.39)	FD10M[32]	Non-Malignant Gastrointestinal Tract Disorders without Interventions, with CC Score 0-2	Non-Elective Short Stay
Dehydration	£1,329.93 (£265.99)	Weighted average KC05G,H,J,K,L,M ,N[32]	Fluid or Electrolyte Disorders, with Interventions, with CC Score, 0-4 and 5+. Fluid or Electrolyte Disorders, without Interventions, with CC Score 0-1, 2-3, 4-6, 7-9 and 10+	Total Healthcare Resource Groups
HRGs: Healthcare Resource Groups; NES: Non-Elective Short Stay				

B11. Priority question: Please clarify/ justify the following aspects regarding the costs/ resource use in the economic analysis. Please:

a) Justify the appropriateness of the numbers for resource use (monitoring) frequency in the post-progression state.

Response: Per the ERG's request in TA737,[6] the monitoring frequencies described in TA378 for the post-progression health state are employed to calculate the post-progression health state cost.[28] These monitoring costs are displayed in Table 36.

Table 36. Post progression monitoring frequencies

Resource Use	Weekly frequency post-progression[6,32]
CT scan	0.08
Blood test	1.00
Kidney	1.00
Hepatic	1.00

b) Justify choice of gamma distribution for NHS reference costs (average for a cohort) over normal distribution.

Response: The gamma distribution is recommended for sampling distribution for costs in the literature.[33,34]

c) Clarify if the unit costs assigned to chemotherapy administrations in Table 44 of the CS are based on the expectation that administration would take place in a day case setting.

Response: The administration cost associated with 'Deliver Simple Parenteral Chemotherapy at First Attendance' is the weighted average of 'Day case and Reg Day/Night', 'Outpatient' and 'Other' for SB12Z from the NHS reference costs.[35] The administration cost associated with 'Deliver Complex Chemotherapy, including Prolonged Infusion Treatment, at First Attendance' is SB14Z day case and reg day/night cost.[35]

d) Provide a table of administration costs applied in the model for the intervention, comparator and subsequent treatments, with columns for

resource, type of administration, NHS Reference code, setting and unit cost.

Response: The administration costs associated with NIVO+CHEMO, CHEMO and PEMBRO+CHEMO are presented in Table 37 and Table 38. Also, please note that a PD-L1 test cost is required upon treatment initiation for NIVO+CHEMO and PEMBRO+CHEMO. This PD-L1 test cost is £42.61 and is applied in the first cycle these treatments are given. The administration costs associated with subsequent treatment are presented in Table 39.

Table 37. NIVO-CHEMO administration costs

Subsequent treatment	Administration type	NHS reference cost	Setting	Unit cost	Note
Nivolumab	Deliver Simple Parenteral Chemotherapy at First Attendance	SB12Z	Weighted average of Day case and Reg Day/Night, Outpatient and Other	£284.05	Admin cost for nivolumab is required every time the treatment is not given in combination with CHEMO. Nivolumab admin cost is captured in fluorouracil admin cost when given in combination.
Cisplatin	-	£0	-	-	Cisplatin admin cost is captured in fluorouracil admin costs as treatments are always given in combination.
Fluorouracil	Deliver Complex Chemotherapy, including Prolonged Infusion Treatment, at First Attendance	£431.72	Day case and Reg Day/Night	£431.72	-

Table 38. CHEMO administration costs

Subsequent treatment	Administration type	NHS reference cost	Setting	Unit cost	Note
Pembrolizumab	-	£0	-	-	Pembrolizumab admin cost is captured in fluorouracil admin costs as treatments are always given in combination.
Cisplatin	-	£0	-	-	Cisplatin admin costs are captured in fluorouracil admin costs as treatments are always given in combination.
Fluorouracil	Deliver Complex Chemotherapy, including Prolonged Infusion Treatment, at First Attendance	£431.72	Day case and Reg Day/Night	£431.72	

Table 39. Subsequent treatment administration costs

Subsequent treatment	Administration type	NHS reference cost	Setting	Unit cost
Nivolumab	Deliver Simple Parenteral Chemotherapy at First Attendance	SB12Z	Weighted average of Day case and Reg Day/Night, Outpatient and Other	£284.05
Taxane: docetaxel				
Taxane: paclitaxel				

e) Clarify which resources are specifically associated with the monitoring requirements of each treatment.

Response: Resource use is health state specific. Therefore, monitoring costs are associated with the patient's health state, not the treatment they are receiving. This is in keeping with TA737.[6]

f) Clarify if the initial cycle cost of administration per model cycle applied in the economic model is the same for all subsequent cycles.

Response: For primary treatments, the 'initial administration costs' are applied every time the treatment is administered (i.e. at initial and subsequent treatment cycles). For example, cisplatin and fluorouracil are administered on day one every 28 days. Accordingly, the initial administration cost of £431.72 is applied every 28 days whilst a patient remains on treatment.

However, due to the inability of a PSM to track individual patients through subsequent lines of treatment, an average cyclical cost has been used for subsequent treatments. This average cyclical cost takes into account both the treatment costs and the administration cost over the treatment cycle, which is applied to every patient receiving the subsequent treatment in a modelled cycle (1-week). Again, using docetaxel as an example, the treatment and administration cost is required every second week. Therefore, to create an average cyclical docetaxel cost, docetaxel's treatment and administration cost must be summed and divided by the treatment cycle (2-weeks).

B12. Priority question: A 'treatment modifier' was applied in the economic model.

a) Please clarify if this is equal to number of occasions where a dose was delayed divided by total number of doses administered.

Response: The treatment modifier is one minus the number of doses delayed divided by the total number of doses received.

b) Please discuss assumptions of dose intensity (using a treatment modifier) for the pembrolizumab plus chemotherapy arm.

Response: There was no data to inform the treatment modifier for the pembrolizumab plus chemotherapy arm (the treatment modifier for pembrolizumab plus chemotherapy was redacted in TA737). Accordingly, the treatment modifier for pembrolizumab plus chemotherapy had to be assumed.

One plausible assumption was to assign a treatment modifier of 1. Under this assumption, no pembrolizumab plus chemotherapy doses are delayed; all doses expected to be received will be received. However, this creates an artificially high treatment cost for the pembrolizumab plus chemotherapy arm, biasing the model results towards the treatment arm. The conservative assumption that the pembrolizumab plus chemotherapy arm has the same treatment modifier as nivolumab plus chemotherapy has been employed to avoid this bias. That being, pembrolizumab, fluorouracil and cisplatin have equivalent treatment modifiers to nivolumab, fluorouracil and cisplatin, respectively. The pembrolizumab plus chemotherapy treatment modifiers displayed in Table 40.

Table 40. Pembrolizumab plus chemotherapy treatment modifier

Treatment		Treatment modifier	Source
Pembrolizumab in combination with chemotherapy (fluorouracil plus cisplatin)	Pembrolizumab	[redacted]	Assumed equivalent to NIVO+CHEMO due to lack of data
	Fluorouracil	[redacted]	
	Cisplatin	[redacted]	

c) **Please clarify that the cost of each drug is adjusted by multiplying by this modifier.**

Response: The cost of each treatment, including both the acquisition cost and the administration cost of treatment, is adjusted by multiplying by the treatment modifier.

d) **Please provide the justification of this precise amount of cost reduction.**

Response: The reduction in doses given and the associated cost is used to match the proportion of patients missing doses for various reasons, for example including co-morbidities, adverse events, patient non-compliance, appointment cancellations. These can be considered reflective of clinical practice and is aligned to the SmPC recommendations on managing adverse events. This approach is common practice in HTAs, and was used in all recent gastro-oesophageal cancer NICE HTAs, including nivolumab for previously treated unresectable, advanced OC (TA707)[22] and pembrolizumab with chemotherapy for untreated advanced OC/GOJC (TA737).[6]

e) **Please explain how this relates to the precise amount of drug delivered in the trial and how this would compare with NHS clinical practice.**

Response: Please see the response to d) above

B13. Please discuss what assumptions have been made for chemotherapy treatments for which dosage is based on body surface area (BSA), to calculate the weighted average cost per dose.

a) **If the standard approach of using the BSA for the deterministic analyses and variations in BSA based on the standard errors (SEs) for the probabilistic analyses has not been used, please justify approach.**

b) **If the standard approach suggested in a) was not used, please include an option in the model to do so.**

Response: To illustrate why body surface area (BSA) has not been included in the model, the BSA has been set at an extreme value of 2. The costs for CHEMO under this extreme BSA value have then been calculated and applied in the model. Table 41 demonstrates how the CHEMO costs change under the extreme BSA.

Table 41. CHEMO treatment costs under original and extreme BSA

Treatment	Treatment cost under original BSA – 1.66m^2	Treatment cost under extreme BSA – 2m^2
Fluorouracil	£15.64	£19.18
Cisplatin	£14.11	£17.46

When the CHEMO treatment costs under the extreme BSA are applied within the model, that being applied to both NIVO+CHEMO and CHEMO, the base case ICER changes by a negligible amount (an increase of £33). The reason for this limited impact on the ICER is twofold. Firstly, in comparison to other costs applied in the model (i.e. the nivolumab treatment cost and fluorouracil administration cost), the treatment costs for CHEMO are small. Therefore, any changes to these costs have a limited impact on the ICER. Secondly, the changes in costs are applied to both the treatment and control arm as both arms require CHEMO and in the same dosage. Therefore, the treatment cost increases by the same amount for both the treatment and control arm. Note that the only reason the ICER changes at all due to the increased CHEMO costs is that patients remain on first-line treatment for longer in the treatment arm. This negligible impact on the ICER under the extreme BSA provides the rationale for BSAs exclusion from the model.

B14. Document B, Section B.3.5.1.1 of the CS states that, "*Additionally, patients in the nivolumab treatment arm are assumed to receive a cost for PD-L1 testing…*"

a) Table 47 does not show that this one-off test cost was applied to all arms. Please clarify if this cost was solely applied to both nivolumab plus chemotherapy and pembrolizumab plus chemotherapy arms.

Response: The PD-L1 test cost was applied to both NIVO+CHEMO and PEMBRO+CHEMO arms within the economic model, but not the chemotherapy arm (reflected in Table 47, Table 48, and Table 49 within the company submission).

b) If yes to a) then please justify this decision.

Response: The PD-L1 test cost was applied to both of these immunotherapies, since they are for a specific PD-L1 positive population (as opposed to chemotherapy which is for the ITT population)

c) **If no to a) please provide a scenario (implemented also in the model), in which the costs for diagnostic testing of PD-L1 status in adults with untreated unresectable metastatic oesophageal squamous cell cancer (OSCC) is applied to both immunotherapy arms.**

Response: Not applicable

Health-related quality of life (HRQoL)

B15. Priority question: Health state utilities were estimated without any on or off nivolumab plus chemotherapy treatment adjustment.

a) **Please estimate a regression model with a covariate for whether on or off treatment.**

Response: Since the PFS time for CheckMate 648 is nearly complete and nearly all the time on treatment is within PFS for both treatment arms (Figure 52 and Figure 53), the difference in the health state utilities on or off nivolumab plus chemotherapy treatment is expected to be small. Therefore, it is irrelevant how the data is split, the estimate of mean utility in PFS is likely to be similar. The approach chosen in the CS is conservative with any negative effects of treatment are captured within the utility analysis.

Figure 52. Overlay PFS and time on treatment for CHEMO arm of CheckMate 648

Figure 53. Overlay of PFS and time on treatment for NIVO+CHEMO of CheckMate 648

b) **Please present all statistical tests for this regression model.**

Response: Since a regression analysis was deemed not necessary, no statistical tests will be presented.

c) **Please conduct a scenario analysis incorporating this adjustment in the economic model.**

Response: Since a regression analysis was deemed not necessary, a scenario analysis in the economic model was not conducted

B16. Appendix O of the CS details that utilities were calculated using two approaches: a progression-based and time-to-death health state model where a mixed model approach was used to account for repeated EuroQol-5 dimension-3 level (EQ-5D-3L) measurements per patient within a health state in estimating the mean values of EQ-5D-3L for each health state in the utility analysis. Please refer to NICE TSD 8 in commenting on the appropriateness of analysis methods and validity of estimates.

Response: An updated utility analysis report is provided with this response. The recommendations of TSD 8 are discussed below.

- *The QALY is the measure of the benefit of treatment*; by the derivation of appropriate health state utility values for the health states used in the economic model, the product of time in state and HSUV provides an estimate of QALYs accrued. Derivation of the HSUVs as described in the utility analysis appendix is compliant with this recommendation.

- *Patient self-report should be used to describe the change in health*; by the use of logitudinal patient reported outcomes from a randomised controlled trial, the analysis is compliant with this recommendation.

- *The EQ-5D should be used to collect data from patients on their health, and a set of values obtained from the UK general population using the time-trade off method applied to generate health-related utilities*; the EQ-5D-3L instrument was used in CheckMate 648 and the dimensions used to estimate utility values via the time-trade off tariff of Dolan et el.{Dolan, 1997 #264} The analysis is compliant with this recommendation.

- *Where it is important, the impact of an intervention on carers can be included and measured using the EQ-5D*; Data were not collected in CheckMate 648 to inform the impact upon carers. Health effects for carers have not been estimated in this submission. This is consistent with the scope of other

contemporary technology appraisals, e.g. TA747. Due to precedence, and as this perspective was not nominated at scoping, the analysis is compliant with this recommendation, as the measurement has not been considered important by the parties involved in the TA.

- *Other preference-based measures of health can be included in sensitivity analysis, if they have been included in the clinical trial/s used to inform the effectiveness estimates*. No additional preference-based measures of health were taken in CheckMate 648, the analysis is compliant with this recommendation as it is not relevant.

- *Consider using an instrument developed for use in children when obtaining health state utility values*. The analysis is compliant with this recommendation as it is not relevant.

Adverse reactions

B17. Document B, Section B.3.4.3 of the CS states that: "*The ten most frequently occurring treatment-related grade 3–4 serious AEs were included in the economic model.*"

Please confirm whether the source of Grade 3/4 treatment-related adverse events (AEs) type and frequency is the nivolumab plus chemotherapy arm in the CheckMate 648 trial.

Response: The company confirms that the source of grade 3-4 adverse event incidence within the economic model (for NIVO+CHEMO and CHEMO arms) was CheckMate 648. For the comparison with PEMBRO+CHEMO, adverse event incidence was sourced from KEYNOTE-590.[30] Adverse event incidence for each comparator are summarised in the table responding to clarification question B18.

Sensitivity analyses

B18. The naming of parameters explored in the probabilistic sensitivity analysis (PSA) is unclear.

a) Please provide a table overview of all parameters used in the model including descriptions, and highlight those that were used in the PSA.

Response: A summary of the parameters included in the PSA is presented in Table 42 to Table 45.

Table 42. Summary of model settings, survival and progression functions applied in the economic model

Variable	Mean value	SE value (if applicable) or uncertainty measurement	Measurement of uncertainty and distribution	Section in Document B
Model settings				
Cycle length	1 week	NA	NA	B.3.2.1
Time horizon	2,080 weeks (40 years)	NA	PSA: NA DSA: 260 to 520 weeks	B.3.2.1
Discounting rate (costs, outcomes)	3.5%	NA	PSA: NA DSA: 0% to 6% costs, 0% to 6% outcomes	B.3.2.1
Baseline parameters				
% Male	[redacted]	[redacted]	PSA: normal distribution DSA: 0% to 100%	B.3.2.2
Age	[redacted]	[redacted]	PSA: beta distribution DSA: 80% to 120% of mean	B.3.2.2
Survival and progression functions				
Overall survival: NIVO + CHEMO	Semi-parametric 6.9 month cut point, log-normal	Confidence intervals	PSA: Described in Section B.3.3.1 DSA: NA	B.3.3.1
Overall survival: CHEMO	Semi-parametric 6.9 month cut point, log-normal	Confidence intervals		
Progression-free survival: NIVO + CHEMO	Semi-parametric 6.9 month cut point, generalised gamma	Confidence intervals		
Progression-free survival: CHEMO	Semi-parametric 6.9 month cut point, Weibull	Confidence intervals		
All-cause mortality	Based on UK lifetables	NA	NA	B.3.3.1.4

AE: adverse events; DSA: deterministic sensitivity analysis; NA: not applicable; PSA: probabilistic sensitivity analysis; SE: standard error.

Table 43. Summary of clinical parameters applied in the economic model

Variable	Mean value	SE value (if applicable) or uncertainty measurement	Measurement of uncertainty and distribution	Section in Document B
Clinical parameters				
First line: Time on treatment	KM data both arms	Confidence intervals	PSA: Described in Section B.3.3.2 DSA: NA	B.3.3.2
Second line: Time on treatment weighted taxane, cyclical discontinuation rate	0.0610	0.0061	PSA: beta distribution DSA: 80% to 120% of mean	B.3.3.2
Second line: Time on treatment weighted, cyclical discontinuation rate	0.0561	0.0056	PSA: beta distribution DSA: 80% to 120% of mean	B.3.3.2
AE incidence NIVO + CHEMO				
Vomiting	[redacted]	[redacted]	PSA: beta distribution DSA: 80% to 120% of mean	B.3.3.3
Hyponatraemia	[redacted]	[redacted]		
Pneumonitis	[redacted]	[redacted]		
Hepatic function abnormal	[redacted]	[redacted]		
Adrenal insufficiency	[redacted]	[redacted]		
Acute kidney injury	[redacted]	[redacted]		
Colitis	[redacted]	[redacted]		
Nausea	[redacted]	[redacted]		
Dehydration	[redacted]	[redacted]		
Febrile neutropenia	[redacted]	[redacted]		
AE incidence chemotherapy				
Vomiting	[redacted]	[redacted]	PSA: beta distribution DSA: 80% to 120% of mean	B.3.3.3
Hyponatraemia	[redacted]	[redacted]		
Pneumonitis	[redacted]	[redacted]		
Hepatic function abnormal	[redacted]	[redacted]		
Adrenal insufficiency	[redacted]	[redacted]		
Acute kidney injury	[redacted]	[redacted]		
Colitis	[redacted]	[redacted]		
Nausea	[redacted]	[redacted]		
Dehydration	[redacted]	[redacted]		
Febrile neutropenia	[redacted]	[redacted]		

AE: adverse events; DSA: deterministic sensitivity analysis; NA: not applicable; PSA: probabilistic sensitivity analysis; SE: standard error.

Table 44. Summary of utilities and disutilities applied in the economic model

Variable	Mean value	SE value (if applicable) or uncertainty measurement	Measurement of uncertainty and distribution	Section in Document B
Utilities				
Pre-progression health state utility	[redacted]	[redacted]	PSA: beta distribution DSA: 80% to 120% of mean	B.3.4.4
Post-progression health state utility	[redacted]	[redacted]	PSA: beta distribution DSA: 80% to 120% of mean	B.3.4.4
End of life utility decrement	[redacted]	[redacted]	PSA: beta distribution DSA: 80% to 120% of mean	B.3.4.4
Adverse event disutilities				
Vomiting	0.048	0.016	PSA: beta distribution DSA: 80% to 120% of mean	B.3.4.4
Hyponatraemia	0.000	0.000		
Pneumonitis	0.037	0.004		
Hepatic function abnormal	0.119	0.012		
Adrenal insufficiency	0.119	0.012		
Acute kidney injury	0.048	0.016		
Colitis	0.047	0.005		
Nausea	0.048	0.016		
Dehydration	0.119	0.012		
Febrile neutropenia	0.090	0.016		

Table 45. Summary of costs applied in the economic model

Variable	Mean value	SE value (if applicable) or uncertainty measurement	Measurement of uncertainty and distribution	Section in Document B
Costs				
First line treatment costs				
Treatment arm: Nivolumab cost per dose	[redacted]	NA	PSA: NA DSA: 80% to 120% of mean	B.3.5.1.1
Treatment and control arm: Fluorouracil cost per dose	£1.77	£0.0012	PSA: Gamma DSA: 80% to 120% of mean	3.5.1.1
	£1.77	£0.0012		
	£1.77	£0.0012		
	£1.77	£0.0012		
	£8.58	£0.0010		
Treatment and control arm: Cisplatin cost per dose	£5.38	£0.0003	PSA: Gamma DSA: 80% to 120% of mean	3.5.1.1
	£8.73	£0.0007		

Variable	Mean value	SE value (if applicable) or uncertainty measurement	Measurement of uncertainty and distribution	Section in Document B
First line treatment modifier				
Treatment arm: Nivolumab	[redacted]	NA	PSA: NA DSA: 80% to 120% of mean	3.5.1.1.1
Treatment arm: Fluorouracil	[redacted]	NA		
Control arm: Treatment arm: Cisplatin	[redacted]	NA		
Control arm: Fluorouracil	[redacted]	NA		
Control arm: Cisplatin	[redacted]	NA		
Number of patients receiving subsequent treatment				
Treatment arm: Nivolumab	[redacted]	NA	NA	3.5.1.2
Treatment arm: Docetaxel	[redacted]	NA		
Treatment arm: Paclitaxel	[redacted]	NA		
Control arm: Nivolumab	[redacted]	NA		
Control arm: Docetaxel	[redacted]	NA		
Control arm: Paclitaxel	[redacted]	NA		
Subsequent treatment costs				
Average cyclical cost: Nivolumab	[redacted]	NA	NA	3.5.1.2
Average cyclical cost: Docetaxel	£129.50	NA		
Average cyclical cost: Paclitaxel	£191.62	NA		
Treatment arm: Weighted average cyclical cost (nivolumab)	£85.36	£17.07	PSA: Gamma DSA: 80% to 120% of mean	
Control arm: Weighted average cyclical cost (taxane: docetaxel and paclitaxel)	£826.80	NA	PSA: NA DSA: 80% to 120% of mean	
Health state costs				
Pre-Progression health state cost	[redacted]	[redacted]	PSA: gamma distribution DSA: 80% to 120% of mean	B.3.5.1.2

Variable	Mean value	SE value (if applicable) or uncertainty measurement	Measurement of uncertainty and distribution	Section in Document B
Post-Progression health state cost	[redacted]	[redacted]	PSA: gamma distribution DSA: 80% to 120% of mean	
Terminal care costs	[redacted]	[redacted]	PSA: gamma distribution DSA: 80% to 120% of mean	B.3.5.1.3
Adverse event costs				
Vomiting	£471.95	£94.39	PSA: gamma distribution DSA: 80% to 120% of mean	B.3.5.2
Hyponatraemia	£1,164.14	£232.83		
Pneumonitis	£1,909.33	£381.87		
Hepatic function abnormal	£2,461.04	£492.21		
Adrenal insufficiency	£2,079.75	£415.95		
Acute kidney injury	£1,961.20	£392.24		
Colitis	£2,426.57	£485.31		
Nausea	£471.95	£94.39		
Dehydration	£1,329.93	£265.99		
Febrile neutropenia	£4,755.76	£951.15		

AE: adverse events; DSA: deterministic sensitivity analysis; NA: not applicable; PSA: probabilistic sensitivity analysis; SE: standard error.

b) Please provide the selection criteria for the parameters to be included in the PSA and deterministic sensitivity analyses (DSA).

The parameters excluded from the PSA are the components that make up first line and second line treatment costs (except for the inputs sourced from eMIT whose treatment costs have some level of uncertainty), lifetables and model settings. These parameters are excluded based on the fact that they are fixed parameters, which do not contain uncertainty with regards to this model. All parameters included in the PSA are done so on the basis that some degree of uncertainty remains.

All parameters except for survival and PAS are included in the DSA. The parameters included in the DSA are those parameters whose variation provides an insight into the key drivers of cost-effectiveness in the model. Whilst survival is a key driver in the model, sensitivity around the choice of extrapolation is explored extensively in scenario analyses in the company submission.

B19. Drug acquisition costs derived from the electronic Market Information Tool (eMIT) are not fixed costs and so they can be a parameter varied in the DSA and PSA using SEs from eMIT. Please include in sensitivity analysis.

Response: The model has been adapted to allow drug acquisition costs from eMIT to vary within the DSA and PSA.

Model validation and transparency

B20. Priority question: Document B, Section B.3.10 of the CS states that, "*A technical review of the cost-effectiveness model was conducted by an independent economist.*"

a) Please provide further details of the questions and results of this validation effort.

Response: The economic model was reviewed by a senior health economist, independent to the team involved in developing the model itself. The approach aligned with established Good Model Validation Practice guidance as presented by ISPOR[1], NICE[2], AdViSHE[3] and TECH-VER.[4] The technical review focussed on various areas including conceptual and internal validation. internal validation comprised:

a. Technical pressure testing (or extreme values analysis) – model input parameters are modified in such a way that their impact on results should be immediately intuitive, enabling rapid identification of errors in modelling logic

b. Directional input testing – modelled clinical input parameters are modified individually and their directional relationship with cost and QALY outcomes evaluated

b) Please confirm whether black and white-box tests to detect modelling errors were conducted.

Response: As previously described, internal validity of the model was tested in line with Büyükkaramikli et al.,[36] and as such included 'black-and-white' tests to detect modelling errors. Some examples include:

a. Setting treatment effects to 0

b. Setting discounting to 0%

c. Setting model inputs equal across treatment arms

d. Setting costs to 0, increasing/decreasing costs per arm

e. Setting utilities to 0, increasing/decreasing utilities per arm

In each case, results were checked to ensure trends and model behaviour were as expected. For example, when discounting was set to 0%, it was checked that discounted costs and QALYs were equivalent to undiscounted; or for increasing costs in the treatment arm only, no impact was observed on costs in the control arm).

c) If no to b), to ensure the internal validity of the model, please complete (if possible, by an independent reviewer) the Technical Verification (TECH-VER) checklist which is a verification checklist to reduce errors in models and improve their credibility (see: Büyükkaramikli, N. C., Rutten-van Mölken, M. P., Severens, J. L., & Al, M. (2019). TECH-VER: A verification checklist to reduce errors in models and improve their credibility. *Pharmacoeconomics*, 37(11), 1391-1408).

Response: Not applicable.

d) Please assess the external validity of model inputs, intermediate outcomes, as well as final outcomes using:

i. Evidence used to develop the economic model

ii. Evidence not used to develop the economic model

Response: As described within the company submission, SLRs were undertaken for economic models (Section 3.1, Appendix H), health-related quality of life (Section 3.4.1, Appendix G), and cost and healthcare resource use (Section 3.5, Appendix I). No relevant UK studies in first line advanced or metastatic OSCC were identified that could be used for external validation.

In terms of validation for model output survival, Table 67 of the company submission (reproduced below in Table 46) explores this against CheckMate 648 trial data. It can be observed that there is only small variation between CheckMate 648 trial data and model output, indicating the model represents the available data well.

Note that validation versus TA submissions is discussed in question B21 below.

Table 46. Comparison of economic model output with CheckMate 648 data

		NIVO+CHEMO			CHEMO		
		PLD	Preferred Survival Curves	Model Output	PLD	Preferred Survival Curves	Model Output
OS	1 year	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	2 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	3 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	5 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	10 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PFS	1 year	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	2 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	3 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	5 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	10 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

CHEMO: chemotherapy; NIVO: nivolumab; OS: Overall survival; PFS: Progression-free survival; PLD: Patient-level data

B21. Please provide cross validations, i.e., comparisons with other relevant NICE TAs such as TA737, and elaborate on the identified differences regarding:

a) Model structure and assumptions

b) Input parameters related to:

i. Clinical effectiveness

ii. Health state utility values

iii. Resource use and costs

iv. Estimated (disaggregated) outcomes per comparator/ intervention

Response: The only relevant NICE appraisal for first line advanced or metastatic OSCC is TA737 for pembrolizumab.[6] As such, model inputs and outputs relating to NIVO+CHEMO cannot be cross-validated. From TA737, data relating to

pembrolizumab + 5FU +cisplatin, and 5FU + cisplatin, have been assessed as these are the relevant treatment regimens in the current appraisal (Table 47).[6]

Table 47. Comparison of the models applied in the company submission and TA737

	Current appraisal	TA737 original company submission[6]
Model structure and assumptions		
Model structure	3 state partitioned survival model (progression-free, progressed disease, death)	3 state partitioned survival model (progression-free, progressed disease, death)
Time horizon	Lifetime	Lifetime
Cycle length	1 week no half-cycle correction	1 week with half-cycle correction
Utility source	CheckMate 648 EQ-5D-3L	KEYNOTE-590 EQ-5D-3L
Cost source	eMIT and BNF for acquisition costs; administration costs, adverse event costs, disease management costs from NHS reference costs	eMIT and BNF for acquisition costs; administration costs, adverse event costs, disease management costs from NHS reference costs
Duration of treatment effect	No treatment waning	No treatment waning in company base case
Treatment pathway	Subsequent treatments in line with clinical practice (based on clinical expert opinion)	Subsequent treatments in line with those from KEYNOTE-590
Safety	Adverse event incidence from CheckMate 648	Adverse event incidence from KEYNOTE-590
Stopping rule	Stopping rule based on treatment specific time on treatment curves	Pembrolizumab not administered beyond 24 months, cisplatin to 6 cycles, 5-FU to 25 cycles
Clinical effectiveness		
PFS efficacy	KM data (CheckMate 648) to 6.9 months, followed by generalised gamma distribution for the treatment arm and Weibull distribution for the control arm	KM data (KEYNOTE-590) to 10 weeks, followed by log-logistic distribution, since first tumour assessment at week 9
OS efficacy	KM data (CheckMate 648) to 6.9 months, followed by generalised lognormal distribution for the treatment and control arm	KM data (KEYNOTE-590) to 40 weeks, followed by log-logistic models, established via clinical validity and AIC/BIC
HRQoL		
Health state utility values	By progression status, [redacted] pre-progression, [redacted] post-progression, with end-of-life decrement ([redacted])	Time-to-death utilities, values redacted
Age-related disutility	Utilities not adjusted by UK general population	Utilities adjusted by UK general population
Resource use and costs		
Time on treatment	Time on treatment curves applied to both arms, based on CheckMate 648 (mean ToT from TA737 used to	Time on treatment applied to both arms, based on KEYNOTE-590

	derive PEMBRO + CHEMO time on treatment curve)	
Relative dose intensity/ treatment modifier	Dose intensity applied to all arms, based on CheckMate 648 for NIVO + CHEMO and CHEMO. PEMBRO + CHEMO assumed equivalent to NIVO + CHEMO	Relative dose intensity applied to both arms, based on KEYNOTE-590. Values redacted
Healthcare resource use	Aligns between treatment and control arms	Aligns between treatment and control arms
Pre-progression healthcare resource use (per cycle)	0.08 CT scan 0.33 full blood count 0.33 renal function test 0.33 hepatic function test 0.25 consultation visit	0.08 CT scan 0.33 full blood count 0.33 renal function test 0.33 hepatic function test 0.25 consultation visit
Post-progression healthcare resource use (per cycle)	0.08 CT scan 1 full blood count 1 renal function test 1 hepatic function test 0.25 consultation visit	0.08 consultation visit
Administration costs for first line treatments*	Cisplatin + 5FU, SB14Z (NHS reference costs) at first attendance Nivolumab, SB12Z (NHS reference costs) on day 15 per cycle, and SB14Z (NHS reference costs) on day 1 per cycle.	In both PEMBRO + cisplatin + 5FU, and cisplatin + 5FU, SB14Z (NHS reference costs) at first attendance
Acquisition costs for first line treatments*	BNF for nivolumab and pembrolizumab, eMIT for chemotherapy components	BNF for pembrolizumab, eMIT for chemotherapy components
Terminal care cost	Based on Georghiou et al. (2014),[37] adjusted for inflation	Based on TA522,[38] adjusted for inflation.
Adverse events	Incidence from CheckMate 648, with the most common grade 3–4 drug-related serious adverse events from all treatment arms included. Incidence for PEMBRO + CHEMO taken from TA737. One-off cost and disutility applied on incidence of adverse event. Adverse events only associated with first line treatment, and only occur on treatment initiation. Costs based on NHS reference costs and literature.[39,40] Disutility based on TAs,[41,42] literature[43] and assumptions.	Incidence from KEYNOTE 590, one-off cost and disutility applied. Cost based on mean duration and NHS reference costs. Utility based on KEYNOTE 590 data, time to death approach.
*detail of costs themselves not incorporated herein, due to updates in NHS reference cost and eMIT databases Note that the data from TA737 presented in this table relates to the original company submission, and not any updates following ERG/NICE review.		

Disaggregated outcomes from TA737 are redacted, and therefore, cannot be compared. Total and incremental LY/QALY/costs from the current appraisal versus

TA737 are explored within B.3.10.3 of the company submission and are reproduced below (Table 48). Overall, predicted LY and costs are broadly comparable.

Table 48. Comparison of outcomes for cisplatin and fluorouracil

		Current appraisal	TA737[44]	
			Company	ERG
Total LYs	CHEMO	[redacted]	1.37	NR
	Intervention	[redacted]	2.13	NR
Incremental costs (£)*		[redacted]	27,165	28,007
Incremental QALYs		[redacted]	0.63	0.54
* applies PAS for intervention arms				

B22. Table 56 in the CS (Document B, Section B.3.6) lists the assumptions applied in the economic model. Please modify this table by providing a column for 'Area'. For example, assumptions made about utilities would be classified under 'HRQoL'.

Response: Please find the table as requested below.

Table 49. Assumptions applied in the economic model

Area	Assumption	Rationale
Baseline parameters	Baseline parameters are derived from CheckMate 648 cohort, which is assumed to be reflective of patients seen in UK clinical practice for the anticipated MA.	Although there may be differences between characteristics in CheckMate 648 and OSCC patients in UK clinical practice, these can be considered small. Sensitivity analyses (probabilistic and deterministic) have been conducted to assess the impact of variability in these parameters.
Model settings/ structure	The model applies a weekly cycle length, which is assumed to be sufficiently granular to accurately reflect costs and benefits when modelling OC.	Previous OC evaluations assessed by NICE had applied weekly cycle lengths, which was considered appropriate by ERG.[6,21,22,45] This cycle length is short enough to reflect the treatment cycles for patients and reflects the frequency of follow-up for patients and reflects the frequency of follow-up for patients and a realistic minimum time during which symptoms or response can change.
Model settings/ structure	To reflect the nature of OC and available evidence, the model assumes that OC phases are consecutive, and	This assumption has been validated by clinicians and is in line with other HTAs and economic analyses assessing the OC population.

Area	Assumption	Rationale
	patients cannot revert to pre-progression from more advanced phases of the disease.	
Efficacy	Identification of most appropriate survival curves describing PFS, and OS inform extrapolation	Extensive analyses have been undertaken to identify appropriate and conservative survival curves describing NIVO+CHEMO efficacy, with reference to the guidance from the NICE Decision Support Unit (DSU)[24] and Bagust and Beale (2014)[46]. The approach and identified survival extrapolations have been validated by clinical and health economic experts. However, to address the uncertainty around this parameter, scenario analyses have been conducted by applying alternative assumptions around extrapolations, as presented in Section B.3.3.1.
Efficacy	Efficacy has been based on BICR-assessed data, rather than investigator-assessed data	During CheckMate 648, the two measures of response of PFS were comparable. However, BICR was designated as the primary endpoint and may be considered slightly more conservative.
Safety	As a simplification, it is assumed that all adverse events occur in the first cycle of treatment.	The majority of patients during CheckMate 648 have discontinued treatment within the current database lock, so that the data can be considered an accurate reflection of the safety profile. AEs are often only observed to occur soon after treatment initiation, so that this may not be well reflected by assuming a constant rate per cycle.
HRQoL	It was assumed that health state utilities, pre-progression, post-progression and the disutility of death, are the same for the treatment and control arm.	This is based on evidence observed during CheckMate 648, described in Section B.3.4.2.
Treatment costs	It was assumed that patients receiving pembrolizumab in combination with chemotherapy experience missing or delayed doses in line with nivolumab during CheckMate 648.	Currently, there is no published data available to inform proportion of received doses of pembrolizumab. As the mechanism of action is similar, this seems an appropriate assumption.
Health state costs	The health state resource use is derived from evidence presented in TA737.	Robust estimates of health state resource use for patients in this setting are not publicly available, given the limited alternative treatment available for which evidence may have previously gathered. In order to provide relevant economic evaluations and facilitate

Area	Assumption	Rationale
		comparison between these appraisals, health state resource use from TA737 is applied.
Treatment pathway	Subsequent treatment for NIVO-CHEMO and PEMBRO+CHEMO is assumed to be single agent taxane (equal use of paclitaxel and docetaxel).	During CheckMate 648, taxane use reflected around 70% of subsequent systemic therapy use, indicating the plausibility of this assumption. Docetaxel and paclitaxel have similar efficacy and cost.
Treatment pathway	Subsequent treatment for CHEMO is assumed to nivolumab monotherapy.	This aligns with the current UK treatment pathway and is aligned with budget impact assumptions applied during TA707.[22]
Safety	AE utility decrement values were assumed for certain AEs.	Values were assumed for those AEs where published data was not available. However, deterministic sensitivity analysis has been presented to show the impact of AE utility decrements.
Efficacy	No treatment waning has been assumed.	Evidence supports a robust and durable treatment effect lasting beyond discontinuation for immunotherapies.[47] Further, during TA737, the committee concluded that all scenarios provided plausible estimates of overall survival and the treatment waning scenarios were not greatly different from those without treatment waning.[6] This is of particular relevance given the low long-term hazard in the CHEMO arm of CheckMate 648.

Other

B23. Priority question: Please provide all details of the communication between the company and clinical and health economic experts. Please include anonymised information about the clinical experts, detailed minutes of the face-to-face meeting and/or teleconference, list of expert recommendations and justifications for clinical assumptions and inputs used in the model.

Response: An advisory board was held on 14 July 2021 by BMS comprising of clinicians and an economist,[9] with the aim of developing insight to support the NICE submission for nivolumab with platinum-based chemotherapy for the treatment of advanced unresectable, recurrent or metastatic previously untreated oesophageal squamous cell carcinoma (OSCC). The board explored key themes developed by BMS around specific issues related to the clinical positioning and economic strategy and shared published results from the CheckMate 648 trial to gain feedback on how they resonated with clinicians and economists.

Details of the attendees and summaries of the discussions held during the meeting are provided in the advisory board report.[9]

A list of the experts' recommendations and justifications for clinical assumptions and inputs used in the model are provided in Table 50.

Table 50. Expert recommendations from the advisory board

Assumption	Justification
UK clinical practice	
Data regarding squamous GC can be considered comparable to OSCC	The REAL2 study,[10] shows that squamous GC and OC are comparable.
There were no additional treatments to be considered for the treatment of advanced, previously untreated OSCC, beyond the doublet and triplet regimens presented.	Confirmed by the clinicians and aligned with NICE guidance at the time of the advisory board.
If nivolumab combination therapy was approved as a first-line treatment, then a nivolumab-containing second-line therapy would not be offered. It was generally believed that a docetaxel or paclitaxel-containing regimen would be offered in the second-line after a nivolumab-containing first-line regimen.	Current NICE clinical guidance and clinical expert opinion.
CheckMate 648	
Eligibility criteria and baseline patient characteristics representative of patients seen in UK clinical practice	Clinical expert opinion
There is no difference between OSCC patients from Asia or Europe.	Clinical expert opinion
The safety profile for nivolumab with chemotherapy was not a concern for the clinicians as they would be expecting AEs with both immunotherapies and chemotherapy and so would select and treat patients accordingly.	Clinical expert opinion
Survival modelling	
The survival data presented from CheckMate 648 aligned with the experts expectations.	Clinical expert opinion
In lethal cancer, patients who survive beyond 18-24 months are considered long-term survivors and would stop immunotherapy at this stage	Product SmPCs and guidance, and expert opinion.
Resource use in patients surviving beyond 24 months would be fairly intensive, as patients may still be symptomatic	Clinical expert opinion
it would be appropriate to use long-term clinical data from other nivolumab indications to validate the hazard profile evolution	Clinical expert opinion
The Weibull and Gompertz estimates were thought to be the most similar to current clinical practice in the UK	Clinical expert opinion
Cost-effectiveness modelling	
Published utility values from a squamous gastric cancer population would be appropriate to include in the model for external validation or to inform post-progression data gaps.	The GO2 trial in upper GI cancer reported utility as a primary endpoint and was suggested as a good source.[48]

Section C: Textual clarification and additional points

Appendices

C1. Appendix C (Nivolumab draft SmPC, provided as a separate document to the main CS) appears to be work in progress, showing several instances of tracked changes. Please provide the final version of this document, without tracked changes.

Response: The final version of the SmPC is provided with this response.

References

1. European Medicines Agency. OPDIVO. 2022. Available at: https://www.ema.europa.eu/en/medicines/human/summaries-opinion/opdivo-5 [Accessed 20/04/2022].
2. Kulangara K, Zhang N, Corigliano E, et al. Clinical Utility of the Combined Positive Score for Programmed Death Ligand-1 Expression and the Approval of Pembrolizumab for Treatment of Gastric Cancer. Arch Pathol Lab Med. 2019;143(3):330-7.
3. Sundar R, Smyth EC, Peng S, et al. Predictive Biomarkers of Immune Checkpoint Inhibition in Gastroesophageal Cancers. Frontiers in Oncology. 2020;10.
4. Hospitals NTLT. Lung cancer testing. 2022. Available at: https://www.leedsth.nhs.uk/a-z-of-services/the-leeds-genetics-laboratory/oncology-genetics/molecular-oncology/lung-cancer-testing/ [Accessed 20/04/22].
5. Sun JM, Shen L, Shah MA, et al. Pembrolizumab plus chemotherapy versus chemotherapy alone for first-line treatment of advanced oesophageal cancer (KEYNOTE-590): a randomised, placebo-controlled, phase 3 study. Lancet. 2021;398(10302):759-71.
6. National Institute for Health and Care Excellence. Pembrolizumab with platinum- and fluoropyrimidine-based chemotherapy for untreated advanced oesophageal and gastro-oesophageal junction cancer [TA737]. 2021. Available at: https://www.nice.org.uk/guidance/ta737/chapter/1-Recommendations [Accessed 09/12/21].
7. Janjigian YY, Shitara K, Moehler M, et al. First-line nivolumab plus chemotherapy versus chemotherapy alone for advanced gastric, gastro-oesophageal junction, and oesophageal adenocarcinoma (CheckMate 649): a randomised, open-label, phase 3 trial. The Lancet. 2021;398(10294):27-40.
8. Kamposioras K, Ntellas P, Nikolaou M, et al. Immunotherapy Efficacy in the Initial Lines of Treatment in Advanced Upper Gastrointestinal Malignancies: A Systematic Review of the Literature. JNCI Cancer Spectr. 2021;5(6).
9. Bristol-Myers Squibb/HEOR Ltd. Nivolumab with platinum-based chemotherapy or ipilimumab for advanced unresectable, recurrent or metastatic previously untreated oesophageal squamous cell carcinoma - Advisory board report [data on file]. 2021.
10. Cunningham D, Starling N, Rao S, et al. Capecitabine and Oxaliplatin for Advanced Esophagogastric Cancer. New England Journal of Medicine. 2008;358(1):36-46.
11. Shyamalee J, Baxter MA, Walsh S, et al. 1720P Real-world outcomes for advanced oesophageal squamous cell carcinoma with supportive care alone. Annals of Oncology. 2021;32:S1194.
12. Jaffe D, DeCongelio M, Gricar J. HSR21-055: Treatment Patterns and Healthcare Resource Utilization in First-Line Therapy for Advanced Esophageal Squamous Cell Carcinoma in Asian and Western Countries. Journal of the National Comprehensive Cancer Network J Natl Compr Canc Netw. 2021;19(3.5):HSR21-055-HSR21-.
13. Bristol-Myers Squibb. Primary Clinical Study Report for Study CA209648: A Randomized Phase 3 Study of Nivolumab plus Ipilimumab or Nivolumab Combined with Fluorouracil plus Cisplatin versus Fluorouracil plus Cisplatin in

Subjects with Unresectable Advanced, Recurrent or Metastatic Previously Untreated Esophageal Squamous Cell Carcinoma. 2021.
14. Doki Y, Ajani JA, Kato K, et al. Nivolumab Combination Therapy in Advanced Esophageal Squamous-Cell Carcinoma. N Engl J Med. 2022;386(5):449-62.
15. Ford HER, Marshall A, Bridgewater JA, et al. Docetaxel versus active symptom control for refractory oesophagogastric adenocarcinoma (COUGAR-02): an open-label, phase 3 randomised controlled trial. The Lancet Oncology. 2014;15(1):78-86.
16. Davidson M, Cafferkey C, Goode EF, et al. Survival in Advanced Esophagogastric Adenocarcinoma Improves With Use of Multiple Lines of Therapy: Results From an Analysis of More Than 500 Patients. Clinical Colorectal Cancer. 2018;17(3):223-30.
17. Shitara K, Van Cutsem E, Bang Y-J, et al. Efficacy and Safety of Pembrolizumab or Pembrolizumab Plus Chemotherapy vs Chemotherapy Alone for Patients With First-line, Advanced Gastric Cancer: The KEYNOTE-062 Phase 3 Randomized Clinical Trial. JAMA oncology. 2020;6(10):1571-80.
18. Muro K, Lordick F, Tsushima T, et al. Pan-Asian adapted ESMO Clinical Practice Guidelines for the management of patients with metastatic oesophageal cancer: a JSMO-ESMO initiative endorsed by CSCO, KSMO, MOS, SSO and TOS. Ann Oncol. 2019;30(1):34-43.
19. Chen S, Zhou K, Yang L, et al. Racial Differences in Esophageal Squamous Cell Carcinoma: Incidence and Molecular Features. BioMed Research International. 2017;2017:1204082.
20. Bristol-Myers Squibb. CheckMate 648 Oct 2021 Summary data [Data on file]. 2021.
21. National Institute for Health and Care Excellence. Nivolumab for adjuvant treatment of resected oesophageal or gastro-oesophageal junction cancer [TA746]. 2021. Available at: https://www.nice.org.uk/guidance/ta746 [Accessed 20/04/2022].
22. National Institute for Health and Care Excellence. Nivolumab for previously treated unresectable advanced or recurrent oesophageal cancer [TA707]. 2021. Available at: https://www.nice.org.uk/guidance/ta707 [Accessed 20/04/2022].
23. Cope S, Chan K, Jansen J. Multivariate Network Meta-Analysis Of Survival Function Parameters. Research Synthesis Methods. 2020.
24. National Institute for Health and Care Excellence Decision Support Unit. Technical Support Document 14: Survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data. 2016. Available at: http://www.nicedsu.org.uk/NICE%20DSU%20TSD%20Survival%20analysis.updated%20March%202013.v2.pdf [Accessed 23/04/2021].
25. Kato K, Sun J, Shah M, et al. Pembrolizumab plus chemotherapy versus chemotherapy as first-line therapy in patients with advanced esophageal cancer: The phase 3 KEYNOTE-590 study. 2020.
26. Guyot P, Ades AE, Ouwens MJ, et al. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. BMC Med Res Methodol. 2012;12:9.
27. Lordick F, Mariette C, Haustermans K, et al. Oesophageal cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2016;27(suppl 5):v50-v7.

28. National Institute for Health and Care Excellence. Ramucirumab for treating advanced gastric cancer or gastro–oesophageal junction adenocarcinoma previously treated with chemotherapy [TA378]. 2016. Available at: https://www.nice.org.uk/guidance/ta378 [Accessed 20/04/2022].
29. Chau Iea. Nivolumab plus ipilimumab or nivolumab plus chemotherapy versus chemotherapy as first-line treatment for advanced esophageal squamous cell carcinoma: first results of the CheckMate 648 study. ASCO annual meeting, Abstract number LBA4001. 2021.
30. clinicaltrials.gov. First-line Esophageal Carcinoma Study With Chemo vs. Chemo Plus Pembrolizumab (MK-3475-590/KEYNOTE-590) (NCT03189719). 2022. Available at: https://clinicaltrials.gov/ct2/show/NCT03189719 [Accessed 21/01/2022].
31. Department of Health. NHS reference costs 2015 to 2016. 2016. Available from: https://www.gov.uk/government/publications/nhs-reference-costs-2015-to-2016.".
32. NHS England. NHS reference costs 2019-2020. 2021. Available at: https://www.england.nhs.uk/national-cost-collection [Accessed 17/01/2022].
33. Oxford Population Health. Decision Modelling for Health Economic Evaluation. Available at: https://www.herc.ox.ac.uk/downloads/decision-modelling-for-health-economic-evaluation [Accessed 20/04/2022].
34. Edlin R, McCabe C, Hulme C, et al. Cost Effectiveness Modelling for Health Technology Assessment. 2015. Available at: https://link.springer.com/book/10.1007/978-3-319-15744-3 [Accessed 20/04/2022].
35. Department of Health and Social Care. NHS reference costs. 2022. Available at: https://www.gov.uk/government/collections/nhs-reference-costs [Accessed 20/04/2022].
36. Büyükkaramikli NC, Rutten-van Mölken MPMH, Severens JL, et al. TECH-VER: A Verification Checklist to Reduce Errors in Models and Improve Their Credibility. PharmacoEconomics. 2019;37(11):1391-408.
37. Georghiou T, Bardsley M. Exploring the cost of care at the end of life. 2014. Available at: https://www.nuffieldtrust.org.uk/research/exploring-the-cost-of-care-at-the-end-of-life [Accessed 21/01/2022].
38. National Institute for Health and Care Excellence. Pembrolizumab for untreated PD-L1-positive locally advanced or metastatic urothelial cancer when cisplatin is unsuitable [TA522]. 2018. Available at: https://www.nice.org.uk/guidance/ta522 [Accessed 28/04/2022].
39. Chauhan R, Lee D. Adrenal Insufficiency: Burden of Disease and Cost of Illness. Value in Health. 2013;16(7):A436.
40. Copley-Merriman C, Stevinson K, Liu FX, et al. Direct costs associated with adverse events of systemic therapies for advanced melanoma: Systematic literature review. Medicine (Baltimore). 2018;97(31):e11736.
41. National Institute for Health and Care Excellence. Nivolumab for previously treated non-squamous non-small-cell lung cancer [TA484]. 2017. Available at: https://www.nice.org.uk/guidance/ta484 [Accessed 28/04/2022].
42. National Institute for Health and Care Excellence. Durvalumab for treating locally advanced unresectable non-small-cell lung cancer after platinum-based chemoradiation [TA578]. 2019. Available at: https://www.nice.org.uk/guidance/ta578 [Accessed 28/04/2022].

43. Nafees B, Stafford M, Gavriel S, et al. Health state utilities for non small cell lung cancer. Health Qual Life Outcomes. 2008;6:84.
44. National Institute for Health and Care Excellence. Technology appraisal guidance [TA737]. Pembrolizumab with platinum- and fluoropyrimidine-based chemotherapy for untreated advanced oesophageal and gastro-oesophageal junction cancer. 2021. Available from: https://www.nice.org.uk/guidance/ta737 [accessed 05/01/2022].
45. National Institute for Health and Care Excellence. Nivolumab with platinum- and fluoropyrimidine-based chemotherapy for untreated HER2-negative advanced gastric, gastro-oesophageal junction or oesophageal adenocarcinoma [ID1465]. 2022. Available from: https://www.nice.org.uk/guidance/indevelopment/gid-ta10352 [accessed 07/03/2022].
46. Bagust A, Beale S. Survival analysis and extrapolation modeling of time-to-event clinical trial data for economic evaluation: an alternative approach. Med Decis Making. 2014;34(3):343-51.
47. Regan MM, Mantia CM, Werner L, et al. Treatment-free survival over extended follow-up of patients with advanced melanoma treated with immune checkpoint inhibitors in CheckMate 067. Journal for ImmunoTherapy of Cancer. 2021;9(11):e003743.
48. Hall PS, Swinson D, Cairns DA, 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 oncology. 2021;7(6):869-77.

The technical team have identified a couple of other questions and would be grateful for your answer:

- Please confirm the formulations used to inform the following unit drug costs:
    - Oxaliplatin
    - Docetaxel
    - Paclitaxel

Drug	Dose required	Formulation
Paclitaxel	166.4 mg	100mg/16.7ml *2
Docetaxel	125 mg	160mg/8ml *1
Oxaliplatin (FOLFOX)	141.4mg	200mg/40ml *1
Oxaliplatin (XELOX)	216.31mg	200mg/40ml *1 plus 50mg/10ml *1

- Please could the company clarify the PAS price per vial as reported in Table 1 of Document A and Table 2 of Document B as the values appear to be the value of the PAS discount itself and do not align with the PAS cost reported in Table 46 of Document B.

PAS = [redacted]%

		List Price	PAS price in CS	Correct PAS price
Doc A, Table 1		Nivolumab: £2,633 per 240 mg vial; £1,097 per 100 mg vial; £439.00 per 40 mg vial.	Nivolumab: £[redacted] per 240 mg vial; £[redacted] per 100 mg vial; £[redacted] per 40 mg vial.	Nivolumab: £[redacted] per 240 mg vial; £[redacted] per 100 mg vial; £[redacted] per 40 mg vial.
Doc B, Table 2		Nivolumab: £2,633 per 240 mg vial; £1,097 per 100 mg vial; £439.00 per 40 mg vial.	Nivolumab: £[redacted] per 240 mg vial; £[redacted] per 100 mg vial; £[redacted] per 40 mg vial.	Nivolumab: £[redacted] per 240 mg vial; £[redacted] per 100 mg vial; £[redacted] per 40 mg vial.
Doc B, Table 46		240 mg - £2,633.00	240mg - £[redacted]	240mg - £[redacted]

The PAS was incorrectly displayed in Doc A, Table 1 and Doc B, Table 2 but correctly in Doc B, Table 46. The PAS was correctly applied in the company's model.

NICE National Institute for Health and Care Excellence

Patient organisation submission

Nivolumab in combination for untreated advanced unresectable recurrent or metastatic oesophageal squamous cell carcinoma [ID2712]

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on conditions and their treatment that is not typically available from other sources.

To help you give your views, please use this questionnaire with our guide for patient submissions.

You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type. [Please note that declarations of interests relevant to this topic are compulsory].

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 10 pages.

About you	
1.Your name	[redacted]

2. Name of organisation	Guts UK Charity
3. Job title or position	[redacted]
4a. Brief description of the organisation (including who funds it). How many members does it have?	Guts UK are a charity that fundraises for research and provides information to help people manage diseases and conditions affecting the digestive tract, liver and pancreas. The charities mission is to • Provide expert information: Information is power! When armed with information, patients can take control of their health and make informed decisions. We do this by information leaflets sent to patients and sold to hospitals, our website and social media accounts. Guts UK also produce a biannual magazine. • Raise public awareness: Guts UK research shows that 58% of people are embarrassed to talk about their digestive condition or symptoms. 51% of people delay seeking advice for their symptoms for over 6 months. When the Guts UK roadshow comes to town, we empower people to seek help. We also fund science of digestion events to increase knowledge. Fund life-changing & life-saving research: Guts UK is the only UK charity funding research into the digestive system from top to tail. It's time the UK got to grips with guts!
4b. Has the organisation received any funding from the manufacturer(s) of the technology and/or comparator products in the last 12 months? [Relevant	To be fully transparent with this process Guts UK are founder members of the Less Survivable Cancers Taskforce (LSCT) and whilst Guts UK have not received any direct funding from the manufacturers in the last 12 months LSCT may have. As LSCT is a separate concern no details of funding amounts can be provided as this is commercially sensitive information.

manufacturers are listed in the appraisal stakeholder list.] If so, please state the name of manufacturer, amount, and purpose of funding.	
4c. Do you have any direct or indirect links with, or funding from, the tobacco industry?	Guts UK has no links at all with the tobacco industry
5. How did you gather information about the experiences of patients and carers to include in your submission?	We asked within support groups for people living with oesophageal cancer and cancer between the stomach and gullet (gastro-oesophageal junction) to get in touch to share their story of living with or caring for someone diagnosed with these cancers. We also asked if anyone had experience of Nivolumab in combination with other chemotherapy for oesophageal cancer (cancer between the stomach and gullet.) We have also developed surveys in the past, but these were not successful in getting responses. Understandably, it is difficult for people to input time into submissions with advanced cancer, so we also searched for qualitative studies for quality of life and life experience of people diagnosed with these cancers to understand their experience. We also interviewed support group leaders who help people living with oesophageal cancers and have lived experience themselves.
Living with the condition	
6. What is it like to live with the condition? What do carers	Oesophageal cancer and cancer between the stomach and gullet are two of six less survivable cancers, for which there are no screening tools to identify them widely used, and as early symptoms are vague, people are frequently diagnosed late, when treatment options are limited. The chance of surviving beyond five years with oesophageal cancer is approximately 15 out of 100 people diagnosed. Often patients and

experience when caring for someone with the condition?	their families have limited time together, as many as 7 in 10 (Humphreys E et al 2020) people are diagnosed at a stage (III or IV) when it has spread to the lymph nodes and has spread to nearby organs and distant body sites. Larsen et al (2020) reported "patients with oesophageal cancer are putting their ordinary lives on hold and experiencing the meal as a battleground during treatment. Patients strive to maintain autonomy, gain control, and take ownership and their suffering was associated with symptoms and side effects of treatment, which affect their and their relatives' social world and relationships." For people with oesophageal cancer swallowing problems can be severe even at times people are unable to swallow their own saliva and this is associated with pain, reflux and indigestion. These symptoms severely affect quality of life, lead to weight loss and fatigue. Not only does eating provoke symptoms but the diet can significantly change not only in texture but food choices are affected by the side effects of treatment. People with cancer also may have a feeding tube and if the cancer is not curable a stent to open the oesophagus and help with swallowing. Fatigue is a major symptom that people with these cancers experience. When I was told, 'You'll feel a bit of fatigue,' you automatically think, 'Ah yeah, so I'll feel a bit tired.' But fatigue is totally different—you have to explain that it's a total knackered—all over. And you haven't done anything, but suddenly you're knackered and you don't know why. And it plays on your mind, where you're saying, 'What's gone wrong now that I'm suddenly like this?' (Bennett et al 2020.) Symptoms have wider impact on quality of life and will affect social activities such as eating with family, enjoyment of food and attending social events. Sharing food and meal provision is an important aspect of family care provision and loss of weight and inability to enjoy meals is often distressing to both the person with cancer and their families and carers. Often people can manage only small portions of food or fluids, if any, and this impacts on eating out as some facilities will not cater for those requirements – some people do not want to make a fuss, so don't go out. With limited lifespan it is extremely important that people living with these cancers enjoy time with their family and controlling tumour progression can help people to participate. Non curative treatments are difficult to tolerate alongside physically debilitating symptoms make it impossible to continue working or take part in social events for some people.

	Awareness of a poor prognosis and the demanding treatment pathways triggered psychological distress, as patients gave expressions of their feelings of vulnerability. (Larson 2020) Bennett AE, O'Neill L, Connolly D, et al. Perspectives of Esophageal Cancer Survivors on Diagnosis, Treatment, and Recovery. *Cancers (Basel)*. 2020;13(1):100. Published 2020 Dec 31. doi:10.3390/cancers13010100 Larsen MK, Schultz H, Mortensen MB, Birkelund R. Patients' Experiences With Illness, Treatment, and Decision-Making for Esophageal Cancer: A Qualitative Study in a Danish Hospital Setting. *Glob Qual Nurs Res*. 2020;7:2333393620935098. Published 2020 Jun 29. doi:10.1177/2333393620935098
Current treatment of the condition in the NHS	
7. What do patients or carers think of current treatments and care available on the NHS?	Current treatments are challenging to experience, and they are not always effective. People with cancer feel that the treatment schedule constantly interrupts their normal everyday life and this is particularly true of chemotherapy (Larsen et al 2020). Decision making regarding treatment can be a burden for some people with respect to complexity of the treatment and side effects, people often have not heard the medical terminology and people will often defer decisions about treatment to their healthcare practitioners (Larsen et al 2020)
8. Is there an unmet need for patients with this condition?	There are few effective treatments for these cancers that are available so yes there is an unmet need. There are relatively few options in advanced disease and is usually chemotherapy, radiotherapy or a combination of both – Nivolumab, being immunotherapy will be an addition to a new type of treatment for these cancers.

Advantages of the technology	
9. What do patients or carers think are the advantages of the technology?	Nivolumab is a different type of treatment that works in a different way to current treatments. Patients are wanting as many options as possible, they are very aware of survival ratio's and know that one type of treatment doesn't fit all. It is very important to them that there are alternatives or additional treatments. The additional treatment does not impact on current chemotherapy treatment time as it is given consecutively with chemotherapy.
Disadvantages of the technology	
10. What do patients or carers think are the disadvantages of the technology?	Immunotherapy may have different side effects to current therapy. The additional treatment does not change treatment time as it is given consecutively with current treatment.
Patient population	
11. Are there any groups of patients who might benefit more or less from the technology than others? If so, please describe them and explain why.	All groups of people will benefit from this treatment. Some however due to age, fitness and other underlining comorbidities might suffer from different side effects.

Equality	
12. Are there any potential equality issues that should be taken into account when considering this condition and the technology?	It might be challenging for hard-to-reach community groups to access information due to language barriers. Inequalities may be particularly true of squamous cell carcinoma as there is an increased risk of this cancer with traditional use in some cultures of areca nut. Culture may also play a part as some cultures may be reluctant to visit their GP or be registered. Also, inequalities in health in respect to cancer mean that people from the most deprived areas are more likely to be diagnosed later as people have reduced ability and opportunity to access healthcare. This is particularly true of oesophageal and stomach cancer.
Other issues	
13. Are there any other issues that you would like the committee to consider?	Yes, these cancers are difficult for GPs to identify or suspect symptoms are due to cancer at an early stage. Quality of life vs treatment all depends on the patients functional fitness and nutritional status, ability to eat or if they are using a feeding tube and also family can provide peer pressure too.
Key messages	
15. In up to 5 bullet points, please summarise the key messages of your submission: • These cancers are less survivable cancers, for which there are no screening tools to identify them which are widely used and they are frequently diagnosed late, when the treatment options are limited. • People with lived experience of these cancers strive to maintain fitness and gain control of their situation and their suffering is associated with symptoms and treatment side effects, which massively affects their quality of life, social experience and relationships with family and carers.	

- With a life limited condition it is extremely important that people living with these cancers enjoy time with their family and this treatment could help people participate and provide them with valuable time.
- This treatment works by a different mechanism and offers another option for treatment where there are currently few options available.
- Patients will always look for hope in new treatments, or trials for themselves and others

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

.....

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☐ **Please tick this box** if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

.....

Nivolumab in combination with fluoropyrimidine- and platinum-based combination chemotherapy for the first-line treatment of adult patients with unresectable advanced, recurrent, or metastatic oesophageal squamous cell carcinoma with tumour cell PD-L1 expression ≥1% [ID2712]

Produced by Kleijnen Systematic Reviews (KSR) Ltd, United Kingdom (UK)

Authors
Susan O'Meara, Reviews Manager, KSR Ltd, UK
Nigel Armstrong, Health Economics Manager, KSR Ltd, UK
Charlotte Ahmadu, Health Economist, KSR Ltd, UK
Evangelos Danopoulos, Systematic Reviewer, KSR Ltd, UK
Lisa Stirk, Information Specialist, KSR Ltd, UK
Jos Kleijnen, Director and Owner, KSR Ltd, UK

Correspondence to
Susan O'Meara, Kleijnen Systematic Reviews Ltd
Unit 6, Escrick Business Park
Riccall Road, Escrick
York, YO19 6FD
United Kingdom

Date completed 08/06/2022

Source of funding: This report was commissioned by the National Institute for Health Research (NIHR) Evidence Synthesis Programme as project number NIHR 13/55/62.

Declared competing interests of the authors None

Acknowledgements None

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

Academic in confidence (AiC) data are highlighted in yellow throughout the report.

Confidential comparator prices are highlighted in green throughout the report.

Any de-personalised data are highlighted in pink throughout the report.

Rider on responsibility for report

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

This report should be referenced as follows:

O'Meara S, Armstrong N, Ahmadu C, Danopoulos E, Stirk L, Kleijnen J. Nivolumab in combination with fluoropyrimidine- and platinum-based combination chemotherapy for the first-line treatment of adult patients with unresectable advanced, recurrent, or metastatic oesophageal squamous cell carcinoma with tumour cell PD-L1 expression ≥1% [ID2712]: a Single Technology Assessment. York: Kleijnen Systematic Reviews Ltd, 2021.

Contributions of authors

Susan O'Meara acted as project lead and systematic reviewer on this assessment, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Nigel Armstrong acted as health economic project lead, critiqued the company's economic evaluation and contributed to the writing of the report. Charlotte Ahmadu acted as health economist on this assessment, critiqued the company's economic evaluation and contributed to the writing of the report. Evangelos Danopoulos acted as systematic reviewer, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Lisa Stirk critiqued the search methods in the submission and contributed to the writing of the report. Jos Kleijnen critiqued the company's definition of the decision problem and their description of the underlying health problem and current service provision, contributed to the writing of the report and supervised the project.

Abbreviations

AE	Adverse events
AIC	Akaike Information Criterion
AiC	Academic in confidence
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
BIC	Bayesian information criterion
BICR	Blinded independent central review
BNF	British National Formulary
BSA	Body surface area
BSC	Best supportive care
CADTH	Canadian Agency for Drugs and Technologies in Health
CEAC	Cost effectiveness acceptability curve
CENTRAL	Cochrane Central Register of Controlled Trials
CHEMO	Chemotherapy
CHMP	Committee for Medicinal Products for Human Use
CI	Confidence interval
CIS	Cisplatin
CiC	Commercial in confidence
CL	Clarification letter
CMH	Cochran-Mantel- Haenszel
CPS	Combined positive score
CRD	Centre for Reviews and Dissemination
CrI	Credible interval
CS	Company submission
CSR	Clinical study report
CT	Computerised tomography
CTCAE	Common Terminology Criteria for Adverse Events
DBL	Database lock
DFS	Disease-free survival
DP	Decision problem
DoR	Duration of response
DSA	Deterministic sensitivity analysis
DSU	Decision Support Unit
ECOG PS	Eastern Cooperative Oncology Group Performance Scale
EMA	European Marketing Authorisation
eMIT	Electronic market information tool
EQ-5D-3L	European Quality of Life-5 dimensions-3 levels
ERG	Evidence Review Group
ESCC	Esophageal squamous cell carcinoma
ESMO	European Society for Medical Oncology
FACT-E	Functional Assessment of Cancer Therapy- Esophagus
FACT-G	Functional Assessment of Cancer Therapy- General
FACT-G7	Functional Assessment of Cancer Therapy- General 7-items
FAD	Final Appraisal Document
FE	Fixing errors
FOLFOX	Fluorouracil and oxaliplatin
FV	Fixing violations
GC	Gastric cancer
GI	Gastrointestinal
HR	Hazard ratio
HRQoL	Health-related quality of life
HTA	Health Technology Assessment
IA	Interim analysis

ICER	Incremental cost effectiveness ratio
IMAEs	Immune-mediated adverse events
IQR	Interquartile range
ITC	Indirect treatment comparison
ITT	Intention to treat
IV	Intravenous
KSR	Kleijnen Systematic Reviews Ltd
K-M	Kaplan-Meier
LYs	Life years
MedDRA	Medical Dictionary for Regulatory Activities
MeSH	Medical Subject Headings
MID	Minimum important difference
MJ	Matters of judgement
N/A	Not applicable
NCI	National Cancer Institute
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
NIVO	Nivolumab
NMA	Network meta-analysis
N/R	Not reported
OC	Oesophageal cancer
OESI	Other events of special interest
OGJ	Oesophagogastric junction
ORR	Objective response rate
OS	Overall survival
OSCC	*Oesophageal squamous cell carcinoma*
PAS	Patient Access Scheme
PEMBRO	Pembrolizumab
PD-L1	Programmed death ligand 1
PFS	Progression-free survival
PFS2	Time to second progression
PRESS	Peer Review of Electronic Search Strategies
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-analyses
PSA	Probabilistic sensitivity analysis
PSM	Partitioned survival model
PSS	Personal Social Services
PT	Preferred terms
QALY	Quality-adjusted life year
QoL	Quality of life
Q2W	Every two weeks
Q4W	Every four weeks
Q6W	Every six weeks
RCT	Randomised controlled trial
RDI	Relative dose intensity
RECIST	Response Evaluation Criteria in Solid Tumours
SAE	Serious Adverse Events
SCC	Squamous cell carcinoma
ScHARRHUD	University of Sheffield School of Health and Related Research Health Utilities Database
SD	Standard deviation
SE	Standard error
SLR	Systematic literature review
SmPC	Summary of Product Characteristics
SOC	System organ class

SoC	Standard of care
STA	Single Technology Appraisal
TA	Technology Assessment
TC	Tumour cell
TECH-VER	Technical Verification
ToT	Time on treatment
TP	Transition probability
TPS	Tumour proportion score
TSST	Time to second subsequent therapy
TTD	Time-to-death
UK	United Kingdom
USA	United States of America
VAS	Visual analogue scale
VBA	Visual Basic for Applications
WTP	Willingness-to-pay
XELOX	Capecitabine + oxaliplatin

Table of Contents

Table of Tables

Table of Figures

1. EXECUTIVE SUMMARY

This summary provides a brief overview of the key issues identified by the Evidence Review Group (ERG) as being potentially important for decision making. If possible, it also includes the ERG's preferred assumptions and the resulting incremental cost-effectiveness ratios (ICERs).

Section 1.1 provides an overview of the key issues. Section 1.2 presents the key model outcomes. Section 1.3 discusses the decision problem, Section 1.4 issues relate to the clinical effectiveness, and Section 1.5 issues relate to the cost effectiveness while a summary in presented in Section 1.6.

Background information on the condition, technology and evidence and information on key as well as non-key issues are in the main ERG report, see Sections 2 (decision problem), 3 (clinical effectiveness) and 4 (cost effectiveness) for more details.

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

1.1 Overview of the ERG's key issues

Table 1.1: Overview of Key Issues

ID2712	Summary of issue	Report Sections
1	Uncertainty as to the appropriate comparators dependent on PD-L1 status	**Executive summary:** Table 1.2 **Main report:** Section 2.3 Section 3.2 Section 3.3 Section 3.4 Section 4
2	There is limited evidence to support the comparability of the PD-L1 ≥10% CPS populations in the two trials used in the ITC analysis.	**Executive summary:** Table 1.3 **Main report:** Section 3.4
3	It is unclear which ITC method, constant HR or time varying HRs formed the base case for the analysis.	**Executive summary:** Table 1.4 **Main report:** Section 3.4.3 Section 4
4	There is uncertainty as to the nature and effectiveness of subsequent therapy.	**Executive summary:** Table 1.5 **Main report:** Section 3.2 Section 4.2.2 Section 4.2.6 Section 4.2.9
5	There is uncertainty as to long term OS and the treatment effect of nivolumab + chemotherapy versus chemotherapy.	**Executive summary:** Table 1.6 **Main report:** Section 3.2 Section 4.2.6

ID2712	Summary of issue	Report Sections
6	There is uncertainty as to how long-term OS for the comparison of nivolumab + chemotherapy versus pembrolizumab + chemotherapy.	**Executive summary:** Table 1.7 **Main report:** Section 3.4 Section 4.2.6
7	There is uncertainty as to how all-cause mortality should be incorporated in the model.	**Executive summary:** Table 1.8 **Main report:** Section 4.2.6.2
8	There is uncertainty as to whether health state utilities should be treatment dependent or incorporate a terminal care decrement.	**Executive summary:** Table 1.9 **Main report:** Section 4.2.8
9	There is uncertainty as to the appropriate method and value of any adjustment to cost due to delayed or missed doses.	**Executive summary:** Table 1.10 **Main report:** Section 4.2.9.3
10	Calculations were missing from the model, which reduces transparency and makes updating difficult.	**Executive summary:** Table 1.11 **Main report:** Section 4.2.9.1 Section 4.2.9.4 Section 5.3
11	Health state costs were estimated from an out-of-date source.	**Executive summary:** Table 1.12 **Main report:** Section 4.2.9
12	Errors, which underestimated the cost of PEMBRO-CHEMO and prevented the PSA for PEMBRO-CHEMO comparison.	**Executive summary:** Table 1.13 **Main report:** Section 6
CPS = combined positive score; HR = hazard ratio; ITC = indirect treatment comparison; OS = overall survival; PD-L1 = programmed death ligand 1		

1.2 Overview of key model outcomes

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

Overall, the technology is modelled to affect QALYs by:

- Increasing overall survival (OS) and thus increasing time alive and delaying terminal care
- Increasing progression-free survival (PFS) and thus increasing time in the higher utility health state

Overall, the technology is modelled to affect costs by:

- Increasing PFS and thus increasing time in the lower cost health state as well as reducing the rate of relatively expensive subsequent immunotherapy
- Increasing OS and thus increasing time alive and delaying terminal care

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

- Choice of OS curve
- How subsequent treatment is modelled in terms of type, effectiveness and cost

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

The decision problem addressed in the company submission (CS) is broadly in line with the final scope issued by NICE. However, there is uncertainty regarding the appropriate comparators according to programmed death ligand 1 (PD-L1) status (Table 1.2).

Table 1.2: Key issue 1: Uncertainty as to the appropriate comparators dependent on PD-L1 status

Report Section	**Section 2.3, Section 3.2, Section 3.3, Section 3.4 and Section 4**
Description of issue and why the ERG has identified it as important	Pembrolizumab is listed as a comparator in the NICE scope. however, the company argue that it is not SoC because it was recommended too recently (20 October 2021) and so, although they conduct a sophisticated ITC, they consign a cost effectiveness analysis to a scenario. The company acknowledge that the appropriate population for the comparison of nivolumab to pembrolizumab is PD-L1 ≥1% TC and ≥10 CPS squamous histology population given that the former is required for nivolumab and the latter for pembrolizumab. However, PD-L1 ≥1% TC status is unknown in the pembrolizumab trial, KEYNOTE-590 and only mixed histology (including adenocarcinoma) PFS data are available. Nevertheless, the ERG agrees with the company that the nivolumab and pembrolizumab datasets are comparable enough (mixed histology might lead to an underestimate of the effectiveness of pembrolizumab) for an ITC and support the general methodological approach taken by the company for this ITC. Given that pembrolizumab is not a comparator in the PD-L1 ≥1% TC and <10 CPS, the only comparator is chemotherapy. The ERG requested that an analysis of CheckMate 648 be performed in this population, but the company refused to do this. Also, the HR for OS for CPS <10% appears to be higher than for PD-L1 ≥1% TC population.
What alternative approach has the ERG suggested?	Pembrolizumab should be a comparator for the PD-L1 ≥1% TC and ≥10 CPS population. A separate analysis of CheckMate 648 trial and cost effectiveness analysis based on this should be conducted for the PD-L1 ≥1% TC and <10 CPS population.
What is the expected effect on the cost effectiveness estimates?	Nivolumab has been shown to be dominated by pembrolizumab. The ICER versus chemotherapy is likely to go up in the PD-L1 ≥1% TC and <10 CPS population.
What additional evidence or analyses might help to resolve this key issue?	A separate analysis of CheckMate 648 trial and cost effectiveness analysis based on this should be conducted for the PD-L1 ≥1% TC and <10 CPS population.

CPS = combined positive score; ERG = Evidence Review Group; HR = hazard ratio; ICER = incremental cost effectiveness ratio; ITC = indirect treatment comparison; NICE = National Institute for Health and Care Excellence; OS = overall survival; PD-L1 = programmed death ligand 1; SoC = standard of care; TC = tumour cells

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

A full summary of the cost effectiveness evidence review conclusions can be found in Section 3.6 of this report. The ERG identified two major concerns with the evidence presented on the clinical effectiveness: the limited evidence of comparability of the PD-L1 ≥10% CPS population in the two trials included in the indirect treatment comparison (ITC) analysis (see Table 1.3), and the lack of clarity on ITC method used as the base case for the analysis (see Table 1.4).

Table 1.3: Key issue 2: Limited evidence of comparability of the PD-L1 ≥10% CPS population in the two trials included in the ITC analysis

Report Section	Section 3.4
Description of issue and why the ERG has identified it as important	The NMA feasibility assessment is based on the comparison of characteristics of populations that are beyond the scope of the analysis. Differences were also identified in study design, patient eligibility and treatment characteristics. The available baseline characteristics for KEYNOTE-590 are for patients in the ITT population which includes patients with SCC or adenocarcinoma, located on the oesophagus or gastroesophageal junction, and patients beyond PD-L1 ≥10% (CPS) expression. The company does not specify if the baseline characteristics of CheckMate-648 presented in the NMA feasibility assessment refer to the study's entire population or the PD-L1 ≥10% (CPS) population.
What alternative approach has the ERG suggested?	Only the populations in the scope of the ITC should be used for the feasibility assessment i.e., OSCC with PD-L1 ≥10% (CPS) expression, where data are available.
What is the expected effect on the cost effectiveness estimates?	The effect on the cost effectiveness is difficult to predict.
What additional evidence or analyses might help to resolve this key issue?	The ERG recommends the use of baseline characteristics of the narrower population from the CheckMate 648 RCT, within the scope of the ITC. The ERG recognises the lack of evidence regarding the KEYNOTE-590 RCT.

CPS = combined positive score; ERG = Evidence Review Group; ITC = indirect treatment comparison; ITT = intention to treat; PD-L1 = programmed death ligand 1; NMA = network meta-analysis; OSCC = Oesophageal squamous cell carcinoma; RCT = randomised controlled trial; SCC = squamous cell carcinoma

Table 1.4: Key issue 3: Lack of clarity on ITC method used as the base case for the analysis

Report Section	Section 3.4.3 and Section 4
Description of issue and why the ERG has identified it as important	It was unclear which ITC method, constant HR or time varying HRs formed the base case for the analysis, based on the CS and clarification letter response. The appropriateness of each method, conceptually and statistically, are based on contradicting assumptions and will affect the fitness of the models as well as their validity in both clinical and cost effectiveness sections.
What alternative approach has the ERG suggested?	The company must clarify which method was used, what were the underlying conceptual assumptions and what statistical tests were used.

Report Section	Section 3.4.3 and Section 4
What is the expected effect on the cost effectiveness estimates?	The effect on the cost effectiveness is difficult to predict.
What additional evidence or analyses might help to resolve this key issue?	In the FAC the company have provided clarification that the time varying method was used.
ERG = Evidence Review Group; HR = hazard ratio; ITC = indirect treatment comparison.	

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

A full summary of the cost effectiveness evidence review conclusions can be found in Section 6.4 of this report. The company's cost effectiveness results are presented in Section 5, the ERG's summary and detailed critique in Section 4 (see ERG comment), and the ERG's amendments to the company's model and results are presented in Section 6. The main ERG results are reproduced using confidential Patient Access Schemes (PASs), i.e. for pembrolizumab, in a confidential appendix. The key issues in the cost effectiveness evidence are discussed in Tables 1.5 to 1.11.

Table 1.5: Key issue 4: There is uncertainty as to the nature and effectiveness of subsequent therapy

Report Section	Section 3.2, Section 4.2.2, Section 4.2.6, and Section 4.2.9
Description of issue and why the ERG has identified it as important	The precise nature of subsequent therapy in NHS clinical practice is unknown. In CheckMate 648, of those who received subsequent therapy, [redacted] and [redacted] of NIVO+CHEMO and CHEMO patients received an anti-PD(-L)1, but in the economic model it is assumed that these proportions are zero and 100% respectively. In TA737, the committee acknowledged that this assumption was probably the best reflection of clinical practice. However, this implies that the treatment effect from the trial is liable to be biased upwards because the NIVO+CHEMO patients who received a subsequent anti-PD(-L)1 will have better outcomes and the CHEMO patients who did not receive a subsequent anti-PD(-L)1 will have worse outcomes than would be expected in clinical practice. There are methods for adjusting for treatment switching as set out in TSD 16 and Ouwens 2021 that could reduce this bias.
What alternative approach has the ERG suggested?	The ERG suggested adjusting the CheckMate 648 outcomes for subsequent anti-PD(-L)1 treatment in order to better reflect clinical practice, but the company did not perform this analysis, appearing to misinterpret the question.
What is the expected effect on the cost effectiveness estimates?	The ICER versus CHEMO is likely to increase.
What additional evidence or analyses might help to resolve this key issue?	An analysis is required of the CheckMate 648 to adjust outcomes for subsequent anti-PD(-L)1 treatment for treatment switching as set out in TSD 16 and Ouwens 2021 that could reduce this bias.
CHEMO = chemotherapy; ERG = Evidence Review Group; ICER = incremental cost-effectiveness ratio; NHS = National Health Service; NIVO+CHEMO = nivolumab + chemotherapy; PD-L1 = programmed death ligand 1; TA = technology appraisal; TSD = technical support document	

Table 1.6: Key issue 5: There is uncertainty as to long term OS and the treatment effect of nivolumab + chemotherapy versus chemotherapy

Report Section	Section 3.2 and Section 4.2.6
Description of issue and why the ERG has identified it as important	The company argued that the reducing hazard rate observed in the CHEMO arm is implausible and, on that basis, choose a semi-parametric modelling approach with a cut-off of 6.9 months, using the K-M data before and a parametric model after this cut-off. Little justification is provided for the implausibility, the most plausible explanation appearing to be the effect of subsequent systemic therapy, especially anti-PD(-L)1. However, the most appropriate method of addressing any bias due to this would be to adjust for treatment switching, but only to better reflect clinical practice, is not performed and might actually reduce the treatment effect, as set out in key issue 4. There is also no clear demonstration of lack of fit of parametric models to the OS data and no consideration of more complex spline-based models for PFS. Landmark analysis of CheckMate 648 and parametric OS functions seem to provide reasonable correspondence not only between CheckMate 648 and parametric extrapolation, but also between these and other trial evidence, casting doubt on the implausibility of the reducing hazard rate. Finally, despite the observation of decreasing CHEMO OS hazard and approximation of survival up to year 3 in the trial, the company reject any treatment waning. In TA737 this was considered reasonable for PEMBRO versus CHEMO and the ERG consider that the evidence of treatment waning from CheckMate 648 even earlier is compelling.
What alternative approach has the ERG suggested?	The ERG base case employs parametric modelling using the default company models and with treatment waning that starts at 2.5 years and gradually increases until there is no treatment effect (HR=1) by year 4.
What is the expected effect on the cost effectiveness estimates?	The ICER increases.
What additional evidence or analyses might help to resolve this key issue?	See key issue 4 regarding subsequent therapy. Spline-based models might also be considered.

CHEMO = chemotherapy; ERG = Evidence Review Group; HR = hazard ratio; ICER = incremental cost-effectiveness ratio; K-M = Kaplan-Meier; OS = overall survival; PD-L1 = programmed death ligand 1; PEMBRO = pembrolizumab; PFS = progression-free survival; TA = technology appraisal

Table 1.7: Key issue 6: There is uncertainty as to long-term OS and PFS for the comparison of nivolumab + chemotherapy versus pembrolizumab + chemotherapy

Report Section	Section 3.4 and Section 4.2.6
Description of issue and why the ERG has identified it as important	The company stated in the ITC HRs for pembrolizumab + chemotherapy versus chemotherapy were applied to the survival curves for chemotherapy to estimate the survival curve for pembrolizumab + chemotherapy in the comparison with nivolumab + chemotherapy. However, in the CS the gamma model for the ITC was presented, in the Appendix C the best

Report Section	Section 3.4 and Section 4.2.6
	fitting model was the log-logistic, but in the clarification letter response the company stated that the Weibull and the lognormal were used to be consistent with the base case semi-parametric models. Also, only one set of survival values were presented in the model, which hinders transparency and the nivolumab + chemotherapy and pembrolizumab + chemotherapy OS curves were found to cross, which is inconsistent with the HRs for nivolumab + chemotherapy versus pembrolizumab + chemotherapy, which are all above 1 up to 48 months.
What alternative approach has the ERG suggested?	The ERG base case employs the default company parametric model for nivolumab + chemotherapy OS and applies the HRs for nivolumab + chemotherapy versus pembrolizumab + chemotherapy to this OS curve using a method as set out by the company in the response to clarification.
What is the expected effect on the cost effectiveness estimates?	LYs and therefore QALYs for pembrolizumab + chemotherapy go up and cost goes down, so that nivolumab + chemotherapy remains less effective, but becomes a little cheaper.
What additional evidence or analyses might help to resolve this key issue?	Clarification on the method of implementation of the HRs from the ITC would be helpful.
CS = company submission; ERG = Evidence Review Group; HR = hazard ratio; ITC = indirect treatment comparison; LY = life year; OS = overall survival; PFS = progression-free survival; QALY = quality-adjusted life year	

Table 1.8: Key issue 7: There is uncertainty as to how all-cause mortality should be incorporated

Report Section	Section 4.2.6.2
Description of issue and why the ERG has identified it as important	The company add all-cause mortality as opposed to using it in the model to prevent implausibly low mortality with any OS extrapolations.
What alternative approach has the ERG suggested?	The ERG has conducted a scenario by removing all-cause mortality.
What is the expected effect on the cost effectiveness estimates?	The ICER decreases slightly.
What additional evidence or analyses might help to resolve this key issue?	Either the company should provide a better justification for the method used or change how all-cause mortality is incorporated in the model.
ERG = Evidence Review Group; ICER = incremental cost-effectiveness ratio; OS = overall survival	

Table 1.9: Key issue 8: There is uncertainty as to whether health state utilities should be treatment dependent or incorporate a terminal care decrement

Report Section	Section 4.2.8
Description of issue and why the ERG has identified it as important	A concern of the ERG is that all of the health state values, as well as the TTD ones were all higher for chemotherapy than nivolumab + chemotherapy, albeit it by a very small amount, but the company chose the treatment-independent ones from the progression-based as opposed to TTD analysis. Also, the PD-L 1≥1% values were not used in the model. Despite stating that a

Report Section	Section 4.2.8
	progression-based analysis was chosen, it is unclear why a terminal care decrement was applied, which would seem consistent with a TTD approach.
What alternative approach has the ERG suggested?	Use of treatment specific utilities including for terminal care decrement.
What is the expected effect on the cost effectiveness estimates?	Uncertain.
What additional evidence or analyses might help to resolve this key issue?	Regression analysis with all three clinically relevant covariates i.e., health state, treatment, and TTD. Reconsideration of the choice of AE disutilities with justification.
AE = adverse event; ERG = Evidence Review Group; PD-L1 = programmed death ligand 1; TTD = time to death	

Table 1.10: Key issue 9: There is uncertainty as to the appropriate method and value of any adjustment to cost due to delayed or missed doses

Report Section	Section 4.2.9.3
Description of issue and why the ERG has identified it as important	The company have reduced the cost of each component of the treatment combination according to what they state is the proportion of doses delayed. It does not seem to make sense to assume that a delayed dose would cost zero and also is inconsistent with an RDI approach that reduces cost according to missed doses, as appears to have been the method used in TA737.
What alternative approach has the ERG suggested?	The ERG has used RDI data from the CSR and a plausible assumption as to how to calculate average RDI per component.
What is the expected effect on the cost effectiveness estimates?	The ICER increased a little.
What additional evidence or analyses might help to resolve this key issue?	Either the company provides better explanation and justification for the method used or uses the RDI approach with the best available data from the trial.
CSR = clinical study report; ERG = Evidence Review Group; ICER = incremental cost-effectiveness ratio; RDI = relative dose intensity; TA = technology appraisal	

Table 1.11: Key issue 10: Calculations were missing from the model, which reduces transparency and makes updating difficult

Report Section	Section 4.2.9.1, Section 4.2.9.4 and Section 5.3
Description of issue and why the ERG has identified it as important	The Data Library tab contains the results of calculations instead of the calculations and original input data. This has made it difficult to interrogate and to update e.g., with more recent costs, or changes to subsequent treatment mix.
What alternative approach has the ERG suggested?	The ERG has made an assumption about how the costs of subsequent treatment were calculated in order to replace with alternatives.
What is the expected effect on the cost effectiveness estimates?	Greater confidence in and updated/better estimates of cost. The effect on the ICER might be small.

