in collaboration with:

ERASMUS UNIVERSITEIT ROTTERDAM
INSTITUTE OF HEALTH POLICY & MANAGEMENT

Maastricht University

Nivolumab for treating recurrent or metastatic squamous-cell carcinoma of the head and neck after platinum-based chemotherapy: CDF review of TA490

Produced by Kleijnen Systematic Reviews Ltd. (KSR) in collaboration with Erasmus University Rotterdam (EUR) and Maastricht University Medical Centre (UMC+)

Authors

Nigel Armstrong, Health Economics Manager, KSR
Bram Ramaekers, Health Economist, Maastricht UMC+
Lloyd Brandts, Health Economist, Maastricht UMC+
Ben Wijnen, Health Economist, Maastricht UMC+, Trimbos Institute Utrecht
Debra Fayter, Reviewer, KSR
Titas Buksnys, Health Economist, KSR
Charlotte Ahmadu, Health Economist, KSR
Vanesa Huertas-Carrera, Reviewer, KSR
Rob Riemsma, Reviews Manager, KSR
Gill Worthy, Statistician, KSR
Kate Misso, Information Specialist Manager, KSR
Steven Duffy, Information Specialist, KSR

Manuela Joore, Health Economist, Professor of Health Technology Assessment & Decision Making, Maastricht UMC+
Jos Kleijnen, Director, KSR Ltd, Professor of Systematic Reviews in Health Care, Maastricht University

Correspondence to Nigel Armstrong, Kleijnen Systematic Reviews
Unit 6, Escrick Business Park
Riccall Road, Escrick
York, UK
YO19 6FD

Date completed 30/03/2020

Source of funding: This report was commissioned by the National Institute for Health Research (NIHR) Health Technology Assessment (HTA) Programme as project number CDF Review TA490 13/04/77.

Declared competing interests of the authors
None.

Acknowledgements
None. Copyright belongs to Kleijnen Systematic Reviews Ltd.

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

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

This report should be referenced as follows:
Armstrong N, Ramaekers BLT, Brandts L, Wijnen B, Fayter D, Buksnys T, Ahmadu C, Huertas-Carrera V, Riemsma R, Worthy G, Misso K, Duffy S, Joore MA, Kleijnen J. Nivolumab for treating recurrent or metastatic squamous-cell carcinoma of the head and neck after platinum-based chemotherapy: CDF review of TA490. York: Kleijnen Systematic Reviews Ltd, 2020.

Contributions of authors
Nigel Armstrong acted as project lead as well as systematic review and health economist on this assessment, critiqued the clinical effectiveness methods and evidence as well as the company's economic evaluation and contributed to the writing of the report. Bram Ramaekers acted as health economic project lead, critiqued the company's economic evaluation and contributed to the writing of the report. Lloyd Brandts, Ben Wijnen, Titas Buksnys and Charlotte Ahmadu acted as health economists on this assessment, critiqued the company's economic evaluation and contributed to the writing of the report. Debra Fayter, Vanessa Huertas-Carrera and Rob Riemsma acted as systematic reviewer on this assessment, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Gill Worthy acted as statistician, critiqued the analyses in the company's submission and contributed to the writing of the report. Kate Misso and Steven Duffy contributed to the writing of the report. Manuela Joore acted as health economists on this assessment, critiqued the company's economic evaluation, contributed to the writing of the report and provided general guidance. 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 event
AIC	Akaike information criterion
ALK	anaplastic lymphoma kinase
ASBI	Average Symptom Burden Index
AUC	area under the curve
BIC	Bayesian information criterion
BICR	blinded independent central review
BMS	Bristol-Myers Squibb
BOR	best objective response
BRAF	B-Raf proto-oncogene
BSC	best supportive care
BTLA	B- and T-lymphocyte attenuator
CD27	cluster of differentiation 27
CD28	cluster of differentiation 28
CD137	cluster of differentiation 137
CDF	Cancer Drugs Fund
chemo	chemotherapy
CHMP	Committee for Medicinal Products for Human Use
CI	confidence interval
CNS	central nervous system
COMP	comparator
CR	complete response
CSR	clinical study report
CT	computed tomography
CTLA-4	cytotoxic T-lymphocyte antigen-4
DMC	Data Monitoring Committee
DOR	duration of response
DoT	duration of treatment
DSA	deterministic sensitivity analysis
DSU	Decision Support Unit
ECG	electrocardiogram
ECOG	Eastern Cooperative Oncology Group
EGFR	epidermal growth factor receptor
EMA	European Medicines Agency
eMIT	electronic market information tool
EQ-VAS	EQ-5D Visual Analogue Scale
ER	endoplasmic reticulum
ERG	Evidence Review Group
EU	European Union
GCP	good clinical practice
GITR	glucocorticoid-induced tumour necrosis factor receptor
GP	general practitioner
HIV	human immunodeficiency virus
HR	hazard ratio
HRG	Healthcare Resource Groups
HRQoL	health-related quality of life
HTA	health technology assessment
HVEM	herpes virus entry mediator
IC	investigator's choice
ICER	incremental cost-effectiveness ratio
IgG4	immunoglobulin G4
IMAE	immune-mediated adverse event
INT	intervention

IO	immuno-oncology
IO-IO	immuno-oncology–immuno-oncology combination therapy
ipi	ipilimumab
IRRC	independent radiology review committee
IV	intravenous/intravenously
IVRS	interactive voice response system
LAG3	lymphocyte-activation gene 3
LCSS	Lung Cancer Symptom Scale
LS	least squares
LY	life-year
LYG	life-year gained
MHC	major histocompatibility complex
MID	minimally important difference
MRI	magnetic resonance imaging
mut/Mb	mutations per megabase
NA	not applicable
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
nivo	nivolumab
NR	not reached
NSCLC	non-small cell lung cancer
NSQ	non-squamous
ORR	objective response rate
OS	overall survival
OX40	tumour necrosis factor receptor superfamily, member 4
PAS	patient access scheme
PD	progressed disease
PD-1	programmed death-1
PDC	platinum doublet chemotherapy
PD-L1	programmed death-ligand 1
PD-L2	programmed death-ligand 2
PF	progression-free
PFS	progression-free survival
PR	partial response
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PS	performance status
PSA	probabilistic sensitivity analysis
PSS	Personal Social Services
PSSRU	Personal Social Services Research Unit
Q12W	every 12 weeks
Q2W	every 2 weeks
Q3W	every 3 weeks
Q6W	every 6 weeks
QALY	quality-adjusted life-year
RANK-L	receptor activator of nuclear factor kappa-B ligand
RCT	randomised controlled trial
RECIST 1.1	Response Evaluation Criteria in Solid Tumors version 1.1
ROC	receiver operating characteristic
ROS1	ROS proto-oncogene 1
SACT	Systemic Anti-Cancer Therapy
SAE	serious adverse event
SIGN	Scottish Intercollegiate Guidelines Network
SLR	systematic literature review
SmPC	summary of product characteristics
SQ	squamous

STA	single technology appraisal
TAP	transporter associated with antigen processing
TCR	T-cell receptor
TIM3	T-cell immunoglobulin and mucin-domain containing-3
TMB	tumour mutational burden
ToE	terms of engagement
TPS	tumour proportion score
TRAE	treatment-related adverse event
TTD	time to treatment discontinuation
TTR	time to response
UK	United Kingdom
US	United States
VAS	visual analogue scale
VISTA	V-domain immunoglobulin suppressor of T-cell activation
WTP	willingness to pay

Table of Contents

Table of Tables

Table of Figures

1. EXECUTIVE SUMMARY

1.1 Critique of the adherence to committees preferred assumptions from the Terms of Engagement (ToE) in the company's submission

The following is a list of the key committee assumptions (preferences) according to the ToE for the Cancer Drugs Fund (CDF) review, each one followed by a statement as to the Evidence Review Group's (ERG's) finding of the extent to which the company submission (CS) has adhered to the committee preferences (See Section 2.2 for more details).

Assumption 0: Nivolumab administered according to a weight base dose (3 mg/kg every two weeks). This was not specified in the ToE, but it might be regarded a tacit assumption. Since the original submission for TA490, on 28 April 2017 the licensed dose of nivolumab has been updated to a flat dose of 240 mg every two weeks (Q2W). The ERG questions the validity of the conclusion by the company that there will be no clinically meaningful difference between weight-based and the specific flat dose of 240 mg in terms of effectiveness and safety given that many patients will have to either increase or decrease dosage.

Assumption 1: Population: adults with recurrent or metastatic squamous-cell carcinoma of the head and neck (SCCHN) that progressed within six months of platinum-based therapy, in either the early or locally advanced disease stage. The ERG notes that there is an apparent discrepancy in that the eligibility criteria for CheckMate 141 include progression at the metastatic or recurrent disease stage. However, there is correspondence between CheckMate 141 and the Systemic Anti-Cancer Therapy (SACT) dataset and the ToE also stated that the CheckMate 141 results are relevant to the population of interest and therefore then this could be considered as tantamount to adherence to the committee's preferred assumption.

Assumption 2: Docetaxel is the comparator of interest. The ERG notes that there appears to be incomplete adherence in that, although it is a comparator in the cost effectiveness analysis, the clinical effectiveness data used to inform this analysis and the clinical effectiveness evidence presented were based on a comparison of nivolumab to investigator choice (IC), i.e. using the all-randomised (full intention to treat) data. Using the all-randomised data, including that from the whole IC arm implies equivalence between docetaxel and methotrexate, which the ToE explicitly rejects. The ERG would therefore argue that the best source of evidence for a comparison with docetaxel should be the subgroup of those chosen to receive docetaxel according to IC (docetaxel subgroup).

Assumption 3: CheckMate 141 data to be used. The ERG can confirm that this assumption was adhered to in the CS, notwithstanding the omission of the docetaxel subgroup.

Assumption 4: Overall survival from CheckMate 141 data updated. The ERG can confirm that this assumption was adhered to in the CS.

Assumption 5: Analysis of the effect of PD L1 expression on updated OS. The ERG can confirm that this assumption was adhered to in the CS.

Assumption 6: No change in model structure. The ERG can confirm that the model structure was unchanged.

Assumption 7: Piecewise model used for extrapolation of survival: timepoint to extrapolate and distribution to be explored. The ERG can confirm that piecewise models were indeed used to extrapolate survival while using alternative cut-off points and two different distributions.

Assumption 8: Continued treatment benefit to be reviewed in light of any new evidence. The ERG notes that the company argued that in light of the new evidence, the assumption of continued treatment benefit (i.e. no treatment waning) was plausible. The ERG, however, preferred to incorporate treatment waning of the nivolumab OS benefit after year 5.

Assumption 9: Quality-of-life benefit of nivolumab cannot be assumed to remain constant. Exploration of the most appropriate utility values should be reviewed in light of any new evidence. The ERG notes that this was only done partly as health state utility values are not updated and it is questionable whether the company's approach to incorporate utility benefit over time appropriately addresses the concerns raised in the ToE.

Assumption 10: The ToE stipulated that the committee considered analyses without a stopping rule are more appropriate for decision-making. However, the appropriateness of a two-years stopping rule should be reviewed in light of any new evidence. The ERG notes that the company stated that based on the time to treatment discontinuation (TTD) extrapolation used in its base-case, [redacted], and a two-year stopping rule has been shown to be clinically plausible during the CDF data collection period. The ERG preferred to exclude the two-year stopping rule, consistent with committee preferences as reported in the ToE.

Assumption 11: ERG amendments will be included (adding the cost and disutility for pneumonitis and using treatment-independent proportions for subsequent treatment). The ERG can confirm that these amendments were included.

1.2 Summary of key issues in the clinical effectiveness evidence

1) Update of CheckMate 141 overall survival (OS) data, according to the ToE: The ERG can confirm that this has been done with the latest data cut being 15 October 2019, i.e. four years follow-up. The results show that the survival advantage of nivolumab over IC was maintained in terms of hazard ratio (HR) and median survival and continued through 36 months and at 48 months. Also, the company provided the up to date data from CheckMate 141 on progression-free survival (PFS) and the ERG can confirm that there is no fundamental change in interpretation: the advantage of nivolumab versus IC in terms of HR and the small advantage of IC versus nivolumab in terms of median survival, were maintained, although neither were statistically significant. Although the ToE did not specify an update in terms of safety, it appears from the company response to clarification, that little has changed in both the number and percentage of AEs between TA490 and the CDF review, which leads to the same conclusion as found by the ERG in TA490, i.e. nivolumab was generally well tolerated by patients in CheckMate 141 compared to IC of therapy, with a lower proportion of patients receiving nivolumab experiencing Grade 3-4 all-causality adverse events (AEs). Given that the committee concluded that the comparator should be docetaxel, the ERG considers that the most appropriate evidence of effectiveness and safety versus docetaxel is that from the docetaxel subgroup, which the company did not provide in either the CS or in response to the clarification letter. The ERG considers that this is a major source of uncertainty that can be reduced by the company.

2) SACT dataset to assess the generalisability of CheckMate 141, according to the ToE: A comparison reveals that UK patients might be slightly older and a small number will have a worse performance status than the patients in the all-randomised population of the CheckMate 141 trial, which might suggest that UK patients do slightly worse than patients in the CheckMate 141 trial. However, although patients in the SACT dataset had a numerically lower median survival than those in the nivolumab arm of Check Mate 141, it is important to remember that this was based on a much shorter median follow-

up and the 95% CIs overlapped. Also, one-year survival was very similar. As mentioned with regards to the comparison between nivolumab and docetaxel, it could be argued that the nivolumab arm of the docetaxel subgroup of CheckMate 141 should be used to compare with the SACT dataset. On the other hand, the all-randomised population might be closer to those patients who would be treated with nivolumab in UK clinical practice. This was the judgment of the committee, who concluded that the CheckMate 141 results (implying the all-randomised population) are relevant to the population of interest, i.e. adults with recurrent or metastatic SCCHN that progressed within six months of platinum-based therapy, in either the early or locally advanced disease stage. Although there is a discrepancy between descriptions of eligibility criteria, those for the SACT dataset could also be regarded as essentially the same as those for CheckMate 141. However, clearly not everyone in CheckMate 141 was found to be eligible for docetaxel according to the IC design. In particular, some were chosen to receive methotrexate, which in the ToE states that it is only for patients who are not fit to have a taxane. The implication of this should be that the population specified for this CDF review and in the SACT dataset should not be aligned with the all-randomised population, but should at least exclude those who would be ineligible for docetaxel. This apparent mismatch between the population and the comparator specified in the ToE does produce some uncertainty in the generalisability of the CheckMate 141, which might be reduced by a comparison of the baseline characteristics and OS in the nivolumab arm of the docetaxel subgroup and the SACT dataset.

3) In terms of PD-L1 status, nivolumab showed an advantage in comparison to IC for both groups, but it was larger for those with PD-L1 ≥ 1% and only statistically significant for this subgroup. However, there was no significant evidence of a treatment and subgroup interaction (p=0.239) and these results should be considered with caution due to the reduced sample sizes and wider confidence intervals. For PFS, HRs were not provided for the PD-L1 subgroups, but the median PFS estimates indicate that there were no significant differences in PFS between nivolumab and IC in patients with PD-L1 <1% or those with PD-L1 ≥ 1%. There was also evidence of only a weak interaction effect. [redacted]

1.3 Summary of the key issues in the cost effectiveness evidence

The company base-case incremental cost effectiveness ratio (ICER) (probabilistic) of nivolumab (with patient access scheme (PAS)) compared with docetaxel was £36,255 per quality-adjusted life-year (QALY) gained. The ERG has incorporated various adjustments to the company base-case. The ERG base-case resulted in an ICER range (probabilistic) of £54,348 to £61,293 per QALY gained for nivolumab (with PAS) versus docetaxel. The most influential adjustments/corrections made by the ERG were:

1) using a generalised gamma distribution for estimating TTD;
2) using treatment independent utilities for PFS and PD health states;
3) including treatment waning of nivolumab OS benefit after year 5 and;
4) excluding the two-year stopping rule.

Additionally, the company explored using SACT data to estimate TTD (i.e. nivolumab treatment duration) in scenario analyses. Compared with the CheckMate 141 trial, the SACT data provides real-world data that might better reflect UK clinical practice. The higher TTD observed in the SACT data resulted in a substantially increased ICER (+£14,198 compared to the CS base-case) highlighting the importance of the TTD assumptions in the model.

The equivalence assumptions between docetaxel and methotrexate as well as between the nivolumab flat dose and weight-based nivolumab can be questioned. Unfortunately, the company did not provide analyses based on the docetaxel subgroup (requested during the clarification phase), nor evidence to support the equivalence assumption between the flat dose and the weight-based dose of nivolumab. An additional area of uncertainty is the extrapolation of the nivolumab quality-of-life benefit over time. Although the company implemented utility decrements related to the time to death, the ERG believes that the committee's concern (i.e. emphasising that quality-of-life benefit cannot be assumed to remain constant over time) is not appropriately addressed. Therefore, the ERG base-case is presented as a range conditional on treatment dependent and treatment independent utilities to address the uncertainty related to the nivolumab utility benefit over time.

The subgroup analyses (based on PD-L1 status) performed by the ERG resulted in ICERs that ranged between £53,152 and £62,895 per QALY gained. It should however be noted that these subgroup analyses did not incorporate any additional costs related to PD-L1, which would be required if PD-L1 testing is not part of UK clinical practice.

In conclusion, the ERG base-case ICERs are estimated to be in the range between £54,348 and £61,293 per QALY gained, reflecting the uncertainty related to nivolumab quality-of-life benefits over time. Uncertainty that was not captured in this range included the equivalence assumptions between docetaxel and methotrexate as well as between the nivolumab flat dose and weight-based nivolumab. Additionally, if the nivolumab treatment duration from the SACT is believed to better reflect UK clinical practice (than TTD from CheckMate 141), this would substantially increase the estimated ICERs.

1.4 Summary of ERG's preferred assumptions and resulting ICER

Table 1.1: ERG analyses (deterministic), nivolumab with PAS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Nivolumab ICER (£/QALY)
Company base-case					
Nivolumab	[redacted]	[redacted]			
Docetaxel	£10,569	0.35	[redacted]	[redacted]	£37,236
1 Company base-case + OS treatment waning[a]					
Nivolumab	[redacted]	[redacted]			
Docetaxel	£10,569	0.35	[redacted]	[redacted]	£45,017
2 Company base-case + generalised gamma model for estimating TTD					
Nivolumab	[redacted]	[redacted]			
Docetaxel	£10,505	0.35	[redacted]	[redacted]	£39,959
3 Company base-case + treatment independent utility					
Nivolumab	[redacted]	[redacted]			
Docetaxel	£10,569	0.38	[redacted]	[redacted]	£41,418
4 Company base-case					

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Nivolumab ICER (£/QALY)
+ excluding the 2-year stopping rule					
Nivolumab	[redacted]	[redacted]			
Docetaxel	£10,569	0.35	[redacted]	[redacted]	£49,018
5 Company base-case + correcting error related to implementation of docetaxel dose intensity					
Nivolumab	[redacted]	[redacted]			
Docetaxel	£10,561	0.35	[redacted]	[redacted]	£37,254
6 ERG base-case 1 Company base-case + OS treatment waning + generalised gamma model for estimating TTD + excluding the 2-year stopping rule					
Nivolumab	[redacted]	[redacted]			
Docetaxel	£10,497	0.35	[redacted]	[redacted]	£53,485
7 ERG base-case 2 Company base-case + OS treatment waning + generalised gamma model for estimating TTD + excluding the 2-year stopping rule + treatment independent utility					
Nivolumab	[redacted]	[redacted]			
Docetaxel	£10,497	0.38	[redacted]	[redacted]	£60,094

ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year; TTD = time to treatment discontinuation
[a]A minimum function was implemented to prevent that PFS would exceed OS (implemented in cells 'Nivolumab Traces'!G11:G370 and 'Docetaxel Traces'!G11:G370)
[b]The following cells were adjusted: Settings!J72:N72, 'Treatment Costs'!N24 and 'Docetaxel Traces'!AU11:AU369

Table 1.2: ERG base-case (probabilistic), nivolumab with PAS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Nivolumab ICER (£/QALY)
6 ERG base-case 1- treatment dependent utility[a]					
Nivolumab	[redacted]	[redacted]			
Docetaxel	£10,556	0.36	[redacted]	[redacted]	£54,348
7 ERG base-case 2 - treatment independent utility[a]					
Nivolumab	[redacted]	[redacted]			
Docetaxel	£10,511	0.38	[redacted]	[redacted]	£61,293

ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year; TTD = time to treatment discontinuation
[a]The PSA produced 1 to 2 errors (#VALUE), these simulations were ignored to calculate the probabilistic means.