Report Section	Section 4.2.9.1, Section 4.2.9.4 and Section 5.3
What additional evidence or analyses might help to resolve this key issue?	The company to incorporate all calculations and input data for these calculations in the model.
ERG = Evidence Review Group; ICER = incremental cost-effectiveness ratio	

Table 1.12: Key issue 11: Health state costs were estimated from an out-of-date source

Report Section	Section 4.2.9
Description of issue and why the ERG has identified it as important	Use of older or incorrect Department of Health Drugs and pharmaceutical eMIT costs with no details of any calculations incorporated in the economic model.
What alternative approach has the ERG suggested?	The ERG has updated all NHS reference costs to 2019/20 costs and eMIT costs to 2021 costs. The ERG has been unable to update the costs in the model as there are no details of calculations.
What is the expected effect on the cost effectiveness estimates?	Estimates are up to date.
What additional evidence or analyses might help to resolve this key issue?	The company to provide updated analyses using more current cost data.
eMIT = electronic marketing information tool; ERG = Evidence Review Group; NHS = National Health Service	

Table 1.13: Key issue 12: Errors, which underestimated the cost of PEMBRO-CHEMO and prevented the PSA for PEMBRO-CHEMO comparison.

Report Section	Section 6
Description of issue and why the ERG has identified it as important	The cost of PEMBRO-CHEMO was underestimated, which could be traced to three errors: 1) a 50% discount on the price of pembrolizumab 2) inappropriate distribution for unit cost of pembrolizumab and fluorouracil, which caused an error in the PSA 3) in the PSA only the cost in the first cycle of pembrolizumab is included. This was traced to an error generated in cell L546 in the 'Survival' tab, related to the estimation of time on treatment (ToT). The ERG notes that this error is not generated in the deterministic case, but is related to a function in the VBA, which estimates a random value from a lognormal given that ToT is estimated using the exponential distribution i.e., with one parameter (the rate). This function requires a mean and standard error, but the cell that should contain the standard error is blank. In fact, it appears that the standard errors and covariance matrices for most of the survival distributions are missing in the 'Survival' tab.
What alternative approach has the ERG suggested?	Errors 1 and 2 were corrected, but the third could not be.
What is the expected effect on the cost effectiveness estimates?	Unlikely to have much of an effect.

Report Section	Section 6
What additional evidence or analyses might help to resolve this key issue?	The company needs to fix any errors.
ERG = Evidence Review Group; NHS = National Health Service	

1.6 Summary of the ERG's view

The estimated ERG base case ICER for NIVO-CHEMO versus CHEMO, based on the ERG preferred assumptions highlighted in Section 6.1, was £49,980 (deterministic) and £49,629 (probabilistic) per QALY gained. The probabilistic ERG base case analyses indicated cost effectiveness probabilities of 0.0%, 0.8% and 52.2% at willingness-to-pay (WTP) thresholds of £20,000, £30,000, and £50,000 per QALY gained.

The estimated ERG base case ICER for NIVO-CHEMO versus PEMBRO-CHEMO, based on the ERG preferred assumptions highlighted in Section 6.1, was £290,554 (SW quadrant) (deterministic) (see Table 1.13). The ERG could not produce a probabilistic value given an error in the PSA (See Table 1.13).

The most influential adjustments were implementing treatment waning from 2.5 to 4 years and using the log-logistic distribution for estimating OS and lognormal for PFS in both NIVO-CHEMO and CHEMO arms. Table 1.13 shows how individual adjustments impact the results, plus the combined effect of all adjustments. The ICER increased most in the ERG scenario analysis (conditional on the ERG base case) using the subsequent therapy mix from the CheckMate 648 trial data.

Table 1.14: Summary of ERG's preferred assumptions and ICER

Technologies	Incremental cost (£)	Incremental QALYs	ICER (£/QALY)
Company's base case			
NIVO-CHEMO versus CHEMO	[redacted]	[redacted]	33,357
Corrected end-of-life utility decrement (no effect on company base case)			
NIVO-CHEMO versus CHEMO	[redacted]	[redacted]	33,357
Matter of judgement 1: use lognormal for PFS and log-logistic for OS (key issue 5)			
NIVO-CHEMO versus CHEMO	[redacted]	[redacted]	38,177
Matter of judgement 2: application of treatment waning from 2.5 to 4 years (key issue 5)			
NIVO-CHEMO versus CHEMO	[redacted]	[redacted]	39,337
Matter of judgement 3: treatment-dependent utility values used (key issue 8)			
NIVO-CHEMO versus CHEMO	[redacted]	[redacted]	34,965
Matter of judgement 4: cost of therapy reduced according to RDI calculated by ERG (key issue 9)			
NIVO-CHEMO versus CHEMO	[redacted]	[redacted]	35,109
ERG base case (Changes 1-4)			
NIVO-CHEMO versus CHEMO	[redacted]	[redacted]	49,017
ERG base case probabilistic (1,000 runs)			
NIVO-CHEMO versus CHEMO	[redacted]	[redacted]	49,629
Company base case (deterministic)			

Technologies	Incremental cost (£)	Incremental QALYs	ICER (£/QALY)
NIVO-CHEMO versus PEMBRO-CHEMO	[redacted]	[redacted]	Dominated
ERG base case (removed 50% PEMBRO discount; log-logistic HRs for NIVO-CHEMO versus PEMBRO-CHEMO with log-logistic OS curve for NIVO-CHEMO) (deterministic)			
NIVO-CHEMO	[redacted]	[redacted]	290,554 (SW quadrant)

CHEMO = chemotherapy; ERG = Evidence Review Group; HR = hazard ratio; ICER = increment cost-effectiveness analysis; NIVO = nivolumab; OS = overall survival; PEMBRO = pembrolizumab; PFS = progression-free survival; QALY = quality-adjusted life year; RDI = relative dose intensity

2. CRITIQUE OF COMPANY'S DEFINITION OF DECISION PROBLEM

Table 2.1: Statement of the decision problem (as presented by the company)

	Final scope issued by NICE	Decision problem addressed in the CS	Rationale if different from the final NICE scope	ERG Comment
Population	Adults with unresectable advanced, recurrent or metastatic, previously untreated OSCC	Adults with unresectable advanced, recurrent or metastatic previously untreated OSCC with tumour cell PD-L1 expression ≥1%.	The evidence provided in this submission is derived from the pivotal CheckMate 648 trial, which demonstrates that the [redacted]. This population is also in line with the expected EMA licensing.	The narrower population considered in the CS is in line with the anticipated marketing authorisation for nivolumab.
Intervention	Nivolumab in combination with fluoropyrimidine- and platinum-based chemotherapy	As per NICE scope	N/A: as specified in the draft SmPC	The intervention is in line with the NICE scope
Comparator(s)	Platinum-based chemotherapy without nivolumab, such as: Doublet treatment with fluorouracil or capecitabine + cisplatin or oxaliplatin Triplet treatment with fluorouracil or capecitabine + cisplatin or oxaliplatin + epirubicin For tumours that express PD-L1 with a CPS of 10 or more: Pembrolizumab with platinum- and fluoropyrimidine-based chemotherapy	Platinum-based chemotherapy without nivolumab, such as: Doublet treatment with fluorouracil or capecitabine + cisplatin or oxaliplatin For tumours that express PD-L1 with a CPS of 10 or more: Pembrolizumab with platinum- and fluoropyrimidine-based chemotherapy	It should be noted that epirubicin-based triplet therapy is not commonly used in UK clinical practice. During TA737, the clinical expert stated that triplet therapy is no longer standard of care as it does not provide additional efficacy and increases toxicity. The committee concluded that a dual chemotherapy regimen would be the appropriate comparator for TA737. This aligns with expert advice provided to BMS. Hence, assessment of epirubicin-based triplet therapy may not be relevant to decision making for this appraisal.	The comparators are in line with the NICE scope apart from the absence of triplet treatment, but the ERG can confirm that the FAD for TA737 reported that it was not considered as standard care[1]. However, although pembrolizumab with chemotherapy is a comparator for the subgroup defined by PD-L1 CPS at least 10 and TC at least one, the data for pembrolizumab with chemotherapy in the

	Final scope issued by NICE	Decision problem addressed in the CS	Rationale if different from the final NICE scope	ERG Comment
			Further, it should also be noted that pembrolizumab was only recently recommended by NICE (October 2021) and is hence not yet SoC	ITC included all CPS at least 10 regardless of TC due to lack of data on the PD-L1 CPS at least 10 and TC at least one subgroup. Also, because pembrolizumab with chemotherapy is only recommended for this subgroup and nivolumab with chemotherapy appears not to be cost effective versus pembrolizumab with chemotherapy, nivolumab with chemotherapy would need to be compared to chemotherapy using data from the other subgroup i.e., TC at least one and CPS less than 10, which has not been done
Outcomes	The outcome measures to be considered include: OS PFS Response rate Adverse effects of treatment HRQoL	As per NICE scope	Not applicable; additional relevant clinical outcomes are presented, including duration of response, objective response rate, complete response rate and partial response rate	The outcomes reported are in line with the NICE scope
Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed	As per NICE scope	N/A – in line with the NICE final scope	The economic analysis is in line with the NICE reference case

	Final scope issued by NICE	Decision problem addressed in the CS	Rationale if different from the final NICE scope	ERG Comment
	in terms of incremental cost per quality-adjusted life year. The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. Costs will be considered from an NHS and PSS perspective. The availability of any commercial arrangements for the intervention, comparator and subsequent treatment technologies will be taken into account			
Subgroups to be considered	If evidence allows subgroups by degree of PD-L1 expression and cancer histology will be considered. If appropriate, the appraisal should include consideration of the costs and implications of additional testing for biological markers but will not make recommendations on specific diagnostic tests or devices. Guidance will only be issued in accordance with the	Pre-defined subgroups are presented for PD-L1 ≥1% and all randomised patients, in line with the NICE scope The costs for PD-L1 screening are included	As 98% of the patients included in CheckMate 648 study histologically have OSCC, no further subgroup analysis was conducted for the purpose of cost effectiveness modelling	No cost effectiveness analyses were presented for any pre-defined subgroups. However, given that pembrolizumab with chemotherapy is only recommended for the subgroup PD-L1 CPS at least 10, the ITC and the cost effectiveness analysis versus pembrolizumab with chemotherapy are effectively in the subgroup of PD-L1 TC at least one and CPS at least

	Final scope issued by NICE	Decision problem addressed in the CS	Rationale if different from the final NICE scope	ERG Comment
	marketing authorisation. Where the wording of the therapeutic indication does not include specific treatment combinations, guidance will be issued only in the context of the evidence that has underpinned the marketing authorisation granted by the regulator			10. As stated above, there should also be a comparison with chemotherapy in the PD-L1 TC at least 1 and CPS less than 10 subgroup.
Special considerations including issues related to equity or equality	None specified.	None identified.	N/A – in line with the NICE final scope	N/A

Based on Table 1 and Section B.1 of the CS[2]

BMS = Bristol-Myers Squibb; CHEMO = chemotherapy; CPS = combined positive score; CS = company submission; EMA = European Marketing Authorisation; ERG = Evidence Review Group; FAD = Final Appraisal Document; HRQoL = health-related quality of life; ITC = indirect treatment comparison; N/A = not applicable; NHS = National Health Service; NICE = National Institute of Health and Care Excellence; NIVO-CHEMO – nivolumab + chemotherapy; OS = overall survival; OSCC = oesophageal squamous cell carcinoma; PD-L1 = programmed death ligand 1; PFS = progression-free survival; PSS = Personal Social Services; SoC = standard of care; SmPC = summary of product characteristics; TA = technology appraisal; TC = tumour cells; UK = United Kingdom

2.1 Population

The population defined in the National Institute for Health and Care Excellence (NICE) final scope is: adults with unresectable advanced, recurrent or metastatic, previously untreated oesophageal squamous cell carcinoma (OSCC).[3] The population in the company submission (CS) is limited to: "*adults with unresectable advanced, recurrent or metastatic previously untreated OSCC with tumour cell PD-L1 expression ≥1%.*"[2] According to the company, the decision problem (DP) addressed in the CS is narrower than that specified in the NICE final scope, and is in line with the population recruited to the CheckMate 648 trial and the expected European Marketing Authorisation (EMA) licensing (Table 1 of the CS).[2]

A marketing authorisation application was submitted to the EMA for nivolumab in combination with fluoropyrimidine- and platinum-based combination chemotherapy for the first-line treatment of adult patients with unresectable advanced, recurrent or metastatic OSCC with tumour cell programmed death ligand 1 (PD-L1) expression ≥1% in [redacted] (Table 2 of the CS).[2] A positive opinion from the Committee for Medicinal Products for Human Use (CHMP) was received by the company in February 2022.[4, 5] The company anticipates regulatory approval and marketing authorisation for nivolumab in combination with fluoropyrimidine- and platinum-based combination chemotherapy for the first-line treatment of adult patients with unresectable advanced, recurrent or metastatic oesophageal squamous cell carcinoma (OSCC) with tumour cell PD-L1 expression ≥1% in [redacted] (Table 2 of the CS).[2]

The clinical effectiveness systematic literature review (SLR) of the CS[2] describes baseline characteristics and outcomes for the subgroup of patients with PD-L1 ≥1% as well as all randomised patients in the included trial (CheckMate 648[6, 7]). However, only patients with PD-L1 combined positive score (CPS) ≥10% from CheckMate 648 were included in the indirect treatment comparison (ITC) for comparison with pembrolizumab, a subpopulation of the group of patients with PD-L1 ≥1%[2] (discussed further in Section 3.4 of this report).

2.2 Intervention

The intervention (nivolumab in combination with fluoropyrimidine- and platinum-based chemotherapy) is in line with the NICE final scope.

The recommended dose of nivolumab is 240 mg every 2 weeks or 480 mg every 4 weeks administered intravenously over 30 minutes in combination with fluoropyrimidine- and platinum-based chemotherapy.[2] Treatment with nivolumab is recommended until disease progression, unacceptable toxicity or up to 24 months in patients without disease progression.[8]

As outlined in the Summary of Product Characteristics (SmPC), management of adverse reactions (including pneumonitis, diarrhoea, colitis, increased aspartate aminotransferase/alanine aminotransferase [AST/ALT], hyperbilirubinemia, increased creatinine, hypothyroidism, hyperthyroidism, adrenal insufficiency, diabetes, hypophysitis, rash, Stevens-Johnson syndrome, toxic epidermal necrolysis or myocarditis) may require dosing delay or permanent discontinuation depending on individual safety and tolerability.[9]

According to the company, assessment of PD-L1 status is not part of current clinical practice for patients with OSCC in the United Kingdom (UK). The company anticipates that this will change in light of the availability of pembrolizumab for patients with PD-L1 CPS ≥10.[1]

2.3 Comparators

The description of the comparators in the NICE scope is as follows:

- *Doublet treatment with fluorouracil or capecitabine plus cisplatin or oxaliplatin*
- *Triplet treatment with fluorouracil or capecitabine plus cisplatin or oxaliplatin plus epirubicin*

For tumours that express PD-L1 with a CPS of 10 or more:

- *Pembrolizumab with platinum- and fluoropyrimidine-based chemotherapy*

The company chose comparators as per the scope except for triplet treatment. This included doublet treatment, which was considered standard care in the Final Appraisal Document (FAD) of Technology Assessment (TA) 737[1]. Doublet treatment was the comparator in the CheckMate 648 trial i.e., fluorouracil 800 mg/m2/day intravenous (IV) Day 1 through Day 5 + cisplatin 80 mg/m2 IV on Day 1 of a 4-week cycle, although clinical experts in TA737 stated that oxaplatin was more commonly used than cisplatin. In fact, the company clinical experts concluded that capecitabine + oxaliplatin (XELOX) would be the primary comparator, given better tolerated and with a shorter infusion time. Therefore, the company conducted a scenario cost effectiveness analysis with XELOX as comparator, assuming only a difference in cost of treatment and that the intervention would be nivolumab added to XELOX. An additional scenario was nivolumab+FOLFOX versus FOLFOX. However, the committee in TA737 also concluded *"that there was comparable efficacy between the different dual regimens and that which combination the model used had little effect on the cost-effectiveness estimate. It therefore concluded that, although it was not reflective of clinical practice, it was appropriate to use [the dual regimen of cisplatin and fluorouracil as used in KEYNOTE-590] for decision making."* (page 20-21)[1].

The company also stated that pembrolizumab was not yet standard of care (SoC) because recommended to recently i.e., October 2021, although an ITC and cost effectiveness analysis based on the ITC were conducted.

ERG comment: The ERG considers that the comparator of doublet therapy for the PD-L1 ≥1% was appropriate. The ERG can see no reason to contest the conclusions of the committee in the FAD for TA737, which means that the doublet treatment in CheckMate 648 is the appropriate chemotherapy comparator. The ERG considers that pembrolizumab ought to be an additional comparator for ≥1% and PD-L1 (CPS) ≥10 because it is in the scope and has been recommended by NICE, regardless of current uptake. In the response to clarification the company confirmed that the population of PD-L1 ≥1% TC and ≥10 CPS population is relevant to both comparators, but the populations of PD-L1 ≥1% TC and less than 10 CPS is relevant to only chemotherapy. On this basis, the ERG requested that the company perform an effectiveness analysis:

- versus pembrolizumab for the PD-L1 ≥1% TC and ≥10 CPS population, to which the company replied that these data are not provided in KEYNOTE-590, precluding such a comparison. The company went on to argue that the PD-L1 ≥1% TC and ≥10 CPS population in CheckMate 648 would *"...enrich the population with those patients likely to have best response by both PD-L1 assessment criteria. As this enriched subgroup would be compared against the published KEYNOTE-590 PD-L1 CPS ≥10%, which would include patients with PD-L1 TC < 1%, this would be a biased comparison."* They also stated that *"Clinically, this subpopulation does not exist as medical decisions which drug to use would be based on CPS or TC."* Nevertheless, the company provided an ITC analysis versus pembrolizumab for the PD-L1 ≥1% TC and ≥10 CPS population, although with a note of caution given that *"...only an overlap analysis could be conducted as PD-L1 TC ≥1% data was not available from KEYNOTE-590 (see Table 6 of updated Appendix L -*

NMA report).” It appears that the term ‘overlap analysis’ simply refers to the discrepancy in PD-L1 status between the two trials. The company also pointed out that the data from KEYNOTE-590 were not from OSCC, but “mixed histology”. The ERG interprets this as referring to the inclusion of adenocarcinoma. The issue of histology was referred to in the FAD for TA737: the clinical experts *“…explained that it is possible that people with squamous cell carcinoma (who appear to be more sensitive to immunotherapies) would benefit more from pembrolizumab than people with adenocarcinoma. However, the magnitude of benefit is smaller between the 2 cancer types when CPS is 10 or more.”* The ERG therefore concludes that the effectiveness of pembrolizumab in the index population i.e., OSCC might have been underestimated in the ITC (see Sections 3.3. and 3.4). The size of this effect is unknown, although an indication might be obtained from the subgroup analysis of KEYNOTE-590. The ERG requested an unredacted version, but this was not available. However, although not mentioned in the response to clarification, Table 6 of updated Appendix L – the network meta-analysis (NMA) report shows that the so-called ‘overlap analysis’ for overall survival (OS) was in the OSCC subgroup, unlike the one for progression-free survival (PFS), which is in the mixed histology population. This was clarified in the FAC.[10] Therefore, for OS, mixed histology is not a problem. In conclusion, lack of full information on PD-L1 status and histology means that this is a key issue.

- versus chemotherapy in the PD-L1 ≥1% TC and less than CPS <10% population, to which the company replied that CheckMate 648 *“is not powered for an analysis that would include fewer patients as suggested when restricting further to patients with PD-L1 ≥ 1% TC and CPS <10%.”* The company also stated that *“the HR for OS for the patient population with CPS <10%, as presented in Table 7.4.2.2-1 of the CSR, is* [redacted] *and taking the confidence interval into account comparable to PD-L1 <1% TC,* [redacted] *as stated in Table 7.4.1.1-1, which suggests that an analysis as requested probably will not have a huge impact on the results.”* However, the ERG does not understand why comparison was made with the PD-L1 <1% tumour cell (TC) population given that this was not the one used in the CS, but the PD-L1 ≥1% TC instead, the hazard ratio (HR) for which at DBL [redacted] is [redacted]. The confidence interval (CI) for this does not include [redacted]. The company did reiterate that chemotherapy is still standard of care (SoC) regardless of PD-L1 status. However, given that pembrolizumab might have replaced chemotherapy in the PD-L1 ≥1% TC and ≥10 CPS population, the lack of comparison with chemotherapy in the PD-L1 ≥1% TC and less than CPS <10% population is a key issue.

Generally, the discrimination between these two populations is not only necessary because of potential differences in effectiveness but is also made feasible in clinical practice by testing for both TC and CPS being likely to be routine, as the company stated in response to clarification: *“It is likely that it will become routine practice to assess both TPS and CPS during the same test to determine which OSCC patients are suitable for either pembrolizumab or nivolumab treatment.”*

2.4 Outcomes

The following outcome measures were listed in the NICE final scope:[3]

- OS
- PFS
- Response rate
- Adverse effects of treatment
- Health-related quality of life (HRQoL)

The company's outcomes as represented in the DP were in line with the NICE final scope.[2] All outcomes listed in the scope were assessed in the CheckMate 648 randomised controlled trial (RCT). Additional outcomes measured in CheckMate 648 included duration of response (DoR), time to second progression (PFS2) time to second subsequent therapy (TSST) as well as further exploratory outcomes including biomarker levels, immunogenicity and pharmacokinetics as described in the trial protocol.[11]

ERG comment: The outcomes presented in the DP were in line with those listed in the NICE final scope.

2.5 Other relevant factors

The NICE final scope stated that "*If evidence allows, subgroups by degree of PD-L1 expression and cancer histology will be considered*" and that "*If appropriate, the appraisal should include consideration of the costs and implications of additional testing for biological markers but will not make recommendations on specific diagnostic tests or devices.*"[3] Within their consideration of the DP the company mentioned that pre-defined subgroups according to PD-L1 ≥1% status were presented (for the CheckMate 648 RCT) and that the cost of screening for PD-L1 status was included (in the economic analysis). In terms of subgroups based on cancer histology, the company stated that: "*As 98% of the patients included in CheckMate 648 study histologically have OSCC, no further subgroup analysis was conducted for the purpose of cost-effectiveness modelling.*"[2]

The company claimed NICE end-of-life criteria for nivolumab with chemotherapy (Section B.2.13.4.3 of Document B).[2] The associated ERG critique can be found in Section 7 of this report.

According to the company, no equality issues related to the use of nivolumab with chemotherapy for the treatment of adults with OSCC were identified or foreseen (CS, Section B.1.4).[2]

ERG comment: The company's "other considerations" presented as part of the DP were in line with those stated in the NICE final scope in terms of population subgroups.

3. CLINICAL EFFECTIVENESS

3.1 Critique of the methods of review(s)

The company conducted a SLR to evaluate the clinical effectiveness (efficacy and safety) of interventions for first-line, advanced unresectable, recurrent or metastatic oesophageal cancer (OC), with a focus on studies recruiting patients with OSCC.[12] Section 3.1 critiques the methods of the review including: the search strategy; study inclusion criteria; data extraction; assessment of risk of bias; and data synthesis.

3.1.1 Searches

The following paragraphs contain summaries and critiques of all searches related to clinical effectiveness presented in the CS.[2] The Canadian Agency for Drugs and Technologies in Health (CADTH) evidence-based checklist for the Peer Review of Electronic Search Strategies (PRESS), was used to inform this critique.[13, 14] The CS[2] was checked against the Single Technology Appraisal (STA) specification for company/sponsor submission of evidence.[15] The ERG has presented only the major limitations of each search strategy in the report.

Appendix E of the CS details the SLR undertaken to identify the clinical effectiveness evidence (efficacy and safety) of interventions for the treatment of unresectable advanced recurrent or metastatic previously untreated OSCC.[12] The searches were conducted in January 2021, and updated in October 2021. The same search strategies were used in the original SLR and the update.

A summary of the sources searched is provided in Table 3.1.

Table 3.1: Data sources for the clinical effectiveness systematic review (as reported in CS)

Resource	Host/Source	Date Ranges	Dates searched
Electronic databases			
MEDLINE	Ovid	1946-12.01.21 1946-01.10.21	14.01.21 04.10.21
Embase	Ovid	1974-Week 2/2021 1974-01.10.21	14.01.21 04.10.21
Embase (conference search)	Ovid	2019-Week 11/2021 2021-04.10.21	17.03.21 05.10.21
CENTRAL	EBM (Ovid)	to Dec 2021 to Oct 2021	14.01.21 04.10.21
Other			
ClinicalTrials.gov	Internet	All to date	Not stated
CENTRAL = Cochrane Central Register of Controlled Trials			

ERG comment:

- Searches were undertaken to identify the clinical effectiveness evidence (efficacy and safety) of interventions for the treatment of unresectable advanced recurrent or metastatic previously untreated OSCC. The CS, Appendix E and the company's response to clarification provided sufficient details for the ERG to appraise the literature searches.[2, 12, 16]
- A good range of databases was searched, and searches for named conferences were conducted via Embase. ClinicalTrials.gov was searched to identify trials with posted results that had not yet been

released in an abstract or full-text publication. No Health Technology Assessment (HTA) resources or other grey literature sources appear to have been searched.

- Searches were well-structured, transparent and reproducible.
- The search strategies combined terms for oesophageal cancer with neoplasm metastasis terms. A good range of subject indexing terms (Medical Subject Headings (MeSH)/EMTREE) and free text was used.
- Trials filters were used to limit the Embase and MEDLINE searches. The filters were not referenced, so it was unclear whether they were published and objectively-derived. These filters contained a facet which aimed to exclude other study designs using the Boolean NOT operator. Although this was conducted cautiously, there is still a risk that potentially relevant records were missed by this approach.
- Separate adverse events (AEs) searches were not performed. Guidance by the Centre for Reviews and Dissemination (CRD) recommends that if searches have been limited by a study design filter, additional searches should be undertaken to ensure that AEs that are long-term, rare or unanticipated are not missed.[17]
- Most searches were not limited by publication date. Conference proceedings searches had a 2019-2021 date limit applied.
- MEDLINE and Embase searches were limited to English language publications only. The ERG was concerned that limiting the searches to English language may have introduced potential language bias. Current best practice states that '*Whenever possible review authors should attempt to identify and assess for eligibility of all possibly relevant reports of trials irrespective of language of publication*'[18] and that '*research related to language bias supports the inclusion of non-English studies in systematic reviews*'.[19, 20]. The company was asked to assess the impact of this restriction (clarification question A8). In their response, the company provided the number of hits for Embase and MEDLINE for the January 2021 search for searches with and without the English language restriction. This indicated that 235 (6%) studies from Embase and 274 (14%) from MEDLINE were omitted on the basis of non-English language and were therefore not screened.[16] This suggests that the possibility of excluding relevant studies from the SLR (language bias) cannot be discounted.
- Appendix E states that '*to further increase search sensitivity, reference lists of relevant SLRs and meta-analyses identified in the database search were searched for relevant citations*'.[12] However, as systematic reviews and meta-analyses were removed by the study design filters in the Embase and MEDLINE searches, it is unclear how these were identified.

3.1.2 Inclusion criteria

As outlined above, the company performed an SLR to evaluate the evidence on clinical effectiveness (efficacy and safety) of interventions for first-line, advanced unresectable, recurrent or metastatic OC, with a focus on studies recruiting patients with OSCC.[12] The study eligibility criteria for the SLR are summarised in Table 3.2 below. Within the population domain, the ERG noted the distinction between global inclusion criteria (i.e., OC in general) and OSCC-specific inclusion criteria.[12]

Table 3.2: Eligibility criteria used in the SLR of clinical effectiveness evidence

	Inclusion criteria	Exclusion criteria
Population	**Global inclusion criteria** Adult patients with previously untreated, unresectable advanced or metastatic OC. **OSCC-specific inclusion criteria**	None stated.

	Inclusion criteria	Exclusion criteria
	Adult patients with previously untreated, advanced or metastatic OSCC.	
Interventions	Any of the following as monotherapy or in combination with one or more of the other treatments: • Nivolumab • Anthracycline • Capecitabine • Carboplatin • Cetuximab • Cisplatin • Docetaxel • Epirubicin • Exaliplatin • Fluorouracil • Ipilimumab • Irinotecan • Leucovorin • Oxaliplatin • Paclitaxel • Pembrolizumab	None stated.
Comparators	• Placebo • Observation • Physician's choice • BSC • Any intervention of interest • Any treatment that facilitates an indirect comparison	• Radiotherapy as monotherapy • Radiotherapy in combination with eligible interventions
Outcomes	**Eligible studies reported at least one of the following outcomes:**[a] • OS • DFS • PFS • TTP • DMFS **In addition, the following outcomes were extracted:** • Any AEs • All-cause grade 3 or 4 AEs • Overall discontinuations	None stated.
Study design	RCTs only	Non-randomised studies
Language restrictions	Studies published in English only	Studies published in languages other than English

	Inclusion criteria	Exclusion criteria
Based on Table 1 of Appendix E of the CS[12] [a]Only efficacy outcomes were used for study selection. If reported, data on all eligible outcomes listed above were extracted.[12] AE(s) = adverse event(s); BSC = best supportive care; CS = company submission; DFS = disease-free survival; DMFS = distant metastasis-free survival; OC = oesophageal cancer; OS = overall survival; OSCC = oesophageal squamous cell carcinoma; PFS = progression-free survival; RCT = randomised controlled trial; SLR = systematic literature review; TTP = time to progression		

ERG comment:

- The population specified for the clinical effectiveness SLR[12] is similar to that defined in the NICE final scope[3] except that recurrent disease is not mentioned within the inclusion criteria for the SLR.[12] However, the company specifies a population subgroup in the DP (OSCC with PD-L1 with TC ≥1%). This population definition is in line with the anticipated marketing authorisation for nivolumab.[2]
- The intervention described in the NICE final scope[3] and DP[2] is more specific than that specified within the SLR study eligibility criteria.[12]
- The comparators described in the NICE final scope[3] are more specific than those listed within the SLR eligibility criteria;[12] and the scope mentions a specific comparator for the population subgroup (OSCC with PD-L1 CPS ≥10).[3]
- The SLR eligibility criteria for comparators do not discuss doublet and triplet therapies[12] which is given consideration in the DP (which concluded that triplet therapy was not favoured).[2] The broad SLR criteria seem to imply that triplet chemotherapy would have been included.[12]
- Some outcomes in the NICE final scope[3] are not mentioned in the SLR eligibility criteria (response rate and HRQoL).[12] However, a separate SLR was conducted for HRQoL (Appendix I).[21]
- The number of included and excluded studies was unclear with discrepancy between the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) diagram and narrative description in the CS.[2] The company clarified the relevant numbers of studies in their response to clarification question A3[16] (further detail is provided in Section 3.2.1 of this report).
- The SLR study design criteria specified RCTs only for inclusion[12] and this approach may have missed relevant data on AEs.
- The study selection process (described on page 9 of Appendix E) is satisfactory.[12]

3.1.3 Critique of data extraction

Two independent reviewers extracted data into a Microsoft Excel workbook. Disagreements were resolved by discussion or if necessary, by consulting a third reviewer. The list of data extracted from each study is described on pages 9 to 10 of Appendix E. Study characteristics (as outlined in the methods) are presented in a series of tables with an accompanying narrative summary on pages 13 to 32 of Appendix E.[12]

ERG comment: The data extraction process is satisfactory and has followed recommended good practice in systematic reviews.[22]

3.1.4 Quality assessment

According to Document B of the CS (Section B.2.5),[2] the quality assessment of studies included in the SLR was based on guidance from the CRD.[17]. However, in Appendix E of the CS[12] a separate assessment was described, using the Cochrane Risk of Bias Tool (original version).[23]

The checklist based on the CRD guidance included criteria relating to the following: randomisation; allocation concealment; baseline comparability; blinding of participants, care providers and outcome assessors; comparability of withdrawals; outcome measurement; and intention-to-treat (ITT) analysis.[17] There was no mention of the number of reviewers involved in applying the instrument, nor of the approach used for deriving an overall risk of bias rating per study. The CRD-based quality assessment checklist was only applied to the CheckMate 648 RCT (as shown in Table 7 of the CS) and not for all studies included in the SLR as stated in Section B.2.5 of the CS.[2]

The Cochrane Risk of Bias Tool (original version) included seven domains relating to: sequence generation; allocation concealment; blinding of participants and study personnel; blinding of outcome assessors; incomplete outcome data; selective outcome reporting; and other sources of bias.[23] Appendix E outlined the following approach for deriving the overall risk of bias for an individual RCT:

- low risk of bias (low risk of bias in all domains)
- unclear risk of bias (unclear risk of bias for one or more domains), and
- high risk of bias (high risk of bias for one or more domains)

Appendix E described the application of the tool, stating that two independent reviewers assessed the risk of bias in the included RCTs. Disagreements were resolved by discussion or if necessary, by consulting a third reviewer.[12]

Twelve RCTs considered relevant to the submission (including CheckMate 648) were assessed using the Cochrane Risk of Bias Tool (original version)[23] as shown in Figure 2 of Appendix E of the CS.[12]

The results of both quality assessment methods together with the ERG critique are shown in Section 3.2.4 of this report.

ERG comment: Although the quality assessment method as executed was not described in full in Document B of the CS,[2] the overall approach was satisfactory and in line with recent recommendations for good practice in systematic reviews.[22] The choice of the Cochrane Risk of Bias Tool (although now superseded by a more recent version[24]) is appropriate for assessing risk of bias in RCTs. However, the risk of bias assessment was conducted at the trial level rather than the outcome level (the latter being suggested by Table 17 of Appendix E[12] and recommended by Cochrane[23]). Furthermore, the specific assessments covered under "*Other sources' of bias*" (an optional domain) should be described as part of the review protocol/methods according to Cochrane.[23] It was apparent from Figure 2 in Appendix E that whilst this domain was evaluated for each included RCT, precisely what was assessed was not explained.[12]

3.1.5 Evidence synthesis

Within the CS,[2] it was stated that pairwise meta-analysis was not undertaken because data from a direct comparison between nivolumab with chemotherapy versus chemotherapy was only available from the CheckMate-648 RCT.[6, 7] However, data from another RCT (KEYNOTE-590[25]) comparing pembrolizumab with chemotherapy versus chemotherapy were combined with CheckMate 648 in an ITC analysis. Further details are provided in Sections 3.3 and 3.4 of this report.

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

3.2.1 Study retrieval

The company confirmed the number of records retrieved and screened as part of their response to the clarification letter (questions A3 and A9).[16] In total, 7,634 records were identified from the combined January 2021 and October 2021 database searches for the CS clinical effectiveness SLR. Following deduplication, 2,489 records were removed, leaving 5,145 for title and abstract screening. As a result of the latter, 4,786 records were excluded, and 362 full-text reports were assessed for eligibility (this included three records identified from sources other than database searching). Of these, 297 records were excluded because of irrelevant: study designs (n=16); populations (n=135); interventions (n=28); outcomes (n=88); or because of other reasons (n=29); or article duplication (n=1). This left 65 records (reporting 43 unique trials) in the "*global SLR*" as described by the company,[12] of which 18 records (reporting 12 unique trials) were considered as potentially relevant to the submission.[16] These 12 RCTs are described below (all recruited patients with OSCC unless otherwise stated):[12]

- **Bleiberg 1997** is a phase II RCT comparing cisplatin versus 5-FU + cisplatin.[26]
- **CheckMate 648** is a three-arm, open-label, phase III RCT comparing nivolumab + fluorouracil and cisplatin (NIVO-CHEMO) versus nivolumab + ipilimumab versus fluorouracil and cisplatin (CHEMO) in patients with unresectable advanced, recurrent or metastatic, previously untreated OSCC.[6]
- **ESCORT-1st** is a phase III, double-blind RCT comparing camrelizumab + cisplatin + paclitaxel versus cisplatin + paclitaxel.[27]
- **Ezdinli 1980** is a three-arm, open-label, phase II RCT comparing 5-FU versus methotrexate versus adriamycin.[28]
- **JUPITER-06** is a phase III, double-blind RCT comparing toripalimab + cisplatin + paclitaxel versus cisplatin + paclitaxel.[29]
- **KEYNOTE-590** is a phase III, double-blind RCT comparing pembrolizumab + cisplatin versus 5-FU + cisplatin in patients with squamous cell carcinoma (SCC) or adenocarcinoma of the oesophagus or gastro-oesophageal junction.[25]
- **Lee 2015** is an open-label, phase II RCT comparing capecitabine + paclitaxel versus capecitabine + cisplatin.[30]
- **Lorenzen 2009** is an open-label, phase II RCT comparing cetuximab + 5-FU + cisplatin versus 5-FU + cisplatin.[31]
- **ORIENT-15** is a phase III, double-blind RCT comparing sintilimab + cisplatin + paclitaxel versus cisplatin + paclitaxel.[32]
- **POWER** is an open-label, phase III RCT comparing panitumumab + 5-FU + cisplatin versus 5-FU + cisplatin.[33]
- **Wang 2017** is a three-arm, open-label, phase II RCT comparing two different dosing schedules of recombinant human lymphotoxin-alpha derivative (rhLTα-Da) + 5-FU + cisplatin versus a control regimen consisting of 5-FU + cisplatin.[34]
- **Yao 2018** compared 5-FU + paclitaxel or cisplatin + paclitaxel versus cisplatin + paclitaxel (methodological details not reported).[35]

Of the above-listed RCTs, only CheckMate 648 provided data on the efficacy and safety of NIVO-CHEMO compared with CHEMO in adults with unresectable advanced, recurrent or metastatic, previously untreated OSCC.[6] As such, this was considered as the only RCT of direct relevance to this appraisal. Further details of CheckMate 648 are summarised in this Section.

A second RCT (KEYNOTE-590)[25] was included in the NMA of the CS[2] and is described in Section 3.3.

ERG comment: The number of records retrieved, screened and included from the database searches for the clinical effectiveness SLR was not clear from the initial CS.[2] However, the pertinent details were clarified by the company's response to the clarification letter (question A3) in which they provided updated PRISMA flow diagrams for the January 2021 and October 2021 database searches.[16]

3.2.2 Summary of details for the CheckMate 648 RCT

The only identified direct data comparison regarding the efficacy and safety of NIVO-CHEMO versus CHEMO in adult patients with unresectable advanced, recurrent or metastatic, previously untreated OSCC was the CheckMate 648 RCT. This phase III, open-label RCT was conducted across 25 countries.

The objective of CheckMate 648 was to ascertain the efficacy and safety of NIVO-CHEMO compared with CHEMO in adult patients with unresectable advanced, recurrent or metastatic, previously untreated OSCC. The primary outcomes were OS and PFS, both in patients with PD-L1 expression ≥1%. Secondary outcomes were: OS and PFS, both in all randomised patients; and objective response rate (ORR) in patients with PD-L1 expression ≥1% as well as all randomised patients. Details of exploratory outcomes as well as further details regarding the design and methods of CheckMate 648 are presented in Table 3.3.

Table 3.3: Trial design and methods of the CheckMate 648 RCT

Parameter	Description
Trial objective	To evaluate the efficacy and safety of NIVO-CHEMO compared with CHEMO in adult patients with unresectable advanced, recurrent or metastatic, previously untreated OSCC.
Trial design	Phase III, multicentre, randomised, open-label trial N=939 patients randomised Randomisation ratio 1:1:1. Randomisation stratification variables: PD-L1 status; region; ECOG PS; number of organs with metastases
Trial registry number	NCT03143153
Trial location	USA, Argentina, Austria, Brazil, Canada, Chile, China, Colombia, Czechia, Denmark, France, Hong Kong, Italy, Japan, Korea, Mexico, Peru, Poland, Romania, Russia, Singapore, Spain, Taiwan, Turkey, UK (five centres in the UK, included [redacted] randomised patients)
Trial status	Ongoing
Population (participant eligibility criteria)	**Eligible:** • Adult patients with unresectable advanced, recurrent or metastatic, previously untreated OSCC • Male or female at least 18 years of age • Must have histologically confirmed SCC or adenosquamous cell carcinoma of the oesophagus (predominant squamous differentiation) • Patients must have unresectable advanced, recurrent or metastatic OSCC • Patients must not be amenable to curative approaches such as definitive chemoradiation and/or surgery • No prior systemic or anticancer therapy given as primary therapy for advanced, metastatic disease. Prior adjuvant, neoadjuvant, or definitive, chemotherapy/radiotherapy/chemoradiotherapy for ESCC was permitted if given as part of curative intent regimen and completed before enrolment. A minimum 24-week recurrence-free period was required after completion of neoadjuvant or adjuvant chemotherapies or after completion of multimodal therapies for locally advanced disease. • ECOG PS of 0 or 1 • Patients must have at least one measurable lesion by CT or MRI per RECIST 1.1 criteria (radiographic tumour assessment must be performed within 28 days prior to randomisation) • Tumour tissues must be provided for biomarker analyses • Patient must have PD-L1 expression classification ≥1% or <1% or indeterminate as determined by the central lab

Parameter	Description
	Not eligible: • Patients must have recovered from the effects of major surgery or significant traumatic injury at least 14 days before randomisation • Prior malignancy requiring active treatment within the previous 3 years except for locally curable cancers that have been apparently cured, such as basal or squamous cell skin cancer, superficial bladder cancer or carcinoma in situ of the prostate, cervix or breast • Patients with active, known or suspected autoimmune disease. Patients with type I diabetes mellitus residual hypothyroidism due to autoimmune thyroiditis only requiring hormone replacement, skin disorders not requiring systemic treatment are permitted to enrol • Patients with a condition requiring systemic treatment with either corticosteroids (>10 mg daily prednisone equivalent) or other immunosuppressive medications within 14 days of start of study treatment. Inhaled or topical steroids and adrenal replacement steroid doses >10 mg daily prednisone equivalent are permitted in the absence of active autoimmune disease • Prior treatment with an anti-PD-1, anti-PD-L1, anti-PD-L2, anti-CD137 or anti-CTLA-4 antibody or any other antibody or drug specifically targeting T-cell co-stimulation or checkpoint pathways
Intervention	NIVO-CHEMO: Nivolumab at a dose of 240 mg administered IV over 30 minutes every 2 weeks (Q2W), with fluorouracil and cisplatin administered every 4 weeks (Q4W) Intervention 1 (n = 321*): NIVO+CHEMO: nivolumab 240 mg Q2W IV + fluorouracil 800 mg/m2/day IV on Day 1 through Day 5 + cisplatin 80 mg/m2 IV Q4W on Day 1 of a 4-week cycle Intervention 2 (n = 325*): NIVO+IPI nivolumab 3 mg/kg Q2W IV + ipilimumab 1 mg/kg every 6 weeks (Q6W) IV. One cycle = 2 weeks. Treat until progression or unacceptable toxicity, to a maximum period of 2 years. Intervention 2 is not part of this submission.
Comparator	CHEMO: Fluorouracil with cisplatin administered Q4W Comparator arm (n = 324*): CHEMO: fluorouracil 800 mg/m2/day IV Day 1 through Day 5 + cisplatin 80 mg/m2 IV Q4W on Day 1 of a 4-week cycle
Permitted concomitant medications	Topical, ocular, intra-articular, intranasal and inhalational corticosteroids, with minimal systemic absorption. Adrenal replacement steroid doses (>10 mg daily prednisone). A brief (<3 weeks) course of corticosteroids for prophylaxis or for treatment of non-autoimmune conditions is permitted. Regular concomitant use of bisphosphonates and RANK-L inhibitors for prevention or reduction of skeletal-related events in patients with bone metastasis is allowed if initiated prior to first dose of study therapy. Palliative radiotherapy was permitted for

Parameter	Description
	patients without evidence of progression per RECIST 1.1 provided the lesions were non-target lesions and this was discussed and approved by the BMS Clinical Trial Physician (Medical Monitor). Patients with evidence of progression per RECIST 1.1 must have met criteria to continue treatment beyond progression in order to resume immunotherapy after palliative local therapy.
Disallowed concomitant medications	The following medications are prohibited during the trial (unless utilised to treat a treatment-related AE): • Immunosuppressive agents • Immunosuppressive doses of systemic corticosteroids (some exemptions – see "Permitted") • Any concurrent anti-neoplastic therapy (i.e., chemotherapy, hormonal therapy, immunotherapy, extensive, non-palliative radiation therapy, or standard or investigational agents for the treatment of OC). • Botanical formulations with an approved indication for cancer treatment [e.g., traditional Chinese medicines]; these should be discontinued (if used) at least 2 weeks prior to randomisation. • Any live/attenuated vaccine (e.g., varicella, zoster, yellow fever, rotavirus, oral polio and measles, mumps, rubella [MMR]) during treatment and until 100 days post last dose
Primary outcomes	OS per BICR for all patients and patients with tumour cell PD-L1 ≥1% PFS per BICR for all patients and patients with tumour cell PD-L1 ≥1%
Secondary outcomes	OS per investigator for all patients and patients with tumour cell PD-L1 ≥1% PFS per investigator for all patients and patients with tumour cell PD-L1 ≥1% ORR in patients with PD-L1 expression ≥1% and in all randomised patients

Key exploratory outcomes	Safety and tolerability:
	• Incidence of:
	  • AEs,
	  • Serious adverse events (SAEs),
	  • AEs leading to discontinuation
	  • AEs leading to dose modification
	  • Select AEs
	  • Immune-mediated AEs
	  • Other events of special interest (OESI)
	  • Deaths
	• Laboratory abnormalities

Parameter	Description
	• PFS as assessed by investigators in patients with PD-L1 expression ≥1% and all randomised patients • ORR as assessed by investigators in patients with PD-L1 expression ≥1% and all randomised patients • DOR as assessed by BICR and by investigators in patients with PD-L1 expression ≥1% and all randomised patients • PFS2/TSST in patients with PD-L1 expression ≥1% and all randomised patients • QoL, measured using the EQ-5D-3L descriptive system and VAS, as well as the FACT-E questionnaire (including the Esophageal Cancer Subscale [ECS] and FACT-G7) Please see the study protocol for further exploratory endpoints, including biomarker analysis, immunogenicity, and pharmacokinetics
Supports marketing authorisation?	Yes
Used in the economic model?	Yes, via an ITC
Rationale for use in the economic model	Source of direct comparative evidence evaluating the efficacy of nivolumab with chemotherapy versus chemotherapy alone in the indicated patient population
Pre-planned subgroups	Age (< 65, ≥ 65 and ≥ 75) Sex Region (Asia and non-Asia) ECOG PS (0 and 1) Number of organs with metastasis (≤ 1 and ≥ 2) Disease stage at current diagnosis Smoking status Alcohol use PD-L1 CPS subgroups: ≥1%, ≥5% and ≥10%

Based on Section B.2.2 and Tables 3, 4, 5 and 6 and figure 7 of the CS.[2]
*Number of randomised patients (from Table 4 of the CS).[2]
AE = adverse event; BICR = blinded independent central review; BMS = Bristol-Myers Squibb; CHEMO = chemotherapy; CS = company submission; CT = computerised tomography; DOR = duration of response; ECOG = Eastern Oncology Cooperative Group; ESCC = esophageal squamous cell carcinoma; ECS = esophageal cancer subscale; EQ-5D-3L = EuroQol-five dimension-three level; FACT-E = Functional Assessment of Cancer Therapy – Esophageal; FACT-G7 = Functional Assessment of Cancer Therapy - General – 7-item version; ITC = indirect treatment comparison; IV = intravenously; MMR = measles, mumps, rubella; MRI = magnetic resonance imaging; NIVO-CHEMO = nivolumab + chemotherapy; NIVO+IPI = nivolumab + ipilimumab; OC = oesophageal cancer; OESI = other events of special interest; ORR = objective

Parameter	Description
response rate; OS = overall survival; OSCC = oesophageal squamous cell carcinoma; PD-L1/2 = programmed death ligand 1/2; PFS = progression-free survival; PFS2 = time to second progression; PS = performance status; Q2W = every two weeks; Q3W = every three week; Q4W = every four weeks, Q6W = every six weeks; RANK-L = Receptor activator of nuclear factor kappa-B ligand; RCT = randomised controlled trial; RECIST = Response Evaluation Criteria in Solid Tumours; SAE = serious adverse event; TSST = time to second subsequent therapy; UK = United Kingdom; USA = United States of America; VAS = visual analogue scale	

3.2.3 Statistical analysis for the CheckMate 648 RCT

The primary, secondary and exploratory outcomes of the CheckMate 648 RCT are listed in Table 3.3 above. The company did not provide further tabulation regarding statistical methods but presented information in a series of narrative Sections in the CS: Sections B.2.4.1 to B.2.4.6 (inclusive).[2] A brief, paraphrased outline of the methods described in Sections B.2.4.1, B.2.4.2 and B.2.4.3 is provided here, followed by tabulation of the information in Sections B.2.4.4, B.2.4.5 and B.2.4.6. Of note, there were two Sections numbered as B.2.4.2 in the CS: Sample size and power calculation (starts on page 36); and Timing of analysis of primary endpoints starts on page 37).

The company provided an overview of the plan for hypothesis testing in Section B.2.4.1 of the CS. Both primary outcomes (OS and PFS, both in patients with PD-L1 ≥1%) were tested first, and in parallel. Secondary outcomes (OS and PFS, both in all randomised patients; and ORR in all randomised patients and in those with PD-L1 ≥1%) were tested only if the corresponding primary outcomes were significant.[2]

Study sample size calculations (described in full in Section B.2.4.2 starting on page 36 of the CS)[2] were based on the primary outcomes and assumed that the prevalence of patients with PD-L1 ≥1% was approximately 50% and that event rates in the experimental arms according to blinded independent central review (BICR) were: OS 15% in patients with PD-L1 ≥1%; OS 10% for PD-L1 <1%; and PFS 0%. The sample size estimation for the comparison between NIVO-CHEMO and CHEMO was as follows:

- [redacted] *PFS events in approximately 313 subjects with PD-L1 ≥1% would provide approximately 90% power to detect an average hazard ratio (HR) of 0.62 with a Type I error of 1.5% (two-sided).*
- [redacted] *OS events in approximately 313 subjects with PD-L1 ≥1% would provide approximately 90% power to detect an average HR of 0.6 with a Type I error of 1% (two-sided).*"

The CS also stipulated the following: "*In case the significance level from the corresponding primary endpoint in patients with PD-L1 ≥1% was passed to the secondary endpoint in all randomised subjects:*

- [redacted] *PFS events in approximately 626 patients (all comers) would provide approximately 90% power to detect an average HR of 0.72 with a Type I error of 1.5% (two sided).*
- [redacted] *OS events in approximately 626 patients (all comers) would provide approximately 94% power to detect an average HR of 0.68 with a Type I error of 1% (two sided).*"

Details of the timing of primary outcomes analysis was provided in Section B.2.4.2 (starting on page 37 of the CS). This included the following statements:[2]

"*Final PFS analysis was planned when* [redacted] *events by BICR were observed among the patients with PD-L1 expression ≥ 1% in the CHEMO arm. This was expected to be reached after approximately 33 months.*

Final OS analysis was planned when [redacted] *events were observed among the patients with PD-L1 expression ≥ 1% in the CHEMO arm. This was expected to be reached after approximately 49 months.*

However, Revised Protocol 05 specified that if the planned number of PFS events per BICR was unlikely to be reached for unforeseen reasons, the final PFS per BICR analysis could occur when at least 12 months minimum follow-up (defined as the time from the date when the last patient was randomised to the clinical cut-off date) was reached. Indeed, the primary analyses of final PFS per BICR and interim

analysis of OS in all randomised subjects with tumour cell PD-L1 expression ≥ 1% were triggered on the basis of achieving 12 months minimum follow-up. Given the study outcomes at that time, the OS interim analysis (IA) is considered as the OS final analysis."

The company describes methods in relation to the "*Protection of Type I error across primary and secondary endpoints*" in Section B.2.4.3 of the CS.[2]

A summary of statistical methods are described in Sections B.2.4.4, B.2.4.5 and B.2.4.6 of the CS[2] and is presented in Table 3.4 below.

Table 3.4: Statistical methods used in the CheckMate 648 RCT

Parameter	Methods
Primary outcomes	
Population	All patients with PD-L1 ≥1%
Treatment comparison	NIVO-CHEMO versus CHEMO
Outcomes	OS assessed by BICR PFS assessed by BICR
Main analysis methods	Two-sided log rank test stratified according to ECOG PS (0 versus 1) and number of organs with metastases (≤1 versus ≥2)* HR estimates for OS and PFS with associated 100(1-alpha)% CI estimated via a stratified Cox model with treatment arm as the only covariate Median OS and PFS per treatment arm estimated using the K-M product-limit method. Median OS and PFS with 95% CI were estimated using a log-log transformed CI for the survival function.
Further analyses	Forest plots for unstratified OS and PFS HR (95% CI) estimates were generated for the following population subgroups: age category; sex; race; region; ECOG PS; weight category; disease stage at initial diagnosis; histological grade at initial diagnosis; histological classification at initial diagnosis; location at initial diagnosis; disease status at current diagnosis; smoking status; alcohol use; number of organs with metastases at baseline; time from initial disease diagnosis to randomisation; prior surgery (excluding biopsy); and prior radiotherapy. OS and PFS rates at 3, 6, 9, and 12 months estimated using K-M estimates on the OS and PFS curves for each randomised arm, with associated two-sided 95% CIs calculated using Greenwood's formula. Minimum follow-up had to be ≥ the timepoint to generate the rate.
Secondary outcomes	
Population/outcomes	OS in all randomised patients PFS by BICR in all randomised patients ORR by BICR in patients with PD-L1 ≥1% ORR by BICR in all randomised patients
Analysis methods	All outcomes analysed by treatment group as randomised. If any of the primary OS and PFS outcomes was significantly superior, the corresponding secondary outcome of OS and PFS per BICR in all randomised participants was compared using a two-sided log-rank test at the allocated significance level, stratified by ECOG PS, number of organs with metastases and PD-L1 expression (≥1% or <1% or indeterminate).

Parameter	Methods
	For each comparison, the HR with its associated two-sided 95% CI was estimated via a stratified Cox model with treatment arm as the only covariate in the model. OS and PFS for each treatment arm were estimated and plotted using the K-M product-limit method. Median OS and PFS with associated two-sided 95% CI were constructed based on a log-log transformed CI for the survival function. Subgroup analyses and analyses at different timepoints as described for the primary outcomes were carried out in all randomised patients. ORR (as assessed by BICR) in patients with PD-L1 ≥1% and in all randomised patients was to be tested only if significance level is passed on them. ORR was computed in each treatment group along with the exact 95% CI using Clopper-Pearson method. An estimate of the difference in ORRs and corresponding 95% CI were calculated using Cochran-Mantel-Haenszel (CMH) methodology and adjusted by the stratification factors. The stratified odds ratios (Mantel-Haenszel estimator) between the treatments were provided along with the 95% CI.
Safety analyses	
Analysis methods	Safety analyses were performed for all treated patients by treatment group, unless otherwise specified. AEs were coded using the Medical Dictionary for Regulatory Activities (MedDRA) Version 23.1. AEs and laboratory values were graded for severity according to the National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE) version 4.0. All on-study AEs, treatment-related AEs, SAEs, and treatment-related SAEs were tabulated using worst grade per NCI CTCAE version 4.0 criteria by System Organ Class (SOC) and Preferred Terms (PT). In the AE summary tables, unless otherwise specified, subjects were counted only once at the PT, only once at the SOC, and only once at the subject level for the counting of total number of subjects with an AE

Based on Sections B.2.4.4, B.2.4.5 and B.2.4.6 of the CS.[2]
*Whilst "*Region (East Asia versus Rest of Asia versus Rest of World)*" was used to stratify randomisation, it was excluded from all stratified analyses due to the small sample size in Rest of Asia (Section B.2.4.4 of the CS).[2]
AE(s) = adverse event(s); BICR = blinded independent central review; CHEMO = fluorouracil and cisplatin chemotherapy; CI = confidence interval; CMH = Cochran-Mantel-Haenszel; CS = company submission; CTCAE – Common Terminology Criteria for Adverse Events; ECOG PS = Eastern Cooperative Oncology Group; HR = hazard ratio; K-M = Kaplan-Meier; MedDRA = Medical Dictionary for Regulatory Activities; NCI = National Cancer Institute; NIVO-CHEMO = nivolumab combined with fluorouracil and cisplatin chemotherapy; ORR = objective response rate; OS = overall survival; PD-L1 = programmed death ligand 1; PFS = progression free survival; PS = Performance Status; PT = preferred term; RCT = randomised controlled trial; SAE(s) = serious adverse event(s); SOC = System Organ Class

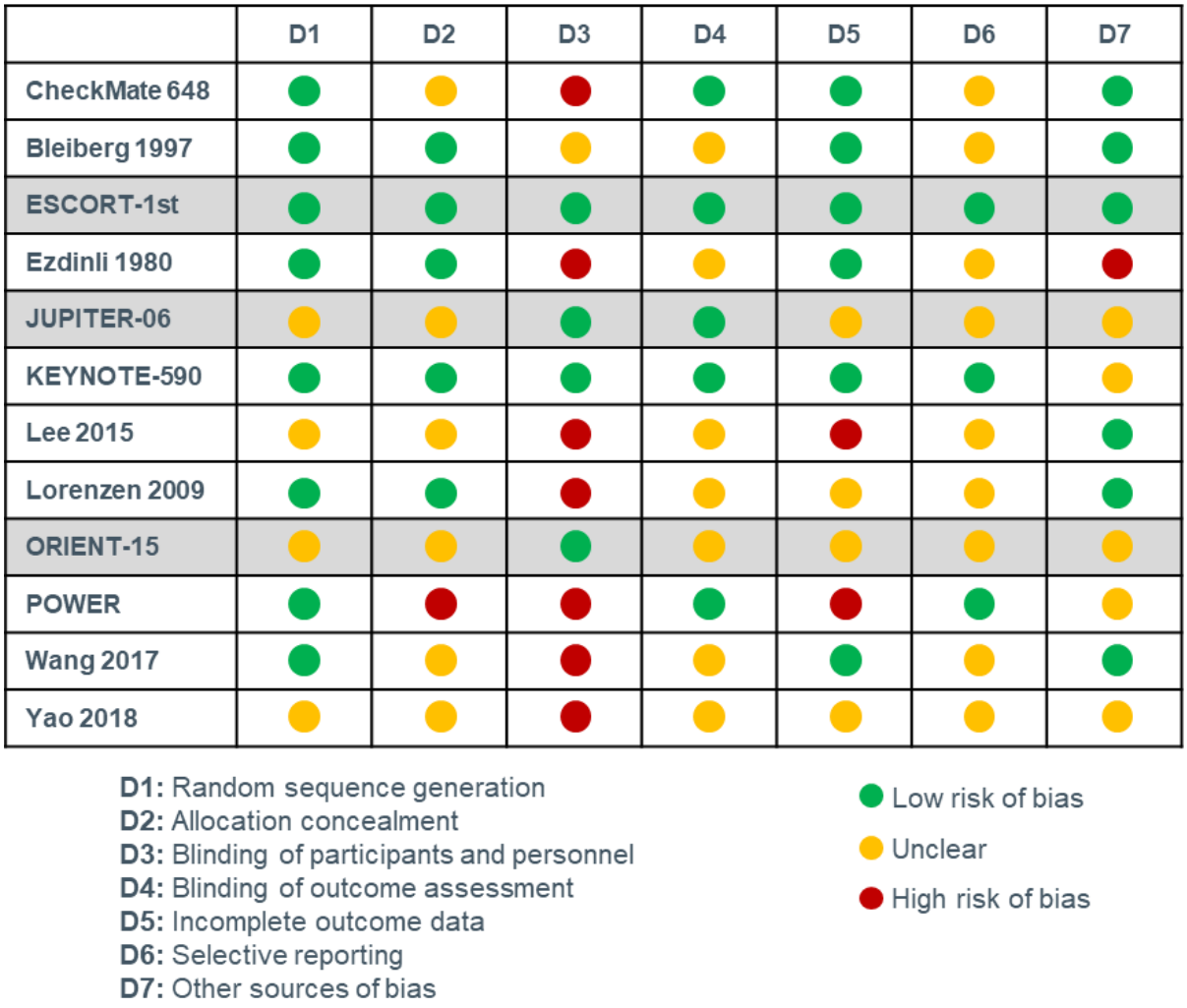
ERG comment: The statistical methods appear to be satisfactory.

3.2.4 Risk of bias in the CheckMate 648 RCT

As outlined in Section 3.1.4 (above), Appendix E of the CS[12] described the use of the Cochrane Risk of Bias Tool (original version) to assess the methodological quality of the 12 RCTs included in the SLR.[23] However, the ERG noted the presentation of two separate methodological quality assessments for the CheckMate 648 RCT in the CS. One was found in Appendix E (page 33)[12] where the quality of CheckMate 648 was presented along with the other 11 trials included in the SLR using the above-mentioned tool (Figure 3.1 below).[23] The second assessment (presented in Table 7 of Document B)[2] focused on CheckMate 648 alone using a tool adapted from the CRD's guidance (Table 3.4 below).[17]

There was no presentation of the overall risk of bias at study or outcome level for either quality assessment tool.

Figure 3.1: Cochrane risk of bias assessment of RCTs included in the SLR



Based on Figure 2 of Appendix E of the CS.[12]
Note: The assessment checklist is the Cochrane Risk of Bias Tool (original version).[23]
Trials identified from update searches are highlighted in grey.
CS = company submission; D1, 2, 3 etc = Domain 1, 2, 3 etc; RCT = randomised controlled trial; SLR = systematic literature review

Table 3.4: CRD quality assessment results for CheckMate 648

	CheckMate 648 (NCT03143153)
Was randomisation carried out appropriately?	Yes, all eligible patients were randomised in a 1:1:1 ratio using interactive response technology. Randomisation was stratified by PD-L1 status (≥1% or <1%), region (East Asia [Japan, Korea, Taiwan], rest of Asia and rest of world), ECOG PS (0 or 1), and the number of organs with metastasis (≤1 or ≥2).
Was the concealment of treatment allocation adequate?	No, the study was open label as a safety measure, so that prompt and accurate assessment of the unique toxicities associated with study treatments could be conducted.

Were the groups similar at the onset of the study in terms of prognostic factors?	Yes, the baseline characteristics of the two treatment arms were generally balanced (see Table 9 of the CS[2] or Section 3.2.6 of this report)
Were the care providers, participants and outcome assessors blind to treatment allocation?	No, the study was open label as a safety measure, so that prompt and accurate assessment of the unique toxicities associated with study treatments could be conducted
Were there any unexpected imbalances in dropouts between groups?	No, a similar number of patients discontinued in both study arms (see Table 8 of the CS[2] or Section 3.2.5 of this report)
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No, all measured outcomes have been reported
Did the analysis include an ITT analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes, an appropriate ITT analysis was conducted and the methods to account for missing data were also appropriate
Based on Table 7 of the CS[2] According to the CS,[2] the above checklist was adapted from CRD's guidance for undertaking reviews in health care.[17] CRD = Centre for Reviews and Dissemination; CS = company submission; ECOG PS = Eastern Cooperative Oncology Group Performance Status; ITT = intention to treat; PD-L1 = programmed death ligand 1	

The ERG performed its own assessment of the CheckMate 648 RCT using the Cochrane Risk of Bias Tool (original version),[23] the results of which are presented in Table 3.5 below.

Table 3.5: ERG's assessment of risk of bias in CheckMate 648

Domain	Risk of bias rating and rationale for judgement
Selection bias	
Random sequence generation	Low risk of bias Patients were randomised in a 1:1:1 ratio using a web-based interactive response technology system implemented by a third party (block size 3). Randomisation was stratified by PD-L1 status (≥1% or <1%), region (East Asia [Japan, Korea, Taiwan], rest of Asia and rest of world), ECOG PS (0 or 1), and the number of organs with metastasis (≤1 or ≥2).[7]
Allocation concealment	Low risk of bias The web registration system was implemented by a third party, ensuring that the assignment sequence was concealed until the treatment allocation was completed.[7]
Performance bias	
Blinding of participants and personnel. Assessments should be made for each main outcome (or class of outcomes).	High risk of bias The study was open label, so investigators were not blind to treatment allocation.[7] Although not explicitly stated, it is assumed that participants were also not blind to treatment allocation.
Detection bias	
Blinding of outcome assessment.	Low risk of bias for primary outcomes (OS and PFS) and other outcomes (ORR and DOR), all evaluated by BICR, and

Assessments should be made for each main outcome (or class of outcomes).	all assessed in all randomised patients and those with PD-L1 TC ≥1%.[7] High risk of bias for patient-reported outcomes (AEs and HRQoL).[7]
Attrition bias	
Incomplete outcome data. Assessments should be made for each main outcome (or class of outcomes).	Unclear risk of bias for the analysis population of all treated patients (defined as all randomised patients who received ≥1 dose of the trial treatment). The proportions of non-treated patients were 3.4% for NIVO-CHEMO and 6.2% for CHEMO (see Table 8 of the CS[2] or Section 3.2.5 of this report). The impact of this difference is unclear.
Reporting bias	
Selective reporting	Low risk of bias All outcomes described in the trial protocol have been reported. However as CheckMate 648 is ongoing, some follow-up data have not yet been reported.[8]
Other bias	
Other sources of bias. Important concerns about bias not addressed in the other domains in the tool.	Not assessed.
The ERG assessed the CheckMate 648 RCT using the Cochrane Risk of Bias Tool (original version).[23] The ERG's assessment was based on the CS,[2] the trial protocol[8] and a published paper of CheckMate 648.[7] AEs = adverse events; BICR = blinded independent central review; CHEMO = fluorouracil and cisplatin chemotherapy; CS = company submission; DOR = duration of response; ECOG PS = Eastern Cooperative Oncology Group Performance Status; ERG = Evidence Review Group; HRQoL = health-related quality of life; NIVO-CHEMO = nivolumab combined with fluorouracil and cisplatin chemotherapy; ORR = objective response rate; OS = overall survival; PD-L1 = programmed death ligand 1; PFS = progression free survival; RCT = randomised controlled trial; TC = tumour cells	

ERG comment:

The company's risk of bias assessment

As outlined above, the company presented two separate methodological assessments of CheckMate 648. The results of these did not entirely agree. In the response to the CRD checklist,[17] blinding of participants, care providers and outcome assessors were considered as a single item and lack of blinding was noted for all parties. However, in the Cochrane Risk of Bias Tool (original version)[23] two separate items were considered (blinding of participants and personnel; and blind of outcome assessors). The former was rated as being at high risk of bias whilst the latter was judged as low risk of bias. There was also disagreement in relation to reporting of outcomes. The response to the CRD checklist suggested a low risk of bias whilst this was unclear in the Cochrane risk of bias assessment.

The ERG noted that a response was entered for the domain of "*Other sources of bias*" in the Cochrane risk of bias assessment, but it was not clear exactly what was being assessed. According to the Cochrane Handbook, the items being assessed under this domain should be pre-specified in the review protocol.[23] Furthermore, the company did not provide an overall rating of risk of bias for CheckMate 648 as indicated in Appendix E of the CS.[12] Finally, the company did not state what CheckMate 648 documents were used as the basis of the assessment.

The ERG's risk of bias assessment

Several points of disagreement were apparent when comparing the company's Cochrane risk of bias assessment with that performed by the ERG.[23] Whilst both the company and the ERG judged the randomisation sequence approach as having low risk of bias, the ERG applied a similar rating to the allocation concealment approach whereas the company found this unclear. Attrition bias was rated as low risk by the company and unclear by the ERG. The risk of reporting bias was rated as unclear by the company and low by the ERG (with an acknowledgement by the latter that some follow-up data were not yet available). The company rated the risk of "*Other bias*"[23] as low whereas the ERG did not assess this because of the lack of clarity about what exactly was being evaluated. The ERG's rationale for the risk of bias judgements made are presented in Table 3.5.

In summary, the ERG rated the CheckMate 648 trial as being at high risk of performance bias however, most other aspects of the trial methods were well-conducted. Summary assessments of the risk of bias for the trial/s at study or at outcome level were not reported. The results of the risk assessment process are meant to be used in deciding the use of scientific evidence according to their potential bias.[22, 24] Although the execution of two formal quality assessments were reported, the CS documents did not indicate how the results of these processes were evaluated and used within the SLR and the ITC analyses.

3.2.5 Patient disposition for the CheckMate 648 RCT

Section B.2.6.1 of the CS[2] outlined the following details concerning patient disposition for the CheckMate 648 RCT. Of 1,358 patients enrolled, 970 were randomised to receive either nivolumab with chemotherapy (n=321), nivolumab with ipilimumab (n=325; treatment arm irrelevant for this submission) or chemotherapy alone (n=324). In the NIVO-CHEMO arm, 11 (3.4%) randomised patients were not treated, compared to 20 (6.2%) in the CHEMO arm.[36] At the database lock in [redacted], 11 (4%) patients in the NIVO-CHEMO arm were continuing treatment, compared to 0 (0%) in the CHEMO arm.[2] A summary of patient disposition is provided in Table 3.6.

Table 3.6: CheckMate 648: patient exposure and disposition ([redacted])

	NIVO-CHEMO	CHEMO
Number of patients (randomised) n	321	324
Number of treated patients n	310	304
Discontinued treatment n (%)	[redacted]	[redacted]
Disease progression (%)	[redacted]	[redacted]
AE related to treatment (%)	[redacted]	[redacted]
AE not related to treatment (%)	[redacted]	[redacted]
Patient request (%)	[redacted]	[redacted]
Other* (%)	[redacted]	[redacted]
Median duration of treatment (range), months	[redacted]	[redacted]
>3 months	[redacted]	[redacted]
>6 months	[redacted]	[redacted]
>9 months	[redacted]	[redacted]
>12 months	[redacted]	[redacted]
Based on Table 8 of the CS[2] which cited "CheckMate 648 [redacted] Summary data" as the primary source.[36]		

	NIVO-CHEMO	CHEMO
*Includes patients still on treatment and patients off treatment continuing in the follow-up period. Note: Percentages are given against the treated population. AE = adverse event; CHEMO = fluorouracil and cisplatin chemotherapy; CS = company submission; DBL = database lock; NIVO-CHEMO = nivolumab combined with fluorouracil and cisplatin chemotherapy		

ERG comment: The median duration of treatment was higher for NIVO-CHEMO compared with CHEMO: median (range) in months [redacted] versus [redacted] respectively. A higher proportion of patients was receiving treatment at each time point (>3, >6, >9 and >12 months) for NIVO-CHEMO compared with CHEMO (Table 3.7).

3.2.6 Baseline data for the CheckMate 648 RCT

The baseline data for CheckMate 648 are shown in Table 3.7.

Table 3.7: Participant characteristics in CheckMate 648 in all randomised patients

Baseline characteristic		NIVO-CHEMO	CHEMO
Cohort size		**321**	**324**
Age	Median (range), years	64 (40-90)	64 (26-81)
Sex	Male n (%)	253 (78.8)	275 (84.9)
Race, n (%)	White	85 (26.5)	84 (25.9)
	Black	1 (0.3)	6 (1.9)
	Asian	227 (71)	227 (70)
	Other	6 (1.9)	6 (1.9)
Geographic location, n (%)	Asia	225 (70)	226 (70)
	Rest of world	96 (29.9)	98 (30.2)
ECOG PS, n (%)	0	150 (46.7)	154 (47.5)
	1	171 (53.3)	170 (52.5)
Histological type, n (%)	Squamous cell carcinoma	311 (96.9)	318 (98.1)
	Adenosquamous cell carcinoma	[redacted]	[redacted]
	Other	[redacted]	[redacted]
Tumour cell PD-L1 expression, n (%)*	≥ 1 %	158 (49.2)	156 (48.4)
	< 1 %	163 (50.8)	166 (51.6)
Disease stage at initial diagnosis, n (%)	Stage I- III	[redacted]	[redacted]
	Stage IV	[redacted]	[redacted]
	Not reported	[redacted]	[redacted]
Disease status at study entry, n (%)	*De novo* metastatic	184 (57.3)	187 (57.7)
	Recurrent – distant	72 (22.4)	60 (18.5)
	Recurrent – loco-regional	21 (6.5)	25 (7.7)
	Unresectable advanced	44 (13.7)	52 (16.0)
Number of organs with metastases, n (%)	≤ 1	158 (49.2)	158 (48.8)
	≥ 2	163 (50.8)	166 (51.2)
Location at initial diagnosis, n (%)	Upper thoracic	[redacted]	[redacted]
	Middle thoracic	[redacted]	[redacted]

Baseline characteristic		NIVO-CHEMO	CHEMO
	Lower thoracic	[redacted]	[redacted]
	Gastroesophageal junction	[redacted]	[redacted]
	Not reported	[redacted]	[redacted]

Based on Table 9 of the CS.[2]
*Does not include indeterminate patients.
CHEMO = fluorouracil and cisplatin chemotherapy; CS = company submission; ECOG PS = Eastern Cooperative Oncology Group Performance Status; NIVO-CHEMO: nivolumab combined with fluorouracil and cisplatin chemotherapy; PD-L1 = programmed death ligand 1

In addition to the tabulated information, the company stated that there were similar proportions of patients aged at least 65 years and below 65 years in the groups receiving NIVO-CHEMO and CHEMO (Section B.2.6.2 of the CS).[2] The supporting data were not shown in the CS but were available from the clinical study report (CSR) (supplemental Table S.3.2.1.2). The presented details supported the company's statement, indicating that the proportions of patients aged at least 65 years and below 65 years were around [redacted] and [redacted] respectively in both treatment groups.[37]

The narrative in Section B.2.6.2 of the CS provided further details on geographical location, stating that: "*Geographically, the largest proportion of patients came from East Asia ([redacted]% in the NIVO-CHEMO arm and [redacted]% in the CHEMO arm), followed by the rest of the world (29.9% and 30.2%, respectively) and the rest of the Asia ([redacted]% and [redacted]%, respectively).*"[2]

Table 3.8 shows the breakdown of PD-L1 status (≥1% or <1% TC) per treatment group at baseline. As part of their response to clarification question A6 (b)[16] the company provided further information by means of a crosstabulation of PD-L1 by CPS status for CheckMate 648 (Table 3.8).

Table 3.8: Frequency of PD-L1 TC by CPS status in CheckMate 648

	NIVO-CHEMO		CHEMO	
	ITT	Tumour cell PD-L1 ≥1%	ITT	Tumour cell PD-L1 ≥1%
ITT	[redacted]	[redacted]	[redacted]	[redacted]
ITT with CPS score	[redacted]	[redacted]	[redacted]	[redacted]
CPS ≥5%	[redacted]	[redacted]	[redacted]	[redacted]
CPS ≥10%	[redacted]	[redacted]	[redacted]	[redacted]

Based on Table 3 of the company's CL response.[16]
CHEMO = fluorouracil and cisplatin chemotherapy; CL = clarification letter; CPS = combined positive score; ITT = intention to treat; NIVO-CHEMO = nivolumab combined with fluorouracil and cisplatin chemotherapy; PD-L1 = programmed death ligand 1; TC = tumour cells

ERG comment: The ERG noted that both treatment groups included a larger proportion of males versus females and that the predominant histological type in both groups was SCC (Table 3.7).

Regarding Table 3.7, the company described the NIVO-CHEMO and CHEMO groups as being comparable at baseline (Section B.2.6.2 of the CS) and this appeared to be the case for most variables.[2] However, the ERG noted that the overall age range was younger (26 to 81 years versus 40 to 90 years) and the proportion of males higher (84.9% versus 78.8%) among patients in the CHEMO group relative to NIVO-CHEMO.

In their response to clarification question A6 (b) the company made the following comment with regard to the degree of overlap between the two biomarker measurement approaches shown in Table 3.9: "*It is acknowledged that there is significant overlap between the PD-L1 TC ≥ 1% and PD-L1 CPS ≥10% populations. Of the* [redacted] *patients in the NIVO-CHEMO arm with tumour cell PD-L1 ≥1% and available CPS data,* [redacted] *also had PD-L1 CPS ≥10%. However,* [redacted] *patients who had PD-L1 CPS ≥10 in the ITT population did not have PD-L1 TC ≥ 1%……demonstrating that not all patients with PD-L1 TC ≥ 1% have PD-L1 CPS ≥10%.*"[2] The ERG notes non-overlap proportions of around 30% in both treatment groups and this does not seem trivial. The ERG remains unclear about the extent of matching between the two measures and the implications for the comparison with pembrolizumab with chemotherapy, as discussed in Sections 2.3, 3.3 and 3.4.

3.2.7 Efficacy results for the CheckMate 648 RCT

3.2.7.1 Overview of primary and secondary outcomes

The minimum follow-up was 20 months at the [redacted] database lock.[2] An overview of the primary outcomes (OS and PFS in patients with TC PD-L1 ≥ 1%) from CheckMate 648 is presented in Table 3.9 and the secondary outcomes (OS and PFS for all randomised patients) are shown in Table 3.10.

Table 3.9: CheckMate 648: overview of primary outcomes in patients with TC PD-L1 ≥ 1%

Endpoint		[redacted]		[redacted]	
		NIVO-CHEMO (n=158)	CHEMO (n=157)	NIVO-CHEMO (n=158)	CHEMO (n=157)
OS	Events, n (%)	[redacted]	[redacted]	98 (62)	121 (77.1)
	Median OS (95% CI), months	[redacted]	[redacted]	15.4 (11.9, 19.5)	9.1 (7.7, 10.0)
	12-month OS rate (95% CI), %	[redacted]	[redacted]	0.54 (0.37, 0.8)	N/A
	HR (99.5% CI)	[redacted]	[redacted]	0.5 (0.4, 0.71)	N/A
	Stratified 2-sided log-rank test p-value	[redacted]	[redacted]	<0.001	N/A
PFS per BICR	Events, n (%)	[redacted]	[redacted]	117 (74.1)	100 (63.7)
	HR (95% CI)	[redacted]	[redacted]	0.7 (0.5, 0.9)	N/A
	Median (95% CI)	[redacted]	[redacted]	6.9 (5.7, 8.3)	4.4 (2.9, 5.8)
	PFS rate (95% CI) at 12 months	[redacted]	[redacted]	25.41 (18.2,22.2)	10.45 (4.7,18.8)
	PFS rate (95% CI) at 18 months	[redacted]	[redacted]	-	-
	Stratified 2-sided log-rank test p-value	[redacted]	[redacted]	0.002	N/A

Endpoint	[redacted]		[redacted]	
	NIVO-CHEMO (n=158)	CHEMO (n=157)	NIVO-CHEMO (n=158)	CHEMO (n=157)
Based on Table 10 of the CS[2] that cited the following as primary sources: CheckMate 648 [redacted] Summary data;[36] and Doki et al. 2022[7] BICR = blinded independent central review; CHEMO = fluorouracil and cisplatin chemotherapy; CI = confidence internal; CS = company submission; DBL = database lock; HR = hazard ratio; N/A = not applicable; NIVO-CHEMO = nivolumab combined with fluorouracil and cisplatin chemotherapy; OS = overall survival; PD-L1 = programmed death ligand 1; PFS = progression-free survival; TC = tumour cells				

Table 3.10: CheckMate 648: overview of secondary outcomes in all randomised patients

Endpoint		[redacted]		[redacted]	
		NIVO-CHEMO (n=321)	CHEMO (n=324)	NIVO-CHEMO (n=321)	CHEMO (n=324)
OS	Events, n (%)	[redacted]	[redacted]	209 (65.1)	232 (71.6)
	Median OS (95% CI), months	[redacted]	[redacted]	13.2 (11.1, 15.7)	10.7 (9.4, 11.9)
	12-month OS rates (95% CI), %	[redacted]	[redacted]	53.53 (47.8,58.9)	44.32 (38.6,49.9)
	HR (99.5% CI)	[redacted]	[redacted]	0.7 (0.6, 1.0)	N/A
	Stratified 2-sided log-rank test p-value	[redacted]	[redacted]	0.0021	N/A
PFS per BICR	Events, n (%)	[redacted]	[redacted]	235 (73.2)	210 (64.8)
	HR (95% CI)	[redacted]	[redacted]	0.8 (0.6, 1.0)	N/A
	Median (95% CI)	[redacted]	[redacted]	5.8 (5.6, 7.0)	5.6 (4.3, 5.9)
	PFS rate (95% CI) at 12 months	[redacted]	[redacted]	23.62 (18.63, 28.95)	16.02 (11.02, 21.86)
	PFS rate (95% CI) at 18 months	[redacted]	[redacted]	-	-
	Stratified 2-sided log-rank test p-value	[redacted]	[redacted]	0.0355	N/A
Based on Table 11 of the CS[2] that cited the following as primary sources: CheckMate 648 [redacted] Summary data;[36] and Doki et al. 2022[7] BICR = blinded independent central review; CHEMO = fluorouracil and cisplatin chemotherapy; CI = confidence internal; CS = company submission; DBL = database lock; HR = hazard ratio; N/A = not applicable; NIVO-CHEMO = nivolumab combined with fluorouracil and cisplatin chemotherapy; OS = overall survival; PFS = progression-free survival.					

3.2.7.2 Overall survival

In the subgroup of patients with tumour cell PD-L1 ≥1%, treatment with NIVO-CHEMO was associated with an increase in OS compared with CHEMO at the [redacted]: median OS 15.4 months versus 9.1 months; HR 0.5 (99.5% CI 0.4 to 0.71). A similar result was observed at the 20-month minimum follow-up ([redacted]): median OS [redacted] months versus [redacted] months; HR: [redacted], Table 3.9).[2]

The median OS estimates for all randomised patients were the same at both timepoints: [redacted] months for NIVO-CHEMO and [redacted] months for CHEMO. The HR estimates suggested no-between group difference at the [redacted] (HR 0.7, 99.5% CI 0.6 to 1.0) and a difference in favour of NIVO-CHEMO at the [redacted] (HR [redacted], Table 3.10).[2]

The corresponding Kaplan-Meier (K-M) survival plots are presented in Figures 3.2 and 3.3.

Figure 3.2: OS in the NIVO-CHEMO and CHEMO arms in patients with TC PD-L1 ≥1%

[redacted]

Based on Figure 9 of the CS[2] that cited CheckMate 648 [redacted] Summary data[36] as the primary source.
Statistical model for HR and p-value: stratified Cox proportional hazard model and stratified log-rank test. Symbols represent censored observations. Stratification factors are ECOG PS (0 versus 1) and number of organs with metastases (≤1 versus ≥2) as recorded in IRT.[2]
CHEMO = fluorouracil and cisplatin chemotherapy; CI = confidence internal; CS = company submission; ECOG PS = Eastern Cooperative Oncology Group Performance Status; HR = hazard ratio; IRT = interactive response technology; NIVO-CHEMO = nivolumab combined with fluorouracil and cisplatin chemotherapy; OS = overall survival; PD-L1 = programmed death ligand 1; TC = tumour cells

Figure 3.3: OS in the NIVO-CHEMO and CHEMO arms in all randomised patients

Based on Figure 10 of the CS[2] that cited CheckMate 648 [redacted] Summary data[36] as the primary source.
Statistical model for HR and p-value: stratified Cox proportional hazard model and stratified log-rank test. Symbols represent censored observations. Stratification factors are ECOG PS (0 versus 1) and number of organs with metastases (≤1 versus ≥2) as recorded in IRT. [2]
CHEMO = fluorouracil and cisplatin chemotherapy; CI = confidence internal; CS = company submission; ECOG PS = Eastern Cooperative Oncology Group Performance Status; HR = hazard ratio; IRT = interactive response technology; NIVO-CHEMO = nivolumab combined with fluorouracil and cisplatin chemotherapy; OS = overall survival

3.2.7.3 Progression-free survival

Treatment with NIVO-CHEMO was associated with increased PFS assessed by BICR in patients with TC PD-L1 ≥1% compared with CHEMO at both timepoints: HR 0.65 (95% CI 0.46 to 0.92) at the [redacted]; and HR [redacted] at the [redacted] (Table 3.9).[2]

Neither of the HR estimates for all randomised patients suggested a between-group difference: HR 0.8 (95% CI 0.6 to 1.0) at the [redacted]; and HR [redacted] at the [redacted] (Table 3.10).[2]

The corresponding K-M survival plots are presented in Figures 3.4 and 3.5.

Figure 3.4: PFS (per BICR) in the NIVO-CHEMO and CHEMO arms in patients with TC PD-L1≥1%

Based on Figure 11 of the CS[2] that cited CheckMate 648 [redacted] Summary data[36] as the primary source.
Statistical model for HR and p-value: stratified Cox proportional hazard model and stratified log-rank test. Symbols represent censored observations. Stratification factors are ECOG PS (0 versus 1), number of organs with metastases (≤1 versus ≥2) and PD-L1 status (≥1 versus <1% or indeterminate) as recorded in IRT.[2]
BICR = blinded independent central review; CHEMO = fluorouracil and cisplatin chemotherapy; CI = confidence internal; CS = company submission; ECOG PS = Eastern Cooperative Oncology Group Performance Status; HR = hazard ratio; IRT = interactive response technology; NIVO-CHEMO = nivolumab combined with fluorouracil and cisplatin chemotherapy; PD-L1 = programmed death ligand 1; PFS = progression-free survival; TC = tumour cells

Figure 3.5: PFS (per BICR) in the NIVO-CHEMO and CHEMO arms in all randomised patients

Based on Figure 12 of the CS[2] that cited CheckMate 648 [redacted] Summary data[36] as the primary source. Statistical model for HR and p-value: stratified Cox proportional hazard model and stratified log-rank test. Symbols represent censored observations. Stratification factors are ECOG PS (0 vs 1), number of organs with metastases (≤1 vs ≥2) and PD-L1 status (≥1 vs <1% or indeterminate) as recorded in IRT.[2]
BICR = blinded independent central review; CHEMO = fluorouracil and cisplatin chemotherapy; CI = confidence internal; CS = company submission; ECOG PS = Eastern Cooperative Oncology Group Performance Status; HR = hazard ratio; IRT = interactive response technology; NIVO-CHEMO = nivolumab combined with fluorouracil and cisplatin chemotherapy; PD-L1 = programmed death ligand 1; PFS = progression-free survival

3.2.7.4 Objective response rate

In patients with PD-L1 TC ≥1% the estimates for the proportion with BICR-assessed ORR were [redacted] ([redacted]) for patients receiving NIVO-CHEMO and [redacted] ([redacted]) for CHEMO. Complete responses assessed by BICR were observed in [redacted] patients in the NIVO-CHEMO arm and [redacted] patients in the CHEMO arm.[2]

The proportion of patients with BICR-assessed ORR among all randomised participants was [redacted] ([redacted]) for NIVO-CHEMO compared with [redacted] ([redacted]) for CHEMO. Complete responses assessed by BICR were observed in [redacted] ([redacted]) patients receiving NIVO-CHEMO and [redacted] ([redacted]) on CHEMO.[2]

Table 3.11 provides an overview of the data on ORR.