Table 1.3: ERG base-case; PD-L1 <1% subgroup (deterministic), nivolumab with PAS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Nivolumab ICER (£/QALY)
6 ERG base-case 1- treatment dependent utility					
Nivolumab	[redacted]	[redacted]			
Docetaxel	£11,048	0.41	[redacted]	[redacted]	£53,152
7 ERG base-case 2 - treatment independent utility					
Nivolumab	[redacted]	[redacted]			
Docetaxel	£11,048	0.43	[redacted]	[redacted]	£62,895
ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year; TTD = time to treatment discontinuation					

Table 1.4: ERG base case; PD-L1 ≥1% subgroup (deterministic), nivolumab with PAS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Nivolumab ICER (£/QALY)
6 ERG base-case 1- treatment dependent utility					
Nivolumab	[redacted]	[redacted]			
Docetaxel	£9,981	0.29	[redacted]	[redacted]	£54,362
7 ERG base-case 2 - treatment independent utility					
Nivolumab	[redacted]	[redacted]			
Docetaxel	£9,981	0.31	[redacted]	[redacted]	£58,926
ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year; TTD = time to treatment discontinuation					

1.5 Summary of exploratory and sensitivity analyses undertaken by the ERG

Table 1.5: ERG scenario (deterministic), nivolumab with PAS for all-randomised population

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Nivolumab ICER (£/QALY)
6 ERG base-case 1- treatment dependent utility + excluding the estimated utility decrements related to time before death					
Nivolumab	[redacted]	[redacted]			
Docetaxel	£10,497	0.36	[redacted]	[redacted]	£50,140
7 ERG base-case 2 - treatment independent utility + excluding the estimated utility decrements related to time before death					
Nivolumab	[redacted]	[redacted]			
Docetaxel	£10,497	0.40	[redacted]	[redacted]	£60,264
ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year; TTD = time to treatment discontinuation					

2. INTRODUCTION AND BACKGROUND

2.1 Background

The ToE for the CDF review states the following:[1] "Nivolumab is recommended for use within the Cancer Drugs Fund as an option for treating squamous cell carcinoma of the head and neck (SCCHN) in adults whose disease has progressed on platinum-based chemotherapy, only if:

- the disease has progressed within 6 months of having chemotherapy
- nivolumab is stopped at 2 years of uninterrupted treatment, or earlier in the event of disease progression and
- the conditions in the managed access agreement are followed."

The committee concluded that based on a PAS of [redacted] and its preferred assumptions the most plausible ICER would fall between £45,000 and £73,600 per QALY (dependent on the time point for extrapolation and treatment-dependent/independent utility values) for the full trial population, irrespective of PD-L1 expression.

Nivolumab was accepted in the CDF on the basis of two main conditions, which formed the managed access agreement:

1) A further discount, i.e. [redacted] commercial access agreement, which implied an ICER of £30,377 and £49,408 per QALY gained depending on the time point used for extrapolation and assuming a 2-year stopping rule.

2) A data collection agreement, reported as follows:[1]

- "The pivotal clinical-effectiveness evidence for nivolumab compared with investigator-choice was taken from the CheckMate 141 trial. This trial is the primary source for data collection under the managed access agreement. 4-year follow-up data would be undertaken based on the trial protocol including the reporting of OS, treatment duration and sub-group analysis by PD-L1 expression level. The company will provide updated evidence on the CheckMate 141 trial.
- Observational data will also be collected for nivolumab during the period of managed access via the systemic anti-cancer therapy (SACT) dataset to support the data collected in the clinical trial. SACT will collect data on OS, duration of therapy and PDL-1 expression. Public Health England will provide a summary of the observational data collected."

The index population is consistent with a subgroup of the licensed indication, i.e. "…recurrent or metastatic squamous cell carcinoma of the head and neck (R/M SCCHN) in adults progressing on or after platinum-based therapy".[2] The recommended dosage of nivolumab in this indication is 240 mg flat dose every two weeks. This is different to the weight-based dose of 3 mg/kg every two weeks that was recommended at the time of the original NICE appraisal for nivolumab in this indication.

2.2 Critique of company's adherence to committees preferred assumptions from the Terms of Engagement

Table 2.1 summarises the key committee assumptions (preferences) according to the ToE for CDF review.[1] It also summarises the extent to which the CS has adhered to the committee preferences.[2] In addition, the ToE state that the end-of-life criteria have been met.

ERG comments:

Assumption 0: Nivolumab dosing

There is a tacit assumption that was not specified in the ToE, which is the nature of the intervention, in particular the dosing regimen, no mention of which was made in the ToE.[1] Since the original submission for TA490, on 28 April 2017 the licensed dose of nivolumab was updated to a flat dose of 240 mg every two weeks (Q2W), rather than the weight-based dose used in the CheckMate 141 trial (3 mg/kg every two weeks).[2]

The company in their submission state that *"Nivolumab flat-dosing regimens are supported by clinical safety data and population pharmacokinetic modelling across many indications, which demonstrated that distributions of nivolumab exposures after 3 mg/kg Q2W and 240 mg Q2W were similar and below the exposures observed with 10 mg/kg Q2W. No clinically meaningful relationship between body weight or nivolumab exposure or nivolumab exposure quartiles and frequency or severity of adverse events was observed. Based on consistent exposure-response relationships across indications, the benefit-risk profile of nivolumab 240 mg Q2W is likely to be similar to 3 mg/kg Q2W, therefore the clinical effectiveness of nivolumab that was demonstrated in CheckMate 141 (weight-based dose) is expected to be generalisable to the use of nivolumab in clinical practice (flat dose)."*(p.9)[2] However, no reference to any source of evidence was provided.

There is the suggestion of some evidence that might provide some support for the use of the flat dose of 240mg from the web-site of the European Medicines Agency (EMA), which states that the introduction of the new dosing regimens of 240 mg every two weeks was based on a "comparison of the exposure-response and safety of nivolumab 3 mg/kg Q2W, 240 mg Q2W in … squamous cell cancer of the head and neck…" (p.11)[3] The summary of product characteristics also states: "Based on modelling of dose/exposure efficacy and safety relationships, there are no clinically significant differences in efficacy and safety between a nivolumab dose of 240 mg every 2 weeks or 3 mg/kg every 2 weeks." (p.26)[4]

The ERG therefore requested empirical evidence from the company with references to support the claim that there will be no meaningful difference in either effectiveness or AE risk between the two methods of dosing, i.e. weight-based and flat dose. However, in response to clarification, the company did not provide any further evidence beyond those produced by the EMA.[3-5] Therefore, the ERG still questions the validity of the conclusion that there will be no clinically meaningful difference between weight-based and the specific flat dose of 240mg in terms of effectiveness and safety given that many patients will have to either increase or decrease dosage.

Assumption 1: Trial population

The committee concluded that, although there are some differences between the trial population and the UK population, the CheckMate 141 results are relevant to the population of interest, i.e. adults with recurrent or metastatic SCCHN that progressed within six months of platinum-based therapy, in either the early or locally advanced disease stage.[1] However, as shown in Table 3.1, the actual CheckMate 141 trial eligibility criteria included the recurrent, or metastatic setting. The ERG notes that it is unclear what difference this might make to the outcomes. However, also shown in Table 3.1, the SACT dataset applied the same additional criteria and therefore one might reasonably conclude that, if the SACT dataset represents clinical practice then the index population should include these additional criteria and also that CheckMate 141 trial is not compromised by this discrepancy.

Assumption 2: Docetaxel comparator

In the ToE, the committee also concluded that the comparator should be docetaxel.[1] They also raised concerns about the generalisability of CheckMate 141 and that it should not be assumed that docetaxel was comparable to the other comparator, methotrexate because it is only for patients who are not fit to have a taxane. The ERG would like to point out that one would therefore expect that the actual population that should be eligible for nivolumab would be only those who might otherwise receive docetaxel. However, it is unclear how this population might be defined precisely, e.g. according to ECOG performance status. There is also no indication from the SACT dataset report that only those eligible for docetaxel were given nivolumab in the CDF. Therefore, it is unclear which of the CheckMate 141 populations would be most representative of UK clinical practice, the all-randomised or the subgroup of patients eligible for docetaxel (who would have been chosen to receive docetaxel according to IC), i.e. those patients who were randomised to docetaxel vs. those who would have received docetaxel according to IC, but who were randomised to nivolumab. The ERG will refer to this subgroup from this point onwards as the 'docetaxel subgroup'.

Nevertheless, in order to assess the comparability of the nivolumab baseline characteristics and outcomes to the SACT dataset (Sections 3.1 and 3.2), it is less clear whether the docetaxel subgroup should be chosen. On the one hand, this would be consistent with the comparator being treated with docetaxel. On the other hand, if the population of the SACT dataset is the same as the CheckMate 141 trial all-randomised population then to exclude patients in the cetuximab or methotrexate subgroups would exclude patients who are also eligible for nivolumab. Nevertheless, the ERG would argue that, on balance, the effectiveness of nivolumab vs. docetaxel should be estimated from the docetaxel subgroup. Although the company used docetaxel as a comparator in the cost effectiveness analysis, it was based on data from the all-randomised population.[2] Because of this, the ERG requested the company to perform analyses in the docetaxel subgroup. The company responded by stating firstly that there was insufficient time to perform these analyses.[5] The company also argued that it had been demonstrated in TA490 that the comparisons using the docetaxel subgroup would have minimal impact on the cost effectiveness results, although no summary measures of treatment effect (e.g. HRs) were presented at that time.[6] This also adds additional uncertainty to the estimated cost effectiveness. The company goes on to present four more arguments against the docetaxel subgroup analyses:

1) the trial was not powered for subgroup analysis by IC. The ERG recognise that this is true, but this is not a reason not to present the analyses, but instead a reason for caution in interpretation.
2) because the choice of intended IC therapy was made prior to randomisation, the analysis of outcomes by individual therapies in the IC arm breaks randomisation and is at risk of selection bias. However, it is precisely because the choice was made before randomisation that there is no selection bias: all patients chosen to have a specific IC were randomised to either that choice of IC or nivolumab.
3) data from the all-randomised IC arm, regardless of specific subgroup, i.e. the all-randomised data, were found in the FAD of TA490 to be appropriate for decision making and that the ToE stipulates no deviation from the committee's preferred assumptions. However, the list of assumptions in the ToE does not explicitly state that only the all-randomised data should be used. The ToE also states, unlike in the FAD, that the comparator should be docetaxel.
4) it would be wrong to focus only on docetaxel as a comparator given that patients not fit enough to take it would receive methotrexate. This is not a reason to not provide the docetaxel subgroup data for a comparison with docetaxel, but instead might be a reason to provide the methotrexate subgroup data for a comparison with methotrexate.

Assumptions 3 to 8 and 11

The ERG can confirm that these assumption were adhered to in the CS, notwithstanding the omission of the docetaxel subgroup and the change in dosing referred to above.

Assumptions 9 and 10

The extent of adherence to these assumptions is discussed in detail in Chapter 4.

Table 2.1: Preferred assumption from ToE

Assumption	Terms of Engagement	Addressed to by the company submission	Rationale if different	ERG comment
Assumption 1	Population: adults with recurrent or metastatic SCCHN that progressed within 6 months of platinum-based therapy, in either the early or locally advanced disease stage.	Incomplete: mismatch with CheckMate 141 trial.	None given.	Probably not a problem. See Chapter 2 for details.
Assumption 2	Docetaxel is the comparator of interest.	Incomplete: Docetaxel subgroup data not presented or used in the cost effectiveness analysis.	None given.	See Chapter 2 for details.
Assumption 3	CheckMate 141 data to be used.	Yes	Not applicable.	See Chapter 3 for details.
Assumption 4	Overall survival from CheckMate 141 data updated	Yes	Not applicable.	See Chapter 3 for details.
Assumption 5	Analysis of the effect of PD L1 expression on updated OS	Yes	Not applicable	See Chapter 3 for details.
Assumption 6	No change in model structure	Yes	Not applicable	See Chapter 4 for details.
Assumption 7	Piecewise model used for extrapolation of survival: timepoint to extrapolate and distribution to be explored.	Yes	Not applicable	See Chapter 4 for details.
Assumption 8	Continued treatment benefit to be reviewed in light of any new evidence.	Yes	Not applicable	See Chapter 4 for details.
Assumption 9	Quality-of-life benefit cannot be assumed to remain constant. Exploration of the most appropriate utility values should be reviewed in light of any new evidence.	Incomplete, health state utility values were not updated and the approach to incorporate utility benefit over time might be debatable.	Not applicable	See Chapter 4 for details.

Assumption	Terms of Engagement	Addressed to by the company submission	Rationale if different	ERG comment
Assumption 10	Appropriateness of a 2-years stopping rule should be reviewed in light of any new evidence.	Incomplete, inclusion of stopping rule might be debatable.	Not applicable	See Chapter 4 for details.
Assumption 11	ERG amendments will be included (adding the cost and disutility for pneumonitis and using treatment-independent proportions for subsequent treatment).	Yes	Not applicable	Not applicable

Source: Based on table of key committee assumptions as reported in the Terms of engagement for CDF review.[1] and the company submission[2]
ERG = evidence review group; CDF = cancer drugs fund

3. CLINICAL EFFECTIVENESS

3.1 Overview of the new clinical evidence

3.1.1 Sources of evidence

The clinical efficacy of nivolumab in the treatment of SCCHN has been investigated in one RCT, CheckMate 141.[2, 7, 8] CheckMate 141 is a phase III, multicentre, open-label, active-controlled randomised trial comparing the efficacy and safety of nivolumab with investigator's choice (IC), which included choice at the clinician's discretion of docetaxel, methotrexate or cetuximab. Its main methodological features are summarised in Table 3.1. The new evidence from this trial is from the latest data cut of the trial (four-year; 15 October 2019).

The other source is the SACT dataset.[9] This was specified in the ToE and created, at the behest of NHS England and NHS Improvement, by Public Health England (PHE), with the purpose of evaluating the real-world treatment effectiveness of nivolumab in the CDF population during the managed access period.[1] It provides evidence on treatment duration, OS and the reasons for stopping treatment (described as 'treatment outcomes') for all patients treated with nivolumab for the same population as in the CheckMate 141 trial.

ERG comment: The SACT dataset permits to some degree a test of the generalisability of the outcomes observed in the CheckMate 141 trial, at least in the nivolumab arm, to UK clinical practice. For this reason, throughout the following sections the ERG will compare these two data sources both to establish comparability of outcomes in terms of design and baseline characteristics and in terms of the outcomes, OS and TTD.

3.1.2 Patient characteristics in CheckMate 141 and SACT

As noted in the previous ERG report, baseline characteristics seemed to be comparable between the two treatment arms of CheckMate 141 (nivolumab and IC), although unsurprisingly, given the IC design, this is not the case between the various treatments (Table 3.2).[6] For example, the percentage of patients who have received at least three lines of therapy is higher for methotrexate and nivolumab than for docetaxel.

The company provided a summary and table comparing the baseline characteristics of the nivolumab arm of the CheckMate 141 trial and the SACT cohort reported by Public Health England. See Table 3.3. Limited information was available concerning the SACT cohort so comparisons can only be made on gender, age, ECOG performance status and PD-L1 scores. It can be seen in Table 3.3 that the number of males was similar in the CheckMate trial and in the SACT cohort (82% versus 81%). Median age in the SACT cohort was slightly older (62 in SACT versus 59 in CheckMate), which was consistent with the larger proportion of those in the older age groups.[2, 5]

As regards ECOG performance status, the numbers with a PS of 0 were fairly similar (20% in CheckMate versus 24% in SACT) but there were more patients with PS of 1 in CheckMate (79% versus 57%). Only one patient (0.4%) in CheckMate had a PS of 2 or more (inclusion criteria for CheckMate was PS of 0 or 1). The SACT cohort had 29 patients with a PS of 2 and four patients with a PS of 3 (7% overall). Thirteen percent of the SACT data were missing so it is possible that some of these patients had a higher PS status. It was not possible to estimate comparability in terms of breakdown of PD-L1 scores as 42% of SACT scores were not recorded. Additionally, both the trial and the SACT cohort had over 30% of scores which could not be quantified.

ERG comment: Although the baseline characteristics between the arms for the all-randomised population are comparable, a comparison of baseline characteristics between the arms for the docetaxel subgroup could be valuable. This was requested as an additional clarification question, which the company did not provide (see Section 2.2).[5] Taking the SACT cohort as being typical of patients to be seen in clinical practice, UK patients might be slightly older and a small number will have a worse performance status than the patients in the all-randomised CheckMate 141 trial. Assuming that other disease characteristics and prior therapies are similar between the two data sources, it might be expected that UK patients do slightly worse than patients in the CheckMate 141 trial. However, this does not appear to be the case looking at the SACT data (see Section 3.2).