Table 3.11: ORR results (per BICR) from the statistical testing hierarch

Endpoint		NIVO-CHEMO (n=158)[a]	CHEMO (n=157)[a]
Patients with tumour cell PD-L1 ≥1%	ORR, %	[redacted]	[redacted]
	95% CI	[redacted]	[redacted]
	Best overall response, %		
	Complete response	[redacted]	[redacted]
	Partial response	[redacted]	[redacted]
	Stable disease	[redacted]	[redacted]
	Progressive disease	[redacted]	[redacted]
	Not evaluable	[redacted]	[redacted]
	Median time to response[b] (range), months	[redacted]	[redacted]
All randomised patients	ORR, %	[redacted]	[redacted]
	95% CI	[redacted]	[redacted]
	Best overall response, %		
	Complete response	[redacted]	[redacted]
	Partial response	[redacted]	[redacted]
	Stable disease	[redacted]	[redacted]
	Progressive disease	[redacted]	[redacted]
	Not evaluable	[redacted]	[redacted]
	Median time to response[b] (range), months	[redacted]	[redacted]

Based on Table 12 of the CS[2] that cited CheckMate 648 [redacted] Summary data[36] as the primary source.
[a]Randomised patients who had target lesion measurements at baseline per BICR assessment
[b]Time to response was defined as the time from the start of treatment to the first objective tumour response
BICR = blinded independent central review; CHEMO = fluorouracil and cisplatin chemotherapy; CI = confidence internal; CS = company submission; NIVO-CHEMO = nivolumab combined with fluorouracil and cisplatin chemotherapy; ORR = objective response rate; PD-L1 = programmed death ligand 1.

The CS also reported results for DoR (Section B.2.6.3.3).[2] The results are not summarised here because this outcome is outside of the NICE final scope.

ERG comment: In Section B.2.13.4 of the CS, the company stated that "*CheckMate 648 provides survival data that may be considered relatively mature*".[2] In the clarification letter, the ERG asked the company how the maturity of the survival data was assessed (question A13). The company replied that: "*Survival data is typically considered mature where the median point has been reached. Per Appendix N Section 4.2.3, as 25.3% of patients remain alive at the end of follow-up, the CheckMate 648 data can be considered relatively mature.*"[16] The company then went on to draw a comparison with the KEYNOTE-062 trial. The ERG would agree that it seems reasonable to conclude that the survival data are relatively mature.

3.2.7.5 Health-related quality of life

Changes in HRQoL were assessed in patients with PD-L1 ≥1% and all randomised participants during the CheckMate 648 RCT using several measurement instruments including: the EQ-5D-3L Utility Index; the European Quality of Life-5 dimensions-3 levels (EQ-5D-3L) visual analogue scale (VAS); FACT-E; FACT-E ECS; FACT-E GP5; and FACT-G7.[2] Details of the results of each assessment are provided in the sections below.

3.2.7.5.1 EQ-5D-3L Utility Index

The EQ-5D-3L measures self-rated health state using five dimensions (mobility, self-care, usual activities, pain/discomfort and anxiety/depression) at three levels (no problems, some problems and extreme problems).[38] The CS states the minimum important difference (MID) threshold as 0.08 but does not provide a reference for this.[2]

The mean (standard deviation (SD)) baseline score among patients with PD-L1≥1% was [redacted] for those in the NIVO-CHEMO arm and [redacted] for CHEMO. The company described the outcome data as follows (Section B.2.6.3.4 of the CS):[2]

"*The mean change from baseline increased during the on-treatment period in both the NIVO-CHEMO arm and the CHEMO arm. These improvements in mean EQ-5D-3L utility index scores were sustained longer and surpassed the minimally important difference (MID) threshold more often in the NIVO-CHEMO arm compared to the CHEMO arm.*"

The mean (SD) baseline score for all randomised patients was [redacted] for those randomised to NIVO-CHEMO and [redacted] for participants on CHEMO. The company summarised the outcome data as follows (Section B.2.6.3.4 of the CS):[2]

"*Improvements in mean EQ-5D-3L utility index scores were sustained longer and surpassed the MID threshold more often in the NIVO-CHEMO arm vs the CHEMO arm. Except for Week 3, mean changes from baseline increased at all on-treatment assessments starting at Week 5 through Week 97 for the NIVO-CHEMO arm and Week 5 through Week 49 for the CHEMO arm. The NIVO-CHEMO arm was above the minimally important difference (MID) threshold (0.08) in Weeks 79, 91, and 97. The CHEMO arm was above the MID threshold at Week 49. Mean decreases from baseline were observed in the NIVO-CHEMO and CHEMO arms at follow-up visits 1 and 2.*"

Figures 3.6 and 3.7 illustrate the scores over time for patients with PD-L1 ≥1% and all randomised patients respectively.

Figure 3.6: Mean changes in EQ-5D-3L utility index score from baseline in patients with PD-L1 ≥1%

Based on Figure 13 of the CS.[2]

Error bars represent standard error of the mean. Only timepoints where data are available for ≥5 patients in each treatment group are plotted. Horizontal reference line indicates the MID considered as a change of ≥0.08 points from baseline.

CHEMO = fluorouracil and cisplatin chemotherapy; CS = company submission; EQ-5D-3L = EuroQol-5 dimensions-3 levels; FU = follow-up; MID = minimal important difference; NIVO-CHEMO = nivolumab combined with fluorouracil and cisplatin chemotherapy; NIVO-IPI = nivolumab combined with ipilimumab; PD-L1 = programmed death ligand 1; W = week

Figure 3.7: Mean changes in EQ-5D-3L utility index score from baseline in all randomised patients

Based on Figure 14 of the CS.[2]

Error bars represent standard error of the mean. Only timepoints where data are available for ≥5 patients in each treatment group are plotted. Horizontal reference line indicates the MID considered as a change of ≥0.08 points from baseline.

CHEMO = fluorouracil and cisplatin chemotherapy; CS = company submission; EQ-5D-3L = EuroQol-5 dimensions-3 levels; FU = follow-up; MID = minimal important difference; NIVO-CHEMO = nivolumab combined with fluorouracil and cisplatin chemotherapy; NIVO-IPI = nivolumab combined with ipilimumab; W = week

3.2.7.5.2 EQ-5D-3L VAS

For the VAS version of EQ-5D-3L, the range of scores is zero to 100, with higher scores reflecting more favourable self-reported health states.[38] The CS states the MID threshold as 7.0 but does not provide a reference for this.[2]

The mean (SD) baseline score for patients with PD-L1 ≥1% was [redacted] for those in the NIVO-CHEMO arm and [redacted] for CHEMO. The company summarised the outcome data included the following statements (Section B.2.6.3.4 of the CS):[2]

"*For patients with PD-L1 ≥1, improvements in mean EQ-5D-3L VAS scores were sustained longer and surpassed the MID threshold more often in the NIVO-CHEMO arm vs the CHEMO arm."*

"Increases above the MID threshold (7.0) were demonstrated at Week 79 for the NIVO-CHEMO arm. Mean decreases from baseline were observed in the NIVO-CHEMO arm at follow-up visit 2 and in the CHEMO arm at follow-up visits 1 and 2."

The mean (SD) baseline score for all randomised patients was [redacted] for those allocated NIVO-CHEMO and [redacted] for participants on CHEMO. The company's summary of the outcome data included the following observations (Section B.2.6.3.4 of the CS):[2]

"*For all randomised patients, improvements in mean EQ-5D-3L VAS scores were sustained longer and surpassed the MID threshold more often in the NIVO-CHEMO arm vs the CHEMO arm.*"

"*Increases above the MID threshold (7.0) were demonstrated at Weeks 91 and 97 for the NIVO-CHEMO arm. Mean decreases from baseline were observed in the NIVO-CHEMO and CHEMO arms at follow-up visits 1 and 2.*"

Figures 3.8 and 3.9 illustrate the scores over time for patients with PD-L1 ≥1% and all randomised patients respectively.

Figure 3.8: Mean changes in EQ-5D-3L VAS score from baseline in patients with PD-L1 ≥1%

Based on Figure 15 of the CS.[2]

Error bars represent standard error of the mean. Only timepoints where data are available for ≥10 patients in each treatment group are plotted. Horizontal reference line indicates the MID considered as a change of ≥7.0 points from baseline.

The ERG noted discrepancies between the text and figure footnotes in the CS in terms of the MID value and the minimum number of patients with data available for the EQ-5D-3L VAS assessment. For the purpose of these figure footnotes, the ERG has shown the values mentioned in the text of the CS.[2]

CHEMO = fluorouracil and cisplatin chemotherapy; CS = company submission; ERG = Evidence Review Group; EQ-5D-3L = EuroQol-5 dimensions-3 levels; FU = follow-up; MID = minimal important difference; NIVO-CHEMO = nivolumab combined with fluorouracil and cisplatin chemotherapy; NIVO-IPI = nivolumab combined with ipilimumab; PD-L1 = programmed death ligand 1; VAS = visual analogue scale; W = week

Figure 3.9: Mean changes in EQ-5D-3L VAS score from baseline in all randomised patients

Based on Figure 16 of the CS.[2]

Error bars represent standard error of the mean. Only timepoints where data are available for ≥5 patients in each treatment group are plotted. Horizontal reference line indicates the MID considered as a change of ≥0.08 points from baseline.

The ERG noted discrepancies between the text and figure footnotes in the CS in terms of the MID value and the minimum number of patients with data available for the EQ-5D-3L VAS assessment. For the purpose of these figure footnotes, the ERG has shown the values mentioned in the text of the CS.[2]

CHEMO = fluorouracil and cisplatin chemotherapy; CS = company submission; ERG = Evidence Review Group; EQ-5D-3L = EuroQol-5 dimensions-3 levels; FU = follow-up; MID = minimal important difference; NIVO-CHEMO = nivolumab combined with fluorouracil and cisplatin chemotherapy; NIVO-IPI = nivolumab combined with ipilimumab; VAS = visual analogue scale; W = week

3.2.7.5.3 FACT-E

The Functional Assessment of Cancer Therapy-Esophagus (FACT-E) measures HRQoL in patients with OC. It includes 44 items covering four subscale domains relating to general well-being (physical well-being, social/family well-being, emotional well-being and functional well-being) and an OC subscale domain. Higher scores represent more favourable HRQoL.[39, 40] The company described the MID threshold as 9.1 but did not provide a reference for this. The company presented results where data were available for ≥10 patients per treatment arm.[2]

In patients with PD-L1 ≥1%, the mean (SD) baseline scores were [redacted] for those on NIVO-CHEMO and [redacted] for CHEMO. The company described the outcome data as follows (Section B.2.6.3.4 of the CS):[2]

"*Except for Week 3, mean changes from baseline increased at all other on-treatment assessments (with ≥ 10 patients) from Week 5 through Week 85 for the NIVO-CHEMO arm and from Week 5 through Week 37 for the CHEMO arm.*"

"*The NIVO-CHEMO arm demonstrated increases above the MID threshold (9.1) from Weeks 31 through 85. Mean decreases from baseline were observed in the NIVO-CHEMO and CHEMO arms at follow-up visits 1 and 2.*"

The mean (SD) baseline scores among all randomised patients were [redacted] for those in the NIVO-CHEMO arm and [redacted] for CHEMO. In the CS, the outcome data were described as follows (Section B.2.6.3.4 of the CS):[2]

"*Except for Week 3, mean changes from baseline increased at all other on-treatment assessments (with ≥ 10 patients) from Week 5 through Week 97 for the NIVO-CHEMO arm and from Week 5 through Week 49 for the CHEMO arm.*"

"*The NIVO-CHEMO arm demonstrated increases above the MID threshold (9.1) at Weeks 43 through 97. Mean decreases from baseline were observed in the NIVO-CHEMO and CHEMO arms at follow-up visits 1 and 2.*"

Illustrative figures were not provided in the CS.[2] A supplemental figure showing the change in FACT-E scores from baseline to week 49 in all randomised patients is presented in a published journal paper of the CheckMate 648 trial.[7]

3.2.7.5.4 FACT-E ECS

The CS described the FACT-E ECS as a disease-specific, 17-item instrument that assesses concerns related to swallowing, vocalisation, breathing, dry mouth, eating, disrupted sleep due to coughing, stomach pain, and weight loss (Section B.2.6.3.4 of the CS). This would appear to be the disease-specific subscale of FACT-E (described above) however, this is not clear from the information provided in the CS.[2]

The mean (SD) baseline scores in patients with PD-L1 ≥1% were [redacted] and [redacted] in the NIVO-CHEMO and CHEMO arms respectively. The company summarised the outcome data with a series of statements as follows (Section B.2.6.3.4 of the CS):[2]

"*Mean changes from baseline for the NIVO-CHEMO arm increased at all on-treatment assessments (with ≥ 10 patients) through Week 85 with increases greater than the MID threshold (4.0) at Weeks 13 and 25 through 85.*"

"*For the CHEMO arm, mean changes from baseline increased at all on-treatment assessments (with ≥ 10 patients) through Week 37 with increases greater than the MID threshold (4.0) at Weeks 13 through Week 37.*"

"*At follow up visits 1 and 2, increases in mean changes from baseline were observed in the NIVO-CHEMO arm at both visits whereas the CHEMO arm showed a decrease at both follow-up visits.*"

"*During survival follow-up visits, mean changes from baseline for the NIVO-CHEMO arm were increased through visit 4 (with ≥ 10 patients). There was an increase greater than the MID threshold (4.0) at follow-up visit 4. Mean changes from baseline for the CHEMO arm were increased during the survival follow-up through follow-up visit 3 (with ≥ 10 patients). Increases greater than the MID threshold (4.0) were seen at survival follow-up visits 1 and 2.*"

The mean (SD) baseline FACT-E ECS scores in all randomised patients were [redacted] for those in the NIVO-CHEMO arm and [redacted] for CHEMO. The company's description of the outcome data was as follows (Section B.2.6.3.4 of the CS):[2]

"*Mean changes from baseline for the NIVO-CHEMO arm increased at all on-treatment assessments (with ≥ 10 subjects) through Week 97, a change greater than the MID (4.0) threshold at Weeks 13 through 97.*"

"*For the CHEMO arm, mean changes from baseline increased at all on-treatment assessments (with ≥ 10 patients) through Week 49, with a change greater than the MID (4.0) threshold at Weeks 25 through 49.*"

"*At follow-up visit 1 and 2, increases in mean changes from baseline were observed in the NIVO-CHEMO group, whereas the CHEMO arm showed a decrease at both follow-up visits.*"

"*Mean changes from baseline for the NIVO-CHEMO arm were increased during the survival follow-up through follow-up visit 5 (with ≥ 10 patients). At survival follow-up visit 4, the increase was greater than the MID (4.0). Mean changes from baseline for the CHEMO arm were increased during the survival follow-up through follow-up visit 6 (with ≥ 10 patients). Increases of greater than 4.0 were seen at survival follow-up visits 2 through 6.*"

No illustrative figures were presented in the CS[2] for the change in FACT-E ECS scores and none were available from the main publication for CheckMate 648.[7]

3.2.7.5.5 FACT-E GP5

The CS describes the FACT-E GP5 as a 5-item patient-reported outcome that measures "*the overall bother associated with the side effects of treatment.*" Information about the baseline and outcome data for patients with PD-L1 ≥1% and all randomised patients was outlined as follows (Section B.2.6.3.4 of the CS):[2]

For patients with PD-L1 ≥1 %:

"*In patients with PD-L1 ≥1, at baseline, patients in the NIVO-CHEMO arm selected "not at all"* [redacted] *of the time and "a little bit"* [redacted] *for a total of* [redacted] *patients identifying as bothered "only a little" or "not at all" by treatment side effects.[41] Except for Week 43, the combined score remained above* [redacted] *during the on-treatment period (with ≥ 10 patients) and went above* [redacted] *multiple times through Week 97.*"

"*Patients in the CHEMO arm had better baseline scores with [redacted] selecting "not at all" and [redacted] "a little bit" (Total = [redacted]). However, the combined score was never above [redacted] during the on-treatment period (with ≥ 10 patients) through Week 49 and dropped under [redacted] at Week 37.*"

For all randomised patients:

"*In all randomised patients, at baseline, patients in the NIVO-CHEMO arm selected "not at all" [redacted] of the time and "a little bit" [redacted] for a total of [redacted] patients identifying as bothered "only a little" or "not at all" by treatment side effects.[41] The combined score remained above [redacted] during the on-treatment period (with ≥ 10 patients) and went above [redacted] multiple times through Week 97.*"

"*At baseline, subjects in the CHEMO arm selected "not at all" [redacted] and [redacted] "a little bit" (combined total = [redacted]). However, the combined total score was never above [redacted] during the on-treatment period (with ≥ 10 patients) through Week 49 and was under [redacted] at multiple time points.*"

Illustrative figures were not provided in the CS.[2] A supplemental figure showing the change in FACT-E GP5 scores from baseline to week 49 in all randomised patients is presented in a published journal paper of the CheckMate 648 trial.[7]

3.2.7.5.6 FACT-G7

The company described Functional Assessment of Cancer Therapy - General (FACT-G) as a 27-item instrument assessing generic, cancer-related symptoms and treatment-related effects covering four well-being domains (physical, social/family, emotional and functional). The FACT-General 7-items (FACT-G7) is an abbreviated form of the FACT-G consisting of seven items which is designed to provide a rapid assessment of general HRQoL in patients with cancer (Section B.2.6.3.4 of the CS). Information about the baseline and outcome data for patients with PD-L1 ≥1% and all randomised patients was outlined as follows (Section B.2.6.3.4 of the CS):[2]

For patients with PD-L1 ≥1 %:

"*In patients with PD-L1 ≥1, at baseline, mean (SD) FACT-G7 scores were similar between the NIVO-CHEMO arm ([redacted]) and CHEMO arm ([redacted]).[41] Except for Week 3, mean changes from baseline increased at all other on-treatment assessments (with ≥10 patients) from Week 5 through Week 85 for the NIVO-CHEMO arm and from Week 5 through Week 37 for the CHEMO arm. Mean decreases from baseline were observed in the NIVO-CHEMO and CHEMO arms at follow-up visits 1 and 2. Except for follow-up visits 1 and 2 in the CHEMO arm, mean change from baseline decreased at all other survival follow-up visits (with ≥10 patients) for both the NIVO-CHEMO and CHEMO arms.*"

For all randomised patients:

"*In all randomised patients, at baseline, mean (SD) FACT-G7 scores for the NIVO-CHEMO arm ([redacted]) were similar to those in the CHEMO arm ([redacted]).[41] Except for Week 3, mean changes from baseline increased at all other on-treatment assessments (with ≥10 subjects) from Week 5 through Week 97 for the NIVO-CHEMO arm and from Week 5 through Week 49 for the CHEMO arm. Mean decreases from baseline were observed in the NIVO-CHEMO and CHEMO arms at follow-up visits 1 and 2. Except for follow-up visit 5 in the CHEMO arm, mean change from baseline decreased at all other survival follow-up visits (with ≥10 patients) for both the NIVO-CHEMO and CHEMO arms.*"

No illustrative figures were presented in the CS[2] for the change in FACT-G7 scores and none were available from the main publication for CheckMate 648.[7]

ERG comment on HRQoL data:

- Going by the information in the CS, baseline scores appeared similar between treatment arms for both patient populations (those with PD-L1 ≥1% and all randomised patients) for measures derived from all reported HRQoL instruments.[2]
- The presentation of outcome data constituted a descriptive narrative summary (sometimes with accompanying figures), focusing on within-group (rather than between-group) differences. For some assessments, there was a focus on whether the MID had been surpassed. Although the MID values were provided for each measurement instrument, supporting references for this information were lacking.[2]
- According to the CS, both treatment arms surpassed the MID during follow-up in both versions of the EQ-5D-3L (Utility Index and VAS) in both patient populations. Results observed at some specific timepoints in the text were not easily discernible from the accompanying figures.[2]
- For the FACT-E assessment, the CS outlined that increases above the MID were seen in the NIVO-CHEMO group only, in both patient populations. However, for FACT-E ECS, increases above the MID were observed in both treatment arms and in both patient populations.[2]
- The MID was not mentioned in the description of outcomes for the FACT-E-GP5 and FACT-G7 assessments in the CS.[2]
- The overall account of HRQoL in the CS[2] lacked detail. Whilst much more detail was available from the supplemental file for the CheckMate 648 CSR,[41] this consisted of tabulation of differences between baseline and different follow-up points within treatment arms for the two patient populations. The within-group comparison statistics included mean score with associated 95% CI and median score with interquartile range (IQR) and range. No p-values or estimates of effect were provided for between-group differences.

3.2.8 Subgroup analyses in CheckMate 648

Subgroup analyses was reported for the outcome of OS alone in the following subgroups:

- Age
- Gender
- Race
- Geographic region per CRF
- Geographic region (Asian/non-Asian)
- Eastern Cooperative Oncology Group Performance Scale (ECOG PS)
- Weight
- Disease Stage
- Histologic grade at initial diagnosis
- Histologic classification
- Location of tumour
- Disease Status
- Smoking status
- Alcohol use
- Number of organs with metastases.
- Time of initial disease diagnosis
- Prior surgery/radiotherapy/systemic therapy
- PD-L1 status

The results are illustrated in Figures 3.10 to 3.13. The company stated that "*Overall, subgroup analyses of OS favoured NIVO-CHEMO over CHEMO (point estimate of HR <1) for all randomised patients*" (p. 62)[2]. Nevertheless, in the following several subgroups the HR were indeed >1: >75 years old, female, Stage I and II as shown in Figure 3.10; recurrent-loco-regional disease status as shown in Figure 3.11; 3 to <5 years from initial disease diagnosis to randomisation and prior radiotherapy as shown in Figure 3.12; and PD-L1 TC status <1% as shown in Figure 3.13. In addition, in as many as 34 subgroups the differences between nivolumab and chemotherapy versus chemotherapy were not significant (Figures 3.10 to 3.13). The degree of overlap between CIs varies across subgroups. For example, between the four age subgroups the overlap is not good, especially for the >75 age group, the same issue is observed between the four disease stages at initial diagnosis (Figure 3.10), the four different locations of the disease at initial diagnosis (Figure 3.11), the time from initial disease diagnosis (Figure 3.12) and PD-L1 TC expression between ≥1% and <1%.

Figure 3.10: Forest plot of subgroup analysis, for age, gender, race, region, ECOG status, weight and disease stage at initial diagnosis, on OS for all randomised patients treated with NIVO-CHEMO and CHEMO

Based on Figure 17 of Document B of the CS[2]

CHEMO = chemotherapy; CI = confidence interval; CS = company submission; NIVO = nivolumab; ECOG PS = Eastern Cooperative Oncology Group Performance Scale; OS = overall survival

Figure 3.11: Forest plot for subgroup analysis, for histologic grade, histologic classification, location, disease status, smoking status, alcohol use, number of organs with metastasis, on OS in all randomised patients treated with NIVO-CHEMO or CHEMO

Based on Figure 18 of Document B of the CS[2]

CHEMO = chemotherapy; CI = confidence interval; CS = company submission; HR = hazard ratio; NIVO = nivolumab; OS = overall survival

Note: HR is not computed for subset category with less than 10 subjects per treatment group

Figure 3.12: Forest plot of subgroup analysis, for time from initial diagnosis to randomisation, prior surgery, prior radiotherapy and prior systemic therapy, on overall survival in all randomised patients treated with NIVO-CHEMO or CHEMO

Based on Figure 19 of Document B of the CS[2]
CHEMO = chemotherapy; CI = confidence interval; CS = company submission; NIVO = nivolumab; OS = overall survival

Figure 3.13: Forest plot of treatment effect on OS by tumour cell PD-L1 TC cut-offs – all randomised patients treated with NIVO-CHEMO or CHEMO

Based on Figure 20 of Document B of the CS[2]

Note: HR is not computed for subset category with less than 10 subjects per treatment group

CHEMO = chemotherapy; CI = confidence interval; CS = company submission; HR = hazard ratio; NIVO = nivolumab; OS = overall survival; PD-L1 = programmed death ligand 1; TC = tumour cell

ERG comment: The company stated that subgroup analyses was executed for both OS and PFS outcomes, but only the OS results are presented. The additional file, provided by the company, containing the results for the [redacted] data cut-off, also included results for only OS subgroup analysis [36].

There is no comment from the company concerning the overwhelmingly high proportion of subgroups that show not-statistically significant differences between the two arms nor the low degree of overlap between the HR CIs of certain subgroups, as mentioned above. The lack of overlap would indicate that the effectiveness of the intervention varies notably between subgroups. The company has chosen to provide the subgroup analysis for the entire randomised population in CheckMate 648 (n=645) which is beyond the scope of this STA. No subgroup analysis was provided within the population expressing PD-L1 TC ≥1%. The significance of this omission is highlighted by the results of the comparison between the PD-L1 TC ≥1% (n=314) and <1% (n=139) populations, where the HRs results are 0.60 (95% CI, 0.47; 0.77) and 1.01 (95% CI, 0.78; 1.30) respectively, and the overlap of the CI is almost non-existent (Figure 3.13). The subgroup analysis presented herein has limited relevance to the scope of the submission.

The results of the subgroup analysis regarding race are discussed in Section 3.2.10 of this report.

3.2.9 Adverse events in the CheckMate 648 RCT

The safety data for nivolumab with chemotherapy come from all the randomised patients in CheckMate 648. Grading of the severity was done according to the National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE) Version 4.0 and AE coding via the Medical Dictionary for Regulatory Activities (MedDRA) Version 23.1. The study focused on AEs of special clinical interest specifically related to the use of nivolumab, grouped in six select AE categories: endocrine, gastrointestinal (GI), hepatic, pulmonary, renal, and skin[37]. Analysis focusing on immune-mediated adverse events (IMAEs) was also conducted including diarrhoea/colitis, hepatitis, pneumonitis, nephritis and renal dysfunction, rash, hypersensitivity/infusion reactions, and endocrine events[37]. Other events of special interest (OESI) under examination included MedDRA preferred terms of myositis/rhabdomyolysis, myocarditis, demyelination, Guillain-Barre syndrome, pancreatitis, uveitis, encephalitis, myasthenic syndrome, and graft versus host disease[37].

The company stated that the treatment was generally well-tolerated with similar rates of AEs and treatment-related AEs across treatment arms. New safety concerns were not raised as the AEs were similar to the ones experienced in populations for alternative nivolumab indications[42-45]. AEs are reported in two categories as overall AE and those with potential immunologic aetiology.

3.2.9.1 Overall adverse events

A summary of the overall AEs is presented in Table 3.12 for the [redacted] DBL.[36] Higher rates of all-causality SAEs were reported in the two arms: [redacted] in NIVO-CHEMO and [redacted] in CHEMO, with treatment-related SAEs of any grade reported for [redacted] and [redacted] of the patients, respectively. In terms of treatment related, any grade AEs, very high rates of [redacted] and [redacted] were reported in both arms, with high rates of Grade 3-4 AEs [redacted] and [redacted] in NIVO-CHEMO and CHEMO, respectively.

Treatment-related AEs leading to discontinuation of treatment were reported in [redacted] (any Grade) and [redacted] (Grade 3-4) of patients in the NIVO-CHEMO arm and [redacted] (any Grade) and [redacted] (Grade 3-4) in the CHEMO arm. The death rates were very high in both arms ([redacted] and [redacted]) but the majority were attributed to disease progression ([redacted] and [redacted]). [redacted] deaths ([redacted])were attributed to study drug toxicity in each arm and [redacted] were considered to be related to nivolumab per investigator.

The company did not provide separate evidence for patients expressing PD-L1 ≥1% (TC) for the [redacted][36]. Table 3.13 presents the results for the PD-L1 ≥1% (TC) population for the earlier data cut-off according to the primary study report.[37, 41]

Table 3.12: Overall AEs in CheckMate 648

Safety parameter	NIVO-CHEMO (n=310)	CHEMO (n=304)
Deaths, n (%)	[redacted]	[redacted]
Primary reason for death		
Disease	[redacted]	[redacted]
Study drug toxicity	[redacted]	[redacted]
Unknown	[redacted]	[redacted]
Other	[redacted]	[redacted]
All-causality AEs		
Any Grade	[redacted]	[redacted]
Grade 3-4	[redacted]	[redacted]
All-causality SAEs		
Any Grade	[redacted]	[redacted]
Grade 3-4	[redacted]	[redacted]
All-causality AEs leading to discontinuation		
Any Grade	[redacted]	[redacted]
Grade 3-4	[redacted]	[redacted]
Treatment-related AEs		
Any Grade	[redacted]	[redacted]
Grade 3-4	[redacted]	[redacted]
Treatment-related SAEs		
Any Grade	[redacted]	[redacted]
Grade 3-4	[redacted]	[redacted]
Treatment-related AEs leading to discontinuation		
Any Grade	[redacted]	[redacted]
Grade 3-4	[redacted]	[redacted]

Based on Table 18 of Document B of the CS[2], [redacted][36].
AE = adverse event; CHEMO = chemotherapy; CS = company submission; NIVO = nivolumab

Table 3.13: Summary of Safety - All Treated Subjects with Tumour Cell PD-L1 ≥ 1% in CheckMate 648

Safety Parameter	NIVO + CHEMO (N=155) (N, %)		CHEMO (N=145) (N, %)	
Deaths	[redacted]		[redacted]	
Primary Reason for Death				
Disease	[redacted]		[redacted]	
Study Drug Toxicity	[redacted]		[redacted]	
Unknown	[redacted]		[redacted]	
Other	[redacted]		[redacted]	
	Any Grade	Grade 3-4	Any Grade	Grade 3-4
All-causality SAEs	[redacted]	[redacted]	[redacted]	[redacted]
Drug-related SAEs	[redacted]	[redacted]	[redacted]	[redacted]
All-causality AEs leading to DC	[redacted]	[redacted]	[redacted]	[redacted]
Drug-Related AEs leading to DC	[redacted]	[redacted]	[redacted]	[redacted]
All-causality AE	[redacted]	[redacted]	[redacted]	[redacted]
Drug related AEs	[redacted]	[redacted]	[redacted]	[redacted]
≥15% of Subjects in any Treatment Arm				
Rash	[redacted]	[redacted]	[redacted]	[redacted]
Pruritus	[redacted]	[redacted]	[redacted]	[redacted]
Diarrhoea	[redacted]	[redacted]	[redacted]	[redacted]
Nausea	[redacted]	[redacted]	[redacted]	[redacted]
Stomatitis	[redacted]	[redacted]	[redacted]	[redacted]
Vomiting	[redacted]	[redacted]	[redacted]	[redacted]
Constipation	[redacted]	[redacted]	[redacted]	[redacted]
Neutrophil count decreased	[redacted]	[redacted]	[redacted]	[redacted]
Fatigue	[redacted]	[redacted]	[redacted]	[redacted]
Malaise	[redacted]	[redacted]	[redacted]	[redacted]
Decreased appetite	[redacted]	[redacted]	[redacted]	[redacted]
Hiccups	[redacted]	[redacted]	[redacted]	[redacted]
Anaemia	[redacted]	[redacted]	[redacted]	[redacted]
All-causality Select AEs by Category				
Gastrointestinal	[redacted]	[redacted]	[redacted]	[redacted]
Hepatic	[redacted]	[redacted]	[redacted]	[redacted]
Pulmonary	[redacted]	[redacted]	[redacted]	[redacted]
Renal	[redacted]	[redacted]	[redacted]	[redacted]
Skin	[redacted]	[redacted]	[redacted]	[redacted]
Hypersensitivity/Infusion Reactions	[redacted]	[redacted]	[redacted]	[redacted]
Drug-Related Select AEs by Category				
Gastrointestinal	[redacted]	[redacted]	[redacted]	[redacted]
Hepatic	[redacted]	[redacted]	[redacted]	[redacted]

Safety Parameter	NIVO + CHEMO (N=155) (N, %)		CHEMO (N=145) (N, %)	
Pulmonary	[redacted]	[redacted]	[redacted]	[redacted]
Renal	[redacted]	[redacted]	[redacted]	[redacted]
Skin	[redacted]	[redacted]	[redacted]	[redacted]
Hypersensitivity/Infusion Reactions	[redacted]	[redacted]	[redacted]	[redacted]
All-causality IMAEs within 100 days of last dose treated with IMM by Category				
Diarrhea/Colitis	[redacted]	[redacted]	[redacted]	[redacted]
Hepatitis	[redacted]	[redacted]	[redacted]	[redacted]
Pneumonitis	[redacted]	[redacted]	[redacted]	[redacted]
Nephritis/Renal Dysfunction	[redacted]	[redacted]	[redacted]	[redacted]
Rash	[redacted]	[redacted]	[redacted]	[redacted]
Hypersensitivity/Infusion Reactions	[redacted]	[redacted]	[redacted]	[redacted]
All-causality Endocrine IMAEs within 100 days of last dose by Category				
Adrenal Insufficiency	[redacted]	[redacted]	[redacted]	[redacted]
Hypophysitis	[redacted]	[redacted]	[redacted]	[redacted]
Hypothyroidism/Thyroiditis	[redacted]	[redacted]	[redacted]	[redacted]
Diabetes Mellitus	[redacted]	[redacted]	[redacted]	[redacted]
Hyperthyroidism	[redacted]	[redacted]	[redacted]	[redacted]
All-causality OESIs within 100 days of last dose with/without IMM by Category				
Pancreatitis	[redacted]	[redacted]	[redacted]	[redacted]
Encephalitis	[redacted]	[redacted]	[redacted]	[redacted]
Myositis/Rhabdomyolysis	[redacted]	[redacted]	[redacted]	[redacted]
Myasthenic Syndrome	[redacted]	[redacted]	[redacted]	[redacted]
Demyelination	[redacted]	[redacted]	[redacted]	[redacted]
Guillain-Barre Syndrome	[redacted]	[redacted]	[redacted]	[redacted]
Uveitis	[redacted]	[redacted]	[redacted]	[redacted]
Myocarditis	[redacted]	[redacted]	[redacted]	[redacted]
Graft Versus Host Disease	[redacted]	[redacted]	[redacted]	[redacted]

Based on Table 8.7.2-1 of the CSR supplemental tables provided in the CS[37], [redacted] data cut-off
AE = adverse event; CHEMO = chemotherapy; CS = company submission; CSR = clinical study report; IMAEs = immune-mediated adverse events; NIVO = nivolumab; OESIs = other events of special interest; PD-L1 = programmed death ligand 1; SAEs = serious adverse events

3.2.9.2 Adverse events with potential immunologic aetiology and other events of special interest

The CS provides additional results for potentially IMAEs and OESI. The most commonly experienced AEs with potential immunologic aetiology in any Grade and all causality in the NIVO-CHEMO arm were: GI, skin and renal ([redacted], [redacted] and [redacted], respectively) and similar in the CHEMO arm: renal, GI and skin ([redacted], [redacted] and [redacted], respectively) (Table 3.14). In terms of treatment-related AEs (with potential immunologic aetiology in any Grade) the most common were renal events at [redacted] and [redacted] for NIVO-CHEMO and CHEMO, respectively (Table 3.15). The most commonly experienced, treatment-related AEs leading to treatment discontinuations (with potential immunologic

aetiology in any Grade) were also renal events at [redacted] and [redacted], for the NIVO-CHEMO and CHEMO arms, respectively (Table 3.16).

Regarding OESI, as defined by the company[37], [redacted] cases (<[redacted]) of uveitis and [redacted] cases (<[redacted]) of myositis/rhabdomyolysis were reported for the patients in the NIVO-CHEMO arm (Table 3.13). No OESI were reported for the patients in the CHEMO arm.

Table 3.14: AEs with potential immunologic aetiology: all causality in CheckMate 648

Safety parameter	NIVO-CHEMO (n=310)		CHEMO (n=304)	
	Any Grade	Grade 3-4	Any Grade	Grade 3-4
Endocrine	[redacted]	[redacted]	[redacted]	[redacted]
Gastrointestinal	[redacted]	[redacted]	[redacted]	[redacted]
Hepatic	[redacted]	[redacted]	[redacted]	[redacted]
Pulmonary	[redacted]	[redacted]	[redacted]	[redacted]
Renal	[redacted]	[redacted]	[redacted]	[redacted]
Skin	[redacted]	[redacted]	[redacted]	[redacted]

Based on Table 19 of Document B of the CS[2], [redacted] data cut-off[36].
AEs = adverse events; CHEMO = chemotherapy; CS = company submission; NIVO = nivolumab

Table 3.15: Treatment-related AEs with potential immunologic aetiology in CheckMate 648

Safety parameter	NIVO-CHEMO (n=310)		CHEMO (n=304)	
	Any Grade	Grade 3-4	Any Grade	Grade 3-4
Endocrine	[redacted]	[redacted]	[redacted]	[redacted]
Gastrointestinal	[redacted]	[redacted]	[redacted]	[redacted]
Hepatic	[redacted]	[redacted]	[redacted]	[redacted]
Pulmonary	[redacted]	[redacted]	[redacted]	[redacted]
Renal	[redacted]	[redacted]	[redacted]	[redacted]
Skin	[redacted]	[redacted]	[redacted]	[redacted]

Based on Table 20 of Document B of the CS[2], [redacted] data cut-off[36].
AEs = adverse events; CHEMO = chemotherapy; CS = company submission; NIVO = nivolumab

Table 3.16: Treatment-related AEs with potential immunologic aetiology leading to discontinuation in CheckMate 648

Safety parameter	NIVO-CHEMO (n=310)		CHEMO (n=304)	
	Any Grade	Grade 3-4	Any Grade	Grade 3-4
Endocrine	[redacted]	[redacted]	[redacted]	[redacted]
Gastrointestinal	[redacted]	[redacted]	[redacted]	[redacted]
Hepatic	[redacted]	[redacted]	[redacted]	[redacted]
Pulmonary	[redacted]	[redacted]	[redacted]	[redacted]
Renal	[redacted]	[redacted]	[redacted]	[redacted]

Safety parameter	NIVO-CHEMO (n=310)		CHEMO (n=304)	
	Any Grade	Grade 3-4	Any Grade	Grade 3-4
Skin	[redacted]	[redacted]	[redacted]	[redacted]
Based on Table 21 of Document B of the CS[2], [redacted] data cut-off[36] AEs = adverse events; CHEMO = chemotherapy; CS = company submission; NIVO = nivolumab				

Table 3.17: OESI summary in CheckMate 648

Safety parameter	NIVO-CHEMO (n=310)		CHEMO (n=304)	
Myasthenic syndrome	[redacted]	[redacted]	[redacted]	[redacted]
Demyelination event	[redacted]	[redacted]	[redacted]	[redacted]
Guillain-Barre syndrome	[redacted]	[redacted]	[redacted]	[redacted]
Pancreatitis event	[redacted]	[redacted]	[redacted]	[redacted]
Uveitis event	[redacted]	[redacted]	[redacted]	[redacted]
Encephalitis event	[redacted]	[redacted]	[redacted]	[redacted]
Myocarditis event	[redacted]	[redacted]	[redacted]	[redacted]
Myositis/rhabdomyolysis event	[redacted]	[redacted]	[redacted]	[redacted]
Graft versus host disease	[redacted]	[redacted]	[redacted]	[redacted]
Based on Table 22 of Document B of the CS[2], [redacted] data cut-off[36] CHEMO = chemotherapy; CS = company submission; NIVO = nivolumab; OESI = other events of special interest				

ERG comment: It is not clear why the company chose to present the AEs in the CS[2] based on the entire population randomised in CheckMate 648 and not the population defined in the scope of this submission. For example, in the CSR[37] and the supplemental tables provided in the CS[41], there are separate results/tables for patients with PD-L1 ≥1% (TC) expression.

The primary CSR of CheckMate 648[37] has an earlier data cut-off date, corresponding to the supplementary tables provided by the company[41] with an [redacted] 'data stamp' (January 2021 data cut-off date), but the results presented in Table 18 of Document B of the CS are from a later data cut-off date of October 2021[36]. Results for the sub-population of interest i.e. PD-L1 ≥1% (TC) have not been reported for the October 2021 data cut-off.[36]

Tables 21 and 22 of Document B of the CS[2] cite evidence from the CheckMate 648 October 2021 Summary data[36]. Nevertheless, these data are not contained in the file provided by the company and therefore could not be verified by the ERG.

3.2.10 Included studies: supporting evidence

Section B.2.13.4 of the CS[2] cites eight references[25, 46-52] intended to support the application of results of the CheckMate 648 RCT[6, 7] to patients seen in routine clinical practice in the UK. The ERG noted that one study was identified as potentially relevant within the CS clinical effectiveness SLR.[25] However, the other seven references[46-52] were not identified through the search strategy used for the CS clinical effectiveness SLR.[12] As part of the company's response to the clarification letter (questions A11

and A12), it emerged that three of these studies[49, 50, 52] had been identified through "*a targeted search*".[16] The approach used for identifying the other four reference is unclear.[46-48, 51]

Initially, the CS[2] compared baseline data (specifically age and ECOG PS) from CheckMate 648[6, 7] with two retrospective cohort studies, both published as conference abstracts (and therefore only providing limited information).[46, 47] One study was UK-based[46] whilst the other included locations across Asia and the West (the United States of America (USA), Canada and Europe).[47] The company summarised the information about this comparison in Section B.2.13.4.1 and Table 23 of the CS.[2] The ERG noted some errors in the tabulation of both comparator studies,[46, 47] whereby pieces of information from different reported populations had been incorrectly conflated.[2] The ERG has provided a corrected version of the pertinent information in Table 3.18 and provides more detail on the observed errors in the ERG comment.

Table 3.18: A comparisons of the baseline characteristics of patients in the CheckMate 648 RCT with those in the Shyamalee et al. 2021 and Jaffe et al. 2022 studies

Study		CheckMate 648[6, 7]		Shyamalee et al. (2021)[46]		Jaffe et al. (2022)[47]
Treatment arm		NIVO-CHEMO	CHEMO	Palliative chemotherapy	BSC alone	1L systemic treatment[a]
Population		Unresectable advanced, recurrent or metastatic, untreated OSCC		Advanced OSCC		Advanced OSCC
Location		International (including UK)		UK		International (including UK)[b]
Cohort size		321	324	22	94	1,049
Age (years)	Median (range)	64 (40 to 90)	64 (26 to 81)	63 (38 to 83)	79 (39 to 95)	-
	Mean (SD)	-	-	-	-	62.9 (10.6)
Sex	Male (%)	78.8	84.9	41	48	82.7
ECOG PS, n (%)	0	150 (46.7)	154 (47.5)	6 (27)	0 (0)	-
	1	171 (53.3)	170 (52.5)	9 (41)	9 (10)	-
	≥2	-	-	3 (14)	8 (8)	-
	Unknown	-	-	4 (18)	77 (82)	-

Based on Table 23 of the CS[2] which is the same as Table 11 of the company's response to the clarification letter[16] and primary sources: tabulated data from Shyamalee et al. 2021[46] and Jaffe et al. 2022[47]

[a]82% and 89% received 1L treatment in overall and western populations respectively; other patients received BSC.[47]

[b]Patients with Asian location 40.2%; patients with western location 59.8%. Western locations include UK but number of UK patients not reported.[47]

1L = first line; BSC = best supportive care; CHEMO = fluorouracil and cisplatin chemotherapy; CS = company submission; ECOG PS = Eastern Cooperative Oncology Group Performance Status; NIVO-CHEMO = nivolumab combined with fluorouracil and cisplatin chemotherapy; OSCC = oesophageal squamous cell carcinoma; RCT = randomised controlled trial; SD = standard deviation; UK = United Kingdom

In Section B.2.13.4 of the CS[2] the company argued that average age of participants was similar between CheckMate 648[6] and the two retrospective cohort studies.[46, 47] The ERG agreed that this was the case between CheckMate 648,[6] Jaffe et al. 2022[47] and the patients receiving palliative chemotherapy in Shyamalee et al. 2021[46] with an average of 64 years across these populations. However, the patients managed with best supportive care (BSC) alone in Shyamalee et al. 2021 were older by comparison (median 79 years). The group receiving BSC alone comprised the majority of reported participants (94/116, 81%) in this study[46] which, despite this study being UK-based, would suggest limited relevance to CheckMate 648.[6, 7] The sex distribution across the three studies was tabulated but not discussed in the CS.[2] The ERG noted the lower proportion of male patients in both populations of the Shyamalee et al. 2021 study (41% for palliative chemotherapy and 48% for BSC alone)[46] compared with CheckMate 648 (82% overall)[6] and Jaffe et al. 2022 (83%).[47] The data on baseline ECOG PS suggested larger proportions of patients with scores of zero in CheckMate 648 compared with the palliative chemotherapy and BSC alone populations in Shyamalee et al. 2021 (47%, 27% and 0% respectively).[6, 7, 46] A similar pattern was seen for the proportions of patients with an ECOG PS score of one (53%, 41% and 10% respectively).[6, 46] However, it should be noted that Shyamalee et al. 2021 reported on a wider range of ECOG PS score categories (including ≥2 and "*unknown*") and scores were classified as "*unknown*" for 82% of patients receiving BSC alone.[46] Data on baseline ECOG PS were not reported for Jaffe et al. 2022.[47]

Section B.2.13.4 of the CS[2] explores the issue of baseline ECOG PS scores further, citing a systematic review by Cheng et al. 2017.[48] This review investigated the effects of novel systemic cancer therapies (in terms of efficacy and toxicity) in patients with reduced versus excellent ECOG PS scores. Sixty-six RCTs recruiting patients with different types (sites) of cancer and evaluating a variety of different therapies were included.[48] The company highlighted the: "*…limited evidence to suggest different outcomes between patients with different performance score*" (Section B.2.13.4 of the CS).[2] The results of the review suggested no difference between reduced versus excellent ECOG PS subgroups for efficacy outcomes (OS and PFS) however, none of the included studies reported results in terms of toxicity outcomes. The ERG noted the low volume of evidence potentially relating to OSCC, with 4/66 (6%) of included RCTs recruiting patients with gastric or gastro-OC. Details of histology were not provided and the number of RCTs recruiting patients with OSCC was not reported.[48] This leaves persisting uncertainty regarding the impact of baseline ECOG PS scores on efficacy and safety outcomes in patients with OSCC.

The company continued their arguments in Section B.2.13.4 of the CS[2] to support the application of results of the CheckMate 648 RCT[6] to patients seen in routine clinical practice in the UK by discussing the baseline data from a further five studies.[25, 49-52] The presentation of this in the CS lacked an overall summary and therefore the ERG requested a tabular presentation of baseline data from CheckMate 648 plus all five comparator studies to facilitate an overall comparison (question A16 in the clarification letter). The company provided further information in response (Table 8 of the clarification letter response),[16] which the ERG used as the basis of a summary table. The ERG included additional details to help provide a suitable context for interpretation, namely: study design; study location; participant selection criteria in terms of cancer site and histology; and PD-L1 status categories used per study (Table 3.19).

It was not clear from the CS whether PD-L1 status data were available for all comparator studies and the ERG asked the company to provide this information (clarification question A17). The company provided details as shown in Table 3.20.[16]

As a further part of their response to clarification question A16,[16] the company provided tabulation of outcomes (median OS) for CheckMate 648[6] and the five comparator studies[25, 49-52] (Table 3.21).

Table 3.19: Baseline characteristics of patients enrolled in CheckMate 648 and five comparator studies

Study name	CheckMate 648[6, 7]		COUGAR-02[49]		Royal Marsden[50]	KEYNOTE-590[25]		CheckMate 649[51]		KEYNOTE-062[52]	
Study level details											
Study design	Phase III RCT		Phase III RCT		Retrospective cohort	Phase III RCT		Phase III RCT		Phase III RCT	
Study location details	International: 182 sites in 26 countries; includes five UK sites (n=[redacted] UK patients)*		30 sites in UK		Single site in UK	International: 168 sites in 26 countries; includes three UK sites (n=22 UK patients)		International: 175 sites in 29 countries; number of UK sites/patients N/R[a]		International: 200 sites in 29 countries; number of UK sites/patients N/R[b]	
Participant selection criteria (cancer site and histology)	Unresectable advanced, recurrent or metastatic, untreated OSCC, regardless of PD-L1 status		Advanced AC of oesophagus, stomach or OGJ		Advanced AC of oesophagus, stomach or OGJ	Unresectable locally advanced, or metastatic, untreated OSCC or oesophageal AC or Siewert type 1 OGJ AC, regardless of PD-L1 status		Unresectable advanced, or metastatic, untreated AC of oesophagus, stomach or OGJ, regardless of PD-L1 expression		Unresectable locally advanced, or metastatic AC of stomach or OGJ, with PD-L1 CPS ≥1	
PD-L1 status categories	PD-L1 TC <1% and PD-L1 TC ≥1%		N/R		N/R	PD-L1 CPS <10 and PD-L1 CPS ≥10		PD-L1 CPS <5 and PD-L1 CPS ≥5; also presented as PD-L1 TC <1% and PD-L1 TC ≥1%		PD-L1 CPS ≥1 but <10; and PD-L1 CPS ≥10	
Arm level details											
Treatment comparison[c]	NIVO-CHEMO	CHEMO	Docetaxel	Active symptom control	Not a comparative design; patients had mix of 1L, 2L	PEMBRO-CHEMO	CHEMO	NIVO-CHEMO	CHEMO	PEMBRO-CHEMO	CHEMO

Study name		**CheckMate 648**[6, 7]		**COUGAR-02**[49]		**Royal Marsden**[50]	**KEYNOTE-590**[25]		**CheckMate 649**[51]		**KEYNOTE-062**[52]	
						and 3L treatments						
Number of patients randomised		321	324	84	84	N=511 included (not randomised)	373	376	789	792	257	250
Male sex (%)		78.8	84.9	82	80	75	82	85	68	71	75.9	71.6
Median age (range) in years		64 (40 to 90)	64 (26 to 81)	65 (28 to 84)	66 (36 to 84)	66 (24 to 90)[d]	64 (28 to 94)	62 (27 to 89)	62 (IQR 54 to 69)	61 (IQR 53 to 68)	62 (22 to 83)	63 (23 to 87)
ECOG PS, n (%)	0	150 (46.7)	154 (47.5)	24 (28)	22 (26)	64 (13)	149 (40)	150 (40)	326 (41)	336 (42)	119 (46)	115 (46)
	1	171 (53.3)	170 (52.5)	46 (55)	50 (60)	276 (54)	223 (60)	225 (60)	462 (59)	452 (57)	138 (53.7)	135 (54.0)
	2	0 (0)	0 (0)	14 (17)	12 (14)	87 (17)	1 (<1)	1 (<1)	1 (<1)	3 (<1)	-	-
	3	0 (0)	0 (0)	0 (0)	0 (0)	1 (<1)	-	-	-	-	-	-
	Not recorded	0 (0)	0 (0)	0 (0)	0 (0)	83 (16)	-	-	0 (0)	1 (<1)	-	-
Disease status, n (%)	Metastatic	184 (57)	187 (58)	73 (87)	74 (88)	443 (87)[e]	344 (92)	339 (90)	757 (96)	756 (95)	243 (95)	235 (94)
	Recurrent, locoregional	21 (7)	25 (8)	-	-	-	-	-	5 (<1)	2 (<1)	-	-
	Recurrent, distant	72 (22)	60 (19)	-	-	-	-	-	-	-	-	-
	Advanced[f]	Unresectable advanced		Locally advanced		Locally advanced (unresectable)	Locally advanced (unresectable)		Locally advanced		-	-
		44 (14)	52 (16)	11 (13)	10 (12)	68 (13)	29 (8)	37 (10)	27 (3)	34 (4)		
Histology, n (%)	SCC	311 (96.9)	318 (98.1)	-	-	-	274 (73)	274 (73)				

Study name		CheckMate 648[6, 7]		COUGAR-02[49]		Royal Marsden[50]	KEYNOTE-590[25]		CheckMate 649[51]		KEYNOTE-062[52]	
	AC	[redacted]	[redacted]	84 (100)	84 (100)	511 (100)	99 (27)	102 (27)	789 (100)	792 (100)	257 (100)	250 (100)
	Other	[redacted]	[redacted]	-	-	-	-	-				

Based on Tables 8 and 10 of the company's response to the clarification letter[16] and primary sources of information.[6, 7, 25, 49-53]
[a]NCT02872116 record suggests that five UK sites were anticipated[54] but no information on the number of UK sites/patients were provided in the paper.[51]
*Of the [redacted] UK patients, [redacted] had PD-L1 >1% expression, [redacted] of whom were in the NIVO+CHEMO arm and [redacted] in the CHEMO arm.[16]
[b]The paper lists one UK study site among the author affiliations but does not provide information on the number of UK patients.[52]
[c]Constituents of chemotherapy varied across studies
[d]Age at diagnosis, not study baseline
[e]Includes 335 patients with *de novo* metastatic disease and 108 patients with relapsed metastatic disease after radical treatment
[f]Exact description of advanced disease varied across studies
1/2/3L = first/second/third line; AC = adenocarcinoma; CHEMO = chemotherapy; CPS = combined positive score; ECOG PS = Eastern Cooperative Oncology Group Performance Status; IQR = interquartile range; NIVO-CHEMO = nivolumab + chemotherapy; N/R = not reported; OGJ = oesophagogastric junction; OSCC = oesophageal squamous cell carcinoma; PD-L1 = programmed death ligand 1; PEMBRO-CHEMO = pembrolizumab + chemotherapy; RCT = randomised controlled trial; SCC = squamous cell carcinoma; TC = tumour cells; UK =-United Kingdom

Table 3.20: Data availability for PD-L1 subgroups in CheckMate 648 and the comparator studies

PD-L1 status	CheckMate 648[6, 7]	COUGAR-02[49]	Royal Marsden[50]	KEYNOTE-590[25]	CheckMate 649[51]	KEYNOTE-062[52]
PD-L1 ≥1% TC	Yes	No	No	No	No	Yes[a]
PD-L1 ≥5% CPS	No	No	No	No	Yes	No
PD-L1 ≥10% CPS	No	No	No	Yes	No	Yes
PD-L1 <10% CPS	No	No	No	Yes	No	No

Based on Table 10 of the company's response to clarification question A17.[16]
[a]This entry is per Table 10 of the company's response to clarification question A17[16] however, this information is incorrect as data on the PD-L1 ≥1% subgroup were available from Table 1 of the trial paper.[51]
CPS = combined positive score; PD-L1 = programmed death ligand 1; TC = tumour cells

Table 3.21: OS of patients enrolled in CheckMate 648 and the five identified comparator studies

Study name	CheckMate 648[6, 7]		COUGAR-02[49]		Royal Marsden[50]	KEYNOTE-590[25]		CheckMate 649[51]		KEYNOTE-062[52]	
Treatment comparison[a]	NIVO-CHEMO	CHEMO	Docetaxel	Active symptom control	Not a comparative design; patients had mix of 1L, 2L and 3L treatment	PEMBRO-CHEMO	CHEMO	NIVO-CHEMO	CHEMO	PEMBRO-CHEMO	CHEMO
N	321	324	84	84	511	373	376	789	792	257	250
Median survival (months)	[redacted]	[redacted]	5.2	3.6	11.5	12.6	9.8	14.4	11.1	10.6	11.1
95% CI	[redacted]	[redacted]	4.1 to 5.9	3.3 to 4.4	10.5 to 12.5	10.2 to 14.3	8.6 to 11.1	13.1 to 16.2	10.0 to 12.1	7.7 to 13.8	9.2 to 12.8

Based on Table 9 of the company's response to clarification question A16[16] and Table 7 of Appendix E of the CS.[12]
[a]Constituents of chemotherapy varied across studies.
1/2/3L = first/second/third line; CHEMO = chemotherapy; CI = confidence interval; CS = company submission; NIVO-CHEMO = nivolumab + chemotherapy; OS = overall survival; PEMBRO-CHEMO = pembrolizumab + chemotherapy

ERG comment:

Identification and presentation of supporting evidence

With the exception of the KEYNOTE-590 RCT,[25] the methods used for identifying and selecting the references providing supporting evidence were not explained in the CS[2] or the response to clarification questions.[16] It is therefore difficult to judge whether other suitable studies could have been overlooked.

The company cited one systematic review[48] and seven primary studies.[25, 46, 47, 49-52] to support the application of results of the CheckMate 648 RCT[6, 7] to patients seen in routine clinical practice in the UK (Section B.2.13.4 of the CS).[2] The ERG noted a series of errors in the company's description of some studies, as follows:

- Incorrect tabulation of data for Shyamalee et al. 2021[46] in Table 23 of Document B:[2] cohort size (n=219) for overall number of patients diagnosed with OSCC during the study period was conflated with median age (63 years) and ECOG PS data for a subset of patients receiving palliative chemotherapy (n=22) and proportion of males (48%) for a second subset of patients managed with BSC alone (n=94). The ERG has tabulated the available information separately for patients receiving palliative chemotherapy and BSC alone (Table 3.18). It appears that full information relating to the n=219 patients identified was not provided in the conference abstract.[46]
- Incorrect description of data from Jaffe et al. 2022 on page 93 of Document B[2] and on page 33 (question A18b) of the company's response to clarification questions:[16] the study reported the mean age across the overall population (covering Asia and the West)[47] whilst the company incorrectly linked this only to the western cohort.[2] Furthermore, the company's response to clarification questions describes Jaffe et al. 2022 as being "*conducted in UK OSCC populations*" (question A14) and there is a similar statement as part of the response to question A18c.[16] This is potentially misleading and it should be noted that the number of UK study sites and patients were not provided in the conference abstract and neither were age data stratified by geographical region. Therefore the ERG presented data for the overall international population as reported in the study reference[47] in Table 3.18.
- Incorrect tabulation of the Cougar-02 study:[49] shown in Table 8 of the company's response to clarification questions (question A16) as recruiting only patients with OC[16] whereas those with adenocarcinoma of the stomach or oesophagogastric junction (OGJ) were also eligible for inclusion.[49] The ERG includes the correct details in Table 3.19.
- Incorrect information for median (range) age in KEYNOTE-062[52] shown in Table 8 of the company's response to clarification question A16:[16] this appears to have been confused with the number and proportions of patients in Asian locations as presented in Table 1 of the study paper.[52] The ERG includes the correct details in Table 3.19.
- Missing information on ECOG PS scores for KEYNOTE-062[52] in Table 8 of the company's response to clarification question A16:[16] the paper reports the number of patients with ECOG PS = 1 in Table 1 and the number with ECOG PS = 0 can be inferred from this since the participant selection criteria (page 1572 of the paper) states that patients had to have a score of 0 or 1 to be included in the trial.[52] The ERG has included this information in Table 3.19.
- Missing information on disease status for the Royal Marsden study[50] in Table 8 of the company's response to clarification question A16:[16] the patients with relapsed metastatic disease after radical treatment were not represented in the Table (meaning that numbers for disease status did not sum to the total N and percentages did not sum to 100); however, a footnote was provided. The ERG has included all information on disease status in Table 3.19.

- Missing information on ECOG PS for the Royal Marsden study[50] in Table 8 of the company's response to clarification question A16:[16] frequencies for score of 3 and "*not recorded*"[50] were not shown therefore the numbers/percentages did not sum to the expected totals. The ERG has included all information on ECOG PS in Table 3.19.
- Missing information on histology for CheckMate 649[51] in Table 8 of the company's response to clarification question A16:[16] all recruited patients had adenocarcinoma,[51] as indicated by the ERG in Table 3.19.
- It is not clear why Table 8 of the company's response to clarification question A16[16] restricted information to the subgroup with PD-L1 ≥1% for the CheckMate 648 RCT[6, 7] whilst data for the overall populations were presented for the comparator studies. The ERG has presented data for all randomised patients in CheckMate 648[6, 7] and has included details on the PD-L1 expression categories used per study in Table 3.20.
- Table 10 of the clarification response (details of data availability for PD-L1 subgroups) included a footnote that incorrectly referred to subgroup data.[16] In addition, the details for the CheckMate 649 RCT were incorrect, suggesting that data for the PD-L1 TC ≥1% subgroup were not available whereas these data were presented in Table 1 of the trial paper.[51]

The ERG has included complete and corrected baseline information for all studies (presented in Table 3.19).

Consideration of ethnicity

In the clarification letter (question A14), the ERG asked the company to discuss the representativeness of the CheckMate 648 trial population to UK clinical practice. In their response,[16] the company referred to consulting UK clinicians during an advisory board meeting.[55] The company reported that:

"*When asked if the baseline characteristics observed in patients randomised in CheckMate 648 were representative of those seen in UK clinical practice, the clinicians agreed that the trial patients were broadly younger than they would expect, but otherwise, the patients were representative of UK OSCC patients.*"[16]

The company also made comments in relation to the balance between Asian and non-Asian participants in CheckMate 648 and how these characteristics might be generalised to patients seen in UK clinical practice:

"*The clinicians agreed during the advisory board that the high proportion of Asian patients in the CheckMate 648 trial (70%) was not an issue when applying the trial data to a UK population. It was explained that in oesophageal adenocarcinoma the imbalance between Asian and non-Asian patients would be an issue as patients are treated over several different lines of therapy. However, this is not the same in OSCC and so should not be considered an issue. It was confirmed that there was no biological reason to consider the populations to be different.*"

"*This is further supported by the data presented in Section B.2.7 where subgroup analysis demonstrated favourable OS for nivolumab with chemotherapy in both Asian and non-Asian populations.*"[16]

Scrutiny of the abovementioned subgroup analyses for Asian and non-Asian populations suggests a high degree of overlap between CIs (Section B.2.7 of the CS): unstratified HR estimates for OS for NIVO-CHEMO versus CHEMO [redacted] and [redacted] for Asian and non-Asian subgroups respectively.[2] These estimates were for all randomised patients and none were provided for the subgroup of patients with PD-L1 ≥1%.

The company completed their response by mentioning that:[16]

"*…the imbalance between Asian and non-Asian patients was not considered to be an issue in the pembrolizumab assessment in OSCC.[1]*"

Consideration of other baseline variables

Table 3.19 shows that two RCTs (CheckMate 648[6, 7] and KEYNOTE-590[25]) recruited UK patients with OSCC. In their response to clarification question A15, the company confirmed that no other studies were identified that provided information on UK patients with OSCC.[16] The other four RCTs in Table 3.19 limited recruitment to patients with adenocarcinoma of the gastro-oesophageal tract.[49-52]

In the CS (Section B.2.13.4), the company outlines arguments around the baseline data for CheckMate 648 being comparable to other studies involving UK populations.[2] Of the seven studies cited to support these arguments, three recruited solely UK participants[46, 49, 50] whilst four were conducted across different countries including the UK.[25, 47, 51, 52] Of the three UK studies, one recruited patients with OSCC the majority of whom (81%) were being managed with BSC alone. These patients were older on average than those in the other studies (median 79 years) and for 82% of patients the ECOG PS was unknown.[46] Participants in the other two UK studies were receiving active treatment for adenocarcinoma of the gastro-oesophageal tract and none had squamous cell histology.[49, 50] These factors would seem to highlight some differences versus CheckMate 648.[6, 7] Among the four international studies, two recruited patients with adenocarcinoma of the gastro-oesophageal tract[51, 52] and one included participants with OSCC but provided only limited information on baseline data because of being published as a conference abstract.[47] The fourth study (the KEYNOTE-590 RCT)[25] enrolled a majority of participants with OSCC (73%) and a larger proportion of patients with metastatic disease relative to CheckMate 648[6, 7] (91% versus 58%). This RCT was selected for inclusion in an ITC as part of the CS[2] and is discussed in more detail in Sections 3.3 and 3.4.

The company's conclusion is that: "*CheckMate 648 baseline characteristics and outcomes are well aligned to the published evidence base, and so can be considered highly relevant to UK clinical practice*" (page 105 of the CS).[2] However, the ERG suggests that more caution is required. The two retrospective cohorts were in populations that were dissimilar (the majority receiving BSC alone)[46] or of unknown comparability (number of UK patients not reported)[47] to CheckMate 648.[6, 7] Of the five remaining primary studies, two were conducted in the UK[49, 50] and the other three were international,[25, 51, 52] with one confirming inclusion of UK patients.[25] Four studies recruited solely patients with adenocarcinoma of different parts of the oesophageal tract[49-52] and one included a 73%/27% split between squamous cell and adenocarcinoma histology[25] (compared with solely OSCC for CheckMate 648).[6, 7] Other potential baseline differences included: a lower proportion of male patients in some comparator studies (around 72%)[51, 52] versus 82% for CheckMate 648;[6, 7] higher proportions of patients with metastatic disease across all five comparator studies (range 87% to 96%)[25, 49-52] versus 58% for CheckMate 648;[6, 7] and smaller proportions with locally advanced disease in some comparator studies (range 3% to 10%)[25, 51] versus 15% in CheckMate 648.[6, 7] Average patient age and distribution of ECOG PS scores appeared broadly comparable between the comparator studies[25, 49-52] and CheckMate 648.[6, 7]

Overall survival estimates

The company incorrectly described the population for KEYNOTE-590 as the OSCC subgroup in Table 9 of the company's response to clarification question A16.[16] However, it was apparent from the numbers per treatment arm that the population was all randomised patients (i.e., it also included those with oesophageal adenocarcinoma and Siewert type 1 OGJ adenocarcinoma).[25]

The ERG noted comparable estimates between the two RCTs including UK patients with OSCC (CheckMate 648[6, 7] and KEYNOTE-590[25]) (Table 3.19). Estimates from the other studies (all for treatments other than nivolumab or pembrolizumab) ranged from 3.6 to 11.5 months.[49-52]

Overall, the ERG remains uncertain about the extent to which results from CheckMate 648 can be generalised to patients seen in UK clinical practice.

3.2.11 Ongoing studies

As confirmed in Section B.2.11 of the CS,[2] CheckMate 648 remains ongoing to further follow-up.

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

One study, KEYNOTE-590, was identified and used in the ITC analyses. The study was identified in the SLR[12]. KEYNOTE-590 compared pembrolizumab with chemotherapy. The ITT population has a diverse mix of patients regarding the type of oesophageal cancer and PD-L1 CPS expression as shown in Table 3.22. The study includes patients with:

- locally advanced unresectable or metastatic oesophageal adenocarcinoma or SCC and
- advanced/metastatic OGJ Siewert type 1 adenocarcinoma.

A comparison between trial characteristics, eligibility criteria and treatment characteristics and baseline characteristics is presented in Section 3.4 in Table 3.26 and 3.27. Baseline characteristics were not available for the OSCC sub-population of KEYNOTE-590 and therefore the entire ITT population was presented in the CS. Efficacy outcomes of interest are presented in Table 3.23. Both studies reported significant improvement on PFS and OS outcomes (HR) of patients receiving combination therapy compared to chemotherapy alone.

KEYNOTE-590 reported participants with PD-L1 CPS <10% and ≥10%. Data for the ITC analysis came from the PD-L1 ≥10% (CPS) population (n=383) (see Table 3.22), as data for PD-L1 CT expression were not available. In order to compare the two studies, only patients with PD-L1 CPS ≥10% from CheckMate 648 were included in the NMA (n=280). PD-L1 status is an integral part of the CS in terms of scope for both population and comparators. The company, responding to a question by the ERG, has clarified that PD-L1 testing is executed in UK practice by using one of the two similar methods of TC/tumour proportion score (TPS) or CPS, which are partially overlapping. TC/TPS, which was used in CheckMate 648, is obtained by dividing the number of PD-L1 stained/positive tumour cells by the total number of viable tumour cells. CPS, on the other hand, is calculated by dividing the number of both tumour and non-tumour PD-L1 expressing cells by the number of all tumour cells[16].

Table 3.22: KEYNOTE-590 sub-populations used in ITC analyses

KEYNOTE- 590 sub-populations	N=
ITT[a]	**749**
OSCC	**548**
Pembrolizumab + hemotherapy	274
Chemotherapy	274
PD-L1 ≥10% (CPS)/ used in PFS ITC	**383**
Pembrolizumab + chemotherapy	186
Chemotherapy	197

KEYNOTE- 590 sub-populations	N=
OSCC, PD-L1 ≥10% (CPS)/ used in OS ITC	**286**
Pembrolizumab + chemotherapy	143
Chemotherapy	143

Based on the data reported in Document B of the CS[2], Appendix L of the CS[56] and updated Appendix L of the CS[57]
[a]includes patients with locally advanced unresectable or metastatic oesophageal adenocarcinoma or SCC and advanced/metastatic esophagogastric junction Siewert type 1 adenocarcinoma
CPS = combined positive score; CS = company submission; ITC = indirect treatment comparison; ITT = intention to treat; OS = overall survival; OSCC = oesophageal squamous cell carcinoma; PD-L1 = programmed death ligand-1; PFS = progression free survival

Table 3.23: Summary of efficacy outcomes of the studies included in the ITC analyses (PD-L1 CPS ≥ 10%.; CheckMate 648 and KEYNOTE-590.

Trial Name	CheckMate 648		KEYNOTE-590	
Treatment Arm	Nivolumab + 5-FU + cisplatin	5-FU + Cisplatin	Pembrolizumab + 5-FU + cisplatin	5-FU + cisplatin
Follow-up (months), median	11.2		10.8	
Sample size	135	145	186 (143[a])	197 (143[a])
OS (months), median (95% CI)	[redacted]	[redacted]	13.96 (11.13–17.74)	8.88 (7.82–10.55)
OS HR (95% CI)	[redacted]	[redacted]	0.563 (0.43–0.74)	--
PFS (months), median (95% CI)	[redacted]	[redacted]	7.43 (6.25–8.19)	5.49 (4.29–6.03)
PFS HR (95% CI)	[redacted]	[redacted]	0.536 (0.43–0.67)	--
Method of Assessment	BICR		IA	

Based on updated Appendix L of the CS[57]
[a] only OSCC patients
5-FU = fluorouracil; BICR = Blinded Independent Central Review; CI = confidence interval; CS = company submission; ECOG PS = Eastern Cooperative Oncology Group Performance Status; HR = hazard ratio; IA = investigator assessed; ITC = indirect treatment comparison; N/R = not reported; OS = overall survival; PFS = progression-free survival

ERG comment:

- According to the company the method used for defining PD-L1 expression in current practice is a matter of choice. It is associated to the site of the disease in the body and is driven by the currently recommended therapies.[16] The scope of the CS as well as the research hypotheses, as defined in the protocol for CheckMate 648[11], were all targeting PD-L1 TC expression ≥ 1%. The choice of TC over CPS is not explored in the CS, which would clarify the rationale and the justification for its use.

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

The company stated that they performed a NMA of OS and PFS with the goal of including pembrolizumab with chemotherapy, as assessed in KEYNOTE-590, as a comparator arm within the CEM. The NMA considered the PD-L1 ≥10% (CPS) population, in line with the population reported in KEYNOTE-590. Therefore, only patients with PD-L1 CPS ≥10% from CheckMate 648 were included in the NMA for comparison with pembrolizumab, a subpopulation of the target population for this submission. The overlapping populations for TC ≥1% and CPS ≥ 10% in CheckMate 648 are summarised on Table 3.24. The proportions of patients that were reported to have a either PD-L1 CPS ≥10% or TC ≥1% expression in both studies is reported in Table 3.25.

Table 3.24: A summary of the overlapping TC ≥1% and CPS ≥ 10% populations in the CheckMate 648 trial

	CPS ≥10	CPS < 10 & NA	Total
Nivolumab - Chemotherapy			
TC ≥1	96	62	158
TC < 1	39	124	163
Total	135	186	321
Chemotherapy			
TC ≥1	100	57	157
TC < 1	45	122	167
Total	145	179	324

Based on Table 14 of Document B of the CS[2]
CPS = combined positive score; CS = company submission; PD-L1 = programmed death ligand-1; TC = tumour cell

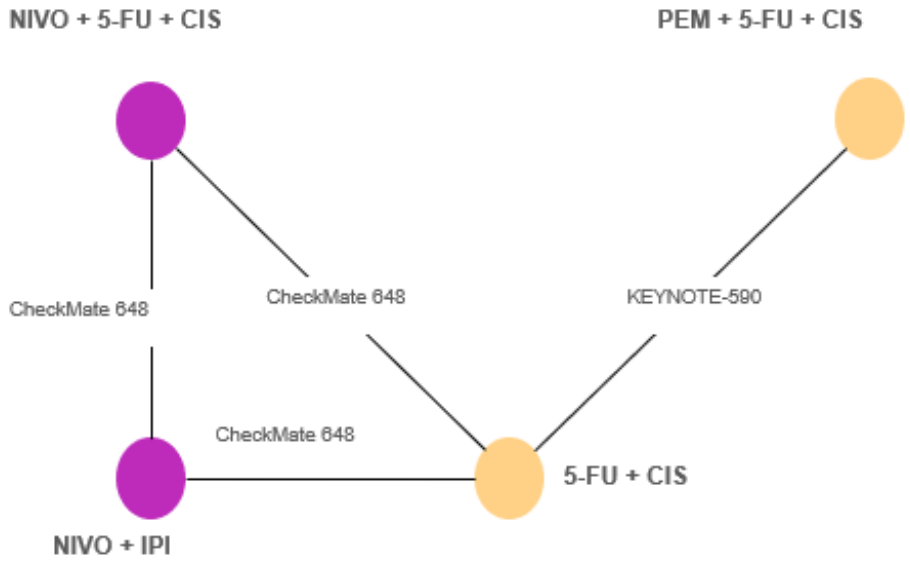
Table 3.25: The proportion of PD-L1 expression reported in KEYNOTE-590 and CheckMate 648

	KEYNOTE-590	CheckMate 648
PD-L1 CPS <10%	46.3%	Can be calculated from PLD
PD-L1 CPS ≥10%	51.1%	Can be calculated from PLD
TC ≥1%	N/R	49.2%
PD-L1 CPS ≥10 and TC ≥1%	N/R	Can be calculated from PLD

Based on Table 4 of the company's response to request for clarification from the ERG[16]
CPS= combined positive score; N/R = not reported; PD-L1 = programmed death ligand-1; PLD = patient-level data; TC = tumour cell

The network diagram of the NMA is illustrated in Figure 3.14. It should be highlighted that the presented network includes intervention 2 of CheckMate 648 (NIVO+IPI), but this arm is not taken into consideration in the ITC. Given that only two trials were included in the so-called NMA, the term ITC, which is also used in Document B of the CS[2], is a better description of the analysis.

Figure 3.14: Network diagram



Based on Figure 21 of Document B of the CS[2]
Note: NIVO + 5-FU + CIS, nivolumab 240 mg Q2W IV + fluorouracil 800 mg/m2/day IV on Day 1 through Day 5 + cisplatin 80 mg/m2 IV on Day 1 of a 4-week cycle; NIVO+IPI, nivolumab 3 mg/kg every 2 weeks (Q2W) intravenously (IV) + ipilimumab 1 mg/kg every 6 weeks (Q6W) IV.
CS = company submission; Q2W = every two weeks, IPI = ipilimumab; IV = intravenously; NIVO = nivolumab

Baseline characteristics were not available for individual patients or subgroups of interest in KEYNOTE-590, thus, the comparability assessment was based on the entire ITT population (n=749) and not the OSCC PD-L1 CPS ≥10% population (n=286). The company stated that the assessment found the populations to be sufficiently similar in terms of both study design and patient baseline characteristics. The company has provided further explanation on how the assessment was executed in an updated Appendix L to the CS[57], as well in their response to request for clarification from the ERG[16]. The summary of the characteristics used for the comparison regarding study design, eligibility criteria of patients and treatment characteristics are presented in Table 3.7. The summary of the baseline characteristics is presented in Table 3.8. CheckMate-648 was an open label trial whereas KEYNOTE-590 was double-blind, both studies were multi-centred, and follow-up time is comparable. The eligibility criteria differ between the two. CheckMate-648 included only SCC located on the oesophagus and allowed prior adjuvant therapy if it was completed ≥6 months prior to enrolment, on the other hand, KEYNOTE-590 included both SCC or adenocarcinoma located in either the oesophagus or the gastroesophageal junction and according to the CS prior treatments were not permitted. Dosing of the regimens was similar across arms, but the duration of cycles differed, leading to more extended treatment durations in KEYNOTE-590. Regarding patients' baseline characteristics CheckMate 648 had a higher proportion of Asian patients (70% in either arm) than KEYNOTE-590 (53% and 52%, in the two arms) and a much lower proportion of patients with metastatic disease (58% versus 91% in both arms combined).

The design, eligibility criteria, treatment characteristics and baseline patients characteristics were used to perform the NMA feasibility assessment which is now reported in the updated Appendix L of the CS.[57] The assessment concluded that the two studies were sufficiently similar. The company acknowledged that the comparison between the baseline characteristics should be interpreted with caution, as they were based on the all-comers population for KEYNOTE-590[16, 57] (see ERG comment below).

Table 3.26: Summary of study design, eligibility criteria and treatment characteristics for CheckMate 648 and KEYNOTE-590

	CheckMate 648	KEYNOTE-590
Study design		
Phase	III	III
Sample size	970	749 (ITT) 548 (SCC)
Masking	Open-label	Double-blind
Geographic locations	Argentina, Australia, Austria, Brazil, Canada, Chile, China, Colombia, Czech Republic, Denmark, France, Hong Kong, Italy, Japan, Mexico, Peru, Poland, Portugal, Romania, Russia, South Korea, Spain, Taiwan, Turkey, UK, US	Argentina, Australia, Brazil, Canada, Chile, China, Colombia, Costa Rica, Denmark, France, Germany, Guatemala, Hong Kong, Japan, Peru, Romania, Russia, South Africa, South Korea, Spain, Thailand, Turkey, UK, US
Follow-up duration, median (months)	11.2	10.8 (ITT)
Eligibility criteria		
Histology	SCC	SCC or adenocarcinoma
Tumour Location	Oesophagus	Oesophagus or gastroesophageal junction
ECOG PS	0–1	0–1
Prior Adjuvant Therapy	Eligible if completed ≥6 months prior to enrolment	N/R[a]
Treatment characteristics		
Intervention	Nivolumab (240 mg, Day 1 and Day 15) + 5-FU (800 mg/m2, Days 1-5) + cisplatin (80 mg/m2, Day 1)	Pembrolizumab (200 mg, Day 1) + 5-FU (800 mg/m2, Days 1-5) + cisplatin (80 mg/m2, Day 1)
Comparator	5-FU (800 mg/m2, Days 1-5) + cisplatin (80 mg/m2, Day 1)	5-FU (800 mg/m2, Days 1-5) + cisplatin (80 mg/m2, Day 1)
Cycle Length	28 days	21 days
Stopping Rules	Nivolumab: 24 months Chemotherapy: 6 cycles	Pembrolizumab: ≤35 cycles Chemotherapy: ≤6 cycles
Treatment Duration, median	NIVO+CHEMO: 5.7 months (IQR: 2.7–10.0) CHEMO: 3.4 months (IQR: 1.3–5.7)	PEMBRO+CHEMO: 7.7 months* (SD: 6.8) CHEMO: 5.8 months* (SD: 4.8)

Based on Tables 2, 3 and 4 of the updated Appendix L of the CS[57]
[a]see ERG comments on allowed prior treatments in Section 3.4

*Mean (not median) value.

CHEMO = chemotherapy; CS = company submission; ECOG PS = Eastern Cooperative Oncology Group Performance Status; ERG = Evidence Review Group; FU = fluorouracil; IQR = interquartile range; ITT = intention to treat; NIVO = nivolumab; N/R = not reported; PEMBRO = pembrolizumab; SCC = squamous cell carcinoma; SD = standard deviation; UK = United Kingdom; US = United States

Table 3.27: Summary of baseline patient characteristics for CheckMate 648 and KEYNOTE 590

Study	CheckMate 648		KEYNOTE-590	
Treatment Arm	**Nivolumab + chemotherapy**	**Chemotherapy**	**Pembrolizumab + chemotherapy**	**Chemotherapy**
Sample Size	321	324	373	376
Age (years), median	64	64	64	62
Asian	Total: 70% East Asia: 57% Rest of Asia: 13%	Total: 70% East Asia: 57% Rest of Asia: 13%	53%	52%
ECOG PS (0)	46%	47%	40%	40%
ECOG PS (1)	54%	53%	60%	60%
Metastatic disease status	57%	58%	92%	90%
Organs with Metastases	≤1: 49% ≥2: 51%	≤1: 49% ≥2: 51%	N/R	N/R
Liver Metastases	N/R	N/R	N/R	N/R
PD-L1 ≥10% (CPS)	135	145	186	197
PD-L1 ≥1% (TC)	158	157	N/R	N/R
PD-L1 ≥10% (CPS) and PD-L1 ≥1% (TC)	96	100	N/R	N/R

Based on Table 9 of Document B of the CS[2], Table 5 of the updated Appendix L of the CS[57] and Table 6 of NICE STA ID 3714[53]
CPS = combined positive score; CS = company submission; ECOG PS = Eastern Cooperative Oncology Group Performance Status; N/R = not reported; PD-L1 = programmed death ligand-1; TC = tumour cell

3.4.1 Estimate of time varying HRs from KEYNOTE-590

The OS and PFS IPD were obtained by digitising the K-M curves for intervention and comparator from KEYNOTE-590 using the method by Guyot 2012[58]. There is an expected margin of error when using this method, but Guyot 2012[58] state that is not significant in terms of introducing systematic error. The company has clarified[16], that the margin of error was taken into consideration in the analysis, but the presentation of the K-M curves at 3-month intervals was at such a level that it was enough to produce satisfactory data.

The company then rejected the proportional hazards assumption based on an examination of Schoenfeld residual, hazard and log cumulative hazard plots, as presented in Appendix L[56]. Therefore, the survival models were fit for each arm independently. Various standard parametric and spline models (hazard, odds and normal; up to three knots for each) were then fitted to the PFS and OS data separately for each of the arms, pembrolizumab with chemotherapy and chemotherapy.

Two partially different populations have been used in the OS and the PFS analysis. Although the population of interest in the NMA is indeed the OSCC PD-L1 ≥10% (CPS) subgroup (n=286), this was

only used for the OS analysis and not for PFS, as these data were not available. Therefore, the broader ITT PD-L1 ≥10% (CPS) population was used instead (n=383). Table 3.4 shows the overlap of the population. The company states that "*Use of mixed histology for the PFS within the PD-L1 ≥10% (CPS) subgroup was based on clinical input from an advisory board and supported by the comparison of hazard ratios (HRs) between the SCC (n=548) and AC (n=241) populations from KEYNOTE-590 with HRs of 0.72 (95% CI 0.60 to 0.88) and 0.74 (95% CI 0.54 to 1.02), respectively.*"[57].

For PFS all parametric models were considered "*indistinguishable*" visually (Figure 9, Appendix L) and statistically (according to AIC and BIC), generalised gamma, log normal, and log-logistic models, all spline models, excluding the 3-knot (odds and normal) were stated to be preferred. Exclusion of the 3-knot (odds and normal) spline model appeared to be based on simplicity*: "…the most simplistic models, 1-knot and 2-knots, should be considered."* (p. 30, Appendix L[56]). The same process was followed for OS and the same conclusions for standard parametric models and spline (odds and normal) were drawn. Unlike with PFS, a preference was also expressed for the 1- and 2-knot spline (hazard) models, again based on simplicity.

Despite preferences, HRs (at 3, 6, 9, 12, 24 and 36 months) for all parametric models were shown in Table 6 for PFS and Table 11 for OS as well as plots of HR versus time up to 30 months in Appendix L[56].

3.4.2 Method of synthesis of CheckMate-648 and KEYNOTE-590

The company stated that a Bayesian random effects model was considered but chose a fixed effect one for the synthesis. The evidence synthesis model was one developed by Cope et al. 2020[59]. In this method, the survival functions that have been fit for OS and PFS for pembrolizumab with chemotherapy versus chemotherapy are then used as inputs in a multivariate NMA. The distribution-specific parameter estimates (HRs) are transformed to a normally distributed scale with accompanying covariance matrix of the transformed parameters. The NMA model in the second step proposed by Cope et al. 2020[59] is based on one specific parametric distribution that is assumed to apply to all arms of all trials within a network of evidence. It is possible to explore alternative parametric distributions as a series of sensitivity analyses, but alternative distributions cannot be combined within one network of evidence, which would violate the transitivity assumption. The company stated that Appendix L contained "*Common distributions used for the analysis of time-to-event data as well as the corresponding survival, hazard functions, link functions, and transformation to linear prediction*" (p. 71). However, this was not the case. The company has now provided the additional models using alternative distributions (Weibull, Gompertz, log-normal, exponential and gamma.) presented in sub-Appendix C to updated Appendix L[60] (see ERG comment below).

The company stated that *"The result of the application of the methods in Cope et al. are differences in each of the survival function parameters between pembrolizumab with chemotherapy and chemotherapy (both from KEYNOTE-590). These differences on the survival function parameters can be applied to chemotherapy as assessed in CheckMate 648 to obtain PFS (IA) and OS over time for pembrolizumab with chemotherapy relative to chemotherapy, as assessed in CheckMate 648."* (p. 71). In their response to the request for clarification by the ERG[16], they reported that the method "*…results in a posterior distribution of survival time distribution parameters and differences between these parameters. These parameter differences are assumed to represent the relative treatment effect of each pair of treatments, and that by applying the parameter difference to the parameters of a model representing outcomes upon the reference treatment, the parametric model predicting outcomes upon the investigational treatment may be formed*." (page 58)[16].

3.4.3 ITC analysis results

3.4.3.1 Pembrolizumab with chemotherapy versus chemotherapy: patients with PD-L1 CPS ≥10

The PFS results (HRs) of the ITC for KEYNOTE-590 between pembrolizumab with chemotherapy versus chemotherapy (mixed histology) are presented in Table 3.28 and Figures 3.2-3.3, while the OS results (squamous histology) are presented in Table 3.29 and Figures 3.4-3.5. The choice of parametric models was based on the Akaike Information Criterion (AIC) ranking. In the PFS results the standard parametric models show significant differences while the spline-hazard, -odds and -normal models have a mix of significant and not significant differences. Across all timepoints pembrolizumab with chemotherapy improved PFS when compared with chemotherapy. The trends vary over time in different models for both outcomes as shown in Figures 3.15 and 3.16.

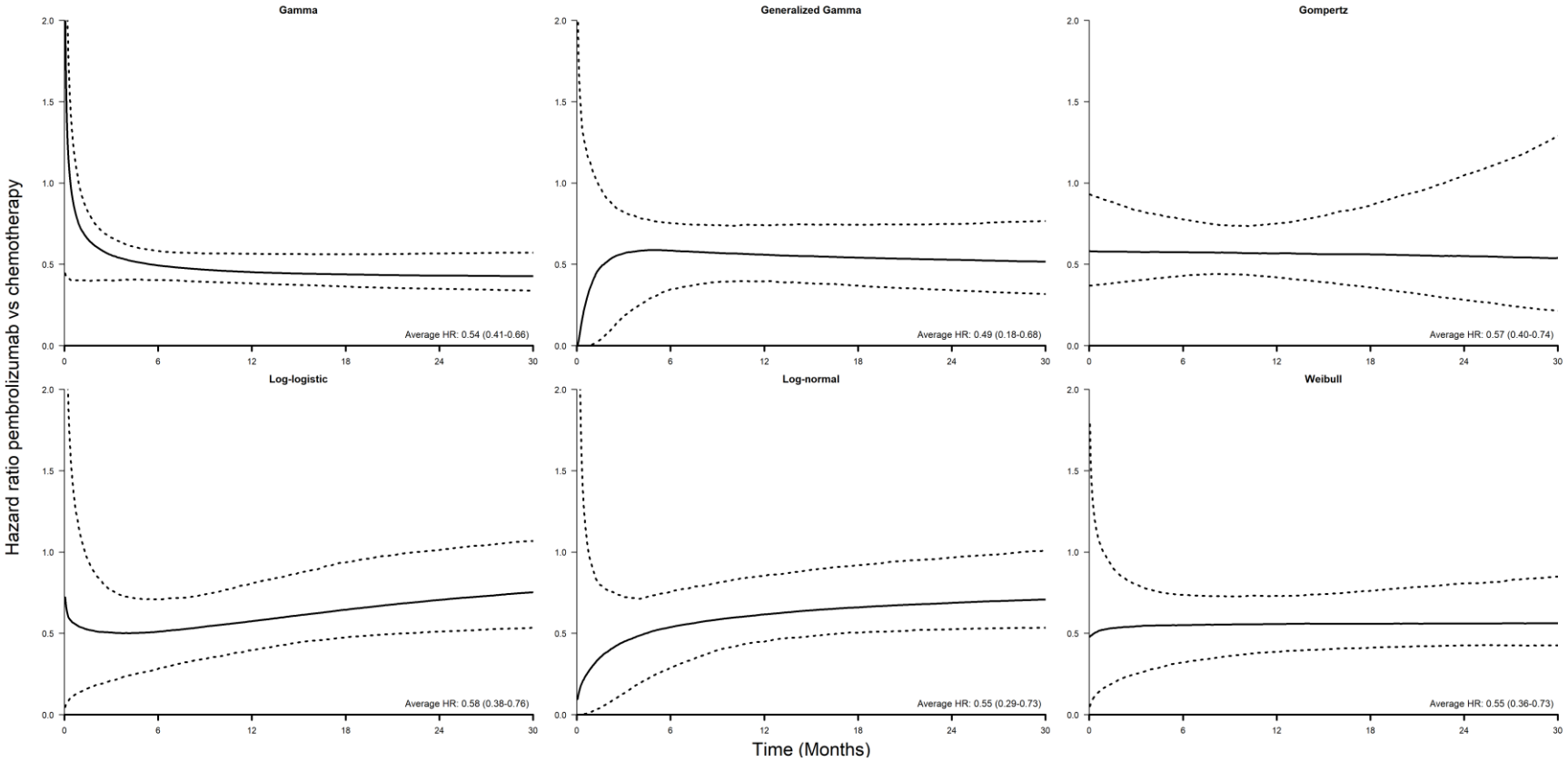
The OS results are somewhat different. Pembrolizumab with chemotherapy again, as in the PFS results, improved OS when compared with chemotherapy across all timepoints and models. On the other hand, both standard and spline parametric model present not significant results. The company states that among the parametric models, only the gamma and generalised gamma models were statistically significant across all time points, but that is not correct. As shown in Table 3.28, the Weibull model also had significant results across all timepoints. The OS HR trends over time vary across models but they are more extreme in the spline models (Figures 3.17 and 3.18).

Table 3.28: ITC PFS results for KEYNOTE-590 patients with PD-L1 ≥10% (CPS) and mixed histology

Model family	Model	HR (95% CrI) for pembrolizumab + chemotherapy versus chemotherapy					
		3 months	6 months	9 months	12 months	24 months	36 months
Standard parametric	Gamma	0.4 (0.56, 0.66)	0.4 (0.49, 0.58)	0.39 (0.47, 0.57)	0.38 (0.45, 0.57)	0.35 (0.43, 0.57)	0.33 (0.42, 0.57)
	Generalised gamma	0.18 (0.57, 0.82)	0.35 (0.59, 0.75)	0.39 (0.57, 0.74)	0.39 (0.56, 0.74)	0.34 (0.53, 0.75)	0.3 (0.51, 0.78)
	Gompertz	0.4 (0.58, 0.83)	0.43 (0.58, 0.78)	0.44 (0.57, 0.74)	0.42 (0.57, 0.75)	0.28 (0.55, 1.05)	0.16 (0.53, 1.63)
	Log-logistic	0.21 (0.5, 0.77)	0.28 (0.51, 0.71)	0.34 (0.54, 0.74)	0.4 (0.58, 0.81)	0.51 (0.71, 1.01)	0.55 (0.79, 1.1)
	Log normal	0.13 (0.45, 0.72)	0.29 (0.54, 0.76)	0.4 (0.58, 0.81)	0.45 (0.62, 0.86)	0.53 (0.69, 0.97)	0.54 (0.72, 1.03)
	Weibull	0.25 (0.54, 0.8)	0.32 (0.55, 0.74)	0.36 (0.55, 0.73)	0.39 (0.56, 0.73)	0.43 (0.56, 0.81)	0.42 (0.57, 0.9)
Spline hazard	1-knot	0.37 (0.59, 0.81)	0.31 (0.53, 0.84)	0.26 (0.48, 0.85)	0.2 (0.44, 0.9)	0.1 (0.38, 1.12)	0.07 (0.35, 1.22)
	2-knot	0.32 (0.58, 1.08)	0.21 (0.43, 0.81)	0.24 (0.48, 0.86)	0.23 (0.53, 1.09)	0.14 (0.66, 2.52)	0.12 (0.71, 3.69)
	3-knot	0.21 (0.61, 1.26)	0.1 (0.4, 0.8)	0.18 (0.43, 1)	0.24 (0.52, 1.07)	0.19 (0.74, 2.97)	0.17 (0.83, 4.2)
Spline odds	1-knot	0.28 (0.59, 1.17)	0.18 (0.41, 0.81)	0.11 (0.34, 0.84)	0.07 (0.31, 0.89)	0.02 (0.31, 1.01)	0.02 (0.33, 1.04)
	2-knot	0.27 (0.57, 1.12)	0.13 (0.4, 0.85)	0.11 (0.42, 0.99)	0.08 (0.45, 1.24)	0.04 (0.55, 1.65)	0.04 (0.6, 1.69)
	3-knot	0.34 (0.73, 1.46)	0.02 (0.21, 0.67)	0.11 (0.37, 0.88)	0.15 (0.56, 1.38)	0.2 (0.92, 2.39)	0.24 (0.98, 2.45)
Spline normal	1-knot	0.3 (0.61, 1.1)	0.21 (0.45, 0.81)	0.12 (0.37, 0.82)	0.08 (0.33, 0.86)	0.03 (0.28, 0.95)	0.02 (0.27, 0.99)
	2-knot	0.25 (0.53, 1)	0.18 (0.44, 0.83)	0.18 (0.47, 0.96)	0.14 (0.49, 1.18)	0.07 (0.55, 1.85)	0.06 (0.56, 2.08)
	3-knot	0.22 (0.67, 1.33)	0.03 (0.27, 0.67)	0.2 (0.38, 0.68)	0.26 (0.52, 0.92)	0.21 (0.78, 2.2)	0.2 (0.85, 2.66)

Based on Table 15 of Document B of the CS[2]
CrI = credible intervals; CPS = combined positive score; CS = company submission; HRs = hazard ratios; ITC = indirect treatment comparison; PFS = progression-free survival: PD-L1 = programmed death ligand-1

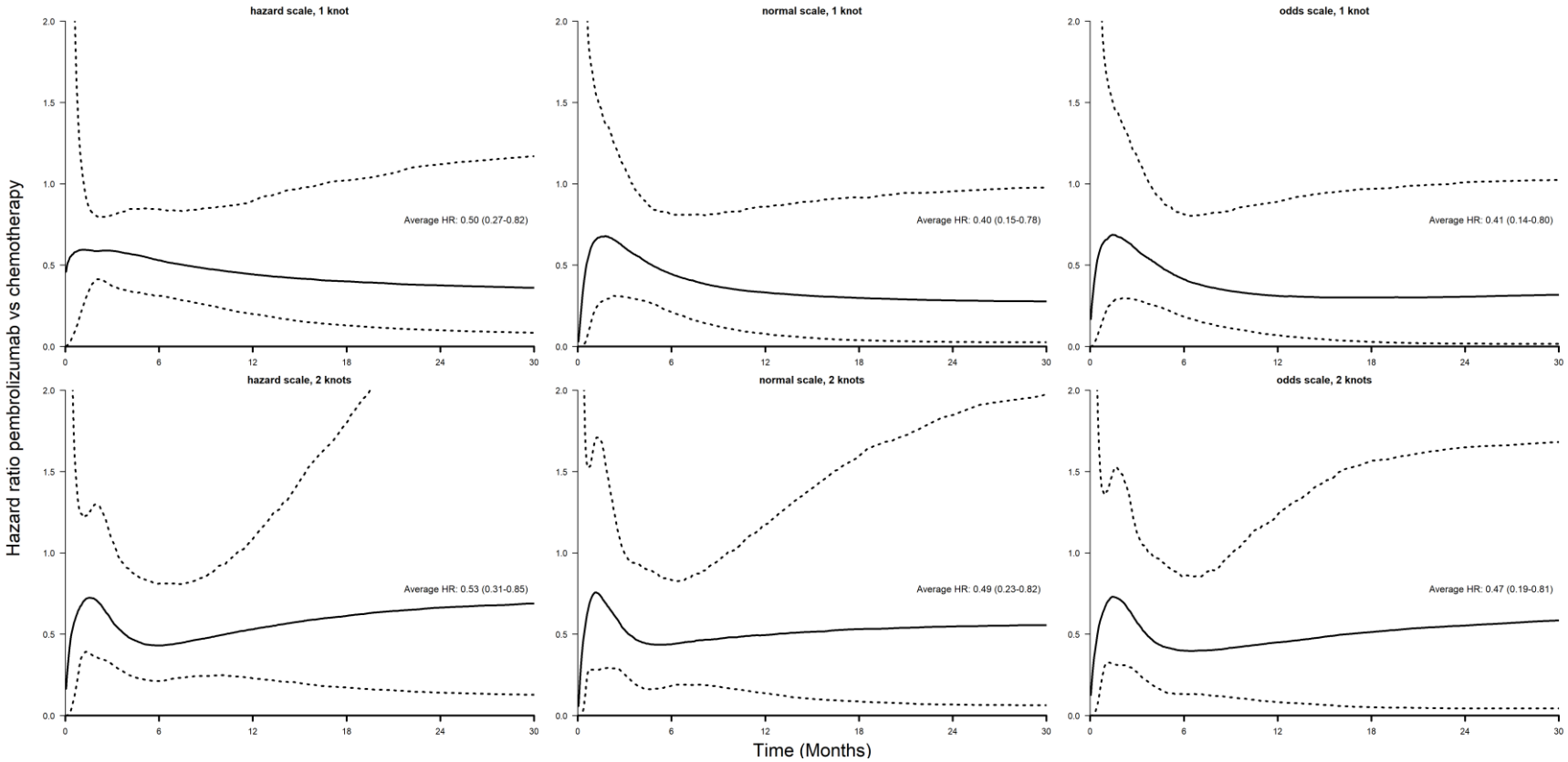
Figure 3.15: Results of the ITC of PFS; standard parametric models, KEYNOTE-590 patients with PD-L1 ≥10% (CPS) and mixed histology



Based on Figure 22 of Document B of the CS[2]

CPS = combined positive score; CS = company submission; HRs = hazard ratios; ITC = indirect treatment comparison; PFS = progression free survival; PD-L1 = programmed death ligand-1

Figure 3.16: Results of the ITC of PFS; spline models, KEYNOTE-590 patients with PD-L1 ≥10% (CPS) and mixed histology



Based on Figure 23 of Document B of the CS[2]

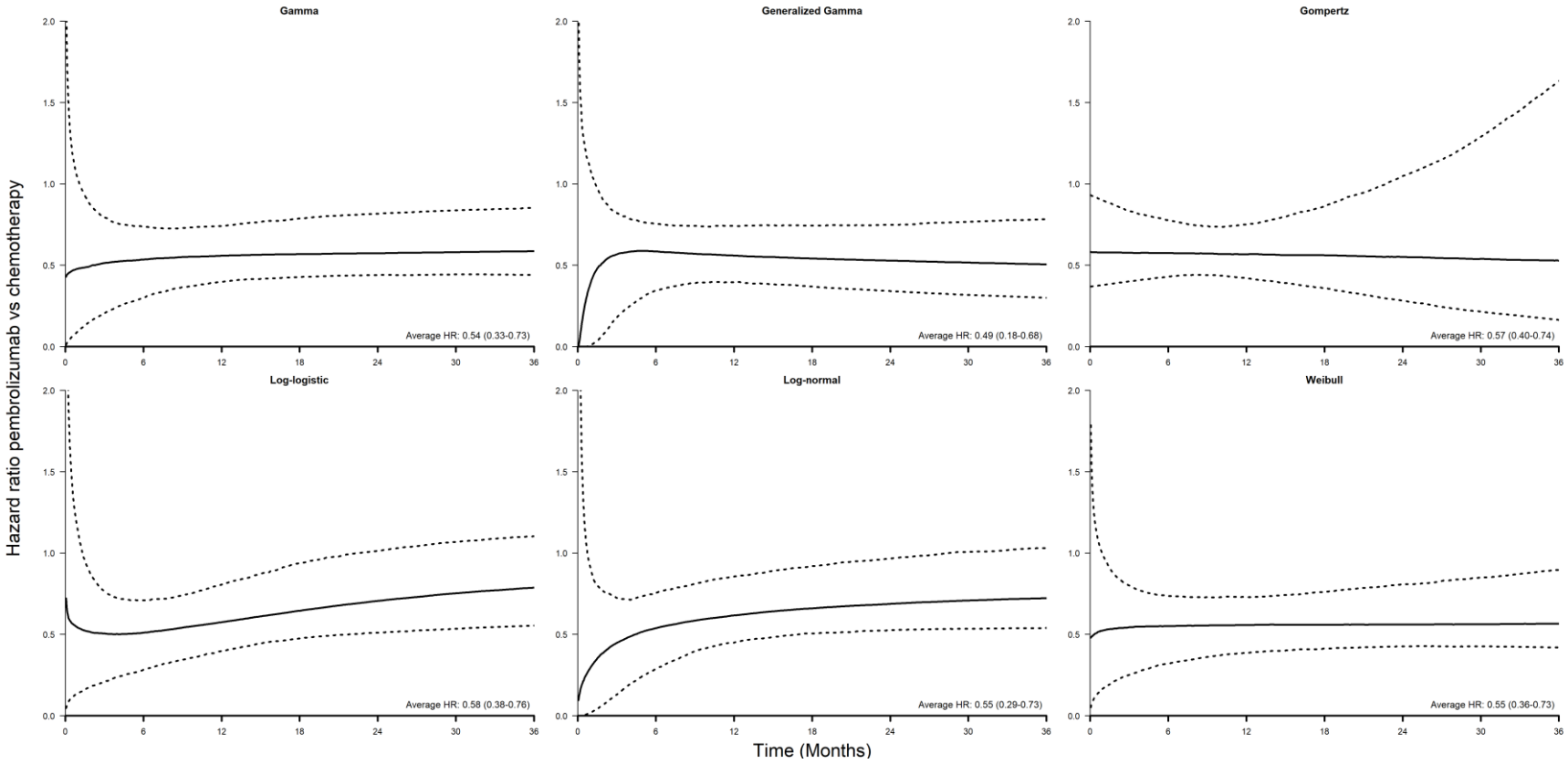
CPS = combined positive score; HRs = hazard ratios; ITC = indirect treatment comparison; PFS = progression free survival: PD-L1 = programmed death ligand-1

Table 3.29: ITC OS results for KEYNOTE-590 patients with PD-L1 ≥10% (CPS) and squamous histology

Model family	Model	HRs (95% CrI) for pembrolizumab + chemotherapy versus chemotherapy					
		3 months	6 months	9 months	12 months	24 months	36 months
Standard parametric	Gamma	0.2 (0.51, 0.79)	0.3 (0.54, 0.74)	0.36 (0.55, 0.73)	0.4 (0.56, 0.74)	0.44 (0.57, 0.82)	0.44 (0.59, 0.85)
	Generalized gamma	0.18 (0.57, 0.82)	0.35 (0.59, 0.75)	0.39 (0.57, 0.74)	0.39 (0.56, 0.74)	0.34 (0.53, 0.75)	0.3 (0.51, 0.78)
	Gompertz	0.4 (0.58, 0.83)	0.43 (0.58, 0.78)	0.44 (0.57, 0.74)	0.42 (0.57, 0.75)	0.28 (0.55, 1.05)	0.16 (0.53, 1.63)
	Log-logistic	0.21 (0.5, 0.77)	0.28 (0.51, 0.71)	0.34 (0.54, 0.74)	0.4 (0.58, 0.81)	0.51 (0.71, 1.01)	0.55 (0.79, 1.1)
	Log normal	0.13 (0.45, 0.72)	0.29 (0.54, 0.76)	0.4 (0.58, 0.81)	0.45 (0.62, 0.86)	0.53 (0.69, 0.97)	0.54 (0.72, 1.03)
	Weibull	0.25 (0.54, 0.8)	0.32 (0.55, 0.74)	0.36 (0.55, 0.73)	0.39 (0.56, 0.73)	0.43 (0.56, 0.81)	0.42 (0.57, 0.9)
Spline hazard	1-knot	0.24 (0.52, 1.13)	0.31 (0.62, 1.19)	0.34 (0.61, 1.08)	0.31 (0.57, 1)	0.08 (0.47, 1.38)	0.05 (0.45, 1.8)
	2-knot	0.22 (0.55, 1.47)	0.3 (0.55, 1)	0.25 (0.54, 1.02)	0.27 (0.55, 1.06)	0.08 (0.66, 2.32)	0.05 (0.69, 3.55)
	3-knot	0.21 (0.52, 1.26)	0.29 (0.61, 1.1)	0.28 (0.56, 0.95)	0.08 (0.48, 1.15)	0.14 (0.75, 2.63)	0.09 (0.84, 4.37)
Spline odds	1-knot	0.24 (0.55, 1.26)	0.28 (0.56, 1.11)	0.24 (0.51, 0.97)	0.18 (0.47, 0.97)	0.05 (0.46, 1.28)	0.05 (0.5, 1.38)
	2-knot	0.18 (0.52, 1.41)	0.24 (0.51, 1.05)	0.2 (0.51, 1.08)	0.22 (0.54, 1.05)	0.05 (0.54, 2.06)	0.03 (0.58, 2.42)
	3-knot	0.29 (0.49, 0.81)	0.35 (0.62, 1.05)	0.26 (0.53, 1)	0.15 (0.42, 0.98)	0.17 (0.67, 1.93)	0.18 (0.75, 2.24)
Spline normal	1-knot	0.2 (0.56, 1.39)	0.28 (0.56, 1.05)	0.27 (0.53, 0.94)	0.2 (0.5, 0.97)	0.1 (0.47, 1.15)	0.08 (0.47, 1.25)
	2-knot	0.15 (0.49, 1.34)	0.27 (0.51, 0.86)	0.25 (0.53, 0.88)	0.35 (0.54, 0.8)	0.07 (0.57, 1.62)	0.04 (0.58, 2.1)
	3-knot	0.27 (0.5, 0.95)	0.23 (0.59, 1.06)	0.25 (0.51, 0.82)	0.08 (0.45, 1.04)	0.13 (0.72, 1.92)	0.09 (0.78, 2.69)

Based on Table 15 of Document B of the CS[2]
CrI = credible intervals; CPS = combined positive score; CS = company submission; HRs = hazard ratios; ITC = indirect treatment comparison; OS = overall survival; PD-L1 = programmed death ligand-1

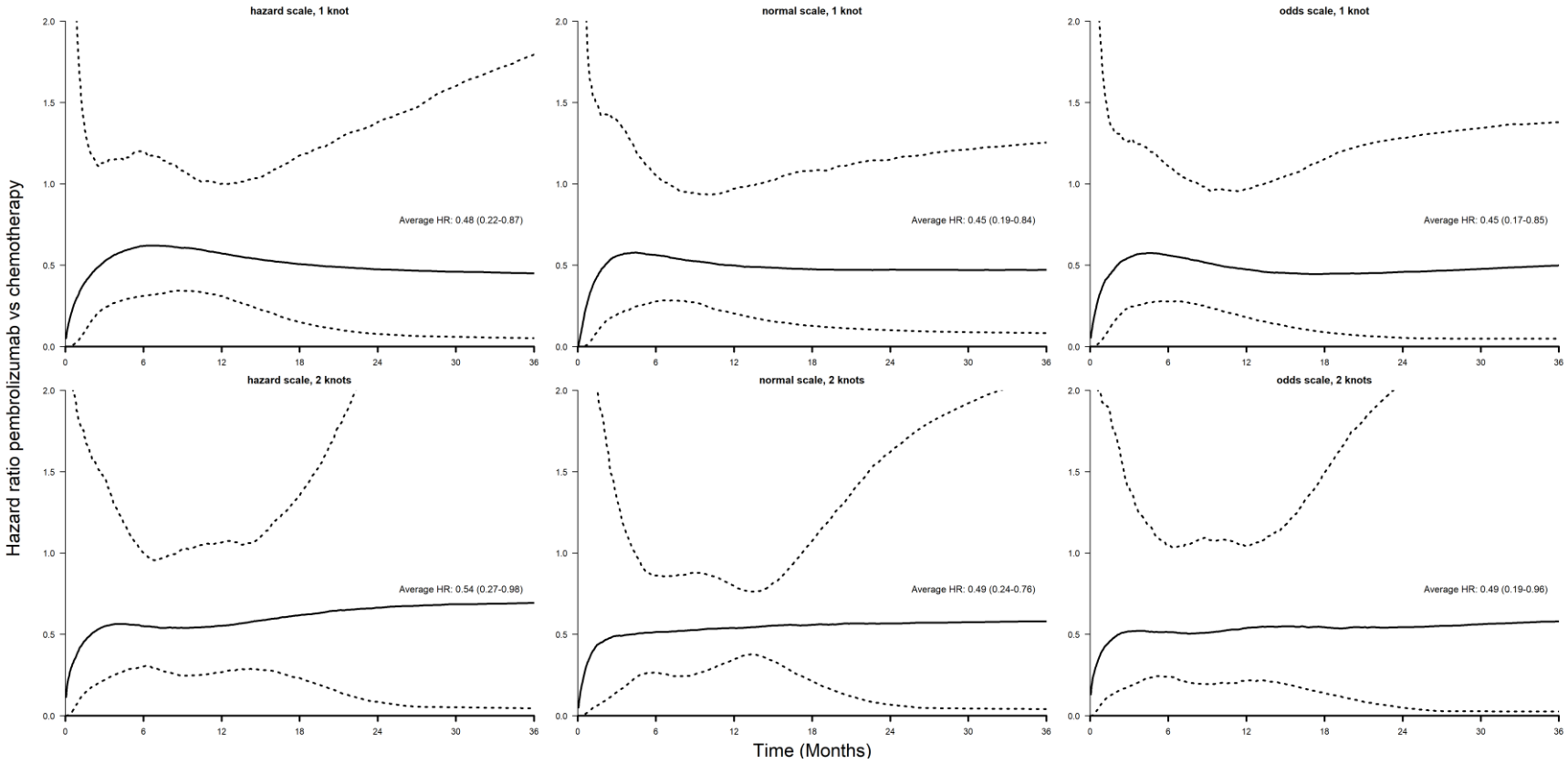
Figure 3.17: Results of the ITC of OS; standard parametric models, KEYNOTE-590 patients with PD-L1 ≥10% (CPS) and squamous histology



Based on Figure 24 of Document B of the CS[2]

CPS = combined positive score; CS = company submission; HRs = hazard ratios; ITC = indirect treatment comparison; OS = overall survival; PD-L1 = programmed death ligand-1

Figure 3.18: Results of the ITC of OS; spline models, KEYNOTE-590 patients with PD-L1 ≥10% (CPS) and squamous histology



Based on Figure 25 of Document B of the CS[2]

CPS = combined positive score; CS = company submission; HRs = hazard ratios; ITC = indirect treatment comparison; OS = overall survival; PD-L1 = programmed death ligand-1

3.4.3.2 Pembrolizumab with chemotherapy vs nivolumab with chemotherapy: patients with PD-L1 CPS ≥10

The results of the ITC for OS of patients treated with pembrolizumab with chemotherapy compared to nivolumab with chemotherapy using the gamma model are presented in Table 3.30. The results illustrate that there are no significant differences between the two interventions across the different time points, indicating that both have a similar effect on OS in the squamous histology population with PD-L1 CPS ≥10 expression. In addition, in point estimates, pembrolizumab with chemotherapy tends to have favourable effects compared to nivolumab with chemotherapy.

Other parametric models have also been fitted, which are now reported in the sub Appendix C of Appendix L[60], showing similar results (see ERG comment below).

Table 3.30: Results of fixed-effects Gamma model for OS, HR over time for patients with PD-L1 CPS ≥10

	HR (95% CrI) for nivolumab + chemotherapy versus comparators at each timepoint (months)									
	3	6	9	12	18	24	30	36	42	48
versus chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
versus pembrolizumab with chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Based on Table 15 of Document B of the CS[2]
[a]estimates based on model extrapolations
CrI = credible intervals; CPS = combined positive score; CS = company submission; HRs = hazard ratios; ITC = indirect treatment comparison; OS = overall survival; PD-L1 = programmed death ligand-1

3.4.3.3 Pembrolizumab with chemotherapy vs chemotherapy: patients with PD-L1 ≥10% CPS and PD-L1 TC ≥1%

An additional ITC analysis was considered including patients with both PD-L1 ≥10% (CPS) as well as PD-L1 ≥1% (TC). The company concluded that such an analysis would not be appropriate because both populations that had PD-L1 TC ≥1% and PD-L1 CPS <10% expressions would have to be removed, which would be an issue for both studies but for different reasons. KEYNOTE-590 does not provide data for the TC PD-L1<1% population, while CheckMate 648 results show diminished efficacy of the drug on TC PD-L1 <1% patients, leading to potential introduction of bias[16]. Nevertheless, the company provided an ITC for CheckMate 648 patients that had both PD-L1 ≥10% CPS as well as PD-L1 ≥1% TC after the request of the ERG. The results are presented in Tables 3.31 and 3.32. Pembrolizumab with chemotherapy improved PFS and OS compared with chemotherapy across all timepoints. Not significant values were reported in only some spline models for all timepoints, except 6 and 9 months regarding PFS HRs, but in some models both parametric and spline model families across all timepoints for OS HRs.

Table 3.31: Results of the ITC of PFS for CheckMate-648 patients with PD-L1 CPS ≥10% and PD-L1 TC ≥1%

Model family	Model	HR (95% CrI) for pembrolizumab + chemotherapy versus chemotherapy					
		3 mths	6 mths	9 mths	12 mths	24 mths	36 mths
Standard parametric	Gamma	[redacted]					
	Generalised gamma	[redacted]					
	Gompertz	[redacted]					
	Log-logistic	[redacted]					
	Log normal	[redacted]					
	Weibull	[redacted]					
Spline hazard	1-knot	[redacted]					
	2-knot	[redacted]					
Spline odds	1-knot	[redacted]					
	2-knot	[redacted]					

Model family	Model	HR (95% CrI) for pembrolizumab + chemotherapy versus chemotherapy					
		3 mths	6 mths	9 mths	12 mths	24 mths	36 mths
Spline normal	1-knot	[redacted]					
	2-knot	[redacted]					

Based on Table 12 of the response to the clarification letter[16]
CPS = combined positive score; CrI = credible intervals; HRs = hazard ratios; ITC = indirect treatment comparison; mths = months; PFS = progression free survival; PD-L1 = programmed death ligand-1

Table 3.32: Results of the ITC of OS for CheckMate-648 patients with PD-L1 CPS ≥10% and PD-L1 TC ≥1%

Model family	Model	HR (95% CrI) for pembrolizumab + chemotherapy versus chemotherapy					
		3 mths	6 mths	9 mths	12 mths	24 mths	36 mths
Standard parametric	Gamma	[redacted]					
	Generalised gamma	[redacted]					
	Gompertz	[redacted]					
	Log-logistic	[redacted]					
	Log normal	[redacted]					
	Weibull	[redacted]					
Spline hazard	1-knot	[redacted]					
	2-knot	[redacted]					
Spline odds	1-knot	[redacted]					
	2-knot	[redacted]					
Spline normal	1-knot	[redacted]					

Model family	Model	HR (95% CrI) for pembrolizumab + chemotherapy versus chemotherapy					
		3 mths	6 mths	9 mths	12 mths	24 mths	36 mths
	2-knot	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Based on Table 12 of the response to the clarification letter[16]
CPS = combined positive score; CrI = credible intervals; HRs = hazard ratios; ITC = indirect treatment comparison; mths = months; OS = overall survival; PD-L1 = programmed death ligand-1

3.4.3.4 Assessment of heterogeneity and uncertainty

The company stated that they did not detect significant between study heterogeneity, enough to prevent the execution of an NMA. Heterogeneity assessment was not executed statistically since fixed-effect models were used in the analyses. The assessment was done only narratively, as described above in the feasibility assessment.

The company acknowledges that there is a number of limitations in the ITC analysis. Each comparison was based on one study which increases uncertainty because the comparison relies on the study populations being the same. Comparability assessment has shown that this is not the case, especially regarding the PD-L1 expression. The ITC limited the analysis to the PD-L1 ≥10% (CPS) population to partly overcome this issue. The sensitivity of the CPS tests used in KEYNOTE-590 is not known, which could introduce further differences in the populations. There are also differences in the proportions of Asian patients included in the two studies, the proportion of patients with metastatic disease and the frequency of chemotherapy administration. The available PFS data from KEYNOTE-590 for the PD-L1 ≥10% (CPS) population refer to the mixed histology population (SCC and adenocarcinoma). The ITC assumed that the data were comparable to the PD-L1 ≥10% (CPS) population of CheckMate 648 who only had SCC histology, which further increased the uncertainly in the analysis. The network had no closed loops, which prevented the assessment of inconsistency and violation of the transitivity assumption.

ERG comment:

- In the updated Appendix L of the CS[57], the company stated that "*KEYNOTE-590 did not allow patients with prior treatment experience while CheckMate 648 allowed patients with prior treatment provided it was completed more than six months prior to trial enrolment resulting in nearly 80% of patients with prior treatment experience in CheckMate 648. It is assumed that these differences do not act as treatment effect modifiers*." This was one of the three assumptions the NMA feasibility assessment was based on. Prior treatment for CheckMate 648 in this setting refers to prior adjuvant, neoadjuvant, or definitive, chemotherapy/radiotherapy/chemoradiotherapy for esophageal squamous cell carcinoma (ESCC)[37].
  Specifically, the primary clinical study report of CheckMate 648 states that "*Prior adjuvant, neoadjuvant, or definitive, chemotherapy/ radiotherapy/ chemoradiotherapy for ESCC was permitted if given as part of curative intent regimen and completed before enrolment. A minimum 24-week recurrence-free period was required after completion of neoadjuvant or adjuvant chemotherapies or after completion of multimodal therapies for locally advanced disease.*"[37](p. 42). Both studies did not allow prior therapy as primary therapy for treatment of advanced or metastatic disease. The company has erroneously reported that KEYNOTE-590[25] did not allow any prior therapy. The study's protocol states in the eligibility criteria that "*Subjects may have received prior neoadjuvant or adjuvant therapy in consideration of following:*

a. *Assessment of disease progression should be confirmed by CT scan. In certain situations, clinical evidence of disease progression such as any new or worsening malignant effusion (documented by ultrasound) and confirmation by pathologic criteria (histology and/or cytology) may be acceptable for assessment.*

b. *Treatment with curative intent, including neoadjuvant/adjuvant treatment, given as chemotherapy or chemoradiotherapy, using standard of care agents or definitive chemoradiation, will count as a line of therapy if disease progression occurs during treatment or within 6 months of cessation of treatment.*

c. *Dose reduction and/or switching of one or more agents due to toxicity/intolerability as deemed clinically appropriate by the investigator will not constitute a new line of therapy*" (p. 45-46)[61]. It is the ERG's opinion that the populations of the two studies are comparable in terms of allowed prior treatments.

- The NMA feasibility assessment is based on the comparison of characteristics of populations that are beyond the scope of the NMA. The available baseline characteristics for KEYNOTE-590 are for patients in the ITT population which includes patients with SCC or adenocarcinoma, located on the oesophagus or gastroesophageal junction, and patients beyond PD-L1 ≥10% (CPS) expression. On the other hand, the company does not specify if the baseline characteristics of CheckMate 648 presented in the NMA feasibility assessment refer to the study's entire population or the PD-L1 ≥10% (CPS) population. A comparison between Table 9 in Document B of the CS[2] and Table 5 in Appendix L of the CS[57] shows that the characteristics of the entire population was probably used. The ERG notes that these comparisons have limitations.
- The results of the feasibility assessment were used to support the feasibility of an NMA between the PD-L1 ≥10% (CPS) populations in CheckMate 648 and KEYNOTE-590. Under the assumption that these population were indeed comparable, the company concluded that "…*there are no differences in treatment effect modifiers between the chemotherapy arms in the CheckMate 648 trial and KEYNOTE-590 and with that, the outcome would be similar to that of a complete network meta-analysis, where randomization is preserved using the principle of transitivity*."[16]. The ERG notes that differences in the features of the trials as well as in patients' characteristics introduce limitations in the ITC results, as outlined in the assessment of heterogeneity and uncertainty section of the report.
- The rejection of the proportional hazards assumption for OS and PFS in the KEYNOTE-590 trial is consistent with the approach taken by both the company and the ERG in TA737[1]. In the response to clarification the company stated that it was incorrectly stated that 3-knot models were *"overly complex"*[16]. They went on to show that there appeared to be no more than two inflexion points in the hazard versus time plots for OS and PFS, thus suggesting that there would be no need for more than 2-knots. Also, they argued that there was little difference in the survival curves between 4- or 5- and 1- to 3- knot models. The ERG concurs with these two conclusions. The ERG did notice that the AIC and Bayesian Information Criterion (BIC) were lowest for the 4- and 5- knot models than the 1- to 3- knot models (see Tables 14 and 15 in the response to clarification) but agree with the company that the difference is so small and visual fit so similar to not require the 4- or 5- knot models.
- In response to clarification the company verified that a random effects model would be inappropriate given the synthesis of only two trials. The company stated that there was an updated version of Appendix L[57]. The ERG has discovered that this document is also labelled as 'NMA report'. It contains none of the evidence in the original Appendix L and described in Section 3.4.1. Instead, it contains a description of a constant HR method of ITC, assuming proportional hazards, as well as the method attributed to Cope et al 2020 referred to in the CS and introduced in Section 3.4.2,

referred to as *"2-step time-varying HR NMA".* The best fit for the survival model for the ITC was stated to have been determined by calculating the sum of the AICs across all CheckMate 648 and KEYNOTE-590 arms, which also include the nivolumab + ipilimumab arm from the CheckMate 648 trial. There is a contradiction between what is reported in the new Appendix L and the original CS in that, whereas the original CS states that the ITC method was that to estimate time varying HRs, Table 6 in the new Appendix L of the CS[57], which is labelled as an overview of analysis scenarios, refers to the constant HR as the primary analysis. The company have clarified in the FAC that the main analysis was the time-varying HR one.[10] Also, the new Appendix L states the following: *"... only the results of the best fitting model for the time-varying HR NMAs are presented in Sections 5.3 and 5.4. The remaining time-varying HR NMAs with alternative parametric distributions are provided in the Appendix in addition to results for the overlap and all-comers analyses."* (page 48). However, Sections 5.3 and 5.4 do not exist, the results apparently being in Sections 5.2.1 and 5.2.2. In these Sections, for the constant HR model, the HR for both OS and PFS for nivolumab + chemotherapy versus pembrolizumab + chemotherapy, is above, although the 95% credible interval (CrI) overlaps 1. For the best fitting (lowest summed AIC) time varying HR model, which is the log-logistic, the HRs at all time points from 3 to 48 months are also above 1 with a 95% CrIs that overlap 1. Survival to all landmark points up to 48 months is also greater for pembrolizumab + chemotherapy.

- The results for time varying HR method for the gamma model are presented in Section B.2.9.2.2. The company were asked why the gamma model was presented to which they replied: *"The fixed effects gamma model was used as an example but is not used in our analysis. This is aligned with other models assessed within the Appendix L (ITC report) of the company submission."* The ERG can verify that the gamma model produces the same pattern of results as the log-logistic i.e., the HRs at all time points from 3 to 48 months are also above 1 with a 95% CrIs that overlap 1, although the point estimates are slightly higher. The ERG can also verify that all other models produce the same pattern, as reported in Appendix C to Appendix L, supplied with the clarification response.
- Given that the index population for this appraisal is PD-L1 TC ≥1%, the ERG requested that an ITC for the PD-L1 ≥10% CPS as well as PD-L1 ≥1% TC be carried out, which the company reported as HRs of pembrolizumab + chemotherapy versus chemotherapy from KEYNOTE-590 in Tables 12 and 13 of the clarification letter response for PFS and OS respectively. The company stated that these results were reproduced from Appendix I of Appendix L. The ERG could not find any appendices in either Appendix L as presented with the clarification response or in the original CS, where it stated that the results were presented in Appendix J, which the ERG has not had sight of despite requesting it. However, in response to another question in the clarification letter the company indicated that another document had been provided, referred to as sub-Appendix A, B, C and I[60], which apparently was referred to as Appendix J in the CS. The ERG can confirm that Tables 12 and 13 correspond to Tables I1 and I3 in this document. Also, in this sub-Appendix and referred to in the new Appendix L are the HR results of the ITC for the PD-L1 ≥10% CPS as well as PD-L1 ≥1% TC, so-called 'overlap' population, which are slightly more favourable for nivolumab + chemotherapy. The point estimates for the best fitting time-varying model for OS, which is again the log-logistic, are all still above 1, but only just and survival is greater to all landmark points, but again only by a small amount. There is a similar trend for the best fitting model for PFS, which his again the log-logistic, with the point estimates for the HRs being less above 1 with one exception at 3 months where it is very slightly below 1. For OS and PFS there is little difference between the log-logistic and the other parametric models, although the point estimates are slightly higher in the others.

- Given the acceptance of failure of the proportional hazards assumption, the ERG considers that the time varying method seems to be the more valid. Given that this requires a single parametric model it also seems reasonable to choose this based on the sum of the AICs, although BICs might have also been used. Also, the effectiveness of nivolumab + chemotherapy versus pembrolizumab + chemotherapy would only be lower with the gamma. Although the most appropriate CheckMate 648 population is the PD-L1 ≥10% CPS as well as PD-L1 ≥1% TC one. However, for comparability the PD- L1 ≥10% CPS from CheckMate 648 as used in the company base case is probably the most appropriate given that the KEYNOTE-590 data is contaminated by some patients with PD-L1 <1%. In conclusion, it was unclear from the CS and the clarification response which ITC method, constant HR or time varying HRs formed the base case, which made this a key issue. The company have since clarified in the FAC that the base case was the time varying method.
- The company has used the expanded ITT PD- L1 ≥10% CPS population for the PFS ITC analyses, based on clinical input from an advisory board and a comparison of HRs between OSCC and adenocarcinoma in the ITT population. This origin of the clinical input was not specifically cited in the updated Appendix L of the CS[57]. The only evidence of an 'advisory board' submitted with the CS has a comment regarding the similarity between squamous histology and adenocarcinoma which is somewhat different. The advisory board report states that "*As OSCC is relatively poorly studied, there are several data gaps in the available evidence. BMS asked the panel whether GC* (Gastric cancer) *data could be considered comparable. The clinicians specified that most GCs are adenocarcinomas, which would not be comparable to OSCC. However, it was accepted that squamous histologies were comparable between tumour locations and therefore, survival outcomes and tumour responses from trials in patients with squamous cell GC could be considered appliable to the OSCC setting*." page 9[55]. Furthermore, the comparison of PFS HRs is of limited use as a judgement cannot be made on the allocation of PD- L1 ≥10% CPS patients in the OSCC or the adenocarcinoma subgroups of the ITT.
- According to both Appendices L[56, 57] the ITC analysis was executed using R (version 3.6.1) and JAGS (version 4.2.0). Regarding R, only the flexsurv package was reported to have been used. The ERG requested that the company would provide further details on any other packages used in the analysis as well as the code and the datasets[62], to which the company responded that "*The packages and code are attached to this response*." (page 41)[16]. The ERG was unable to locate any such evidence in the response and the accompanying files. Therefore, an assessment of the tools used in the analysis was not feasible.
- The ERG inquired why the company did not consider the use of fractional polynomial models for the ITC analysis. The company has responded that they were indeed considered for survival modelling of independent treatment arms. The results for both OS and PFS outcomes were nearly identical to the standard parametric and spline models or "*…or generated models with long tails that were considered clinically implausible for chemotherapy survival.*"[16]. The company has provided the graphical illustrations of the models in their response to request for clarification from the ERG[16].
- The ERG asked the company to elaborate on the Bayesian framework that was used for the NMA analysis to which the company replied "*The RTEs of pembrolizumab + chemotherapy versus chemotherapy were synthesized in a Bayesian framework. For a given parametric survival distribution, the data was the arm-level scale and shape parameters, the likelihood was a multivariate Normal distribution, and the parameters of interest are the relative treatment effects, i.e., scale and shape parameter d's. Parameters have been provided as part of the response to A29b. Normal non-informative prior distributions with a mean of 0 and a variance of 10,000 were used for the relative treatment effect parameters estimated*."(pages 74-75)[16]

3.5 Additional work on clinical effectiveness undertaken by the ERG

Not applicable.

3.6 Conclusions of the clinical effectiveness section

The CS and response to clarification provided sufficient details for the ERG to appraise the literature searches conducted to identify studies on the treatment of unresectable advanced recurrent or metastatic previously untreated OSCC.[2, 16] Searches were conducted in January 2021 and updated in October 2021. Searches were transparent and reproducible, and comprehensive strategies were used. A good range of databases was searched, although additional grey literature resources may have been useful. Overall, the ERG has no major concerns about the literature searches conducted, although separate AEs searches and non-restriction to English language publications may have retrieved additional relevant references.

The study selection criteria for participants, interventions, comparators and outcomes in the clinical effectiveness SLR[12] generally encompassed those in the NICE final scope.[3] Study selection was restricted to English language reports and this may have meant that relevant evidence was missed. Restriction to RCTs may have resulted in some data on AEs being overlooked.

The data extraction process was satisfactory and in line with recommended good practice in systematic reviews.[22]

The process for assessing risk of bias in the included studies was satisfactory overall although there was a discrepancy between information in different parts of the CS (Document B and Appendix E).[2, 12] The CheckMate 648 RCT was assessed twice: with the Cochrane Risk of Bias Tool (original version)[23] as part of a group of RCTs considered relevant to the submission; and in isolation using a checklist adapted from the CRD guidance on undertaking systematic reviews.[17] The review process (e.g., number of reviewers involved) was described and found satisfactory for use of the Cochrane Risk of Bias Tool but no such information was provided for the assessment with the CRD checklist. It was not clear why CheckMate 648 was assessed twice.

The number of studies retrieved, screened and included was unclear from the initial CS[2] but was elucidated by the company's response to the clarification letter.[16] Twelve unique RCTs were identified as being relevant to the SLR of which one (CheckMate 648)[6, 7] served as the main source of evidence and another (KEYNOTE-590)[25] provided comparison data for an ITC.

CheckMate 648 was an international, phase III, open-label RCT comparing the efficacy and safety of NIVO-CHEMO versus CHEMO in adult patients with unresectable advanced, recurrent or metastatic, previously untreated OSCC. The primary outcomes focused on a subgroup of patients with PD-L1 TC ≥1% whilst the secondary outcomes included all randomised patients.[6, 7] The ERG rated CheckMate 648 as being at high risk of performance bias (being open-label) whilst most other aspects of the trial methods were well-conducted. Baseline variables were comparable between the two treatment arms with the exception of age range and sex distribution (younger patients and more males for CHEMO). At the 20-month minimum follow-up, OS was more favourable for NIVO-CHEMO compared with CHEMO in all randomised patients and for the subgroup with PD-L1 TC ≥1%. At the same timepoint, PFS assessed by BICR was more favourable among patients with PD-L1 TC ≥1% but no between-group difference was aparent for all randomised patients. Changes in HRQoL were assessed in both populations using the EQ-5D-3L Utility Index, the EQ-5D-3L VAS, FACT-E, FACT-E ECS, FACT-E GP5 and FACT-G7. Baseline scores were comparable between treatment groups. Interpretation of HRQoL outcomes was hindered by a focus on within-group differences and lack of information to substantiate MID values. More deaths were observed among patients on CHEMO compared with NIVO-CHEMO: [redacted] versus [redacted] for all randomised patients; and [redacted] versus [redacted] for patients with PD-L1 TC ≥1%. All-cause SAEs were lower among those receiving CHEMO compared with NIVO-CHEMO: 43% versus 60% for all randomised patients; and 46% versus 56% for patients with PD-L1

TC ≥1%. All AEs of potential immunological aetiology were higher among patients on NIVO-CHEMO versus CHEMO. Subgroup analysis was presented for OS, only for the entire randomized population. Poor overlap in HR CIs was observed in several subgroups including the PD-L1 TC ≥1% and <1% groups.

The ITC analysis was based on a NMA consisting of only two RTCs CheckMate 648 and KEYNOTE-590; the latter providing data for the intervention of pembrolizumab with chemotherapy over chemotherapy alone. A rationale for the use of KEYNOTE-590 in the ITC analysis was not provided in the CS. However, in the FAC the company stated: "*KEYNOTE 590 was the only trial evaluating pembrolizumab plus chemotherapy in this indication identified by SLR…".*[10] Regarding PD-L1 expression, KEYNOTE-590 only provided CPS ≥10% and not TC ≥1% data and as such this population was used in the NMA for both studies. There was some overlap with the population defined in the scope of the STA (PD-L1 TC ≥1%) but not complete. In addition, KEYNOTE-590 only provided PFS results for population of mixed histology (OSCC and adenocarcinoma) which were used in the NMA.

The NMA feasibility assessment was based on the baseline characteristics of the entire ITT population in KEYNOTE-590 and not the sub-populations used in the NMA. Furthermore, differences between the two studies were identified in study design (open-label versus double-blinded) and methods, for example, the frequency of chemotherapy administration and the use of different tests for PD-L1 expression. Overall, there is limited evidence on the comparability of the two studies.

Fixed-effects models were used in the analysis. The survival functions were first fit for OS and PFS for pembrolizumab with chemotherapy versus chemotherapy and then used as inputs in a multivariate NMA. The NMA models tested a series of parametric distributions, but the final model was based on one that was assumed to apply to all arms within the network. The choice of model was based on an AIC ranking. Although the company reported the rejection of the proportional hazard assumption for OS and PFS in KEYNOTE-590, it was unclear from the CS and clarification letter response whether constant or time varying HRs formed the base case of the analysis. The FAC has clarified that the time varying method was used in the base case.[10] Nevertheless, the OS results of the NMA did not show significant differences between the two double regimens while point estimates favored pembrolizumab with chemotherapy over nivolumab with chemotherapy in the OSCC PD-L1 ≥10% CPS population. An additional analysis including only patients with PD-L1 ≥10% CPS and PD-L1 TC ≥1%, pembrolizumab with chemotherapy improved PFS and OS compared with chemotherapy across all timepoints, with a mix of significant and not significant values in the range of fitted models. Limitations of the NMA were acknowledged by the company regarding study design and population comparability.

4. COST EFFECTIVENESS

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

This Section pertains mainly to the review of cost effectiveness analysis studies. However, the search Section (4.1.1) also contains summaries and critiques of other searches related to cost effectiveness presented in the CS. Therefore, the following Section includes searches for the cost effectiveness analysis review, measurement, and evaluation of health effects as well as for cost and healthcare resource identification, measurement, and valuation.

4.1.1 Searches performed for cost effectiveness section

The following paragraphs contain summaries and critiques of all searches related to cost effectiveness presented in the CS.[2] The CADTH evidence-based checklist for the PRESS, was used to inform this critique.[13, 14] The CS[2] was checked against the STA specification for company/sponsor submission of evidence.[15] The ERG has presented only the major limitations of each search strategy in the report.

Appendix H of the CS details a literature review undertaken to identify economic evaluations to inform an economic model comparing treatments for advanced OC.[63] The strategy was also designed to retrieve relevant studies on cost and resource use. Searches were conducted in April 2021.

A summary of the sources searched is provided in Table 4.1.

Table 4.1: Data sources for economic evaluations (as reported in CS)

Resource	Host/Source	Date Ranges	Dates searched
Electronic databases			
MEDLINE	Ovid	1946-27.04.21	28.04.21
Embase	Ovid	1974-Week 16/2021	28.04.21
Northern Light	Ovid	2019-Week 14/2021	28.04.21
DARE NHS EED HTA	CRD website	to 28.04.21	28.04.21
CRD = Centre for Reviews and Dissemination; CS = company submission; DARE = Database of Abstracts of Reviews of Effects; HTA = Health Technology Assessment database; NHS EED = NHS Economic Evaluation Database			

Appendix I of the CS details a literature review undertaken to identify evidence on HRQoL for patients with advanced OC.[21] Searches were conducted in April 2021.

A summary of the sources searched is provided in Table 4.2.

Table 4.2: Data sources searched for HRQoL studies (as reported in CS)

Resource	Host/Source	Date Ranges	Dates searched
Electronic databases			
MEDLINE	Ovid	1946-23.04.21	27.04.21
Embase	Ovid	1974-Week 16/2021	27.04.21
CENTRAL	Cochrane Library	to 27.04.21	27.04.21
Northern Light	Ovid	2019-Week 14/2021	27.04.21

Resource	Host/Source	Date Ranges	Dates searched
ScHARRHUD	Internet	to 27.04.21	27.04.21
CENTRAL = Cochrane Central Register of Controlled Trials; CS = company submission; HRQoL = health-related quality of life; ScHARRHUD = University of Sheffield School of Health and Related Research Health Utilities Database			

ERG comment:

- Searches were undertaken in April 2021 to identify economic, HRQoL and healthcare resource use/cost data on OSCC. The CS, Appendices H and I, and the company's response to clarification provided sufficient details for the ERG to appraise the literature searches.[2, 16, 21, 63] The company's response to clarification (question A5) states that update searches for the economic SLR were conducted on 22 October 2021.[16] As no strategies for the update were provided in the CS or response to clarification, the ERG has been unable to critique these searches and they are not referenced in the tables above.
- A good range of databases was searched, and searches for named conferences were conducted via Northern Light. No HTA resources or other grey literature sources appear to have been searched.
- Most searches were not limited by publication date. Conference proceedings searches had a 2019-2020 date limit as no relevant conferences had taken place in 2021 at the time of searching.
- Searches were well structured, transparent and reproducible.
- The search strategies contained terms for OC. Study design filters were applied to the searches of MEDLINE and Embase in order to limit the results to relevant economic evaluations, healthcare resource use/cost data and HRQoL/utilities studies. The filters were not referenced, so it was unclear whether they were published and objectively-derived but appear comprehensive. A good range of subject indexing terms (MeSH/EMTREE) and free text was used.

4.1.2 Inclusion/exclusion criteria

In- and exclusion criteria for the review on economic evaluations and HRQoL studies were not clearly presented in the CS but have been expounded on in Appendix H[63] and Appendix I[21] of the CS, respectively.

ERG comment: The HRQoL SLR allowed for studies published in the English language from 2011 and onwards to be included in the review. The patient population was also restricted to *"patients that have progressed after first-line therapy are also of interest."*[21] Eligible economic evaluations were also restricted to only studies published in the English language, in the economic evaluations SLR. Concerning cost and resource studies, the CS states that, *"studies describing costs and healthcare resource use for patients with OSCC were identified systematically, during the cost-effectiveness SLR."*[2] Appendix J[64] also states that, *"Appendix H includes the search strategies which cover the costs and resource use terms."* However, the PICOS outlined in Table 1 of Appendix H[63] only allowed for 'modelled direct/indirect costs' and 'incremental costs' and additionally, the restrictions on study design do not allow for primary research publications. Furthermore, the list of databases searched, and the date when the searches were executed is inconsistent with Appendix H reporting. It is unclear if Document B[2] alludes to a consequent cost effectiveness SLR undertaken following the economic evaluation SLR and HRQoL SLR. The ERG could not identify an SLR in the submission that was powered to identify relevant cost and healthcare resource use studies.

4.1.3 Conclusions of the cost effectiveness review

A total of 23 unique studies describing full economic evaluations of interventions aimed at managing previously untreated advanced or metastatic OSCC were identified. Of these, nine studies were prioritised for extraction, whereas the remaining 14 studies evaluated non-pharmacological interventions, which were not deemed relevant to the objective of this SLR. These studies were summarised in Table 2 of Appendix H.[63]

ERG comment: The CS and response to clarification provided sufficient details for the ERG to appraise the literature searches conducted to identify economic, HRQoL and cost data on advanced OC. Searches were conducted in April 2021.[2, 16] Searches were transparent and reproducible, and comprehensive strategies were used. A good range of databases was searched, although additional grey literature resources may have been useful. Overall, the ERG has no major concerns about the literature searches conducted.

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

4.2.1 NICE reference case checklist

Table 4.3: NICE reference case checklist

Element of HTA	Reference case	ERG comment on CS
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	In line with reference case
Perspective on costs	NHS and PSS	In line with reference case
Type of economic evaluation	Cost utility analysis with fully incremental analysis	In line with reference case
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	In line with reference case
Synthesis of evidence on health effects	Based on systematic review	In line with reference case
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of HRQoL in adults.	In line with reference case
Source of data for measurement of HRQoL	Reported directly by patients and/or carers	In line with reference case
Source of preference data for valuation of changes in HRQoL	Representative sample of the UK population	In line with reference case
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	In line with reference case
Evidence on resource use and costs	Costs should relate to NHS and PSS resources and should be	In line with reference case

Element of HTA	Reference case	ERG comment on CS
	valued using the prices relevant to the NHS and PSS	
Discounting	The same annual rate for both costs and health effects (currently 3.5%)	In line with reference case

CS = company submission; EQ-5D = European Quality of Life-5 dimensions; ERG = Evidence Review Group; HTA = Health Technology Assessment; HRQoL = health-related quality of life; NHS = National Health Service; NICE = National Institute for Health and Care Excellence; PSS = Personal Social Services; QALY = quality-adjusted life year; UK = United Kingdom

4.2.2 Model structure

The analysis was based on a three-health state partitioned survival model (PSM), using a cycle time of 1 week (without half-cycle correction) to accommodate the administration cycles for therapies considered in the model, and capture a realistic minimum time during which the symptoms or responses can change in UK clinical practice. The model was developed in Microsoft Excel 365, and Visual Basic for Applications (VBA) was employed to handle inputs and update results.

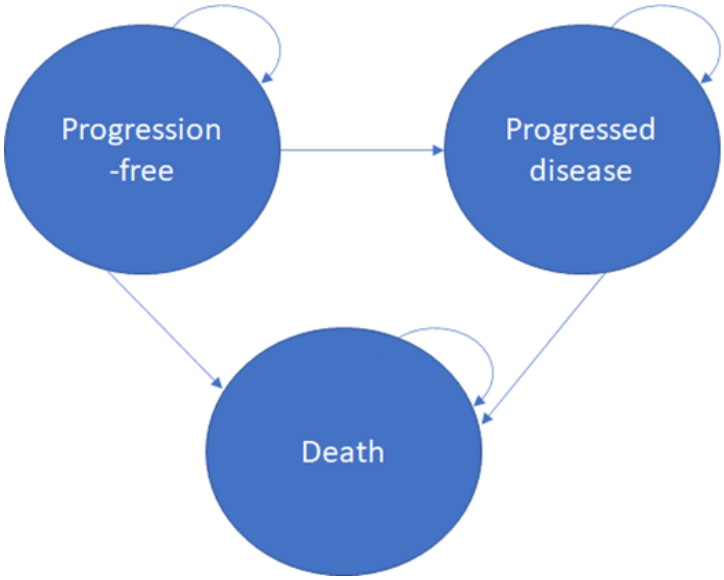
The states in the model are mutually exclusive and represent progression-free disease, post-progression and death, stratified by on-treatment versus discontinued (Figure 4.1). These health states reflect disease severity and determine use of healthcare resources, HRQoL and mortality rates. This model structure is consistent with the approaches adopted in previous published economic evaluations for OC.

Health state occupancy in the PSM is determined by OS and PFS survival curves. Each first-line treatment has unique survival curves which determines the time spent in each health state. Each first-line treatment also has a unique time on treatment (ToT) curve which determines how patients move through lines of treatment. The model assumes that progression phases are consecutive, which means patients are not able to revert to pre-progression from more advanced phases of the disease. Although patients may be able to respond to therapy following progression, patients are still considered to have a higher hazard and an increased resource use.

Patients enter the model and can receive NIVO-CHEMO or a comparator treatment. Following treatment cessation, patients receive a subsequent line of therapy. As a simplifying assumption, it is assumed that patients may not discontinue this final line of therapy.

In each health state, patients accrue treatment costs based on drug acquisition and administration, and health care resources use costs while in that health state, based on disease monitoring and management. Utilities are applied per health state, and a disutility is applied as a one-off utility decrement in the four cycles before death. No treatment waning effect was applied.

Figure 4.1: Model structure



Source: Based on Figure 31 of the CS[2]
CS = company submission

ERG comment: The main concern of the ERG relates to the modelling of subsequent treatments, as set out below.

Modelling of subsequent treatments - As the duration of subsequent treatment was not modelled via a separate health state, the ERG asked the company to justify this aspect of the model structure by providing reference to previous TAs and NHS clinical practice. The company in response to clarification stated that, *"Patient discontinuation of first line treatment is based on treatment specific time on treatment (ToT) curves. Patients then discontinue from second line treatment to no treatment based on a treatment specific cyclical discontinuation rate derived from the average ToT for the corresponding treatment in TA707."*[16] They stated that as PSM is unable to explicitly track individual patients over time through subsequent lines of treatment (i.e. does not explicitly capture how long specific patients have spent within progressed disease, nor how long they have spent on subsequent treatments), *"…discontinuation from second line treatment cannot be tracked or explicitly captured within the model, and so cannot be health state specific. Best supportive care (BSC) would be the relevant a third-line treatment, with patients who discontinue second line treatment receiving BSC until death. The recent NICE appraisal for nivolumab in previously treated oesophageal cancer (TA707) utilised BSC for subsequent therapies in each arm, reflecting clinical practice."*[16] However, they chose not to adopt this as, *"…the inclusion of BSC would create a bias towards the control arm, the arm which provides lower survival… BSC components are palliative as opposed to curative, and therefore are implicitly encompassed by the cost of terminal care as opposed to a subsequent line of treatment."*[16] They reiterated that, *"These assumptions are in keeping with TA737 which only explicitly incorporated one line of subsequent treatments and did not explicitly incorporate further discontinuation to BSC. Additionally, the options of subsequent treatment (further discussed in answer to B2 below) align with UK clinical practice and have been validated by UK clinicians."*[16]

Assumptions regarding subsequent treatment - The ERG asked for further clarification on the mentioned 'budget impact modelling assumptions during NICE TA707 on subsequent therapies' that dictated for patients on the CHEMO arm receiving nivolumab monotherapy. In their response to clarification, the company stated that, *"Budget impact assumptions from TA707 are not publicly*

available. However, within this submission for second line nivolumab in OSCC, nivolumab displaced the majority of taxane use. This indicates that, where nivolumab is applicable at second line, nivolumab would replace the use of taxanes. Therefore, within the chemotherapy arm of the company submission, nivolumab is used as the subsequent treatment (as opposed to single agent taxanes which are used in the NIVO-CHEMO arm).”[16]

Section B.3.2.1.2 of the CS[2] states that: *“Following treatment cessation, patients receive a subsequent line of therapy. As a simplifying assumption, it is assumed that patients may not discontinue this final line of therapy.”* The ERG asked the company to justify why the approach of specifying a maximum number of treatment cycles was not taken, for each subsequent treatment. In response, the company clarified that, *“This sentence within Document B of the Company Submission is erroneous, patients do discontinue second line therapy. As previously described, due to the PSM approach, time in health state cannot be tracked for subsequent health states. Therefore, a treatment cycle-based approach cannot be used. Instead, second line time on treatment is incorporated for subsequent therapies, reflecting the second line nivolumab OSCC submission. Mean time on treatment is used for subsequent therapies. This data is used to calculate and adjust weekly acquisition and administration subsequent treatment costs accordingly. Hence, although number of treatment cycles cannot be incorporated for subsequent therapies, time on treatment for subsequent therapies is captured.”*[16] The ERG therefore can confirm that discontinuation from subsequent therapy is at a per cycle probability based on TA707, which is 0.056 and 0.061 for nivolumab and chemotherapy respectively (see Table 4.24).

Proportion receiving subsequent treatment - As 10% of patients with PD-L1 ≥1% on the chemotherapy arm go on to receive nivolumab monotherapy in the CheckMate 648 trial, the ERG asked for clarification on the modelled 56.7% of patients on the CHEMO arm who go on to receive nivolumab monotherapy. In response the company clarified that, *“The proportions of patients who go on to subsequent therapy ([redacted] in the NIVO-CHEMO arm, [redacted] in the CHEMO arm) are sourced from CheckMate 648 (sourced from the latest DBL, [redacted]).”*[16] In fact, that figure for CHEMO refers to systemic therapy and the equivalent for NIVO-CHEMO was updated from the original CS value of [redacted] (see Table 4.4).[2, 16]

Choice of subsequent treatment - As solely systemic therapies (nivolumab monotherapy or single agent taxanes) were included in the economic model as subsequent treatments, the ERG asked the company to justify why radiotherapy and surgery (second line treatments in the CheckMate 648 trial) were not included, and to discuss the clinical experts’ opinion on subsequent therapies for patients who have progressed on study treatments. Concerning the decision to exclude radiotherapy and surgery as subsequent therapies in the model, in response the company stated that, *“Radiotherapy and surgery were not included as subsequent treatments as these are considered palliative and not curative within the UK and therefore are encompassed implicitly within the cost of terminal care as opposed to a subsequent line of treatment. This approach is in keeping with TA737, where neither radiotherapy nor surgery were incorporated as subsequent treatments.”*[16] Concerning the clinical experts’ opinion on subsequent therapies, the company explained that, *“During an advisory board meeting conducted by BMS, clinicians specialising in the treatment of OSCC in the UK stated that if nivolumab combination therapy was approved as a first-line treatment, then they would not offer an immunotherapy-containing second-line therapy. It was generally believed that a docetaxel or paclitaxel-containing regimen would be offered in the second-line after a nivolumab-containing first-line regimen. This is in-line with current ESMO guidance, which recommends taxanes as monotherapy in second-line therapy for advanced or metastatic OSCC.”*[16] They supplemented their argument with reflections from TA737 stating that, *“During the NICE appraisal of pembrolizumab with chemotherapy for untreated oesophageal and gastro-oesophageal cancer (TA737), it was deemed preferable to give treatment with a PD-L1 inhibitor*

early in the treatment pathway. During the appraisal, clinical experts explained that because pembrolizumab and nivolumab were both PD-L1 inhibitors, it would not be suitable to give nivolumab as a second-line treatment after pembrolizumab with chemotherapy and stated that it was likely that immunotherapy is more effective when used earlier."[16] In conclusion, the company stated that, *"Therefore, according to current NICE guidelines and clinical expert feedback from an advisory board conducted by BMS and from previous NICE appraisals, the second-line therapy for patients with advanced OSCC who have progressed on current first-line treatment, would be nivolumab or taxanes for patients who have received fluoropyrimidine and platinum-based combination therapy first-line. In patients who receive a PD-1 inhibitor first-line, only taxanes would be offered as second-line therapy."*[16]

In TA737, patients on the PEMBRO-CHEMO arm were modelled to receive a range of subsequent therapies including platinum-based cisplatin and oxaliplatin, taxanes (docetaxel and paclitaxel), fluorouracil, and irinotecan, reflecting the subsequent therapy distribution seen with the KEYNOTE-590 trial.[53] As can be seen in Table 4.4, in the CheckMate 648 trial, patients received a range of non- anti-PD1 systemic anticancer therapies with [redacted] of patients on the NIVO-CHEMO arm receiving subsequent taxane therapy and [redacted] of patients on the CHEMO arm receiving subsequent nivolumab therapy. This is in contrast to the [redacted] of patients on the NIVO-CHEMO arm modelled to receive single agent taxane subsequent treatment after progression, and [redacted] of patients on the CHEMO arm modelled to receive nivolumab subsequent treatment. In ATTRACTION-3 (TA707), patients with OC who were refractory or intolerant to combination therapy with fluoropyrimidine and platinum-based drugs (subsequent treatments in CheckMate 648) on the comparator arm were given taxane monotherapy as second line treatment.[45] The ERG remains uncertain on if taxane monotherapy and nivolumab monotherapy can be considered as the only relevant standard second line treatments for NIVO-CHEMO and CHEMO respectively in patients with advanced OSCC in NHS clinical practice.

Table 4.4: Subsequent cancer therapy, CheckMate 648 trial, PD-L1 ≥1%

	Number of subjects (%)	
	NIVO-CHEMO N = 158	**CHEMO N = 157**
Subjects with any subsequent therapy (%)	[redacted]	[redacted]
Subjects who received subsequent systemic therapy (%)	[redacted]	[redacted]
Anti-PD 1	[redacted]	[redacted]
Nivolumab	[redacted]	[redacted]
Other subsequent systemic therapy (%)	[redacted]	[redacted]
Fluorouracil	[redacted]	[redacted]
Cisplatin	[redacted]	[redacted]
Paclitaxel	[redacted]	[redacted]
Docetaxel	[redacted]	[redacted]
Oxaliplatin	[redacted]	[redacted]
Carboplatin	[redacted]	[redacted]
Irinotecan	[redacted]	[redacted]

Adapted from CheckMate 648 CSR Tables S.6.23.1 and S.10.B.1[41]
CHEMO = chemotherapy; CSR = clinical study report; NIVO = nivolumab

Implications of subsequent treatment - The ERG further asked the company to discuss the implications on effectiveness of patients in either NIVO-CHEMO or PEMBRO-CHEMO arms receiving only single-agent taxanes whilst those in the CHEMO arm receive only nivolumab monotherapy as opposed to what was actually received in the CheckMate 648 trial. In response, the company explained that *"According to the latest DBL, [redacted] of the NIVO-CHEMO patients received subsequent therapy, of which, [redacted] received anti-PD(-L)1 and [redacted] received other systemic anticancer therapy with some patients receiving a combination of anti-PD(-L)1 and other systemic therapy. In contrast, [redacted] of the patients in the CHEMO arm received subsequent therapy with [redacted] receiving anti-PD(-L)1 and [redacted] receiving other systemic therapy. In the company's economic model, the NIVO-CHEMO and PEMBRO-CHEMO patients would receive only taxanes as subsequent therapy, whereas the CHEMO patients would receive nivolumab. None of those patients would receive a combination therapy. The approach in the economic model is more conservative as patients in the CHEMO arm would highly benefit from a subsequent treatment with nivolumab. In contrast, there would be a slight underestimation of the subsequent treatment effectiveness in the NIVO-CHEMO and PEMBRO-CHEMO arms if all patients would receive taxanes subsequently and none a PD(-L)1 treatment. It should be noted that all but one patient in the NIVO-CHEMO arm that received subsequent systemic therapy received other systemic anticancer therapy so the implications for this treatment arm should be marginal. The conservative approach chosen overestimates the effectiveness in the CHEMO arm and slightly underestimates the effectiveness in the NIVO-CHEMO and potentially PEMBRO-CHEMO arm leading to a higher ICER."*[16] According to the FAC, the figures of [redacted] and [redacted] should be replaced by [redacted] and [redacted][10] However, the company cites Table S.10.B1, which is for the ITT population, but the source used for Table 4.4, which is for the PD-L1 ≥1%, is Table S.10.B2. The ERG therefore presumes that the correct values according to the latest DBL ([redacted]) remain [redacted] and [redacted]. Given the difference in choice of subsequent therapy between the trial and what might be NHS clinical practice, the ERG requested an analysis of OS and PFS in both arms of CheckMate 648 adjusting for switching to anti-PD-1/PD-L1 therapies by reference to TSD 16.[65] The company were also requested to conduct scenario analyses using adjusted data in the economic model, including variation in the proportion of patients who experience the treatment effect of anti-PD-1/PD-L1 therapies to better reflect NHS clinical practice.[66] However, no adjustment for switching was performed by the company, who appeared to misinterpret the question as being about unplanned pre-progression switching even though it was a sub question of one where subsequent therapy was the subject and to which the company responded accordingly in every other sub question.[16] The ERG strongly disagrees that the company's approach is conservative given that greater use and no use of a PD(-L)1 treatment following CHEMO and NIVO-CHEMO in clinical practice would respectively probably increase and decrease effectiveness in clinical practice compared to CheckMate 648. The ERG in TA737 also made the comment that, although estimating the cost of subsequent therapy according to the what was observed in the KEYNOTE-590 trial did not match clinical practice, it did provide consistency with effectiveness as estimated from the trial, with which this ERG agrees. Therefore, although ideally the effectiveness data should be adjusted to better match clinical practice, as a second best solution the ERG have provided a scenario where costs have been adjusted to better match the trial (see Section 4.2.9.4).

4.2.3 Population

The economic evaluation considered the use of nivolumab in combination with fluoropyrimidine- and platinum-based combination chemotherapy for the anticipated indication for the first line treatment of adult patients with unresectable advanced, recurrent or metastatic OSCC with tumour cell PD-L1 expression ≥1%.

The patient profile of the population subgroup is aligned to that of the PD-L1 positive population subgroup (PD-L1 expression of ≥1%) in the CheckMate 648 trial. The baseline age (standard error (SE)) of the population subgroup is [redacted], with a proportion male (SE) of [redacted]. This included the NIVO-IPI arm of intervention 2 which was not a part of this submission.

The ERG in its clarification letter asked the company to confirm that the demographic parameters used in the economic model (namely age and proportion of males) is indicative of UK clinical practice. In response, the company stated that, *"The median age and proportion male in CheckMate 648 and other oesophageal clinical trials are displayed in Table XX. The Cougar-2 trial, a UK specific clinical trial, has baseline characteristics that lie close to the baseline characteristics of CheckMate 648. Accordingly, the demographic parameters used in the model, taken from CheckMate 648, are representative of the UK clinical practice. Additionally, TA737 utilised data from KEYNOTE-590, with a median age of 62.4 years old, and 83.4% male; both of which are closely aligned to CheckMate 648 data (62.6 years old, 81.8% male). Within the TA737 submission, the ERG agreed that age and proportion male were representative of the target population. This further highlights the generalisability of the demographics used within the company submission herein."*[16]

Table 4.5: Comparison between CheckMate 648 and other OC clinical trials

Trial	Treatment	Age (years)		Proportion male (%)
		Median	Range	
CheckMate 648	NIVO-CHEMO	64	40–85	79%
	CHEMO	62	28–81	83%
Cougar-2	Docetaxel	65	28–84	82%
	Active symptom control	66	36–84	80%
KEYNOTE-590	Pembrolizumab + CHEMO	64	28–94	82%
	CHEMO	62	27–89	85%
CheckMate 649	NIVO-CHEMO	62	54-69	68%
	CHEMO	61	53-68	71%
KEYNOTE-062	Pembrolizumab + CHEMO	62	22–83	76%
	CHEMO	62.5	23–87	72%

Source: Table 26 of CL response[16]
CHEMO = chemotherapy; NIVO = nivolumab

ERG comment: The main concerns of the ERG (as previously highlighted in Sections 2.1 and 2.3 of this report) relate to a) how PD-L1 status has been determined; and b) the population being narrower than the scope.

a) For the CheckMate 648 trial, Appendix M[67] confirmed that, *"Specifically, the subgroup of patients with tumours with a PD-L1 expression of ≥ 1%, determined by the tumour proportion score (TPS)"* whilst PD-L1 status in the PEMBRO-CHEMO arm was assessed using CPS.
b) The population in CheckMate 648 was limited to patients with ECOG PS of 0 or 1.

4.2.4 Interventions and comparators

No explicit statements were made in Section B.3.2.3 of the CS[2] about the dosing regimen for the intervention and comparators considered in the company's model. However, going by the drug acquisition dosing regimen in Tables 46, 47, and 48 of the CS, the intervention considered in the CS was nivolumab + chemotherapy (fluorouracil + cisplatin). Assuming body surface area (BSA) of 1.66 m^2, 240 mg of nivolumab, 1331 mg (6,656 mg over 5 days) of fluorouracil and 133 mg of cisplatin is expected to be received by a dosing regimen of nivolumab 240 mg, on Day 1 Q2W + fluorouracil 800 mg/m^2, on Day 1 through Day 5 Q4W + cisplatin 80 mg/m2, on Day 1 Q4W. This is in line with the CheckMate 648 trial drug dosing regimen.

The comparators considered was doublet chemotherapy (fluorouracil + cisplatin) received in a dosing regimen of fluorouracil 800 mg/m2, on Day 1 through Day 5 Q4W + cisplatin 80 mg/m2, on Day 1 Q4W, which was also in line with the comparator arm of the CheckMate 648 trial. In a scenario analysis, pembrolizumab + chemotherapy (fluorouracil + cisplatin) received in a dosing regimen of pembrolizumab 200 mg, on Day 1 Q3W + fluorouracil 800 mg/m2, on Day 1 through Day 5 Q4W + cisplatin 80 mg/m^2, on Day 1 Q4W, which was reported to be in line with the KEYNOTE-590 trial, was also compared. The company also conducted scenario cost effectiveness analyses with capecitabine + oxaliplatin (XELOX) or fluorouracil and oxaliplatin (FOLFOX) as comparator, but assuming only a difference in cost of treatment and that the intervention would be nivolumab added to the comparator.

The NICE final scope also specified a triplet chemotherapy combination of fluorouracil or capecitabine + cisplatin or oxaliplatin + epirubicin. This comparator was not considered in the model as the company stated that, *"…that epirubicin-based triplet therapy is not commonly used in UK clinical practice. During TA737, the clinical expert stated that triplet therapy is no longer standard of care as it does not provide additional efficacy and increases toxicity. The committee concluded that a dual chemotherapy regimen would be the appropriate comparator for TA737… Hence, assessment of epirubicin-based triplet therapy may not be relevant to decision making for this appraisal."*[2]

ERG comment: The appropriateness of comparators is discussed in Section 2.3: in summary, the ERG considers that PEMBRO-CHEMO and CHEMO, where CHEMO is fluorouracil + cisplatin, are the most appropriate comparators.

4.2.5 Perspective, time horizon and discounting

The analysis is performed from the NHS and PSS perspective. Discount rates of 3.5% are applied to both costs and benefits. The model cycle length is four weeks with a lifetime time horizon (40 years) and a half-cycle correction was not applied.

ERG comment: Nothing of note.

4.2.6 Treatment effectiveness and extrapolation

For the base case comparison, clinical data to inform PFS and OS for NIVO-CHEMO and CHEMO were derived from CheckMate 648 with extrapolation of survival data from the study undertaken with reference to the guidance from the NICE Decision Support Unit (DSU) and Bagust and Beale (2014), the full description being reported in Appendix N.[68, 69] PFS was based on the BICR-assessment. All analyses in Appendix N were based on the PD-L1 ≥1% CheckMate 648 ([redacted] DBL population. For the comparison with PEMBRO-CHEMO, the company conducted a scenario analysis arguing that it is still not SoC (see Section 2.3) and they used what they described as an "HR approach".[2] Efficacy

inputs were derived from the ITC, described in Section 3.3 and 3.4, were applied to the CHEMO survival curves to generate the PEMBRO-CHEMO survival curves.

4.2.6.1 NIVO-CHEMO versus CHEMO: OS and PFS

The proportional hazard assumption was deemed to be violated due to log cumulative hazard curves not being parallel, as shown in Figure 34 of the CS.[2] The company also identified that, whilst the smoothed OS hazard for NIVO-CHEMO seemed to slightly increase toward the end of the follow-up period, that for CHEMO seemed to be decreasing and, depending on the method, toward general population mortality, as shown in Appendix N:[69] this was cited as the reason for using a semi-parametric approach with K-M data until 6.9 months for both OS and PFS. In Appendix N it was stated that the "…benefit [of NIVO-CHEMO versus CHEMO] was [redacted] at 12 months, 18 months and 24 months, but was [redacted] at 36 months due to [redacted] hazards after 24 months for the CHEMO arm and [redacted] hazards after 24 months for the NIVO-CHEMO arm.[69] OS outcomes for CHEMO remained at approximately [redacted] at 3.5 years, which may not be clinically plausible." (page 12) An explanation did seem to be provided: "This may be related to the use of subsequent treatment in the chemotherapy and may also be driven by outcomes in responder patients." (page 13) This prompted a scenario that was 'response-stratified' i.e., extrapolating OS and PFS per response status (complete response, partial response, stable disease, progressive disease/unable to determine, complete response/partial response). Elsewhere in Appendix N it was also stated that: "When characterising CheckMate 648, it was found the response was strongly associated with long-term outcomes (as described in Section 4.1.1.1)." (page 22) In Section 4.1.1.1 an analysis of response status at any time point and in those who survived to 24 months was presented in Table 12. Also, the proportion of patients in each arm who received subsequent therapy was presented (see Table 4.4), which showed it was higher for CHEMO, much of which was explained by increased receipt of anti-PD1/PD-L1 therapy: 15% versus 6%. The precise time of 6.9 months for the semi-parametric approach of 6.9 months appeared to have only the following justification: *"A number of potential cut points were considered, avoiding assessment windows due to the rapid change in hazard near the model start time implied by these periods. As a compromise between maximisation of data for use in extrapolation and removal of the largest hazard discontinuities, a time of 6.9 months was chosen. This timepoint avoids the sharp change in hazard observed in the first six months for NIVO-CHEMO (Figure 18) and CHEMO (Figure 16)."* (p. 50)[69]

OS, Parametric approach

For OS, Landmark analyses, at 1, 2, 3 and 3.5 years, only of OS for PD-L1 ≥1% (TC) from CheckMate 648 were presented in Table 4.6 as well as Tables 22 to 26 of Appendix N by responder status.[69] Parametric survival curves are shown in Figure 4.2. Those up to 10 years for parametric extrapolations of OS were presented in Tables 4.7 and 4.8.

Table 4.6: CheckMate 648 Landmark OS in patients with TC PD-L1 ≥1% from [redacted]

OS	1 year	2 years	3 years	3.5 years
CHEMO (N=157)	[redacted]	[redacted]	[redacted]	[redacted]
NIVO-CHEMO (N=158)	[redacted]	[redacted]	[redacted]	[redacted]

Based on Table 18, Appendix N
CHEMO = chemotherapy; CI = confidence interval; NIVO = nivolumab; OS = overall survival; PD-L1 = programmed death ligand 1; TC = tumour cell

Figure 4.2: CheckMate 648 OS patients with TC PD-L1 ≥1% from [redacted] - OS, CHEMO: Standard statistical models overlaid upon K-M

Based on Figure 20, Appendix N
AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; K-M = Kaplan-Meier; OS = overall survival; PD-L1 = programmed death ligand 1; TC = tumour cell

Figure 4.3: CheckMate 648 OS patients with TC PD-L1 ≥1% from [redacted] - OS, NIVO-CHEMO: Standard statistical models overlaid upon K-M

Based on Figure 22, Appendix N
AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; K-M = Kaplan-Meier; OS = overall survival; PD-L1 = programmed death ligand 1; TC = tumour cell

Table 4.7: CheckMate 648 Landmark OS for parametric models of patients with TC PD-L1 ≥1% from [redacted]: CHEMO

Survival Model	Exp.	Weibull	Gompertz	L.Logistic	Lognormal	Gen.Gamma
1 year	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
2 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
3 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
5 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
10 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Based on Table 20, Appendix N
CHEMO = chemotherapy; CI = confidence interval; Exp = exponential; Gen.Gamma = generalised gamma; L.Logistic = log-logistic; OS = overall survival; PD-L1 = programmed death ligand 1

Table 4.8: CheckMate 648 Landmark OS for parametric models of patients with TC PD-L1 ≥1% from [redacted]: NIVO-CHEMO

Survival Model	Exp.	Weibull	Gompertz	L.Logistic	Lognormal	Gen.Gamma
1 year	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
2 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
3 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
5 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
10 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Based on Table 21, Appendix N
CHEMO = chemotherapy; CI = confidence interval; DBL = database lock; Exp = exponential; Gen.Gamma = generalised gamma; L.Logistic = log-logistic; NIVO = nivolumab; OS = overall survival; PD-L1 = programmed death ligand 1; TC – tumour cell

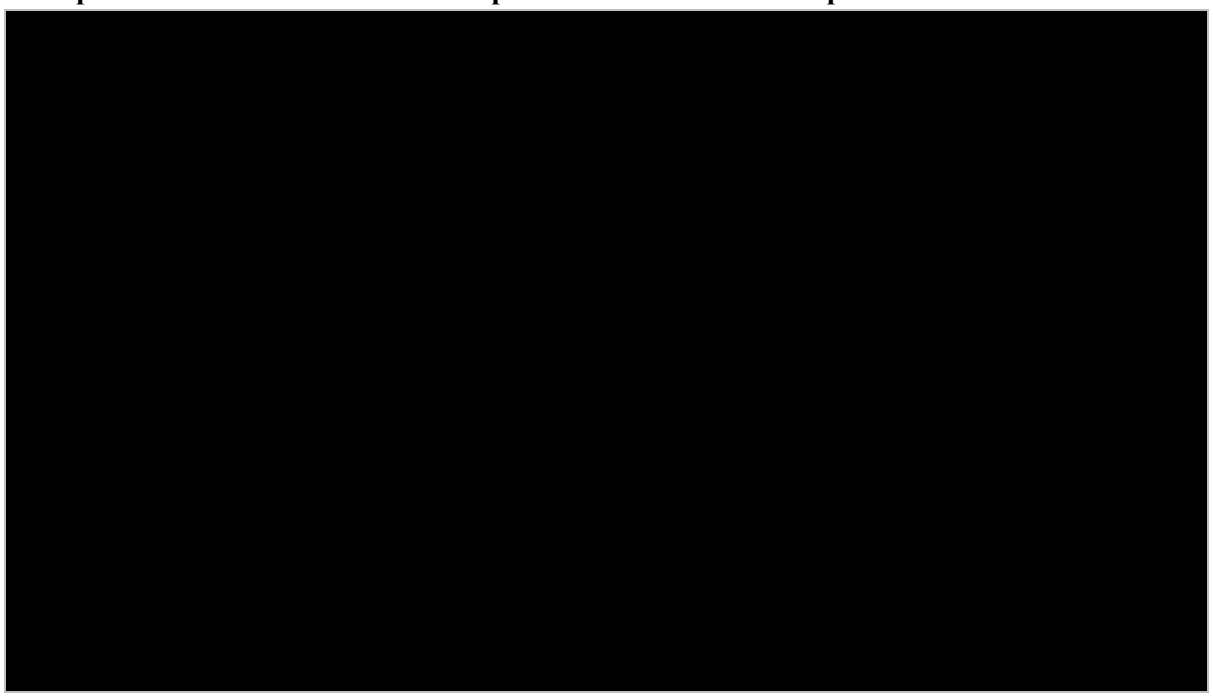
For OS, for the NIVO-CHEMO and CHEMO arms, the lognormal and log-logistic respectively were chosen. For CHEMO this was based on the best statistical fit (AIC and BIC) and 'face validity'. For NIVO-CHEMO, this was based on the Weibull and generalised gamma being ruled out because of poor

visual fit for the first 6 months despite having the best statistical fit (AIC and BIC) and "overemphasising events in the tail of the data where uncertainty is greatest" (page 48, Appendix N) respectively.

OS, semi-parametric approach

OS parametric curves and measures of statistical fit are shown in Figures 4.4 and 4.5.

Figure 4.4: CheckMate 648 in patients with TC PD-L1 ≥1% from [redacted], CHEMO: Semi-parametric OS models overlaid upon K-M – 6.9 months cut point



Based on Figure 24, Appendix N

AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; DBL = database lock; K-M = Kaplan-Meier; OS = overall survival; PD-L1 = programmed death ligand 1; TC = tumour cell

Figure 4.5: CheckMate 648 in patients with TC PD-L1 ≥1% from [redacted], NIVO-CHEMO: Semi-parametric OS models overlaid upon K-M – 6.9 months cut point

Based on Figure 25, Appendix N

AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; DBL = database lock; K-M = Kaplan-Meier; OS = overall survival; PD-L1 = programmed death ligand 1; TC = tumour cell

For OS, for both NIVO-CHEMO and CHEMO arm, the lognormal was chosen. For CHEMO this was based on Gompertz not having a plausible "survival extrapolation" even though it provided the best statistical fit. For NIVO-CHEMO, this was based on the mean survival for generalised gamma, Weibull and exponential being too low in comparison to the restricted mean survival in CheckMate 648, which was 17.90 months.[69] No landmark analyses were provided.

PFS, parametric approach

Landmark analyses of PFS are shown in Table 4.9.

Table 4.9: CheckMate 648 Landmark PFS (BICR) survival in patients with TC PD-L1 ≥1% from [redacted]

PFS (BICR)	1 year	2 years	3 years	3.5 years
CHEMO (N=157)	[redacted]	[redacted]	[redacted]	[redacted]
NIVO-CHEMO (N=158)	[redacted]	[redacted]	[redacted]	[redacted]

Based on Table 30, Appendix N
BICR = blinded independent central review; CHEMO = chemotherapy; DBL = database lock; NIVO = nivolumab; PFS = progression-free survival; PD-L1 = programmed death ligand 1

PFS parametric curves and measures of statistical fit are shown in Figures 4.6 and 4.7.

Figure 4.6: CheckMate 648 PD-L1 ≥1% from [redacted] - BICR-assessed PFS, CHEMO: Standard statistical models overlaid upon K-M

Based on Figure 71, Appendix N
AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; BICR = blinded independent central review; CHEMO = chemotherapy; DBL = database lock; K-M = Kaplan-Meier; PFS = progression-free survival; PD-L1 = programmed death ligand 1

Figure 4.7: CheckMate 648 PD-L1 ≥1% from October 2021 DBL - BICR-assessed PFS, NIVO-CHEMO: Standard statistical models overlaid upon K-M

Based on Figure 73, Appendix N

AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; BICR = blinded independent central review; CHEMO = chemotherapy; DBL = database lock; K-M = Kaplan-Meier; NIVO = nivolumab; PFS = progression-free survival; PD-L1 = programmed death ligand 1

For PFS, the company selected the lognormal for both arms due to best statistical fit, although the company stated that none provided a good visual fit.

PFS, semi-parametric approach

PFS curves and measures of statistical fit are shown in Figures 4.8 and 4.9.