Table 3.1: Summary of methodology of CheckMate 141 trial and SACT dataset

Trial name	CheckMate 141	SACT dataset
Location	International: 55 study sites across 15 countries in North America (USA and Canada), South America, Europe and Asia. Five study sites were included in the UK, with a total of 34 patients randomised to study treatment at UK sites.[10]	UK
Design	Multicentre, open-label, phase III randomised controlled trial	Observational study
Eligibility criteria for participants	*Key inclusion criteria:* • Males and females ≥18 years of age with an ECOG performance status of 0 or 1 • Histologically confirmed R/M SCCHN (oral cavity, pharynx, larynx), stage III/IV and not amenable to local therapy with curative intent (surgery or radiation therapy with or without chemotherapy) • Tumour progression or recurrence within 6 months of last dose of platinum therapy in the adjuvant, primary, recurrent, or metastatic setting • Measurable disease by CT or MRI per RECIST 1.1 criteria[11] • Documentation of p-16 positive or p-16 negative disease to determine HPV-p16 status of tumour for SCCHN of the oropharynx • Availability of tumour samples for PD-L1 expression analysis *Key exclusion criteria:* Active, known or suspected autoimmune disease Systemic treatment with either corticosteroids or other immunosuppressive medications (within 14 days of study drug administration) Active brain metastases or leptomeningeal metastases	*Key inclusion criteria:* • ECOG performance status of 0 or 1 and would otherwise be potentially fit for docetaxel-based or methotrexate-based 2nd line chemotherapy • Histologically confirmed R/M SCCHN not amenable to local therapy with curative intent. (surgery and/or radiation therapy with or without chemotherapy.) • Tumour progression or recurrence within 6 months of last dose of platinum therapy (*as adjuvant chemotherapy; neo-adjuvant chemotherapy; concurrent with radiotherapy; or palliative chemotherapy for recurrent or metastatic disease) • Not received prior treatment with an anti-PD-1, anti-PD-L1, anti-PD-L2, anti-CD137, or anti-cytotoxic T-lymphocyte-associated antigen-4 (CTLA-4) antibody • Every effort has been made for the patient to have PD-L1 testing with an approved and validated test to determine the Tumour Proportion Score (TPS)

Trial name	CheckMate 141	SACT dataset
	Histologically confirmed R/M carcinoma of the nasopharynx, SCC of unknown primary, and salivary gland or non-squamous histologies (e.g. mucosal melanoma) Prior treatment with an anti-PD-1, anti-PD-L1, anti-PD-L2, anti-CTLA-4 antibody, or any other antibody or drug specifically targeting T-cell co-stimulation or immune checkpoint pathways	
Trial drugs and method of administration	*Nivolumab group (n=240)* Nivolumab, i.v. infusion, 3 mg/kg, Q2W Four patients randomised to the nivolumab arm did not receive ≥1 dose of study treatment. *Investigator's choice (IC) (n=121)* Patients were randomised to the IC arm and received one of the three possible therapies at the discretion of the investigator (see list below). Docetaxel (30 mg/m^2, i.v. infusion, QW) Methotrexate (40 mg/m^2, i.v. infusion, QW) Cetuximab (400 mg/m^2, i.v. infusion, once, then 250 mg/m^2, i.v., QW) Treatment in both arms was continued until progression, unacceptable toxicity, or withdrawal of consent. Patients in the nivolumab arm were permitted to continue treatment beyond investigator-assessed RECIST 1.1-defined progression if they were experiencing a clinical benefit, as determined by the investigator, and were tolerating the study drug. Dose reductions were not permitted for nivolumab but were allowed for therapies in the IC arm. Dose delays were permitted in both trial arms.	*Nivolumab only (n=556)* Nivolumab (i.v. infusion, Q2W) Dosing started as weight base (3 mg/kg) and then changed to a flat dose (240 mg) in response to the licence. Six patients did not receive treatment and 44 patients died before treatment.
Primary outcomes	*Overall survival (OS)* Patients were followed up continuously whilst on study treatment and then every 3 months until death, loss to follow-up, or	*OS* *Treatment duration (TTD)*

Trial name	CheckMate 141	SACT dataset
	withdrawal of study consent after patients discontinued study treatment.	
Secondary and other outcomes	*Secondary endpoints:* Progression-free survival (PFS) Time to discontinuation (TTD) Objective response rate (ORR) *Exploratory endpoints:* Duration of response (DOR) Time to response (TTR) Safety Patient-reported outcomes (PROs) assessed using EORTC QLQ-C30 and QLQ-H&N35 questionnaires, as well as the EQ-5D-3L questionnaire	Reason for stopping treatment ('Treatment outcomes for patients that have ended treatment')
Subgroups	A pre-planned exploratory subgroup analysis of OS by treatment group and PD-L1 expression (≥1% or <1%) was conducted.	A subgroup analysis of OS by PD-L1 expression level was conducted.
Duration of study and follow-up	The study was initiated on the 29th May 2014 with the last patient's last visit on 6th November 2015 and the clinical database locked on the 18th December 2015. At this data cut-off point, the median duration of follow-up was 5.3 months (range, 0.0–16.8) and 4.6 months (range, 0.0–15.2) in the nivolumab and IC arms, respectively.	Entry to the SACT dataset from 13 October 2017 to 12 May 2019. A snapshot of SACT data was taken on 5 October 2019 and made available for analysis on the 14 October 2019. The snapshot includes SACT activity up to the 30 June 2019. Tracing the patients' vital status was carried out on 11 October 2019 using the personal demographics service (PDS).[9] The median follow-up time was 83.5 days.

Source: CS,[2] and SACT dataset report.[9] except *provided in an e-mail from NICE.[12]

AEs = adverse events; CS = company submission; CT = computerised tomography; CTLA-4 = cytotoxic T-lymphocyte-associated protein 4; DMC = Data Monitoring Committee; DOR = duration of response; ECOG = Eastern Cooperative Oncology Group; EORTC QLQ-C30 and H&N35 = European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30 and Head and Neck 35; EQ-5D-3L = 3-level EuroQoL 5-Dimensions; HPV = human papillomavirus; HRQoL = health-related quality of life; i.v. = intravenous; IC = investigator's choice; IDMC = independent data monitoring committee; IVRS = interactive voice response system; MRI = magnetic resonance imaging; ORR = objective response rate; OS = overall survival; PD-L1 = programmed death ligand 1; PD-L2 = programmed death ligand 2; PFS = progression-free survival; PROs = patient-reported outcomes; Q2W = once every two weeks; QW = once weekly; RECIST 1.1 = Response Evaluation Criteria In Solid Tumours version 1.1; R/M = recurrent or metastatic; SCC = squamous-cell carcinoma; SCCHN = squamous-cell carcinoma of the head and neck; TTR = time to response; UK = United Kingdom; USA = United States of America

Table 3.2: Baseline characteristics of patients in the all-randomised population in CheckMate 141 by individual therapy[a]

Characteristic	Nivolumab (n=240)	IC (n=121)	Docetaxel (n=54)	Methotrexate (n=52)	Cetuximab (n=15)
Demographics					
Age, median years (range)	59.0 (29–83)	61.0 (28–78)	61.0 (28–74)	61.0 (32–78)	57.0 (39–78)
Age categorisation, n (%)					
<65	172 (71.7)	76 (62.8)	34 (63.0)	32 (61.5)	10 (66.7)
≥65 and <75	56 (23.3)	39 (32.2)	20 (37.0)	16 (30.8)	3 (20.0)
≥75	12 (5.0)	6 (5.0)	0	4 (7.7)	2 (13.3)
Male, n (%)	197 (82.1)	103 (85.1)	45 (83.3)	44 (84.6)	14 (93.3)
Race, n (%)					
White	196 (81.7)	104 (86.0)	50 (92.6)	41 (78.8)	13 (86.7)
Black/African American	10 (4.2)	3 (2.5)	0	2 (3.8)	1 (6.7)
Asian	29 (12.1)	14 (11.6)	4 (7.4)	9 (17.3)	1 (6.7)
Other	5 (2.1)	0	0	0	0
Region, n (%)					
North America	101 (42.1)	44 (36.4)	12 (22.2)	19 (36.5)	13 (86.7)
Europe	109 (45.4)	62 (51.2)	37 (68.5)	25 (48.1)	0
Rest of the world	30 (12.5)	15 (12.4)	5 (9.3)	8 (15.4)	2 (13.3)
Tobacco use, n (%)					
Current/former	191 (79.6)	85 (70.2)	40 (74.1)	35 (67.3)	10 (66.7)
Never	39 (16.3)	31 (25.6)	11 (20.4)	15 (28.8)	5 (33.3)
Unknown	10 (4.2)	5 (4.1)	3 (5.6)	2 (3.8)	0
Disease characteristics					
Site of primary tumour, n (%)[b]					
Oral cavity	108 (45.0)	67 (55.4)	29 (53.7)	31 (59.6)	7 (46.7)

Characteristic	Nivolumab (n=240)	IC (n=121)	Docetaxel (n=54)	Methotrexate (n=52)	Cetuximab (n=15)
Pharynx	92 (38.3)	36 (29.8)	19 (35.2)	11 (21.2)	6 (40.0)
Larynx	34 (14.2)	15 (12.4)	5 (9.3)	8 (15.4)	2 (13.3)
Other	6 (2.5)	3 (2.5)	1 (1.9)	2 (3.8)	0
HPV p-16 status, n (%)					
Positive	63 (26.3)	29 (24.0)	16 (29.6)	9 (17.3)	4 (26.7)
Negative	50 (20.8)	36 (29.8)	19 (35.2)	15 (28.8)	2 (13.3)
Not tested[c]	127 (52.9)	56 (46.3)	19 (35.2)	28 (53.8)	9 (60.0)
Prior therapy					
Number of lines of prior systemic cancer therapy, n (%)					
1	106 (44.2)	58 (47.9)	29 (53.7)	21 (40.4)	8 (53.3)
2	80 (33.3)	45 (37.2)	19 (35.2)	19 (36.5)	7 (46.7)
≥3	54 (22.5)	18 (14.9)	6 (11.1)	12 (23.1)	0
ECOG PS (%)					
0	49 (20.4)	23 (19.0)	Not reported		
1	189 (78.8)	94 (77.7)			
≥ 2	1 (0.4)	3 (2.5)			
Not reported	1 (0.4)	1 (0.8)			

Source: ERG report for TA490.[6]

Notes: [a]The investigator had to indicate which IC agent he or she would use if the subject were randomised to the IC arm. This information was recorded in the IVRS system prior to randomisation; [b] Each was not subcategorised to capture a more precise primary tumour site (e.g., oropharynx); [c] Baseline 'unknown' HPV status included 180 patients who were not tested (per protocol, HPV status testing was only required for patients with oropharyngeal disease), 2 patients whose sample was collected after baseline, and 1 nivolumab subject who was tested for HPV, but had a non-evaluable test result.

CSR = clinical study report; ECOG PS = Eastern Cooperative Oncology Group performance status; HPV= human papillomavirus; IC= investigator's choice; IVRS= interactive voice response system

Table 3.3: Baseline characteristics of patients in CheckMate 141 compared to the SACT data cohort study

Characteristic	CheckMate 141: Nivolumab (n = 240)	SACT data cohort study (n = 506)
Male, n (%)	197 (82)	411 (81)
Age, median years	59	62
Age categorisation, n (%)		
< 40	14 (6)	15 (3)
40 – 49	18 (8)	39 (8)
50 – 59	90 (38)	145 (29)
60 – 69	87 (36)	194 (38)
70 – 79	29 (12)	104 (21)
80 +	2 (1)	9 (2)
Performance status, n (%)		
0	49 (20)	122 (24)
1	189 (79)	286 (57)
≥ 2	1 (0.4)	33 (7)
Missing	1 (0.4)	65 (13)
PD-L1 score, n (%)		
< 1	73 (30)	55 (11)
≥ 1	88 (37)	52 (10)
Can't be quantified	79 (33)	189 (37)
Not recorded	0	210 (42)

Source: Company submission; Company response to clarification; Public Health England Data Review[2, 5, 9]
Notes: Percentages may not total 100 due to rounding.
PD-L1 = programmed death ligand 1; SACT = Systemic Anti-Cancer Therapy

3.2 Results of the new clinical evidence

3.2.1 Overall survival

An overview of OS in the previous data cut (20th September 2016**)** and new data cut (15th October 2019**)** of CheckMate 141 **and** the SACT data is provided in Table 3.4. From the table it can be seen, that as in the earlier data from CheckMate 141, there is an OS advantage to nivolumab in terms of HR (0.6858 [95% CI, 0.5483 to 0.8579; p<0.001]). The advantage is very similar, albeit slightly greater with the more mature data. Median OS was similar between the earlier and later data cuts of the CheckMate 141 data, the point estimates being identical and showing a longer survival in the nivolumab arm (7.72 months [95% CI: 5.68 to 8.77]) versus the IC arm (5.06 months [95% CI: 4.04 to 6.24]).

The later data cut of the CheckMate 141 trial provides fuller data for 24-month survival and data for 36- and 48-month survival as shown in Table 3.4. The data showed that the survival advantage of nivolumab was maintained at 36 months (10.3% [95% CI: 6.8 to 14.7] versus 2.5% [95% CI: 0.7, to 6.6) and at 48 months (8.0% [95% CI: 4.9 to 12.0] versus 1.7% [95% CI: 0.3, to 5.4).

The Kaplan-Meier (KM) plot for OS based on the latest data cut is presented in Figure 3.1. The IC and nivolumab Kaplan-Meier OS curves overlapped until approximately Month 4 and then separated, favouring nivolumab.

In terms of comparison to the SACT dataset, median OS on nivolumab is higher in CheckMate 141 than the 6.5 months (95% CI: 5.6 to 7.6) of the SACT dataset, although this is reported to be based on a median follow up of only 83.5 days.[9] However, one-year survival rates were similar between the nivolumab arm of the latest CheckMate 141 data and the SACT database (33.4%[95% CI: 27.5 to 39.5]) and 34% [95% CI 29 to 38]).

In terms of OS according to PD-L1 status, for patients with a PD-L1 < 1%, the HR was below 1 and those receiving nivolumab had a longer median survival (6.51 months [95% CI: 4.37 to 11.73]) than those in the IC group (5.45 months [95% CI: 3.68 to 8.54]) but neither of these outcomes were statistically significant. For patients with a PD-L1 ≥ 1%, the HR was lower and statistically significant and median survival was statistically significantly longer with nivolumab (8.15 months [95% CI: 6.67 to 9.53]) than with IC (4.60 months [95% CI: 3.81 to 5.78]). However, as reported in the response to clarification, the interaction between treatment and PD-L1 status in the Cox proportional hazards model was not statistically significant (p = 0.239) indicating that there was no evidence that the treatment effect differed between the different PD-L1 status subgroups.[5] The company indicate that the interpretation of analyses of the PD-L1 subgroups should be made with caution due to the smaller sample sizes (116 for PD-L1 <1% and 157 for ≥ 1%) and wider 95% CI for the HR. The OS curves according to PD-L1 status are presented in Figures 3.2 and 3.3.

.

Table 3.4: Overall survival in the all-randomised population in CheckMate 141 and SACT

Outcome[a]	CheckMate141 20th September 2016		CheckMate141 15th October 2019		SACT 11th October 2019
	Nivolumab (n=240)	IC (n=121)	Nivolumab (n=240)	IC (n=121)	Nivolumab (n=506)
Deaths, n (%)	[redacted]	[redacted]	218 (90.8)	118 (97.5)	335/506 (66.2)
Median OS, months (95% CI)	[redacted]	[redacted]	7.72 (5.68, 8.74)	5.06 (4.04, 6.24)	6.5 (5.6, 7.6)
HR for death with nivolumab (95% CI)	0.70 (97.73% CI: 0.51, 0.96)*		0.6858 (0.5483, 0.8579)		NA
1-year survival rate, % (95% CI)	[redacted]	[redacted]	33.4 (27.5, 39.5)	19.4 (12.9, 26.9)	34 (29, 38)
18-month survival rate, % (95% CI)	[redacted]	[redacted]	22.1 (17.0, 27.6)	8.4 (4.3, 14.3)	NR
24-month survival rate, % (95% CI)	[redacted]	[redacted]	16.8 (12.3, 21.9)	5.9 (2.6, 11.1)	NR
36-month survival rate, % (95% CI)	[redacted]	[redacted]	10.3 (6.8, 14.7)	2.5 (0.7, 6.6)	NR
48-month survival rate, % (95% CI)	[redacted]	[redacted]	8.0 (4.9, 12.0)	1.7 (0.3, 5.4)	NR

Source: Tables 5 and 8 CS[2] except *ERG report for TA490.[6]
HR = hazard ratio; IC = investigator choice; NA = not available

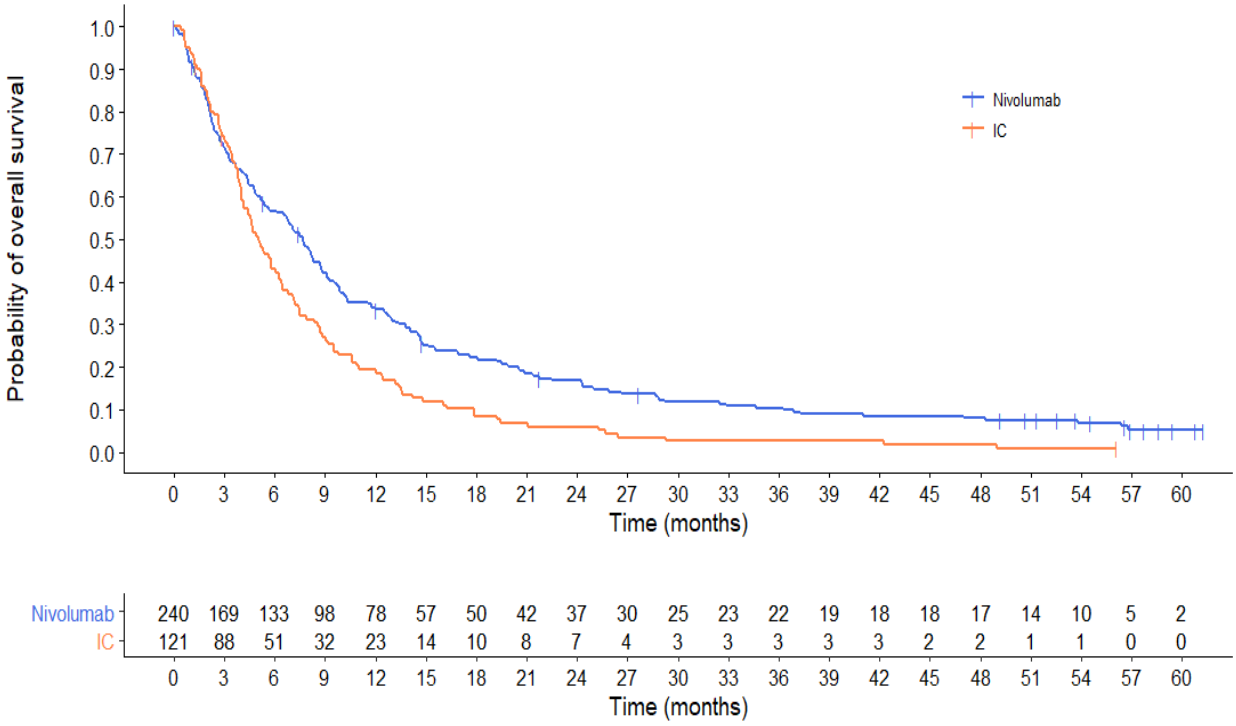
Table 3.5: Overall survival according to PD-L1 status in CheckMate 141 and SACT

Outcome[a]	CheckMate141 PD-L1 <1% 15 October 2019		CheckMate141 PD-L1 ≥1% 15 October 2019		SACT 11th October 2019
	Nivolumab (n=76)	IC (n=40)	Nivolumab (n=96)	IC (n=61)	Nivolumab (n=506)
Deaths, n (%)	72/76 (94.7)	40/40 (100)	87/96 (90.6)	60/61 (98.4)	NR
Median OS, months (95% CI)	6.51 (4.37, 11.73)	5.45 (3.68, 8.54)	8.15 (6.67, 9.53)	4.60 (3.81, 5.78)	NR
HR for death with nivolumab (95% CI; p-value)*	0.7429 (0.5015, 1.101; p=0.138)		0.5397 (0.3857, 0.7554; p<0.001)		NR

Source: Tables 8, 9, of the CS and Table 5 of the CS appendix.[2, 13]

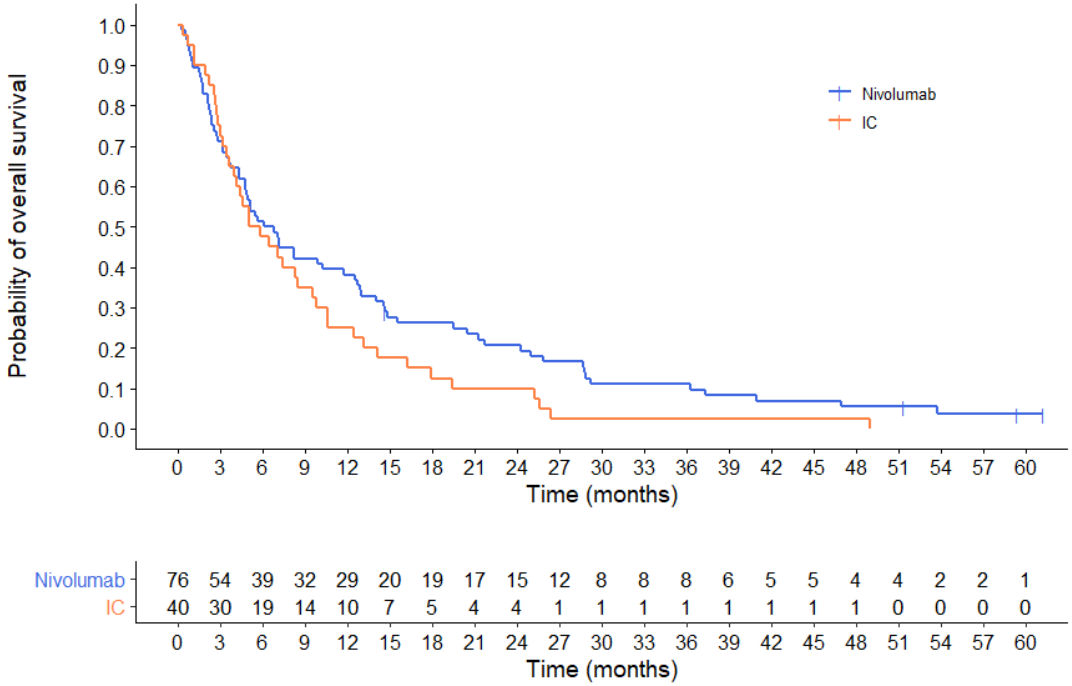
* Computed using unstratified Cox proportional hazards model with treatment group as the sole covariate.