Figure 4.8: CheckMate 648 PD-L1 ≥1% [redacted], CHEMO: Semi-parametric PFS (BICR) models overlaid upon K-M – 6.9 months cut point

AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; BICR = blinded independent central review; CHEMO = chemotherapy; DBL = database lock; K-M = Kaplan-Meier; PFS = progression-free survival; PD-L1 = programmed death ligand 1

Figure 4.9: CheckMate 648 PD-L1 ≥1% [redacted], NIVO-CHEMO: Semi-parametric OS models overlaid upon K-M – 6.9 months cut point

AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; BICR = blinded independent central review; CHEMO = chemotherapy; DBL = database lock; K-M = Kaplan-Meier; NIVO = nivolumab; OS = overall survival; PD-L1 = programmed death ligand 1

For PFS, the Weibull and generalised gamma were chosen for CHEMO and NIVO-CHEMO respectively. The former was chosen according to best statistical fit by first ruling out the Gompertz, log-logistic and lognormal because of *"implausibly long mean PFS".*(page 93, Appendix N)[69] The latter was chosen according to best statistical fit after ruling out the Weibull because of "poor visual fit" and the Gompertz and log-logistic "mean PFS outcomes that could not be considered plausible." No external validation was reported, the only justification being for the NIVO-CHEMO arm: *"There are no other studies with which to validate the results for extrapolation of the NIVO-CHEMO arm other than the informing trial, CheckMate 648."* (page 125, CS)[2]

ERG comment: There seemed to be a lack of justification for choosing 6.9 months as the cut-point for the semi-parametric approach. Therefore, in the clarification letter, the ERG requested a justification for why precisely 6.9 months was chosen and, given the apparent inflexion point of about 6 months in the smoothed hazard plot for chemotherapy, to provide scenario analyses for later cut-points, including 12 months and 20 months (minimum follow-up in the trial). The company were also asked to provide a cost effectiveness model that permits choice of cut-point for OS as well as all fully parametric and semi-parametric models for both OS and PFS. In response, the company did not provide any further justification, but simply quoted Appendix N.[16] However, they did provide plots of semi-parametric PFS and OS curves with a 12-month cut-point and the functionality in the model as requested by the ERG. Spline based models were not considered by the company.[70]

Given the clear inflexion point in the smoothed hazard OS plot for the CHEMO arm in CheckMate 648 with a decreasing hazard thereafter, which was not observed in the NIVO-CHEMO arm, the ERG agrees with the company that subsequent therapy might provide at least part of the explanation, in particular the increased receipt of anti-PD1/PD-L1 therapy in the CHEMO arm. However, it was not clear to the ERG why a response stratified approach would be an appropriate reaction to this observation. The ERG therefore requested an analysis of OS and PFS in both arms of CheckMate 648 adjusting for switching to anti-PD-1/PD-L1 therapies by reference to TSD 16.[65] The company were also requested to conduct scenario analyses using adjusted data in the economic model, including variation in the proportion of patients who experience the treatment effect of anti-PD-1/PD-L1 therapies to better reflect NHS clinical practice.[66] However, no adjustment for switching was performed by the company (see Section 4.2.2).[16]

For OS, considering fully parametric models, the choice of log-logistic for CHEMO appears appropriate given that it did have the best statistical fit (AIC and BIC) and no worse visual fit. The ERG would also add that landmark analysis reveals a reasonable fit to the trial data up to 3 years. The choice of lognormal for NIVO-CHEMO seemed to make less sense given that, after the Weibull, the log-logistic provides the best statistical fit and seems to provide a better visual fit, not too dissimilar to the generalised gamma. However, there seemed to be no reference to the landmark analyses, which show that the log-logistic not only provides the best statistical fit, but also a good approximation to survival in the trial at 1, 2 and 3 years. The generalised gamma also seems to provide a reasonable fit according to the landmark analyses. In fact, the model was programmed to default to the log-logistic using the fully parametric approach.

For OS, considering the semi-parametric approach, the lognormal for CHEMO seems appropriate given that it provides the best statistical fit and reasonable visual fit. However, the ERG questions the choice of lognormal for NIVO-CHEMO: several models provide both a better statistical and apparently better visual fit, including the Weibull, which has the next highest mean survival after the lognormal. No landmark analyses were provided by the company for any piecewise model, despite a request in the clarification letter: instead the survival curves were reproduced.[16] Although the ERG preferred the Weibull for NIVO-CHEMO in the semi-parametric approach, it was noted that implementing the

Weibull for NIVO-CHEMO with the lognormal for CHEMO caused the transition probability (TP) ratio calculated in the model to exceed 1, implying that the former was less effective for much of the time horizon. This was regarded as implausible by the ERG and mitigated by setting the HR effectively to 1 by replacing the TP for nivolumab with the one for CHEMO at every point the HR exceeded 1 (see Section 6).

The ERG notes that there are no plots of HR over time between nivolumab + chemotherapy and any comparator. Therefore, the company were asked to provide plots of HRs over time from the smoothed hazards from the K-M data for NIVO-CHEMO versus all comparators, including CHEMO and PEMBRO-CHEMO, as well as HR plots over time for all extrapolations (parametric and semi-parametric). The company were asked to discuss the validity of the choice of most appropriate extrapolation in the context of the comparison with the HRs from the smoothed hazards. In response the company provided the HRs for each extrapolation for NIVO-CHEMO and CHEMO, which did show that for some parametric functions, the HR continued to remain below 1 and in some cases was decreasing i.e., the treatment effect was increasing. The ERG consider that this is clinically questionable. The company's reasoning behind not applying a treatment waning effect (HR reaching 1 at some point) was that *"there is now long-term evidence of a robust and durable treatment effect lasting beyond discontinuation for immunotherapies."*[2] However, given the possibility of some extrapolations producing HRs that are clinically questionable and that might not fit well what is observed in the trial, the ERG implemented a treatment waning function in the Microsoft Excel model, which works by increasing the TPs calculated from the OS Markov trace using a linear relationship between the ratio of TPs (NIVO-CHEMO/CHEM) and number of weeks such that there is no change at a starting point and then a ratio of 1 at a finishing point. The starting and finishing point can be set by the user and 5 and 7 years were considered for the ERG base case, consistent with the choice of the ERG in TA737 and which was considered acceptable by the committee.[1] However, the observation that survival at 3.5 years was so similar between the two arms (see Table 4.6) and the apparent deceasing hazard with CHEMO, notwithstanding the company's argument that this seems implausible, the ERG prefers a starting and finishing point of 2.5 and 4 years respectively for the ERG base case. Despite a request by the ERG, landmark analyses were not provided for any of the extrapolations: instead only the survival plots were reproduced and no reference to external validation was made in the clarification response.[16] However, survival of at least what is observed for CHEMO in the CheckMate 648 trial is supported by comparability between 2-year survival of 12% in CheckMate 648 and 13.2%, 14.8% and 14% in the three trials cited by the company in Appendix N as evidence of outcomes with chemotherapy, by Chau et al 2009 (adenocarcinoma), Shankaran et al 2021 (adenocarcinoma) and Lyu et al 2018 (OSCC) respectively.[69, 71-73] In fact, although the company suggests that survival might be better with adenocarcinoma, 2-year survival was comparable in the Lyu et al 2018 study where all patients had OSCC.

In terms of PFS, the ERG agree with the company's choice of lognormal for the parametric models and the opinion of poor visual fit at least for CHEMO. The semi-parametric models did provide a better visual fit of course up to 6.9 months, although beyond this it is not convincing. Nevertheless, the choice of Weibull for CHEMO seems reasonable. On the other hand, it is unclear why the Weibull was considered to have a poor visual fit and so the ERG prefers this to the generalised gamma for NIVO-CHEMO given its better statistical fit (see Section 6).

Given the huge number of potential scenarios and lack of justification, choice of method or model for extrapolation of survival, both PFS and OS, for chemotherapy and nivolumab + chemotherapy, is a key issue. Also, given the lack of justification for the cut-point or the semi-parametric approach, and the lack of consideration of spline-based models, the ERG prefers the parametric approach with the log-

logistic for OS and lognormal for PFS for both arms and treatment waning between 2.5 and 4 years for OS. It should be noted that the 12-month cut point analyses might also have had some merit, especially for PFS given the relatively poor fit, but the model would not run these analyses, displaying the following message: "Please ensure all input profiles are appropriately specified".

4.2.6.2 PEMBRO-CHEMO OS and PFS

For details on the ITC used to estimate OS and PFS see Sections 3.3 and 3.4.

ERG comment: For comparison with PEMBRO-CHEMO, as stated in Section 3.4, the ERG considers that the method used by the company in its base case, at least as reported in the CS, for this comparison i.e., time varying HRs (as opposed to fixed HR) based on the PD- L1 ≥10% CPS from CheckMate 648 is probably the most appropriate. In Appendix L it was stated that the best fitting model for the ITC was the log-logistic, although HRs for the gamma were presented (see Section 3.4). However, although many other parametric functions were considered for the ITC (see Section 3.4), none of these could be implemented in the model. On the other hand, it seems that no matter which one is chosen, PEMBRO-CHEMO would be more effective. The ERG also requested landmark analyses, including for the PD-L1 ≥1% (TC) and PD-L1 ≥10% (CPS) population for pembrolizumab for KEYNOTE-590 as well as NIVO-CHEMO and CHEMO and for each parametric or semi-parametric extrapolation as well as external validation for CHEMO. In response to clarification the company provided some of these analysis for the K-M data as shown in Tables 4.10 and 4.11.

Table 4.10: PFS for landmark analysis for PEMBRO + CHEMO relative to CHEMO

PFS	1 year	2 years	3 years	3.5 years
CHEMO KEYNOTE-590	8.3%	3.4%	-	-
PEMBRO + CHEMO KEYNOTE-590	29.9%	15.3%	-	-
CHEMO CheckMate 648	[redacted]	[redacted]	[redacted]	[redacted]
PEMBRO + CHEMO CheckMate 648	[redacted]	[redacted]	[redacted]	[redacted]

Based on Table 18, clarification response.
CHEMO = chemotherapy; PEMBRO = pembrolizumab; PFS = progression free survival

Table 4.11: OS for landmark analysis for PEMBRO + CHEMO relative to CHEMO

	CheckMate 648				KEYNOTE-590	
	PD-L1 ≥1% (TC)		PD-L1 ≥10% (CPS)		PD-L1 ≥10% (CPS)	
Time (Months)	CHEMO	NIVO + CHEMO	CHEMO	NIVO + CHEMO	CHEMO	PEMBRO + CHEMO
6	[redacted]	[redacted]	[redacted]	[redacted]		
12	[redacted]	[redacted]	[redacted]	[redacted]	37.7%	54.5%
18	[redacted]	[redacted]	[redacted]	[redacted]		
24	[redacted]	[redacted]	[redacted]	[redacted]	17.0%	31.9%
30	[redacted]	[redacted]	[redacted]	[redacted]		

36	[redacted]	[redacted]	[redacted]	[redacted]		

Based on Table 20, clarification response.
CHEMO = chemotherapy; CPS = combined positive score; NIVO = nivolumab; OS = overall survival; PEMBRO = pembrolizumab; PD-L1 = programme death ligand 1; TC = tumour cell

The ERG are able to report values estimated in the model of 55.6% versus 53.9% and [redacted] for PEMBRO-CHEMO versus NIVO-CHEMO for OS at 12 and 24 months respectively. However, one would not expect a perfect correspondence between the values for PEMBRO-CHEMO in Table 4.11 and those from the model given the ITC method that uses the treatment effect (PEMBRO-CHEMO versus CHEMO) from KEYNOTE-590. However, although not clear in the CS, the comparison with PEMBRO-CHEMO still assumes the same base case survival extrapolation for NIVO-CHEMO i.e., semi-parametric with 6.9-month cut-off and generalised gamma for PFS and lognormal for OS. The model also states that a semi-parametric model was used for PEMBRO-CHEMO. In response to clarification, the company stated that OS and PFS for PEMBRO-CHEMO were estimated in the model by applying the HRs from the ITC to the CHEMO survival estimates using the parametric function that was used in the semi-parametric CHEMO survival model i.e., Weibull for PFS and lognormal for OS. This contrasts with the ITC, which was supposed to have used a single parametric model to estimate the HRs between NIVO-CHEMO, PEMBRO-CHEMO and PEMBRO-CHEMO, one for OS and another for PFS. The ERG also noted that the OS curves for NIVO-CHEMO and PEMBRO-CHEMO cross in the model, which is inconsistent with the HRs between them always being greater than 1 up to 48 months (see Section 3.4). The ERG therefore implemented some functionality in the model to permit the estimation PEMBRO-CHEMO OS in a way that is more consistent, by applying these HRs to the NIVO-CHEMO OS using the method as described in the clarification response to questions A32:

1. Evaluate accumulated hazard to each model timestep using the relationship $H(t) = -\ln\big(S(t)\big)$
2. Calculate mean hazard experienced during timestep as $h(t) = H(t) - \big(H(t-1)\big)$
3. Multiply this hazard by the reported point HR at nearest time ≤ t for t ≤ 36 months, else hold at HR for t = 36 months

This then prevents the curve crossing, but also allows the combination of any survival extrapolation for NIVO-CHEMO, including the company or ERG base case with HRs based on any survival function, as reported in Appendix C. However, the ERG has presented the results for only one combination in Section 6 because, as expected given that all HRs are a little above 1, there is little difference in outcomes. Nevertheless, some uncertainty in precisely how survival is estimated for PEMBRO-CHEMO means that this is a key issue.

4.2.6.3 All-cause mortality

On the basis that "individual randomised into clinical trials are likely to be slightly younger and healthier than the overall oesophageal cancer patient populations" (page 126, CS), general population lifetable values were combined with the OS extrapolation multiplicatively. The company acknowledged that this would imply double counting, but that it would have a negligible effect on cost effectiveness because: *"…effect applies equally to all comparators".* (p. 126, CS)[2]

ERG comment: The ERG is unclear why mortality was increased beyond what was observed in the trial or its extrapolation without any comparison between the trial and actual clinical practice either in terms of patient characteristics or actual survival. report. This is therefore a key issue. The ERG therefore chose to remove the additional mortality by setting the lifetable values to zero in the ERG base case.

4.2.6.4 Time to treatment discontinuation

To estimate time to treatment discontinuation, the ToT K-M data of CheckMate 648 for NIVO-CHEMO and CHEMO and the PEMBRO-CHEMO arm of KEYNOTE-590 with a mean ToT of 33.67 weeks.

ERG comment: The company stated that in CheckMate 648 there was a stopping rule at 24 months so that patients with a complete response would not receive subsequent treatment.[2] However, examination of the ToT curve for NIVO-CHEMO shows that three patients were still at risk at 24 months and remained so up to 48 months. The effect of this on effectiveness is probably small but needs some clarification by the company.

4.2.7 Adverse events

The 10 most frequently occurring treatment-related Grade 3/4 SAEs were included in the economic model as reductions in total utility. Table 4.12 details the proportion of the cohort receiving the costs and utility decrements associated with that AE.

ERG comment: In its clarification letter, the ERG asked the company to confirm it the source of Grade 3/4 treatment-related AEs type and frequency is the nivolumab + chemotherapy arm in the CheckMate 648 trial. In response, the company stated that, *"The company confirms that the source of grade 3-4 adverse event incidence within the economic model (for NIVO-CHEMO and CHEMO arms) was CheckMate 648. For the comparison with PEMBRO-CHEMO, adverse event incidence was sourced from KEYNOTE-590."*[16]

Table 4.12: CheckMate 648 Grade 3–4 treatment-related SAEs rates

Adverse Event	NIVO-CHEMO (n = 310)			CHEMO (n = 304)			PEMBRO-CHEMO (n = 370)		
	n	%	SE	n	%	SE	n	%	SE
Source	**CheckMate 648**						**KEYNOTE-590**		
Total patients with an event	[redacted]						47	12.70 %	1.92%
Vomiting	[redacted]						9	2.43%	0.80%
Hyponatraemia	[redacted]						0	0.00%	0.00%
Pneumonitis	[redacted]						12	3.24%	0.92%
Hepatic function abnormal	[redacted]						0	0.00%	0.00%
Adrenal insufficiency	[redacted]						0	0.00%	0.00%
Acute kidney injury	[redacted]						11	2.97%	0.88%
Colitis	[redacted]						0	0.00%	0.00%
Nausea	[redacted]						0	0.00%	0.00%
Dehydration	[redacted]						6	1.62%	0.66%
Febrile neutropenia	[redacted]						9	2.43%	0.80%

Based on Table 38 of the CS[2]
CHEMO = chemotherapy; CS = company submission; NIVO = nivolumab; PEMBRO = pembrolizumab; SAEs = serious adverse events; SE = standard error

4.2.8 Health-related quality of life

4.2.8.1 Health-related quality of life data identified in the review

According to the CS, the SLR, which is reported in detail in Appendix I, identified six studies reporting utility values consistent with the NICE reference case i.e., as measured by the EQ-5D-3L.[21] Out of these, the company considered that none were in the correct population i.e., first-line advanced ESCC, three being in advanced ESCC, but not first line.

ERG comment: The ERG considers that the eligibility criteria, which were reported in Table 1 of Appendix I, were appropriate to obtain utility values suitable for the scope.[21] The methods of study selection and extraction were also appropriate. However, values obtained do seem to lack applicability to the scope and the model structure given the discrepancy in population and that no values were reported by health state i.e., pre- or post-progression.

4.2.8.2 Health state utility values

The company estimated utility values per health state, pre- and post-progression from CheckMate 648 in the base case, both for the ITT population and PD-L1 ≥1%, although the former were used in the model.[2] A so-called 'end-of-life' decrement was also applied to the per state values for the last four cycles before death. The company stated that this *"…represents the deterioration of the condition, and thus the reduction in quality of life, in the time prior to death for a patient with OC"* (p. 137) A summary of all utility values used in the cost effectiveness analysis is provided in Table 4.13.

An additional analysis according to time-to-death (TTD) was also performed, the results being presented in Table 40 of the CS.[2]

Table 4.13: Health state utility and AE utility decrement values

Health state	Utility value, mean (SE)	Reference	Justification
Pre-Progression	[redacted]	CheckMate 648	Not explicitly stated but appeared to be only appropriate values.
Post-Progression	[redacted]		
End-of-life utility decrement	[redacted]		
Vomiting	0.048 (0.016)	Nafees 2008	None provided.
Hyponatraemia	0 (0)	TA484	
Pneumonitis	0.037 (0.004)	TA578	
Hepatic function abnormal	0.119 (0.012)	Assumption	
Adrenal insufficiency	0.119 (0.012)	Assumption	
Acute kidney injury	0.048 (0.016)	Assumption	
Colitis	0.047 (0.005)	Assumption	
Nausea	0.048 (0.016)	Nafees 2008	
Dehydration	0.119 (0.012)	Assumption	
Febrile neutropenia	0.09 (0.016)	Nafees 2008	

Based on Tables 41 and 42, CS[2]
AE = adverse event; CS = company submission; SE = standard error

ERG comment: A concern of the ERG is that all the health state values, as well as the TTD ones were all higher for chemotherapy than nivolumab + chemotherapy, albeit by a very small amount, but the company chose the treatment-independent ones. Also, the PD-L1 ≥1% were not used in the model. It is also unclear why the end-of-life decrement was applied. The ERG would recommend that a regression analysis be performed with all three clinically relevant covariates i.e., health state, treatment, and TTD. The TA737 committee concluded that it preferred the progression-based utilities for decision making because the values were more plausible, although it considered that any of the TTD, progression-based and interaction (combining TTD and progression status) approaches could be appropriate.[1]

It is also unclear how the AE disutility values were chosen. No mention was made of AE disutilities in the FAD for TA737, but method that was used was to add Grade 3/4 AE status in the regression model used to estimate the health state values.[1]

Uncertainty is which utility values are the most appropriate mean that this is a key issue.

4.2.9 Resources and costs

The following cost categories were included in the analysis: drug acquisition and administration costs, subsequent treatment costs, monitoring and disease management costs, terminal care costs, and AEs costs.

Unit prices were mostly based on the NHS reference costs, British National Formulary (BNF) 2020, and the Department of Health Drugs and pharmaceutical electronic market information tool (eMIT) - NHS reference costs 2015/16 and eMIT 2021.[2]

In each health state, patients accrued treatment costs based on drug acquisition and administration, and health care resources use costs while in that health state, were based on disease monitoring and management.[2]

4.2.9.1 Resource use and costs data identified in the review

Although the CS[2] states that, *"studies describing costs and healthcare resource use for patients with OSCC were identified systematically, during the cost-effectiveness SLR,"* none of the nine unique studies identified by the cost effectiveness SLR (Appendix H) were cost studies.[63]

The CS[2] further states that, *"Costs have been sourced from the relevant UK literature and NHS reference costs. Where values for standard errors are not available, a default value of 20% of the mean has been used."* It remains unclear if any of the 'relevant UK literature' used to source costs, were identified by the review.

4.2.9.2 Drug acquisition and administration costs

The costs of each therapy were applied with each cycle where treatment was continued, and this included drug procurement and administration costs. Tables 46, 47, and 48 in the CS[2] detailed drug acquisition and administration unit costs for each treatment arm- CHEMO, NIVO-CHEMO, and PEMBRO-CHEMO, respectively. Table 4.14 details the acquisition costs for drugs used in all treatment arms. The ERG in its clarification letter requested for a list of administration costs applied in the model for the intervention, comparator and subsequent treatments, with columns for resource, type of administration, NHS reference code, setting and unit cost. This information has been summarised in Table 4.15.

Table 4.14: Drug acquisition costs

Drug	Formulation	Acquisition cost	Source
Capecitabine	150 mg tablets pack size 60	£4.43	eMIT database[74]
	300 mg tablets pack size 60	£7.77	
	500 mg tablets pack size 120	£26.30	
Cisplatin	100 mg/100 ml solution for infusion vials	£8.73	eMIT database[74]
	50 mg/50 ml solution for infusion vials	£5.38	
Fluorouracil	1 g/20 ml (5%) solution for infusion vial	£2.35	eMIT database[74]
	2.5 g/100 ml (2.5%) solution for infusion vial	£3.79	
	2.5 g/50 ml (5%) solution for infusion vial	£4.01	
	500 mg/10 ml (5%) solution for infusion vial	£1.77	
	5 g/100 ml (5%) solution for infusion vials	£8.58	
Nivolumab*	240 mg/24 ml concentrate for solution for infusion vial	£2,633	BNF 2020[75]
Pembrolizumab*	100 mg/4 ml concentrate for solution for infusion vial	£2,630	BNF 2020[75]

*Patient Access Schemes available
Adapted from Table 45 of CS[2]
BNF = British National Formulary; CS = company submission; eMIT = electronic market information tool

Table 4.15: Drug administration costs

Subsequent treatment	Administration type	NHS reference cost	Setting	Unit cost	Note
CHEMO and PEMBRO-CHEMO arms					
Pembrolizumab	-	£0	-	-	Pembrolizumab admin cost is captured in fluorouracil admin costs as treatments are always given in combination.
Cisplatin	-	£0	-	-	Cisplatin admin costs are captured in fluorouracil admin costs as treatments are always given in combination.
Fluorouracil	Deliver complex chemotherapy, including prolonged infusion treatment, at first attendance	SB14Z[76]	Day case and reg day/night	£431.72	

Subsequent treatment	Administration type	NHS reference cost	Setting	Unit cost	Note
NIVO-CHEMO arm					
Nivolumab	Deliver simple parenteral chemotherapy at first attendance	SB12Z[76]	Weighted average of day case and reg day/night, Outpatient and other	£284.05	Admin cost for nivolumab is required every time the treatment is not given in combination with CHEMO. Nivolumab admin cost is captured in fluorouracil admin cost when given in combination.
Cisplatin	-	£0	-	-	Cisplatin admin cost is captured in fluorouracil admin costs as treatments are always given in combination.
Fluorouracil	Deliver complex chemotherapy, including prolonged infusion treatment, at first attendance	SB14Z[76]	Day case and reg day/night	£431.72	-
Subsequent treatment arm					
Nivolumab	Deliver simple parenteral chemotherapy at first attendance	SB12Z[76]	Weighted average of day case and reg day/night, outpatient and other	£284.05	-
Taxane: docetaxel					
Taxane: paclitaxel					
Adapted from Tables 33, 37 and 38 of clarification letter response[16] CHEMO = chemotherapy; NHS = National Health Service; NIVO = nivolumab; PEMBRO = pembrolizumab					

ERG comment: The ERG in its clarification letter asked the company for clarification on if the initial cycle cost of administration per model cycle applied in the economic model is the same for all subsequent cycles. In their response[16] the company stated that, *"For primary treatments, the 'initial administration costs' are applied every time the treatment is administered (i.e. at initial and subsequent treatment cycles) … However, due to the inability of a PSM to track individual patients through subsequent lines of treatment, an average cyclical cost has been used for subsequent treatments. This average cyclical cost takes into account both the treatment costs and the administration cost over the treatment cycle, which is applied to every patient receiving the subsequent treatment in a modelled cycle (1-week) … using docetaxel as an example, the treatment and administration cost is required every second week. Therefore, to create an average cyclical docetaxel cost, docetaxel's treatment and administration cost must be summed and divided by the treatment cycle (2-weeks)."*

The company[16] also clarified that *"PD-L1 test cost is required upon treatment initiation for NIVO-CHEMO and PEMBRO-CHEMO. This PD-L1 test cost is £42.61 and is applied in the first cycle these treatments are given."* It is unclear if this cost is for TPS PD-L1 testing or CPS testing. Although PD-

L1 testing is not currently routine for OCs in NHS clinical practice, it is expected that when adopted, this test will be conducted along with other histological tests.[53] Thus, the ERG considers this one-off cost application to immunotherapy treatment arms to be appropriate.

The ERG noted that there was an error in the model, which was the inclusion of a figure of 50 in the cell F417 in the Data Library tab given that his appears to be the value of a discount on the price of pembrolizumab (referred to as PAS Discount in the model). The company base case presented in the CS depends on the value being 50. This was corrected by setting the value to 0 (see Section 6). The PAS discount for pembrolizumab supplied to the ERG was applied in the confidential appendix.

Treatment costs (with Patient Access Scheme)

In the company's response to clarification, the Patient Access Scheme (PAS) drug acquisition cost for Nivolumab 240 mg/24 ml concentrate for solution for infusion vial was provided: £2,633.00 (PAS cost: [redacted]).

4.2.9.3 Dose intensity (correcting for missed doses)

Treatment modifiers were applied to the acquisition and administration costs, accounting for missed or delayed doses during the CheckMate 648 trial. The costs applied and their sources/assumptions have been summarised in Table 4.16.

Table 4.16: Treatment modifiers

Treatment		Treatment modifier	Source
Nivolumab in combination with chemotherapy (fluorouracil + cisplatin)	Nivolumab	[redacted]	CheckMate 648 trial[77]
	Fluorouracil	[redacted]	
	Cisplatin	[redacted]	
Chemotherapy (fluorouracil + cisplatin)	Fluorouracil	[redacted]	
	Cisplatin	[redacted]	
Pembrolizumab in combination with chemotherapy (fluorouracil + cisplatin)	Pembrolizumab	[redacted]	Assumed equivalent to NIVO-CHEMO due to lack of data
	Fluorouracil	[redacted]	
	Cisplatin	[redacted]	

Adapted from Table 49 of CS[2]
CS = company submission; CHEMO = chemotherapy; NIVO = nivolumab

ERG comment: The ERG requested for clarification on certain issues surrounding the application of a treatment modifier in the economic model. Those questions and the company's response[16] to them have been summed below:

- Clarification on if the treatment modifier is equal to the number of occasions where a dose was delayed divided by total number of doses administered as was stated in the CS: "*The treatment modifier is one minus the number of doses delayed divided by the total number of doses received."*
- Clarification on the assumptions of dose intensity for the PEMBRO + CHEMO arm: "*There was no data to inform the treatment modifier for the pembrolizumab plus chemotherapy arm (the treatment modifier for pembrolizumab plus chemotherapy was redacted in TA737). Accordingly, the treatment modifier for pembrolizumab plus chemotherapy had to be assumed. One plausible assumption was to assign a treatment modifier of 1. Under this assumption, no pembrolizumab plus chemotherapy doses are delayed; all doses expected to be received will be received. However, this creates an artificially high treatment cost for the pembrolizumab plus chemotherapy arm, biasing the model results towards the treatment arm. The conservative assumption that the pembrolizumab*

plus chemotherapy arm has the same treatment modifier as nivolumab plus chemotherapy has been employed to avoid this bias. That being, pembrolizumab, fluorouracil and cisplatin have equivalent treatment modifiers to nivolumab, fluorouracil and cisplatin, respectively."

- Clarification on if the cost of each drug was adjusted by multiplying by the modifier: "*The cost of each treatment, including both the acquisition cost and the administration cost of treatment, is adjusted by multiplying by the treatment modifier."*
- Justification on the precise amount of cost reduction and how the precise amount of drug delivered in the trial would compare with NHS clinical practice: "*The reduction in doses given and the associated cost is used to match the proportion of patients missing doses for various reasons, for example including co-morbidities, adverse events, patient non-compliance, appointment cancellations. These can be considered reflective of clinical practice and is aligned to the SmPC recommendations on managing adverse events."*

The ERG considers that it is unclear how cost needs to be reduced due to delayed dosing, as opposed to missed doses. The factor that the company use seems to imply that there should be zero cost for both missed and delayed. However, this would appear to be an underestimate of the cost. It is also not clear that assuming the same value for PEMBRO-CHEMO is conservative: relative dose intensity (RDI) was estimated in TA737 and so it would be better to use these values, although the ones for TA737 are redacted.[1] The ERG examined the CSR for information on missed or delayed dosing or RDI and discovered a table (6.1-1) that presented the distribution of patients across RDI (percentage) ranges.[37] The ERG has used these data, assuming that the percentage of patients for each range can be assumed to apply to the midpoint of the range e.g., for nivolumab in the NIVO-CHEMO arm it is assumed that the 67.4% of patients with an RDI between 90% and 110% refers to an RDI of 100%. The values so obtained appear to be quite different to those used as 'treatment modifiers', which constitutes a key issue. These RDI values have been used in the ERG base case (see Section 6.2).

4.2.9.4 Subsequent treatment costs

In the economic model, the assumption that patients who discontinue their initial therapy will consequently be eligible to receive a subsequent therapy, was applied. In the CheckMate 648 trial, not all patients received subsequent treatment because of patient comorbidities, fitness, or due to the stopping rule at 24 months so that patients with a complete response would not receive subsequent treatment.[2] The proportion of patients receiving subsequent therapy in the model to reflect this, has been presented in Table 4.17. The company in the CS[2] stated that this assumption reflected clinical practice where *"OC patients who discontinue their first-line therapy are likely to receive a subsequent therapy, with the possible subsequent therapies determined by the treatment they received in the first-line."*

Cyclical second-line average costs were calculated, with the frequency of each second-line treatment being dependent on the first-line treatment. These have been summarised in Table 4.18. The resultant average weighted costs applied in the model (for proportion of patients who will receive subsequent therapy) have been summarised in Table 4.19. All therapies assumed wastage. To prevent an implausible accrual of second-line treatment costs, a functionality which uses the median ToT data for second-line treatments to derive a cyclical second-line treatment discontinuation rate (0.056 in the model) that will enable patients to be moved from second-line treatment to no treatment, was applied in the economic model. This rate was weighted based on the frequency of use of treatment in the second line and combined to form an average second line cyclical discontinuation rate, both for the treatment and control arms. The median ToT of 12 weeks for nivolumab and 11 weeks for taxanes, was sourced from TA707.[45]

ERG comment:

- The ERG raised several questions on the modelling of duration of subsequent treatment, choice of agents for subsequent therapy, source of evidence for proportion of patients receiving subsequent therapies, and the non-application of specifying a maximum number of subsequent treatment cycles. The company's responses to these questions have been addressed in Section 4.2.2 of this report. Given that immunotherapies are expected to have a longer duration of treatment compared to other chemotherapies, not assuming the same median duration of second-line treatment for taxanes and nivolumab appears to be correct.
- The ERG finds the application of a median ToT estimate from TA707 in the economic model to limit subsequent therapy to a finite period in both treatment arms, to be an acceptable approach.

As described in Section 4.4.4, in the absence of effectiveness data that better match clinical practice, the ERG has constructed a scenario to better match cost to effectiveness as estimated from CheckMate 648. Given that calculations are not shown in the model, the ERG has used the values for proportion of patients who received subsequent therapy (Table 4.17) and the weighted average treatment costs (Table 4.19) to first calculate the unweighted per patient costs for taxanes and nivolumab and then reweight these values assuming the proportion of PD(L-)1 use observed in each arm of the trial (see Section 6.2).

Table 4.17: Subsequent therapy proportion, CheckMate 648 [redacted]

Treatment arm	Subsequent treatment	Proportion of patients
NIVO-CHEMO	Single agent taxane; assumed equal use of docetaxel and paclitaxel	[redacted]
CHEMO	Nivolumab monotherapy	[redacted]
PEMBRO-CHEMO	Single agent taxane; assumed equal use of docetaxel and paclitaxel	Aligned with NIVO-CHEMO

Based on Table 24 of CL response[16]
CHEMO = chemotherapy; CL = clarification letter; NIVO = nivolumab; PEMBRO = pembrolizumab

Table 4.18: Subsequent treatment costs, all arms

	Nivolumab	Taxane: docetaxel	Taxane: paclitaxel	Source
Dosing regimen	240 mg, on Day 1 Q2W	75 mg/m2, on Day 1 Q2W	100 mg/m2, on Day 1 every week for 6 weeks, followed by a 2-week break	ATTRACTION-3[78]
Dose received	240 mg	125 mg	166 mg	ATTRACTION-3[78]
Unit cost	£2,633.00	£17.95	£14.44	Drug acquisition costs
Admin method	Intravenous	Intravenous	Intravenous	ATTRACTION-3[78]
Admin cost	£284.05	£284.05	£284.05	Administration costs
Average cyclical cost	£1,458.52	£129.50	£191.62	Derived

Adapted from Table 50 of CS[2]
CS = company submission; Q2W = every two weeks

Table 4.19: Weighted average subsequent treatment costs

	Second-line treatment frequency			Second-line weighted average cyclical cost
	Nivolumab	Taxane: docetaxel	Taxane: paclitaxel	
NIVO-CHEMO	0%	26.58%	26.58%	£85.36*
CHEMO	56.7%	0%	0%	£826.80
PEMBRO-CHEMO	0%	24.7%	24.7%	£85.36*

Adapted from Table 51 of CS[2]
*Based on clarification response, due to change from 49.4% to [redacted] subsequent therapy.[16] (£79.26 in the original CS).[2]
CHEMO = chemotherapy; CS = company submission; NIVO = nivolumab; PEMBRO = pembrolizumab

4.2.9.5 Disease monitoring and management costs

The frequency of resource use in each health state was sourced through the literature using TA737, and the cost for each resource use was sourced from the NHS reference costs 2019-2020. The total cyclical health state costs used in the economic model have been publishes in Table 4.20. The company[16] clarified that *"monitoring costs are associated with the patient's health state, not the treatment they are receiving."* They further stated that this is in line with the approach taken in TA737.

Table 4.20: Weekly health state costs

Resource Use	Cost	Currency Code	Description	Setting	Weekly frequency pre-progression	Weekly frequency post-progression
CT scan	£103.31	RD25Z[76]	Computerised tomography scan of three areas, without contrast	Diagnostic imaging: weighted average of imaging: direct access, imaging: outpatient and imaging: other	0.08	0.08
Blood test	£2.53	DAPS05[76]	Haematology	Directly accessed pathology	0.33	1.00

Resource Use	Cost	Currency Code	Description	Setting	Weekly frequency pre-progression	Weekly frequency post-progression
				services		
Kidney	£33.80	WH15Z[76]	Special screening, examinations or other genetic disorders	Directly accessed diagnostic services	0.33	1.00
Hepatic	£33.80	WH15Z[76]	Special screening, examinations or other genetic disorders	Directly accessed diagnostic services	0.33	1.00
Consultant	£203.14	WF01A[76]	Non-admitted face-to-face attendance, follow-up	Consultant led: medical oncology	0.25	0.25
Total cost (SE)					£82.77 (£16.55)	£129.52 (£25.90)

Adapted from Table 53 of CS[2] and Table 34 of CL response[16]
CL = clarification letter; CS = company submission; CT = computed tomography; SE = standard error

ERG comment: Concerning monitoring frequencies, the company[16] clarified that, *"Per the ERG's request in TA737, the monitoring frequencies described in TA378 for the post-progression health state are employed to calculate the post-progression health state cost."* The ERG considers the choice of disease monitoring and management cost categories to be appropriate for the disease condition, and also reflective of TA737.[53]

4.2.9.6 Terminal care costs

A terminal care cost was applied as a one-off cost in the economic model to represent the management, monitoring, and resource use for patients with OC in the months leading up to death. This cost of £9,171.92 (SE: £1,834.38), was sourced from Georghiou et al. 2014.[79] The company considered radiotherapy and surgery (subsequent treatments used in CheckMate 648) to be palliative treatments that would implicitly be encompassed within the cost of terminal care.[16]

4.2.9.7 Adverse event costs

The CS reported that that the most common Grade 3/4 drug-related SAEs rates reported during the CheckMate 648 trial, were incorporated into the economic model. Each AE required a specific cost of management in the cycle in which the AE occurred and had a specific AE utility decrement applied additively to the health state utility values in the cycle in which the AE occurred. These costs have been summarised in Table 4.21 while the AE utility decrements have been discussed in Section 4.2.8 of this report.

ERG comment: In the economic model, AE management costs appear to have been applied in cycle 1 as a one-off cost.

Table 4.21: AE costs

AE	AE cost (SE)	Currency code	Description	Setting
Vomiting	£471.95 (£94.39)	FD10M[80]	Non-malignant gastrointestinal tract disorders without interventions, with CC Score 0-2	Non-elective short stay
Hyponatraemia	£1,164.14 (£232.83)	KC05H[80]	Fluid or electrolyte disorders, with interventions, with CC Score 0-4	Non-elective short stay
Pneumonitis	£1,909.33 (£381.87)	Weighted average DZ111K,L,M,N,P,Q,R,S,T,U,V[80]	Lobar, atypical or viral pneumonia, with multiple interventions, with CC Score 0-8, 9-13 and 14+. Lobar, atypical or viral pneumonia, with single intervention, with CC Score 0-7, 8-12 and 13+. Lobar, atypical or viral pneumonia, without interventions, with CC Score 0-3, 4-6, 7-9, 10-13 and 14+.	Total HRGs
Hepatic function abnormal	£2,461.04 (£492.21)	Weighted average GC01C,D,E,F[80]	Liver failure disorders with multiple interventions. Liver failure disorders with single intervention. Liver failure disorders without interventions, with CC Score 0-4 and 5+.	Total HRGs
Adrenal insufficiency	£2,079.75 (£415.95)	Chauhan 2013[81]	Not provided	Not provided
Acute kidney injury	£1,961.20 (£392.24)	Weighted average LA07H,J,K,L,M,N,P[80]	Acute kidney injury with interventions, with CC Score 0-5, 6-10 and 11+. Acute kidney injury without interventions, with CC Score 0-3, 4-7, 8-11 and 12+.	Total HRGs

AE	AE cost (SE)	Currency code	Description	Setting
Colitis	£2,426.57 (£485.31)	Copley-Merriman 2018[82]	Not provided	Not provided
Nausea	£471.95 (£94.39)	FD10M[80]	Non-malignant gastrointestinal tract disorders without interventions, with CC Score 0-2	Non-Elective Short Stay
Dehydration	£1,329.93 (£265.99)	Weighted average KC05G,H,J,K,L,M,N[80]	Fluid or electrolyte disorders, with interventions, with CC Score, 0-4 and 5+. Fluid or electrolyte disorders, without interventions, with CC Score 0-1, 2-3, 4-6, 7-9 and 10+.	Total HRGs
Febrile neutropenia	£4,755.76 (£951.15)	Copley-Merriman 2018[82]	Not provided	Not provided

Adapted from Table 54 of CS[2] and Table 35 of CL response[16]
AE = adverse event; CS = company submission; CL = clarification letter; HRGs = Healthcare Resource Groups; NHS = National Health Service; SE = standard error

ERG comment: The main concerns of the ERG relate to: a) use of old costs and inconsistent referencing; and b) terminal care costs

a) Administration costs were sourced from NHS reference 2015/16 costs rather than newer, up-to-date NHS reference 2019/20 costs and eMIT 2020 costs were chosen over eMIT 2021 costs. It is unclear why the company has preferred to use older costs. The ERG would prefer that all NHS costs are updated to the latest available reference costs. There were several inconsistencies in referencing of cost sources, for example, The CS references NHS reference 2019/20 costs for all costs of resources used in disease management e.g., computerised tomography (CT) scan, blood test, etc, whilst in Table 34 of the clarification letter response, costs are referenced NHS reference 2015/16 costs. The ERG examined all costs listed in the CS and clarification letter response and produced Table 4.22 using the provided currency codes and detailed its preferred costs and sources. Although the effect any update is likely to be small, given the lack of calculations in the model, none of the original values could be located and so the ERG has been unable to test the effect in the model. This is therefore part of a key issue, which is the lack of calculations based on original input data in the model.
b) The CS states that a one-off terminal care cost sourced from Georghiou et al. 2014[79] and adjusted to account for inflation: £9,171.92 (SE: £1,834.38) was applied to the economic model. This cost is significantly higher than the £7630.19 cost applied in TA737. Despite stating that this cost was *'externally sourced and published cost,'* no details on the inflation adjustment were provided, neither was what index was used and what the source of index is. The ERG remains retains concerns about how inflation was undertaken. As the CS only stated that the *"Terminal care costs represent the management, monitoring and resource use for patients with OC in the months prior to death,"* without giving further details on the period of time, it was difficult for the ERG to ascertain how applicable these generic end-of-life care costs would be

to OSCC. The ERG in TA737 also expressed some concern that the cost, which was actually £7,795.01 in the model, might be too high given the inclusion of radiotherapy.[1] Given the lack of justification for the cost used, the ERG did consider a scenario using £7,795.01 and considered that because it made little difference to the ICER, it would not be regarded as a key issue.

Table 4.22: Costs applied in the model, ERG preference

Resource	Description	Currency code	Company source and unit cost	ERG preferred source and unit cost
Administration costs				
Deliver simple parenteral chemotherapy at first attendance	Day case and reg day/night	SB12Z	NHS reference cost 2015/16 (weighted average)- £284.05	NHS reference cost 2019/20- £299.61
Deliver complex chemotherapy, including prolonged infusion treatment, at first attendance	Day case and reg day/night	SB14Z	NHS reference cost 2015/16- £431.72	NHS reference cost 2019/20- £431.72
Drug acquisition costs				
Capecitabine	150 mg tablets pack size 60	DHA224	eMIT 2020 - £4.43	eMIT 2021 - £4.43
	300 mg tablets pack size 60	DKE068	eMIT 2020 - £7.77	eMIT 2021 - £7.77
	500 mg tablets pack size 120	DHA225	eMIT 2020 - £26.30	eMIT 2021 - 26.30
Cisplatin	100 mg/100 ml solution for infusion vials	DHA010	eMIT 2020 - £8.73	eMIT 2021 - £8.97
	50 mg/50 ml solution for infusion vials	DHA011	eMIT 2020 - £5.38	eMIT 2021 - £6.03
Fluorouracil	1 g/20 ml (5%) solution for infusion vial	DHA265	eMIT 2020 - £2.35	eMIT 2021 - £2.25
	2.5 g/100 ml (2.5%) solution for infusion vial	DHA024	eMIT 2020 - £3.79	eMIT 2021 - £4.32
	2.5 g/50 ml (5%) solution for infusion vial	DHA102	eMIT 2020 - £4.01	eMIT 2021 - £4.21
	500 mg/10 ml (5%) solution for infusion vial	DHA240	eMIT 2020 - £1.77	eMIT 2021 - £2.86
	5 g/100 ml (5%) solution for infusion vials	DHA137	eMIT 2020 - £8.58	eMIT 2021- £9.2
Nivolumab	240 mg/24 ml concentrate for solution for infusion vial	N/A	BNF - £2,633	BNF- £2,633

Resource	Description	Currency code	Company source and unit cost	ERG preferred source and unit cost
Pembrolizumab	100 mg/4 ml concentrate for solution for infusion vial	N/A	BNF - £2,630	BNF- £2,630
Docetaxel	160 mg/8 ml solution for infusion vials (20 mg/ml)	DHC046	eMIT 2020 - £17.95	eMIT 2021- £17.38
Paclitaxel	300 mg/50 ml solution for injection vials	DHA210	eMIT 2020 - £14.44	eMIT 2021- £15.97
Disease monitoring costs				
CT scan	CT scan of three areas, without contrast	RD25Z	NHS 2019/20 - £103.31	NHS 2019/20 - £103.34
Blood test	Haematology	DAPS05	NHS 2019/20 - £2.53	NHS 2019/20 - £2.53
Kidney	Special screening, examinations or other genetic disorders	WH15Z	NHS 2019/20 - £33.80	NHS 2019/20 - £33.80
Hepatic	Special screening, examinations or other genetic disorders	WH15Z	NHS 2019/20 - £33.80	NHS 2019/20 - £33.80
Consultant	Non-admitted face-to-face attendance, follow-up	WF01A: service code 370	NHS 2019/20 - £203.14	NHS 2019/20 - £200

Based on CS,[2] CL response,[16] eMIT database,[74] BNF 2020[75] and NHS Reference costs 2019-2020[80]
BNF = British National Formulary; CS = company submission; CL = clarification letter; CT = computerised tomography; ERG = Evidence Review Group; eMIT = electronic market information tool; N/A = not available; NHS = National Health Service

4.2.10 Summary of base case analysis inputs and assumptions

The ERG asked the company to make the model input parameters and the probability distributions used in the sensitivity analyses, assumptions, and areas where these assumptions were applied in the economic model more transparent. These were provided by the company and can be seen in Tables 4.24 to 4.28.

Table 4.23: Summary of model settings, survival and progression functions applied in the economic model

Variable	Mean value	SE value (if applicable) or uncertainty measurement	Measurement of uncertainty and distribution	Section in Document B
Model settings				
Cycle length	1 week	N/A	N/A	B.3.2.1

Variable	Mean value	SE value (if applicable) or uncertainty measurement	Measurement of uncertainty and distribution	Section in Document B
Time horizon	2,080 weeks (40 years)	N/A	PSA: NA DSA: 260 to 520 weeks	B.3.2.1
Discounting rate (costs, outcomes)	3.5%	N/A	PSA: N/A DSA: 0% to 6% costs, 0% to 6% outcomes	B.3.2.1
Baseline parameters				
% Male	[redacted]	[redacted]	PSA: normal distribution DSA: 0% to 100%	B.3.2.2
Age	[redacted]	[redacted]	PSA: beta distribution DSA: 80% to 120% of mean	B.3.2.2
Survival and progression functions				
OS: NIVO-CHEMO	Semi-parametric 6.9 month cut point, log-normal	CIs	PSA: Described in Section B.3.3.1 DSA: N/A	B.3.3.1
OS: CHEMO	Semi-parametric 6.9 month cut point, log-normal	CIs		
PFS: NIVO-CHEMO	Semi-parametric 6.9 month cut point, generalised gamma	CIs		
PFS: CHEMO	Semi-parametric 6.9 month cut point, Weibull	CIs		
All-cause mortality	Based on UK lifetables	N/A	N/A	B.3.3.1.4

Based on Table 42 of CL response[16]
CHEMO = chemotherapy; CIs = confidence intervals; CL = clarification letter; DSA = deterministic sensitivity analysis; N/A = not applicable; NIVO = nivolumab; OS = overall survival; PFS = progression-free survival; PSA = probabilistic sensitivity analysis; SE = standard error; UK = United Kingdom

Table 4.24: Summary of clinical parameters applied in the economic model

Variable	Mean value	SE value (if applicable) or uncertainty measurement	Measurement of uncertainty and distribution	Section in Document B
Clinical parameters				
First line: ToT	K-M data both arms	CIs	PSA: described in Section B.3.3.2 DSA: N/A	B.3.3.2
Second line: ToT weighted taxane, cyclical discontinuation rate	0.0610	0.0061	PSA: beta distribution DSA: 80% to 120% of mean	B.3.3.2

Variable	Mean value	SE value (if applicable) or uncertainty measurement	Measurement of uncertainty and distribution	Section in Document B
Second line: ToT weighted, cyclical discontinuation rate	0.0561	0.0056	PSA: beta distribution DSA: 80% to 120% of mean	B.3.3.2
AE incidence NIVO-CHEMO				
Vomiting	[redacted]	[redacted]	PSA: beta distribution DSA: 80% to 120% of mean	B.3.3.3
Hyponatraemia	[redacted]	[redacted]		
Pneumonitis	[redacted]	[redacted]		
Hepatic function abnormal	[redacted]	[redacted]		
Adrenal insufficiency	[redacted]	[redacted]		
Acute kidney injury	[redacted]	[redacted]		
Colitis	[redacted]	[redacted]		
Nausea	[redacted]	[redacted]		
Dehydration	[redacted]	[redacted]		
Febrile neutropenia	[redacted]	[redacted]		
AE incidence chemotherapy				
Vomiting	[redacted]	[redacted]	PSA: beta distribution DSA: 80% to 120% of mean	B.3.3.3
Hyponatraemia	[redacted]	[redacted]		
Pneumonitis	[redacted]	[redacted]		
Hepatic function abnormal	[redacted]	[redacted]		
Adrenal insufficiency	[redacted]	[redacted]		
Acute kidney injury	[redacted]	[redacted]		
Colitis	[redacted]	[redacted]		
Nausea	[redacted]	[redacted]		
Dehydration	[redacted]	[redacted]		
Febrile neutropenia	[redacted]	[redacted]		

Based on Table 43 of CL response[16]

AE = adverse events; CHEMO = chemotherapy; CI = confidence intervals; CL = clarification letter; DSA = deterministic sensitivity analysis; K-M = Kaplan-Meier; N/A = not applicable; NIVO = nivolumab; PSA = probabilistic sensitivity analysis; SE = standard error; ToT = time on treatment

Table 4.25: Summary of utilities and disutilities applied in the economic model

Variable	Mean value	SE value (if applicable) or uncertainty measurement	Measurement of uncertainty and distribution	Section in Document B
Utilities				
Pre-progression health state utility	[redacted]	[redacted]	PSA: beta distribution DSA: 80% to 120% of mean	B.3.4.4
Post-progression health state utility	[redacted]	[redacted]	PSA: beta distribution DSA: 80% to 120% of mean	B.3.4.4

Variable	Mean value	SE value (if applicable) or uncertainty measurement	Measurement of uncertainty and distribution	Section in Document B
End-of-life utility decrement	[redacted]	[redacted]	PSA: beta distribution DSA: 80% to 120% of mean	B.3.4.4
AE disutilities				
Vomiting	0.048	0.016	PSA: beta distribution DSA: 80% to 120% of mean	B.3.4.4
Hyponatraemia	0.000	0.000		
Pneumonitis	0.037	0.004		
Hepatic function abnormal	0.119	0.012		
Adrenal insufficiency	0.119	0.012		
Acute kidney injury	0.048	0.016		
Colitis	0.047	0.005		
Nausea	0.048	0.016		
Dehydration	0.119	0.012		
Febrile neutropenia	0.090	0.016		

Based on Table 44 of CL response[16]
AE = adverse events; CL = clarification letter; DSA = deterministic sensitivity analysis; PSA = probabilistic sensitivity analysis; SE = standard error

Table 4.26: Summary of costs applied in the economic model

Variable	Mean value	SE value (if applicable) or uncertainty measurement	Measurement of uncertainty and distribution	Section in Document B
First line treatment costs				
Treatment arm: nivolumab cost per dose	[redacted]	N/A	PSA: N/A DSA: 80% to 120% of mean	B.3.5.1.1
Treatment and control arm: fluorouracil cost per dose	£1.77	£0.0012	PSA: gamma DSA: 80% to 120% of mean	3.5.1.1
	£1.77	£0.0012		
	£1.77	£0.0012		
	£1.77	£0.0012		
	£8.58	£0.0010		
Treatment and control arm: cisplatin cost per dose	£5.38	£0.0003	PSA: Gamma DSA: 80% to 120% of mean	3.5.1.1
	£8.73	£0.0007		
First line treatment modifier				
Treatment arm: nivolumab	[redacted]	N/A	PSA: NA DSA: 80% to 120% of mean	3.5.1.1.1
Treatment arm: fluorouracil	[redacted]	N/A		
Control arm: treatment arm: cisplatin	[redacted]	N/A		

Variable	Mean value	SE value (if applicable) or uncertainty measurement	Measurement of uncertainty and distribution	Section in Document B
Control arm: fluorouracil	[redacted]	N//A		
Control arm: cisplatin	[redacted]	NA		
Number of patients receiving subsequent treatment				
Treatment arm: nivolumab	[redacted]	N/A	N/A	3.5.1.2
Treatment arm: docetaxel	[redacted]	N/A		
Treatment arm: paclitaxel	[redacted]	N/A		
Control arm: nivolumab	[redacted]	N/A		
Control arm: docetaxel	[redacted]	N/A		
Control arm: paclitaxel	[redacted]	N/A		
Subsequent treatment costs				
Average cyclical cost: nivolumab	[redacted]	N/A	N/A	3.5.1.2
Average cyclical cost: docetaxel	£129.50	N/A		
Average cyclical cost: paclitaxel	£191.62	N/A		
Treatment arm: weighted average cyclical cost (nivolumab)	£85.36	£17.07	PSA: Gamma DSA: 80% to 120% of mean	
Control arm: weighted average cyclical cost (taxane: docetaxel and paclitaxel)	£826.80	NA	PSA: NA DSA: 80% to 120% of mean	
Health state costs				
Pre-progression health state cost	[redacted]	[redacted]	PSA: gamma distribution DSA: 80% to 120% of mean	B.3.5.1.2
Post-progression health state cost	[redacted]	[redacted]	PSA: gamma distribution DSA: 80% to 120% of mean	
Terminal care costs	[redacted]	[redacted]	PSA: gamma distribution DSA: 80% to 120% of mean	B.3.5.1.3
AE costs				
Vomiting	£471.95	£94.39	PSA: gamma distribution	B.3.5.2
Hyponatraemia	£1,164.14	£232.83		
Pneumonitis	£1,909.33	£381.87		

Variable	Mean value	SE value (if applicable) or uncertainty measurement	Measurement of uncertainty and distribution	Section in Document B
Hepatic function abnormal	£2,461.04	£492.21	DSA: 80% to 120% of mean	
Adrenal insufficiency	£2,079.75	£415.95		
Acute kidney injury	£1,961.20	£392.24		
Colitis	£2,426.57	£485.31		
Nausea	£471.95	£94.39		
Dehydration	£1,329.93	£265.99		
Febrile neutropenia	£4,755.76	£951.15		

Based on Table 45 of CL response[16]
AE = adverse events; CL = clarification letter; DSA = deterministic sensitivity analysis; N/A = not applicable; PSA = probabilistic sensitivity analysis; SE = standard error

Table 4.27: Assumptions applied in the model

Area	Assumption	Rationale
Baseline parameters	Baseline parameters are derived from CheckMate 648 cohort, which is assumed to be reflective of patients seen in UK clinical practice for the anticipated MA.	Although there may be differences between characteristics in CheckMate 648 and OSCC patients in UK clinical practice, these can be considered small. Sensitivity analyses (probabilistic and deterministic) have been conducted to assess the impact of variability in these parameters.
Model settings/ structure	The model applies a weekly cycle length, which is assumed to be sufficiently granular to accurately reflect costs and benefits when modelling OC.	Previous OC evaluations assessed by NICE had applied weekly cycle lengths, which was considered appropriate by ERG. This cycle length is short enough to reflect the treatment cycles for patients and reflects the frequency of follow-up for patients and reflects the frequency of follow-up for patients and a realistic minimum time during which symptoms or response can change.
Model settings/structure	To reflect the nature of OC and available evidence, the model assumes that OC phases are consecutive, and patients cannot revert to pre-progression from more advanced phases of the disease.	This assumption has been validated by clinicians and is in line with other HTAs and economic analyses assessing the OC population.
Efficacy	Identification of most appropriate survival curves describing PFS, and OS inform extrapolation	Extensive analyses have been undertaken to identify appropriate and conservative survival curves describing NIVO-CHEMO efficacy, with reference to the guidance from the NICE DSU and Bagust and Beale et al 2014. The approach and identified survival extrapolations have been validated by clinical and health economic experts. However, to address the uncertainty

Area	Assumption	Rationale
		around this parameter, scenario analyses have been conducted by applying alternative assumptions around extrapolations, as presented in Section B.3.3.1.
Efficacy	Efficacy has been based on BICR -assessed data, rather than investigator-assessed data.	During CheckMate 648, the two measures of response of PFS were comparable. However, BICR was designated as the primary endpoint and may be considered slightly more conservative.
Safety	As a simplification, it is assumed that all AEs occur in the first cycle of treatment.	The majority of patients during CheckMate 648 have discontinued treatment within the current database lock, so that the data can be considered an accurate reflection of the safety profile. AEs are often only observed to occur soon after treatment initiation, so that this may not be well reflected by assuming a constant rate per cycle.
HRQoL	It was assumed that health state utilities, pre-progression, post-progression and the disutility of death, are the same for the treatment and control arm.	This is based on evidence observed during CheckMate 648, described in Section B.3.4.2.
Treatment costs	It was assumed that patients receiving pembrolizumab in combination with chemotherapy experience missing or delayed doses in line with nivolumab during CheckMate 648.	Currently, there is no published data available to inform proportion of received doses of pembrolizumab. As the mechanism of action is similar, this seems an appropriate assumption.
Health state costs	The health state resource use is derived from evidence presented in TA737.	Robust estimates of health state resource use for patients in this setting are not publicly available, given the limited alternative treatment available for which evidence may have previously gathered. In order to provide relevant economic evaluations and facilitate comparison between these appraisals, health state resource use from TA737 is applied.
Treatment pathway	Subsequent treatment for NIVO-CHEMO and PEMBRO-CHEMO is assumed to be single agent taxane (equal use of paclitaxel and docetaxel).	During CheckMate 648, taxane use reflected around 70% of subsequent systemic therapy use, indicating the plausibility of this assumption. Docetaxel and paclitaxel have similar efficacy and cost.
Treatment pathway	Subsequent treatment for CHEMO is assumed to nivolumab monotherapy.	This aligns with the current UK treatment pathway and is aligned with budget impact assumptions applied during TA707.
Safety	AE utility decrement values were assumed for certain AEs.	Values were assumed for those AEs where published data was not available. However, deterministic sensitivity analysis has been presented to show the impact of AE utility decrements.

Area	Assumption	Rationale
Efficacy	No treatment waning has been assumed.	Evidence supports a robust and durable treatment effect lasting beyond discontinuation for immunotherapies. Further, during TA737, the committee concluded that all scenarios provided plausible estimates of overall survival and the treatment waning scenarios were not greatly different from those without treatment waning. This is of particular relevance given the low long-term hazard in the CHEMO arm of CheckMate 648.

Based on Table 49 of the CL response[16]
AEs = adverse events; BICR = blinded independent central review; CL = clarification letter; CHEMO = chemotherapy; DSU = Decision Support Unit; ERG = Evidence Review Group; HTA = Health Technology Assessment; NICE = National Institute of Health and Care Excellence; NIVO = nivolumab; OC = oesophageal cancer; OS = overall survival; OSCC = oesophageal squamous cell carcinoma; PFS = progression-free survival; TA = technology appraisal; UK = United Kingdom

5. COST EFFECTIVENESS RESULTS

5.1 Company's cost effectiveness results

The CS base case cost effectiveness results have been summarised in Table 5.1. For patients treated with chemotherapy, the model predicted 1.281 discounted life years (LYs) with an accrual of 0.917 discounted QALYs. Nivolumab use with chemotherapy was estimated to result in an additional 0.717 discounted QALYs (total 1.634 QALYs) and an additional 0.998 discounted LYs (total 2.280 LYs).

Table 5.1: Overview of base case results (with PAS; discounted), NIVO-CHEMO versus CHEMO

	NIVO-CHEMO	CHEMO	Incremental
LYs	[redacted]	[redacted]	[redacted]
QALYs	[redacted]	[redacted]	[redacted]
Total costs (£)	[redacted]	[redacted]	[redacted]
ICER (£/QALY)			£33,272

Based on Table 57 of CS[2]
CHEMO – chemotherapy; CS = company submission; ICER = incremental cost-effectiveness ratio; LYs = life years; NIVO = nivolumab; PAS = Patient Access Scheme; QALY = quality-adjusted life year

The company provided updated cost effectiveness analyses using more recent subsequent treatment costs and updated proportion of patients receiving a subsequent treatment in response to the ERG's clarification question. These have been presented in Table 5.2 for a NIVO-CHEMO versus CHEMO comparison, and Table 5.3 for a NIVO-CHEMO versus PEMBRO-CHEMO comparison in [redacted], 6.9 month cut-off point.

Table 5.2: NIVO-CHEMO versus CHEMO, [redacted], 6.9 month cut-point, PD-L1 ≥1%

	NIVO-CHEMO	CHEMO	Incremental
LYs	[redacted]	[redacted]	[redacted]
QALYs	[redacted]	[redacted]	[redacted]
Total costs (£)	[redacted]	[redacted]	[redacted]
ICER (£/QALY)			£33,357

Based on Table 27 of CL response[16]
CHEMO – chemotherapy; CL = clarification letter; DBL = database lock; ICER = incremental cost-effectiveness ratio; LYs = life years; NIVO = nivolumab; PD-L1 = programmed death ligand 1; QALY = quality-adjusted life year

Table 5.3: NIVO-CHEMO versus PEMBRO + CHEMO, [redacted], 6.9 month cut-point, PD-L1 ≥1%

	NIVO-CHEMO	PEMBRO + CHEMO	Incremental
LYs	[redacted]	[redacted]	[redacted]
QALYs	[redacted]	[redacted]	[redacted]
Total costs (£)	[redacted]	[redacted]	[redacted]
ICER (£/QALY)			-£5,594

Based on Table 28 of CL response[16]

	NIVO-CHEMO	PEMBRO + CHEMO	Incremental
CHEMO – chemotherapy; CL = clarification letter; DBL = database lock; ICER = incremental cost-effectiveness ratio; LYs = life years; NIVO = nivolumab; PD-L1 = programmed death ligand 1; PEMBRO = pembrolizumab; QALY = quality-adjusted life year			

Total discounted costs associated with nivolumab, and chemotherapy were predicted to be £51,334. Incremental costs were predicted to be £27,494 compared to chemotherapy alone, under base case assumptions. The resulting ICER estimate for nivolumab with chemotherapy versus chemotherapy alone was £33,272 per QALY gained. Therefore, the base case ICER is below the £50,000 per QALY willingness to pay threshold. Table 5.4 details the base case parameters and results.

Table 5.4: Detailed base case analysis results

	NIVO-CHEMO	CHEMO
Patient level survival (undiscounted)		
Median PFS (years)	[redacted]	0.383
Mean PFS (years)	[redacted]	0.576
Median OS (years)	[redacted]	0.747
Mean OS (years)	[redacted]	1.382
Patient-level progression		
Time in pre-progression (years)	[redacted]	0.576
Time in post-progression (years)	[redacted]	0.807
Costs (with PAS)		
Health state costs	[redacted]	£7,290
Treatment costs	[redacted]	£11,355
AE costs for initial therapy	[redacted]	£82
Terminal care costs	[redacted]	£8,768
Total costs	[redacted]	£27,494
Health benefits		
Health state QALYs	[redacted]	0.931
AE utility	[redacted]	-0.0001
Time-to-death utility	[redacted]	-0.0142
Total QALYs	[redacted]	0.917
Total LYs (undiscounted)	[redacted]	1.382
Incremental results		
Incremental total costs	-	[redacted]
Incremental QALYs	-	[redacted]
Incremental LYs (undiscounted)	-	[redacted]
Cost/QALY	**-**	**£33,272**
Based on Table 58 of CS[2] AE = adverse event; CHEMO = chemotherapy; CS = company submission; LY = life years; NIVO = nivolumab; OS = overall survival; PAS = Patient Access Scheme; PFS = progression-free survival; QALY = quality-adjusted life year		

A full incremental analysis was also presented in the company's response to clarification, this has been presented in Table 5.5.

Table 5.5: Fully incremental analysis, base case population (NIVO-CHEMO versus PEMBRO-CHEMO versus CHEMO, [redacted]

Treatment	Total costs (discounted, £)	Total QALYs (discounted)	ICER (£/QALY)/result	ERG corrected ICER
NIVO-CHEMO	[redacted]	[redacted]	-	Dominated
PEMBRO + CHEMO	[redacted]	[redacted]	£29,204*	£29,204*
CHEMO	[redacted]	[redacted]	Dominated	-

Based on Table 29 in CL response[16]
*ICER versus CHEMO
CL = clarification letter; CHEMO = chemotherapy; DBL = database lock; ERG = Evidence Review Group; ICER = incremental cost-effectiveness ratio; NIVO = nivolumab; PEMBRO = pembrolizumab; QALY = quality-adjusted life year

5.2 Company's sensitivity and scenario analyses

5.2.1 Probabilistic sensitivity analysis

The company conducted a probabilistic sensitivity analysis (PSA) in which all input parameters were sampled simultaneously from their corresponding probability distributions over 1,000 iterations.

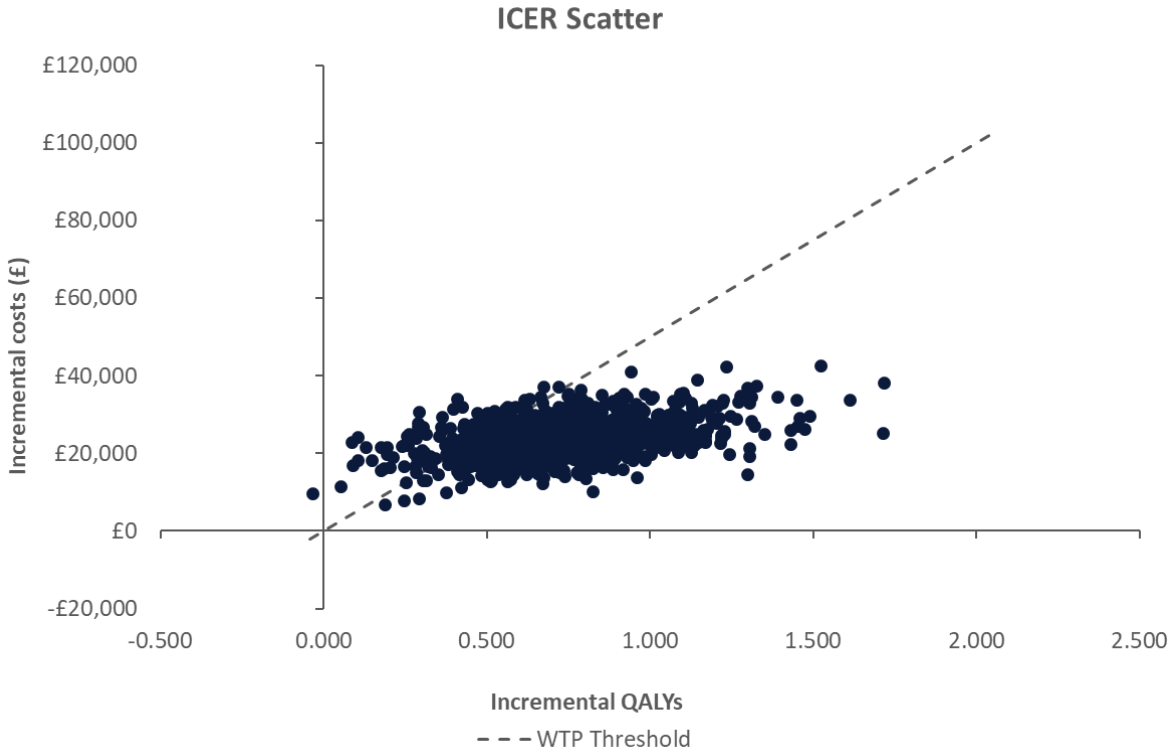
The average PSA results are summarised in Table 5.6 and presented on a cost effectiveness plane in Figure 5.1, from which a cost effectiveness acceptability curve (CEAC) was calculated and plotted in Figure 5.2. Both the cost effectiveness plane and CEAC plots are based on NIVO-CHEMO versus CHEMO.

Table 5.6: PSA results (combined complete response/partial response)

	NIVO-CHEMO	CHEMO	Incremental
LYs	[redacted]	[redacted]	[redacted]
QALYs	[redacted]	[redacted]	[redacted]
Total costs (£)	[redacted]	[redacted]	[redacted]
ICER (£/QALY)			£32,736

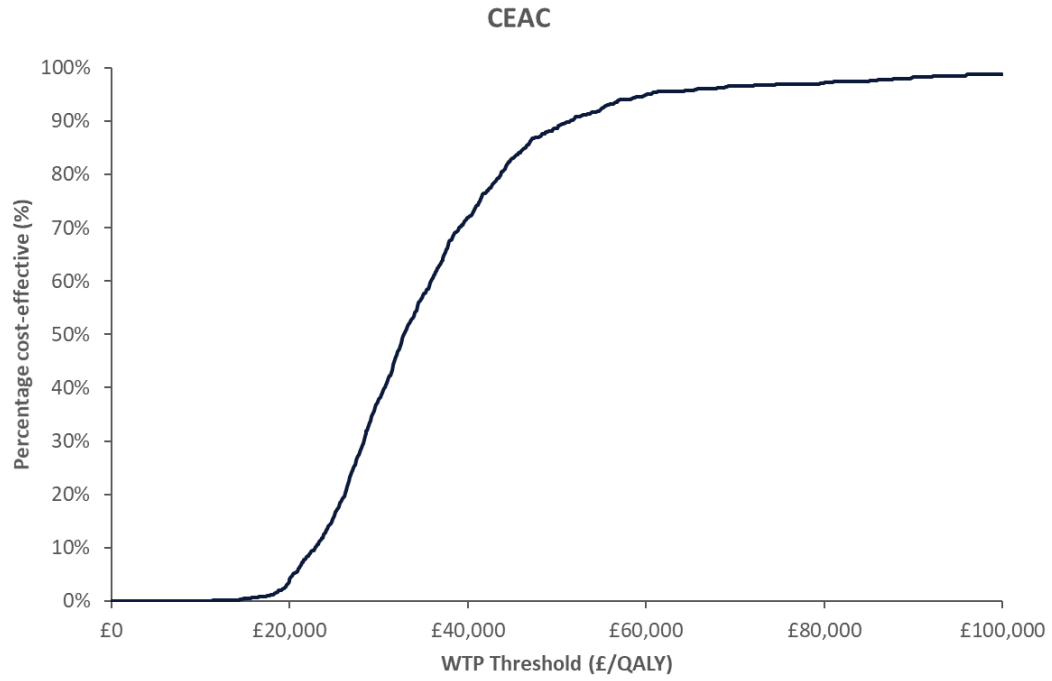
Based on Table 59 of CS[2]
CS = company submission; ICER = incremental cost-effectiveness ratio; LYs = life years; PSA = probability sensitivity analysis; QALY = quality-adjusted life year

Figure 5.1: PSA cost effectiveness plane



Based on Figure 45 of CS[2]
CS = company submission; ICER = incremental cost-effectiveness ratio; PSA = probability sensitivity analysis; QALY = quality-adjusted life year; WTP = willingness-to-pay

Figure 5.2: PSA cost effectiveness acceptability curve



Based on Figure 46 of CS[2]
CEAC = cost effectiveness acceptability curve; CS = company submission; PSA = probability sensitivity analysis; QALY = quality-adjusted life year; WTP = willingness-to-pay

5.2.2 Deterministic sensitivity analyses

The company also conducted deterministic sensitivity analyses (DSAs) where key parameters which influence the results and conclusions of the decision problem to the greatest degree, were individually varied at lower and upper bounds of values that were deemed plausible by the company. These are summarised in Table 5.7.

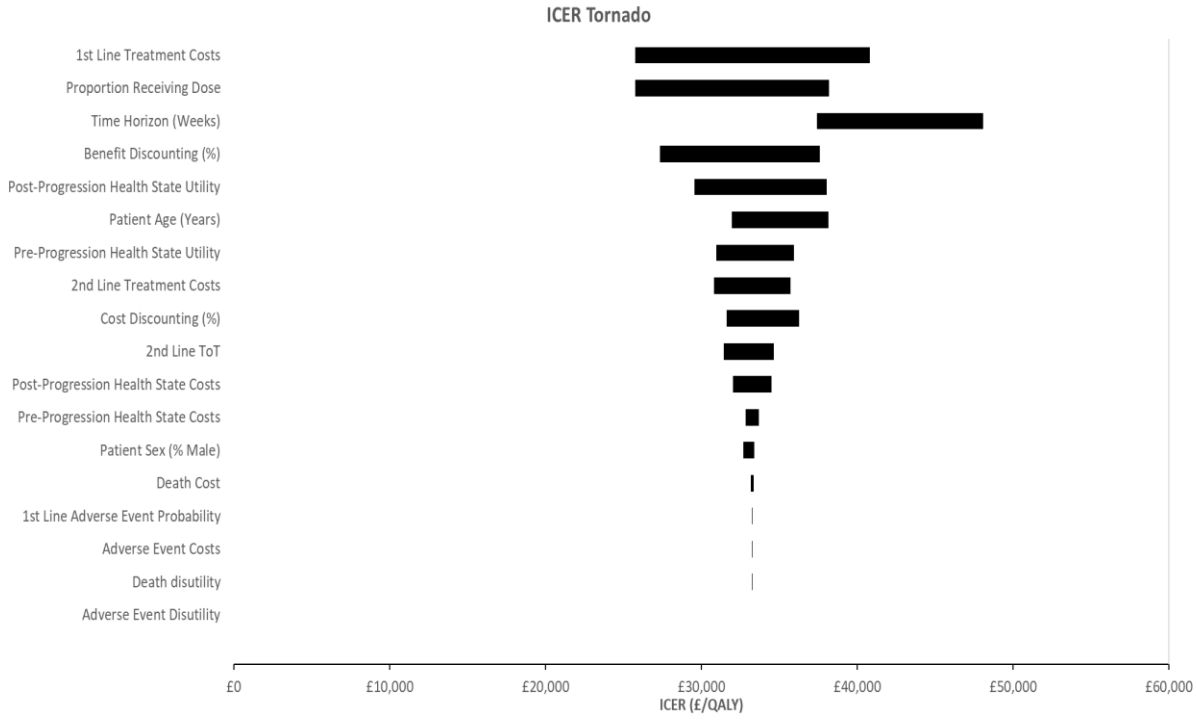
Table 5.7: Parameters and values included in the company's DSA

Parameter	Parameter variation	ERG comment
Time horizon	260 weeks (5 year) and 520 weeks (10 years)	Model time horizon is 40 years, 5 and 10 years not sensitive. Suitable for scenario analysis
Discounting: costs	0% and 6%	Agree, commonly used
Discounting: benefits	0% and 6%	Agree, commonly used
Baseline characteristics: age	± 20%, impacting on all-cause mortality	Agree, commonly used
Baseline characteristics: sex	0% and 100% male, impacting on all-cause mortality	Arbitrary
Health state costs: pre-progression and post-progression	± 20%	Agree, 80% to 120% of mean appropriate
Health state costs: terminal care costs	± 20%	Agree, 80% to 120% of mean appropriate
Initial treatment costs	± 20%	Agree, 80% to 120% of mean appropriate
Subsequent treatment costs	± 20%	Agree, 80% to 120% of mean appropriate
AE costs	± 20%	Agree, 80% to 120% of mean appropriate
Health state utility: pre-progression and post-progression	± 20%	Agree, 80% to 120% of mean appropriate
End-of-life utility	± 20%	Agree, 80% to 120% of mean appropriate
AE disutility	± 20%	Agree, 80% to 120% of mean appropriate
Second line ToT	± 20%	Agree, 80% to 120% of mean appropriate
Treatment modifier: proportion receiving dose	± 20%	Agree, 80% to 120% of mean appropriate
AE probability	± 20%	Agree, 80% to 120% of mean appropriate
Subsequent treatment ToT	± 20%	Agree, 80% to 120% of mean appropriate

Based on Section B.3.8.2 of CS[2]
AE = adverse event; CS = company submission; DSA = deterministic sensitivity analysis; ERG = Evidence Review Group; ToT = time on treatment

The results of the DSA for all scenarios defined in Table 5.7 have been presented in Table 60 of the CS.[2] In comparing NIVO-CHEMO versus CHEMO, all ICERs were between £25,000 per QALY and £50,000 per QALY gained. The results for this comparison were summarised by the company in the form of a tornado diagram as shown in Figure 5.3. This shows that the parameters with the greatest impact were first-line treatment costs, proportion of patients receiving a dose and the post-progression health state utility.

Figure 5.3: DSA tornado diagram for NIVO-CHEMO versus CHEMO



Based on Figure 47 of CS[2]
CHEMO = chemotherapy; CS = company submission; DSA = deterministic sensitivity analysis; NIVO = nivolumab; QALY = quality-adjusted life year; ICER = incremental cost-effectiveness ratio

5.2.3 Scenario analyses

The company conducted several scenario analyses to evaluate the effect of modelling assumptions on the ICER in Section B.3.8.3 of the CS.[2] These included:

- Applying alternative survival extrapolations in comparing NIVO-CHEMO to CHEMO: All scenarios increased the ICER when compared to the base case.
- Testing the impact of alternative comparators: ICERs for the additional alternative comparators, FOLFOX, XELOX and cisplatin + capecitabine, decrease compared to the base case ICER.
- Removing the treatment modifier: The ICER increased but remained below the £50,000/ QALY WTP threshold.
- Testing the impact of alternative utility assumptions: No significant change in the ICER when time to death utilities were removed.

5.2.4 Additional scenario analyses requested by the ERG

The ERG in its clarification letter asked the company to conduct scenario analyses using adjusted data in the economic model, including variation in the proportion of patients who experience the treatment effect of anti-PD-1/PD-L1 therapies to better reflect NHS clinical practice. In response,[16] the company

stated that *"Since we believe that it is not necessary to undertake analyses to adjust for treatment switching, scenario analyses will not be conducted either."*

A scenario analysis incorporating an on/off NIVO-CHEMO treatment adjustment in the economic model was also requested. In response,[16] the company stated that, *"Since a regression analysis was deemed not necessary, a scenario analysis in the economic model was not conducted."*

5.2.5 Conclusions from company's sensitivity and scenario analyses

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

- First-line treatment costs.
- Proportion receiving a dose.
- Post-progression health state utility.

ERG comment: The ERG in its clarification letter asked the company to provide the selection criteria for the parameters to be included in the PSA and DSA. In response,[16] the company stated that, *"The parameters excluded from the PSA are the components that make up first line and second line treatment costs (except for the inputs sourced from eMIT whose treatment costs have some level of uncertainty), lifetables and model settings. These parameters are excluded based on the fact that they are fixed parameters, which do not contain uncertainty with regards to this model. All parameters included in the PSA are done so on the basis that some degree of uncertainty remains. All parameters except for survival and PAS are included in the DSA. The parameters included in the DSA are those parameters whose variation provides an insight into the key drivers of cost-effectiveness in the model. Whilst survival is a key driver in the model, sensitivity around the choice of extrapolation is explored extensively in scenario analyses in the company submission."* The ERG is satisfied with the company's applied criteria.

5.3 Model validation and face validity check

Validation efforts conducted on the economic model were discussed in Section B.3.10.1 of the CS.[2]

5.3.1 Face validity assessment

In the CS[2], the company stated that, *"the relevance of the model structure and assumptions were validated through consultation with UK clinicians."*

ERG comment: It was unclear if external clinical and health economic experts were consulted to ensure that the model had clinical validity. To probe further, the ERG in its clarification letter asked the company to provide details of the communication between the company and clinical and health economic experts, and also list their recommendations and justifications for clinical assumptions and inputs used in the model. In response, the company[16] stated that, *"An advisory board was held on 14 July 2021 by BMS comprising of clinicians and an economist, with the aim of developing insight to support the NICE submission for nivolumab with platinum-based chemotherapy for the treatment of advanced unresectable, recurrent or metastatic previously untreated oesophageal squamous cell carcinoma (OSCC). The board explored key themes developed by BMS around specific issues related to the clinical positioning and economic strategy and shared published results from the CheckMate 648 trial to gain feedback on how they resonated with clinicians and economists."*

Table 5.8: Expert recommendations from the advisory board

Assumption	Justification
UK clinical practice	
Data regarding squamous GC can be considered comparable to OSCC	The REAL2 study,[83] shows that squamous GC and OC are comparable.
There were no additional treatments to be considered for the treatment of advanced, previously untreated OSCC, beyond the doublet and triplet regimens presented.	Confirmed by the clinicians and aligned with NICE guidance at the time of the advisory board.
If nivolumab combination therapy was approved as a first-line treatment, then a nivolumab-containing second-line therapy would not be offered. It was generally believed that a docetaxel or paclitaxel-containing regimen would be offered in the second-line after a nivolumab-containing first-line regimen.	Current NICE clinical guidance and clinical expert opinion.
CheckMate 648	
Eligibility criteria and baseline patient characteristics representative of patients seen in UK clinical practice	Clinical expert opinion
There is no difference between OSCC patients from Asia or Europe.	Clinical expert opinion
The safety profile for nivolumab with chemotherapy was not a concern for the clinicians as they would be expecting AEs with both immunotherapies and chemotherapy and so would select and treat patients accordingly.	Clinical expert opinion
Survival modelling	
The survival data presented from CheckMate 648 aligned with the experts expectations.	Clinical expert opinion
In lethal cancer, patients who survive beyond 18-24 months are considered long-term survivors and would stop immunotherapy at this stage	Product SmPCs and guidance, and expert opinion.
Resource use in patients surviving beyond 24 months would be fairly intensive, as patients may still be symptomatic	Clinical expert opinion
it would be appropriate to use long-term clinical data from other nivolumab indications to validate the hazard profile evolution	Clinical expert opinion
The Weibull and Gompertz estimates were thought to be the most similar to current clinical practice in the UK	Clinical expert opinion

Assumption	Justification
Cost effectiveness modelling	
Published utility values from a squamous gastric cancer population would be appropriate to include in the model for external validation or to inform post-progression data gaps.	The GO2 trial in upper GI cancer reported utility as a primary endpoint and was suggested as a good source.
Based on Table 50 of the CL response[16] AEs = adverse events; CL = clarification letter; GI = gastrointestinal; GC = gastric cancer; NICE = National Institute for Health and Care Excellence; OSCC = oesophageal squamous cell carcinoma; SmPC = summary of product characteristics	

5.3.2 Technical verification

The company stated that a technical review of the cost effectiveness model was conducted by an independent economist.

ERG comment: Although the company[2] stated that, *"quality control was undertaken, whereby a cell-by-cell verification process was conducted to allow checking of all input calculation, formulae and visual basic code,"* as the economic model did not display input calculations and formulae (in the data library), the ERG was unable to verify these statements and felt that additional information was needed to understand if appropriate technical verification had been conducted and thus asked the company to provide further details of this validation effort, to confirm if this technical review consisted of black and white-box tests to detect modelling errors, and if not, to complete the Technical Verification (TECH-VER) checklist.

In their response to clarification,[16] the company assured the ERG that the *"…internal validity of the model was tested in line with Büyükkaramikli et al., and as such included 'black-and-white' tests to detect modelling errors. Some examples include:*

a. *Setting treatment effects to 0*
b. *Setting discounting to 0%*
c. *Setting model inputs equal across treatment arms*
d. *Setting costs to 0, increasing/decreasing costs per arm*
e. *Setting utilities to 0, increasing/decreasing utilities per arm*

In each case, results were checked to ensure trends and model behaviour were as expected. For example, when discounting was set to 0%, it was checked that discounted costs and QALYs were equivalent to undiscounted; or for increasing costs in the treatment arm only, no impact was observed on costs in the control arm)." They stated[16] that this effort was carried out by an independent senior health economist and that *"The technical review focussed on various areas including conceptual and internal validation. internal validation comprised:*

a. *Technical pressure testing (or extreme values analysis) – model input parameters are modified in such a way that their impact on results should be immediately intuitive, enabling rapid identification of errors in modelling logic*
b. *Directional input testing – modelled clinical input parameters are modified individually and their directional relationship with cost and QALY outcomes evaluated."*

The ERG is satisfied with the thoroughness and appropriateness of the company's TECH-VER process of the economic model. However, there is a lack of transparency in certain aspects of the model:

1) the cost calculations with none of these being available in the workbook such that it was impossible for the ERG to check or update some of the costs (see Section 4.2.9).
2) the implementation of the ITC to estimate the survival of PEMBRO-CHEMO (see Section 4.2.6).

The uncertainty as result of this lack of transparency is therefore a key issue.

5.3.3 Comparisons with other technology appraisals

The ERG in its clarification letter asked the company to provide cross validations with other relevant NICE TAs regarding model structure and assumptions, and input parameters related to clinical effectiveness, health state utility values, resource use and costs, and estimated (disaggregated) outcomes per comparator/intervention. In response,[16] the company stated that, *"The only relevant NICE appraisal for first line advanced or metastatic OSCC is TA737 for pembrolizumab. As such, model inputs and outputs relating to NIVO-CHEMO cannot be cross-validated. From TA737, data relating to pembrolizumab + 5FU +cisplatin, and 5FU + cisplatin, have been assessed as these are the relevant treatment regimens in the current appraisal."* Concerning outcomes, they published Table 66 from the CS and stated,[16] *"Disaggregated outcomes from TA737 are redacted, and therefore, cannot be compared. Total and incremental LY/QALY/costs from the current appraisal versus TA737 are explored within B.3.10.3 of the company submission… Overall, predicted LY and costs are broadly comparable."* See Table 5.9 for relevant comparisons.