Figure 3.1: Kaplan-Meier plot for overall survival in CheckMate 141



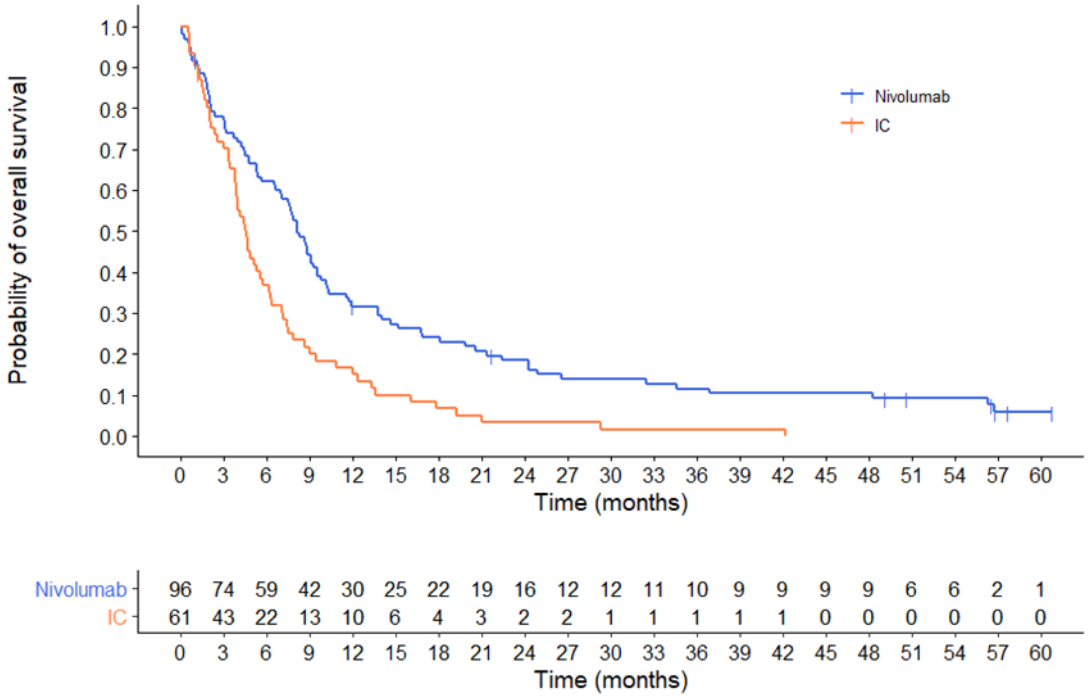
Data cut-off: 15 October 2019
Abbreviations: IC: investigator's choice.
Source: Company submission, Figure 1.[2]

Figure 3.2: Kaplan-Meier plot for overall survival for patients with the PD-L1 <1% in CheckMate 141



CheckMate 141 data cut-off: 15 October 2019
Abbreviations: CI: confidence interval; PD-L1: programmed death ligand 1.
Source: Company submission, Figure 5.[2]

Figure 3.3: Kaplan-Meier plot for overall survival for patients with the PD-L1 ≥1% in CheckMate 141



CheckMate 141 data cut-off: 15 October 2019
Abbreviations: CI: confidence interval; PD-L1: programmed death ligand 1.
Source: Company submission, Figure 6.[2]

ERG comment:

- The committee had specific concerns about the OS benefit beyond two years and expected to see further evidence. In relation to this the ERG noted that the company presented data from the 15 October 2019 data cut which had a minimum follow up of 48.2 months. Results presented above showed that the survival advantage of nivolumab over IC was maintained in terms of HR and median survival and continued through 36 months and at 48 months.
- However, the ToE stipulated that docetaxel should be the comparator and so the ERG requested in the clarification letter for analyses in the docetaxel subgroup to be presented, the response to which was not to provide these (see Section 2.2).[5]
- Although patients in the SACT data set had a lower median survival (6.5 vs. 7.7 months) than those in the nivolumab arm of Check Mate 141, it is important to note that this was based on a much shorter median follow-up of 83.5 days and the 95% CIs overlapped. Also, one-year survival rates were very similar (34% and 33.4%).
- The committee also had concerns regarding the evidence of the benefit of nivolumab for those with PD-L1 expression < 1%. CheckMate 141 was not powered to detect differences in benefit according to PD-L1 status. However, the company presented data according to PD-L1 status as requested by the committee based on the updated 15 October 2019 data cut providing four-year results. This showed that patients with a PD-L1 < 1% had a reduced hazard of death on nivolumab compared with IC but this was not statistically significant. For patients with a PD-L1 ≥ 1% the hazard of death was significantly reduced with nivolumab. However, there was no significant evidence of a treatment and subgroup interaction (p = 0.239) and these results should be considered with caution due to the reduced sample sizes and wider confidence intervals.

3.2.2 Progression-free survival

An overview of PFS in the previous data cut (20 September 2016**)** and new data cut (15 October 2019**)** of CheckMate 141 is presented in Table 3.6 From the table it can be seen, that, as for OS, there was little change with the HR of 0.82 (0.65, 1.02; p=0.0766) showing a slightly greater advantage for nivolumab than previously. [redacted], showing a shorter median PFS in the nivolumab arm than in the IC arm (2.04 months [95% CI:1.91 to 2.14] versus 2.33 months [1.94, 3.06]).

As explained by the company, there was delayed separation of the Kaplan-Meier curves using the CheckMate 141 data (see Figure 3.4) which showed that by six months the estimated PFS rate was higher in the nivolumab arm than the IC arm.(20.4% [95% CI:15.4 to 26.0] versus 10.2% [95% CI: 5.2 to 17.2]).

Progression-free survival data was not required to be collected in the SACT data set.

In terms of PFS according to PD-L1 status, the PD-L1 <1% group receiving nivolumab had a shorter median PFS than those in the IC arm (1.95 months [95% CI: 1.87 to 2.14 versus 2.68 months [95% CI: 1.97, 4.63]) (see Table 3.7). The PD-L1 ≥ 1% group receiving nivolumab had a numerically longer PFS than those in the IC arm (2.14 months [95% CI: 1.97 to 3.45 versus 1.97 months [95% CI: 1.84, 3.06]). The ERG requested the results of analyses including an interaction term between treatment and PD-L1 status. The company response to clarification showed that there was weak evidence of an interaction (p=0.077) indicating that the treatment effect of nivolumab differed between the groups based on PD-L1 status, although the HRs were not reported.[5]

Table 3.6: Progression Free Survival in the all-randomised population in CheckMate 141

Outcome[a]	CheckMate141 20 September 2016		CheckMate141 15 October 2019	
	Nivolumab (n=240)	IC (n=121)	Nivolumab (n=240)	IC (n=121)
Events, n (%)	[redacted]	[redacted]	214 (89.2)	104 (86.0)
Median PFS, months (95% CI)	[redacted]	[redacted]	2.04 (1.91, 2.14)	2.33 (1.94, 3.06)
HR for progression or death with nivolumab (95% CI; p-value)	[redacted]		0.82 (0.65, 1.02; p=0.0766)	
6-month PFS rate, % (95% CI)	[redacted]	[redacted]	20.4 (15.4, 26.0)	10.2 (5.2, 17.2)
1-year PFS rate, % (95% CI)	[redacted]	[redacted]	9.5 (6.0, 14.0)	2.6 (0.5, 8.0)
18-month PFS rate, % (95% CI)	[redacted]	[redacted]	8.5 (5.2, 12.8)	NA
24-month PFS rate, % (95% CI)	[redacted]	[redacted]	7.5 (4.5, 11.7)	NA

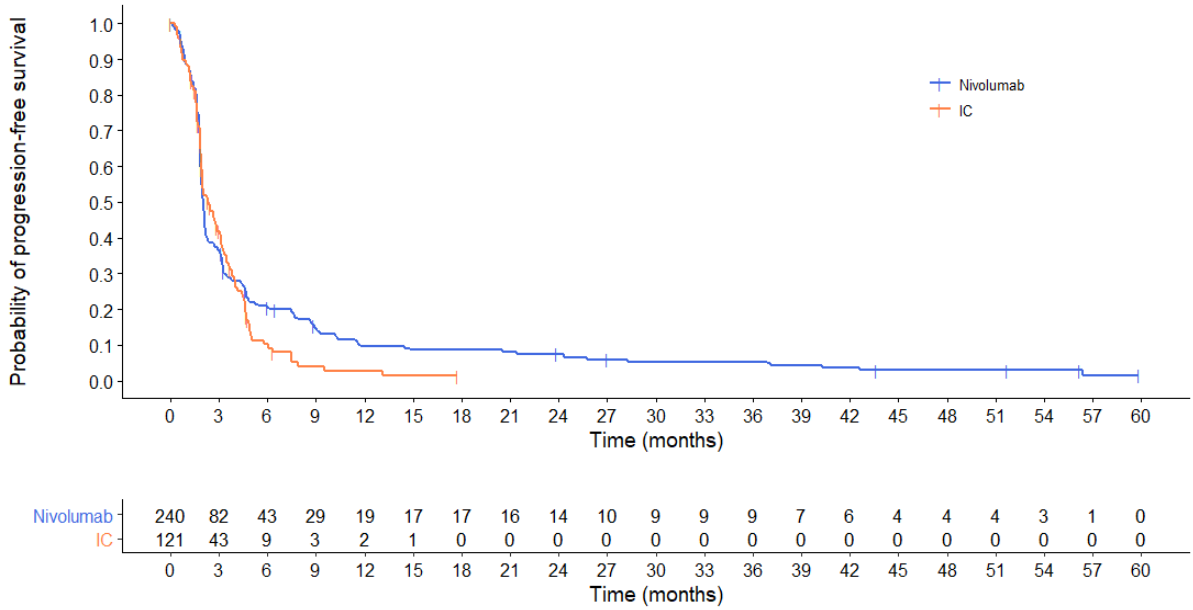
Source: Table 6 CS;[2] 1.1 addendum ERG report.[6]
HR = hazard ratio; IC = investigator choice; NA = not assessed; PFS = progression-free survival

Table 3.7: Progression Free Survival by PD-L1 status

Outcome[a]	CheckMate141 PD-L1 < 1%		CheckMate141 PD-L1 ≥ 1%	
	Nivolumab (n=240)	IC (n=121)	Nivolumab (n=240)	IC (n=121)
Events, n (%)	69/76 (90.8)	36/40 (90.0)	88/96 (91.7)	54/61 (88.5)
Median PFS, months (95% CI)	1.95 (1.87, 2.14)	2.68 (1.97, 4.63)	2.14 (1.97, 3.45)	1.97 (1.84, 3.06)
HR	NR		NR	

Source: Table 10 CS.[2]
HR = hazard ratio; IC = investigator choice; NA = not assessed; PFS = progression-free survival

Figure 3.4: Kaplan-Meier plot for progression-free survival in the all-randomised population in CheckMate 141



Data cut-off: 15 October 2019
Abbreviations: IC: investigator's choice.
Source: Company submission, Figure 2.[2]

ERG comment:

- Concerns about PFS were not specifically mentioned in the ToE and PFS data were not required to be collected in the SACT dataset. However the company provided the up to date data (15 October 2019) from CheckMate 141 on PFS and the ERG can confirm that there was no fundamental change in the conclusion that the PFS advantage to nivolumab versus IC in terms of HR, although not statistically significant, was maintained and the advantage to IC in terms of median survival, although small, was also maintained.
- The company was not explicitly required to present data by PD-L1 status for PFS and as stated before CheckMate 141 was not powered to detect differences by PD-L1 status. HRs were not provided for PFS for the PD-L1 subgroups, but the median PFS estimates indicated that there were no significant differences in PFS between nivolumab and IC in patients with PD-L1 <1% or those with PD-L1 ≥ 1%.

3.2.3 Time to treatment discontinuation

The latest CheckMate 141 data cut provides data from a minimum follow-up of 48.2 months (representing 36.8 additional months of follow-up). At the time of this data cut-off, the company stated that 13 patients in the nivolumab arm and one patient in the IC arm were still alive and in follow-up, with [redacted] still on treatment. Median TTD was similar between the CheckMate 141 earlier data cut and the later data cut. It was also similar between nivolumab and IC arms in the trial [redacted] (Table 3.8). The company showed in Kaplan-Meier curves that there was separation of the curves favouring nivolumab from approximately [redacted] months.

The SACT data showed a longer median TTD of 3.0 months (95% CI: 2.7 to 3.3) with no overlap in the 95% CIs. The company also noted that at six months 28% of SACT patients were still receiving treatment as opposed to [redacted]% of the CheckMate 141 patients and at 12 months 17% of patients in the SACT database were still receiving treatment as opposed to [redacted]% of the CheckMate 141 patients.

For PD-L1 < 1% median TTD in CheckMate 141 was virtually identical between treatment groups and similar to the overall result at [redacted] for nivolumab versus [redacted] for IC (Table 3.9). In the PD-L1 ≥ 1 group the median TTD was higher in the nivolumab group than the IC group of CheckMate 141 at [redacted] The response to clarification showed that there was a statistically significant interaction (p=0.0208) in the Cox proportional hazards model between treatment and PD-L1 subgroup indicating that the treatment effect was different in patients with PD-L1 < 1% compared to ≥ 1%, although the HRs were not reported.[5]

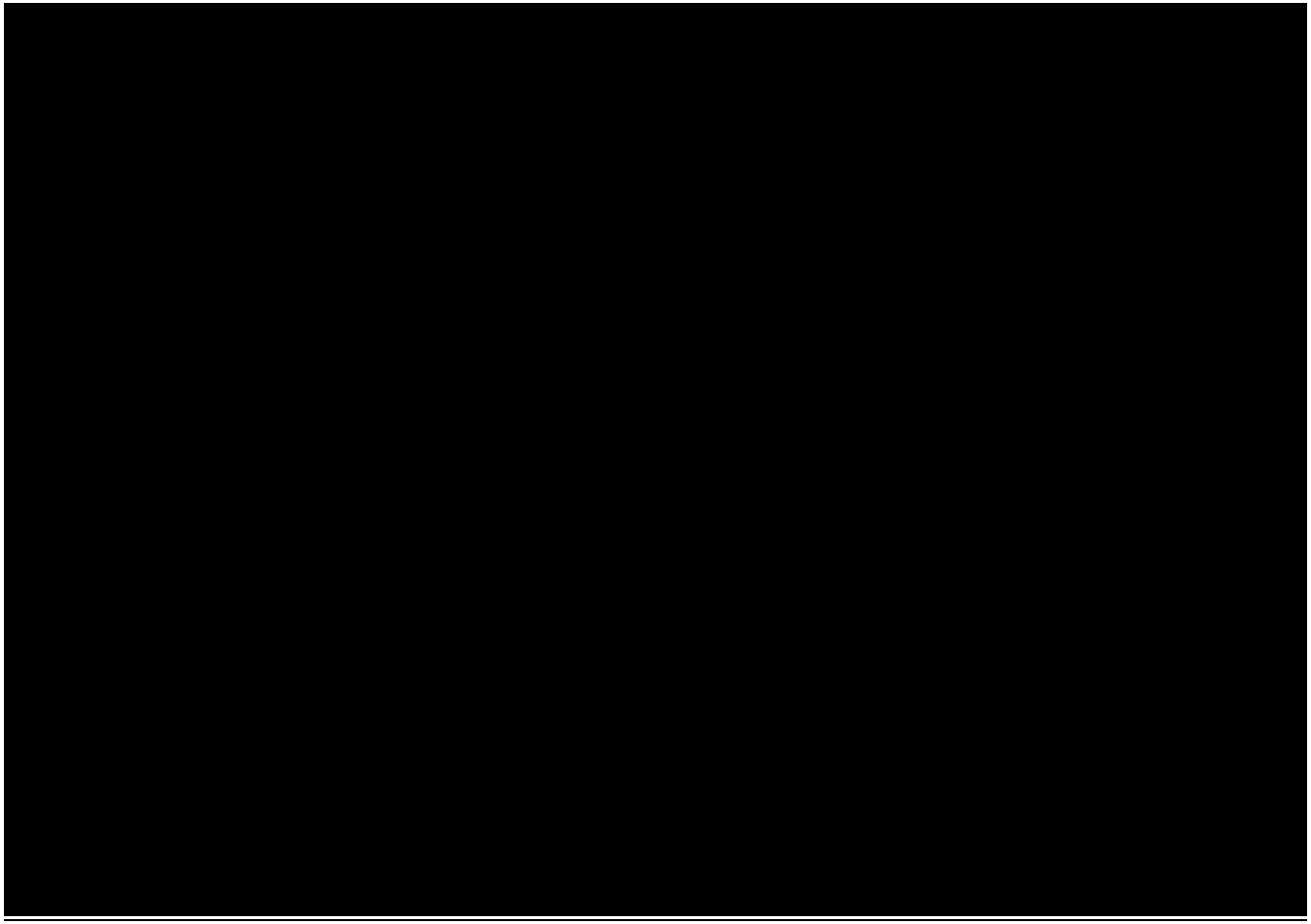
Table 3.8: Time to treatment discontinuation in CheckMate 141 and SACT

Outcome[a]	CheckMate141 20 September 2016		CheckMate141 15 October 2019		SACT 11 October 2019
	Nivolumab (n=240)	IC (n=121)	Nivolumab (n=240)	IC (n=121)	Nivolumab (n=506)
Events, n (%)	[redacted]	[redacted]	[redacted]	[redacted]	394/506
Median TTD, months (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]	3.0 (2.7, 3.3)
Source: Tables 7 CS; Addendum to ERG report.[2, 6]					

Table 3.9: Time to treatment discontinuation by PD-L1 status in CheckMate 141 and SACT

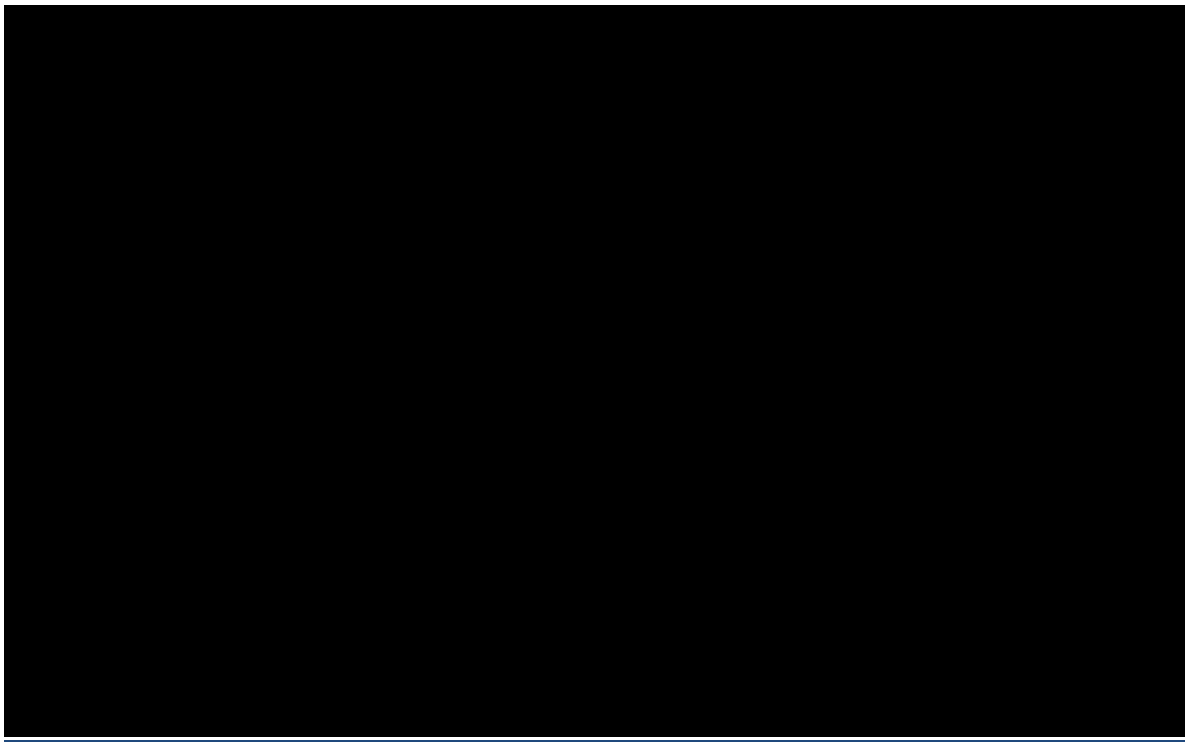
Outcome[a]	CheckMate141 PD-L1 <1%		CheckMate141 PD-L1 ≥1%		SACT 11 October 2019
	Nivolumab (n=73)	IC (n=38)	Nivolumab (n=88)	IC (n=61)	Nivolumab (n=506)
Events, n (%)	[redacted]	[redacted]	[redacted]	[redacted]	NR
Median TTD, months (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]	NR
Source: Table 11 of the CS.[2]					

Figure 3.5: Kaplan-Meier comparing time to discontinuation in CheckMate 141 and the SACT database



CheckMate 141 data cut-off: 15 October 2019
Abbreviations: SACT: Systemic Anti-Cancer Therapy.
Source: Company submission, Figure 12;[2] Public Health England report[9]

Figure 3.6: Kaplan-Meier plot for time to treatment discontinuation for patients with the PD-L1 <1% in CheckMate 141



CheckMate 141 data cut-off: 15 October 2019
Abbreviations: CI: confidence interval; PD-L1: programmed death ligand 1.
Source: Company submission, Figure 9.[2]

Figure 3.7: Kaplan-Meier plot for time to treatment discontinuation for patients with the PD-L1 ≥1% in CheckMate 141

CheckMate 141 data cut-off: 15 October 2019
Abbreviations: CI: confidence interval; PD-L1: programmed death ligand 1.
Source: Company submission, Figure 10.[2]

ERG comment

- Concerns about TTD were not specifically mentioned in the ToE. However, the company provided the up to date data (15 October 2019) from CheckMate 141 on TTD which the ERG has presented above and the ERG noted that median TTD was similar between the earlier and later data cuts of the CheckMate 141 data.
- However, [redacted], the median TTD was shorter than in the SACT data (three months). It is unclear to the ERG why this was and what the implications for generalisability of the effectiveness of nivolumab in terms of OS or PFS might be. OS seemed to be slightly shorter in the SACT dataset, although this was very uncertain. It might seem to indicate that more drug needed to be given to obtain the same OS, but this is unclear.
- The company was not explicitly required to present data by PD-L1 status for TTD and, as stated before, CheckMate 141 was not powered to detect differences by PD-L1 status. [redacted]

3.2.4 Health-related quality of life

The committee had requested an exploration of the most appropriate utility values in the light of new evidence. However, the company used the EQ-5D data from the 20 September 2016 data cut of the CheckMate 141 trial to analyse how utility might change over time and how utility might change with respect to how close patients were from death. Details of the generation of the utility values and a discussion of their appropriateness can be found in the cost-effectiveness section of this report.

3.2.5 Adverse effects of treatment

No specific requirements were asked of the company regarding an update of AE data and the SACT study did not collect such data either. For completeness of reporting the ERG asked the company to provide AE data from the 15 October 2019 data cut as per the original submission. Table 3.10 provides a high-level summary, which compares the new with the September 2016 data cut.

Table 3.10: Summary of adverse events from CheckMate 141

Adverse event, n (%)	Nivolumab (n=236) September 2016		IC (n=111) September 2016		Nivolumab (n=236) 15 October 2019		IC (n=111) 15 October 2019	
	Any grade	Grade 3-4	Any grade	Grade 3-4	Any grade	Grade 3-4	Any grade	Grade 3-4
All causality AEs	232 (98.3)	113 (47.9)	109 (98.2)	69 (62.2)	232 (98.3)	117 (49.6)	109 (98.2)	70 (63.1)
Drug-related AEs	146 (61.9)	36 (15.3)	88 (79.3)	40 (36.0)	146 (61.9)	37 (15.7)	88 (79.3)	41 (36.9)

Source: Company response to clarification.[5]
AEs = adverse events; CS = company submission; IC = investigator's choice

The most frequently reported grade 3-4 AEs in the nivolumab arm were also reported in the response to clarification, which the ERG can confirm were those found to be most common during TA490.[5, 14] These are (15 October 2019 vs. September 2016 data cuts):

- Anaemia: (17, 7.2%) vs. (15, 6.4%),
- dyspnoea (13, 5.5%) vs. (13, 5.5%),
- hyponatraemia (13, 5.5%) vs. (11, 4.7%),
- pneumonia (12, 5.1%) vs. (11, 4.7%) and
- malignant neoplasm progression (11, 4.7%) vs. (11, 4.7%)

ERG comment: It appears that little has changed in both the number and percentage of AEs between TA490 and the CDF review, which leads to the same conclusion as found by the ERG in TA490, i.e. nivolumab was generally well tolerated by patients in CheckMate 141 compared to IC of therapy, with a lower proportion of patients receiving nivolumab experiencing Grade 3-4 all-causality AEs.

3.3 Summary of the new clinical effectiveness evidence according to the terms of engagement for the CDF review

The ToE stated that OS from CheckMate 141 data was to be updated. The ERG can confirm that this has been done with the latest data cut being 15 October 2019, i.e. four years follow-up. The results show that the survival advantage of nivolumab over IC was maintained in terms of HR and median survival and continued through 36 months and at 48 months. Also, the company provided the up to date data from CheckMate 141 on PFS and the ERG can confirm that the numerical advantage to nivolumab versus IC was maintained. Although the ToE did not specify an update in terms of safety, the ERG asked the company to provide up to date AE data and, according to the clarification letter response, it appears that little has changed in both the number and percentage of AEs between TA490 and the CDF review, which leads to the same conclusion as found by the ERG in TA490, i.e. nivolumab was generally well tolerated by patients in CheckMate 141 compared to IC of therapy, with a lower proportion of patients receiving nivolumab experiencing Grade 3-4 all-causality AEs.

However, given that the committee also concluded that the comparator should be docetaxel, the ERG considers that the most appropriate evidence of effectiveness and safety versus docetaxel is that from the docetaxel subgroup, which the company did not provide in either the CS or in response to the clarification letter (see Section 2.2). The ERG considers that this is a major source of uncertainty that can be resolved by the company.

The SACT dataset, created as a result of the ToE, permits to some degree a test of the generalisability of the outcomes observed in the CheckMate 141 trial in the nivolumab arm to UK clinical practice, at least in terms of the outcomes that were analysed from it, i.e. OS and TTD. Indeed, a comparison reveals that UK patients might be slightly older and a small number will have a worse performance status than the patients in the all-randomised CheckMate 141 trial, which might suggest that UK patients do slightly worse than patients in the CheckMate 141 trial. However, although patients in the SACT data set had a numerically lower median survival than those in the nivolumab arm of Check Mate 141, it is important to remember that this was based on a much shorter median follow-up and the 95% CIs overlapped. Also, one-year survival was very similar. As mentioned with regards to the comparison between nivolumab and docetaxel, it could be argued that the nivolumab arm of the docetaxel subgroup of CheckMate 141 should be used to compare with the SACT dataset. On the other hand, the all-randomised population might be closer to those patients who would be treated with nivolumab in UK clinical practice. This was the judgment of the committee, who concluded that the CheckMate 141 results (implying the all-randomised population) are relevant to the population of interest, i.e. adults

with recurrent or metastatic SCCHN that progressed within six months of platinum-based therapy, in either the early or locally advanced disease stage. Although there is a discrepancy between descriptions of eligibility criteria, those for the SACT dataset could also be regarded as essentially the same as those for CheckMate 141. However, clearly not everyone in CheckMate 141 was found to be eligible for docetaxel according to the IC design. In particular, some were chosen to receive methotrexate, which in the ToE states that it is only for patients who are not fit to have a taxane. The implication of this should be that the population specified for this CDF review and in the SACT dataset should not be aligned with the all-randomised population, but should at least exclude those who would be ineligible for docetaxel. This apparent mismatch between the population and the comparator specified in the ToE does produce some uncertainty in the generalisability of the CheckMate 141, which might be reduced by a comparison of the baseline characteristics and OS in the nivolumab arm of the docetaxel subgroup and the SACT dataset.

In terms of PD-L1 status, nivolumab showed an advantage in terms of OS in comparison to IC for both groups, but it was larger for those with PD-L1 ≥ 1% and only statistically significant in terms of the HR for this subgroup. However, there was no significant evidence of a treatment and subgroup interaction (p=0.239) and these results should be considered with caution due to the reduced sample sizes and wider confidence intervals. For PFS, HRs were not provided for the PD-L1 subgroups, but the median PFS estimates indicate that there were no significant differences in PFS between nivolumab and IC in patients with PD-L1 <1% or those with PD-L1 ≥ 1%. There was also evidence of only a weak interaction effect.

[redacted]

4. COST EFFECTIVENESS

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

4.1.1 Model structure

The model structure was unchanged from the TA490 CS and consisted of a cohort-based partitioned survival model with three mutually exclusive health states: progression-free (PF), progressed disease (PD) and death.[2, 6] Disease progression was defined by Response Evaluation Criteria in Solid Tumors version 1.1, which was also used in the CheckMate 141 trial. Moreover, TTD was incorporated while allowing treatment continuation after progression in both treatment arms.

Costs and disutilities associated with AEs were estimated per episode and applied only once, at the beginning of the first cycle. This was based on the proportion of patients in each treatment arm experiencing each AE. A four week cycle length was used. The model was programmed in Excel.

ERG comment: According to the ToE for CDF review, the company's model structure is suitable for decision making and it was anticipated that the model structure would not change for the CDF review.[1] Moreover, in its original ERG report (for TA490), the ERG stated that "The model structure is similar to other oncology assessments and seems appropriate for the current decision problem".[6]

4.1.2 Population

The cost effectiveness analysis considers patients with R/M SCCHN who have progressed within six months after platinum-based therapy. The company states this is consistent with the study population of the CheckMate 141 trial, because this population underpins the marketing authorisation and is a distinct subset of the population whose disease has progressed after platinum-based therapy.

In the ToE, the committee further concluded that there was evidence of nivolumab's benefit in patients with a PD-L1 expression of 1% or more, but that the benefit was less convincing for those with a PD-L1 expression of less than 1%.[1] As a consequence, the committee expected the updated OS evidence from Checkmate 141 to include analyses by PD-L1 expression. The company provided additional subgroup analyses according to PD-L1 expression level.[2]

ERG comment: The focus on the study population of the CheckMate 141 trial is consistent with the committee preferences stating that the committee concluded that although there are some differences between the trial population and the UK population, the CheckMate 141 results are relevant to the UK population.

4.1.3 Interventions and comparators

As described in Section 2.2, since the original submission for TA490, the licensed dose of nivolumab has been updated to a flat dose of 240 mg every two weeks (Q2W), rather than the weight-based dose used in the CheckMate 141 trial (3 mg/kg every two weeks). The recommended dosage of nivolumab in this indication is 240 mg flat dose every two weeks. The licence also specifies that nivolumab treatment should be continued until treatment is no longer tolerated or clinical benefit is no longer observed. This latter aspect of anticipated use with nivolumab is reflected through the use of the TTD curve to model time on treatment instead of the PFS curve.

According to the company, in the UK, treatment in the platinum-refractory setting would most likely be with a taxane (docetaxel or paclitaxel), or methotrexate if a taxane was clinically inappropriate due to tolerability issues or prior taxane therapy.[2] Single-agent docetaxel is predominantly used in UK clinical practice, although paclitaxel may also be used for patients who are not fit enough to receive

treatment with docetaxel and have not received prior taxane therapy.[6] However, as stated in Section 2.2, the ToE specifies docetaxel as the main comparator of interest. In the cost effectiveness model, it is assumed that docetaxel is administrated at a dose of 75mg/m^2 every three weeks.

ERG comment: Based on the available evidence, it seems reasonable to assume docetaxel (75mg/m^2) and docetaxel (30 mg/m^2 as in IC of checkmate trial) are equally effective. It is however questionable whether the nivolumab flat dose can be assumed equally effective to weight-based nivolumab (see section 2.2) and whether the effectiveness of docetaxel, the main comparator according to the ToE, is equally effective as the IC from CheckMate 141 (see section 4.1.5).

4.1.4 Perspective, time horizon and discounting

The analysis was conducted from the perspective of the NHS and PSS in England and Wales over a time horizon of 20 years. Costs and outcomes were discounted by 3.5%.

ERG comment: This is in line with the NICE reference case.

4.1.5 Treatment effectiveness and extrapolation

Multiple parametric time-to-event models were used to estimate:

- OS;
- PFS and;
- TTD.

These were estimated based on the nivolumab arm and the investigator's choice (IC) arm of the CheckMate 141 trial (data cut-off: October 15 2019). The IC arm did include treatment with docetaxel, methotrexate and cetuximab. The estimated OS, PFS and TTD based on the IC arm were assumed by the company to be applicable to docetaxel, methotrexate and paclitaxel.

The following parametric survival distributions were examined using goodness-of-fit statistics and visual inspection:

- Exponential
- Weibull
- Gamma
- Gompertz
- Log-normal
- Log-logistic
- Generalised-gamma
- Spline models (using 1- and 2-knots)

In addition to the standard parametric and spline models, the company did also explore piecewise models to estimate OS and PFS. This was consistent with the ToE indicating that a piecewise model is expected to be used to extrapolate OS.[1] The piecewise models consisted of the Kaplan-Meier curves up to a specific cut-off, followed by extrapolation for OS using Exponential (cut-offs: 20, 28, 36, 48, 96 weeks) or Log-normal (cut-offs: 20, 36, 48, 96 weeks) distributions while for PFS the piecewise models were extrapolated using Exponential (cut-offs: 12, 16, 20, 28 weeks) or Weibull (cut-off: 12 weeks) distributions.

For OS the proportional hazards assumption did not hold (CS Figure 13; non-parallel lines that cross/overlap), for PFS and TTD this is unclear for the new data-cut. It should however be noted that

the proportional hazards assumption did not hold for PFS and TTD in the original submission (i.e. based on the September 2016 data-cut). The company estimated all parametric time-to-event models independently for nivolumab and IC. The goodness-of-fit statistics for the parametric time-to-event models are presented in Table 4.1. In this table, the lowest AIC/BIC is printed in bold.

Selection of model for overall survival

To select the piecewise model for OS, the visual fit to the Kaplan-Meier curves was considered by the company. Based on this visual assessment, the company considered that the piecewise log-normal distribution provided a better fit than the Exponential distribution and selected the 96-week cut-off point to maximise the use of the observed data (Figures 4.1 and 4.2). Additionally, the company considered the standard parametric survival models to provide plausible alternative models to estimate OS, particularly the log-normal (had the best goodness-of-fit statistics) and log-logistic distributions were considered to be plausible candidates (Figure 4.3).

Long-term waning of overall survival treatment effect

The company preferred to assume no treatment waning, given the maturity of the CheckMate 141 trial data (compared with the September 2016 data cut-off) and since the log cumulative hazard plot for OS indicated diverging curves towards the end of the follow-up period (Figure 4.2). The company stated that, if this trend would continue, the assumption of treatment waning at five-year is not valid.

Selection of model for progression free survival

As per TA490, the company selected the generalised gamma model for estimating PFS as this distribution had a reasonable visual fit, had one of the best statistical fit (when excluding spline models) and did not result in logical inconsistencies (i.e. that PFS was predicted to be higher than OS). The spline models provided a better statistical fit for nivolumab than the standard parametric models, but the best fitting curves often produced logical inconsistencies. Excluding the spline models, the log-normal and log-logistic models provided the best statistical fit for IC but were associated with a poor visual fit to the observed data for nivolumab in the long term. See Figure 4.4 for the visual fit to the Kaplan-Meier curves.

Selection of model for time to treatment discontinuation

For nivolumab, the two-spline normal model provided the best statistical fit and a reasonable visual fit to the observed data, and was thus considered to be more plausible for extrapolation of TTD than the generalised gamma model used in TA490. The two-spline model also predicted a reasonable estimate of mean TTD when compared to PFS (i.e. mean TTD and mean PFS were similar). [redacted] See Figure 4.5 for the visual fit to the Kaplan-Meier curves.

Table 4.1: Summary of goodness-of-fit data (all-randomised population)

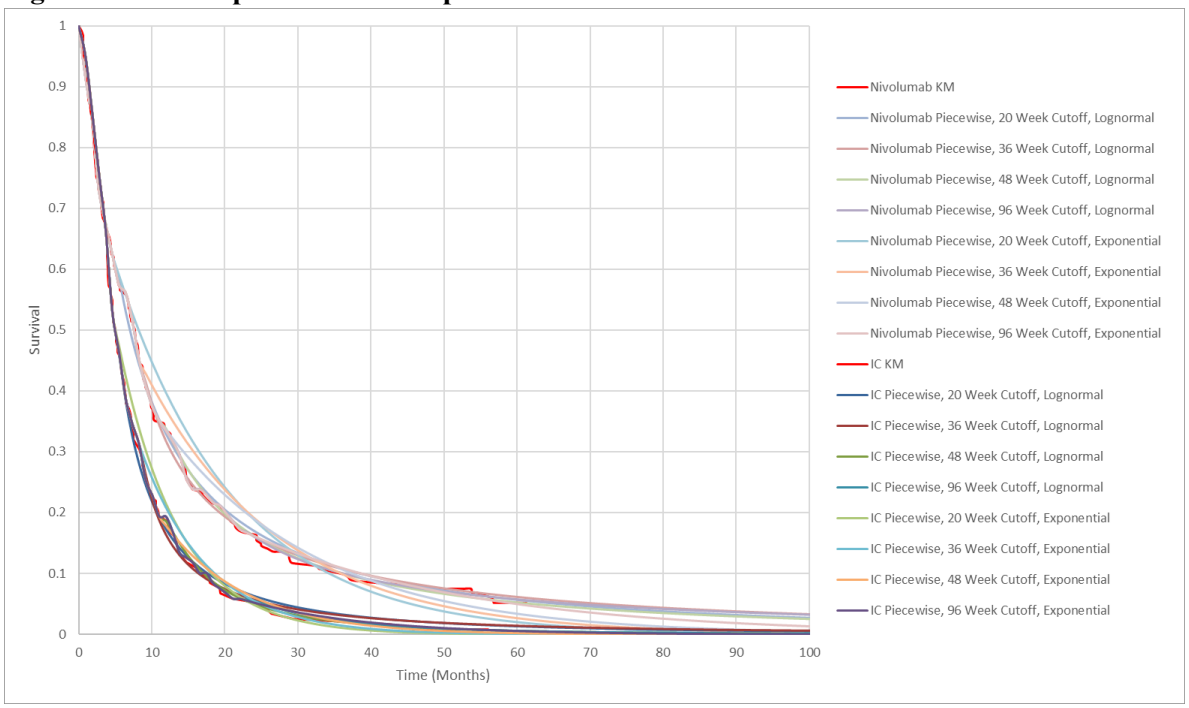
	OS		PFS		TTD	
Distribution	**AIC**	**BIC**	**AIC**	**BIC**	**AIC**	**BIC**
Nivolumab						
Exponential	1576.347	1579.828	1189.575	1193.056	1239.736	1243.200
Weibull	1564.828	1571.789	1164.921	1171.882	1183.841	1190.768
Gamma	1571.444	1578.406	1184.336	1191.298	1202.061	1208.988
Gompertz	1546.749	1553.711	1106.591	1113.552	1164.232	1171.159
Log-normal	**1540.163**	**1547.124**	1073.288	1080.249	1182.226	1189.154
Log-logistic	1542.166	1549.127	1054.897	1061.858	1160.668	1167.596
Generalised-gamma	1542.155	1552.597	1051.098	1061.540	1171.362	1181.753
Spline models:						
1-Spline Hazard	1544.033	1554.475	1034.038	1044.480	1167.889	1178.281
2-Spline Hazard	1545.414	1559.337	1031.208	1045.130	1152.755	1166.611
1-Spline Odds	1544.082	1554.524	**1021.233**	**1031.675**	1155.359	1165.751
2-Spline Odds	1543.426	1557.349	1022.361	1036.283	1148.706	1162.561
1-Spline Normal	1542.105	1552.547	1038.624	1049.066	1166.073	1176.464
2-Spline Normal	1544.113	1558.036	1027.264	1041.187	**1147.494**	**1161.349**
IC						
Exponential	729.503	732.298	460.787	463.583	419.022	421.732
Weibull	730.838	736.430	446.402	451.994	418.167	423.587
Gamma	728.217	733.809	438.978	444.570	419.407	424.826
Gompertz	729.083	734.674	461.184	466.775	418.815	424.234
Log-normal	**713.309**	**718.901**	433.239	438.830	458.579	463.998
Log-logistic	713.485	719.077	**430.911**	**436.502**	439.908	445.327
Generalised-gamma	715.275	723.662	434.690	443.077	419.038	427.167
Spline models:						
1-Spline Hazard	715.287	723.674	434.421	442.808	416.997	425.126
2-Spline Hazard	717.127	728.310	435.534	446.717	**411.662**	422.500
1-Spline Odds	715.426	723.814	432.689	441.076	413.240	**421.369**
2-Spline Odds	717.326	728.509	434.637	445.820	414.945	425.784
1-Spline Normal	715.207	723.594	434.211	442.599	413.987	422.115
2-Spline Normal	716.381	727.565	434.917	446.100	434.917	445.755

Source: Based on CS Appendix B and the economic model

Note: the lowest AIC/BIC is printed in bold.