Table 5.9: Comparison of the CS and TA737 economic model

	Current appraisal	TA737 original CS
Model structure and assumptions		
Model structure	Three state partitioned survival model (progression-free, progressed disease, death)	Three state partitioned survival model (progression-free, progressed disease, death)
Time horizon	Lifetime	Lifetime
Cycle length	One week no half-cycle correction	One week with half-cycle correction
Utility source	CheckMate 648 EQ-5D-3L	KEYNOTE-590 EQ-5D-3L
Cost source	eMIT and BNF for acquisition costs; administration costs, AE costs, disease management costs from NHS reference costs	eMIT and BNF for acquisition costs; administration costs, AE costs, disease management costs from NHS reference costs
Duration of treatment effect	No treatment waning	No treatment waning in company base case
Treatment pathway	Subsequent treatments in line with clinical practice (based on clinical expert opinion)	Subsequent treatments in line with those from KEYNOTE-590
Safety	AE incidence from CheckMate 648	AE incidence from KEYNOTE-590
Stopping rule	Stopping rule based on treatment specific time on treatment curves	Pembrolizumab not administered beyond 24 months, cisplatin to 6 cycles, 5-FU to 25 cycles
Clinical effectiveness		
PFS efficacy	K-M data (CheckMate 648) to 6.9 months, followed by generalised gamma distribution for the treatment	K-M data (KEYNOTE-590) to 10 weeks, followed by log-logistic

	Current appraisal	TA737 original CS
	arm and Weibull distribution for the control arm	distribution, since first tumour assessment at week 9
OS efficacy	K-M data (CheckMate 648) to 6.9 months, followed by generalised lognormal distribution for the treatment and control arm	K-M data (KEYNOTE-590) to 40 weeks, followed by log-logistic models, established via clinical validity and AIC/BIC
HRQoL		
Health state utility values	By progression status, [redacted] pre-progression, [redacted] post-progression, with end-of-life decrement ([redacted])	TTD utilities, values redacted
Age-related disutility	Utilities not adjusted by UK general population	Utilities adjusted by UK general population
Resource use and costs		
Time on treatment	ToT curves applied to both arms, based on CheckMate 648 (mean ToT from TA737 used to derive PEMBRO + CHEMO time on treatment curve)	ToT applied to both arms, based on KEYNOTE-590
Relative dose intensity/ treatment modifier	Dose intensity applied to all arms, based on CheckMate 648 for NIVO-CHEMO and CHEMO. PEMBRO + CHEMO assumed equivalent to NIVO-CHEMO	Relative dose intensity applied to both arms, based on KEYNOTE-590. Values redacted.
Healthcare resource use	Aligns between treatment and control arms	Aligns between treatment and control arms
Pre-progression healthcare resource use (per cycle)	0.08 CT scan 0.33 full blood count 0.33 renal function test 0.33 hepatic function test 0.25 consultation visit	0.08 CT scan 0.33 full blood count 0.33 renal function test 0.33 hepatic function test 0.25 consultation visit
Post-progression healthcare resource use (per cycle)	0.08 CT scan 1 full blood count 1 renal function test 1 hepatic function test 0.25 consultation visit	0.08 consultation visit
Administration costs for first line treatments*	Cisplatin + 5FU, SB14Z (NHS reference costs) at first attendance Nivolumab, SB12Z (NHS reference costs) on day 15 per cycle, and SB14Z (NHS reference costs) on day 1 per cycle.	In both PEMBRO + cisplatin + 5FU, and cisplatin + 5FU, SB14Z (NHS reference costs) at first attendance
Acquisition costs for first line treatments*	BNF for nivolumab and pembrolizumab, eMIT for chemotherapy components	BNF for pembrolizumab, eMIT for chemotherapy components
Terminal care cost	Based on Georghiou et al. 2014, adjusted for inflation	Based on TA522, adjusted for inflation.

	Current appraisal	TA737 original CS
AEs	Incidence from CheckMate 648, with the most common Grade 3–4 drug-related SAEs from all treatment arms included. Incidence for PEMBRO + CHEMO taken from TA737. One-off cost and disutility applied on incidence of AE. AEs only associated with first line treatment, and only occur on treatment initiation. Costs based on NHS reference costs and literature. Disutility based on TAs, literature, and assumptions.	Incidence from KEYNOTE 590, one-off cost and disutility applied. Cost based on mean duration and NHS reference costs. Utility based on KEYNOTE 590 data, time to death approach.

Adapted from Table 47 in CL response[16]
*Detail of costs themselves not incorporated herein, due to updates in NHS reference cost and eMIT databases
Note that the data from TA737 presented in this table relates to the original CS, and not any updates following ERG/NICE review.
AEs = adverse events; AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; BNF = British National Formulary; CHEMO = chemotherapy; CL = clarification letter; CS = company submission; CT = computed tomography; ERG = Evidence Review Group; eMIT = electronic market information tool; Eq-5D-3L = European Quality of Life-5 dimensions-3 levels; K-M = Kaplan-Meier; NHS = National Health Service; NICE = National Institute for Health and Care Excellence; NIVO = nivolumab; PEMBRO = pembrolizumab; SAEs = serious adverse events; TA = Technology Assessment; ToT = time on treatment; UK = United Kingdom

5.3.4 Comparison with external data

As referred to in Section 4.2.6, in the CS no external validation for survival extrapolation was reported, the only justification being for the NIVO-CHEMO arm: *"There are no other studies with which to validate the results for extrapolation of the NIVO-CHEMO arm other than the informing trial, CheckMate 648."* (p. 125, CS)[2] The company therefore performed a validation exercise, which summarised a comparison of the company's economic model output with CheckMate 648 trial data across patient-level data, preferred survival curves, and model outputs (Table 5.10).

Table 5.10: Comparison of economic model output with CheckMate 648 data

		NIVO + CHEMO			CHEMO		
		PLD	Preferred Survival Curves	Model Output	PLD	Preferred Survival Curves	Model Output
OS	1 year	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	2 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	3 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	5 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	10 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PFS	1 year	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	2 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	3 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	5 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	10 years	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

	NIVO + CHEMO			CHEMO		
	PLD	Preferred Survival Curves	Model Output	PLD	Preferred Survival Curves	Model Output
Based on Table 67, CS.[2] CHEMO = chemotherapy; CS = company submission; NIVO = nivolumab; N/R = not reported; OS = overall survival; PFS = progression-free survival; PLD = patient-level data						

ERG comment: The ERG in its clarification letter asked the company to assess the external validity of model inputs, intermediate outcomes, and final outcomes using evidence not used to develop the model. The company referred to the cross-validation effort discussed in Section 5.3.3 of this report, but not did provide further comparisons with external data not used to develop the model. The ERG has already compared the patient level (K-M) data to the extrapolations based on each of the company's preferred survival curves in Section 4.2.6. It is worth noting that there are also K-M data at 3.5 years (see Tables 4.6 and 4.9. Unfortunately, there are discrepancies for PFS, which the ERG cannot explain, but generally the values at 3.5 years show much less of a difference between the arms than would be the case with the preferred survival curves/according to the model output at 5 and 10 years, especially for OS. This supports the ERG's conclusion that there is likely to be a warning of treatment effect at least for OS and that this probably begins before 3.5 years.

6. EVIDENCE REVIEW GROUP'S ADDITIONAL ANALYSES

6.1 Exploratory and sensitivity analyses undertaken by the ERG

Based on all considerations in the preceding Sections of this ERG report, the ERG defined a new base case. This base case included multiple adjustments to the original base case presented in the previous Sections. These adjustments made by the ERG form the ERG base case and were subdivided into three categories (derived from Kaltenthaler 2016)[84]:

- Fixing errors (FE) (correcting the model where the company's submitted model was unequivocally wrong)
- Fixing violations (FV) (correcting the model where the ERG considered that the NICE reference case, scope or best practice had not been adhered to)
- Matters of judgement (MJ) (amending the model where the ERG considers that reasonable alternative assumptions are preferred)

6.1.1 ERG base case

Adjustments made by the ERG, to derive the ERG base case (using the CS base case as a starting point) are listed below. Table 6.1 shows how individual adjustments impact the results plus the combined effect of all abovementioned adjustments simultaneously, resulting in the ERG base case. The 'FE' adjustments were combined, and the other ERG analyses were performed also incorporating these 'FE' adjustments given the ERG considered that the 'FE' adjustments corrected unequivocally wrong issues.

Fixing errors

Key issue 9. Health state costs were estimated from an out-of-date source (Section 4.2.9). However, values could not be updated because the calculations are not provided in the model.

1. The FAC identified another error, which was that the model did not allow for different end-of-life utility decrements for the treatment and control arm: this has now also been corrected.

The ERG also noted three errors in the comparison with PEMBRO-CHEMO:

- 50% discount was incorrectly applied to the price of pembrolizumab, which was removed in the comparison with PEMBRO-CHEMO.
- Cells L356 and L366 incorrectly contained functions related to PSA distributions: this was both inappropriate given that they related to unit costs (pembrolizumab and fluorouracil), which are not uncertain and caused an error due to 'NA' instead of values for any distribution parameters. This can be corrected by replacing with references to the cells containing the unit cost values.
- In the PSA only the cost in the first cycle of pembrolizumab is included. This was traced to an error generated in cell L546 in the 'Survival' tab, related to the estimation of time on treatment (ToT). The ERG notes that this error is not generated in the deterministic case, but is related to a function in the VBA, which estimates a random value from a lognormal given that ToT is estimated using the exponential distribution i.e., with one parameter (the rate). This function requires a mean and standard error, but the cell that should contain the standard error is blank. In fact, it appears that the standard errors and covariance matrices for most of the survival distributions are missing in the 'Survival' tab. These errors are therefore not fixable by the ERG.

Fixing violations

None.

Matters of judgement

2. Key issue 5. There is uncertainty as to long term OS and the treatment effect of nivolumab + chemotherapy versus chemotherapy (Sections 3.2 and 4.2.6). Parametric functions (lognormal for PFS, log-logistic for OS) chosen (select 'Scenario: Parametric survival approach' from dropdown menu in cell F20, Model Control).
3. Key issue 5. There is uncertainty as to long term OS and the treatment effect of nivolumab + chemotherapy versus chemotherapy (Sections 3.2 and 4.2.6). Treatment waning (2.5 to 4 years) chosen (put a value of 1 in cell S7, Outcomes Trace).
4. Key issue 8. There is uncertainty as to whether health state utilities should be treatment dependent (Section 4.2.8). Treatment dependent values used (copy values in cells C54:C59 to F54:F59 in Data Library).
5. Key issue 9. There is uncertainty as to the cost of doses of medication received. ERG calculated RDI used (copy values in cells C237:C239 to F237:F239 and C419:C420 to F419:F420 in Data Library).

6.1.2 ERG exploratory scenario analyses

The ERG performed the following exploratory scenario analyses to explore the impact of alternative assumptions conditional on the ERG base case.

Exploratory scenario analyses

6. Key issue 4: There is uncertainty as to the nature and effectiveness of subsequent therapy (Section 3.2, 4.2.2, 4.2.6 and 4.2.9). Choice made according to CheckMate 648 (copy values in cell C448 to F448 and C470 to F47 in Data Library to adjust the cost; copy values in C459 to F459 and C481 to F481 to adjust the discontinuation rate).
7. Key issue 7: There is uncertainty as to how all-cause mortality should be incorporated in the model (Section 4.2.6). Additional all-cause mortality removed (copy values in J15:J116 to H15:H116 in Life Tables).

6.1.3 ERG subgroup analyses

Comparison to PEMBRO-CHEMO is assumed to be in the PD-L1 ≥10% CPS subgroup of the PD-L1 ≥10% TP population.

The ERG has also provided a new base case based on key issue 6: There is uncertainty as to how long-term OS and PFS for the comparison of nivolumab + chemotherapy versus pembrolizumab + chemotherapy. This involves estimating OS for PEMBRO-CHEMO by applying the log-logistic HRs for NIVO-CHEMO versus PEMBRO-CHEMO to the company base case parametric OS curve for NIVO-CHEMO, which is the log-logistic (put a value of 2 in cell S7, Outcomes Trace; select Scenario: 'Alternative comparator - Pembrolizumab (HR approach)' from dropdown menu in cell F20 and 'Parametric – NIVO + CHEMO, OS, Log-logistic' from dropdown menu in cell F32).

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

In Section 6.1 the ERG base case was presented, which was based on various changes compared to the company base case. Table 6.1 shows how individual changes impact the results plus the combined effect of all changes simultaneously. The exploratory scenario analyses are presented in Table 6.3. These are all conditional on the ERG base case. The analyses numbers in Tables 6.2 and 6.3 correspond to the numbers reported in Section 6.1. The submitted model file contains technical details on the analyses performed by the ERG (e.g., the "ERG" sheet provides an overview of the cells that were altered for each adjustment).

Table 6.1: Deterministic/probabilistic ERG base case for NIVO-CHEMO versus CHEMO

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
CS base case					
NIVO-CHEMO	[redacted]	[redacted]	[redacted]	[redacted]	£33,357
CHEMO	[redacted]	[redacted]			
1. Corrected end-of-life utility decrement (no effect on company base case)					
NIVO-CHEMO	[redacted]	[redacted]	[redacted]	[redacted]	£33,357
CHEMO	[redacted]	[redacted]			
2. Parametric functions (lognormal for PFS, log-logistic for OS) chosen					
NIVO-CHEMO	[redacted]	[redacted]	[redacted]	[redacted]	£38,177
CHEMO	[redacted]	[redacted]			
3. Treatment waning (2.5 to 4 years) chosen					
Nivolumab, CHEMO	[redacted]	[redacted]	[redacted]	[redacted]	£39,337
CHEMO	[redacted]	[redacted]			
4. Treatment dependent utility values used					
NIVO-CHEMO	[redacted]	[redacted]	[redacted]	[redacted]	£34,965
CHEMO	[redacted]	[redacted]			
5. Cost of therapy reduced according to RDI calculated by ERG					
NIVO-CHEMO	[redacted]	[redacted]	[redacted]	[redacted]	£35,109
CHEMO	[redacted]	[redacted]			
ERG base case (1-5)					
NIVO-CHEMO	[redacted]	[redacted]	[redacted]	[redacted]	£49,017
CHEMO	[redacted]	[redacted]			
ERG base case (1-5) (probabilistic)					
NIVO-CHEMO	[redacted]	[redacted]	[redacted]	[redacted]	£49,629
CHEMO	[redacted]	[redacted]			

CHEMO = chemotherapy; CS = company submission; ERG = Evidence Review Group; ICER = incremental cost-effectiveness ratio; NIVO = nivolumab; OS = overall survival; PFS = progression-free survival; QALY = quality-adjusted life year; RDI = relative dose intensity

Table 6.2: Deterministic scenario analyses (conditional on ERG base case)

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
ERG base case					
NIVO-CHEMO	[redacted]	[redacted]	[redacted]	[redacted]	£49,017
CHEMO	[redacted]	[redacted]			
6. Subsequent therapy mix from trial data					
NIVO-CHEMO	[redacted]	[redacted]	[redacted]	[redacted]	£65,019
CHEMO	[redacted]	[redacted]			

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
7. Remove additional all-cause mortality					
NIVO-CHEMO	[redacted]	[redacted]	[redacted]	[redacted]	£47,459
CHEMO	[redacted]	[redacted]			
CHEMO = chemotherapy; ERG = Evidence Review Group; ICER = incremental cost-effectiveness ratio; NIVO = nivolumab; QALY = quality-adjusted life year					

Table 6.3: Deterministic/probabilistic ERG base case for NIVO-CHEMO versus PEMBRO-CHEMO

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
Company base case (deterministic)					
NIVO-CHEMO	[redacted]	[redacted]	[redacted]	[redacted]	Dominated
PEMBRO-CHEMO	[redacted]	[redacted]			
Fixing error: removal of 50% PAS discount for pembrolizumab					
NIVO-CHEMO	[redacted]	[redacted]	[redacted]	[redacted]	£307,447 (SW quadrant)
PEMBRO-CHEMO	[redacted]	[redacted]			
ERG base case (log-logistic HRs for NIVOL-CHEMO vs. PEMBRO-CHEMO with log-logistic OS curve for NIVOL-CHEMO) (deterministic)					
NIVO-CHEMO	[redacted]	[redacted]	[redacted]	[redacted]	£290,554 (SW quadrant)
PEMBRO-CHEMO	[redacted]	[redacted]			
ERG base case (probabilistic)					
NIVO-CHEMO	Unable to be run due to an error that could not be fixed.				
PEMBRO-CHEMO					

6.3 ERG's preferred assumptions

The estimated ERG base case ICER for NIVO-CHEMO versus CHEMO, based on the ERG preferred assumptions highlighted in Section 6.1, was £49,017 (deterministic) and £49,629 (probabilistic) per QALY gained. The probabilistic ERG base case analyses indicated cost effectiveness probabilities of 0.0%, 0.8% and 52.2% at WTP thresholds of £20,000, £30,000, and £50,000 per QALY gained.

The estimated ERG base case ICER for NIVO-CHEMO versus PEMBRO-CHEMO, based on the ERG preferred assumptions highlighted in Section 6.1, was £290,554 (SW quadrant) (deterministic). The probabilistic ERG base case analyses could not be calculated due to errors in the model.

6.4 Conclusions of the cost effectiveness section

The company conducted an economic analysis, which was in line with the NICE reference case. In the base case NIVO-CHEMO was compared to CHEMO as in the CheckMate 648 trial i.e., doublet treatment with fluorouracil or capecitabine + cisplatin for the PD-L1 ≥1% TC population as in the NICE scope.[2] In scenario analyses comparisons were made with other types of CHEMO with only the drug costs being changed and PEMBRO-CHEMO in the PD-L1 ≥1% TC and ≥10 CPS population based on an ITC (see Sections 3.3 and 3.4). The ERG considers, based on the FAD for TA737 that the comparison with CHEMO as in CheckMate 648 is appropriate.[1] The ERG also considers that PEMBRO-CHEMO should be included as comparator for PD-L1 ≥1% TC and ≥10 CPS in line with the recommendation of TA737, although only for squamous histology in line with the NICE scope. The ERG does have some concerns regarding the comparability of the KEYNOTE-590 and CheckMate 648 trials (see Section 3.3 and 3.4), particularly given the lack of squamous histology PFS data for the former. However, the ERG is supportive of the ITC methodology, and it is likely that the treatment effect on PFS for PEMBRO-CHEMO might be lower in the mixed histology population. Given that PEMBRO-CHEMO is not a comparator in the PD-L1 ≥1% TC and <10 CPS, the only comparator is CHEMO. The ERG requested that an analysis of CheckMate 648 be performed in this population, but the company refused to do this.[16] Therefore, choice of appropriate comparator as dependent on PD-L1 status is a key issue.

The main concern of the ERG regarding the comparison with CHEMO relates to the modelling of subsequent treatments. The precise nature of subsequent therapy in NHS clinical practice is unknown. In CheckMate 648, of those who received subsequent therapy, [redacted] and [redacted] of NIVO-CHEMO and CHEMO patients received an anti-PD(-L)1, but in the economic model it is assumed that these proportions are zero and 100% respectively.[2] In TA737, the committee acknowledged that this assumption was probably the best reflection of clinical practice.[1] However, this implies that the treatment effect from the trial is liable to be biased upwards because the NIVO-CHEMO patients who received a subsequent anti-PD(-L)1 will have better outcomes and the CHEMO patients who did not receive a subsequent anti-PD(-L)1 will have worse outcomes than would be expected in clinical practice. There are methods for adjusting for treatment switching as set out in TSD 16 and Ouwens 2021 that could reduce this bias, which the ERG recommended, but the company did not employ.[16, 65, 66] Given the likely bias, the nature and effectiveness of subsequent treatment is a key issue.

The company argued that the reducing hazard rate observed in the CHEMO arm is implausible and, on that basis, choose a semi-parametric modelling approach with a cut-off of 6.9 months, using the K-M data before and a parametric model after this cut-off.[2] Little justification is provided for the implausibility, the most plausible explanation appearing to be the effect of subsequent systemic therapy, especially anti-PD(-L)1. However, the most appropriate method of addressing any bias due to this would be to adjust for treatment switching, but only to better reflect clinical practice, is not performed and might actually reduce the treatment effect, as described above. There is also no clear demonstration of lack of fit of parametric models to the OS data and no consideration of more complex spline-based models for PFS. Landmark analysis of CheckMate 648 and parametric OS functions seem to provide reasonable correspondence not only between CheckMate 648 and parametric extrapolation, but also between these and other trial evidence, casting doubt on the implausibility of the reducing hazard rate. Finally, despite the observation of decreasing CHEMO OS hazard and approximation of survival up to year 3 in the trial, the company reject any treatment waning.[2] In TA737 this was considered reasonable for PEMBRO versus CHEMO and the ERG consider that the evidence of treatment waning from CheckMate 648 even earlier is compelling.[1] Given lack of justification of the survival extrapolation methods employed by the company, the observed outcome in CheckMate 648 and its plausibility relative to other trials, OS extrapolation is a key issue.

The main concern regarding the comparison with PEMBRO-CHEMO was in how the HRs estimated using the ITC were used in the economic model. The company stated in the clarification letter that the ITC HRs for pembrolizumab + chemotherapy versus chemotherapy were applied to the survival curves for chemotherapy to estimate the survival curve for pembrolizumab + chemotherapy in the comparison with nivolumab + chemotherapy.[16] However, in the CS the gamma model for the ITC was presented, in the Appendix C the best fitting model was the log-logistic, but in the clarification letter response the company stated that the Weibull and the lognormal were used to be consistent with the base case semi-parametric models.[2, 9, 16] Also, only one set of survival values were presented in the model, which hinders transparency and the nivolumab + chemotherapy and pembrolizumab + chemotherapy OS curves were found to cross, which is inconsistent with the HRs for nivolumab + chemotherapy versus pembrolizumab + chemotherapy, which are all above 1 up to 48 months. Therefore, for the base case the ERG employed the default company parametric model for NIVO-CHEMO OS, which is the log-logistic, and applied the log-logistic HRs for NIVO-CHEMO versus PEMBRO-CHEMO to this OS curve using a method as set out by the company in the response to clarification.[16]

The company chose a progression-based, as opposed to TTD approach to utility estimation.[2] However, a concern of the ERG is that all the health state values, as well as the TTD ones were all higher for chemotherapy than NIVO-CHEMO, albeit by a very small amount, but the company chose the treatment-independent ones from the progression-based as opposed to TTD analysis. Also, the PD-L1 ≥1% values were not used in the model. Despite stating that a progression-based analysis was chosen, it is unclear why an end-of-life decrement was applied, which would seem consistent with a TTD approach. Finally, choice of values for AE decrements was not justified. Therefore, utility estimation is a key issue.

The general approach to costing was appropriate, but adjustments the cost of the intervention and comparator treatments were reduced according to a 'treatment modifier', which lacked face validity and was as opposed to the use of RDI, as in TA737.[1] Also, outdated NHS reference costs were employed with no details of any calculations incorporated in the model.[2] Therefore, these are key issues.

Results showed that the ICER for NIVO-CHEMO was £33,357 post-clarification, only slightly different to in the original CS. Given that life years gained with CHEMO were lower than two years at 1.28 years and the incremental gain was nearly a year, it appears that the EOL criteria might be fulfilled, which would imply that NIVO-CHEMO is cost-effective vs. CHEMO. However, NIVO-CHEMO was dominated by PEMBRO-CHEMO and so it would not be cost-effective in the PD-L1 ≥1% TC and ≥10 CPS population. The ERG base case increased the ICER for NIVO-CHEMO vs. CHEMO, but it remained below £50,000 (£49,017 (deterministic) and £49,629 (probabilistic)): life years gained were also greater than 3 months at [redacted] years. The ICER might be further increased on using the PD-L1 ≥1% TC and <10 CPS population data and if adjusted to better reflect subsequent therapy use in clinical practice. For the comparison with PEMBRO-CHEMO, given that all ITC parametric functions resulted in HRs that favoured PEMBRO-CHEMO and no reason to believe that there was any bias in either subsequent treatments, utilities or costs, the ERG base case only varied in how the OS HRs were incorporated. The ERG also had to make a correction in the model, which was to remove a 50% discount on pembrolizumab, which the company did not mention in any documents in the CS. The ERG base case showed that NIVO-CHEMO was still less effective, but possibly less costly notwithstanding the actual PAS discount for pembrolizumab (see confidential appendix). However, further clarification on the method of implementation of the HRs by the company might be helpful and any errors in the PSA need to be fixed.

7. END-OF-LIFE

The company propose a case to apply NICE end-of-life criteria for NIVO-CHEMO in the treatment of patients with OSCC (Section B.2.13.4.3 of the CS).[2] The details of this are summarised in Table 7.1.

Table 7.1: End-of-life criteria

Criterion	Data available	Reference in submission (Section and page number)
The treatment is indicated for patients with a short life expectancy, normally less than 24 months	Available therapies in patients with untreated, unresectable, advanced, recurrent or metastatic OSCC are associated with poor outcomes, although data describing this patient population are limited. Based on available data, median OS for platinum-based chemotherapy, 5-fluorouracil and cisplatin, as observed during CheckMate 648, was 10.7 months.	B.2.6.3.1
There is sufficient evidence to indicate that the treatment offers an extension to life, normally of at least an additional 3 months, compared with current NHS treatment	The mean OS is more representative of the survival benefit associated with nivolumab with chemotherapy. However, it is acknowledged that extrapolated outputs are subject to uncertainty due to the potential variation in extrapolations. However, when data are restricted to the observed period, restricted mean OS is [redacted] months in the nivolumab with chemotherapy arm and [redacted] months in the chemotherapy arm, providing [redacted] months of survival benefit. Based on model output, mean OS extrapolated over a life-time horizon was [redacted] years in the nivolumab with chemotherapy arm and 1.4 years in the chemotherapy arm (an improvement of [redacted] years). Based on this evidence, it can be concluded that end-of-life criteria are met.	B.2.6.3.1
Based on Table 30 of the CS[2] CS = company submission; NHS = National Health Service; OS = overall survival; OSCC = oesophageal squamous cell carcinoma		

ERG comment: The first criterion seems to be met. In fact, the cited median OS estimate refers to all randomised patients and the value for the subgroup with PD-L1 TP ≥1% is lower at [redacted] months (see Table 3.9). This CheckMate 648 OS estimate for the same doublet chemotherapy is comparable to the estimate of 9.8 months for KEYNOTE-590, which is in a similar population (some UK patients recruited within an international trial and majority 73% have squamous cell histology). Median OS was only slightly longer i.e. 11 months for doublet chemotherapy (fluorouracil + cisplatin) in the study by Lyu et al. in OSCC included by the company in Appendix N.[69, 73] This criterion would also be met if pembrolizumab with chemotherapy was the comparator i.e. in the PD-L1 TP ≥1% and CPS ≥10% subgroup given a median OS of 12.6 months (see Table 3.23).

The second criterion also seems to be met in comparison with chemotherapy with a difference in median OS of 6 months ([redacted] – see Table 3.9). This translates into a gain of [redacted] years after extrapolation in the economic model: although reduced in the ERG base case to [redacted] years, this is still greater than 3 months. However, in comparison to pembrolizumab with chemotherapy for the PD-L1 TP ≥1% and CPS TP ≥10% subgroup, this criterion would almost certainly not be met given that

nivolumab with chemotherapy in the PD-L1 TP ≥1% population had a median OS of [redacted] versus 13.96 for pembrolizumab with chemotherapy in the PD-L1 CPS ≥10% OSCC population, which translates into HRs greater than 1 in the ITC and a loss in LYs gained in the economic model.

8. REFERENCES

[1] National Institute for Health and Care Excellence. *Pembrolizumab with platinum- and fluoropyrimidine-based chemotherapy for untreated advanced oesophageal and gastro-oesophageal junction cancer [TA737]*. London: NICE, 2021 Available from: https://www.nice.org.uk/guidance/ta737/chapter/1-Recommendations

[2] Bristol-Myers Squibb Pharmaceuticals Ltd. *Nivolumab in combination with fluoropyrimidine- and platinum-based combination chemotherapy for the first-line treatment of adult patients with unresectable advanced, recurrent, or metastatic oesophageal squamous cell carcinoma with tumour cell PD-L1 expression ≥1% [ID2712]: Submission to National Institute of Health and Care Excellence. Single technology appraisal (STA): Document B - Company evidence submission*: Bristol-Myers Squibb Pharmaceuticals Ltd., 2022. 173p.

[3] National Institute for Health and Care Excellence. *Nivolumab with platinum-based chemotherapy or ipilimumab for untreated advanced, unresectable or recurrent or squamous cell oesophageal cancer. Final scope*. London: NICE, 2022. 4p. Available from: https://www.nice.org.uk/guidance/gid-ta10572/documents/final-scope

[4] Bristol Myers Squibb. Bristol Myers Squibb receives European Commission approval for Opdivo (nivolumab) with chemotherapy as first-line treatment for patients with unresectable advanced, recurrent or metastatic esophageal squamous cell carcinoma with tumor cell PD-L1 expression ≥ 1%. Press Release 04/05/22 [Internet]. 2022 [accessed 24.5.22]. Available from: https://news.bms.com/news/corporate-financial/2022/Bristol-Myers-Squibb-Receives-European-Commission-Approval-for-Opdivo-nivolumab-with-Chemotherapy-as-First-Line-Treatment-for-Patients-with/default.aspx

[5] Ono Pharmaceutical Co. Ltd. Bristol Myers Squibb receives positive CHMP opinion recommending approval for Opdivo (nivolumab) plus Yervoy (ipilimumab) for first-line treatment of patients with unresectable advanced, recurrent or metastatic esophageal squamous cell carcinoma with tumor cell PD-L1 expression ≥ 1%. Press Release 28/02/22 [Internet]. 2022 [accessed 24.5.22]. Available from: https://www.ono-pharma.com/sites/default/files/en/news/press/enews_Opdivo%2BYervoy1L%20ESC%20CHMP%20positive%20opinionl%20CM-648%2020220228.pdf

[6] Bristol Myers Squibb. *A randomized phase 3 study of nivolumab plus ipilimumab or nivolumab combined with fluorouracil plus cisplatin versus fluorouracil plus cisplatin in subjects with unresectable advanced, recurrent or metastatic previously untreated esophageal squamous cell carcinoma [Data on file]*, 2020

[7] Doki Y, Ajani JA, Kato K, Xu J, Wyrwicz L, Motoyama S, et al. Nivolumab combination therapy in advanced esophageal squamous-cell carcinoma. *N Engl J Med* 2022;386(5):449-462.

[8] Bristol Myers Squibb. *Clinical Protocol CA209648 [Data on file]*, 2020

[9] Bristol-Myers Squibb Pharmaceuticals Ltd. *Nivolumab in combination with fluoropyrimidine- and platinum-based combination chemotherapy for the first-line treatment of adult patients with unresectable advanced, recurrent, or metastatic oesophageal squamous cell carcinoma with tumour cell PD-L1 expression ≥1% [ID2712]: Submission to National Institute of Health and Care Excellence. Single technology appraisal (STA): Appendix C - Summary of product characteristics*: Bristol-Myers Squibb Pharmaceuticals Ltd., 2022. 46p.

[10] National Institute for Health and Care Excellence. *Nivolumab in combination with fluoropyrimidine- and platinum-based combination chemotherapy for the first-line treatment of adult patients with unresectable advanced, recurrent, or metastatic oesophageal squamous cell carcinoma*

with tumour cell PD-L1 expression ≥1% [ID2712]: ERG report – factual accuracy check and confidential information check. London: NICE, 2022. 32p.

[11] Bristol Myers Squibb. *Clinical Protocol CA209648: A randomized phase 3 study of nivolumab plus ipilimumab or nivolumab combined with fluorouracil plus cisplatin versus fluorouracil plus cisplatin in subjects with unresectable advanced, recurrent or metastatic previously untreated esophageal squamous cell carcinoma [Data on file]*, 2020

[12] Bristol-Myers Squibb Pharmaceuticals Ltd. *Nivolumab in combination with fluoropyrimidine- and platinum-based combination chemotherapy for the first-line treatment of adult patients with unresectable advanced, recurrent, or metastatic oesophageal squamous cell carcinoma with tumour cell PD-L1 expression ≥1% [ID2712]: Submission to National Institute of Health and Care Excellence. Single technology appraisal (STA): Appendix E - Systematic literature review of clinical evidence evaluating first-line therapies for patients with advanced or metastatic esophageal cancer*: Bristol-Myers Squibb Pharmaceuticals Ltd., 2022. 64p.

[13] McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS Peer Review of Electronic Search Strategies: 2015 Guideline statement. *J Clin Epidemiol* 2016;75:40-6.

[14] Canadian Agency for Drugs and Technologies in Health. *PRESS - Peer Review of Electronic Search Strategies: 2015 Guideline Explanation and Elaboration (PRESS E&E) [Internet]*. Ottawa: CADTH, 2016 [accessed 22.3.22] Available from: https://www.cadth.ca/resources/finding-evidence/press

[15] National Institute for Health and Care Excellence. *Single Technology Appraisal: company evidence submission template [Internet]*. London: NICE, 2015 [accessed 22.3.22]. 28p. Available from: http://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisals/specification-for-company-submission-of-evidence-2015-version.docx

[16] National Institute for Health and Care Excellence. *Nivolumab in combination with fluoropyrimidine- and platinum-based combination chemotherapy for the first-line treatment of adult patients with unresectable advanced, recurrent, or metastatic oesophageal squamous cell carcinoma with tumour cell PD-L1 expression ≥1% [ID2712]: Response to request for clarification from the ERG*. London: NICE, 2022. 159p.

[17] Centre for Reviews and Dissemination. *Systematic Reviews: CRD's guidance for undertaking reviews in health care [Internet]*. York: University of York, 2009 [accessed 23.3.22] Available from: http://www.york.ac.uk/inst/crd/SysRev/!SSL!/WebHelp/SysRev3.htm

[18] Morrison A, Polisena J, Husereau D, Moulton K, Clark M, Fiander M, et al. The effect of English-language restriction on systematic review-based meta-analyses: a systematic review of empirical studies. *Int J Technol Assess Health Care* 2012;28(2):138-44.

[19] Egger M, Zellweger-Zähner T, Schneider M, Junker C, Lengeler C, Antes G. Language bias in randomised controlled trials published in English and German. *Lancet* 1997;350(9074):326-9.

[20] Lefebvre C, Glanville J, Briscoe S, Littlewood A, Marshall C, Metzendorf MI, et al. Chapter 4: Searching for and selecting studies. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al., editors. *Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019)*: Cochrane, 2019. Available from: http://www.training.cochrane.org/handbook

[21] Bristol-Myers Squibb Pharmaceuticals Ltd. *Nivolumab in combination with fluoropyrimidine- and platinum-based combination chemotherapy for the first-line treatment of adult patients with unresectable advanced, recurrent, or metastatic oesophageal squamous cell carcinoma with tumour cell PD-L1 expression ≥1% [ID2712]: Submission to National Institute of Health and Care*

Excellence. Single technology appraisal (STA): Appendix I - Systematic literature review of the health-related quality of life for advanced esophageal cancer: Bristol-Myers Squibb Pharmaceuticals Ltd., 2022. 39p.

[22] Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, et al. *Cochrane handbook for systematic reviews of interventions version 6.3 (updated February 2022) [Internet]*: Cochrane, 2022 [accessed 4.3.22] Available from: https://training.cochrane.org/handbook

[23] Higgins JPGS. *Cochrane Handbook for Systematic Reviews of Interventions*: Wiley, 2008.

[24] 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.

[25] Sun JM, Shen L, Shah MA, Enzinger P, Adenis A, Doi T, et al. Pembrolizumab plus chemotherapy versus chemotherapy alone for first-line treatment of advanced oesophageal cancer (KEYNOTE-590): a randomised, placebo-controlled, phase 3 study. *Lancet* 2021;398(10302):759-771.

[26] Bleiberg H, Conroy T, Paillot B, Lacave AJ, Blijham G, Jacob JH, et al. Randomised phase II study of cisplatin and 5-fluorouracil (5-FU) versus cisplatin alone in advanced squamous cell oesophageal cancer. *Eur J Cancer* 1997;33(8):1216-1220.

[27] Luo H, Lu J, Bai Y, Mao T, Wang J, Fan Q, et al. Effect of camrelizumab vs placebo added to chemotherapy on survival and progression-free survival in patients with advanced or metastatic esophageal squamous cell carcinoma: the ESCORT-1st randomized clinical trial. *JAMA* 2021;326(10):916-925.

[28] Ezdinli EZ, Gelber R, Desai DV, Falkson G, Moertel CG, Hahn RG. Chemotherapy of advanced esophageal carcinoma: Eastern Cooperative Oncology Group experience. *Cancer* 1980;46(10):2149-2153.

[29] Xu RH, Wang F, Cui C, Yao J, Zhang Y, Wang G, et al. 1373MO JUPITER-06: A randomized, double-blind, phase III study of toripalimab versus placebo in combination with first-line chemotherapy for treatment naive advanced or metastatic esophageal squamous cell carcinoma (ESCC). *Ann Oncol* 2021;32:S1041.

[30] Lee SJ, Kim S, Kim M, Lee J, Park YH, Im Y-H, et al. Capecitabine in combination with either cisplatin or weekly paclitaxel as a first-line treatment for metastatic esophageal squamous cell carcinoma: a randomized phase II study. *BMC Cancer* 2015;15(1):1-7.

[31] Lorenzen S, Schuster T, Porschen R, Al-Batran SE, Hofheinz R, Thuss-Patience P, et al. Cetuximab plus cisplatin–5-fluorouracil versus cisplatin–5-fluorouracil alone in first-line metastatic squamous cell carcinoma of the esophagus: a randomized phase II study of the Arbeitsgemeinschaft Internistische Onkologie. *Ann Oncol* 2009;20(10):1667-1673.

[32] Shen L, Lu ZH, Wang JY, Shu YQ, Kong L, Yang L, et al. LBA52 Sintilimab plus chemotherapy versus chemotherapy as first-line therapy in patients with advanced or metastatic esophageal squamous cell cancer: first results of the phase III ORIENT-15 study. *Ann Oncol* 2021;32:S1330.

[33] Moehler M, Maderer A, Thuss-Patience PC, Brenner B, Meiler J, Ettrich TJ, et al. Cisplatin and 5-fluorouracil with or without epidermal growth factor receptor inhibition panitumumab for patients with non-resectable, advanced or metastatic oesophageal squamous cell cancer: a prospective, open-label, randomised phase III AIO/EORTC trial (POWER). *Ann Oncol* 2020;31(2):228-235.

[34] Wang FH, Wang Y, Sun GP, Chen JH, Lin YC, Liu W, et al. Efficacy and safety of recombinant human lymphotoxin-α derivative with cisplatin and fluorouracil in patients with metastatic esophageal squamous cell carcinoma: a randomized, multicenter, open-label, controlled, phase 2b trial. *Cancer* 2017;123(20):3986-3994.

[35] Yao YW, He YF, Han XH, Ji CS, Hu B. Clinical research of individualized therapy in advanced esophageal cancer based on the ERCC1 C8092A genotype. *Oncol Lett* 2018;16(2):2539-2548.

[36] Bristol Myers Squibb. *CheckMate 648 Oct 2021 Summary data [Data on file]*, 2021

[37] Bristol Myers Squibb. *Primary clinical study report for study CA209648: a randomized phase 3 study of nivolumab plus ipilimumab or nivolumab combined with fluorouracil plus cisplatin versus fluorouracil plus cisplatin in subjects with unresectable advanced, recurrent or metastatic previously untreated esophageal squamous cell carcinoma [Data on file]*, 2021. 275p.

[38] EuroQol. EQ-5D [Internet]. 2022 [accessed 24.5.22]. Available from: https://euroqol.org/eq-5d-instruments/

[39] FACIT.org. FACT-E: Functional Assessment of Cancer Therapy - Esophageal [Internet]. 2022 [accessed 24.5.22]. Available from: https://www.facit.org/measures/FACT-E

[40] Kidane B, Ali A, Sulman J, Wong R, Knox JJ, Darling GE. Health-related quality of life measure distinguishes between low and high clinical T stages in esophageal cancer. *Ann Transl Med* 2018;6(13):270.

[41] Bristol Myers Squibb. *CA209648 BMS-936558 Clinical study report supplemental tables [Data on file]*, 2021

[42] National Institute for Health and Care Excellence. *Nivolumab in combination with ipilimumab for treating advanced melanoma [TA400]*. London: NICE, 2016 Available from: https://www.nice.org.uk/guidance/ta400

[43] National Institute for Health and Care Excellence. *Nivolumab for previously treated advanced renal cell carcinoma [TA417]*. London: NICE, 2016 Available from: https://www.nice.org.uk/guidance/ta417

[44] National Institute for Health and Care Excellence. *Nivolumab for treating advanced (unresectable or metastatic) melanoma [TA384]*. London: NICE, 2016 Available from: https://www.nice.org.uk/guidance/ta384

[45] National Institute for Health and Care Excellence. *Nivolumab for previously treated unresectable advanced or recurrent oesophageal cancer [TA707]*. London: NICE, 2021 Available from: https://www.nice.org.uk/guidance/ta707

[46] Shyamalee J, Baxter MA, Walsh S, Spender L, Adamson DJA, Petty R. 1720P Real-world outcomes for advanced oesophageal squamous cell carcinoma with supportive care alone. *Ann Oncol* 2021;32:S1194.

[47] Jaffe D, DeCongelio M, Gricar J. HSR21-055: Treatment patterns and healthcare resource utilization in first-line therapy for advanced esophageal squamous cell carcinoma in Asian and Western countries. *J Natl Compr Canc Netw* 2021;19(3.5):HSR21-055-HSR21-055.

[48] Cheng S, Qureshi M, Pullenayegum E, Haynes A, Chan KKW. Do patients with reduced or excellent performance status derive the same clinical benefit from novel systemic cancer therapies? A systematic review and meta-analysis. *ESMO Open* 2017;2(4):e000225.

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

[50] Davidson M, Cafferkey C, Goode EF, Kouvelakis K, Hughes D, Reguera P, et al. Survival in advanced esophagogastric adenocarcinoma improves with use of multiple lines of therapy: results from an analysis of more than 500 patients. *Clin Colorectal Cancer* 2018;17(3):223-230.

[51] Janjigian YY, Shitara K, Moehler M, Garrido M, Salman P, Shen L, et al. First-line nivolumab plus chemotherapy versus chemotherapy alone for advanced gastric, gastro-oesophageal junction, and oesophageal adenocarcinoma (CheckMate 649): a randomised, open-label, phase 3 trial. *Lancet* 2021;398(10294):27-40.

[52] Shitara K, Van Cutsem E, Bang Y-J, Fuchs C, Wyrwicz L, Lee K-W, et al. Efficacy and safety of pembrolizumab or pembrolizumab plus chemotherapy vs chemotherapy alone for patients with first-line, advanced gastric cancer: the KEYNOTE-062 phase 3 randomized clinical trial. *JAMA Oncol* 2020;6(10):1571-1580.

[53] National Institute for Health and Care Excellence. *Pembrolizumab with platinum-based chemotherapy for untreated advanced oesophageal cancer [ID3741]. Committee Papers*. London: NICE, 2021 Available from: https://www.nice.org.uk/guidance/ta737/documents/committee-papers

[54] Bristol-Myers Squibb, Ono Pharmaceutical Co Ltd. Efficacy study of nivolumab plus ipilimumab or nivolumab plus chemotherapy against chemotherapy in stomach cancer or stomach/esophagus junction cancer. NCT02872116. In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). 2020 [accessed 30.5.22]. Available from: https://ClinicalTrials.gov/show/NCT02872116

[55] Still E, Jones B. *Nivolumab with platinum-based chemotherapy or ipilimumab for advanced unresectable, recurrent or metastatic previously untreated esophageal squamous cell carcinoma. Advisory Board Report. 16 July 2021 [Data on file]*: Health Economics and Outcomes Research (HEOR) Ltd., 2021. 19p.

[56] Bristol-Myers Squibb Pharmaceuticals Ltd. *Nivolumab in combination with fluoropyrimidine- and platinum-based combination chemotherapy for the first-line treatment of adult patients with unresectable advanced, recurrent, or metastatic oesophageal squamous cell carcinoma with tumour cell PD-L1 expression ≥1% [ID2712]: Submission to National Institute of Health and Care Excellence. Single technology appraisal (STA): Appendix L - Nivolumab-containing regimens for the treatment of first-line advanced or metastatic esophageal cancer - ITC report October 2021 database lock*: Bristol-Myers Squibb Pharmaceuticals Ltd., 2022. 57p.

[57] Bristol-Myers Squibb Pharmaceuticals Ltd. *Nivolumab in combination with fluoropyrimidine- and platinum-based combination chemotherapy for the first-line treatment of adult patients with unresectable advanced, recurrent, or metastatic oesophageal squamous cell carcinoma with tumour cell PD-L1 expression ≥1% [ID2712]: Submission to National Institute of Health and Care Excellence. Single technology appraisal (STA): Appendix L (06.05.22 update) - Network meta-analysis report - October 2021 database lock. Version 1*: Bristol-Myers Squibb Pharmaceuticals Ltd., 2022. 51p.

[58] Guyot P, Ades AE, Ouwens MJ, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. *BMC Med Res Methodol* 2012;12:9.

[59] Cope S, Chan K, Jansen JP. Multivariate network meta-analysis of survival function parameters. *Res Synth Methods* 2020;11(3):443-456.

[60] Bristol-Myers Squibb Pharmaceuticals Ltd. *Nivolumab in combination with fluoropyrimidine- and platinum-based combination chemotherapy for the first-line treatment of adult patients with unresectable advanced, recurrent, or metastatic oesophageal squamous cell carcinoma with tumour cell PD-L1 expression ≥1% [ID2712]: Submission to National Institute of Health and Care Excellence. Single technology appraisal (STA): Sub appendices A, B, C, I to Appendix L (06.05.22 update)*: Bristol-Myers Squibb Pharmaceuticals Ltd., 2022. 269p.

[61] Sun JM, Shen L, Shah MA, Enzinger P, Adenis A, Doi T, et al. Pembrolizumab plus chemotherapy versus chemotherapy alone for first-line treatment of advanced oesophageal cancer (KEYNOTE-590): a randomised, placebo-controlled, phase 3 study. Supplementary online content. *Lancet* 2021;398(10302).

[62] National Institute for Health and Care Excellence. *Nivolumab in combination with fluoropyrimidine- and platinum-based combination chemotherapy for the first-line treatment of adult patients with unresectable advanced, recurrent, or metastatic oesophageal squamous cell carcinoma with tumour cell PD-L1 expression ≥1% [ID2712]: Clarification questions*. London: NICE, 2022. 22p.

[63] Bristol-Myers Squibb Pharmaceuticals Ltd. *Nivolumab in combination with fluoropyrimidine- and platinum-based combination chemotherapy for the first-line treatment of adult patients with unresectable advanced, recurrent, or metastatic oesophageal squamous cell carcinoma with tumour cell PD-L1 expression ≥1% [ID2712]: Submission to National Institute of Health and Care Excellence. Single technology appraisal (STA): Appendix H - Systematic literature review of economic evaluations for the treatment of esophageal cancer*: Bristol-Myers Squibb Pharmaceuticals Ltd., 2022. 35p.

[64] Bristol-Myers Squibb Pharmaceuticals Ltd. *Nivolumab in combination with fluoropyrimidine- and platinum-based combination chemotherapy for the first-line treatment of adult patients with unresectable advanced, recurrent, or metastatic oesophageal squamous cell carcinoma with tumour cell PD-L1 expression ≥1% [ID2712]: Submission to National Institute of Health and Care Excellence. Single technology appraisal (STA): Appendix J*: Bristol-Myers Squibb Pharmaceuticals Ltd., 2022

[65] Latimer NR, Abrams KR. *NICE DSU Technical Support Document 16: Adjusting survival time estimates in the presence of treatment switching*. Sheffield: Decision Support Unit, ScHARR, 2014. 57p. Available from: http://nicedsu.org.uk/wp-content/uploads/2016/03/TSD16_Treatment_Switching.pdf

[66] Ouwens M, Darilay A, Zhang Y, Mukhopadhyay P, Mann H, Ryan J, et al. Assessing the influence of subsequent immunotherapy on overall survival in patients with unresectable stage III non-small cell lung cancer from the PACIFIC study. *Curr Ther Res Clin Exp* 2021;95:100640.

[67] Bristol-Myers Squibb Pharmaceuticals Ltd. *Nivolumab in combination with fluoropyrimidine- and platinum-based combination chemotherapy for the first-line treatment of adult patients with unresectable advanced, recurrent, or metastatic oesophageal squamous cell carcinoma with tumour cell PD-L1 expression ≥1% [ID2712]: Submission to National Institute of Health and Care Excellence. Single technology appraisal (STA): Appendix M1 - Cost-effectiveness of nivolumab in combination with chemotherapy versus chemotherapy for first-line treatment of oesophageal cancer - Technical Report*: Bristol-Myers Squibb Pharmaceuticals Ltd., 2022. 42p.

[68] Bagust A, Beale S. Survival analysis and extrapolation modeling of time-to-event clinical trial data for economic evaluation: an alternative approach. *Med Decis Making* 2014;34(3):343-51.

[69] Bristol-Myers Squibb Pharmaceuticals Ltd. *Nivolumab in combination with fluoropyrimidine- and platinum-based combination chemotherapy for the first-line treatment of adult patients with unresectable advanced, recurrent, or metastatic oesophageal squamous cell carcinoma with tumour cell PD-L1 expression ≥1% [ID2712]: Submission to National Institute of Health and Care Excellence. Single technology appraisal (STA): Appendix N - Survival analysis report*: Bristol-Myers Squibb Pharmaceuticals Ltd., 2022. 128p.

[70] Rutherford MJ, Lambert PC, Sweeting MJ, Pennington R, Crowther MJ, Abrams KR, et al. *NICE DSU Technical Support Document 21: Flexible methods for survival analysis*. Sheffield: Decision Support Unit, ScHARR, 2020. 97p. Available from: http://nicedsu.org.uk/wp-content/uploads/2020/11/NICE-DSU-Flex-Surv-TSD-21_Final_alt_text.pdf

[71] Chau I, Norman AR, Cunningham D, Oates J, Hawkins R, Iveson T, et al. The impact of primary tumour origins in patients with advanced oesophageal, oesophago-gastric junction and gastric adenocarcinoma - individual patient data from 1775 patients in four randomised controlled trials. *Ann Oncol* 2009;20(5):885-91.

[72] Shankaran V, Xiao H, Bertwistle D, Zhang Y, You M, Abraham P, et al. A comparison of real-world treatment patterns and clinical outcomes in patients receiving first-line therapy for unresectable advanced gastric or gastroesophageal junction cancer versus esophageal adenocarcinomas. *Adv Ther* 2021;38(1):707-720.

[73] Lyu J, Li T, Wang Q, Li F, Diao P, Liu L, et al. Outcomes of concurrent chemoradiotherapy versus chemotherapy alone for stage IV esophageal squamous cell carcinoma: a retrospective controlled study. *Radiat Oncol* 2018;13(1):233.

[74] Department of Health and Social Care. Drugs and pharmaceutical electronic market information tool (eMIT). 2020 [accessed 21.2.22]. Available from: https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit

[75] British National Formulary. Nivolumab. 2020. Available from: https://bnf.nice.org.uk/drug/nivolumab.html#indicationsAndDoses

[76] Department of Health. *NHS reference costs 2015 to 2016*, 2016 Available from: https://www.gov.uk/government/publications/nhs-reference-costs-2015-to-2016."

[77] Chau I, et al. Nivolumab plus ipilimumab or nivolumab plus chemotherapy versus chemotherapy as first-line treatment for advanced esophageal squamous cell carcinoma: first results of the CheckMate 648 study. In: ASCO annual meeting, 2021:LBA4001.

[78] Kato K, Cho BC, Takahashi M, Okada M, Lin CY, Chin K, et al. Nivolumab versus chemotherapy in patients with advanced oesophageal squamous cell carcinoma refractory or intolerant to previous chemotherapy (ATTRACTION-3): a multicentre, randomised, open-label, phase 3 trial. *Lancet Oncol* 2019;20(11):1506-1517.

[79] Georghiou T, Bardsley M. Exploring the cost of care at the end of life. 2014. Available from: https://www.nuffieldtrust.org.uk/research/exploring-the-cost-of-care-at-the-end-of-life

[80] NHS England. *NHS reference costs 2019-2020*, 2021 [accessed 17.1.22] Available from: https://www.england.nhs.uk/national-cost-collection

[81] Chauhan R, Lee D. Adrenal insufficiency: burden of disease and cost of illness. *Value Health* 2013;16(7):A436.

[82] Copley-Merriman C, Stevinson K, Liu FX, Wang J, Mauskopf J, Zimovetz EA, et al. Direct costs associated with adverse events of systemic therapies for advanced melanoma: Systematic literature review. *Medicine (Baltim)* 2018;97(31):e11736.

[83] 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(1):36-46.

[84] Kaltenthaler E, Carroll C, Hill-McManus D, Scope A, Holmes M, Rice S, et al. The use of exploratory analyses within the National Institute for Health and Care Excellence single technology appraisal process: an evaluation and qualitative analysis. *Health Technol Assess* 2016;20(26):1-48.

National Institute for Health and Care Excellence

Centre for Health Technology Evaluation

ERG report – factual accuracy check and confidential information check

Nivolumab in combination for untreated advanced unresectable recurrent or metastatic oesophageal squamous cell carcinoma cancer [ID2712]

'*Data owners will be asked to check that confidential information is correctly marked in documents created by others in the technology appraisal process before release; for example, the technical report and ERG report.*' (Section 3.1.29, Guide to the processes of technology appraisals).

You are asked to check the ERG report to ensure there are no factual inaccuracies or errors in the marking of confidential information contained within it. The document should act as a method of detailing any inaccuracies found and how they should be corrected.

If you do identify any factual inaccuracies or errors in the marking of confidential information, you must inform NICE by the end of **17 June 2022** using the below comments table.

All factual errors will be highlighted in a report and presented to the Appraisal Committee and will subsequently be published on the NICE website with the committee papers.

Please underline all confidential information, and separately highlight information that is submitted as 'commercial in confidence' in turquoise, all information submitted as 'academic in confidence' in yellow, and all information submitted as 'depersonalised data' in pink.

General note

BMS would like to clarify aspects of nomenclature relating to PD-L1 testing to ensure consistency in the following document.

PD-L1 testing can be broadly divided into two methods:

- Tumour cell (TC/TPS) method
    - TC scores are obtained by dividing the number of PD-L1 stained tumour cells by the total number of viable tumour cells and multiplying by 100.
    - TC scores are reported as a percentage.
- Combined positive score (CPS) method
    - CPS evaluates the number of PD-L1-stained cells (tumour cells, lymphocytes, macrophages) relative to all viable tumour cells.
    - CPS scores are obtained by dividing the number of PD-L1 stained cells (tumour cells and immune cells) and multiplying by 100.
    - CPS scores are reported as a number.

While KEYNOTE-590 reported only CPS-based subgroups, CheckMate 648 reported primarily TC-based subgroups.

We would like to emphasise that these two methods are based on different measurements and should not be used together. Due to the differences between these methods, the following cases are possible:

- A patient may have a tumour with PD-L1 TC ≥ 1% and no PD-L1 expression on the immune cells, so they would be PD-L1 CPS <10. Therefore, they would be eligible to receive nivolumab, but not eligible to receive pembrolizumab
- A patient may present with a tumour with no tumour cell staining for PD-L1 (PD-L1 TC <1), but have PD-L1 CPS >10 on immune cells. Therefore, this patient would be eligible for treatment with pembrolizumab, but not with nivolumab

Currently, there is no data available to estimate the likelihood of each scenario, but given the methodology of each scoring system, it is entirely possible.

Issue 1 Appropriateness of comparators

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
Section 2.3 page 31 The ERG report states: "…given that pembrolizumab might have replaced chemotherapy in the PD-L1 ≥1% TC and ≥10 CPS population…"	This statement should be deleted or amended to reflect that around half of patients are likely to be receiving pembrolizumab within five years.	This statement does not align with NICE budgetary assumptions. Per the resource impact template provided to support TA737,[1] only 51% of eligible patients will have switched to receiving pembrolizumab-based regimens by year 5 (2026). Further, it is likely that this market share will be gathered over time, so not all 491 patients will be receiving pembrolizumab in 2022. Hence, this statement is currently inaccurate, based on NICE assumptions.	Not a factual inaccuracy – this statement does not imply any particular degree of replacement and any replacement implies, at least for these patients, pembrolizumab and not chemotherapy is the comparator.
Table 1.2, page 16 The ERG report states that: *"The ERG requested that an analysis of CheckMate 648 be performed in this population, but the company refused to do this.*"	The text should be updated to read: The ERG requested that an analysis of CheckMate 648 be performed in this population, *however, this analysis was not possible due to data limitations.*	Despite NICE advice, it is likely that many patients with PD-L1 CPS ≥10 continue to initiate chemotherapy, as discussed above in line with NICE budgetary assumptions.[1] As such, chemotherapy continues to be a relevant comparator regardless of CPS score. In response to the ERG clarification questions, the company investigated the possibility of conducting an analysis versus chemotherapy in the PD-L1 TC ≥1% and CPS <10 population. However, the CheckMate 648 study was not powered for an analysis that would	Not a factual inaccuracy. The ERG would also like to point out that "Providing outcomes for small subgroups" does not run the risk of increasing uncertainty but reduces uncertainty relative to not providing any outcomes for subgroups when those subgroups are relevant to the decision problem.

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
		include fewer patients, as suggested when restricting to patients with PD-L1 TC ≥1% and CPS <10. As outlined in Document B Table 28, only [redacted] patients in the NIVO+CHEMO and [redacted] in the CHEMO arm meet the definition of PD-L1 TC ≥1% and CPS <10. Hence, this reduced population is not expected to be informative. Providing outcomes for small subgroups with wide confidence intervals runs the risk of increasing uncertainty, rather than addressing remaining uncertainty.	
Table 1.2, page 16 The ERG report states that: *"The company acknowledge that the appropriate population is PD-L1 ≥1% TC and ≥10 CPS squamous histology population given that the former is required for nivolumab and the latter for pembrolizumab."*	The text should be updated to read: The company acknowledge that the appropriate population *for the comparison of nivolumab to pembrolizumab* is PD-L1 ≥1% TC and ≥10 CPS squamous histology given that the former is required for nivolumab and the latter for pembrolizumab.	This phrasing is unclear as it suggests that the company believe this to be the relevant population for the appraisal, when the company only acknowledge that this is the relevant population when comparing pembrolizumab to nivolumab due to the different patient populations they are recommended in. When comparing against chemotherapy, the PD-L1 TC ≥1% OSCC population is the appropriate population, as chemotherapy is a comparator for all OSCC patients.	Amended.

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
Table 1.2, page 16 The ERG report states that: *"The ICER versus chemotherapy is likely to go up in the PD-L1 ≥1% TC and <10 CPS population."*	This statement should be amended to: "The ICER impact is uncertain, but it is plausible that the ICER versus chemotherapy is likely to go up in the PD-L1 ≥1% TC and <10 CPS population."	The ERG has noted that the HR for OS for the patient population with PD-L1 CPS <10 is [redacted]. As a result, the ERG suggests that clinical effectiveness for nivolumab will be reduced and ICERs will be increased. However, cost-effectiveness outcomes will be more influenced by the hazard profile in the longer term follow up than by the overall hazard ratio. As such, the impact suggested by the ERG is plausible, but highly uncertain.	Not a factual inaccuracy. The ERG would also point out that, based on the PD-L1 ≥1% TC analysis, there is no evidence to suggest that the pattern of change of the HR over time is likely to be more favourable to nivolumab in the PD-L1 ≥1% TC and <10 CPS than in the main PD-L1 ≥1% TC population.
Table 1.2 page 16 and Section 2.3 page 30 The ERG state: "however, the company argue that it is not SoC because it was recommended too recently (20 October 2021)" and "The company also stated that pembrolizumab was not yet standard of care (SoC) because recommended to recently i.e., October 2021..."	This statement should be amended to reflect the likely timeline for implementation of NICE advice.	It is acknowledged that NICE advice was published in October 2021. However, uptake and implementation of NICE advice is not immediate and it is likely that many patients with PD-L1 CPS ≥10 continue to initiate chemotherapy. As such, chemotherapy continues to be a relevant comparator.	Not a factual inaccuracy – pembrolizumab not replacing chemotherapy for every patient does not imply that it cannot be SoC just as chemotherapy might still also be considered to be SoC for some patients.
Section 2.3, page 31	This sentence should be removed.	It is unclear how the mixed histology population included in KEYNOTE-590 would have impacted the	Not a factual inaccuracy – the ERG made this inference based on the clinical expert

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
The ERG report states that: *"The ERG therefore concludes that the effectiveness of pembrolizumab in the index population i.e., OSCC might have been underestimated in the ITC"*		effectiveness of pembrolizumab. While OSCC patients may appear to be more sensitive to immunotherapies,[2] patients with adenocarcinoma tend to have worse outcomes in response to chemotherapy.[3] Therefore, the difference in outcomes between pembrolizumab and chemotherapy may appear larger than would be observed with an OSCC only population.	opinion expressed in the FAD for TA737, which in the context of the generalisability of the KEYNOTE-590 trial as a whole suggests that they were referring to the treatment effect of pembrolizumab vs. chemotherapy and not only outcomes with pembrolizumab being better in squamous cell carcinoma. This is notwithstanding the uncertainty.
Section 2.3, page 31 The ERG report states that: "*However, although not mentioned in the response to clarification, Table 6 of updated Appendix L – the network meta-analysis (NMA) report shows that the so-called 'overlap analysis' for overall survival (OS) was in the OSCC subgroup, unlike the one for progression-free survival (PFS), which is in the mixed histology population. Therefore, it appears, but is not entirely clear that at least for OS mixed histology is not a problem."*	This sentence should be updated for clarity.	In Section B.2.9.2.1 of the Company Submission, the company states that for PFS (IA), PD-L1 ≥10% (CPS) is only reported for the mixed histology population, however, OS is reported in the OSCC population. Likewise, in this section, the results presented for PFS refer to the mixed histology population, whereas the results for OS are for the OSCC population. Therefore, the analysis was conducted in the relevant OSCC population where this data was available, and was only conducted in the mixed histology population when the OSCC data was not presented.	The ERG was referring to the lack of explicit mention that OS analysis was in the OSCC population. Nevertheless, this has been amended to recognise the clarification provided by the company.

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
Section 2.3, page 31 The ERG report states that: *"However, given that pembrolizumab might have replaced chemotherapy in the PD-L1 ≥1% TC and ≥10 CPS population,"*	This sentence should be removed.	It is unlikely that pembrolizumab will fully replace chemotherapy in the PD-L1 TC ≥1% and CPS ≥10 population. In the pembrolizumab submission, TA737,[2] pembrolizumab is positioned alongside doublet chemotherapy as an alternative therapeutic option, but does not claim that it will replace chemotherapy in this population. It is likely that some patients with PD-L1 ≥10 will still receive chemotherapy.	Not a factual inaccuracy – this statement does not imply any particular degree of replacement and any replacement implies, at least for these patients, pembrolizumab and not chemotherapy is the comparator.
Section 3.3 Page 92 The ERG report states that: "*Document B of the CS does not provide a specific rationale for the use of KEYNOTE-590 for the ITC analyses. The OSCC SLR executed by the company, identified 11 unique studies, apart from CheckMate 648 (see Table 3 of Appendix E[4]). Four studies reported PD-L1 status: CheckMate 648, ORIENT-15[5], ESCORT-1st[6] and KEYNOTE-590. ORIENT-15[5] did not provide population sizes for patients expressing PD-L1. ESCORT-1st provided subgroup data for patients with PD-L1 TPS expression with cut-offs of 1%, 5%, and 10% which facilitates a direct*	This section should be updated.	The ITC analysis was conducted to include pembrolizumab with chemotherapy as a comparator within the cost-effectiveness modelling and therefore, compared pembrolizumab with chemotherapy to the chemotherapy arm of CheckMate 648. While the studies referred to by the ERG report PD-L1 status, ORIENT-15 and ESCORT-1st only consider chemotherapy treatment and do not include a nivolumab or pembrolizumab comparison, which was the aim of the ITC conducted. Additionally, the treatments investigated are not relevant to NHS clinical practice.	This section has been deleted.

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
comparison to the population of CheckMate 648, since the same method was used (TC/TPS). According to the company the intervention (camrelizumab) was not part of current treatment pathways. Nevertheless, it was part of the subsequent systemic therapies received in CheckMate 648. The control arm of ESCORT-1st is an accepted SoC (CIS + PAC). The study was found to be similar to CheckMate 648 in terms of trial phase, histology, and ECOG PS while the main identified difference was that it was conducted at Asian study sites alone. The CS did not clarify why the study was not included in an ITC for clinical effectiveness."			

Issue 2 Indirect treatment comparison

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
Section 1.4, page 17, Table 1.4 The ERG report states that: "*It is unclear which ITC method, constant HR or time varying HRs formed the base case for the analysis*."	This text should be removed.	The NMA report (Updated appendix L, clarification questions) gives the following objective: "*Perform an NMA to estimate the relative treatment effect parameters*	Not a factual inaccuracy – as stated in the ERG report, ."…,Table 6 in the new Appendix L of the CS, which is labelled as an overview of analysis scenarios, refers to the constant HR as the primary analysis".

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
Section 3.6, page 114 The ERG report states that: "*it is unclear whether constant or time varying HRs formed the base case of the analysis*"		*representing the time-varying hazard ratios for PFS and OS*" In addition, the response to question A32 in the clarification letter describes the use of the time-varying hazard ratios to modify the survival models used in the economic model, after the following line: "*In practice, the adjustment of the survival curves by these time-varying hazard ratios was undertaken offline. To do so, the following process was undertaken*"…	Nevertheless, this has been amended to recognise the clarification provided in the FAC by the company.
Section 3.4, page 93-113 The ERG report has been written with respect to results from both the indirect treatment comparison report (Appendix L, company submission) and NMA report (Updated Appendix L and sub-appendices A, B, C and I)	Reference only the results of the NMA report (Updated Appendix L and sub-appendices) in this section	Within Appendix L of the company submission, an analysis was described to inform a time-varying hazard ratio between the arms of KEYNOTE 590 for the OS and PFS outcomes. The methodology of this report was used in informing the indirect treatment comparison between NIVO-CHEMO and PEMBRO-CHEMO, however, the results given within this report were preliminary pending the inclusion of CheckMate 848 within the network. The analysis including CheckMate 648 within the network was	Not a factual inaccuracy. The company seem to have provided no reason to believe that the method and results presented in the original Appendix L are not additional to those presented in the updated Appendix L and essentially part of the same analysis, regardless of whether labelled as ITC or NMA, which the ERG refers to anyway for clarity. To simply refer to one as "ITC" and the other as "NMA" would be misleading as it would

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
		undertaken and results applied in the company submission. The analysis described in Updated Appendix L at clarification questions and sub-appendices A, B, C and I include these results used to inform the cost-effectiveness evaluation between NIVO-CHEMO and PEMBRO-CHEMO. For the purposes of clarity, the company proposes that the original Appendix L describing the methodology be referred to as the "ITC" report, and the updated Appendix L be referred to as the "NMA" report, as the difference between these analyses is the inclusion of multiple studies within the network.	suggest they referred to two different analyses.
Section 2.3, page 31 The ERG report states that: "*However, although not mentioned in the response to clarification, Table 6 of updated Appendix L – the network meta-analysis (NMA) report shows that the so-called 'overlap analysis' for overall survival (OS) was in the OSCC subgroup, unlike the one for progression-free survival (PFS), which is in the mixed histology population.*	This sentence should be updated for clarity.	In Section B.2.9.2.1 of the Company Submission, the company states that for PFS (IA), "PD-L1 ≥10% (CPS)" [sic: See general note] is only reported for the mixed histology population, however, OS is reported in the OSCC population. Likewise, in this section, the results presented for PFS refer to the mixed histology population, whereas the results for OS are for the OSCC population. Therefore, the analysis was	The ERG was referring to the lack of explicit mention that OS analysis was in the OSCC population. Nevertheless, this has been amended to recognise the clarification provided by the company.

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
Therefore, it appears, but is not entirely clear that at least for OS mixed histology is not a problem."		conducted in the relevant OSCC population where this data was available, and was only conducted in the mixed histology population when the OSCC data was not presented. Impact: No impact	
Section 2.3, page 31 The ERG report states that: *"The ERG therefore concludes that the effectiveness of pembrolizumab in the index population i.e., OSCC might have been underestimated in the ITC"*	This sentence should be clarified.	It is unclear how the mixed histology population included in KEYNOTE-590 would have impacted the effectiveness of pembrolizumab. While OSCC patients may appear to be more sensitive to immunotherapies,[2] patients with adenocarcinoma tend to have worse outcomes in response to chemotherapy.[3] Therefore, the difference in outcomes between pembrolizumab and chemotherapy may appear larger than would be observed with an OSCC only population. Impact: No impact	Not a factual inaccuracy – the ERG made this inference based on the clinical expert opinion expressed in the FAD for TA737, which in the context of the generalisability of the KEYNOTE-590 trial as a whole suggests that they were referring to the treatment effect of pembrolizumab vs. chemotherapy and not only outcomes with pembrolizumab being better in squamous cell carcinoma. This is notwithstanding the uncertainty.
Section 3.3 Page 92 The ERG report states that: "*Document B of the CS does not provide a specific rationale for the use of KEYNOTE-590 for the ITC analyses. The OSCC SLR executed by the company, identified 11 unique studies, apart*	This section should be updated.	The ITC analysis was conducted to include pembrolizumab with chemotherapy as a comparator within the cost-effectiveness modelling and therefore, compared pembrolizumab with chemotherapy to the chemotherapy arm of	This section has been deleted.

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
from CheckMate 648 (see Table 3 of Appendix E[4]). Four studies reported PD-L1 status: CheckMate 648, ORIENT-15[5], ESCORT-1st[6] and KEYNOTE-590. ORIENT-15[5] did not provide population sizes for patients expressing PD-L1. ESCORT-1st provided subgroup data for patients with PD-L1 TPS expression with cut-offs of 1%, 5%, and 10% which facilitates a direct comparison to the population of CheckMate 648, since the same method was used (TC/TPS). According to the company the intervention (camrelizumab) was not part of current treatment pathways. Nevertheless, it was part of the subsequent systemic therapies received in CheckMate 648. The control arm of ESCORT-1st is an accepted SoC (CIS + PAC). The study was found to be similar to CheckMate 648 in terms of trial phase, histology, and ECOG PS while the main identified difference was that it was conducted at Asian study sites alone. The CS did not clarify why the study was not included in an ITC for clinical effectiveness."		CheckMate 648. While the studies referred to by the ERG report PD-L1 status, ORIENT-15 and ESCORT-1st only consider chemotherapy treatment and do not include a nivolumab or pembrolizumab comparison, which was the aim of the ITC conducted. Impact: No impact	

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
Section 3.4.3, page 111 The ERG report states that: *"There is a contradiction between what is reported in the new Appendix L and the original CS in that, whereas the original CS states that the ITC method was that to estimate time varying HRs, Table 6 in the new Appendix L of the CS, which is labelled as an overview of analysis scenarios, refers to the constant HR as the primary analysis."*	This text should be removed.	Whilst the company acknowledges that the report section labels are not aligned, the primary analysis referred to is the constant HR analysis *and* the best fitting model, where both the constant and time-varying analyses were co-primary outcomes, with the constant HR being presented as contextual information and for use as an optional modelling parameter. The time-varying analysis is presented with the intention of fulfilling the analytical objective of "perform[ing] an NMA to estimate the relative treatment effect parameters representing time-varying hazard ratios for PFS and OS". There is no contradiction in intent between these documents. Impact: No further impact	Not a factual inaccuracy. Nevertheless, this has been amended to recognise the clarification provided by the company.
Section 3.6, page 114 The ERG report states that: *"A rationale for the use of KEYNOTE-590 in the ITC analysis was not provided*"	This text should be removed.	The NMA report introduces the purpose of the ITC as "inform[ing] a key comparator of interest within the CEM, particularly pembrolizumab + chemotherapy." As KEYNOTE 590 was the only trial evaluating pembrolizumab plus chemotherapy in this indication identified by SLR, its rationale for use is implicit and limitations discussed extensively in	Not a factual inaccuracy – the ERG could not find an explicit statement to justify its inclusion. Nevertheless, this has been amended to recognise the clarification provided by the company.

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
		the submission and clarification documents. Impact: No further impact	
Section 3.6, page 114 The ERG report states that: "*Regarding PD-L1 expression, KEYNOTE-590 only provided CPS ≥10% data and as such this population was used in the NMA for both studies*"	This should be amended to read: *Regarding PD-L1 expression, KEYNOTE-590 only provided outcomes data for CPS ≥10 and CPS <10, and as such the CPS ≥10 population was used in the NMA for both studies.*	KEYNOTE 590 did not provide full data (e.g. baseline characteristics) disaggregated per CPS {≥|<}10 subgroup. All assessments of trial comparability were necessarily upon the ITT population within KEYNOTE 590. Impact: No impact	Not a factual inaccuracy – this was in contrast to TC ≥1%: amended to improve clarity.
Section 3.6, page 114 The ERG report states that: "*Serious limitations of the NMA were acknowledged by the company regarding study design and population comparability*"	The word "serious" should be removed	The company described the limitations of the analysis in the NMA report (Updated appendix L, clarification questions). It did not characterise these limitations as "serious". Impact: No impact	Amended accordingly.

Issue 3 Impact of subsequent therapies

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
Table 1.5 page 18 The ERG state that the likely outcome of adjusting the CheckMate 648 data to reflect subsequent treatment would be that "The ICER versus CHEMO is likely to increase."	This statement should be amended to reflect the uncertainty in amending this analysis.	During CheckMate 648, subsequent treatment, and subsequent PD-L1 use, was more prevalent in the CHEMO arm. Adjusting for PD-L1 usage would likely have a larger impact on the clinical outcomes for CHEMO than for the NIVO+CHEMO arm. Even if modelling separate second-line nivolumab usage, using data from ATTRACTION-3, may result in lower survival outcomes; although this study demonstrated a benefit for nivolumab, median survival remained short in the nivolumab arm (10.9 months).[7] This impact on comparative effectiveness may be ameliorated by reduced costs in the CHEMO arm as a result of changes in assumed second-line nivolumab usage. However, as a result, overall impact on the ICER would be difficult to predict.	Not a factual inaccuracy.
Table 1.5 Page 19 The suggested analysis from the ERG will have an impact on the feasibility of an ITC and the plausibility of ITC outcomes.	This section should include a reflection on the impact of an ITC versus pembrolizumab.	The suggested analysis would amend outcomes for the NIVO+CHEMO versus CHEMO population. However, no associated amendment is possible for the pembrolizumab-treated population.	Not a factual inaccuracy. The company make a reasonable point that any mismatch between subsequent therapy use in KEYNOTE-590 and NHS clinical practice might

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
		Further, KEYNOTE-590 was undertaken during a slightly different time period to CheckMate 648, so that the impact of PD-L1 inhibitors would be slightly difference. Hence, this analysis would have an impact on feasibility and plausibility of outcomes.	also have implications for outcomes with pembrolizumab. However, any lack of discussion by the ERG of this does not mean that what the ERG have written about the comparison with chemotherapy is incorrect. The ERG responded to the data on subsequent therapy use in the CheckMate-648 trial, which suggested that there could be a substantial discrepancy versus NHS clinical practice, which could imply an efficacy bias, which could be adjusted for using the methods recommended by the ERG. These data were not available to the ERG for pembrolizumab, and it would seem that adjustment would be limited, although the method by Ouwens 2021 does seem to suggest that a treatment effect for subsequent immunotherapy might be estimated from CheckMate 648 and applied to KEYNOTE-590.
Table 1.5 page 18	The sentence should read:	Of the 14 patients in the NIVO+CHEMO arm who received subsequent PD-L1 inhibitors during	The ERG checked the CSR and note that the figures in Table S.10.B1, which is the

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
The ERG state: "In CheckMate 648, of those who received subsequent therapy, [redacted] and [redacted] of NIVO+CHEMO and CHEMO patients received an anti-PD(-L)1"	"In CheckMate 648, of those who received subsequent therapy, [redacted] and [redacted] of NIVO+CHEMO and CHEMO patients received an anti-PD(-L)1"	CheckMate 648, 12 received subsequent nivolumab. The values are corrected based on Table S.10.B2 (page 2085) of the CSR supplementary tables.	source used for Table 4.4, are [redacted] and [redacted]. The company also erroneously cite the figures of 14 and 12, which are actually reported for the nivolumab + ipilimumab arm in Table S.10.B.2. Therefore, the ERG presume that [redacted] and [redacted] are the correct figures for the latest DBL (presumably [redacted], as provided in the clarification letter response. However, this still requires clarification and the provision of an updated CSR for the October[redacted] DBL.
Section 4.2.2, page 121 The ERG report describes in text: "*As can be seen in Table 4.4, in the CheckMate 648 trial, patients received a range of non- anti-PD1 systemic anticancer therapies*" And the caption of Table 4.4 reads: "*Table 4.4: Subsequent cancer therapy, CheckMate 648 trial*"	It should be clarified that these data refer to the PD-L1 expressing subgroup of CheckMate 648	The source tables within the CSR supplement refer to the PD-L1 subgroup. This is the relevant subgroup, and the error is solely in labelling.	Amended.
Section 4.2.2, page 122 The ERG report states:	This should be re-worded to remove the implication of mis-interpretation.	The relevant clarification question was B2(g), which reads as follows:	Not a factual inaccuracy and if the question had been unclear, clarification would

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
"*However, no adjustment for switching was performed by the company, who appeared to misinterpret the question as being about unplanned pre-progression switching even though it was a sub question of one where subsequent therapy was the subject and to which the company responded accordingly in every other sub question*"		"*Please conduct an analysis of OS and PFS in both arms of CheckMate 648 adjusting for switching to anti-PD-1/PD-L1 therapies by reference to TSD 16.*" This question is unclear as to the direction of adjustment, and the most common purpose of treatment switching analysis, of the type described in TSD16, is to compensate for switching from the control treatment to the investigational treatment in crossover trials, which was the case discussed. The company now understands the intent of the question was to undertake a switching analysis in the opposite direction, but this could not be inferred from the question as asked.	have been provided if requested, including at during the Clarification telcon.

Issue 4 Survival analysis

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
Table 1.6 Page 19 The ERG state: "Little justification is provided for the implausibility, the most plausible explanation appearing	This statement should be removed or amended to reflect that the evidence for second-line PD-L1 inhibitors would not provide a plausible reason for several of the optimistic long-term extrapolations.	Nivolumab is the only PD-L1 inhibitor with a marketing authorisation for second-line OSCC. Further, during CheckMate 648, nivolumab was by far the most commonly used subsequent PD-L1	Not a factual inaccuracy – this is a matter of opinion.

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
to be the effect of subsequent systemic therapy, especially anti-PD(-L)1."		inhibitor. As such, nivolumab clinical effectiveness can be considered highly representative of the impact of PD-L1 inhibitors as a subsequent treatment in OSCC. As outlined in the survival analysis report, several of the CHEMO extrapolation models predicted long-term extrapolations with limited or zero disease-related hazard to the end of the time horizon. During ATTRACTION-3, only 20% of patients receiving nivolumab remained alive two years after first dose of study therapy. Based on this evidence, even if all potential patients in the CHEMO arm had received a PD-L1 inhibitor during CheckMate 648, a larger disease-related hazard would have been plausible. Based on the limited usage ([redacted]), predications of very low long term hazard above the assumed lifetable baseline can be considered implausible.	
Table 1.6 Page 19 The ERG state: "Despite the observation of decreasing CHEMO OS hazard and approximation of survival up	This statement should be amended to reflect the long-term predictions from survival curves selected by both BMS and the ERG.	Survival curves selected by BMS and by the ERG reflect a reduced benefit for NIVO+CHEMO versus CHEMO over time. Introduction of scenarios where treatment waning is applied on top of these survival	Not a factual inaccuracy – it is true that some survival curve selection effectively cause implicit waning, but this is not the same as a consideration of and implementation of the

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
to year 3 in the trial, the company reject any treatment waning"		curves can be considered "double counting" of any potential effect.	most plausible timing and degree of waning.
Table 1.7, page 19 The ERG state: "There is uncertainty as to how long-term OS and PFS for the comparison of nivolumab + chemotherapy versus pembrolizumab + chemotherapy"	There is a missing word or words in this statement.	This statement is unclear.	Amended.
Section 4.2.6, page 135 The ERG report states that: "I*t is unclear why the Weibull was considered to have a poor visual fit and so the ERG prefers this to the generalised gamma for NIVO-CHEMO given its better statistical fit*"	The word "better" should be replaced by "similar" or "more parsimonious".	The generalised gamma model, having an additional parameter versus the Weibull model, has an AIC less than 2 points greater than the Weibull, implying a greater log-likelihood. Thus, the statistical fit of the generalised gamma is better. The difference in AIC is also not great enough to suspect over-fitting by the generalised gamma model.	Not a factual inaccuracy – the term 'statistical fit' relates to the statistic regardless of the relative contribution to it of precision and parsimony.
Section 4.2.6.3, page 137 The ERG report states that: "*The ERG therefore chose to remove the additional morality by setting the lifetable values to zero in the ERG base case.*"	Consider the impact of this change on PSA samples	The change made by the ERG has limited impact in the base case, although a very small fraction of patients do survive to age 100 on both arms under their base case. However, it is possible that the model will predict implausible high survival when sampling survival model parameters under PSA.	Not a factual inaccuracy – it is also unclear how this scenario could lead to implausible high survival, but this could be addressed by implementing lifetable values as limit instead of adding them.

Issue 5 Health state utilities

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
Key issue 8, section 1.5, Pages 21-22 Within the discussion section of key issue 8, the ERG recommend the following approach "Use of treatment specific utilities including for terminal care decrement." The ERG also imply that the utility values should be sourced from the PD-L1 positive population from the utility analysis through the following statement "Also, the PD-L 1≥1% values were not used in the model." There are no problems with this alternative approach. However, the application of the approach in the CEM is incorrect in two ways: 1. The application of the approach when calculating the end of life utility decrement. The end of life utility decrement is calculated as the utility values in the post-progression disease state minus the utility value is the 30 days before death occurs (time-to-death of 0-30 days). This is carried out in the ERG CEM, however, the	The end of life utility decrement has been updated in the CEM so that the post-progression health state utility and the time to death utility are sourced from the same population (PD-L1 population) from the utility analysis, such that: • NIVO-CHEMO end of life utility decrement is now 0.138 (0.672 – 0.534) • CHEMO end of life utility decrement is now 0.182 (0.694 – 0.512) The SE values have been updates in the CEM have been updated based on the new health state utility values, such that: • NIVO-CHEMO pre-progression SE is now 0.017 • NIVO-CHEMO post-progression SE is now 0.020 • NIVO-CHEMO end of life utility decrement SE is now 0.014 (0.138/10) • CHEMO pre-progression SE is now 0.017 • CHEMO post-progression SE is now 0.022 • CHEMO end of life utility decrement SE is now 0.018 (0.182/10)	Correct of health state utility values, and corresponding SE, employed in the ERG CEM.	Corrected (including the error in the model in order to allow for different end-of-life utility decrements for the treatment and control arm).

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
post-progression health state utility value is sourced from the PD-L1 positive population (NIVO+CHEMO: 0.672; CHEMO:0.694), and the utility value in the 30 days before death is sourced from the All Randomized Population (NIVO+CHEMO: 0.484; CHEMO:0.517) – this lacks consistency. Accordingly, the end-of-life utility decrements employed in the model for both NIVO-CHEMO (0.188) and CHEMO (0.177) are incorrect. 2. There has been no consideration for how changing the health state utility values will impact the SE values. Accordingly, the SE values employed in the CEM for the health state utility values are incorrect. Note: An issue was identified in the CEM, in that the model doesn't allow for different end-of-life utility decrements for the treatment and control arm (the control arm end-of-life utility decrement is always	Note: The issue identified in the CEM that doesn't allow for different end-of-life utility decrements for the treatment and control arm has been corrected to generate these results.		

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
equivalent to that of the treatment arm).			

Issue 6 Economic modelling

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
Section 6.2, page 177 The results for scenario 3 (treatment-dependent utility values used), presented in table 6.1, are incorrect. This results from the incorrect end-of-life utility decrements employed in the CEM (as discussed in key issue 5). Accordingly, the total QALYs (NIVO-CHEMO: [redacted]; CHEMO: [redacted]; incremental: [redacted]), and the resultant ICER (£35,124) are incorrect.	The CEM has been corrected to the correct end-of-life utility decrements. Accordingly, the results in Table 6.1, scenario 3, need to be corrected, specifically: • The ICER needs to be changed to £35,165 • The total QALYs for NIVO-CHEMO need to be changed to [redacted] • The total QALYs for CHEMO need to be changed to [redacted] • The incremental QALYs need to be changed to [redacted] Note: The issue identified in the CEM that doesn't allow for different end-of-life utility decrements for the treatment and control arm has been corrected to generate these results.	Correction of the scenario 3 in Table 6.1.	Corrected, although the company appear to have made a mistake: • The ICER needs to be changed to £35,965 • The total QALYs for NIVO-CHEMO need to be changed to [redacted] • The total QALYs for CHEMO need to be changed to [redacted] • The incremental QALYs need to be changed to [redacted]
Section 6.2, page 177 The results for scenario 4 (Cost of therapy reduced according to RDI calculated by ERG), presented in Table 6.1, are incorrect.	The CEM has been corrected to the correct RDI values. Accordingly, the results in Table 6.1, scenario 4, need to be corrected, specifically:	Correction of the scenario 4 in Table 6.1.	Corrected.

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
Specifically, the weighting has been performed by the column immediately to the left in the calculation sheet, rather than with reference to the "central" RDI values (i.e. column "A") This error results in the total costs (NIVO-CHEMO: £[redacted]; CHEMO: £[redacted]; incremental: £[redacted]) being incorrect. Accordingly, ICER (£33,479) is incorrect.	• The ICER needs to be changed to £35,109. • The total costs for NIVO-CHEMO need to be changed to £[redacted] • The total costs for CHEMO need to be changed to £[redacted] • The incremental costs need to be changed to £[redacted]		
Section 6.2, page 177 The results for the ERG base case, presented in Table 6.1, are incorrect. This is a result of the calculation error in the RDI values employed in the model. This results in the total costs (NIVO-CHEMO: £[redacted]; CHEMO: £[redacted]; incremental: £[redacted]) being incorrect. Moreover, given issue 5, the incorrect end-of-life utility decrement employed in the CEM, the total QALYs (NIVO-CHEMO: [redacted]; CHEMO: [redacted]; incremental: [redacted]) are incorrect. Accordingly, the ICER (£46,859) is incorrect.	The CEM has been corrected to the correct RDI values and end-of-life utility decrements. Accordingly, the results in Table 6.2, scenario 6, need to be corrected, specifically: • The ICER needs to be changed to £[redacted] • The total costs for NIVO-CHEMO need to be changed to £[redacted] • The total costs for CHEMO need to be changed to £[redacted] • The incremental costs need to be changed to £[redacted]	Correction of the ERG base case in Table 6.1.	Corrected.

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
	• The total QALYs for NIVO-CHEMO need to be changed to [redacted] • The total QALYs for CHEMO need to be changed to [redacted] • The incremental QALYs need to be changed to [redacted] Note: The issue identified in the CEM that doesn't allow for different end-of-life utility decrements for the treatment and control arm has been corrected to generate these results.		
Section 6.2, page 177 The results for scenario 6 (the removal of additional all-cause mortality), presented in Table 6.2, are incorrect. This is a result of the calculation error in the RDI values employed in the model. This results in the total costs (NIVO-CHEMO: £[redacted]; CHEMO: £[redacted]; incremental: £[redacted]) being incorrect. Moreover, given issue 5, the incorrect end-of-life utility decrement employed in the CEM, the total QALYs (NIVO-CHEMO: [redacted]; CHEMO: [redacted]; incremental: [redacted]) are incorrect. Accordingly, the ICER (£45,190) is incorrect.	The CEM has been corrected to the correct RDI values and end-of-life utility decrements. Accordingly, the results in Table 6.2, scenario 6, need to be corrected, specifically: • The ICER needs to be changed to £[redacted] • The total costs for NIVO-CHEMO need to be changed to £[redacted] • The total costs for CHEMO need to be changed to £[redacted] • The incremental costs need to be changed to £[redacted]	Correction of scenario 6 results in Table 6.2.	Corrected.