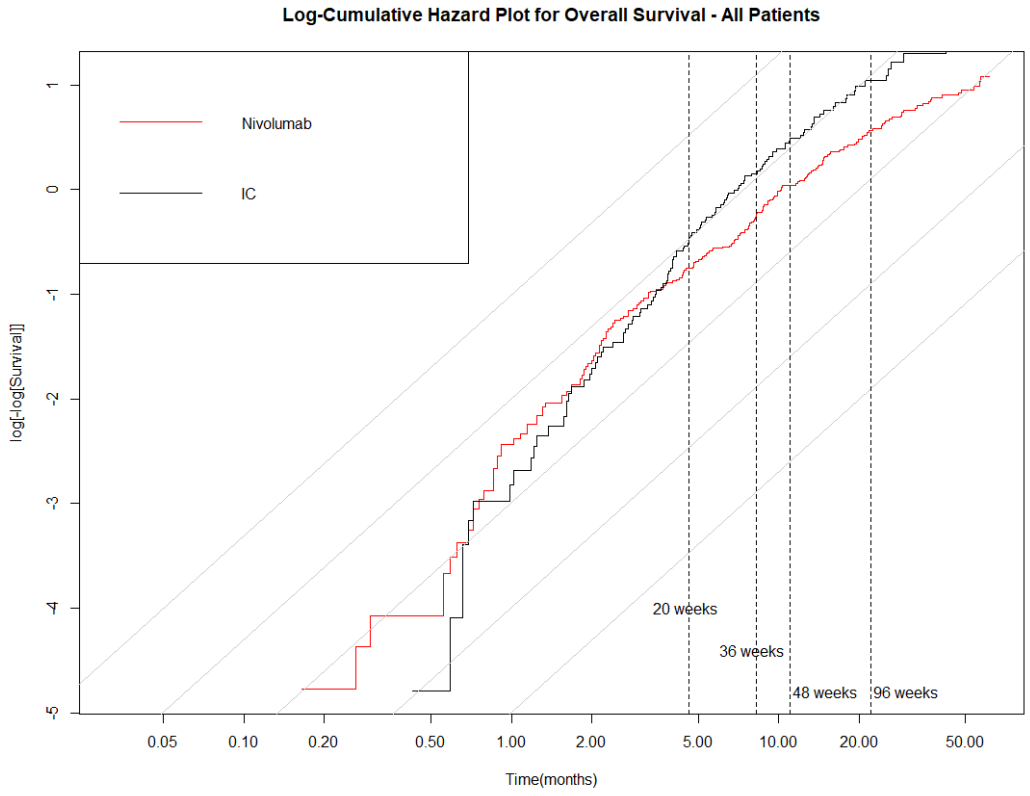
AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; CS = company submission; IC = investigator's choice; OS = overall survival; PFS = progression-free survival; TTD = time to treatment discontinuation

Figure 4.1: OS Kaplan-Meier with piecewise models



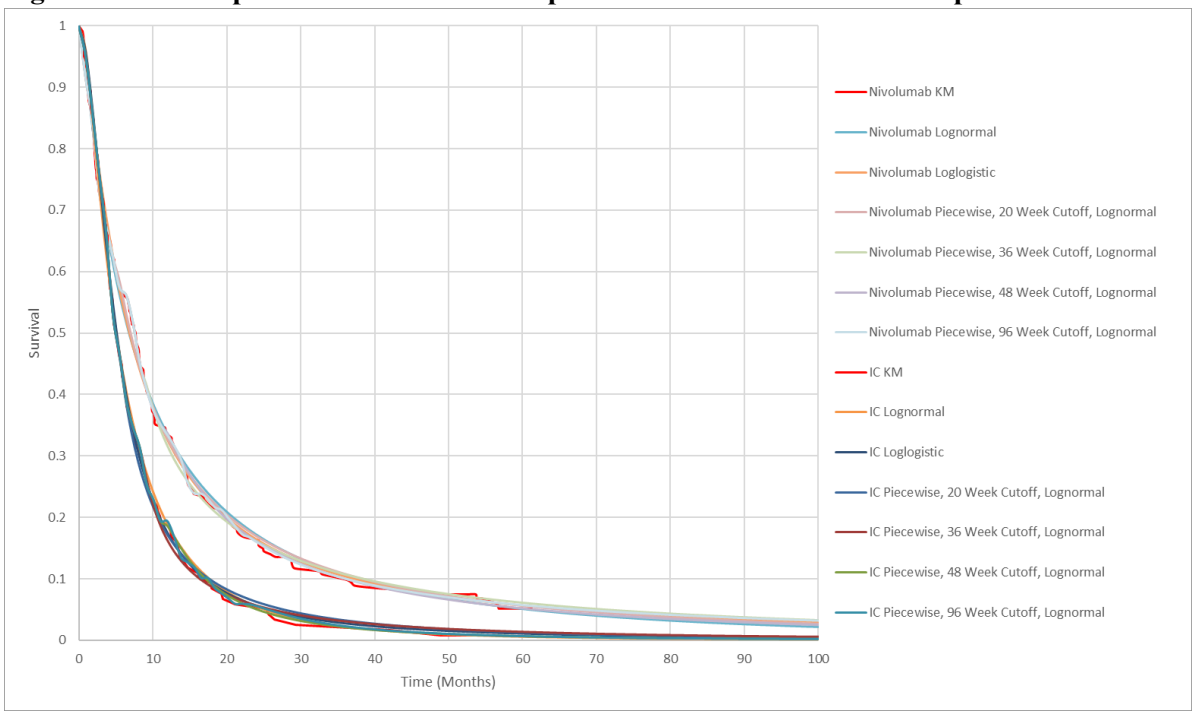
Source: CS Figure 14[2]

Figure 4.2: Log cumulative hazard plot for overall survival



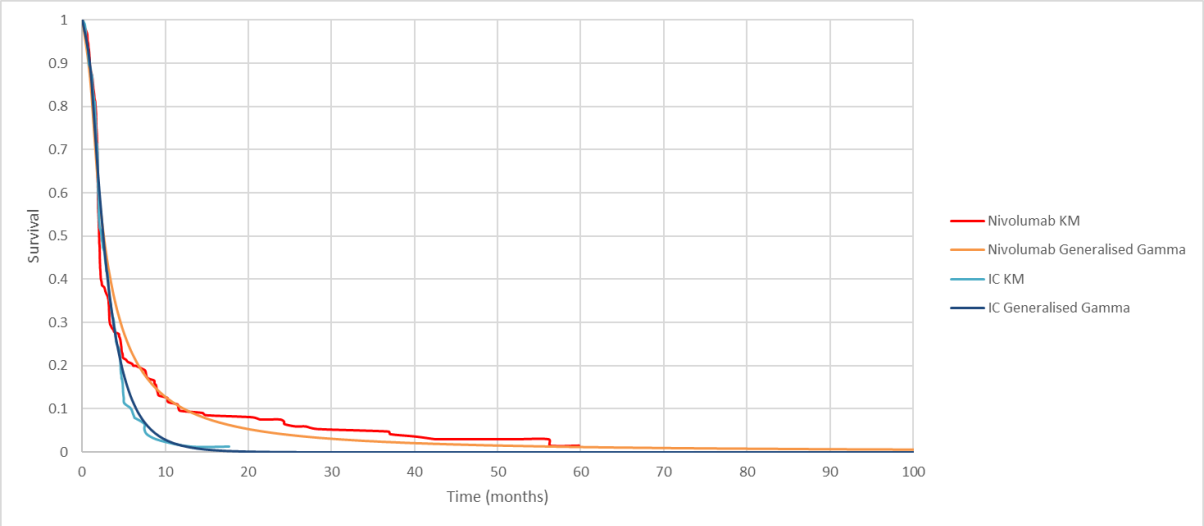
Source: CS Figure 13[2]

Figure 4.3: OS Kaplan-Meier with selected piecewise model and alternative parametric models



Source: CS Figure 15[2]

Figure 4.4: PFS Kaplan-Meier with generalised Gamma model



Source: CS Figure 16[2]

Figure 4.5: TTD Kaplan-Meier with generalised Gamma and two-spline normal model

[redacted]

Source: CS Figure 17

Plausibility of selected distribution for extrapolation

The company did not report on the plausibility of the selected distributions for extrapolation.

Selection of model for patient subgroups based on PD-L1 <1% and ≥1%

[redacted]

For patients with PD-L1 <1% and ≥1% receiving nivolumab, the piecewise method was used to extrapolate OS from the latest data cut of the CheckMate 141 trial. As for the overall population, the log-normal piecewise models produced a better fit compared to piecewise models using the exponential distribution. Piecewise models using a week 48 cut-off provided a reasonable fit to the observed data in both PD-L1 <1% and ≥1% subgroups. The week 96 cut-off piecewise models were not used as extrapolations at this later cut-off point were based on few patients in each of the subgroups.

To extrapolate PFS for nivolumab (PD-L1 <1% subgroup), the generalised gamma model was selected for extrapolation of PFS, providing good visual fit (and best statistical fit of non-spline models). The spline models provided better statistical fit than the standard parametric models, but the best fitting curves often produced logical inconsistencies when compared to the preferred extrapolation for OS. For the PD-L1 ≥1% subgroup, the log-logistic model provided the best statistical fit but a poor visual fit to the observed data. The one-spline hazards model provided reasonable statistical and visual fit, [redacted], and was thus selected for use in the model.

To extrapolate TTD for nivolumab (PD-L1 ≥1% subgroup), the two-spline normal model provided the best statistical fit. However, the one-spline odds model provided a better visual fit to the observed data compared to the one-spline odds model, [redacted], and was thus selected for use in the model.

Tables 4.2 and Table 4.3 provide an overview of the goodness-of-fit data for the patient subgroups based on PD-L1 <1% and ≥1%. Table 4.4 provides an overview of the company preferred approaches to estimate OS, PFS and TTD.

Table 4.2: Summary of goodness-of-fit data (PD-L1 <1% subgroup)

	OS		PFS		TTD	
Distribution	**AIC**	**BIC**	**AIC**	**BIC**	**AIC**	**BIC**
Nivolumab						
Exponential	523.061	525.391	382.266	384.597	372.696	375.000
Weibull	521.397	526.058	370.017	374.678	367.723	372.331
Gamma	523.027	527.688	379.939	384.600	371.248	375.856
Gompertz	518.899	523.560	340.312	344.973	362.022	366.630
Log-normal	**514.495**	**519.157**	330.201	334.862	365.298	369.906
Log-logistic	517.230	521.892	317.282	321.944	357.779	**362.387**
Generalised-gamma	516.495	523.487	312.145	319.137	363.601	370.513
Spline models:						
1-Spline Hazard	517.110	524.103	303.342	310.334	361.395	368.307
2-Spline Hazard	516.808	526.131	304.969	314.292	359.192	368.408
1-Spline Odds	519.069	526.061	292.756	**299.748**	358.682	365.594
2-Spline Odds	517.343	526.666	**291.913**	301.236	357.682	366.898
1-Spline Normal	516.485	523.478	301.060	308.052	362.587	369.499
2-Spline Normal	517.139	526.462	517.139	526.462	**356.983**	366.200
IC						
Exponential	258.516	**260.204**	170.794	172.483	167.034	**168.698**
Weibull	260.161	263.539	168.161	171.538	167.801	171.128
Gamma	259.592	262.969	166.981	170.359	167.945	171.272
Gompertz	260.471	263.849	170.969	174.347	168.473	171.800
Log-normal	**257.796**	261.173	**166.113**	**169.491**	180.353	183.681
Log-logistic	258.502	261.880	167.172	170.550	171.449	174.776
Generalised-gamma	259.286	264.353	167.858	172.924	169.800	174.790
Spline models:						
1-Spline Hazard	259.223	264.289	167.859	172.925	169.608	174.598
2-Spline Hazard	261.034	267.789	169.768	176.524	167.844	174.498
1-Spline Odds	260.390	265.456	169.162	174.228	**166.443**	171.433
2-Spline Odds	261.636	268.391	170.964	177.719	168.035	174.689
1-Spline Normal	259.376	264.443	167.923	172.989	167.055	172.045
2-Spline Normal	261.335	268.091	169.906	176.662	169.906	176.560

Source: Based on CS Appendix B and the economic model

Note: the lowest AIC/BIC is printed in bold.

AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; CS = company submission; IC = investigator's choice; OS = overall survival; PFS = progression-free survival; TTD = time to treatment discontinuation

Table 4.3: Summary of goodness-of-fit data (PD-L1 ≥1% subgroup)

	OS		PFS		TTD	
Distribution	**AIC**	**BIC**	**AIC**	**BIC**	**AIC**	**BIC**
Nivolumab						
Exponential	645.037	647.601	397.867	400.344	544.974	547.538
Weibull	641.707	646.835	399.604	404.559	516.025	521.154
Gamma	644.015	649.144	398.397	403.352	522.798	527.927
Gompertz	635.344	640.473	398.474	403.428	513.679	518.808
Log-normal	637.062	642.191	388.003	392.957	519.720	524.849
Log-logistic	**634.880**	**640.009**	**387.144**	**392.098**	512.473	517.601
Generalised-gamma	637.890	645.583	389.960	397.392	514.839	522.532
Spline models:						
1-Spline Hazard	637.880	645.573	387.619	395.051	514.887	522.580
2-Spline Hazard	637.412	647.669	389.879	399.788	510.672	520.930
1-Spline Odds	636.398	644.091	387.811	395.243	509.638	**517.331**
2-Spline Odds	638.195	648.453	389.690	399.599	510.117	520.375
1-Spline Normal	637.380	645.073	389.888	397.320	511.956	519.649
2-Spline Normal	637.898	648.155	389.544	399.453	**509.520**	519.778
IC						
Exponential	352.238	354.349	223.310	225.421	188.395	190.438
Weibull	353.307	357.529	209.443	213.665	188.092	192.178
Gamma	351.796	356.018	**208.213**	**212.435**	189.581	193.667
Gompertz	354.055	358.277	216.230	220.452	**183.339**	**187.425**
Log-normal	**346.405**	**350.627**	211.391	215.612	210.804	214.890
Log-logistic	347.544	351.765	210.795	215.017	204.963	209.049
Generalised-gamma	348.282	354.615	210.183	216.516	184.181	190.310
Spline models:						
1-Spline Hazard	348.730	355.062	210.140	216.472	184.847	190.977
2-Spline Hazard	350.620	359.063	212.193	220.637	186.771	194.943
1-Spline Odds	349.279	355.612	211.530	217.863	191.835	197.964
2-Spline Odds	351.263	359.706	212.982	221.425	189.356	197.528
1-Spline Normal	348.181	354.513	210.215	216.547	203.170	209.300
2-Spline Normal	349.882	358.325	212.125	220.569	188.068	196.240

Source: CS Appendix B[13] and the economic model.

Note: the lowest AIC/BIC is printed in bold.

AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; CS = company submission; IC = investigator's choice; OS = overall survival; PFS = progression-free survival; TTD = time to treatment discontinuation

Table 4.4: Summary selected parametric survival models

	Selected extrapolations		
	OS	**PFS**	**TTD**
Total population (original assessment; TA490)			
Nivolumab	Piecewise log-normal (different cut offs)[a]	Generalised gamma	Generalised gamma
IC	Piecewise log-normal (different cut offs)[a]	Generalised gamma	Generalised gamma
Total population (current assessment)			
Nivolumab	Piecewise log-normal 96-week cut off	Generalised gamma	2-spline normal
IC	Piecewise log-normal 96-week cut off	Generalised gamma	[redacted]
PD-L1 <1%			
Nivolumab	Piecewise log-normal 48-week cut off	Generalised gamma	[redacted]
IC	Kaplan-Meier data	Kaplan-Meier data	[redacted]
PD-L1 ≥1%			
Nivolumab	Piecewise log-normal 48-week cut off	1 spline hazards	1 spline odds
IC	Kaplan-Meier data	Kaplan-Meier data	[redacted]

Source: Company submission.[2]
IC = investigator's choice; OS = overall survival; PFS = progression-free survival; TTD = time to treatment discontinuation
[a]The log-normal (had the best goodness-of-fit statistics) and log-logistic distributions were considered to be plausible candidates by the company

ERG comment: The main concerns of the ERG relate to: a) the generalisability of the IC arm to docetaxel; b) equivalence of nivolumab flat dose and weight-based nivolumab; c) treatment waning assumptions for OS; d) estimation of OS; e) use of fully parametric models and; f) estimation of TTD.

a) As stated in the ToE, docetaxel is the comparator of interest in the CDF review. Effectiveness of docetaxel was however informed by the IC arm from CheckMate 141. The IC arm consists of docetaxel, methotrexate and cetuximab. Therefore, the ERG (as also stated in the ERG report for TA490) would ideally prefer to use treatment specific effectiveness estimates in its base-case (i.e. using docetaxel specific data). Main reasons for this preference are 1) the potential impact on the relative treatment effect of nivolumab (see published subgroup analyses of CheckMate 141 indicating the relative treatment effect is for nivolumab is less in the docetaxel subgroup); 2) in the TA490 guidance, the committee noted these subgroup results and indicated that the committee was not persuaded by the company's assumption that docetaxel is equivalent to methotrexate (see also ToE) and; 3) given cetuximab which is not considered by clinical experts to be established practice in England (according to TA490 guidance). Therefore, the ERG requested (clarification question B1) that the company would use the subgroup of patients (from CheckMate 141) who were randomised to docetaxel versus. those who would be eligible to receive docetaxel according to IC, but who were randomised to nivolumab to inform the economic model. Unfortunately, the company did not provide these analyses.

b) As highlighted in Sections 2.2 and 4.1.3, it is unclear whether the nivolumab flat dose can be assumed equally effective to weight-based nivolumab and thus to what the degree the CheckMate 141 nivolumab (relative) effectiveness estimates are generalisable to the currently used nivolumab flat dose.
c) The company assumed no treatment waning of nivolumab effectiveness. However, the (smoothed) hazard rate of nivolumab and IC seem to converge (indicating similar mortality probabilities for both treatments, see clarification response Figure 2), this converging trend might potentially occur earlier if continued nivolumab treatment after two years was not allowed in the CheckMate 141 trial (i.e. if the two-year stopping rule for nivolumab was reflected in the clinical data). Therefore, the ERG include treatment waning of nivolumab OS benefit after year 5 (assuming similar mortality probabilities for both treatments after year 5).
d) In response to clarification question B2, the company provided different distributions (than Exponential and log-normal) to extrapolate OS using the piecewise model with a 96-week cut off. Based on the AIC (clarification response Table 10) the generalised gamma distribution seemed to be an appropriate candidate to extrapolate OS (given its lower AIC for IC). However, after inspection of the piecewise generalised gamma 96-week cut off curve, it seemed implausible for IC (given the mortality probability was 100% at a certain point). Therefore, the ERG would, based on the AIC, agree with the log-normal distribution to extrapolate OS using the piecewise model with a 96-week cut off. However, it should be noted that the selection of the approach to extrapolate OS is not informed by external validation (neither expert opinion nor external data) of the extrapolated OS. Hence, the plausibility of the extrapolated OS might be considered uncertain.
e) Although the committee clearly indicated that a piecewise model is expected to be used to extrapolate OS, the ERG agrees with the company that fully parametric models are still considered to provide plausible alternative to extrapolate OS. Therefore, it should be noted that the company explored fully parametric models to extrapolate OS in scenario analyses (CS Table 22), using log-normal and log-logistic distributions (both increasing the estimated ICERs).
f) The company used the two-spline normal (nivolumab) and the [redacted] (IC) to estimate TTD. The generalised gamma distribution was the preferred distribution to model TTD in TA490 and, according to the ERG, there is no clear justification to deviate from this. Additionally, [redacted]. Given the above, the ERG preferred to use the generalised gamma distribution to estimate TTD (for both nivolumab and IC) in the ERG base-case.

4.1.6 Adverse events

The approach to incorporate the impact of AEs on costs and utility was similar to TA490, i.e. incorporated in the first cycle of the model (once only). Any all-cause Grade 3 or 4 AE were included if the incidence was ≥5% in either arm of the CheckMate 141 trial. Subsequently clinical expert opinion was sought to validate these AEs and to confirm that no AEs with a meaningful cost or disutility had been omitted using these criteria. Based on clinical expert feedback dysphagia, nausea and vomiting and anorexia were incorporated as well. Additionally, pneumonitis was included based on ERG preferences.

ERG comment: The ERG considers the 'once only' approach not to be in line with best practices but does not regard this as a priority issue because the impact on the incremental outcomes is most likely minimal.

4.1.7 Health-related quality of life

EQ-5D-3L data from the CheckMate 141 trial

In TA490, treatment-dependent health state utilities for the PF and PD states were derived from the EQ-5D-3L data collected from patients in the CheckMate 141 trial and analysed using mixed models in which progression status with and without treatment arm were included as covariates (see FAD Committee Papers 5; BMS additional evidence submitted in response to ACD; Appendix 1).[15] The company conducted no further analyses to estimate utility based on progression status. See Table 4.5 for the utility values used by the company (regression model 6; treatment dependent).

Table 4.5: Utility values estimated based on the CheckMate 141 trial (as per TA490)

	Nivolumab	IC	Difference
Regression model 6 (treatment dependent)			
Progression-free	[redacted]	[redacted]	[redacted]
Progressed disease	[redacted]	[redacted]	[redacted]
Regression model 7 (treatment independent)			
Progression-free	[redacted]	[redacted]	[redacted]
Progressed disease	[redacted]	[redacted]	[redacted]

Source: CS and FAD Committee Papers 5. BMS additional evidence submitted in response to ACD, Appendix 1[15]
IC = investigator's choice; OS = overall survival

Duration of nivolumab quality-of-life benefit

According to the ToE for CDF review, the committee was concerned that the abovementioned utility values were associated with significant uncertainty and that quality-of-life benefit cannot be assumed to remain constant over time.

In the ToE it was stated that the most appropriate utility values lie between the treatment-dependent (regression model 6) and the treatment-independent (regression model 7) estimates. It is noteworthy that in one of the TA490 ERG addenda, the ERG explored the use of a disutility of [redacted] (difference in post progression utility between nivolumab and IC) for patients that discontinued nivolumab treatment as an alternative scenario (i.e. assuming treatment independent utility values after treatment discontinuation).[6] Also, in this addenda, the ERG wondered why the company did not opt to use regression Model 1 or Model 2 (adding a covariate for being off treatment), given the lower AIC. These models indicate the post-progression utility difference between the two treatments of [redacted] is potentially an overestimation given that this is [redacted] when considering the model with the lowest AIC.

To incorporate time dependency, the company used CheckMate 141 trial data to estimate utility decrements (both treatment-dependent and treatment-independent) related to time before death (CS Table 15). Using this approach utility decrements are applied for the proportion of patients who are predicted to die within the next three model cycles, with separate decrements applied based on whether patients are one (0–28 days), two (29–56 days) or three (57–84 days) cycles from death.

Adverse event utility decrements

Consistent with TA490, utility decrements were applied separately for each AE and were applied once during the first cycle of the model, based on the proportion of patients in each treatment arm experiencing each AE.

ERG comment: The main concerns of the ERG relate to: a) the health state utilities not being updated using the latest CheckMate141 data cut-off (15 October 2019); b) incorporating time dependency of nivolumab utility benefit.

a) In the ToE, the committee emphasised that it "was concerned that the utility values calculated by the company's mixed model approach were associated with significant uncertainty". In clarification question B7 the ERG requested the company to provide updated utilities based on progression status using the latest data from the CheckMate 141 trial (data cut-off: 15 October 2019). However, the company did not provide these. In response to clarification question B7 the company does state that "Whilst the number of observations has increased since the earlier data cut, there were very few additional observations in the IC arm ([redacted]) and at Week 57, [redacted] in the nivolumab arm were still in the study and able to complete an EQ-5D questionnaire.".[5] Although the ERG would have preferred updated utilities based on progression status, the ERG agrees with the company that the impact, given the limited number of additional observations, might be rather small.
b) In the ToE for CDF review NICE stated that it expected the quality-of-life benefit to not remain constant over time and that the appropriate utility values should be reviewed in light of any new evidence. The company tried to address this by applying decrements in utility based on the proportion of patients who are predicted to die within the next three model cycles (so last three months only). Whilst this approach may account, to some extent, for decreasing health state utilities over time (see CS Table 15), according to the ERG this does not address the committee's concerns regarding the nivolumab quality of life (treatment) over time. According to the ERG, it would have been more intuitive to use time since start/ stop treatment (rather than time to death) to address this concern. In the PD state patients in the nivolumab arm have a large treatment benefit compared to patients in the IC arm ([redacted] utility difference). As stated in the ERG report for TA490 (and highlighted above), the ERG wonders why the company did not opt to use a regression in which a covariate for being off treatment was added. This could then in turn be used for patients that discontinued nivolumab treatment (i.e. assuming treatment independent utility values after treatment discontinuation), as done in regression Model 1 or Model 2 (which had a better AIC than the currently used regression models). This would remove the constant quality of life benefit of treatment over time, which would have addressed the concerns highlighted in the ToE. Hence, to reflect the uncertainty, the ERG explored two base-cases, one with treatment-dependent utilities (based on regression model 6; Table 4.5), and one with treatment-independent utilities (based on regression model 7; Table 4.5). Additionally, the company's approach to obtain utility decrements related to time to death was not completely clear (e.g. what data cut-off was used, the number of observations included, details regarding the regression model), the ERG excluded the utility decrements related to time to death in scenario analyses.

4.1.8 Resources and costs

Resource use and costs included in the CS model were based on data from the CheckMate 141, previous technology appraisals and published sources identified in the SLR of TA490.

Intervention and comparators' costs and resource use

Treatment costs

Drug acquisition costs were obtained from the British National Formulary for nivolumab and from the electronic market information tool for IC drugs. A PAS ([redacted]) has been agreed for nivolumab.

The dosing frequency for docetaxel, methotrexate and paclitaxel are provided in Table 4.6. Since the original submission for TA490, the licensed dose of nivolumab has been updated to a flat dose of 240 mg every two weeks, rather than the weight-based dose used in the CheckMate 141 trial (3 mg/kg every two weeks). The flat dose approximates the exposures achieved with 3 mg/kg in patients weighing 80 kg.

Table 4.6: Treatment costs

	Dosage	Treatment costs (per 28-day cycle)	Administration costs (per 28-day cycle) [b]	Monitoring costs (per 28-day cycle) [b]
Nivolumab (flat dose)	240 mg Q2W	[redacted]	£371.06	£190.79
Nivolumab (weight based)[a]	3 mg/kg Q2W	[redacted]	£371.06	£190.79
Docetaxel [a]	75 mg/m^2 Q3W	£33.32	£247.37	£190.79
Methotrexate [a]	40 mg/m^2 QW	£48.76	£742.12	£190.79
Paclitaxel [a]	80 mg/m^2 QW	£68.84	£742.12	£190.79

Source: CS and Economic model.[2]
[a]Mean weight and BSA were based on the population of European patients reported in CheckMate 141 ([redacted] respectively).
[b]All therapies included in the model are intravenously-administered and therefore assumed to incur the same administration costs per administration.
IC = investigator's choice; OS = overall survival

No vial sharing was assumed for all therapies. A reduction in dose intensity was included in the base-case based on the proportion of doses received that were delayed in CheckMate 141. Dose intensity was estimated to be [redacted] for nivolumab, docetaxel and methotrexate, respectively. This calculation relied on the assumption that a dose delay was equivalent to a single missed dose for nivolumab, methotrexate or docetaxel – in CheckMate 141 (i.e. the drug cost would not be incurred by the NHS), the average dose delay was [redacted] days for nivolumab, [redacted] days for methotrexate and [redacted] days for docetaxel. The reduction in dose intensity calculated for docetaxel ([redacted]) was also applied to paclitaxel, in the absence of data for paclitaxel from CheckMate 141. Although the committee considered analyses without a stopping rule are more appropriate for decision-making (based on ToE), the company applied a two-year stopping rule.

Subsequent systemic therapy

In the base-case analysis, the proportion of patients who received subsequent systemic therapy post-discontinuation was assumed to be treatment independent, in line with ERG preferences (ERG report for TA490) and the ToE.

Health state and event costs

Health state and event costs were implemented as per TA490. Health state costs consisted of costs related to the PF and PD health states as well as event costs related to progression (one oncologist visit and one CT scan in order to confirm disease progression) and death (terminal care cost).

Adverse event costs

As per TA490, the costs per episode of treating AEs were sourced from currency codes for NHS reference costs and assumptions used in previous appraisals.

ERG comment: a) the validity of the TTD assumptions for UK clinical practice; b) incorporating dose intensity when calculating docetaxel treatment costs and; c) the two-year stopping rule.

- **a)** Compared with the CheckMate 141 trial, the SACT data provides real-world data that might better reflect UK clinical practice. To this extent, the ERG requested the company to provide a scenario analysis using the SACT data to estimate time to TTD for nivolumab (clarification question B6). In their response, the company stated that "TTD in the SACT cohort was generally higher than that observed in the CheckMate 141 trial, as shown in the company evidence submission. The use of TTD data from the SACT cohort in the cost-effectiveness analysis therefore produces a higher estimate of the ICER than the base-case analysis (i.e. using data from CheckMate 141) due to the increased costs related to treatment that are accrued in the nivolumab arm." The substantial increase in the ICER (+£14,198 compared to the CS base-case) highlights the importance of the TTD assumptions in the model and may be subject to a large degree of uncertainty. Hence, if the nivolumab treatment duration from the SACT is believed to better reflect UK clinical practice, this would substantially increase the estimated ICERs (both those presented as the ERG as well as CS base-case).
- **b)** In the calculation of treatments costs for docetaxel, when assuming no vial sharing, the company included the average dose intensity in their calculation of the number of required vials per mg/m2 group. As dose intensity is related to doses that are missed (rather than the number of vials per mg/m2 group), the dose intensity should rather be applied to the calculated docetaxel costs per administration. Hence, the ERG corrected the implementation of dose intensity, resulting in per cycle costs for docetaxel of £30.39 (instead of £33.32 per cycle; see Table 4.8).
- **c)** The company incorporated a two-year stopping rule to nivolumab. However, according to the ToE, the committee considered analyses without a stopping rule as more appropriate for decision-making. Moreover, excluding the two-year stopping rule is consistent with the CheckMate 141 trial data used to estimate effectiveness. The justification by the company to include the stopping rule is minimal (i.e. [redacted], and a two-year stopping rule has been shown to be clinically plausible during the CDF data collection period). Therefore, the ERG excluded the two-year stopping rule in its base-case

5. COST EFFECTIVENESS RESULTS

5.1 Company's cost effectiveness results

The company cost effectiveness results are described for the all-randomised population and patient subgroup based on PD-L1 status. First, the company stated that they have replicated the key cost effectiveness results (cost effectiveness (C-E) analysis 1) used in the committee's decision-making at the point of CDF entry (i.e. data cut-off: September 2016). Second, the company provided cost effectiveness results (C-E analysis 2) that incorporated data collected during the CDF data collection period (i.e. data cut-off: October 2019), which included the committee's preferred assumptions for decision-making at the point of CDF entry. Third, the company provided a revised base-case analysis C-E analysis 3). These cost effectiveness results incorporated data collected during the CDF data collection period, plus any associated changes to the company's preferred assumptions, as stated in Table 5.1. For the cost effectiveness analyses a flat dose of 240 mg every two weeks (Q2W) nivolumab was used.

Table 5.1: Key model assumptions and inputs

Model input and cross reference	C-E analysis 1 (Original assumptions)	C-E analysis 2	C-E analysis 3
OS, PFS and TTD data source	CheckMate 141 (Data cut-off: 20 September 2016)	CheckMate 141 (Data cut-off: 15 October 2019)	CheckMate 141 (Data cut-off: 15 October 2019)
OS extrapolation	Nivolumab and IC: piecewise with log-normal (20, 36 and 48 week cut-off points)	Nivolumab and IC: piecewise with log-normal (20, 36 and 48 week cut-off points)	Nivolumab and IC: piecewise with log-normal (96-week cut-off point)
Long-term treatment waning effect	Treatment waning at 5 years included	Treatment waning at 5 years included	Treatment waning at 5 years excluded
PFS extrapolation	Nivolumab and IC: generalised gamma	Nivolumab and IC: generalised gamma	No change
TTD extrapolation	Nivolumab and IC: generalised gamma	Nivolumab and IC: generalised gamma	Nivolumab: 2-spline normal IC: [redacted]
Utility values	**Treatment-specific** PF nivolumab: [redacted] PD nivolumab: [redacted] PF IC: [redacted] PD IC: [redacted] **Treatment independent** PF: [redacted] PD: [redacted]	**Treatment-specific** PF nivolumab: [redacted] PD nivolumab: [redacted] PF IC: [redacted] PD IC: [redacted] **Treatment independent** PF: [redacted] PD: [redacted]	**Treatment-specific** PF nivolumab: [redacted] PD nivolumab: [redacted] PF IC: [redacted] PD IC: [redacted] With time-to-death utility decrements applied

Model input and cross reference	C-E analysis 1 (Original assumptions)	C-E analysis 2	C-E analysis 3
Stopping rule	2-year stopping rule included	2-year stopping rule included	No change
ERG's amendments to the company's model	Adding the cost and disutility for pneumonitis and using treatment-independent proportions for subsequent treatment	Adding the cost and disutility for pneumonitis and using treatment-independent proportions for subsequent treatment	No change

Source: Based on CS Table 17.[2]
ACD: Appraisal Consultation Document; ERG: Evidence Review Group; FAD: Final Appraisal Determination; IC: investigator's choice; OS: overall survival; PD: progressed disease; PF: progression free; PFS: progression-free survival; TTD: time to treatment discontinuation

5.1.2 Overall population

Replication of the key cost effectiveness results used in committee's decision-making at the point of CDF entry

The company used a piecewise model using the log-normal distribution to model OS, extrapolated from 20, 36 and 48 weeks (estimated based on September 2016 data cut-off). The company used both treatment-dependent and treatment-independent utility values. The analyses include a PAS discount of [redacted]% to the list price of nivolumab. The ICER per QALY gained for nivolumab versus docetaxel ranged from £45,874 to £67,555 depending on the cut-off (20, 36, or 48 weeks) and utility (treatment-specific, or treatment independent) used (Table 5.2).

Table 5.2: Cost effectiveness analysis 1: Replication of analysis that demonstrated plausible potential for cost effectiveness at CDF entry (with PAS) – overall population, flat dose

Technologies	Incr. costs (£)	Incr. QALYs	ICER (£/QALY gained)	Incr. costs (£)	Incr. QALYs	ICER (£/QALY gained)	Incr. costs (£)	Incr. QALYs	ICER (£/QALY gained)
Piecewise log-normal cut-off point:	**20 weeks**			**36 weeks**			**48 weeks**		
Treatment-specific utility									
Docetaxel	[redacted]	[redacted]	£45,874	[redacted]	[redacted]	£41,304	[redacted]	[redacted]	£53,634
Paclitaxel	[redacted]	[redacted]	£42,252	[redacted]	[redacted]	£38,065	[redacted]	[redacted]	£49,363
Methotrexate	[redacted]	[redacted]	£43,215	[redacted]	[redacted]	£38,925	[redacted]	[redacted]	£50,498
Treatment-independent utility									
Docetaxel	[redacted]	[redacted]	£58,448	[redacted]	[redacted]	£52,528	[redacted]	[redacted]	£67,555
Paclitaxel	[redacted]	[redacted]	£53,833	[redacted]	[redacted]	£48,409	[redacted]	[redacted]	£62,175
Methotrexate	[redacted]	[redacted]	£55,059	[redacted]	[redacted]	£49,503	[redacted]	[redacted]	£63,604

Source: Based on CS Table 17.[2]

CS = company submission; ICER = incremental cost effectiveness ratio; LY = life-years; PAS = Patient Access Scheme; QALYs = quality-adjusted life years; incr. = incremental

ERG comment: As stipulated in the ToE, the company should provide a replication of the key cost effectiveness results used in the committee's decision-making at the point of CDF entry. The ICERs abovementioned results (reported in CS Table 17 and Appendix D Table 15) do not appear to be in line with the ICERs reported in the Final Appraisal Document or ToE for nivolumab compared with docetaxel (i.e. these ICERs do not range between either £45,000 and £73,600 or, as per the commercial access agreement, £30,377 and £49,408 per quality-adjusted life year gained). After clarification (response to question B12) from the company, it became clear that the differences were due the application of the higher [redacted]% PAS discount and/ or the application of the two-year stopping rule. Based on these clarifications, the ERG was able to reproduce the ICER used in the committee's decision-making at the point of CDF entry.

Cost effectiveness results that incorporate data collected during the CDF data collection period, with the assumptions used in committee's decision-making at the point of CDF entry

The company used a piecewise model using the log-normal distribution to model OS, extrapolated from 20, 36 and 48 weeks. The company used both treatment-dependent and treatment-independent utility values. The analyses included a PAS discount of [redacted]% to the list price of nivolumab. The ICER per QALY gained for nivolumab versus docetaxel ranged from £41,906 to £55,051 depending on the cut-off (20, 36, or 48 weeks) and utility (treatment-specific, or treatment independent) used (Table 5.3).