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
	• The total QALYs for NIVO-CHEMO need to be changed to [redacted] • The total QALYs for CHEMO need to be changed to [redacted] • The incremental QALYs need to be changed to [redacted] Note: The issue identified in the CEM that doesn't allow for different end-of-life utility decrements for the treatment and control arm has been corrected to generate these results.		
Section 6.2, page 177 Subsequent therapy mix from trial data is calculated incorrectly, meaning the ICER reported in Table 6.2 for scenario 5 is incorrect. Two issues exist: • The total number of patients who receive subsequent therapy and the number of patients who receive anti-PD-L1 therapy as a subsequent therapy is sourced from an outdated database lock. • When calculating the number of patients who receive anti-PD-L1 as a subsequent therapy,	Souring patient numbers from the most recent database lock and removing double-counting of nivolumab when calculating the number of patients who receive anti-PD-L1 as a subsequent therapy results in the following second-line treatment costs: • Intervention: £200.37 • Control: £289.43 Weighting subsequent therapy discontinuation rate based on the number of patients receiving anti-PD-L1 as a subsequent therapy and those who do not receive anti-PD-L1 as a subsequent therapy results in the following second-line cyclical discontinuation rate: • Intervention: 0.06016 • Control: 0.05965 Implementing these updates, and correcting the RDI calculation error and the incorrect end-	Correction of scenario 5 results in Table 6.2.	Given the lack of calculations in the Workbook, the ERG had to try to back-calculate the original cost unweighted by subsequent therapy use. This was done by taking the values for weighted cost and % subsequent treatment (see ERG report Tables 4.17 (from clarification letter response) and 4.19 (from Table 51, CS). The % subsequent treatment values ([redacted]) are the ones provided by the company in response to clarification as being from the latest DBL. The ERG erroneously double counted by adding the % PD-L1 and nivolumab and used the original %PD-L1

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
nivolumab is double-counted, resulting in an inflated value. Moreover, there is no consideration of the impact on the subsequent therapy weekly discontinuation rate when using a subsequent therapy mix. Due to the above errors, and the calculation error in the RDI values employed in the model, the total costs in Table 6.2 for scenario 5 (NIVO-CHEMO: £[redacted]; CHEMO: £[redacted]; incremental: £[redacted]) are incorrect. Moreover, given issue 5, the incorrect end-of-life utility decrement employed in the CEM, the total QALYs (NIVO-CHEMO: [redacted]; CHEMO: [redacted]; incremental: [redacted]) are incorrect. Accordingly, ICER (£59,951) is incorrect.	of-life utility decrements, results in a change to the ICER. Accordingly, the results in Table 6.2, scenario 5, need to be corrected, specifically: • The ICER needs to be changed to £65,468 • The total costs for NIVO-CHEMO need to be changed to £[redacted] • The total costs for CHEMO need to be changed to £[redacted] • The incremental costs need to be changed to £[redacted] • The total QALYs for NIVO-CHEMO need to be changed to [redacted] • The total QALYs for CHEMO need to be changed to [redacted] • The incremental QALYs need to be changed to [redacted] Note: The issue identified in the CEM that doesn't allow for different end-of-life utility decrements for the treatment and control arm has been corrected to generate these results.		from the CS: this has been corrected with the latest DBL figures for PD-L1 ([redacted], [redacted]) from the response to clarification letter. The new weighted cost values are very similar to the ones given by the company: • Intervention: £200.86 • Control: £289.53 Because of the similarity and because the company presumably have access to the original data in order to perform their calculations, the ERG have replaced the values they have calculated with the ones calculated by the company. The ERG has also been able to reproduce the figures for discontinuation rate: • Intervention: 0.06016 • Control: 0.05965 Therefore, based on these figures provided by the company the ERG have re-run this scenario.

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
			Unfortunately, there is a discrepancy between the company and ERG results. As already referred to in Key issue 10, this highlights the need for all calculations to be provided by the company in the model.
Section 6.2, page 177 The probabilistic ERG base case does not consider uncertainty around the parametric survival curves. Moreover, the wrong end-of-life utility decrement, and the wrong SE values for the health state utilities and end-of-life utility, are employed in the model (as discussed in issue 5). Furthermore, there is a calculation error in the RDI values employed in the model. This resulting in the estimated QALYs and costs for NIVO-CHEMO and CHEMO, and the resultant ICER, presented in table 6.1 for the probabilistic ERG base case being outputted incorrectly. Specifically, the total costs for the probabilistic ERG base case (NIVO-CHEMO: [redacted]; CHEMO:	• The results of the ERG base case probabilistic results presented in Table 6.1 need to be corrected	BMS has identified multiple errors in the ERG's modelling approach that impact the PSA; although the PSA results will need to be corrected, there was insufficient time to provide this analysis	The ERG does not understand what is meant by: “The probabilistic ERG base case does not consider uncertainty around the parametric survival curves.” It is the case, as identified in the ERG report, that the model omitted data required for estimating uncertainty pertaining to the survival curves, but this would affect any analysis including the ones by the company.

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
[redacted] incremental [redacted]), the total QALYs (NIVO-CHEMO: [redacted]; CHEMO: [redacted] incremental: [redacted]) and the resultant ICER [redacted]) are incorrect.			
Section 4.2.9.7, page 148 The ERG report states that: *"ERG comment: In the economic model, AE management costs appear to have been applied in cycle 1 as a one-off cost of £171,524."*	This sentence should be updated to say: "ERG comment: In the economic model, AE management costs appear to have been applied in cycle 1 as a one-off cost of £171,524 *for NIVO-CHEMO*."	This statement is only true for the treatment arm (NIVO + CHEMO). The control arms have different AE profiles (different incident rates for all of the modelled AEs), resulting in a different one-off cost applied in the first cycle.	Amended.
Section 4.2.9.4 , page 145 - 146 In Table 4.19, the second-line treatment frequency of Taxane: docetaxel and Taxane: Paclitaxel, when either NIVO-CHEMO or PEMBRO-CHEMO is a first-line treatment, is outdated.	The values described in Table 4.19 should be amended to the most up-to-date values. That is changing the values from 24.7% to 26.58%.	Provides a clear understanding of how the values in the column titled 'Second-line weighted average cyclical cost' are derived.	Corrected.
Section 4.2.9.4 , page 145 The ERG report states that: *"In the economic model, the assumption that patients who discontinue their initial therapy will*	The sentence should be amended to say: "…will consequently be eligible to receive a subsequent therapy, was applied"	This statement implies that all patients who discontinue first-line therapy receive a second-line treatment; this is incorrect as not all patients receive a subsequent therapy. Instead, the model adopts	Amended.

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
consequently receive a subsequent therapy, was applied."		the assumption that patients who discontinue their initial therapy will consequently be eligible to receive subsequent therapy.	
Section 4.2.9.3, page 144 The ERG report states that: *"It is also not clear that assuming the same value for PEMBRO-CHEMO is conservative: relative dose intensity (RDI) was estimated in TA737 and so it would be better to use these values"*	The sentence should be removed.	The relative dose intensity values used in TA737[2] are redacted and so cannot be used to inform the PEMBRO+CHEMO treatment modifier. Moreover, a similar approach for assuming equivalent dose intensity (or RDI) between the treatment and control arm is adopted in the TA737[2] where data gaps exist. Specifically, oxaliplatin is assumed to have equivalent RDI to cisplatin, and capecitabine and epirubicin are assumed to have equivalent RDI 5-FU. Accordingly, assuming an equivalent dose intensity between NIVO + CHEMO and PEMBRO + CHEMO is the most conservative approach available.	Amended.
Section 4.2.2, page 118 The ERG report states that: *"As a simplifying assumption, it is assumed that patients may not discontinue this final line of therapy."*	The sentence should be amended to say: "Patients discontinue second-line therapy based on an average cyclical discontinuation rate".	In order to account for the fact that patients do not remain on second-line treatment indefinitely, and so avoiding the implausible accrual of second-line treatment costs, an average weekly cyclical discontinuation rate, derived from the average time on treatment value sourced from TA707,[8] moving	Not a factual inaccuracy – this is what was stated in the CS. The ERG report goes on to quote the clarification response, which states that this statement was erroneous and that patients do discontinue subsequent therapy. However, a

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
		patients on to no treatment, is employed in the model.	sentence has been added to provide clarification as to the source and values.
Table 1.11 page 21 The ERG report states: "Calculations were missing from the model, which reduces transparency and makes updating difficult"	This section should be clarified	Cost calculations were provided to the ERG, documented in the technical report and Document B. It is acknowledged that these calculations were not integrated into the modelling, which did not support the ERG in amending calculations. However, this is due to the complex nature of the timings for several of these drug regimens and health state costs. External calculations enabled the model to be more flexible to support analysis requests made at clarification and technical engagement. While the company maintains this approach as appropriate, economic models provided at technical engagement will include within the model the calculations for these costs; however, these calculations will not be linked to model inputs.	Not a factual inaccuracy. The ERG would also maintain that it is not best modelling practice to not include all calculations from input parameters in the model.
Table 1.2, page 16 The ERG report states that: *"Nivolumab has been shown to be dominated by pembrolizumab.*	This section should be updated to reflect the uncertainty in the analysis.	As discussed below, the discount upon pembrolizumab should be removed for the results informing this claim. In addition, it should be noted that the incremental QALYs for this analysis are very small and are more indicative of equivalence	Not a factual inaccuracy – the ERG reported the results of the company base case and then implemented the correction to remove the discount.

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
		than a difference in clinical benefit. As such, this uncertainty should be reflected in the reporting.	
Section 4.2.6, Page 136 The ERG report states with respect to parametric survival models: "*On the other hand, it seems that no matter which one is chosen, PEMBRO-CHEMO would be more effective and probably less costly and thus NIVO-CHEMO would still be dominated*"	This text should be removed	This conclusion is drawn in the context of the publicly available results, which the ERG notes were produced with the inclusion of a 50% discount on the price of pembrolizumab. Removing this discount gives results in the south-west quadrant of the CE plane, with the ICER for NIVO-CHEMO cost-effective.	Amended.
ERG modified cost-effectiveness model (NIVO-CHEMO vs CHEMO scenarios) Also: **Section 1.6, Table 1.14, page 23** Also: **Section 6.2, Table 6.1, page 174** The RDIs calculated for ERG case 4 and ERG base case are incorrect – the weighting has been performed by the column immediately to the left in the calculation sheet, rather than with	The corrected values for NIVO-CHEMO are: • Nivolumab: 92.48% • Cisplatin: 86.15% • Fluorouracil: 88.91% The corrected values for CHEMO are: • Cisplatin: 91.44% • Fluorouracil: 94.23%	Correction of ERG RDI values.	Corrected.

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
reference to the “central” RDI values (i.e. column “A”)			

Issue 7 Clinical effectiveness

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
Section 3.1.1, Page 34 The ERG report states that: *“…additional searches should be undertaken to ensure that AEs that are long-term, rare or unanticipated are not missed.”*	This sentence should be amended	The AEs for nivolumab are well-characterised in line with other indications for nivolumab, which has been extensively studied.[8-11] Therefore, long-term, rare or unanticipated AEs would be characterised by the EMA.	Not a factual inaccuracy.
Section 3.3, page 91, The ERG report states that: *“TC/TPS, which was used in KEYNOTE-590,”*	This should be removed to say: “TC/TPS is obtained by dividing…”	This is incorrect as KEYNOTE-590 used PD-L1 CPS to define their patient population and not TC/TPS.	Amended.

Issue 8 Other typographical errors and inaccuracies

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
Section 2.3, Page 31 The ERG states: “The company did reiterate that chemotherapy is still system organ class (SOC) regardless of PD-L1 status”	This sentence should be updated to: “The company did reiterate that chemotherapy is still standard of care (SOC) regardless of PD-L1 status”	BMS do not believe that this statement was made during the submission or clarification questions. BMS believe that this is a typographic error and SOC should be defined as standard of care.	Amended.

<table>
<tr><th>Description of problem</th><th>Description of proposed amendment</th><th>Justification for amendment</th><th>ERG comment</th></tr>
<tr><td>Section 3.1.2 Page 36
The ERG states:
“However, the company specifies a population subgroup in the DP (OSCC with PD-L1 with CPS≥10).”</td><td>This sentence should be updated to:
“However, the company specifies a population subgroup in the DP (OSCC with PD-L1 with TC≥1%).”</td><td>BMS do not believe that this statement was made during the submission or clarification questions. BMS believe that this is a typographic error.</td><td>Amended.</td></tr>
<tr><td>Section 3.3 Page 91
The ERG states:
“TC/TPS, which was used in KEYNOTE-590, is obtained by …”</td><td>This sentence should be updated to:
“TC/TPS, which was used in CheckMate 648, is obtained by …”</td><td>BMS assumes that this is a typographic error as the CPS method was used in KEYNOTE-590 opposed to the TC method in CheckMate 648.</td><td>Amended.</td></tr>
<tr><td>Table 3.26, Page 95
The ERG state within the eligibility criteria for CheckMate 648:<table><tr><td>Prior Adjuvant Therapy</td><td>Chemotherapy completed ≥6 months prior to enrolment</td></tr></table></td><td>This should be amended for clarity:<table><tr><td>Prior Adjuvant Therapy</td><td>Eligible if completed ≥6 months prior to enrolment</td></tr></table></td><td>The current wording is misleading, implying that patients were eligible only if prior adjuvant therapy had been received. However, [redacted] of enrolled patients had received no prior systemic therapy.
Patients who had received prior adjuvant or neoadjuvant treatment were eligible, but a recurrence-free period of 24 weeks was required after completion of therapy.</td><td>Amended.</td></tr>
<tr><td>Section 3.2, Table 3.26 in the ERG report gives the median duration of treatment in KEYNOTE-590:</td><td>The footnote to Table 3.26 should be updated to indicate that “*” indicates a mean rather than median</td><td>Typographical error in copying from NMA report: The value quoted from Table 4 of the NMA</td><td>Amended.</td></tr>
</table>

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
"PEMBRO+CHEMO: 7.7 months (SD: 6.8)* *CHEMO: 5.8 months* (SD: 4.8)"*		report is indicated as a mean duration of treatment. Impact: No impact.	
Section 3.4.1, page 97 The ERG report states that: *"Despite preferences, HRs (at 3, 6, 9, 12, 24 and 36 months) for all parametric models were shown in Table 6 for PFS and Table 11 for OS as well as plots of HR versus time up to 30 years in Appendix L"*	The text should be amended such that the word "years" reads "months"	Typographical error in copying from Appendix L of submission documents.	Amended.
Section 6.2 Table 6.2: Table subsection 6. Has heading "Remove additional all-cause morality"	"morality" should be corrected to "mortality"	Typographical error	Amended.
Section 7, Page 178-179 The ERG states: "However, in comparison to pembrolizumab with chemotherapy for the PD-L1 TP ≥1% and CPS TP ≥10% subgroup, this criterion would almost certainly not be met given that nivolumab with chemotherapy in the PD-L1 TP ≥1% population had a median OS of [redacted] versus 12.6 for pembrolizumab with chemotherapy in the PD-L1 CPS ≥10% population,	The statement should be amended to reflect the OSCC CPS≥10 population. "However, in comparison to pembrolizumab with chemotherapy for the PD-L1 TP ≥1% and CPS TP ≥10% subgroup, this criterion would almost certainly not be met given that nivolumab with chemotherapy in the PD-L1 TP ≥1% population had a median OS of [redacted] versus 13.9 for pembrolizumab with chemotherapy in the PD-L1 CPS ≥10% population, which translates into HRs greater than 1 in the ITC	KEYNOTE-590 reported a median OS of 13.9 months for OSCC patients with CPS≥10 receiving pembrolizumab; median OS of 12.6 months was reported for OSCC patients regardless of PD-L1 status. The median OS for OSCC patients with CPS≥10 is likely to be the correct comparison in this situation.	Corrected – the ERG note that the figure for median OS for pembrolizumab with chemotherapy is actually 13.96, as reported in the original Appendix L. Also, the figure for nivolumab with chemotherapy has been corrected to [redacted] for the PD-L1 CPS ≥10% subgroup, as reported in the updated Appendix L. Finally, the whole of Table

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
which translates into HRs greater than 1 in the ITC and a loss in LYs gained in the economic model"	and a loss in LYs gained in the economic model"		3.23 has been changed to report the results for the PD-L1 CPS ≥10% population.
Section 4.2.3, Table 4.5, Page 123 Table 26 of CL response	*Table 24 of CL response*	Typographical error, should read: *"Table 24 of CL response"*	The ERG believes that the table footnote is correct (as per page 84 and Table 26 of the clarification letter response). Therefore, not amended.
Section 4.2.6 Page 136, Table 4.11 Based on Table 20, clarification response.	*Table 18, clarification response*	Typographical error, should read: *"Table 18, clarification response"*	The ERG believes that the table footnote is correct (as per page 61 and Table 20 of the clarification letter response). Therefore, not amended.
Section 4.2.6 Page 136, Table 4.10 Based on Table 18, clarification response.	*Table 16, clarification response*	Typographical error, should read: *"Table 16, clarification response"*	The ERG believes that the table footnote is correct (as per page 60 and Table 18 of the clarification letter response). Therefore, not amended.
Section 4.2.9, Page142, Table 4.15 Adapted from Tables 33, 37 and 38 of clarification letter response[12]	*Adapted from Tables 35, 36 and 37 of clarification letter response*	Typographical error, should read: *"Adapted from Tables 35, 36 and 37 of clarification letter response"*	The ERG believes that the table footnote is correct (as per pages 121,126 and 126-7 and Tables 33, 37 and 38 respectively of the clarification letter

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
			response). Therefore, not amended.
Section 4.2.9, Page 145, Table 4.17 Based on Table 24 of CL response	*Based on Table 22 of the CL response*	Typographical error, should read: *"Based on Table 22 of the CL response"*	The ERG believes that the table footnote is correct (as per page 80 and Table 24 of the clarification letter response). Therefore, not amended.
Section 4.2.9, Page 145, Table 4.17 CHEMO = chemotherapy; CL = clarificatin letter; NIVO = nivolumab; PEMBRO = pembrolizumab	*CL= clarification letter*	Typographical error, should read: *"CL= clarification letter"*	Amended.
Section 5.1, Page 159, Table 5.2 Based on Table 27 of CL response	*Based on Table 25 of the CL response*	Typographical error, should read: *"Based on Table 25 of the CL response"*	The ERG believes that the table footnote is correct (as per page 85 and Table 27 of the clarification letter response). Therefore, not amended.
Section 5.1, Page 159, Table 5.3 Based on Table 27 of CL response	*Based on Table 26 of the CL response*	Typographical error, should read: *"Based on Table 26 of the CL response"*	The ERG agrees that the table footnote is incorrect but does not agree that this should be Table 26 of the clarification letter response. Now amended to read 'Table 28' as per page 85 of the clarification letter response.

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
Section 5.1, page 161, Table 5.5 Based on Table 29 in CL response	*Based on Table 27 of the CL response*	Typographical error, should read: *"Based on Table 27 of the CL response"*	The ERG believes that the table footnote is correct (as per page 86 and Table 29 of the clarification letter response). Therefore, not amended.
Section 5.3.1, page 166, Table 5.8 Based on Table 50 of the CL response	*Based on Table 48 of the CL response*	Typographical error, should read: *"Based on Table 48 of the CL response"*	The ERG believes that the table footnote is correct (as per page 154 and Table 50 of the clarification letter response). Therefore, not amended.

Confidential marking

Location of incorrect marking	Description of incorrect marking	Amended marking	ERG comment
Section 2, Table 2.1, Page 25 **Section 2.1, Page 29** **Section 4.2.3, Page 122.**	The indication is no longer commercial in confidence as the licence has now been granted and published.	Marking should be removed	Amended.
Section 3.2.5, Page 52,	Academic in confidence marking missing from text	This should read: *"median (range) in months* [redacted] ([redacted] to [redacted]) *versus months* [redacted] ([redacted] to [redacted])*respectively"*	Amended.

Location of incorrect marking	Description of incorrect marking	Amended marking	ERG comment
Section 3.2.5, Page 53,	Academic in confidence marking missing from text	This should read: *"were around [redacted]% and [redacted]% respectively"*	Amended.
Section 3.2.7, Table 3.11, Page 60	Commercial in confidence marking missing from text	The table footnote should read: *"Based on Table 12 of the CS that cited CheckMate 648 [redacted] Summary data as the primary source."*	Amended.
Section 3.2.8, Page 75	Mixed academic and commercial in confidence marking when it should all be commercial in confidence	The text should read: *"The additional file, provided by the company, containing the results for the [redacted] data cut-off,"*	Amended.
Section 3.2.9.1, Page 75	Commercial in confidence marking missing from text	The text should read: *"A summary of the overall AEs is presented in Table 3.12 for the [redacted] DBL"*	Amended.
Section 3.2.9, Page 76	Mixed academic and commercial in confidence marking when it should all be commercial in confidence	The text should read: *"The company did not provide separate evidence for patients expressing PD-L1 ≥1% (TC) for the [redacted]"*	Amended.
Section 3.2.9, Table 3.12, Page 76	Mixed academic and commercial in confidence marking when it should all be commercial in confidence	The footnote should read: *"Based on Table 18 of Document B of the CS, [redacted]"*	Amended.

Location of incorrect marking	Description of incorrect marking	Amended marking	ERG comment
Section 3.2.9, Table 3.12, Page 77-78	Table is incorrectly marked commercial in confidence, instead of academic in confidence	All values in the table should be marked in yellow, opposed to blue. Only [redacted] data cut-off in the footnote should be marked in blue.	The ERG believes that the company intended to mention Table 3.13 here (as opposed to Table 3.12). The requested amendments have been made to Table 3.13.
Section 3.2.9, Tables 3.14, 3.15, 3.16 and 3.17, Page 79-80	Commercial in confidence marking missing from text	The footnote in each table should read: *"[redacted] data cut-off"*	Amended.
Section 3.2.9, Page 80	Commercial in confidence marking missing from text as information was taken from the CSR which is confidential.	Text should read: "… corresponding to the supplementary tables provided by the company with an [redacted] 'data stamp'	Amended.
Section 3.4.3.2, Table 3.30, Page 106	Academic in confidence marking missing in table	The values in the table should be marked in yellow which was missed in Document B of CS.	Amended.
Section 3.6, Page 114	Academic in confidence marking missing from text	Text should read: *"More deaths were observed among patients on CHEMO compared with NIVO-CHEMO: [redacted]% versus [redacted]% for all randomised patients; and [redacted]% versus [redacted]% for patients with PD-L1 TC ≥1%."*	Amended.

Location of incorrect marking	Description of incorrect marking	Amended marking	ERG comment
Section 4.2.6, page 125	Academic in confidence marking is not required as this is published in Kato et al.[7]	Text should read: *"…of anti-PD1/PD-L1 therapy: 15% versus 6%."*	Amended.

References

1. National Institute for Health and Care Excellence. Putting NICE guidance into practice. Resource impact template: Pembrolizumab with platinum- and fluoropyrimidine-based chemotherapy for untreated advanced oesophageal and gastro-oesophageal junction cancer (TA737). 2021. Available at: https://www.nice.org.uk/guidance/ta737/resources/resource-impact-template-excel-9263083501 [Accessed 15/06/2022].
2. National Institute for Health and Care Excellence. Pembrolizumab with platinum- and fluoropyrimidine-based chemotherapy for untreated advanced oesophageal and gastro-oesophageal junction cancer [TA737] 2022. Available at: https://www.nice.org.uk/guidance/ta737/history [Accessed 15/06/2022].
3. Davidson M, Chau I, Cunningham D, et al. Impact of tumour histological subtype on chemotherapy outcome in advanced oesophageal cancer. World J Gastrointest Oncol. 2017;9(8):333-40.
4. Bristol-Myers Squibb Pharmaceuticals Ltd. Nivolumab in combination with fluoropyrimidine- and platinum-based combination chemotherapy for the first-line treatment of adult patients with unresectable advanced, recurrent, or metastatic oesophageal squamous cell carcinoma with tumour cell PD-L1 expression ≥1% [ID2712]: Submission to National Institute of Health and Care Excellence. Single technology appraisal (STA): Appendix E - Systematic literature review of clinical evidence evaluating first-line therapies for patients with advanced or metastatic esophageal cancer. Bristol-Myers Squibb Pharmaceuticals Ltd., 2022.
5. Shen L, Lu ZH, Wang JY, et al. LBA52 Sintilimab plus chemotherapy versus chemotherapy as first-line therapy in patients with advanced or metastatic esophageal squamous cell cancer: first results of the phase III ORIENT-15 study. Annals of Oncology. 2021;32:S1330.
6. Luo H, Lu J, Bai Y, et al. Effect of camrelizumab vs placebo added to chemotherapy on survival and progression-free survival in patients with advanced or metastatic esophageal squamous cell carcinoma: the ESCORT-1st randomized clinical trial. Jama. 2021;326(10):916-25.
7. Kato K, Cho BC, Takahashi M, et al. Nivolumab versus chemotherapy in patients with advanced oesophageal squamous cell carcinoma refractory or intolerant to previous chemotherapy (ATTRACTION-3): a multicentre, randomised, open-label, phase 3 trial. Lancet Oncol. 2019;20(11):1506-17.
8. National Institute for Health and Care Excellence. Nivolumab for previously treated unresectable advanced or recurrent oesophageal cancer [TA707]. 2021. Available at: https://www.nice.org.uk/guidance/ta707 [Accessed 25/06/2021].
9. National Institute for Health and Care Excellence. Nivolumab in combination with ipilimumab for treating advanced melanoma [TA400]. 2016. Available at: https://www.nice.org.uk/guidance/ta400 [Accessed 25/06/2021].
10. National Institute for Health and Care Excellence. Nivolumab for previously treated advanced renal cell carcinoma [TA417]. 2016. Available at: [Accessed 25/06/2021].
11. National Institute for Health and Care Excellence. Nivolumab for treating advanced (unresectable or metastatic) melanoma [TA384]. 2016. Available at: [Accessed 25/06/2021].
12. National Institute for Health and Care Excellence. Nivolumab in combination with fluoropyrimidine- and platinum-based combination chemotherapy for the first-line treatment of adult patients with unresectable advanced, recurrent, or metastatic oesophageal squamous cell carcinoma with tumour cell PD-L1 expression ≥1% [ID2712]: Response to request for clarification from the ERG. London: NICE, 2022.

NICE National Institute for Health and Care Excellence

Technical engagement response form

Nivolumab platinum-based chemotherapy for untreated advanced oesophageal cancer with tumour cell PD-L1 expression ≥1% [ID2712]

As a stakeholder you have been invited to comment on the external assessment group (EAG) report for this appraisal.

Your comments and feedback on the key issues below are really valued. The EAG report and stakeholders' responses are used by the appraisal committee to help it make decisions at the appraisal committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

Information on completing this form

We are asking for your views on key issues in the EAG report that are likely to be discussed by the committee. The key issues in the EAG report reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the EAG report.

You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise.

If you would like to comment on issues in the EAG report that have not been identified as key issues, you can do so in the 'Additional issues' section.

If you are the company involved in this appraisal, please complete the 'Summary of changes to the company's cost-effectiveness estimates(s)' section if your response includes changes to your cost-effectiveness evidence.

Please do not embed documents (such as PDFs or tables) because this may lead to the information being mislaid or make the response unreadable. Please type information directly into the form.

Technical engagement response form

Do not include medical information about yourself or another person that could identify you or the other person.

We are committed to meeting the requirements of copyright legislation. If you want to include journal articles in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.

Please underline all confidential information, and separately highlight information that is submitted under [redacted], all information submitted under [redacted], and all information submitted under [redacted] in pink. If confidential information is submitted, please also send a second version of your comments with that information replaced with the following text: 'academic/commercial in confidence information removed'. See the Guide to the processes of technology appraisal (sections 3.1.23 to 3.1.29) for more information.

Deadline for comments by **5pm** on **28 July 2022**. Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

Technical engagement response form

Nivolumab platinum-based chemotherapy for untreated advanced oesophageal cancer with tumour cell PD-L1 expression ≥1% [ID2712]

NICE National Institute for Health and Care Excellence

About you

Table 1. About you

Your name	Reya Sharma
Organisation name: stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder, please leave blank)	Bristol-Myers Squibb Ltd
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	N/A

PD-L1 testing

As highlighted during factual accuracy check of the EAG report, PD-L1 testing can be broadly divided into two methods:

- Tumour cell (TC/TPS) method
    - TC scores are obtained by dividing the number of PD-L1 stained tumour cells by the total number of viable tumour cells and multiplying by 100.
    - TC scores are reported as a percentage.
- Combined positive score (CPS) method
    - CPS evaluates the number of PD-L1-stained cells (tumour cells, lymphocytes, macrophages) relative to all viable tumour cells.
    - CPS scores are obtained by dividing the number of PD-L1 stained cells (tumour cells and immune cells) and multiplying by 100.
    - CPS scores are reported as a number.

While KEYNOTE-590 reported only CPS-based subgroups, CheckMate 648 reported primarily TC-based subgroups.

Ahead of technical engagement, BMS contacted expert pathologists to determine likely clinical practice around TC and CPS testing. The pathologists highlighted that these are essentially both independent tests for assessing PD-L1 expression. While there is overlap in these tests, this overlap is not complete. Additionally, there should be no perceived linear relationship between expression levels: patients with PD-L1 TC ≥ 1% should not be seen as lower PD-L1 expression than PD-L1 CPS ≥ 10.

As an example, a theoretical patient may have:

- PD-L1 tumour cell expression but no PD-L1 expression on the immune cells. This example would mean that PD-L1 scoring could go either way depending on the amount of tumour cell expression. Therefore, this patient could be eligible for either nivolumab or pembrolizumab.

- No PD-L1 tumour cell staining (TC < 1%) but have PD-L1 expression on immune cells (CPS ≥ 10) Therefore, this patient would be eligible for treatment with pembrolizumab, but not with nivolumab.

As the nivolumab licence specifies usage in patients with PD-L1 TC ≥ 1%, the PD-L1 TC test is a mandatory requirement if nivolumab were to be prescribed, while CPS score would be irrelevant and would not be calculated as it does not inform whether nivolumab should be administered.

Pathologists noted that clinicians should typically request the treatment of choice and this would then guide the relevant test, so that testing method would not be a barrier to use of nivolumab in clinical practice. Further, pathologists would prefer TC testing as this is less time consuming to undertake. Pathologists would not want to do both tests, as this will require staining and counting of two sets of slides, taking up time and resource in a busy department.

Updated patient access scheme

Ahead of addressing the key issues presented in the Technical Engagement response, Bristol-Myers Squibb wish to highlight an update to the patient access scheme (PAS). For clarity, all analyses presented in this response apply this updated PAS. The impact of this update is described briefly below and in appendices.

The agreed PAS for nivolumab has been updated from [redacted]% to [redacted]% impacting on vial costs as follows:

Nivolumab costs without PAS[1]

- £2,633.00 per 240 mg (24 mL) vial;
- £1,097.00 per 100 mg (10 mL) vial;
- £439.00 per 40 mg (4 mL) vial.

Nivolumab costs with PAS

- [redacted] per 240 mg (24 mL) vial;
- [redacted] per 100 mg (10 mL) vial;
- [redacted] per 40 mg (4 mL) vial.

This updated PAS has been applied within this response. For reference, the impact of the updated PAS on the company preferred base case pre-technical engagement is presented in Table 2.

Table 2. Impact of updated PAS on cost-effectiveness analysis

			Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
NIVO-CHEMO versus CHEMO	Company base case post CQs	Original PAS	[redacted]	[redacted]	£33,357
		Updated PAS	[redacted]	[redacted]	£31,826
	ERG base case	Original PAS	[redacted]	[redacted]	£49,017
		Updated PAS	[redacted]	[redacted]	£46,599
NIVO-CHEMO versus PEMBRO-CHEMO	Company base case post CQs*	Original PAS	[redacted]	[redacted]	£307,447 (CE region)
		Updated PAS	[redacted]	[redacted]	£320,254(CE region)
	ERG base case†	Original PAS	[redacted]	[redacted]	£104,998 (CE region)
		Updated PAS	[redacted]	[redacted]	£109,039 (CE region)

CE: cost-effectiveness; CQs: clarification questions; ERG: evidence review group; ICER: incremental cost-effectiveness ratio; PAS: patient access scheme; QALY: quality-adjusted life year.
** Discount upon Pembrolizumab removed – see Key Issue 1*
† Corrected application of hazard ratio – see Appendix A, Correction to ERG model

Modification of subsequent therapy outcomes

Additionally, BMS would like to highlight a modification of subsequent therapy outcomes. Based on ERG's request to adjust the outcomes or models thereof to improve prediction of outcomes in the cohort receiving subsequent therapy, analyses were conducted with results presented in Table 3. The Rank-Preserving Structural Failure Time Model was used as suggested in Ouwens et al.[2] as well as a model-based evaluation. For this scenario, the economic model was modified such that patients on CHEMO would discontinue first-line therapy at a rate determined by a model of time to subsequent therapy or death, informed by data from CheckMate 648.

Table 3. Impact of modification of subsequent therapy outcomes

			Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
NIVO-CHEMO versus CHEMO	Updated Company base case	Updated base case	[redacted]	[redacted]	£27,106
		RPSFTM-scenario	[redacted]	[redacted]	£29,253
		Model-based ATTRACTION-3 scenario	[redacted]	[redacted]	£39,868

Key issues for engagement

All: Please use the table below to respond to the key issues raised in the EAG report.

Table 4 Key issues

Key issue	Does this response contain new evidence, data or analyses?	Response
Key issue 1: Uncertainty as to the appropriate comparators dependent on PD-L1 status **Report section** Executive summary: Table 1.7 Main report: Section 3.4, Section 4.2.6	Yes – new cost-effectiveness scenario analyses are presented describing the comparison of NIVO-CHEMO with PEMBRO-CHEMO	**Comparison of NIVO-CHEMO versus PEMBRO-CHEMO** As shown in Table 25 in Appendix A, a scenario exploring equal efficacy of NIVO-CHEMO and PEMBRO-CHEMO was undertaken within the updated ERG model showing an ICER in the south west quadrant of the cost-effectiveness (CE) plane. The results are presented for the updated company base case, including all updates that were incorporated based on the key issues outlined in this response form, using the survival models as outlined in the company submission (semi-parametric approach: generalised gamma and lognormal for NIVO-CHEMO progression free survival [PFS] and overall survival [OS], respectively and Weibull and lognormal for PEMBRO-CHEMO PFS and OS, respectively). Additionally, a scenario with alternative survival models was conducted with a fully parametric approach and log-logistic extrapolation for NIVO-CHEMO OS, PEMBRO-CHEMO PFS, and PEMBRO-CHEMO OS with the result that NIVO-CHEMO dominates PEMBRO-CHEMO
	Yes – subgroup analysis for PD-L1 ≥ 1% TC and <10 CPS population	**Comparison of NIVO-CHEMO versus CHEMO for PD-L1 TC ≥ 1% and CPS < 10**

Key issue	Does this response contain new evidence, data or analyses?	Response
		The ERG has requested an analysis comparing NIVO-CHEMO versus CHEMO in CheckMate 648 patients with PD-L1 ≥ 1% TC and < 10 CPS population. Clinically, even though the two types of testing methods are related, the subsequent populations derived from those tests are not transposable. Ideally, the oncologist would outline the preferred treatment pathway and the pathologist would action the appropriate test that is validated for that treatment. Pathologists contacted by BMS would not want to do both tests, as this will require staining and counting of two sets of slides, taking up time and resource in a busy department. Further, as stated during the response to clarification questions, the sample size calculation of CheckMate 648 is based on PD-L1 TC ≥ 1% and not CPS, that the study is not powered for a subgroup analysis of PD-L1 TC ≥ 1% and CPS < 10%. However, in order to address uncertainty, this subgroup analysis is provided in the Issue 1 – Comparison of NIVO-CHEMO versus CHEMO in PD-L1 ≥ 1% TC and <10 CPS population. This was used to inform two scenarios: one reflecting the impact on the company base case analysis and one reflecting the impact on the ERG base case analysis. OS, PFS and time on treatment (ToT) were updated to reflect the subgroup of interest. Additionally, the nivolumab patient access was updated to reflect that described in the updated patient access scheme section. All other parameters remained as per the respective base case analyses. As shown in Table 10 within the appendix, economic modelling of survival data for this subgroup marginally decreased QALY accrual in the NIVO-CHEMO and CHEMO arms. The resulting ICERs were higher than both base case analyses but remained below a willingness-to-pay threshold of £50,000/QALY.
	No	**Relevance of chemotherapy as a comparator**

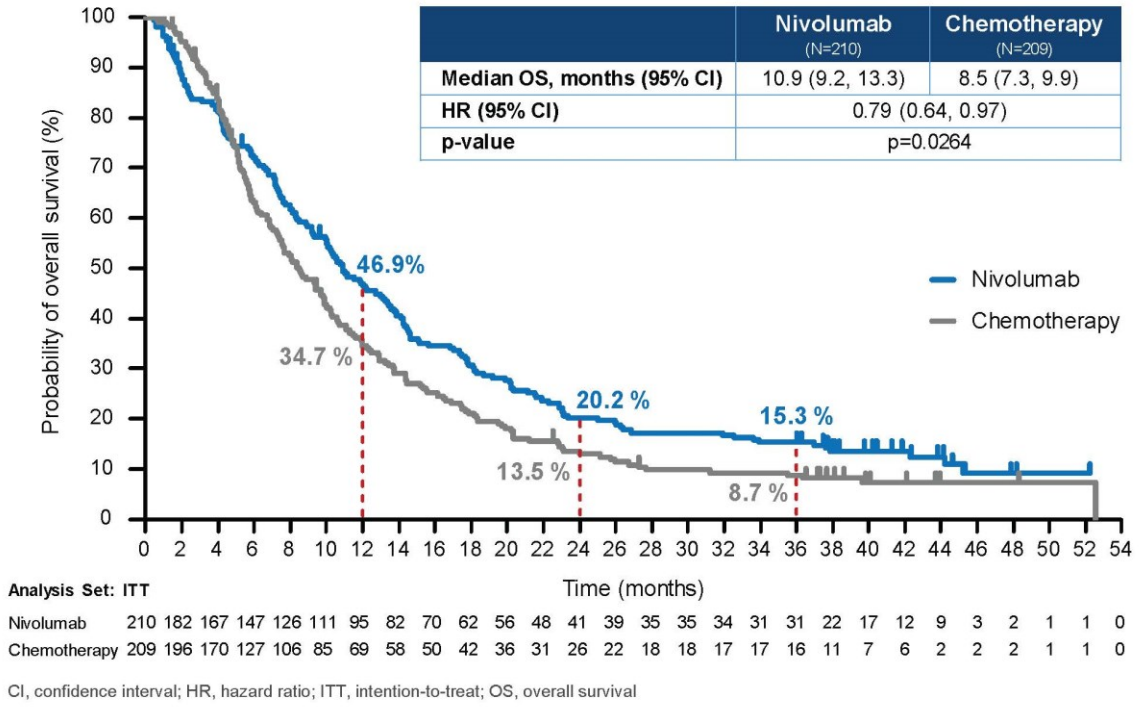
Key issue	Does this response contain new evidence, data or analyses?	Response
		In Appendix A, BMS has presented an analysis comparing NIVO-CHEMO versus CHEMO for patients with PD-L1 TC ≥ 1% and CPS < 10. Although these outcomes remain cost-effective, BMS wants to highlight that this should not be considered the base case population for the comparison with chemotherapy. Despite NICE advice, it is likely that many patients with PD-L1 CPS ≥10 continue to initiate chemotherapy, in line with NICE budgetary assumptions.[3] Per the resource impact template provided to support TA737,[3] only 51% of eligible patients will have switched to receiving pembrolizumab-based regimens by year 5 (2026). Further, it is likely that this market share will be gathered over time, so not all 51% patients will be receiving pembrolizumab in 2022. Hence, at the time of NICE decision making, chemotherapy will remain the relevant treatment choice for the population of interest.
Key issue 2: There is limited evidence to support the comparability of the PD-L1 ≥10% CPS populations in the two trials used in the ITC analysis. **Report section** Executive summary: Table 1.8 Main report: Section 4.2.6.2	No	The ITC report has been updated to include a comparison of the PD-L1 CPS ≥ 10 population from CheckMate 648 (provided as Appendix B to this document). As can be seen, baseline characteristics from this CheckMate 648 subgroup are broadly comparable to the CheckMate 648 overall population, and hence does not affect the conclusions or introduce uncertainty in the indirect comparison.

Key issue	Does this response contain new evidence, data or analyses?	Response
Key issue 3: It is unclear which ITC method, constant HR or time varying HRs formed the base case for the analysis. **Report section** Executive summary: Table 1.9 Main report: Section 4.2.8	No	The updated EAG report notes that the company have provided clarification that the time varying method was used. Additionally, it should be clarified that "statistical significance" in the ITC report refers to the credibility of confidence intervals. No specific statistical tests were run.
Key issue 4: There is uncertainty as to the nature and effectiveness of subsequent therapy. **Report section** Executive summary: Table 1.10 Main report: Section 4.2.9.3	No	The ERG recommends use of methods described in Technical Support Document (TSD) 16 by Ouwens et al. to adjust outcomes for subsequent anti-PD-1/L1 treatment in the NIVO-CHEMO arm.[2,4] It should be noted that TSD 16 states *"…this TSD focuses upon treatment switching from the control group onto the experimental treatment – we do not consider in detail situations in which experimental group patients switch onto the control treatment, or where patients randomized to either group receive other post-study treatments*." In the current case, patients from an intervention arm are initiating subsequent therapy that was not part of the study (nivolumab monotherapy). Thus, the methods described in TSD 16 do not apply to the current context. However, if we considered the method as a means for adjusting survival in the NIVO-CHEMO arm, TSD 16 describes three methods that are used to adjust for crossover: rank preserving structural failure time (RPSFT), inverse probability of censoring weights (IPCW), and two-stage adjustment. Although these types of analyses are challenging an attempt was made but the results need to be interpreted with caution due to the low patient numbers.

Key issue	Does this response contain new evidence, data or analyses?	Response
		First, the outcomes to reflect modelled subsequent therapies were adjusted. In a scenario the economic model was extended with outcomes from ATTRACTION-3 to adjust OS outcomes for CHEMO patients assumed to receive anti-PD-1 (see: Adjustment of outcomes to reflect modelled subsequent therapies). The outcomes upon NIVO monotherapy in ATTRACTION-3 were deemed to be sufficiently similar to those of patients in CheckMate 648 to enable use of outcomes models derived from this trial to inform expected post second line survival of NIVO-receiving patients in the modelled population (Table 11). An adjustment of OS by RPSFTM is presented (see: RPSFTM adjustment), showing agreement with this model scenario. A model-based evaluation of outcomes was planned in light of the similarity in outcomes for patients receiving subsequent anti-PD-1 therapy in CheckMate 648 and the NIVO arm of ATTRACTION-3. For this scenario, the economic model was modified such that patients on CHEMO would discontinue first-line therapy at a rate determined by a model of time to subsequent therapy or death, informed by data from CheckMate 648. They would then be modelled as at risk to a constant hazard of death informed by the mean survival time from the nivolumab arm of the economic model informing TA707,[5] such that patients in this state would have the same mean survival time as in that economic model. The cost-effectiveness results for the RPSFTM and model-based ATTRACTION-3 scenarios are presented in Table 12. In both cases the ICER remains cost-effective.
	No	As part of the response to this key issue, the modelling approach for the cost of second-line therapy was re-evaluated. The new approach, as described in Appendix A, is considered to be more reflective of clinical practice and reduces the level of uncertainty regarding the costing approach for subsequent therapy**Error! Reference source not found.**. This resulted in a better

Key issue	Does this response contain new evidence, data or analyses?	Response
		representation of the mean time on treatment for patients receiving second-line anti-PD-1 therapy, decreasing the incremental costs between NIVO-CHEMO and CHEMO.
Key issue 5: There is uncertainty as to long term OS and the treatment effect of nivolumab + chemotherapy versus chemotherapy. **Report section** Executive summary: Table 1.11 Main report: Section 4.2.9.1, Section 4.2.9.4, Section 5.3	No	**Plausibility of long-term extrapolations in CHEMO arm** As noted in the survival analysis appendix provided with the company submission (Appendix N), improving the fit to the observed data reduces the hazard in the long-term extrapolation for the CHEMO arm and as a result predicts implausible clinical outcomes. The ERG has commented that little justification is provided for the implausibility, with the most convincing explanation appearing to be the effect of subsequent systemic therapy, especially anti-PD-L1. However, this does not take into account published evidence for PD-L1 inhibitors. Nivolumab was the most commonly used PD-L1 inhibitor during CheckMate 648 and can be considered the most likely to be used in second-line treatment in UK clinical practice, based on TA707.[5] Evidence for nivolumab as a second-line treatment in oesophageal cancer was primarily derived from ATTRACTION-3. During this trial, the median survival for nivolumab-treated patients was 10.9 months (Figure 1). At 12 months, only 46.9% of patients in the nivolumab arm remained alive at 12 months.[6] At three years, 15.3% of patients in the nivolumab arm remain alive. During CheckMate 648, the median duration of CHEMO treatment was [redacted] months in the ITT population, with all patients discontinued by [redacted] months. Based on the ATTRACTION-3 evidence, it can be considered plausible that at least half of these patients would die within one year of initiating second-line nivolumab treatment, with approximately one-fifth alive at two years. At the most optimistic assumption, the ATTRACTION-3 nivolumab arm could be considered the highest possible upper bound for extrapolation of the CheckMate 648 CHEMO arm. This would be considered a highly optimistic assumption, as not all patients discontinuing CHEMO in CheckMate

Key issue	Does this response contain new evidence, data or analyses?	Response
		648 received PD-L1 inhibitors and not all patients in clinical practice would receive second-line nivolumab therapy. As a result of this optimistic assumption, the following OS and PFS extrapolations for CHEMO can be considered clinically implausible without further investigation: • Extrapolations predicting a proportion of the population with no remaining hazard of progression or disease-related death. • Extrapolations predicting ≥ 10% of patients alive at four years. Although this does not impact the ERG's curve selection, it is important to note this point, as this excludes several extrapolations that fit well to the later part of the observed data. As an example, the CHEMO semi-parametric log-logistic and Gompertz fits would not be considered clinically plausible on this basis. Aligned with this assumption, spline models were not considered as a solution to this issue, as the extrapolations fitting well to the tail of the observed data provided implausible extrapolations.

Key issue	Does this response contain new evidence, data or analyses?	Response
		 **Figure 1. ATTRACTION-3 overall survival[6]**
	No	**Discussion of survival modelling approach** The OS Kaplan-Meier curves demonstrate a clear difference in hazard profile between the treatment arms. During the first 6 months, the NIVO-CHEMO and the CHEMO KM curves are very

Key issue	Does this response contain new evidence, data or analyses?	Response
		similar. After 6 months, patients in the CHEMO arm have a higher hazard than the NIVO-CHEMO arm. As a result, the NIVO-CHEMO (Figure 35) hazard rate peaks in the first 2–3 months, whereas in the CHEMO (Figure 36) arm, the hazard continually increases until reaching a peak at approximately 8–9 months. Parametric models were explored but did not adequately reflect the observed change in hazard. As a result, BMS believes that the company base case analysis provides the best prediction of long-term outcomes for NIVO-CHEMO. However, the ERG base case analysis may provide a more conservative approach.
	No	The ERG has employed treatment waning, which has precedent in technology assessment, but is otherwise unjustified and arbitrary.Firstly, the time point of waning "start" is arbitrary. As demonstrated in the figure below, both the company base case and the ERG base case without artificial waning demonstrate an increasing hazard ratio (decreasing treatment effect). Applying the artificial waning assumption at 2.5 years is within the domain of the observed survival outcomes in CheckMate 648, and this time should be delayed until at least the end of the observed period. In the light of sustained benefit in other indications (described above) this may under-estimate the benefit of nivolumab, and later artificial waning points should be considered. In addition, the time period over which waning occurs is arbitrary, and in the ERG base case demonstrates a substantial deviation from the hazard ratio trajectory prior to that point. Although the company disagree with incorporating treatment waning, scenarios are presented within this response. Based on previous appraisals and the ongoing appraisal in first-line gastro-oesophageal cancer treatment waning should be applied at 5 years the earliest.[7-10] Alternative scenarios were explored within the ERG version of the economic model, reported in Appendix A (Issue 5 – Impact of long-term treatment effect between NIVO-CHEMO and CHEMO).

Key issue	Does this response contain new evidence, data or analyses?	Response
		Cost-effectiveness increases with later implementation and with longer duration of waning. A case where waning is delayed until 4 years (end of trial period) and closes the hazard ratio at 10 years, to approximate the gradient of the hazard ratio at the point of waning start, is considered to be an improvement on the modelling assumptions of the ERG scenario. All scenarios considered were found to be cost-effective. **Figure 2. Alternative waning assumptions applied to the ERG model base case. Hazard ratio of overall survival using economic model output traces**
Key issue 6: There is uncertainty as to how long-term OS for the comparison of nivolumab + chemotherapy versus pembrolizumab + chemotherapy. **Report section** Executive summary: Table 1.7 Main report: Section 3.4, Section 4.2.6	No	PEMBRO-CHEMO curves were generated external to the economic model. The hazard ratios from the ITC were applied to the CHEMO survival curves to generate the PEMBRO-CHEMO survival curves by scaling the hazards of the survival curve in a piecewise constant manner, per the time-varying hazard ratios reported in the NMA report. This calculation has been included within the economic model. In so doing, it was revealed that there were stochastic differences between the estimated hazard ratios in the analysis originally used to perform the adjustment, and those given within the NMA report; for consistency, the values from the final NMA report have been used in the final base case and for probabilistic sensitivity analysis. The impact of this is described in Appendix A.
Key issue 7: There is uncertainty as to how all-cause mortality should be incorporated in the model.	No	The long-term extrapolation of the survival models derived from the CheckMate 648 trial are constrained only by the behaviour of the statistical distributions which are assumed. By extrapolating these distributions, we not only assume that the distribution chosen correctly describes the distribution of survival times due to the processes observed in the trial period, but also that these hazard-generating processes are well described by these distributions through the

Key issue	Does this response contain new evidence, data or analyses?	Response
Report section Executive summary: Table 1.8 Main report: Section 4.2.6.2		unobserved period. In practice, plausible models of mortality hazard must always predict increasing hazards as patients approach old age, in spite of clinical trial data showing a medium-term real reduction in marginal mortality hazard over the population. To enforce a plausible minimum rate of mortality, hazards due to a life table matched to the economic model patient are assumed additional to the hazards predicted by the overall survival model. This decision was made due to the relatively high hazard of mortality during the observed period relative to the matched general population, at a baseline age of only 62.6 years. This matched hazard is low enough that fitting of the extrapolative models in a relative survival framework (i.e. assuming a baseline hazard due to lifetables when fitting the parameters of the survival models) would result in very little difference to the fitted parameters. Double-counting due to application of both the parametric overall survival model and the lifetable survival curve is thus minimal, as demonstrated by consideration of the model input and output traces, plotted in *Figure 3. A plausible separation of hazards between the general population and the modelled population is maintained into very long-term extrapolation by this method. **Figure 3: Hazard of mortality over each weekly model cycle from model input, lifetable, and output traces, company base case**
Key issue 8: There is uncertainty as to whether health state utilities should be treatment dependent or incorporate a terminal care decrement.	Yes	An updated utility analysis was run, but the impact on the finding of the previous analysis was minimal.

Key issue	Does this response contain new evidence, data or analyses?	Response
Report section Executive summary: Table 1.9 Main report: Section 4.2.8		
Key issue 9: There is uncertainty as to the appropriate method and value of any adjustment to cost due to delayed or missed doses. **Report section** Executive summary: Table 1.10 Main report: Section 4.2.9.3	No	In the submission documents, the company presented an analysis based upon the proportion of doses missed or delayed during the trial in order to inform the costs of medication incurred during modelled time on first-line treatment. For each component of the treatment regimen, the total number of doses delayed divided by the total number of doses delivered modified the mean cost of treatment acquisition. By contrast, the ERG presented an analysis using summaries of patient-level relative dose intensity (RDI) from the CheckMate 648 trial. As an estimate of the mean cost of treatment per unit time on treatment, a mean derived by this method is compromised twofold: 1. The mean-of-mean approach gives equal weight to each patient, rather than each unit of patient exposure time. If RDI is associated with time on treatment (e.g. if early discontinuers would be more likely to have low RDI) then this biases the mean versus a grand mean approach. 2. The ERG used grouped data and assumed the midpoint of each RDI group to apply to all individuals in that group. The mean RDI of each group was not known, and so this estimate may be biased. The company used patient-level data from CheckMate 648 to resolve the second compromise and applied the results to the ERG model. As these results are an expression of the intent of the ERG analysis with access to higher-quality data, these results supersede the ERG's base case. In addition, an analysis was undertaken whereby the mean after weighting by patient-level time on treatment (as used to derive the economic model input curve) was calculated. This weighted

Key issue	Does this response contain new evidence, data or analyses?	Response
		mean resolves the first compromise of the ERG method and is incorporated into the new company base case.
Key issue 10: Calculations were missing from the model, which reduces transparency and makes updating difficult. **Report section** Executive summary: Table 1.11 Main report: Section 4.2.9.1, Section 4.2.9.4, Section 5.3	No	Cost calculations were provided to the ERG, documented in the technical report and Document B. It is acknowledged that these calculations were not integrated into the modelling, which did not support the ERG in amending calculations. However, this is due to the complex nature of the timings for several of these drug regimens and health state costs. External calculations enabled the model to be more flexible to support analysis requests made at clarification and technical engagement. While the company maintains this approach as appropriate, economic models provided at technical engagement will include within the model the calculations for these costs; however, these calculations will not be linked to model inputs.
Key issue 11: Health state costs were estimated from an out-of-date source. **Report section** Executive summary: Table 1.12 Main report: Section 4.2.9	Yes	Updated analyses are presented in Appendix A. The drug acquisition and additional costs were updated and the results are presented in Table 20.

Key issue	Does this response contain new evidence, data or analyses?	Response
Key issue 12: Errors, which underestimated the cost of PEMBRO-CHEMO and prevented the PSA for PEMBRO-CHEMO comparison. **Report section** Executive summary: Table 1.13 Main report: Section 6	Yes	Updated analyses are presented in Appendix A. Table 21 provides the results of updating the pembrolizumab cost upon the company base case after Clarification Questions, i.e. prior to application of the updated nivolumab PAS or any other changes. Doing so moves nivolumab to the cost-effective quadrant, although the incremental benefits between the two treatments are small.

Additional issues

All: Please use the table below to respond to additional issues in the EAG report that have not been identified as key issues. Please do **not** use this table to repeat issues or comments that have been raised at an earlier point in this appraisal (for example, at the clarification stage).

Table 5 Additional issues from the EAG report

Issue from the EAG report	Relevant section(s) and/or page(s)	Does this response contain new evidence, data or analyses?	Response
None			

NICE National Institute for Health and Care Excellence

Summary of changes to the company's cost-effectiveness estimate(s)

Company only: If you have made changes to the base-case cost-effectiveness estimate(s) in response to technical engagement, please complete the table below to summarise these changes. Please also provide sensitivity analyses around the revised base case. If there are sensitivity analyses around the original base case which remain relevant, please re-run these around the revised base case.

Table 6 Changes to the company's cost-effectiveness estimate

Key issue(s) in the EAG report that the change relates to	Company's base case before technical engagement	Change(s) made in response to technical engagement	Impact on the company's base-case incremental cost-effectiveness ratio (ICER)
Company base case at clarification stage	Not applicable	Not applicable	ICER versus CHEMO: £33,357 ICER versus PEMBRO-CHEMO: -£5,594 (not CE)
Key Issue 12 – Pembrolizumab cost update	Assumed discount upon pembrolizumab	PEMBRO-CHEMO costs updated to remove assumed discount upon pembrolizumab	ICER versus PEMBRO-CHEMO: -£307,447 (CE region, SW quadrant)
Updated Patient Access Scheme (NB: included in all subsequent analyses)	Not applicable	Not applicable	ICER versus CHEMO: £31,826 ICER versus PEMBRO-CHEMO: £320,254 (CE region, SW quadrant)

Key issue(s) in the EAG report that the change relates to	Company's base case before technical engagement	Change(s) made in response to technical engagement	Impact on the company's base-case incremental cost-effectiveness ratio (ICER)
Key issue 4 – impact of subsequent therapies	Second-line treatment at list price, discontinued at rate defined by median time on treatment	New PAS applied to nivolumab, costs assuming mean time on treatment from TA 707 applied as lump sum on discontinuation	ICER versus CHEMO: £28,607 ICER versus PEMBRO-CHEMO: £320,428 (CE region, SW quadrant)
Key issue 6 - There is uncertainty as to how long-term OS for the comparison of nivolumab + chemotherapy versus pembrolizumab + chemotherapy	Adjustment of CHEMO survival curves to PEMBRO-CHEMO survival performed offline using early data from ITC	Incorporation of adjustment within economic model, update to hazard ratio table source from NMA report	ICER versus PEMBRO-CHEMO: £1,153,385 (CE region, SW quadrant)
Key issue 10: There is uncertainty as to the appropriate method and value of any adjustment to cost due to delayed or missed doses	Dose modification based upon total number of doses delayed in CheckMate 648	Dose modification based upon time-on-treatment weighted mean relative dose intensity in CheckMate 648	ICER versus CHEMO: £30,183 ICER versus PEMBRO-CHEMO: £298,853 (CE region, SW quadrant)
Key issue 11: Health state Costs were estimated from an out-of-date source.	ICER versus chemotherapy: £33,357 ICER versus pembrolizumab: -£5,594	Costs updated	ICER versus CHEMO: £32,123 ICER versus PEMBRO-CHEMO: £327,407 (CE region, SW quadrant)
Company's base case following technical engagement (or revised base case)	**Incremental QALYs:** **Versus CHEMO:** [redacted] **Versus PEMBRO-CHEMO:**	**Incremental costs:** **Versus CHEMO:** [redacted] **Versus PEMBRO-CHEMO:**	**Please provide company revised base-case ICER** **Versus CHEMO:** **£27,106**

Key issue(s) in the EAG report that the change relates to	Company's base case before technical engagement	Change(s) made in response to technical engagement	Impact on the company's base-case incremental cost-effectiveness ratio (ICER)
	[redacted]	[redacted]	**Versus PEMBRO-CHEMO: £1,107,961 (CE region, SW quadrant)**

Sensitivity analyses around revised base case

PSA and DSA around updated PEMBRO-CHEMO model are presented in Issue 12 – Pembrolizumab cost update and PSA, and a further exploration of the structural uncertainty around the indirect treatment comparison informing this scenario is undertaken in Issue 1 – Comparison of NIVO-CHEMO and PEMBRO-CHEMO. Sensitivity analyses exploring the survival modelling approach were undertaken as per the company submission, with results presented in Table 7. PSA using the updated base case is reported in Appendix A Additional Sensitivity Analysis

Table 7. Scenario analyses - per company submission

Scenarios	Incremental costs (£)	Incremental LYs	Incremental QALYs	ICER (£/QALY)
vs CHEMO – Updated base case analysis	[redacted]	[redacted]	[redacted]	**£27,106**
Scenario: vs CHEMO – parametric survival approach	[redacted]	[redacted]	[redacted]	**£30,618**
Scenario: vs CHEMO – alternative semi-parametric models	[redacted]	[redacted]	[redacted]	**£48,963**
CHEMO: chemotherapy; ICER, incremental cost-effectiveness ratio; LYs: life years; NIVO: nivolumab; QALYs, quality-adjusted life years				

References

1. British National Formulary. Nivolumab solution for infusion. 2021. Available at: https://bnf.nice.org.uk/medicinal-forms/nivolumab.html [Accessed 31 January 2021].
2. Ouwens M, Darilay A, Zhang Y, et al. Assessing the Influence of Subsequent Immunotherapy on Overall Survival in Patients with Unresectable Stage III Non-Small Cell Lung Cancer from the PACIFIC Study. Curr Ther Res Clin Exp. 2021;95:100640.
3. National Institute for Health and Care Excellence. Putting NICE guidance into practice. Resource impact template: Pembrolizumab with platinum- and fluoropyrimidine-based chemotherapy for untreated advanced oesophageal and gastro-oesophageal junction cancer (TA737). 2021. Available at: https://www.nice.org.uk/guidance/ta737/resources/resource-impact-template-excel-9263083501 [Accessed 15/06/2022].
4. Latimer NR, Abrams KR. NICE DSU technical support document 16: adjusting survival time estimates in the presence of treatment switching. 2014.
5. National Institute for Health and Care Excellence. Nivolumab for previously treated unresectable advanced or recurrent oesophageal cancer (TA707). 2021. Available at: https://www.nice.org.uk/guidance/ta707 [Accessed 15/06/2022].
6. Chin K, Kato K, Cho BC, et al. Three-year follow-up of ATTRACTION-3: A phase III study of nivolumab (Nivo) in patients with advanced esophageal squamous cell carcinoma (ESCC) that is refractory or intolerant to previous chemotherapy. Journal of Clinical Oncology. 2021;39(3_suppl):204-.
7. National Institute for Health and Care Excellence. Nivolumab for treating advanced (unresectable or metastatic) melanoma (TA384). 2016. Available at: https://www.nice.org.uk/guidance/ta384 [Accessed 15/06/2022].
8. National Institute for Health and Care Excellence. Nivolumab in combination with ipilimumab for treating advanced melanoma (TA400). 2016. Available at: https://www.nice.org.uk/guidance/ta400 [Accessed 15/06/2022].
9. National Institute for Health and Care Excellence. Nivolumab for adjuvant treatment of completely resected melanoma with lymph node involvement or metastatic disease (TA558). 2019. Available at: https://www.nice.org.uk/guidance/ta558 [Accessed 15/06/2022].
10. National Institute for Health and Care Excellence. Nivolumab with platinum- and fluoropyrimidine-based chemotherapy for untreated HER2-negative advanced gastric, gastro-oesophageal junction or oesophageal adenocarcinoma [ID1465]. 2022. Available at: https://www.nice.org.uk/guidance/indevelopment/gid-ta10352 [Accessed 28/07/2022].
11. Bond S. RDocumentation: rpsftm. 2020. Available from: https://www.rdocumentation.org/packages/rpsftm/versions/1.2.7.

Appendix A: additional analyses

Correction to ERG model in Pembrolizumab plus chemotherapy scenario

When implementing the analyses described in this response, an error was discovered in the ERG's implementation of the scaling of the chemotherapy OS model by time-varying hazard ratio to represent Pembrolizumab OS. This error was sufficient to result in overall survival > 1 in the Pembrolizumab arm, and so correction of this error was undertaken prior to further analysis.

The formula in Cell S2100 of the ERG model was:

=IF(ERG_1=2,IFERROR(EXP(J11+LN(Q10)),Q11),Q2100)

With this formula copied down the "OS" trace until maximum model timestep. Here, J11, labelled "Pembro_HR" contained the difference in accumulated hazard of mortality due to the OS model on the NIVO-CHEMO arm, multiplied by time-varying hazard ratio calculated on the "Data_ERG" sheet:

=IF(I11>208,1,HLOOKUP(I11+13,Data_ERG!B$3:K$4,2))*(-LN(Q11)+LN(Q10))

This is intended to represent the cyclic accumulated hazard of mortality in the pembrolizumab arm. However, the hazard ratios are expressed as the ratio of nivolumab plus chemotherapy hazards to pembrolizumab plus chemotherapy in the reference table, and it is thus necessary to use the reciprocal of the hazard ratio, in this first correction:

=IF(I11>208,1,(1/HLOOKUP(I11+13,Data_ERG!B$3:K$4,2)))*(-LN(Q11)+LN(Q10))

A second correction is necessary to the formula in Cell S2100. In this original formulation, it is added to LN(Q10), i.e. -1 times the accumulated hazard *upon nivolumab* to the previous model cycle. As the intention is to represent the survival due to the accumulation of hazard upon the scaled (pembrolizumab) arm, the formula was corrected to:

=IF(ERG_1=2,IFERROR(EXP(-(J11-LN(S2099))),S2099),Q2100)

The ERG base case for NIVO-CHEMO versus PEMBRO-CHEMO is thus corrected (Table 8). In the corrected model, there is a greater difference in QALYs between treatments, as in the original version where the deviation between the overall

survival curves was limited. Nevertheless, there is no change in the conclusion that NIVO-CHEMO is cost-effective versus PEMBRO-CHEMO.

Table 8. Impact of correction to ERG model in Pembrolizumab plus chemotherapy

			Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
NIVO-CHEMO versus PEMBRO-CHEMO	ERG base case	EAG report	[redacted]	[redacted]	£290,554(CE region)
		Corrected	[redacted]	[redacted]	£104,998 (CE region)

Issue 1 – Comparison of NIVO-CHEMO versus CHEMO in PD-L1 ≥ 1% TC and <10 CPS population

The ERG has requested an analysis comparing NIVO-CHEMO versus CHEMO in CheckMate 648 patients with PD-L1 ≥ 1% TC and <10 CPS population. Clinically, this subpopulation does not exist as medical decisions over which drug to use would be based on CPS or TC, rather than both. Pathologists contacted by BMS would not want to do both tests, as this will require staining and counting of two sets of slides, taking up time and resource in a busy department. Further, as stated during the response to clarification questions and outlined on page 4 of this response, the sample size calculation of CheckMate 648 is based on PD-L1 TC ≥ 1% and not CPS, so that the study is not powered for a subgroup analysis of PD-L1 TC ≥ 1% and CPS < 10%.

However, in order to address uncertainty, this subgroup analysis is provided below. This was used to inform two scenarios: one reflecting the impact of the selected survival curves on the company base case analysis and one reflecting the impact on the ERG base case analysis. Overall survival (OS), progression-free survival (PFS) and time on treatment (ToT) were updated to reflect the subgroup of interest. Additionally, the nivolumab patient access was updated to reflect that described in the updated patient access scheme section. All other parameters remained as per the respective base case analyses in the company submission.

Survival curve selection

CHEMO overall survival

Figure 4 and Figure 5 present the parametric and semi-parametric extrapolations, respectively, for CHEMO overall survival. No fully parametric models provided a good fit to the available data, particularly in the first six months and after 30 months. The exponential, Weibull and Gompertz semi-parametric fits provided a good visual fit to the data and very similar goodness-of-fit statistics. The Gompertz semi-parametric fit was selected, as this predicted slightly improved outcomes for CHEMO.

Figure 4. CheckMate 648 PD-L1 TC ≥ 1% and CPS < 10: CHEMO overall survival (parametric extrapolation)

Figure 5. CheckMate 648 PD-L1 TC ≥ 1% and CPS < 10: CHEMO overall survival (semi-parametric extrapolation – cutpoint at 6.9 months)

CHEMO progression-free survival

Figure 6 and Figure 7 present the parametric and semi-parametric extrapolations, respectively, for CHEMO PFS. As can be seen, the parametric extrapolations provide poor fits to the data. The semi-parametric Weibull function provided a good visual fit to the data, while predicting a plausible long-term extrapolation. As a result, this function was selected.

Figure 6. CheckMate 648 PD-L1 TC ≥ 1% and CPS < 10: CHEMO progression-free survival (parametric extrapolation)

Figure 7. CheckMate 648 PD-L1 TC ≥ 1% and CPS < 10: CHEMO progression-free survival (semi-parametric extrapolation – cutpoint at 6.9 months)

CHEMO time on treatment

Kaplan-Meier estimates of time on treatment (presented in Figure 8) were complete at the end of the trial follow-up period, in that the number of patients at risk of discontinuation at the end of follow-up was 0. As such the Kaplan-Meier curves themselves were used in the model to estimate time on treatment, ensuring complete consistency with the clinical trial data.

Figure 8. CheckMate 648 PD-L1 TC≥1% and CPS<10: CHEMO time on treatment (Kaplan-Meier)

NIVO-CHEMO overall survival

Figure 9 and Figure 10 present the parametric and semi-parametric extrapolations, respectively, for NIVO-CHEMO overall survival. Aligned with the approach used in the CHEMO arm, a semi-parametric approach was selected. Of these extrapolations, Weibull provided a good visual fit to the data and the lowest goodness-of-fit statistics. Further, the predicted extrapolation was among the most conservative options. As such, this extrapolation was selected.

Figure 9. CheckMate 648 PD-L1 TC≥1% and CPS<10: NIVO-CHEMO overall survival (parametric extrapolation)

Figure 10. CheckMate 648 PD-L1 TC≥1% and CPS<10: NIVO-CHEMO overall survival (semi-parametric extrapolation – cutpoint at 6.9 months)

NIVO-CHEMO progression-free survival

Figure 11 and Figure 12 present the parametric and semi-parametric extrapolations, respectively, for NIVO-CHEMO PFS. Aligned with the approach used in the CHEMO arm, a semi-parametric approach was selected. The semi-parametric lognormal function provided a good visual fit to the data and the optimal goodness-of-fit statistics, while predicting a plausible long-term extrapolation. As a result, this function was selected.

[redacted]

Figure 11. CheckMate 648 PD-L1 TC≥1% and CPS<10: NIVO-CHEMO progression-free survival (parametric extrapolation)

[redacted]

Figure 12. CheckMate 648 PD-L1 TC≥1% and CPS<10: NIVO-CHEMO progression-free survival (semi-parametric extrapolation – cut-off point at 6.9 months)

NIVO-CHEMO time on treatment

Kaplan-Meier estimates of time on treatment (presented in Figure 13) were complete at the end of the trial follow-up period, in that the number of patients at risk of discontinuation at the end of follow-up was 0. As such the Kaplan-Meier curves themselves were used in the model to estimate time on treatment, ensuring complete consistency with the clinical trial data.

[redacted]

Figure 13. CheckMate 648 PD-L1 TC≥1% and CPS<10: NIVO-CHEMO time on treatment Kaplan-Meier

Cost-effectiveness analysis

Following selection of the survival extrapolations, two scenarios were undertaken: one reflecting the impact on the company base case analysis and one reflecting the impact on the ERG base case analysis. OS, PFS and ToT were updated to reflect the subgroup of interest (Table 9). Additionally, the nivolumab patient access was

updated to reflect that described in the updated patient access scheme section. All other parameters remained as per the respective base case analyses.

Table 9. CheckMate 648 PD-L1 ≥1% TC and <10 CPS survival extrapolations

	CHEMO	NIVO-CHEMO
OS	Semi-parametric (6.9 month cut point); Gompertz	Semi-parametric (6.9 month cut point); Weibull
PFS (BICR)	Semi-parametric (6.9 month cut point); Weibull	Semi-parametric (6.9 month cut point); Lognormal
ToT	Kaplan-Meier	Kaplan-Meier

As shown in Table 10 economic modelling of survival data for this subgroup marginally decreased QALY accrual in the NIVO-CHEMO and CHEMO arms. The resulting ICER was lower than the company's base case analysis and remains well below a willingness-to-pay threshold of £50,000/QALY.

Table 10. Cost-effectiveness results: impact of PD-L1 ≥1% TC and <10 CPS subgroup analysis on company updated base case analyses

		Total Costs (£)	Total QALYs	ICER (£/QALY)	Updated Base case ICER (£/QALY)
Impact on company base case analysis	Nivolumab, cisplatin and fluorouracil	[redacted]	[redacted]	**£29,717**	**£27,106**
	Cisplatin and fluorouracil	[redacted]	[redacted]		
	Incremental	[redacted]	[redacted]		

Issue 4 – Impact of subsequent therapies

Adjustment of outcomes to reflect modelled subsequent therapies

In the economic model, it is assumed that a fraction of all patients, on both arms, will commence a second line therapy after discontinuing first line therapy. This fraction is based upon the fraction of patients in CheckMate 648 who commenced a subsequent therapy and is set according to the treatment arm. In the trial, these patients received various therapies, in some cases multiple treatment lines. In the economic model, a single second line of therapy is modelled as the dominant driver of differential costs. For patients receiving NIVO-CHEMO, the fraction of patients receiving second line therapy are all assumed to receive taxane monotherapy, with costs split equally between docetaxel and paclitaxel. For patients receiving CHEMO, the fraction of patients receiving second line therapy are all assumed to receive nivolumab monotherapy.

As this treatment mix differs from that observed in the trial, the ERG has requested that an attempt be made to adjust the outcomes or models thereof to improve prediction of outcomes in a cohort receiving the costed second-line treatments. The adjustments may be considered in two parts:

1. Adjustment of NIVO-CHEMO overall survival to remove the effect of subsequent PD-1 inhibitors used in the CheckMate 648 trial that are not part of the modelled treatment pathway

2. Adjustment of CHEMO overall survival to extend the effect of subsequent PD-1 inhibitors to all patients receiving subsequent therapy, whilst within the trial a small traction of CHEMO patients went on to receive PD-1 inhibitors.

The data was inspected to determine the plausibility and necessity of adjustment for these two parts. Data from the ATTRACTION-3 trial, which tested nivolumab monotherapy versus taxanes in second-line oesophageal cancer, was used to provide context to this exploration.

Technical engagement response form

Multiple methods for adjusting for subsequent therapies exist in the literature, including those mentioned in TSD 16,[4] which implies further development of methods, one of which (the two-step method) is demonstrated in Ouwens et al. (2021).[2] As models of overall survival after commencement of second-line therapy are described in TA707, informed by ATTRACTION-3 data, methods of adjustment within the economic model were considered, as well as patient-level adjustment by methods such as Rank-Preserving Structural Failure Time Models (RPSFTM).

Outcomes in CheckMate 648 and ATTRACTION-3

To determine the applicability of ATTRACTION-3 data to the modelled population, a visual comparison was made of Kaplan-Meier estimated survival after commencement of second-line systemic therapy (for patients with a known start date of second line therapy) in the PD-L1 ≥ 1% subgroup of CheckMate 648 and the same subgroup of ATTRACTION-3. Data was stratified by randomised trial arm in both cases, and for CheckMate 648, also by whether the patient received a PD-1 inhibitor at any subsequent treatment line.

Table 11 and Figure 14 demonstrate that outcomes upon PD-1 inhibitors for CHEMO receiving patients matched well those in ATTRACTION-3. By contrast, for CHEMO patients receiving subsequent therapies without PD-1 inhibitors, outcomes were worse than those receiving taxanes in ATTRACTION-3. As ATTRACTION-3 was a randomised, controlled trial, its treatment effect would be expected to be less biased than that demonstrated by the observational CHEMO arm, and it is thus probable that there is some selection bias, and it is likely that the survival curve suggested by this arm would over-estimate outcomes for the patients not receiving anti-PD-1 therapy in the counterfactual. This is supportive of the use of model-based adjustment techniques or patient-level adjustment using conditional models. The NIVO-CHEMO arm shows a heterogenous profile with respect to patients not receiving additional PD-1 therapy, demonstrating higher hazard than the taxane arm of ATTRACTION-3 initially, then lower, sufficient that the two survival curves almost meet at 18 months. By contrast, for patients receiving additional anti-PD-1 therapy,

survival is initially elevated versus ATTRACTION-3, then decreases. Some of this effect will be immortal bias due to selecting patients receiving anti-PD-1 therapy at any subsequent line; a necessary measure as only 3 patients received anti-PD1 therapy at second line in the NIVO-CHEMO arm, and 8 in lines 2 or 3. By contrast, 20/22 patients receiving anti-PD-1 therapy on the CHEMO arm were doing so by line 3, so the bias should be more limited in this arm. Nevertheless, it is difficult to draw conclusions for NIVO-CHEMO. There is insufficient evidence to suggest different outcomes on second-line nivolumab for anti-PD-1 experienced and anti-PD-1 naïve patients, and RPSFTM may be possible.

The outcomes upon nivolumab monotherapy in ATTRACTION-3 were deemed to be sufficiently similar to those of patients in CheckMate 648 to enable use of outcomes models derived from this trial to inform expected post second line survival of nivolumab-receiving patients in the modelled population.

Table 11. Summary of OS from second line therapy - CheckMate 648 and ATTRACTION-3

Population	N	Events	Median (95% CI) [months]
ATTRACTION-3 Taxane	[redacted]	[redacted]	[redacted]
CM648 CHEMO No subsq. anti PD-1	[redacted]	[redacted]	[redacted]
CM648 NIVO-CHEMO No subsq. antiPD-1	[redacted]	[redacted]	[redacted]
ATTRACTION-3 Nivolumab	[redacted]	[redacted]	[redacted]
CM648 CHEMO subsq. antiPD-1	[redacted]	[redacted]	[redacted]
CM648 NIVO-CHEMO subsq. antiPD-1	[redacted]	[redacted]	[redacted]
CM648 population is limited to patients with known second line systemic therapy start date. All patients in PD-L1 ≥ 1% (TC) subgroup of respective trials.Median by Kaplan-Meier estimator with Greenwood confidence intervals *CM648: CheckMate 648; subsq. antiPD-1: anti-PD1 therapy on any treatment line after first.*			

[redacted]

Figure 14. Kaplan-Meier of OS from second-line therapy – CheckMate 648 and ATTRACTION-3

RPSFTM adjustment

Whilst it was considered a risk that the treatment effect of nivolumab as a second-line therapy may be overestimated by use of the CheckMate 648 data, the RPSFTM method was attempted as part of a series of scenario analyses.

In brief, RPSFTM attempts to adjust the overall survival times of patients receiving an intervention to match those of those who do not receive that intervention by use of an acceleration factor (ψ), which scales up or down the survival times of the patients. In the general case, this acceleration factor does not apply for the whole time that a patient is followed up, but for the period where they are exposed to the intervention of interest, so their survival is unmodified prior to receiving the therapy, and accelerated or decelerated subsequently. ψ is sought to minimise the difference between the arms of the trial, using the trial randomisation to assist in finding an unbiased ψ. A statistic, such as that used for the log-rank test of survival difference (used here), is minimised to inform the optimisation problem.

The most significant assumption of the RPSFTM method is that it assumes the same treatment effect can apply to all patients, at all times. In this case, this implies that the same proportional delay to expected time of death would apply to the population without prior systemic therapy, and to the population who have been previously treated with CHEMO. Heterogeneity of response cannot be introduced by this treatment effect, as it serves only to proportionally extend or shorten individual survival, and so the hazard function of the modified data may not be representative of a population with mixed response.

An RPSFTM model was fitted using the “rpsftm” package in the “R” software environment.[11] Re-censoring was not used, as this did not appear to be correctly implemented at the time of analysis (v1.2.7) and gave counter-intuitive results. Exposure was considered to be the time from first exposure to PD-1 inhibitor until end of follow-up or death. Thus, patients on the NIVO-CHEMO arm were considered

exposed for their full follow-up, whilst patients on CHEMO were considered exposed from the start time of any subsequent line containing an anti-PD-1.

ψ was estimated as -0.550, meaning that after receiving NIVO-CHEMO, the expected time of death of a patient would be extended by 1/exp(ψ) = 1.73 times. The model was applied to the same data, in order to remove the PD-1 treatment effect from both arms and demonstrate the model had fitted correctly, by overlapping the adjusted survival estimates (Figure 15)

Figure 15. RPSFTM model - adjusted Kaplan-Meier, full removal of PD-1 effect

The data were then adjusted to remove the effect of only subsequent IO in the NIVO-CHEMO arm, by defining exposure in the same manner as CHEMO; i.e. from second line + exposure until death. The CHEMO arm was adjusted in the opposite direction, decelerating time from commencement of second line therapy until death. The resulting adjusted Kaplan-Meier curve is shown in Figure 16. Adjustment to NIVO-CHEMO results in a very small reduction in survival which is not sustained beyond 30 months; it can be assumed that these longest-living patients received no subsequent anti-PD-1 therapy. By contrast, the CHEMO arm shows some uplift, sufficient to re-order some events and censors and cause the tail of the curve to run slightly lower than previously. The progression into the “plateau” for this adjusted curve is less gradual than with the unadjusted data, which caused concern for model fitting.

Figure 16. RPSFTM model - adjusted Kaplan-Meier, removal of subsequent PD-1 from NIVO-CHEMO, deceleration of CHEMO after start of subsequent systemic therapy.

The plausibility of these adjustments was considered by comparing the post second line survival of the adjusted data to the ATTRACTION-3 survival upon the therapies to which the data were adjusted (Figure 14). These show the expected reduction of

NIVO-CHEMO survival towards the taxane arm and increase of CHEMO survival towards the nivolumab arm. The profile of the CHEMO arm adjusted data was noted to differ substantially from the ATTRACTION-3 nivolumab profile, which was considered to be due to the lack of heterogenous response in the RPSFTM model.

[redacted]

Figure 17 RPSFTM Adjusted Kaplan-Meier of OS from second line therapy - CheckMate 648 and ATTRACTION-3

To use these data for economic evaluation, parametric survival models were fitted as for the unadjusted data.

It was clear from the adjusted KM that there was minimal difference for NIVO-CHEMO. When fitting semi-parametric models to these adjusted data, a greater mean was predicted, which lacked plausibility. Therefore, the original base case models are used in the RPSFTM scenario.

For the CHEMO arm, it was considered that plausibility must be maintained under consideration that TA707 predicted a mean survival of 1.650 years after commencement of second line nivolumab, compared to 0.997 on taxane, with an increment of 0.653 years. Applying this to 56.67% of the cohort (the proportion modelled as commencing second-line therapy), an increase on the order of 0.370 life years (4.44 months) from the original estimation would be considered plausible, ignoring the contribution of patients already receiving PD-1 inhibitors on subsequent therapy lines. Given that the presented first-line base case predicts [redacted] undiscounted life years ([redacted] months) for CHEMO ([redacted] than the ERG scenario, which predicts [redacted] life years), an upper bound for plausible mean survival after adjustment would be [redacted] life years ([redacted] months).

For further context, applying this simple treatment effect to the [redacted] patients in the CheckMate 648 NIVO-CHEMO PD-L1 ≥ 1% subgroup would be expected to reduce mean survival by [redacted] life years ([redacted] months), under the strong assumption

that the treatment effect of PD-1 inhibitors in experienced patients is similar to that in naïve patients. This is [redacted] of the [redacted] undiscounted life years predicted in the base case, which goes some way to explaining why random variation in the patient-level data resulted in a lower mean when attempting to fit models to the adjusted data, overwhelming the small “signal” of the adjustment.

Figure 18 shows the result of this fitting. In light of the plausible mean established above, only lognormal model was considered feasible in this scenario.

Figure 18. Semi-parametric fits to RPSFTM-adjusted OS data, CHEMO arm of CheckMate 648 (PD-L1 ≥ 1%)

An RPSFTM scenario was thus implemented in the economic model, using the company base case versus CHEMO, and modifying the CHEMO OS curve to use the semi-parametric lognormal model fitted to the RPSFTM adjusted data. Results are shown in Table 12.

Model-based evaluation

In light of the similarity in outcomes for patients receiving subsequent anti-PD-1 therapy in CheckMate 648 and the nivolumab arm of ATTRACTION-3 (Figure 14), a model-based evaluation of outcomes was conducted. For this scenario, the economic model was modified such that patients on CHEMO would discontinue first-line therapy at a rate determined by a model of time to subsequent therapy or death, informed by data from CheckMate 648. At this time, 56.67% of patients would commence second-line therapy and be costed according to the mean time on treatment. They would then be modelled as at risk to a constant hazard of death informed by the mean survival time from the nivolumab arm of the economic model informing TA707, such that patients in this state would have the same mean survival time as in that economic model. The ratio of patients in pre and post progression on subsequent treatment used the ratio evaluated at the same time in the un-modified

NIVO-CHEMO arm. The remaining (1-56.67%) of patients die at the time of evaluation.

A parametric model fitted to the newly defined outcome of time to subsequent therapy or death was fitted. The majority of events occurred within the first year, followed by a reduced rate (Figure 19). The log-logistic model was chosen as this had the best statistical goodness of fit per AIC and BIC and represented well the rate of change of hazard in the second year.

This model and an exponential rate giving a mean of 1.650 years for post second-line survival were used as inputs for the modified economic model. Results are presented in section Issue 4 – Impact of subsequent therapies: Cost-effectiveness results.

Figure 19. Models of time to death or second line treatment, CHEMO

Cost-effectiveness results

Table 12 presents the results of the RPSFTM informed model scenario, and the scenario whereby post-second-line mortality was informed by ATTRACTION-3 nivolumab. As expected, by modelling the superior outcomes in ATTRACTION-3, which represent a much larger treatment effect upon the CheckMate 648 population than that observed in ATTRACTION-3, reduces the incremental benefit of NIVO-CHEMO. Nevertheless, in both cases the ICER remains cost-effective, with only a small increase in the RPSFTM scenario.

Table 12. Cost-effectiveness results: impact of modification of subsequent therapy outcomes

			Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
NIVO-CHEMO versus CHEMO	Updated Company base case	Updated base case	[redacted]	[redacted]	£27,106
		RPSFTM-scenario	[redacted]	[redacted]	£29,253
		Model-based ATTRACTION-3 scenario	[redacted]	[redacted]	£39,868

Amendment to modelling subsequent treatment costs

In order to reduce uncertainty around the impact of subsequent treatments on the decision problem, the representation of second-line treatments within the economic model was re-evaluated.

In the first instance, it was noted, from the committee papers of TA707, that the company model reported similar median time on treatment for taxanes and nivolumab (0.211 years and 0.230 years, respectively), but different means (0.291 years versus 0.496). As nivolumab is associated with heterogenous response, this was considered a plausible difference. Mean time on treatment determines the mean cost to the payer, and therefore this input was considered to be more appropriate than the median.

Survival outcomes for patients in CheckMate 648 (PD-L1 ≥ 1%) who received subsequent systemic therapy were plotted stratified by treatment arm and the use of any PD-1 inhibitor in any subsequent systemic therapy line. These plots overlaid the ATTRACTION-3 PD-L1 ≥ 1% overall survival outcomes (Figure 14). Outcomes for patients randomised to first-line CHEMO who received subsequent PD-1 inhibitors were well matched by the outcomes from the NIVO arm of ATTRACTION-3, supporting the hypothesis that the modelled time on treatment for these patients would be appropriate for the CheckMate 648 CHEMO population.

To improve the representation of the costs associated with these treatments within the economic model, the following modifications were made:

- The costs of these therapies were re-evaluated – in particular, it was assumed that all eligible patients would be receiving nivolumab at the new PAS price. This price was converted to a cyclic (weekly) cost and scaled by the proportion of the total cohort expected to receive a second line therapy, as informed by the rates of subsequent treatment in CheckMate 648. Thus, the mean per patient-cycle cost of taxane therapy was reduced to 53.16% of its full cost, to be applied as second line costs to PD-1 inhibitor first line model arms; the mean per patient-cycle cost of nivolumab was likewise reduced to 56.69% of its full cost, to be applied to CHEMO first-line model arms.

- These costs were converted to an expected lifetime cost, by multiplying the mean per patient-cycle cost by the mean time on treatment, sourced from TA707.

- This expected lifetime cost was applied to all patients at the time of discontinuation from first line therapy. As the mean time on treatment on nivolumab was less than half a year, the discounting error incurred by lumping these costs is low.

The impact of these changes to the modelling of subsequent treatment costs is shown in Table 13. The incremental cost between NIVO-CHEMO and CHEMO is reduced due to the higher modelled mean time on treatment in the CHEMO arm, resulting in a lower ICER and increased cost-effectiveness for NIVO-CHEMO at the first line.

Table 13. Cost-effectiveness results: updated subsequent treatment costs

			Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
NIVO-CHEMO versus CHEMO	Company base case post CQs	Original	[redacted]	[redacted]	£31,826
		Updated costs	[redacted]	[redacted]	£28,664
	ERG base case	Original	[redacted]	[redacted]	£46,599
		Updated costs	Not applied to ERG model		
NIVO-CHEMO versus PEMBRO-CHEMO	Company base case post CQs	Original	[redacted]	[redacted]	£320,254(CE region)
		Updated costs	[redacted]	[redacted]	£320,428(CE region)
	ERG base case	Original	[redacted]	[redacted]	£109,039 (CE region)
		Updated costs	Not applied to ERG model		

Issue 5 – Impact of long-term treatment effect between NIVO-CHEMO and CHEMO

The following scenarios were explored in the economic model and results are presented in Table 14:

- Company updated base case
- ERG base case updated with new nivolumab PAS, updated costs, time on treatment weighted RDI – no treatment waning
- ERG updated base case – waning between 2.5 and 4.0 years
- ERG updated base case – waning between 4.0 and 5.5 years (end of follow-up)
- ERG updated base case – waning between 6.0 and 7.5 years ("7-year wane" scenario)
- ERG updated base case – waning between 4.0 and 10.0 years (end of follow-up, matched hazard ratio gradient)

Table 14. Cost-effectiveness results: treatment waning scenarios

			Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
NIVO-CHEMO versus CHEMO	Updated Company base case	Updated base case	[redacted]	[redacted]	£27,106
	Updated ERG base case	Updated, no Issue 5 – Impact of long-term treatment effect between NIVO-CHEMO and CHEMO	[redacted]	[redacted]	£36,423
		Updated, waning 2.5 – 4.0 years	[redacted]	[redacted]	£42,016
		Updated, waning 4.0 – 5.5 years	[redacted]	[redacted]	£39,364

		Updated, waning 6.0 – 7.5 years	[redacted]	[redacted]	£38,041
		Updated, waning 4.0 – 10.0 years	[redacted]	[redacted]	£38,113

Issue 6 – Impact of Pembrolizumab parameter changes

Amendment of time-varying hazard ratios between Pembrolizumab and Chemotherapy

In the company submission, the tables of time-varying hazard ratio were used to scale the chemotherapy survival curves, and so provide outcomes models for pembrolizumab. In incorporating the scaling within the economic model, the tables have been updated by the appropriate values from the "overlap" analysis described in the NMA report. The results are presented in Table 15.

The ERG's base case did not solely use these hazard ratios; rather, hazard ratios with a reference arm of NIVO-CHEMO were used to scale the NIVO-CHEMO OS curve to form the PEMBRO-CHEMO OS curve. As such, the ERG base case results are not updated by this change.

Table 15. Cost-effectiveness results: updated PEMBRO-CHEMO hazard ratios

			Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
NIVO-CHEMO versus PEMBRO-CHEMO	Company base case post CQs	Original	[redacted]	[redacted]	£320,254 (CE region)
		Updated hazard ratios	[redacted]	[redacted]	£1,153,385 (CE region)

Issue 10 – Impact of updated RDI calculation

- ERG Scenario: Mean of patient RDIs based upon assuming midpoint of category from CSR table

- ERG scenario update: Mean of RDI using all patients with at least 1 dose of named therapy component in CheckMate 648 (PD-L1 ≥ 1%), as of the [redacted] DBL. This is a direct improvement on ERG data, where the same assumptions apply.

- ERG scenario improvement: Weighted mean of RDI where weight is by patient time on treatment for all patients with at least 1 dose of named therapy component in CheckMate 648 (PD-L1 ≥ 1%) as of the [redacted] DBL. This is an approximation to the grand mean RDI, i.e all doses dispensed / all doses expected, based upon total time on treatment. This modifies the company base case.

Table 16. Original RDI, mean RDI of PLD, and mean weighted RDI of PLD

Treatment		Delayed dose modifier	Original RDI	RDI – mean of patient-level data	RDI – mean of patient-level data weighted by time on treatment
Nivolumab in combination with chemotherapy (fluorouracil plus cisplatin)	Nivolumab	[redacted]	[redacted]	[redacted]	[redacted]
	Fluorouracil	[redacted]	[redacted]	[redacted]	[redacted]
	Cisplatin	[redacted]	[redacted]	[redacted]	[redacted]
Chemotherapy (fluorouracil plus cisplatin)	Fluorouracil	[redacted]	[redacted]	[redacted]	[redacted]
	Cisplatin	[redacted]	[redacted]	[redacted]	[redacted]
Pembrolizumab in combination with chemotherapy (fluorouracil plus cisplatin)	Pembrolizumab	[redacted]	[redacted]	[redacted]	[redacted]
	Fluorouracil	[redacted]	[redacted]	[redacted]	[redacted]
	Cisplatin	[redacted]	[redacted]	[redacted]	[redacted]
* The ERG scenario versus PEMBRO-CHEMO used the original delayed dose modifier for both treatment arms.					

Cost-effectiveness outcomes

The time-on-treatment weighted RDI estimate increases the second-line costs in CHEMO relative to NIVO-CHEMO, decreasing the incremental cost and improving the ICER (Table 17).

Technical engagement response form

Nivolumab platinum-based chemotherapy for untreated advanced oesophageal cancer with tumour cell PD-L1 expression ≥1% [ID2712]

Table 17. Cost-effectiveness results: impact of updated RDI calculation

			Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
NIVO-CHEMO versus CHEMO	Company base case post CQs	Original – delayed dose modifier	[redacted]	[redacted]	£31,826
		Updated RDI – Time on treatment weighted	[redacted]	[redacted]	£30,183
	ERG base case	Original RDI	[redacted]	[redacted]	£46,599
		Updated RDI – unweighted	[redacted]	[redacted]	£42,959
		Updated RDI – Time on treatment weighted	[redacted]	[redacted]	£41,550
NIVO-CHEMO versus PEMBRO-CHEMO	Company base case post CQs	Original – delayed dose modifier	[redacted]	[redacted]	£320,254 (CE region)
		Updated RDI – Time on treatment weighted	[redacted]	[redacted]	£298,853 (CE region)
	ERG base case	Original – delayed dose modifier	[redacted]	[redacted]	£109,039 (CE region)
		Updated RDI - unweighted	[redacted]	[redacted]	£105,386 (CE region)
		Updated RDI – Time on treatment weighted	[redacted]	[redacted]	£102,286 (CE region)
All results with updated PAS for nivolumab					

Issue 11 – Impact of updated costs

Updated cost inputs

The updated drug acquisition costs are presented in Table 18 and the additional costs in Table 19.

Table 18.Updated drug acquisition costs

Resource	Description	Currency code	Original input	Updated input
Cisplatin	100 mg/100 ml solution for infusion vials	DHA010	eMIT 2020 - £8.73	eMIT 2021 - £8.97
	50 mg/50 ml solution for infusion vials	DHA011	eMIT 2020 - £5.38	eMIT 2021 - £6.03
Fluorouracil	1 g/20 ml (5%) solution for infusion vial	DHA265	eMIT 2020 - £2.35	eMIT 2021 - £2.25
	2.5 g/100 ml (2.5%) solution for infusion vial	DHA024	eMIT 2020 - £3.79	eMIT 2021 - £4.32
	2.5 g/50 ml (5%) solution for infusion vial	DHA102	eMIT 2020 - £4.01	eMIT 2021 - £4.21
	500 mg/10 ml (5%) solution for infusion vial	DHA240	eMIT 2020 - £1.77	eMIT 2021 - £2.86
	5 g/100 ml (5%) solution for infusion vials	DHA137	eMIT 2020 - £8.58	eMIT 2021- £9.2
Capecitabine	150 mg tablets pack size 60	DHA224	eMIT 2020 - £4.43	eMIT 2021 - £4.43
	300 mg tablets pack size 60	DKE068	eMIT 2020 - £7.77	eMIT 2021 - £7.77
	500 mg tablets pack size 120	DHA225	eMIT 2020 - £26.30	eMIT 2021 - 26.30
Docetaxel	160mg/8ml solution for infusion (20mg/ml)	DHC046	eMIT 2020 - £17.95	eMIT 2021 - £17.38
	Docetaxel 20mg/1ml solution for infusion vials (20mg/ml)	DHC025	eMIT 2020 - £3.77	eMIT 2021 - £3.56
	80mg/4ml solution for infusion vials (20mg/ml)	DHC029	eMIT 2020 - £9.13	eMIT 2021 - £8.90
Paclitaxel	100mg/16.7ml solution for infusion vials	DHA145	eMIT 2020 - £7.22	eMIT 2021 - £8.06
	150mg/25ml solution for infusion vials	DHA297	eMIT 2020 - £12.41	eMIT 2021 - £10.15
	300mg/50ml solution for injection vials	DHA210	eMIT 2020 - £17.66	eMIT 2021 - £15.97
	30mg/5ml solution for infusion vials	DHA144	eMIT 2020 - £4.41	eMIT 2021 - £4.15

Table 19. Additional cost updates

Resource	Description	Currency code	Company source and unit cost	ERG preferred source and unit cost
Administration costs				
Deliver simple parenteral chemotherapy at first attendance	Day case and reg day/night	SB12Z	NHS reference cost 2015/16 (weighted average)- £284.05	NHS reference cost 2019/20- £299.61
Deliver complex chemotherapy, including prolonged infusion treatment, at first attendance	Day case and reg day/night	SB14Z	NHS reference cost 2015/16- £431.72	NHS reference cost 2019/20- £431.72
Disease monitoring costs				
CT scan	CT scan of three areas, without contrast	RD25Z	NHS 2019/20 - £103.31	NHS 2019/20 - £103.34
Blood test	Haematology	DAPS05	NHS 2019/20 - £2.53	NHS 2019/20 - £2.53
Kidney	Special screening, examinations or other genetic disorders	WH15Z	NHS 2019/20 - £33.80	NHS 2019/20 - £33.80
Hepatic	Special screening, examinations or other genetic disorders	WH15Z	NHS 2019/20 - £33.80	NHS 2019/20 - £33.80
Consultant	Non-admitted face-to-face attendance, follow-up	WF01A: service code 370	NHS 2019/20 - £203.14	NHS 2019/20 - £200

Cost-effectiveness outcomes

As shown in Table 20**Error! Reference source not found.**, updating these cost inputs had limited impact. The resulting ICERs remained below a willingness-to-pay threshold of £50,000/QALY.

Table 20. Cost-effectiveness results: impact of updates to modelled costs

			Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
NIVO-CHEMO versus CHEMO	Company base case post CQs	Original	[redacted]	[redacted]	£33,357
		Updated costs	[redacted]	[redacted]	£32,123
	ERG base case	Original*	[redacted]	[redacted]	£46,599
		Updated costs	[redacted]	[redacted]	£47,093
NIVO-CHEMO versus PEMBRO-CHEMO	Company base case post CQs	Original	[redacted]	[redacted]	£320,254 (CE region)
		Updated costs	[redacted]	[redacted]	£327,407 (CE region)
	ERG base case	Original	[redacted]	[redacted]	£109,039 (CE region)
		Updated costs	[redacted]	[redacted]	£111,300 (CE region)
*All analyses use updated PAS					

Issue 12 – Pembrolizumab cost update and PSA

Table 21 provides the results of updating the pembrolizumab cost upon the company base case after Clarification Questions, i.e. prior to application of the updated nivolumab PAS or any other changes. Doing so moves nivolumab from dominated to the cost-effective quadrant, although the incremental benefits between the two treatments are small.

Table 21. Cost-effectiveness results: impact of updates to pembrolizumab cost

		Total Costs (£)	Total QALYs	ICER (£/QALY)	Base case ICER* (£/QALY)
Impact on company base case analysis	Nivolumab, cisplatin and fluorouracil	[redacted]	[redacted]	£307,447 (CE region)	-£5,594 (not CE)
	Pembrolizumab, cisplatin and fluorouracil	[redacted]	[redacted]		
	Incremental	[redacted]	[redacted]		
*Before Nivolumab PAS update and update of pembrolizumab costs					

Updates to the economic model were made to allow for greater transparency of input derivation (Issue 6 – Impact of Pembrolizumab parameter changes). Deterministic

(Table 22) and probabilistic (Table 23) sensitivity analysis were run in the PEMBRO-CHEMO scenario.

The results presented here incorporate all of the improvements to the PEMBRO-CHEMO scenario generated as part of this technical engagement, i.e.:

- Update to pembrolizumab cost (Issue 12 – Pembrolizumab cost update and PSA)
- Update to nivolumab PAS (Updated patient access scheme)
- Updated costs (Issue 11 – Impact of updated costs)
- Update to method of modelling second line therapy costs (**Error! Reference source not found.**)
- Update to pembrolizumab hazard ratios to match NMA report (Issue 6 – Impact of Pembrolizumab parameter changes)
- Update to calculation of treatment modifier by weighted RDI (Issue 10 – Impact of updated RDI calculation)

Survival models are per the company submission, i.e.:

- NIVO-CHEMO PFS – Semi-parametric, 6.9 month cut, generalised gamma
- NIVO-CHEMO OS – Semi-parametric, 6.9 month cut, lognormal
- NIVO-CHEMO ToT – Kaplan-Meier
- PEMBRO-CHEMO PFS – Semi-parametric, 6.9 month cut, Weibull (CHEMO reference) scaled by Weibull NMA HRs
- PEMBRO-CHEMO OS – Semi-parametric, 6.9 month cut, lognormal (CHEMO reference) scaled by lognormal NMA HRs
- PEMBRO-CHEMO Exponential ToT (no variance assumed)

Second line therapy was assumed to be taxane monotherapy for 53.16% of the cohort, for a mean duration of 15.18 weeks, for both arms.

Table 22. Cost-effectiveness results: vs PEMBRO-CHEMO deterministic base case

		Total Costs (£)	Total QALYs	ICER (£/QALY)
Impact on company base case analysis	Nivolumab, cisplatin and fluorouracil	[redacted]	[redacted]	£1,107,961 (CE region, SW quadrant)
	Pembrolizumab, cisplatin and fluorouracil	[redacted]	[redacted]	
	Incremental	[redacted]	[redacted]	
ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year Company base case, updated nivolumab PAS, updated PEMBRO-CHEMO case, updated costs, amendment of modelling subsequent treatment costs, updated RDI,				

Table 23. Probabilistic sensitivity analysis results: NIVO-CHEMO versus PEMBRO-CHEMO

	NIVO-CHEMO	PEMBRO-CHEMO	Incremental
Life years	[redacted]	[redacted]	[redacted]
QALYs	[redacted]	[redacted]	[redacted]
Total costs (£)	[redacted]	[redacted]	[redacted]
ICER (£/QALY)			£6,146,030 (CE region, SW quadrant)
Percentage CE			91.9%
ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year Company base case, updated nivolumab PAS, updated PEMBRO-CHEMO case, updated costs, amendment of modelling subsequent treatment costs, updated RDI,			

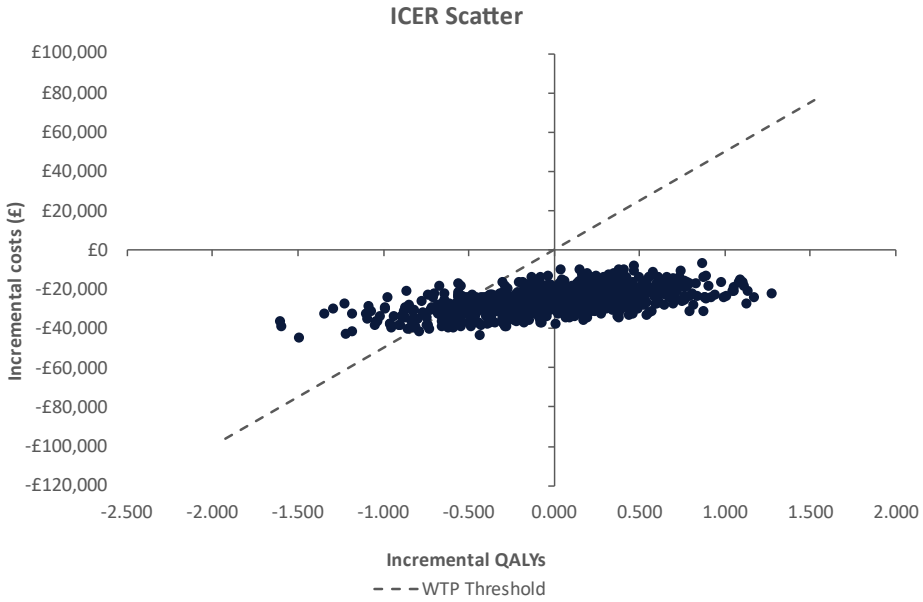


Figure 20. Scatterplot of probabilistic results: NIVO-CHEMO versus PEMBRO-CHEMO – updated base case

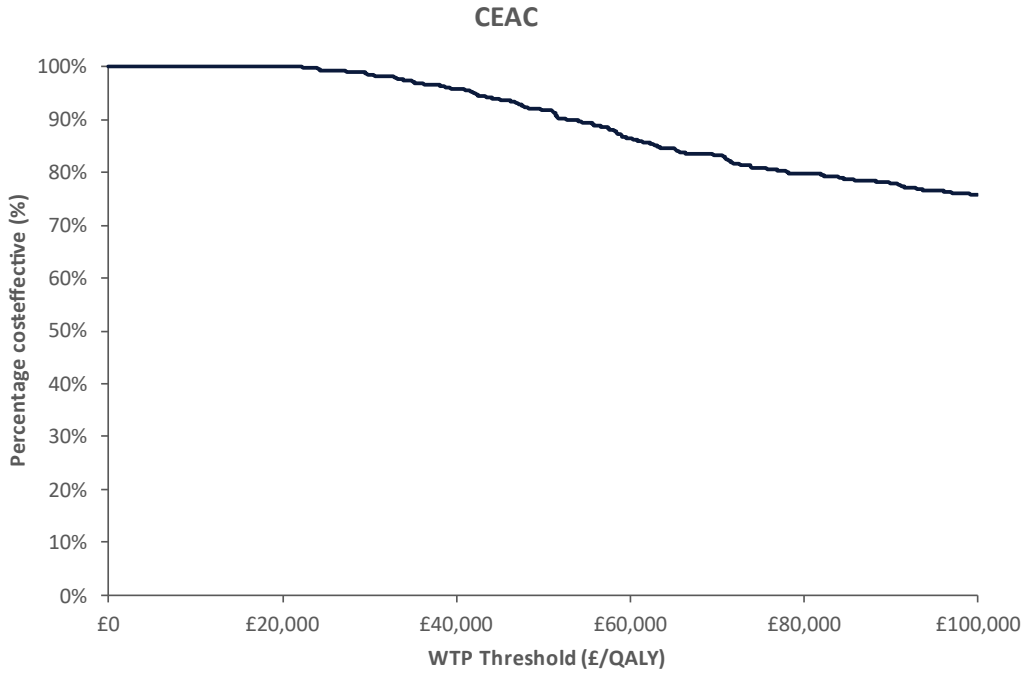


Figure 21. Cost-effectiveness acceptability curve: NIVO-CHEMO versus PEMBRO-CHEMO – updated base case

The deterministic sensitivity analysis results reveal that modification to the treatment cost (directly or by proportion receiving dose) impacts the incremental cost, as expected, and so the ICER. The ICER is most sensitive to changes in the pre and post progression state utility values (Table 24). In the base case, the total utility from both states is well matched across therapies, but the balance between pre- and post-progression is slightly different. Therefore, by varying only one of these components, the balance is lost, and so the ICER is reduced as the incremental QALYs grow absolutely larger, though the total difference in benefits between therapies remains small. The DSA supports the conclusion that due to the low difference in QALYs between these therapies, the ICER is highly sensitive to assumptions regarding this benefit, but NIVO-CHEMO remains cost-effective against PEMBRO-CHEMO in either the south-west or south-east (dominant) quadrant.

Table 24. Deterministic sensitivity results: NIVO-CHEMO versus PEMBRO-CHEMO

Scenario	Parameter variation	Incremental		ICER (£/QALY)
		Costs	QALY	
Base case analysis		[redacted]	[redacted]	**£1,107,961**
Time horizon (weeks	Lower	[redacted]	[redacted]	£700,905
	Upper	[redacted]	[redacted]	£812,391
Cost discount rate (%)	Lower	[redacted]	[redacted]	£1,096,966
	Upper	[redacted]	[redacted]	£1,109,537
Benefit discount rate (%)	Lower	[redacted]	[redacted]	£1,542,244
	Upper	[redacted]	[redacted]	£986,735
Patient age (years)	Lower	[redacted]	[redacted]	£1,210,320
	Upper	[redacted]	[redacted]	£918,140
Patient sex (% male)	Lower	[redacted]	[redacted]	£1,145,161
	Upper	[redacted]	[redacted]	£1,100,456
Pre-progression health state cost	Lower	[redacted]	[redacted]	£1,067,361
	Upper	[redacted]	[redacted]	£1,148,549
Post-progression health state cost	Lower	[redacted]	[redacted]	£1,178,862
	Upper	[redacted]	[redacted]	£1,037,053
Terminal care cost	Lower	[redacted]	[redacted]	£1,107,626
	Upper	[redacted]	[redacted]	£1,108,298

1st Line treatment costs	Lower	[redacted]	[redacted]	£857,263
	Upper	[redacted]	[redacted]	£1,358,664
2nd Line treatment costs	Lower	[redacted]	[redacted]	£1,107,957
	Upper	[redacted]	[redacted]	£1,107,973
Adverse event costs	Lower	[redacted]	[redacted]	£1,107,105
	Upper	[redacted]	[redacted]	£1,108,825
Pre-progression health state utility	Lower	[redacted]	[redacted]	-£172,531
	Upper	[redacted]	[redacted]	£131,559
Post-progression health state utility	Lower	[redacted]	[redacted]	£134,776
	Upper	[redacted]	[redacted]	-£178,107
End of life disutility	Lower	[redacted]	[redacted]	£1,107,360
	Upper	[redacted]	[redacted]	£1,108,569
Adverse event disutility	Lower	[redacted]	[redacted]	£1,107,413
	Upper	[redacted]	[redacted]	£1,108,517
Proportion receiving dose	Lower	[redacted]	[redacted]	£899,089
	Upper	[redacted]	[redacted]	£1,335,308
1st Line adverse event probability	Lower	[redacted]	[redacted]	£1,106,553
	Upper	[redacted]	[redacted]	£1,109,378
2nd Line Time on treatment	Lower	[redacted]	[redacted]	£1,107,975
	Upper	[redacted]	[redacted]	£1,107,958

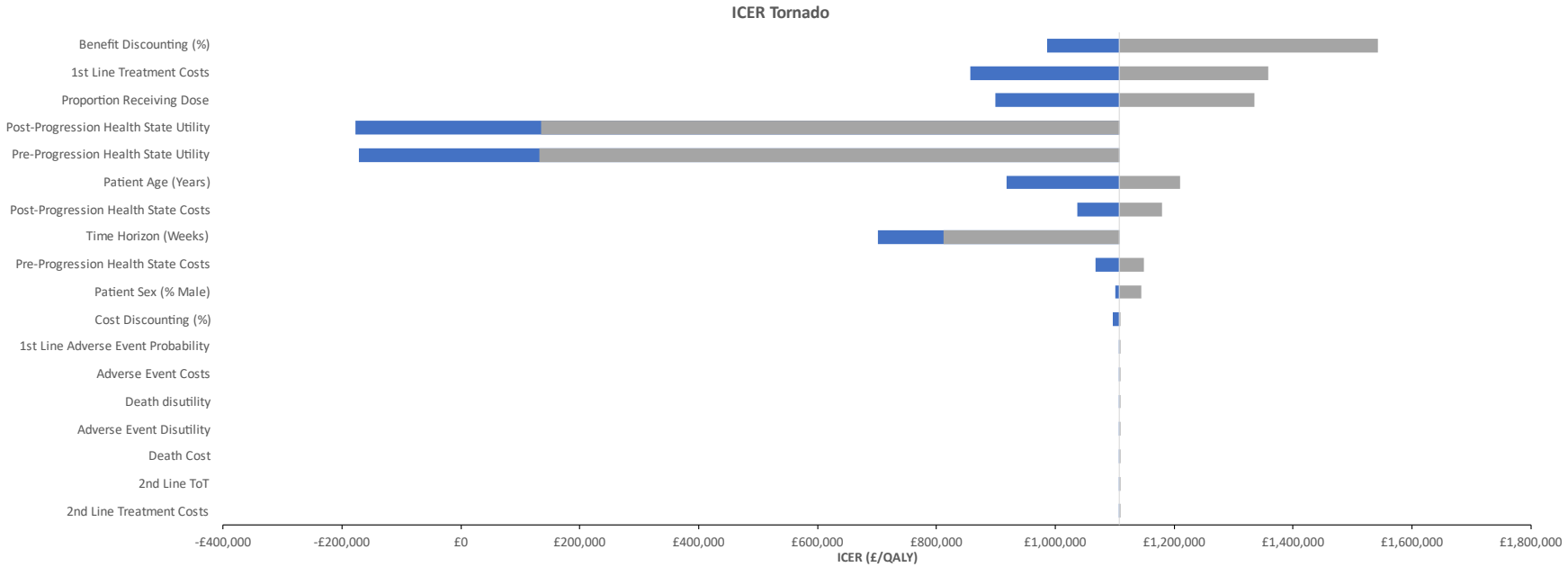


Figure 22. Deterministic sensitivity analysis tornado diagram – NIVO-CHEMO vs PEMBRO-CHEMO

Issue 1 – Comparison of NIVO-CHEMO and PEMBRO-CHEMO

Alternative scenarios for the comparison of NIVO-CHEMO and PEMBRO-CHEMO are presented in Table 25.

Updated company base case

- As described in Issue 12 – Pembrolizumab cost update and PSA, a number of updates have been made to the company base case comparison with PEMBRO-CHEMO, summarised in: Update to pembrolizumab cost (Issue 12 – Pembrolizumab cost update and PSA)
- Update to nivolumab PAS (Updated patient access scheme)
- Updated costs (Issue 11 – Impact of updated costs)
- Update to method of modelling second line therapy costs (**Error! Reference source not found.**)
- Update to pembrolizumab hazard ratios to match NMA report (Issue 6 – Impact of Pembrolizumab parameter changes)
- Update to calculation of treatment modifier by weighted RDI (Issue 10 – Impact of updated RDI calculation)

Survival models are per the company submission, i.e.:

- NIVO-CHEMO PFS – Semi-parametric, 6.9 month cut, generalised gamma
- NIVO-CHEMO OS – Semi-parametric, 6.9 month cut, lognormal
- NIVO-CHEMO ToT – Kaplan-Meier
- PEMBRO-CHEMO PFS – Semi-parametric, 6.9 month cut, Weibull (CHEMO reference) scaled by Weibull NMA HRs
- PEMBRO-CHEMO OS – Semi-parametric, 6.9 month cut, lognormal (CHEMO reference) scaled by lognormal NMA HRs
- PEMBRO-CHEMO Exponential ToT (no variance assumed)

Second line therapy was assumed to be taxane monotherapy for 53.16% of the cohort, for a mean duration of 15.18 weeks, for both arms.

Technical engagement response form

Alternative survival models

The updated company base case was modified by the following survival models:

- NIVO-CHEMO OS – Fully parametric, log-logistic
- PEMBRO-CHEMO PFS – Fully parametric, log-logistic CHEMO survival model scaled by log-logistic NMA HRs
- PEMBRO-CHEMO OS – Fully parametric, log-logistic CHEMO survival model scaled by log-logistic NMA HRs

These models were chosen as satisfying both the ERG's scenario selection, and the NMA models that are highlighted within the NMA report as being best fitting across the network.

This scenario was also explored in probabilistic sensitivity analysis.

Equal efficacy

The updated company base case was modified by setting the comparator OS, PFS and ToT curves equal to the NIVO-CHEMO curves.

Updated ERG's model

The ERG model was updated with the correction detailed in Correction to ERG model in Pembrolizumab plus chemotherapy scenario, and the following changes were made:

- Update to nivolumab PAS
- Update to costs
- Update to weighted RDI

Survival models were as per the model used in the EAG report, i.e. as per the company base case excepting:

- NIVO-CHEMO OS – Fully parametric, log-logistic
- PEMBRO-CHEMO PFS – Per the offline scaling of the CHEMO semi-parametric Weibull PFS curve with Weibull NMA HRs
- PEMBRO-CHEMO OS – scaling of the used NIVO-CHEMO OS curve by time-varying hazard ratios from log-logistic model in the NMA report

Equal efficacy – Updated ERG model

As with the company model, a scenario exploring equal efficacy was undertaken within the updated ERG model.

Table 25. Cost-effectiveness results: impact of alternative NIVO-CHEMO versus PEMBRO-CHEMO comparison approaches

		Incremental Costs (£)	Incremental QALYs	ICER (£/QALY)
Impact on company base case analysis	Updated company submission NMA approach	[redacted]	[redacted]	£1,107,961 (CE region)
	Alternative survival models	[redacted]	[redacted]	-£270,695 (Dominates)
	Equal efficacy	[redacted]	[redacted]	(Dominates)
Impact on ERG base case analysis	Updated ERG model	[redacted]	[redacted]	£133,862 (CE region)
	Equal efficacy	[redacted]	[redacted]	(Dominates)

Table 26. Probabilistic sensitivity analysis results: Alternative survival models

	NIVO-CHEMO	PEMBRO-CHEMO	Incremental
Life years	[redacted]	[redacted]	[redacted]
QALYs	[redacted]	[redacted]	[redacted]
Total costs (£)	[redacted]	[redacted]	[redacted]
ICER (£/QALY)			-£275,994 (Dominates)
Percentage CE			99.4%

ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year

Company base case, updated nivolumab PAS, updated PEMBRO-CHEMO case, updated costs, amendment of modelling subsequent treatment costs, updated RDI,

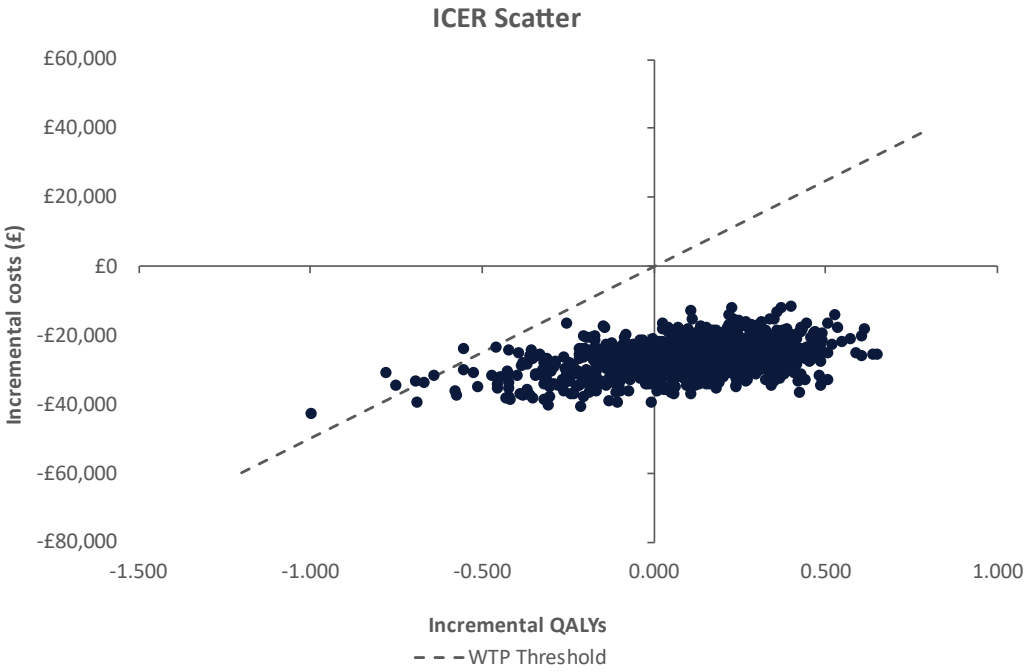


Figure 23. Scatterplot of probabilistic results: NIVO-CHEMO versus PEMBRO-CHEMO – Alternative survival models

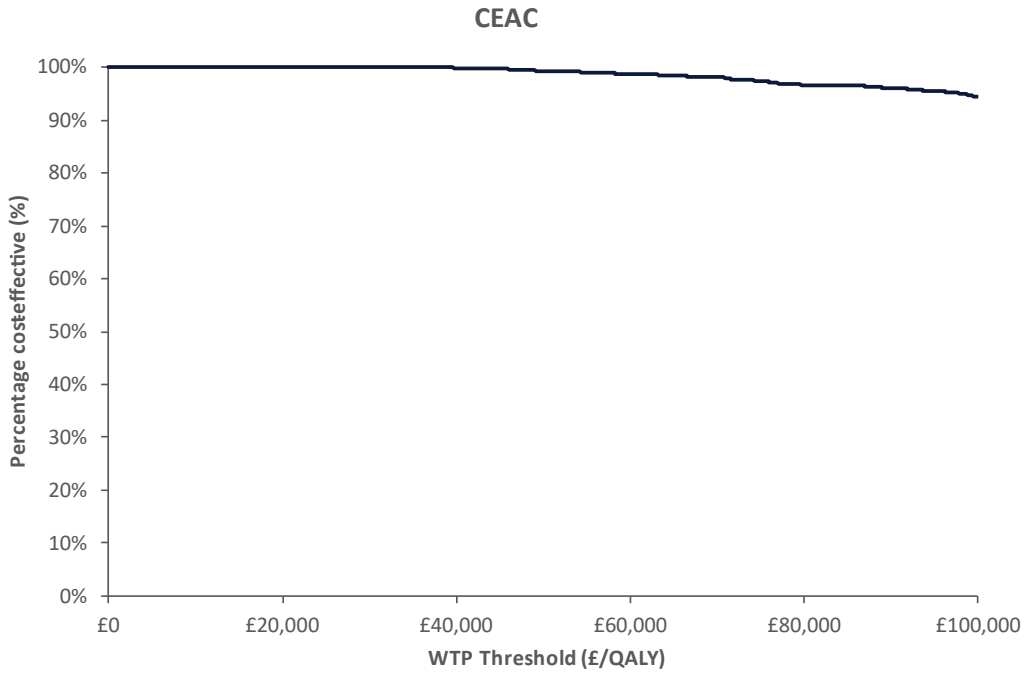


Figure 24. Cost-effectiveness acceptability curve: NIVO-CHEMO versus PEMBRO-CHEMO – alternative survival models

These results are highly supportive of the conclusion that NIVO-CHEMO is likely to be cost-effective versus PEMBRO-CHEMO, and that efficacy differences between the two therapies are minimal, with incremental costs being quite insensitive to the assumed outcomes distribution.

Additional Sensitivity Analysis

NIVO-CHEMO vs CHEMO PSA

Table 27. Probabilistic sensitivity analysis results: NIVO-CHEMO versus CHEMO – updated base case

	NIVO-CHEMO	PEMBRO-CHEMO	Incremental
Life years	[redacted]	[redacted]	[redacted]
QALYs	[redacted]	[redacted]	[redacted]
Total costs (£)	[redacted]	[redacted]	[redacted]
ICER (£/QALY)			£26,288
Percentage CE			94.0%

ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year

Company base case, updated nivolumab PAS, updated costs, amendment of modelling subsequent treatment costs, updated RDI.

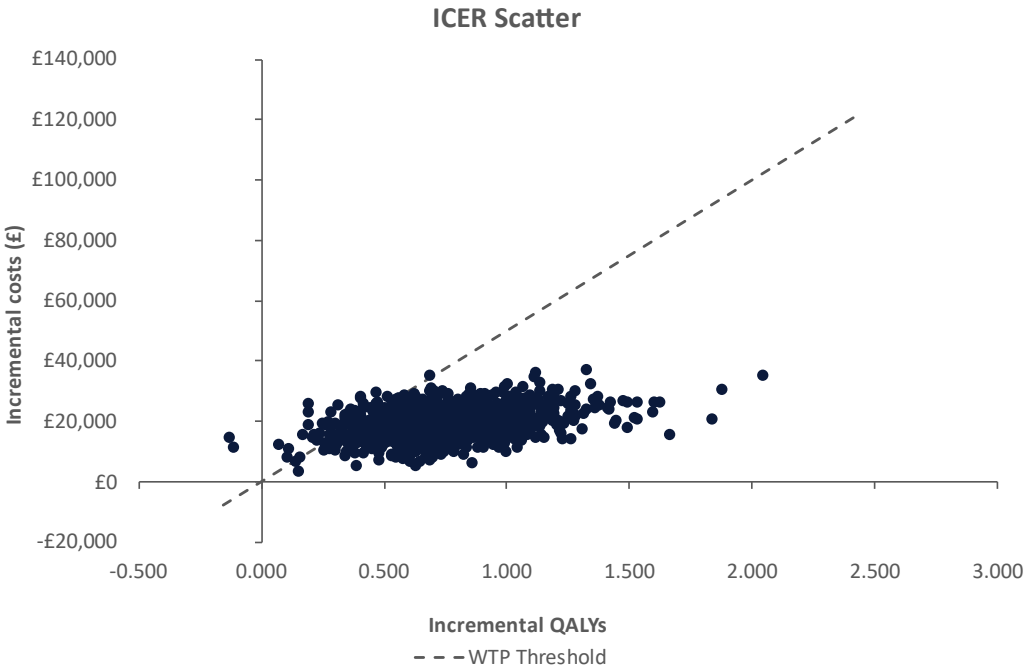


Figure 25. Scatterplot of probabilistic results: NIVO-CHEMO versus CHEMO – updated base case

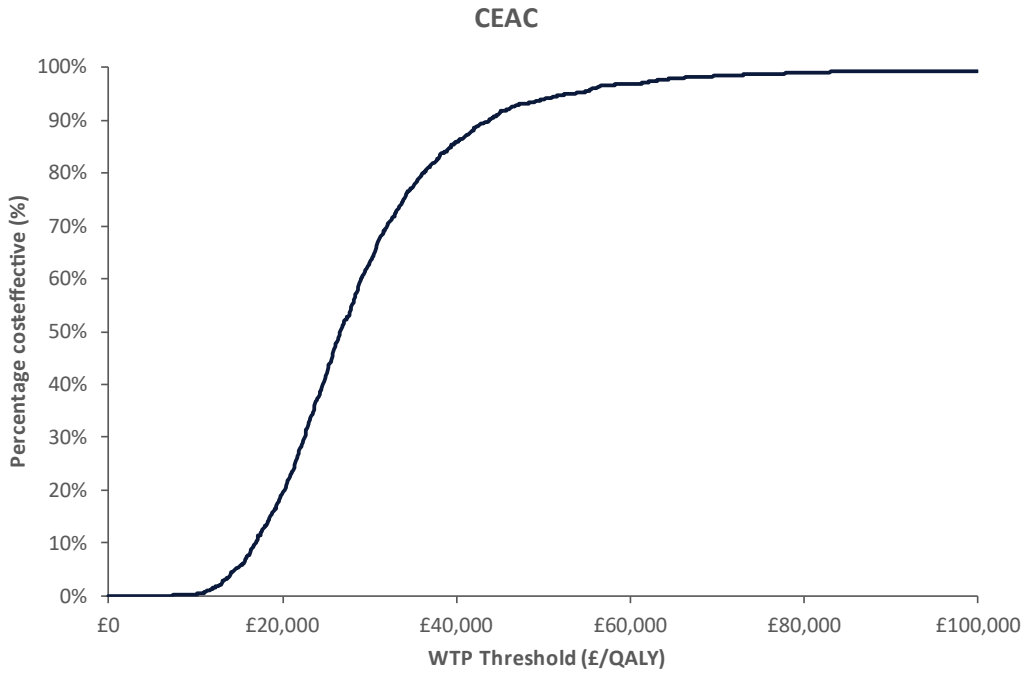


Figure 26. Cost-effectiveness acceptability curve: NIVO-CHEMO versus PEMBRO-CHEMO – updated base case

Additional appendices

Appendix B: updated NMA report

NICE National Institute for Health and Care Excellence

Technical engagement response form

Nivolumab platinum-based chemotherapy for untreated advanced oesophageal cancer with tumour cell PD-L1 expression ≥1% [ID2712]

As a stakeholder you have been invited to comment on the external assessment group (EAG) report for this appraisal.

Your comments and feedback on the key issues below are really valued. The EAG report and stakeholders' responses are used by the appraisal committee to help it make decisions at the appraisal committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

Information on completing this form

We are asking for your views on key issues in the EAG report that are likely to be discussed by the committee. The key issues in the EAG report reflect the areas where there is uncertainty in the evidence, and because of this the cost effectiveness of the treatment is also uncertain. The key issues are summarised in the executive summary at the beginning of the EAG report.

You are not expected to comment on every key issue but instead comment on the issues that are in your area of expertise.

If you would like to comment on issues in the EAG report that have not been identified as key issues, you can do so in the 'Additional issues' section.

If you are the company involved in this appraisal, please complete the 'Summary of changes to the company's cost-effectiveness estimates(s)' section if your response includes changes to your cost-effectiveness evidence.

Please do not embed documents (such as PDFs or tables) because this may lead to the information being mislaid or make the response unreadable. Please type information directly into the form.

Technical engagement response form

NICE National Institute for Health and Care Excellence

Do not include medical information about yourself or another person that could identify you or the other person.

We are committed to meeting the requirements of copyright legislation. If you want to include journal articles in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.

Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.

Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, and all information submitted under 'depersonalised data' in pink. If confidential information is submitted, please also send a second version of your comments with that information replaced with the following text: 'academic/commercial in confidence information removed'. See the Guide to the processes of technology appraisal (sections 3.1.23 to 3.1.29) for more information.

Deadline for comments by the end of **8 August 2022**. Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Thank you for your time.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

Technical engagement response form

Nivolumab platinum-based chemotherapy for untreated advanced oesophageal cancer with tumour cell PD-L1 expression ≥1% [ID2712]

NICE National Institute for Health and Care Excellence

About you

Table 1 About you

Your name	[redacted]
Organisation name: stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder, please leave blank)	Stakeholder: Merck Sharp & Dohme UK Limited
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	**N/A**

NICE National Institute for Health and Care Excellence

Key issues for engagement

All: Please use the table below to respond to the key issues raised in the EAG report.

Table 2 Key issues

Key issue	Does this response contain new evidence, data or analyses?	Response
Key issue 1: Uncertainty as to the appropriate comparators dependent on PD-L1 status **Report section** Executive summary: Table 1.7 Main report: Section 3.4, Section 4.2.6	Yes/No	Please provide your response to this key issue, including any new evidence, data or analyses
Key issue 2: There is limited evidence to support the comparability of the PD-L1 ≥10% CPS populations in the two trials used in the ITC analysis. **Report section** Executive summary: Table 1.8	Yes/No	Please provide your response to this key issue, including any new evidence, data or analyses

NICE National Institute for Health and Care Excellence

Main report: Section 4.2.6.2		
Key issue 3: It is unclear which ITC method, constant HR or time varying HRs formed the base case for the analysis. **Report section** Executive summary: Table 1.9 Main report: Section 4.2.8	Yes/No	Please provide your response to this key issue, including any new evidence, data or analyses
Key issue 4: There is uncertainty as to the nature and effectiveness of subsequent therapy. **Report section** Executive summary: Table 1.10 Main report: Section 4.2.9.3	Yes/No	Please provide your response to this key issue, including any new evidence, data or analyses
Key issue 5: There is uncertainty as to long term OS and the treatment effect of nivolumab + chemotherapy versus chemotherapy. **Report section** Executive summary: Table 1.11 Main report: Section 4.2.9.1, Section 4.2.9.4, Section 5.3	Yes/No	Please provide your response to this key issue, including any new evidence, data or analyses

NICE National Institute for Health and Care Excellence

Key issue 6: There is uncertainty as to how long-term OS for the comparison of nivolumab + chemotherapy versus pembrolizumab + chemotherapy. **Report section** Executive summary: Table 1.7 Main report: Section 3.4, Section 4.2.6	Yes/No	Please provide your response to this key issue, including any new evidence, data or analyses
Key issue 7: There is uncertainty as to how all-cause mortality should be incorporated in the model. **Report section** Executive summary: Table 1.8 Main report: Section 4.2.6.2	Yes/No	Please provide your response to this key issue, including any new evidence, data or analyses
Key issue 8: There is uncertainty as to whether health state utilities should be treatment dependent or incorporate a terminal care decrement. **Report section** Executive summary: Table 1.9	Yes/No	Please provide your response to this key issue, including any new evidence, data or analyses

NICE National Institute for Health and Care Excellence

Main report: Section 4.2.8		
Key issue 9: There is uncertainty as to the appropriate method and value of any adjustment to cost due to delayed or missed doses. **Report section** Executive summary: Table 1.10 Main report: Section 4.2.9.3	Yes/No	Please provide your response to this key issue, including any new evidence, data or analyses
Key issue 10: Calculations were missing from the model, which reduces transparency and makes updating difficult. **Report section** Executive summary: Table 1.11 Main report: Section 4.2.9.1, Section 4.2.9.4, Section 5.3	Yes/No	Please provide your response to this key issue, including any new evidence, data or analyses
Key issue 11: Health state costs were estimated from an out-of-date source. **Report section** Executive summary: Table 1.12 Main report: Section 4.2.9	Yes/No	Please provide your response to this key issue, including any new evidence, data or analyses

NICE National Institute for Health and Care Excellence

Key issue 12: Errors, which underestimated the cost of PEMBRO-CHEMO and prevented the PSA for PEMBRO-CHEMO comparison. **Report section** Executive summary: Table 1.13 Main report: Section 6	Yes/No	Please provide your response to this key issue, including any new evidence, data or analyses

NICE National Institute for Health and Care Excellence

Additional issues

All: Please use the table below to respond to additional issues in the EAG report that have not been identified as key issues. Please do **not** use this table to repeat issues or comments that have been raised at an earlier point in this appraisal (for example, at the clarification stage).

Table 3 Additional issues from the EAG report

Issue from the EAG report	Relevant section(s) and/or page(s)	Does this response contain new evidence, data or analyses?	Response
Additional issue 1: Factual clarification of TA737 details	Section 5.3.3, Table 5.9 , Page 168	No	It is stated that the source of utility data in TA737 was the KEYNOTE-590 EQ-5D 3L. MSD wish to clarify that patients in the trial were administered the 5L version of the questionnaire, and the responses were mapped to the 3L scores using the appropriate algorithm.
Additional issue 2: Factual clarification of TA737 details	Section 5.3.3, Table 5.9 , Page 168	No	When discussing the stopping rule implemented in the TA737 company model, it is stated that 5-FU was administered for up to 25 cycles. MSD wish to clarify this should be corrected to 35 cycles.

NICE National Institute for Health and Care Excellence

Summary of changes to the company's cost-effectiveness estimate(s)

Company only: If you have made changes to the base-case cost-effectiveness estimate(s) in response to technical engagement, please complete the table below to summarise these changes. Please also provide sensitivity analyses around the revised base case. If there are sensitivity analyses around the original base case which remain relevant, please re-run these around the revised base case.

Table 4 Changes to the company's cost-effectiveness estimate

Key issue(s) in the EAG report that the change relates to	Company's base case before technical engagement	Change(s) made in response to technical engagement	Impact on the company's base-case incremental cost-effectiveness ratio (ICER)
Insert key issue number and title as described in the EAG report	Briefly describe the company's original preferred assumption or analysis	Briefly describe the change(s) made in response to the EAG report	Please provide the ICER resulting from the change described (on its own), and the change from the company's original base-case ICER.
Insert key issue number and title as described in the EAG report	…	…	[INSERT / DELETE ROWS AS REQUIRED]
Company's base case following technical engagement (or revised base case)	**Incremental QALYs: [QQQ]**	**Incremental costs: [£££]**	**Please provide company revised base-case ICER**

Sensitivity analyses around revised base case
[PLEASE DESCRIBE HERE]

Technical engagement response form

Nivolumab in combination with fluoropyrimidine- and platinum-based combination chemotherapy for the first-line treatment of adult patients with unresectable advanced, recurrent, or metastatic oesophageal squamous cell carcinoma with tumour cell PD-L1 expression ≥1% [ID2712]

ERG response to the technical engagement response form

Produced by Kleijnen Systematic Reviews (KSR) Ltd, United Kingdom (UK)

Authors
Susan O'Meara, Reviews Manager, KSR Ltd, UK
Nigel Armstrong, Health Economics Manager, KSR Ltd, UK
Charlotte Ahmadu, Health Economist, KSR Ltd, UK
Evangelos Danopoulos, Systematic Reviewer, KSR Ltd, UK
Lisa Stirk, Information Specialist, KSR Ltd, UK
Jos Kleijnen, Director and Owner, KSR Ltd, UK

Correspondence to
Susan O'Meara, Kleijnen Systematic Reviews Ltd
Unit 6, Escrick Business Park
Riccall Road, Escrick
York, United Kingdom
YO19 6FD

Source of funding: This report was commissioned by the National Institute for Health Research (NIHR) Evidence Synthesis Programme as project number NIHR 13/55/62.

Declared competing interests of the authors None

Acknowledgements None

Commercial in confidence (CiC) data are highlighted in blue throughout the report.
Academic in confidence (AiC) data are highlighted in yellow throughout the report.
Confidential comparator prices are highlighted in green throughout the report.
Any de-personalised data are highlighted in pink throughout the report.

.

Abbreviations

1L	First line
AE	Adverse events
AIC	Akaike Information Criterion
AiC	Academic in confidence
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
BIC	Bayesian information criterion
BICR	Blinded independent central review
BNF	British National Formulary
BSA	Body surface area
BSC	Best supportive care
CADTH	Canadian Agency for Drugs and Technologies in Health
CEAC	Cost effectiveness acceptability curve
CENTRAL	Cochrane Central Register of Controlled Trials
CHEMO	Chemotherapy
CHMP	Committee for Medicinal Products for Human Use
CI	Confidence interval
CIS	Cisplatin
CiC	Commercial in confidence
CL	Clarification letter
CMH	Cochran-Mantel- Haenszel
CPS	Combined positive score
CRD	Centre for Reviews and Dissemination
CrI	Credible interval
CS	Company submission
CSR	Clinical study report
CT	Computerised tomography
CTCAE	Common Terminology Criteria for Adverse Events
DBL	Database lock
DFS	Disease-free survival
DP	Decision problem
DoR	Duration of response
DSA	Deterministic sensitivity analysis
DSU	Decision Support Unit
ECOG PS	Eastern Cooperative Oncology Group Performance Scale
EMA	European Marketing Authorisation
eMIT	Electronic market information tool
EQ-5D-3L	European Quality of Life-5 dimensions-3 levels
ERG	Evidence Review Group
ESCC	Esophageal squamous cell carcinoma
ESMO	European Society for Medical Oncology
FACT-E	Functional Assessment of Cancer Therapy- Esophagus
FACT-G	Functional Assessment of Cancer Therapy- General
FACT-G7	Functional Assessment of Cancer Therapy- General 7-items
FAD	Final Appraisal Document
FE	Fixing errors
FOLFOX	Fluorouracil and oxaliplatin
FV	Fixing violations
GC	Gastric cancer
GI	Gastrointestinal
HR	Hazard ratio
HRQoL	Health-related quality of life
HTA	Health Technology Assessment

IA	Interim analysis
ICER	Incremental cost effectiveness ratio
IMAEs	Immune-mediated adverse events
IQR	Interquartile range
ITC	Indirect treatment comparison
ITT	Intention to treat
IV	Intravenous
KSR	Kleijnen Systematic Reviews Ltd
K-M	Kaplan-Meier
LYs	Life years
MedDRA	Medical Dictionary for Regulatory Activities
MeSH	Medical Subject Headings
MID	Minimum important difference
MJ	Matters of judgement
N/A	Not applicable
NCI	National Cancer Institute
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
NIVO	Nivolumab
NMA	Network meta-analysis
N/R	Not reported
OC	Oesophageal cancer
OESI	Other events of special interest
OGJ	Oesophagogastric junction
ORR	Objective response rate
OS	Overall survival
OSCC	Oesophageal squamous cell carcinoma
PAS	Patient Access Scheme
PEMBRO	Pembrolizumab
PD-L1	Programmed death ligand 1
PFS	Progression-free survival
PFS2	Time to second progression
PRESS	Peer Review of Electronic Search Strategies
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-analyses
PSA	Probabilistic sensitivity analysis
PSM	Partitioned survival model
PSS	Personal Social Services
PT	Preferred terms
QALY	Quality-adjusted life year
QoL	Quality of life
Q2W	Every two weeks
Q4W	Every four weeks
Q6W	Every six weeks
RCT	Randomised controlled trial
RDI	Relative dose intensity
RECIST	Response Evaluation Criteria in Solid Tumours
SAE	Serious Adverse Events
SCC	Squamous cell carcinoma
ScHARRHUD	University of Sheffield School of Health and Related Research Health Utilities Database
SD	Standard deviation
SE	Standard error
SLR	Systematic literature review
SmPC	Summary of Product Characteristics

SOC	System organ class
SoC	Standard of care
STA	Single Technology Appraisal
TA	Technology Assessment
TC	Tumour cell
TE	Technical Engagement
TECH-VER	Technical Verification
ToT	Time on treatment
TP	Transition probability
TPS	Tumour proportion score
TSST	Time to second subsequent therapy
TTD	Time-to-death
UK	United Kingdom
USA	United States of America
VAS	Visual analogue scale
VBA	Visual Basic for Applications
WTP	Willingness-to-pay
XELOX	Capecitabine + oxaliplatin

Introduction

This document is the Evidence Review Group's (ERG's) critique of comments and additional data provided by the company as part of the technical engagement (TE) process for nivolumab in combination with fluoropyrimidine- and platinum-based combination chemotherapy for the first line (1L) treatment of adults with unresectable advanced, recurrent or metastatic, previously untreated oesophageal squamous cell carcinoma (OSCC) with tumour cell programmed death ligand 1 (PD-L1) expression of at least one percent.[1] A confidential appendix with updated ERG base-case results has been provided as part of the ERG's critique of the company's TE response.

Key issues

Key issue 1: Uncertainty as to the appropriate comparators dependent on PD-L1 status

Given that chemotherapy would still be standard of care (SoC) regardless of PD-L1 status (Section 2.3, ERG report[2]), and thus only chemotherapy would be an appropriate comparator in the PD-L1 ≥1% tumour cell (TC) and <10 combined positive score (CPS) subgroup population, the ERG reiterated its request that an analysis of nivolumab combined with chemotherapy (NIVO-CHEMO) versus chemotherapy (CHEMO) in the CheckMate 648 trial be performed in this population. The company in their TE response provided new analyses in form of a separate analysis of the CheckMate 648 trial and cost effectiveness analysis for the PD-L1 ≥1% TC and <10 CPS population comparing NIVO-CHEMO versus CHEMO. Figures 1 and 2 compare CHEMO overall survival (OS) parametric and semi-parametric curves; Figures 3 and 4 compare CHEMO progression-free survival (PFS) parametric and semi-parametric curves; and Figure 5 presents the CHEMO time on treatment (ToT) Kaplan-Meier curve in the CheckMate 648 PD-L1 ≥1% TC and <10 CPS subgroup. Figures 6 and 7 compare NIVO-CHEMO OS parametric and semi-parametric curves; Figures 8 and 9 compare NIVO-CHEMO PFS parametric and semi-parametric curves; and Figure 10 presents the NIVO-CHEMO ToT Kaplan-Meier curve in the CheckMate 648 PD-L1 ≥1% TC and <10 CPS subgroup. For OS, Gompertz and Weibull semi-parametric curves were chosen for CHEMO and NIVO-CHEMO respectively. For PFS, Weibull and lognormal semi-parametric functions were chosen for CHEMO and NIVO-CHEMO respectively. Kaplan-Meier curves were used to estimate ToT. Table 1 presents the cost-effectiveness analysis results for this subgroup which is below a willingness-to-pay threshold of £50,000 per quality-adjusted life year (QALY).

In their preamble to discussion of the Key Issues, the company provided an outline of PD-L1 expression testing methods, namely the TC and CPS approaches. The company explained that they had contacted expert pathologists ahead of TE to explore likely clinical practice around the two testing methods. The company stated that the expert pathologists "*highlighted that these are essentially both independent tests for assessing PD-L1 expression. While there is overlap in these tests, this overlap is not complete. Additionally, there should be no perceived linear relationship between expression levels: patients with PD-L1 TC ≥ 1% should not be seen as lower PD-L1 expression than PD-L1 CPS ≥ 10.*"[1] The company went on to make the following statements:

"*As the nivolumab licence specifies usage in patients with PD-L1 TC ≥ 1%, the PD-L1 TC test is a mandatory requirement if nivolumab were to be prescribed, while CPS score would be irrelevant and would not be calculated as it does not inform whether nivolumab should be administered.*

Pathologists noted that clinicians should typically request the treatment of choice and this would then guide the relevant test, so that testing method would not be a barrier to use of nivolumab in clinical practice. Further, pathologists would prefer TC testing as this is less time consuming to undertake.

Pathologists would not want to do both tests, as this will require staining and counting of two sets of slides, taking up time and resource in a busy department."[1]

The ERG noted that no references were provided for the statements made and no documentation of the contact with expert pathologists was included with the company's TE response dossier.

ERG comment: The ERG acknowledges the arguments proposed by the company.

The ERG notes that the company's assertion (as part of the TE response) that pathologists would only want to undertake a single test is at odds with the company's response to clarification question A7 which suggests that the TC and CPS tests could be performed simultaneously: *"It is likely that it will become routine practice to assess both tumour proportion score (TPS) and CPS during the same test to determine which OSCC patients are suitable for either pembrolizumab or nivolumab treatment."* (p. 15).[3] The comments in the company's TE response indicate a preference for the TC test and this seems to imply that that identification of those eligible to be treated with pembrolizumab would be overlooked. Furthermore, the company's comment that "*treatment of choice….would then guide the relevant test*"(p. 5)[1] is contrary to the usual ideas about patient management pathways, whereby a diagnostic or screening test would be expected to guide considerations of treatment and not the other way round.[4] From the overall information, the ERG remains uncertain about the expectation in clinical practice (i.e., would choice of treatment precede testing or vice versa), particularly in light of the lack of referencing and absence of transparency in relation to the discussions with expert pathologists.

Survival curve selection- Although the company states that, *"no fully parametric models provided a good fit to the available data, particularly in the first six months and after 30 months,"* (p. 31)[1] it could be argued that the Weibull fully parametric function for CHEMO OS, Log-logistic for CHEMO PFS, Weibull for NIVO-CHEMO OS and Lognormal for NIVO-CHEMO PFS are exceptions and as no well-founded justification to the choice of the 6.9 month cut-off previously highlighted in the ERG report,[2] was provided, the ERG preference remains with the parametric approach. As the pre-load option to run analyses for this subgroup was not made available in the updated ERG base-case model, the ERG could not provide its analyses for this subgroup.

Table 1 Updated company base-case cost-effectiveness results: PD-L1 ≥1% TC and <10 CPS subgroup analysis

		Total Costs (£)	Total QALYs	ICER (£/QALY)	Updated Base case ICER (£/QALY)
Impact on company base case analysis	Nivolumab, cisplatin and fluorouracil	[redacted]	[redacted]	**£29,717**	**£27,106**
	Cisplatin and fluorouracil	[redacted]	[redacted]		
	Incremental	[redacted]	[redacted]		

Based on Table 10 of TE response[1]
Abbreviations: ICER = Incremental cost effectiveness ratio; QALY = quality-adjusted life year

Figure 1. CheckMate 648 PD-L1 TC ≥ 1% and CPS < 10: CHEMO overall survival (parametric extrapolation)

Based on Figure 4 of TE response[1]

Abbreviations: AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; K-M = Kaplan-Meier; OS = overall survival; PD-L1 = programmed death ligand 1; TC = tumour cell

Figure 2. CheckMate 648 PD-L1 TC ≥ 1% and CPS < 10: CHEMO overall survival (semi-parametric extrapolation – cutpoint at 6.9 months)

Based on Figure 5 of TE response[1]

Abbreviations: AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; K-M = Kaplan-Meier; OS = overall survival; PD-L1 = programmed death ligand 1; TC = tumour cell

Figure 3. CheckMate 648 PD-L1 TC ≥ 1% and CPS < 10: CHEMO progression-free survival (parametric extrapolation)

Based on Figure 6 of TE response[1]

Abbreviations: AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; K-M = Kaplan-Meier; OS = overall survival; PD-L1 = programmed death ligand 1; TC = tumour cell

Figure 4. CheckMate 648 PD-L1 TC ≥ 1% and CPS < 10: CHEMO progression-free survival (semi-parametric extrapolation – cutpoint at 6.9 months)

Based on Figure 7 of TE response[1]

Abbreviations: AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; K-M = Kaplan-Meier; OS = overall survival; PD-L1 = programmed death ligand 1; TC = tumour cell

Figure 5. CheckMate 648 PD-L1 TC≥1% and CPS<10: CHEMO time on treatment (Kaplan-Meier)

Based on Figure 8 of TE response[1]

Abbreviations: AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; K-M = Kaplan-Meier; OS = overall survival; PD-L1 = programmed death ligand 1; TC = tumour cell

Figure 6. CheckMate 648 PD-L1 TC≥1% and CPS<10: NIVO-CHEMO overall survival (parametric extrapolation)

Based on Figure 9 of TE response[1]

Abbreviations: AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; K-M = Kaplan-Meier; OS = overall survival; PD-L1 = programmed death ligand 1; TC = tumour cell

Figure 7. CheckMate 648 PD-L1 TC≥1% and CPS<10: NIVO-CHEMO overall survival (semi-parametric extrapolation – cutpoint at 6.9 months)

Based on Figure 10 of TE response[1]

Abbreviations: AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; K-M = Kaplan-Meier; OS = overall survival; PD-L1 = programmed death ligand 1; TC = tumour cell

Figure 8. CheckMate 648 PD-L1 TC≥1% and CPS<10: NIVO-CHEMO progression-free survival (parametric extrapolation)

Based on Figure 11 of TE response[1]

Abbreviations: AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; K-M = Kaplan-Meier; OS = overall survival; PD-L1 = programmed death ligand 1; TC = tumour cell

Figure 9. CheckMate 648 PD-L1 TC≥1% and CPS<10: NIVO-CHEMO progression-free survival (semi-parametric extrapolation – cut-off point at 6.9 months)

Based on Figure 12 of TE response[1]

Abbreviations: AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; K-M = Kaplan-Meier; OS = overall survival; PD-L1 = programmed death ligand 1; TC = tumour cell

Figure 10. CheckMate 648 PD-L1 TC≥1% and CPS<10: NIVO-CHEMO time on treatment Kaplan-Meier

Based on Figure 13 of TE response[1]

Abbreviations: AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; K-M = Kaplan-Meier; OS = overall survival; PD-L1 = programmed death ligand 1; TC = tumour cell

Key issue 2: There is limited evidence to support the comparability of the PD-L1 ≥10% CPS populations in the two trials used in the ITC analysis.

The company updated the indirect treatment comparison (ITC)/network meta-analysis (NMA) report to include the baseline characteristics of the PD-L1 CPS ≥ 10% population from CheckMate 648.[5, 6] These additional data are presented in Table 2, along with additional data for KEYNOTE-590 regarding both arms (pembrolizumab plus chemotherapy and chemotherapy). They state that the baseline characteristics of the subgroups are broadly comparable to the CheckMate 648 overall population, and hence do not affect the conclusions or introduce uncertainty in the indirect comparison.

ERG comment: The ERG agrees that the baseline characteristics of the PD-L1 ≥10% (CPS) subgroups are comparable to the overall population of CheckMate 648. Regarding KEYNOTE-590, the conclusions that we can draw from this additional data are limited since only three characteristics were reported (proportion of Asian patients, Eastern Co-operative Oncology Group [ECOG] status and metastases) and these were only available for both arms of the trial combined.

Table 2: Summary of baseline patient characteristics for CheckMate 648 and KEYNOTE 590

Study	CheckMate 648		CheckMate 648 (PD-L1 ≥10% CPS population)		KEYNOTE-590		KEYNOTE-590 (ESCC and PD-L1 ≥10% CPS population)
Treatment Arm	Nivolumab + chemotherapy	Chemotherapy	Nivolumab + chemotherapy	Chemotherapy	Pembrolizumab + chemotherapy	Chemotherapy	Both treatment arms
Sample Size	321	324	135	145	373	376	286
Age (years), median	64	64	65	64	64	62	N/R
Asian	Total: 70% East Asia: 57% Rest of Asia: 13%	Total: 70% East Asia: 57% Rest of Asia: 13%	Total: 72% East Asia: 59% Rest of Asia: 13%	Total: 75% East Asia: 60% Rest of Asia: 15%	53%	52%	68%
ECOG PS (0)	46%	47%	50%	47%	40%	40%	40%
ECOG PS (1)	54%	53%	50%	52%	60%	60%	60%
Metastatic disease status	57%	58%	N/R	N/R	92%	90%	N/R
Organs with Metastases	≤1: 49% ≥2: 51%	≤1: 49% ≥2: 51%	≤1 sites: 47% ≥2 sites: 53%	≤1 sites: 52% ≥2 sites: 48%	N/R	N/R	92%
Liver Metastases	26%	26%	28%	23%	N/R	N/R	N/R

Based on Table 9 of Document B of the CS[7], Table 5 of the updated Appendix L of the CS[8], Table 6 of NICE STA ID 3714[9] and Table 5 of the updated NMA report version 3[6]
CPS = combined positive score; CS = company submission; ECOG PS = Eastern Cooperative Oncology Group Performance Status; N/R = not reported; PD-L1 = programmed death ligand-1

Key issue 3: It is unclear which ITC method, constant HR or time varying HRs formed the base case for the analysis.

The company clarified that the time varying method was used. In addition, they also highlighted that "*"statistical significance" in the ITC report refers to the credibility of confidence intervals. No specific statistical tests were run."*(p.11)[1]

ERG comment: The ERG had also requested that the company would discuss the underlying conceptual assumptions that were connected to the ITC method of choice. The company did not provide this discussion.

Key issue 4: There is uncertainty as to the nature and effectiveness of subsequent therapy.

The company provided results to illustrate the impact of modifying subsequent therapy outcomes, demonstrated in Table 3.

Table 3 Impact of modification of subsequent therapy outcomes

			Incremental costs (£)	**Incremental QALYs**	**ICER (£/QALY)**
NIVO-CHEMO versus CHEMO	Updated Company base case	Updated base case	[redacted]	[redacted]	£27,106
		RPSFTM-scenario	[redacted]	[redacted]	£29,253
		Model-based ATTRACTION-3 scenario	[redacted]	[redacted]	£39,868
Adapted from Table 3 of TE response[1] Abbreviations: CHEMO = chemotherapy; ICER = incremental cost-effectiveness ratio; NIVO = nivolumab; QALY = quality-adjusted life year; RPSFTM = Rank-Preserving Structural Failure Time Model.					

ERG comment: The ERG concurs that the method utilised by the company method has some validity. However, NICE TSD 16[10] suggests other valid methods that the company could have used for treatment switching such as the two-stage method. It is noteworthy that the company's incremental cost-effectiveness ratio (ICER) is underestimated without the addition of treatment switching. As employing treatment switching causes the ICER to go up significantly, it is uncertain how important the addition of treatment switching could be.

Key issue 5: There is uncertainty as to long term OS and the treatment effect of nivolumab + chemotherapy versus chemotherapy.

The company presented six scenario analyses exploring the impact of treatment waning on cost-effectiveness results as presented in Table 4. They stated that, *"Cost-effectiveness increases with later implementation and with longer duration of waning. A case where waning is delayed until 4 years (end of trial period) and closes the hazard ratio at 10 years, to approximate the gradient of the hazard ratio at the point of waning start, is considered to be an improvement on the modelling assumptions of the ERG scenario."* (p. 17)[1].

Table 4 NIVO-CHEMO vs. CHEMO cost-effectiveness results: treatment waning scenarios

			Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
NIVO-CHEMO versus CHEMO	Updated Company base case	Updated base case	[redacted]	[redacted]	£27,106
	Updated ERG base case	Updated, impact of long-term treatment effect between NIVO-CHEMO and CHEMO	[redacted]	[redacted]	£36,423
		Updated, waning 2.5 – 4.0 years	[redacted]	[redacted]	£42,016
		Updated, waning 4.0 – 5.5 years	[redacted]	[redacted]	£39,364
		Updated, waning 6.0 – 7.5 years	[redacted]	[redacted]	£38,041
		Updated, waning 4.0 – 10.0 years	[redacted]	[redacted]	£38,113

Based on Table 14 of TE response[1]
Abbreviations: CHEMO = chemotherapy; ERG = Evidence Review Group; NIVO = nivolumab; ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year

ERG comment: In the ERG base-case, treatment waning was applied from 2.5 to 4 years, as per the company's scenario analyses. The ERG remains unconvinced that a treatment waning effect of 4 to 10 years is an improvement to the ERG's base case, being that there is sufficient evidence in the landmark analysis that treatment waning appears earlier on.

The ERG re-ran the company's explored scenarios for treatment waning and were unable to reproduce the results displayed in Table 4. Table 5 presents the results derived by the ERG.

Scenarios explored in Table 5:

- Company updated base case
- ERG base case updated with new nivolumab PAS, updated costs, time on treatment weighted RDI – no treatment waning
- ERG updated base case – waning between 2.5 and 4.0 years
- ERG updated base case – waning between 4.0 and 5.5 years (end of follow-up)
- ERG updated base case – waning between 6.0 and 7.5 years ("7-year wane" scenario)
- ERG updated base case – waning between 4.0 and 10.0 years (end of follow-up, matched hazard ratio gradient)

Table 5 ERG updated NIVO-CHEMO vs. CHEMO cost-effectiveness results: treatment waning scenarios

			Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
NIVO-CHEMO versus CHEMO	Updated Company base case	Updated base case	[redacted]	[redacted]	£27,106
	Updated ERG base case (updated costs, new PAS, wtd RDI update)	No treatment waning	[redacted]	[redacted]	£36,225
		Waning 2.5 – 4.0 years	[redacted]	[redacted]	£41,764
		Waning 4.0 – 5.5 years	[redacted]	[redacted]	£39,138
		Waning 6.0 – 7.5 years	[redacted]	[redacted]	£37,826
		Waning 4.0 – 10.0 years	[redacted]	[redacted]	£37,898

Abbreviations: CHEMO = chemotherapy; ERG = Evidence Review Group; NIVO = nivolumab; ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year; wtd = treatment weighted

Key issue 6: There is uncertainty as to long-term OS for the comparison of nivolumab + chemotherapy versus pembrolizumab + chemotherapy.

The company have stated that *"there were stochastic differences between the estimated hazard ratios in the analysis originally used to perform the adjustment, and those given within the NMA report,"* (p. 18)[1] and for consistency have updated cost-effectiveness results using PEMBRO-CHEMO hazard ratios from the "overlap" analysis described in the NMA report. This is presented in Table 6.

Table 6 NIVO-CHEMO vs. PEMBRO-CHEMO cost-effectiveness results: updated PEMBRO-CHEMO hazard ratios

			Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
NIVO-CHEMO versus PEMBRO-CHEMO	Company base case post CQs	Original	[redacted]	[redacted]	£320,254 (CE region)
		Updated hazard ratios	[redacted]	[redacted]	£1,153,385 (CE region)

Adapted from Table 15 of TE response[1]
Abbreviations: CE = cost effective, CQs = clarification questions; ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year

ERG comment: Insufficient information was provided on what changes have been made for the ERG to make any further comments.

Key issue 7: There is uncertainty as to how all-cause mortality should be incorporated in the model.

ERG comment: The ERG removed additional all-cause mortality in a deterministic scenario analysis. The company's commentary seems to acknowledge that there is double counting, however minimal.

Although the ERG base-case remains the same i.e., with all-cause mortality added, the ERG remains unconvinced of the necessity of all-cause mortality being applied in an economic model.

Key issue 8: There is uncertainty as to whether health state utilities should be treatment dependent or incorporate a terminal care decrement.

The company stated that *"an updated utility analysis was run, but the impact on the finding of the previous analysis was minimal."* (p. 21)[1]

ERG comment: As a matter of judgement, the ERG employed treatment-dependent health state utility values in its ERG base case.[2] As a regression analysis with all three clinically relevant covariates i.e., health state, treatment, and time to death (TTD), as advised in the ERG report, was not submitted as part of the technical engagement response, the ERG is unable to comment on the plausibility of the company's statement.

Key issue 9: There is uncertainty as to the appropriate method and value of any adjustment to cost due to delayed or missed doses.

The ERG presented analyses using summaries of patient-level relative dose intensity (RDI) from the CheckMate 648 trial.[2] The company identified two compromises with this method and proposed to update the ERG's analyses by solving one of the compromises, or improve upon the ERG's preferred method, by modifying their analysis and updating their base case. Table 7 presents RDI values and Table 8, the impact of different RDI calculations on cost-effectiveness results.

- ERG scenario update: Mean of RDI using all patients with at least 1 dose of named therapy component in CheckMate 648 (PD-L1 ≥ 1%), as of the [redacted]_DBL. This is a direct improvement on ERG data, where the same assumptions apply.
- ERG scenario improvement: Weighted mean of RDI where weight is by patient time on treatment for all patients with at least 1 dose of named therapy component in CheckMate 648 (PD-L1 ≥ 1%) as of the [redacted]_DBL. This is an approximation to the grand mean RDI, i.e., all doses dispensed / all doses expected, based upon total time on treatment. This modifies the company base case.

ERG comment: As a matter of judgement concerning the cost of doses of medication received, the ERG calculated RDI to estimate the mean cost of treatment per unit time on treatment. The ERG acknowledges the compromises highlighted by the company but struggles to comprehend what changes have been made and the justification behind them i.e., the need to reweight by time-on-treatment. It remains unclear why the company have made the changes that they have. The ERG's stance is to continue with the ERG base-case (see confidential appendix for impact of RDI on base-case results).

Table 7 Original RDI, mean RDI of PLD, and mean weighted RDI of PLD

Treatment		Delayed dose modifier	Original RDI	RDI – mean of patient-level data	RDI – mean of patient-level data weighted by time on treatment
Nivolumab in combination with chemotherapy (fluorouracil plus cisplatin)	Nivolumab	[redacted]	[redacted]	[redacted]	[redacted]
	Fluorouracil	[redacted]	[redacted]	[redacted]	[redacted]
	Cisplatin	[redacted]	[redacted]	[redacted]	[redacted]
Chemotherapy (fluorouracil plus cisplatin)	Fluorouracil	[redacted]	[redacted]	[redacted]	[redacted]
	Cisplatin	[redacted]	[redacted]	[redacted]	[redacted]
Pembrolizumab in combination with chemotherapy (fluorouracil plus cisplatin)	Pembrolizumab	[redacted]	[redacted]	[redacted]	[redacted]
	Fluorouracil	[redacted]	[redacted]	[redacted]	[redacted]
	Cisplatin	[redacted]	[redacted]	[redacted]	[redacted]

Based on Table 16 of TE response[1]
Abbreviations: RDI = relative dose intensity
* The ERG scenario versus PEMBRO-CHEMO used the original delayed dose modifier for both treatment arms.

Table 8 Cost-effectiveness results: impact of updated RDI calculation

			Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
NIVO-CHEMO versus CHEMO	Company base case post CQs	Original – delayed dose modifier	[redacted]	[redacted]	£31,826
		Updated RDI – Time on treatment weighted	[redacted]	[redacted]	£30,183
	ERG base case	Original RDI	[redacted]	[redacted]	£46,599
		Updated RDI – unweighted	[redacted]	[redacted]	£42,959
		Updated RDI – Time on treatment weighted	[redacted]	[redacted]	£41,550
NIVO-CHEMO versus PEMBRO-CHEMO	Company base case post CQs	Original – delayed dose modifier	[redacted]	[redacted]	£320,254 (CE region)
		Updated RDI – Time on treatment weighted	[redacted]	[redacted]	£298,853 (CE region)

			Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
	ERG base case	Original – delayed dose modifier	[redacted]	[redacted]	£109,039 (CE region)
		Updated RDI - unweighted	[redacted]	[redacted]	£105,386 (CE region)
		Updated RDI – Time on treatment weighted	[redacted]	[redacted]	£102,286 (CE region)
Based on Table 17 of TE response[1] **Abbreviations**: RDI = relative dose intensity **Note**: All results with updated PAS for nivolumab					

Key issue 10: Calculations were missing from the model, which reduces transparency and makes updating difficult.

The company stated that, *"external calculations enabled the model to be more flexible to support analysis requests made at clarification and technical engagement,"* (p. 22)[1] and included the cost calculations within the model provided. These calculations are not linked to model inputs.

ERG comment: An input calculation sheet was provided in the company's updated economic model.

Key issue 11: Health state costs were estimated from an out-of-date source.

The company provided updates for drug acquisition costs (Table 9), additional costs (Table 10) and cost-effectiveness results (Table 11).

Table 9 Updated drug acquisition costs

Resource	Description	Currency code	Original input	Updated input
Cisplatin	100 mg/100 ml solution for infusion vials	DHA010	eMIT 2020 - £8.73	eMIT 2021 - £8.97
	50 mg/50 ml solution for infusion vials	DHA011	eMIT 2020 - £5.38	eMIT 2021 - £6.03
Fluorouracil	1 g/20 ml (5%) solution for infusion vial	DHA265	eMIT 2020 - £2.35	eMIT 2021 - £2.25
	2.5 g/100 ml (2.5%) solution for infusion vial	DHA024	eMIT 2020 - £3.79	eMIT 2021 - £4.32
	2.5 g/50 ml (5%) solution for infusion vial	DHA102	eMIT 2020 - £4.01	eMIT 2021 - £4.21
	500 mg/10 ml (5%) solution for infusion vial	DHA240	eMIT 2020 - £1.77	eMIT 2021 - £2.86

Resource	Description	Currency code	Original input	Updated input
	5 g/100 ml (5%) solution for infusion vials	DHA137	eMIT 2020 - £8.58	eMIT 2021- £9.2
Capecitabine	150 mg tablets pack size 60	DHA224	eMIT 2020 - £4.43	eMIT 2021 - £4.43
	300 mg tablets pack size 60	DKE068	eMIT 2020 - £7.77	eMIT 2021 - £7.77
	500 mg tablets pack size 120	DHA225	eMIT 2020 - £26.30	eMIT 2021 - 26.30
Docetaxel	160mg/8ml solution for infusion (20mg/ml)	DHC046	eMIT 2020 - £17.95	eMIT 2021 - £17.38
	Docetaxel 20mg/1ml solution for infusion vials (20mg/ml)	DHC025	eMIT 2020 -£3.77	eMIT 2021 - £3.56
	80mg/4ml solution for infusion vials (20mg/ml)	DHC029	eMIT 2020 - £9.13	eMIT 2021 - £8.90
Paclitaxel	100mg/16.7ml solution for infusion vials	DHA145	eMIT 2020 - £7.22	eMIT 2021 - £8.06
	150mg/25ml solution for infusion vials	DHA297	eMIT 2020 - £12.41	eMIT 2021 - £10.15
	300mg/50ml solution for injection vials	DHA210	eMIT 2020 - £17.66	eMIT 2021 -£15.97
	30mg/5ml solution for infusion vials	DHA144	eMIT 2020 - £4.41	eMIT 2021 - £4.15
Adapted from Table 18 of TE response[1] **Abbreviations**: eMIT = Electronic market information tool.				

Table 10 Additional cost updates

Resource	Description	Currency code	Company source and unit cost	ERG preferred source and unit cost
Administration costs				
Deliver simple parenteral chemotherapy at first attendance	Day case and reg day/night	SB12Z	NHS reference cost 2015/16 (weighted average)- £284.05	NHS reference cost 2019/20- £299.61
Deliver complex chemotherapy, including prolonged infusion treatment, at first attendance	Day case and reg day/night	SB14Z	NHS reference cost 2015/16- £431.72	NHS reference cost 2019/20- £431.72
Disease monitoring costs				
CT scan	CT scan of three areas, without contrast	RD25Z	NHS 2019/20 - £103.31	NHS 2019/20 - £103.34

Resource	Description	Currency code	Company source and unit cost	ERG preferred source and unit cost
Blood test	Haematology	DAPS05	NHS 2019/20 - £2.53	NHS 2019/20 - £2.53
Kidney	Special screening, examinations or other genetic disorders	WH15Z	NHS 2019/20 - £33.80	NHS 2019/20 - £33.80
Hepatic	Special screening, examinations or other genetic disorders	WH15Z	NHS 2019/20 - £33.80	NHS 2019/20 - £33.80
Consultant	Non-admitted face-to-face attendance, follow-up	WF01A: service code 370	NHS 2019/20 - £203.14	NHS 2019/20 - £200

Adapted from Table 19 of TE response[1]
Abbreviations: CT = computed tomography; ERG = Evidence Review Group; NHS = National Health Service.

Table 11 Cost-effectiveness results: impact of updated modelled costs

			Incremental costs (£)	**Incremental QALYs**	**ICER (£/QALY)**
NIVO-CHEMO versus CHEMO	Company base case post CQs	Original	[redacted]	[redacted]	£33,357
		Updated costs	[redacted]	[redacted]	£32,123
	ERG base case	Original*	[redacted]	[redacted]	£46,599
		Updated costs	[redacted]	[redacted]	£47,093
NIVO-CHEMO versus PEMBRO-CHEMO	Company base case post CQs	Original	[redacted]	[redacted]	£320,254 (CE region)
		Updated costs	[redacted]	[redacted]	£327,407 (CE region)
	ERG base case	Original	[redacted]	[redacted]	£109,039 (CE region)
		Updated costs	[redacted]	[redacted]	£111,300 (CE region)

Adapted from Table 20 of TE response[1]
Abbreviations: CQ = clarification questions; ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year
*All analyses use updated PAS

ERG comment: The company presented a table with updated drug acquisition and additional costs. These costs mostly mirrored the ERG's preferred source and unit cost (as presented in Section 4.2.9 of the ERG report[2]) with the exception of administration costs which the company appears to favour an outdated NHS reference 2015/2016 source without justification. However, in cell C83 (input calculation sheet), the company appears to have made the correction to use the ERG preferred cost.

Key issue 12: Errors, which underestimated the cost of PEMBRO-CHEMO and prevented the PSA for PEMBRO-CHEMO comparison.

The ERG identified three errors in the comparison with PEMBRO-CHEMO- incorrect application of a 50% discount to the price of pembrolizumab, inappropriate PSA distributions for pembrolizumab and fluorouracil unit costs, and the inclusion of the cost of pembrolizumab only in the first cycle. These errors could not be fixed by the ERG, and thus the PSA for NIVO-CHEMO vs. PEMBRO-CHEMO could not be run. These issues were addressed and the company provided an updated cost-effectiveness analysis as presented in Table 12.

ERG comment: These issues raised by the ERG do not appear to have been fixed in the ERG base case model as the cost incurred the pembrolizumab arm in the PSA (using the updated costs preload option) was still severely underestimated, with NIVO-CHEMO being dominated. This was not so when the PSA for NIVO-CHEMO vs. PEMBRO-CHEMO was run in the company's base-case model.

Table 12 Cost-effectiveness results: updated PEMBRO-CHEMO cost

		Total Costs (£)	Total QALYs	ICER (£/QALY)	Base case ICER* (£/QALY)
Impact on company base case analysis	Nivolumab, cisplatin and fluorouracil	[redacted]	[redacted]	£307,447 (CE region)	-£5,594 (not CE)
	Pembrolizumab, cisplatin and fluorouracil	[redacted]	[redacted]		
	Incremental	[redacted]	[redacted]		

Based on Table 21 of TE response[1]
Abbreviations: CE = cost effective; ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year
*Before Nivolumab PAS update and update of pembrolizumab costs

References

[1] National Institute for Health and Care Excellence. *Nivolumab platinum-based chemotherapy for untreated advanced oesophageal cancer with tumour cell PD-L1 expression ≥1% [ID2712]: Technical engagement response form*. London: NICR, 2022. 76p.

[2] O'Meara S, Armstrong N, Ahmadu C, Danopoulos E, Stirk L, Kleijnen J. *Nivolumab in combination with fluoropyrimidine- and platinum-based combination chemotherapy for the first-line treatment of adult patients with unresectable advanced, recurrent, or metastatic oesophageal squamous cell carcinoma with tumour cell PD-L1 expression ≥1% [ID2712]: a Single Technology Assessment*. York: Kleijnen Systematic Reviews Ltd., 2022. 187p.

[3] National Institute for Health and Care Excellence. *Nivolumab in combination with fluoropyrimidine- and platinum-based combination chemotherapy for the first-line treatment of adult patients with unresectable advanced, recurrent, or metastatic oesophageal squamous cell carcinoma with tumour cell PD-L1 expression ≥1% [ID2712]: Response to request for clarification from the ERG*. London: NICE, 2022. 159p.

[4] Centre for Reviews and Dissemination. *Systematic Reviews: CRD's guidance for undertaking reviews in health care [Internet]*. York: University of York, 2009 [accessed 23.3.22] Available from: http://www.york.ac.uk/inst/crd/SysRev/!SSL!/WebHelp/SysRev3.htm

[5] Gricar J, McDonald L. *Nivolumab-containing regimens for the treatment of first-line advanced or metastatic esophageal cancer. Network meta-analysis report October 2021 database lock. Version 1*: PRECISIONheor, 2022. 51p.

[6] Gricar J, McDonald L. *Nivolumab-containing regimens for the treatment of first-line advanced or metastatic esophageal cancer. Network meta-analysis report October 2021 database lock. Version 3*: PRECISIONheor, 2022. 313p.

[7] Bristol-Myers Squibb Pharmaceuticals Ltd. *Nivolumab in combination with fluoropyrimidine- and platinum-based combination chemotherapy for the first-line treatment of adult patients with unresectable advanced, recurrent, or metastatic oesophageal squamous cell carcinoma with tumour cell PD-L1 expression ≥1% [ID2712]: Submission to National Institute of Health and Care Excellence. Single technology appraisal (STA): Document B - Company evidence submission*: Bristol-Myers Squibb Pharmaceuticals Ltd., 2022. 173p.

[8] Bristol-Myers Squibb Pharmaceuticals Ltd. *Nivolumab in combination with fluoropyrimidine- and platinum-based combination chemotherapy for the first-line treatment of adult patients with unresectable advanced, recurrent, or metastatic oesophageal squamous cell carcinoma with tumour cell PD-L1 expression ≥1% [ID2712]: Submission to National Institute of Health and Care Excellence. Single technology appraisal (STA): Appendix L (06.05.22 update) - Network meta-analysis report - October 2021 database lock. Version 1*: Bristol-Myers Squibb Pharmaceuticals Ltd., 2022. 51p.

[9] National Institute for Health and Care Excellence. *Pembrolizumab with platinum-based chemotherapy for untreated advanced oesophageal cancer [ID3741]. Committee Papers*. London: NICE, 2021 Available from: https://www.nice.org.uk/guidance/ta737/documents/committee-papers

[10] Latimer NR, Abrams KR. *NICE DSU Technical Support Document 16: Adjusting survival time estimates in the presence of treatment switching*. Sheffield: Decision Support Unit, ScHARR, 2014. 57p. Available from: http://nicedsu.org.uk/wp-content/uploads/2016/03/TSD16_Treatment_Switching.pdf