Table 5.3: Cost effectiveness analysis 2: Analysis that demonstrated plausible potential for cost effectiveness at CDF entry – incorporating updated clinical evidence (with PAS) – overall population, flat dose

Technologies	Incr. costs (£)	Incr. QALYs	ICER (£/QALY gained)	Incr. costs (£)	Incr. QALYs	ICER (£/QALY gained)	Incr. costs (£)	Incr. QALYs	ICER (£/QALY gained)
Piecewise log-normal cut-off point:	**20 weeks**			**36 weeks**			**48 weeks**		
Treatment-specific utility									
Docetaxel	[redacted]	[redacted]	£43,959	[redacted]	[redacted]	£41,906	[redacted]	[redacted]	£45,793
Paclitaxel	[redacted]	[redacted]	£40,644	[redacted]	[redacted]	£38,757	[redacted]	[redacted]	£42,333
Methotrexate	[redacted]	[redacted]	£41,527	[redacted]	[redacted]	£39,596	[redacted]	[redacted]	£43,255
Docetaxel	[redacted]	[redacted]	£53,510	[redacted]	[redacted]	£50,728	[redacted]	[redacted]	£55,051
Paclitaxel	[redacted]	[redacted]	£49,474	[redacted]	[redacted]	£46,916	[redacted]	[redacted]	£50,892
Methotrexate	[redacted]	[redacted]	£50,550	[redacted]	[redacted]	£47,932	[redacted]	[redacted]	£52,000

Source: Based on CS Table 18.[2]

CS = company submission; ICER = incremental cost effectiveness ratio; LY = life-years; PAS = Patient Access Scheme; QALYs = quality-adjusted life years; incr. = incremental

ERG comment: As stipulated in the ToE, the company should provide a replication of the key cost effectiveness results that incorporate data collected during the CDF data collection period, with the assumptions used in committee's decision making at the point of CDF entry. Because the results of the replication (cost effectiveness analysis 1) was not consistent with the original results (section above), the validity of the results (cost effectiveness analysis 2) was unclear. However, the ERG was able to replicate the original results after clarification of the company (section above). Therefore, the ERG considers the results of cost effectiveness analysis 2 to be reproducible (using the cost effectiveness estimates at CDF entry as starting point).

Cost effectiveness results that incorporate data collected during the CDF data collection period plus any associated changes to the company's preferred assumptions.

The analyses included a PAS discount of [redacted]% to the list price of nivolumab. The increased QALYs and costs for nivolumab resulted in ICERs of £37,236, £34,186, and £35,019 per QALY gained versus docetaxel, paclitaxel and methotrexate, respectively (Table 5.4).

Table 5.4: New company base-case results (nivolumab with PAS) – overall population

Treatment	Total costs	Total LYs	Total QALYs	Incremental costs	Incremental LYs	Incremental QALYs	ICER
Nivolumab	[redacted]	1.31	[redacted]				
Docetaxel	£10,569	0.67	0.35	[redacted]	0.65	[redacted]	£37,236
Paclitaxel	£12,000	0.67	0.35	[redacted]	0.65	[redacted]	£34,186
Methotrexate	£11,609	0.67	0.35	[redacted]	0.65	[redacted]	£35,019

Source: Based on CS Table 19.[2]
CS = company submission; ICER = incremental cost effectiveness ratio; LY = life-years; PAS = Patient Access Scheme; QALYs = quality-adjusted life years

ERG comment: It is noteworthy that in the CS base-case the majority of the estimated QALY gain (~65%) is attributable to the period after disease progression has been confirmed. This implies that additional benefit continues to accrue to patients whose disease has progressed. The plausibility of the proportion of post-progression gains is unclear to the ERG.

5.1.3 Patients with PD-L1 <1% and ≥1%

As requested in the ToE, the company provided cost-effectiveness results of nivolumab versus docetaxel for the PD-L1 expression subgroups (PD-L1<1%, and PD-L1≥1%) (Table 5.5). The results for the revised base-case (cost effectiveness analysis 3) incorporate the inputs and assumptions as described in Table 5.L1.

According to the company, the clinical effectiveness results by PD-L1 status could not demonstrate a statistically significant difference between the subgroups in the treatment effect on OS. Therefore, the company stated that the evidence is such that the overall population should be considered as the patient population within the CDF review.

The revised base-case analyses (cost effectiveness analysis 3) (Table 5.5) resulted in ICERs of £46,309 and £36,163 per QALY gained for the subgroups PD-L1<1% and PD-L1≥1%, respectively.

Table 5.5: Summary of cost effectiveness analyses and revised base-case (with PAS) versus docetaxel only – PD-L1 subgroups, flat dose.

Analysis		ICER (£/QALY gained) versus docetaxel	
Utility values		**Treatment-specific**	**Treatment-independent**
PD-L1 <1%			
Cost effectiveness analysis 1, flat dose			
Piecewise log-normal cut-off point	20 weeks	£39,218	£53,242
	36 weeks[a]	-	-
	48 weeks	£65,154	£102,195
Cost effectiveness analysis 2, flat dose			
Piecewise log-normal cut-off point	20 weeks	£42,558	£54,341
	36 weeks[a]	-	-
	48 weeks	£47,982	£61,729
Cost effectiveness analysis 3, flat dose		£46,309	-
PD-L1 ≥1%			
Cost effectiveness analysis 1, flat dose			
Piecewise log-normal cut-off point	20 weeks	£43,647	£51,809
	36 weeks	£35,882	£41,020
	48 weeks	£41,581	£47,714
Cost effectiveness analysis 2, flat dose			
Piecewise log-normal cut-off point	20 weeks	£42,945	£49,710
	36 weeks	£42,061	£48,051
	48 weeks	£44,045	£50,253
Cost effectiveness analysis 3, flat dose		£36,163	-

Source: Based on CS Table 20.2

aAs noted in FAD Committee Papers 8; appendix, with 2-year stopping rule, the extrapolation of OS using the piecewise model with the 36-week cut-off point was not considered plausible by the company, particularly for the PD-L1 <1% subgroup. This cut-off point creates a kink in the shape of the survival curve for IC which causes the IC curve to cross the nivolumab curve and produce a plateau after 3 years. The resulting survival curve is therefore wholly clinically implausible given the known prognosis for patients with R/M SCCHN after platinum therapy. ICERs have therefore not been presented from the PD-L1 <1% subgroup using the 36-week cut-off point.

ICER, incremental cost-effectiveness ratio; LYG, life years gained; PAS: Patient Access Scheme; PD-L1: programmed death ligand 1; QALYs, quality-adjusted life years; R/M SCCHN: recurrent/metastatic squamous cell carcinoma of the head and neck.

ERG comment: According to Table 13 of the CS, the PD-L1 score for patients was not recorded for 42% (n=210) of the SACT data cohort study population. The ERG is concerned that testing for PD-L1 expression is not part of usual care for treating SCCHN patients within the UK. This would mean that if nivolumab would only be accepted for treating patients according to their PD-L1 expression level, additional testing on PD-L1 expression would be required, which will lead to additional costs related to

nivolumab. However, in response to clarification question B10, the company argues that PD-L1 testing is standard clinical practice in the UK, when required.

5.2. Company's sensitivity analyses

The company presented probabilistic sensitivity analysis (PSA), deterministic sensitivity analysis (DSA) and deterministic scenario analysis.

Probabilistic sensitivity analysis

A probabilistic sensitivity analysis (PSA) with 1,000 iterations was conducted. In each iteration, the model inputs were randomly drawn from the specified distributions. Whenever available, the standard error of the selected distribution was obtained directly from the same data source that informed the mean value. In the absence of data on the variability around a particular value, it was varied by ±15%.

The base-case results using PSA are presented in Table 5.6 and resulted in slightly lower ICERs than those presented for the new deterministic company base-case. The ICERs were £36,255, £33,340 and £34,059 for nivolumab versus docetaxel, paclitaxel and methotrexate, respectively.

Table 5.6: Revised base-case results (average probabilistic) (with PAS) – overall population, flat dose

Treatment	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER
Nivolumab	[redacted]	[redacted]			
Docetaxel	£10,574	0.36	[redacted]	[redacted]	£36,255
Paclitaxel	£11,983	0.36	[redacted]	[redacted]	£33,340
Methotrexate	£11,638	0.36	[redacted]	[redacted]	£34,059

Source: Based on CS Table 21.[2]
CS = company submission; ICER = incremental cost effectiveness ratio; PAS = Patient Access Scheme; QALYs = quality-adjusted life years

The company provided incremental cost effectiveness planes and cost effectiveness acceptability curves (CEACs; CS Figures 18 and 19). The company reported a [redacted] probability of nivolumab (with PAS) being cost effective at a threshold of £50,000 per QALY.

Deterministic sensitivity analysis

The company conducted DSA by varying all parameters for which there were single input values in the model. Whenever available, values were varied using the standard error obtained directly from the same data source that informed the mean value. In the absence of data on the variability around a particular value, it was varied by ±20%.

The DSA results are presented using tornado diagrams with the top 10 drivers of cost effectiveness in CS Figure 20. The company identified the following parameters as the main influential parameters on the cost effectiveness (in order of importance):

1. Nivolumab treatment frequency
2. Nivolumab utility value – Progressed disease
3. Nivolumab utility value – Progression free
4. Comparator utility value – Progressed disease
5. Comparator utility value – Progression free

6. Nivolumab administration cost
7. Nivolumab monitoring cost
8. Docetaxel administration cost
9. Docetaxel monitoring cost
10. Docetaxel treatment frequency

Deterministic scenario analysis

The company performed various deterministic scenario analyses, see Table 5.7.

Table 5.7: Deterministic scenario analyses performed by the company – overall population, flat dose

Scenario		Scenario detail	ICER vs docetaxel (£/QALY gained)	Impact on base-case ICER
	Base-case		£37,236	-
1	Alternative OS assumption	Piecewise log-normal 48-week cut-off for OS extrapolation	£40,167	+£2,931
2	Alternative OS assumption	Fully parametric log-normal	£41,158	+£3,922
3	Alternative OS assumption	Fully parametric log-logistic	£38,896	+£1,660
4	Treatment-dependent utility values	Treatment-dependent utility values No time-to-death utility decrements	£35,340	-£1,896
5	Treatment-independent utilities	Treatment-independent utility values Time-to-death utility decrements	£41,418	+£4,182
6	Treatment-independent utilities	Treatment-independent utility values No time-to-death utility decrements	£41,537	+£4,301
7	No stopping rule	2-year stopping rule is not applied	£49,018	+£11,782
8	Treatment waning (5 years)	Treatment waning applied from 5 years	£45,014	+£7,778
9	Treatment waning (7 years)	Treatment waning applied from 7 years	£41,639	+£4,403
10	Treatment waning (10 years)	Treatment waning applied from 10 years	£39,214	+£1,978
11	"Partial" treatment waning (5 years)	Treatment waning applied from 5 years for [redacted]% of patients only	£41,821	+£4,585
12	"Partial" treatment waning (7 years)	Treatment waning applied from 7 years for [redacted]% of patients only	£39,921	+£2,685
13	"Partial" treatment waning (10 years)	Treatment waning applied from 10 years for [redacted]% of patients only	£38,472	+£1,237

Source: Based on CS Table 22.[2]
Abbreviations: ICER: incremental cost effectiveness ratio; OS: overall survival.

The results of the scenario analyses are summarised in Table 5.7, showing that alternative OS assumptions, stopping rule, treatment-independent utilities, and treatment waning effects had a strong impact on the base-case ICER. The most influential scenarios were 1) removing the two-year stopping

rule (scenario 7; impact on base-case ICER: +£11,782), 2) implementing treatment waning from five years (scenario 8; base-case ICER: +£7,778), and 3) implementing partial treatment waning from five years (scenario 11; base-case ICER: +£4,585)

ERG comment: In addition to the sensitivity analyses provided in the CS (based on "cost effectiveness analysis 3", revised company base-case). The company also provided sensitivity analyses for "cost effectiveness analysis 2" (updated committee preferred base-case) in response to clarification question B13.

5.3 Model validation and face validity check

The company did not report on the validity of the economic model.

ERG comment: The ERG was able to reproduce the results mentioned in ToE (the committee preferred ICER range £45,000 and £73,600 per QALY per QALY gained). Moreover, the changes implemented related to updating of input parameters and not to the model structure. Therefore, the ERG believes that the internal validation described in TA490 (detecting no major errors) is still valid. However, also the ERG's concerns in TA490 regarding the lack of external validation hampers the interpretation of the cost effectiveness.

6. EVIDENCE REVIEW GROUP'S ADDITIONAL ANALYSES

6.1 Exploratory and sensitivity analyses undertaken by the ERG

Based on all considerations from Chapter 4, the ERG defined a new base-case. This base-case included multiple adjustments to the company base-case presented in the CS. These adjustments mainly consisted of adjustments that could be categorised as matters of judgement (amending the model were the ERG considers that reasonable alternative assumptions are preferred):

1. Include treatment waning of nivolumab OS benefit after year 5
   The (smoothed) hazard rate of nivolumab and IC seem to converge (indicating similar mortality probabilities for both treatments), this converging trend might potentially occur earlier if continued nivolumab treatment after two years was not allowed in the CheckMate 141 trial (i.e. if the two-year stopping rule for nivolumab was reflected in the clinical data) (Section 4.1.5).
2. Using the generalised gamma model for estimating TTD
   The generalised gamma distribution was the preferred distribution to model TTD in TA490 and, according to the ERG, there is no clear justification to deviate from this. Additionally [redacted] *(Section 4.1.5).*
3. Include both treatment dependent and treatment independent utility
   Although the company attempted to incorporate time dependent utility values, the time to death utility are more likely to reflect the declining utility towards the end of life than reflecting a nivolumab quality-of-life benefit that is not constant over time (Section 4.1.7).
4. Excluding the two-year stopping rule
   According to the ToE, the committee considered analyses without a stopping rule are more appropriate for decision-making. Moreover, excluding the two-year stopping rule is consistent with the CheckMate 141 trial data used to estimate effectiveness. The justification by the company to include the stopping rule is minimal (i.e. [redacted]*, and a two-year stopping rule has been shown to be clinically plausible during the CDF data collection period).*
5. Correcting error related to implementation of docetaxel dose intensity
   The ERG corrected an error related to the implementation of dose intensity for calculating docetaxel treatment costs (Section 4.1.8).

In addition, the following scenario analyses were performed:

1. Excluding the estimated utility decrements related to time before death

For the PD-L1 subgroups the following adjustments were implemented:

1. Using the piecewise log-normal 48-week cut off for estimating OS (i.e. [redacted]).
2. Using the generalised gamma model for estimating PFS (i.e. [redacted]).
3. Using the one-spline normal and generalised gamma models for estimating TTD for the PD-L1 <1% and PD-L1 ≥1% subgroups respectively (i.e. [redacted]).

These distributions were selected given the reasonable AIC and since these did not produce logical inconsistencies between TTD and OS.

For all three adjustments, [redacted].

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

Correcting the docetaxel dose intensity error as well as excluding the estimated utility decrements related to time before death (while assuming treatment dependent utility) have a minor impact on the estimated ICER. The other adjustments have a more pronounced impact on the estimated ICER (Tables 6.1-6.5).

6.3 ERG's preferred assumptions

Table 6.1: ERG analyses (deterministic), nivolumab with PAS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Nivolumab ICER (£/QALY)
Company base-case					
Nivolumab	[redacted]	[redacted]			
Docetaxel	£10,569	0.35	[redacted]	[redacted]	£37,236
1 Company base-case + OS treatment waning[a]					
Nivolumab	[redacted]	[redacted]			
Docetaxel	£10,569	0.35	[redacted]	[redacted]	£45,017
2 Company base-case + generalised gamma model for estimating TTD					
Nivolumab	[redacted]	[redacted]			
Docetaxel	£10,505	0.35	[redacted]	[redacted]	£39,959
3 Company base-case + treatment independent utility					
Nivolumab	[redacted]	[redacted]			
Docetaxel	£10,569	0.38	[redacted]	[redacted]	£41,418
4 Company base-case + excluding the 2-year stopping rule					
Nivolumab	[redacted]	[redacted]			
Docetaxel	£10,569	0.35	[redacted]	[redacted]	£49,018
5 Company base-case + correcting error related to implementation of docetaxel dose intensity					
Nivolumab	[redacted]	[redacted]			
Docetaxel	£10,561	0.35	[redacted]	[redacted]	£37,254

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Nivolumab ICER (£/QALY)
6 ERG base-case 1 Company base-case + OS treatment waning + generalised gamma model for estimating TTD + excluding the 2-year stopping rule					
Nivolumab	[redacted]	[redacted]			
Docetaxel	£10,497	0.35	[redacted]	[redacted]	£53,485
7 ERG base-case 2 Company base-case + OS treatment waning + generalised gamma model for estimating TTD + excluding the 2-year stopping rule + treatment independent utility					
Nivolumab	[redacted]	[redacted]			
Docetaxel	£10,497	0.38	[redacted]	[redacted]	£60,094

ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year; TTD = time to treatment discontinuation
[a]A minimum function was implemented to prevent that PFS would exceed OS (implemented in cells 'Nivolumab Traces'!G11:G370 and 'Docetaxel Traces'!G11:G370)
[b]The following cells were adjusted: Settings!J72:N72, 'Treatment Costs'!N24 and 'Docetaxel Traces'!AU11:AU369

Table 6.2: ERG scenario (deterministic), nivolumab with PAS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Nivolumab ICER (£/QALY)
6 ERG base-case 1- treatment dependent utility + excluding the estimated utility decrements related to time before death					
Nivolumab	[redacted]	[redacted]			
Docetaxel	£10,497	0.36	[redacted]	[redacted]	£50,140
7 ERG base-case 2 - treatment independent utility + excluding the estimated utility decrements related to time before death					
Nivolumab	[redacted]	[redacted]			
Docetaxel	£10,497	0.40	[redacted]	[redacted]	£60,264

ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year; TTD = time to treatment discontinuation

Table 6.3: ERG base-case (probabilistic), nivolumab with PAS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Nivolumab ICER (£/QALY)
6 ERG base-case 1- treatment dependent utility[a]					
Nivolumab	[redacted]	[redacted]			
Docetaxel	£10,556	0.36	[redacted]	[redacted]	£54,348

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Nivolumab ICER (£/QALY)
7 ERG base-case 2 - treatment independent utility[a]					
Nivolumab	[redacted]	[redacted]			
Docetaxel	£10,511	0.38	[redacted]	[redacted]	£61,293

ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year; TTD = time to treatment discontinuation

[a]The PSA produced 1 to 2 errors (#VALUE), these simulations were ignored to calculate the probabilistic means.

Table 6.4: ERG base-case; PD-L1 <1% subgroup (deterministic), nivolumab with PAS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Nivolumab ICER (£/QALY)
6 ERG base-case 1- treatment dependent utility					
Nivolumab	[redacted]	[redacted]			
Docetaxel	£11,048	0.41	[redacted]	[redacted]	£53,152
7 ERG base-case 2 - treatment independent utility					
Nivolumab	[redacted]	[redacted]			
Docetaxel	£11,048	0.43	[redacted]	[redacted]	£62,895

ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year; TTD = time to treatment discontinuation

Table 6.5: ERG base case; PD-L1 ≥1% subgroup (deterministic), nivolumab with PAS

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Nivolumab ICER (£/QALY)
6 ERG base-case 1- treatment dependent utility					
Nivolumab	[redacted]	[redacted]			
Docetaxel	£9,981	0.29	[redacted]	[redacted]	£54,362
7 ERG base-case 2 - treatment independent utility					
Nivolumab	[redacted]	[redacted]			
Docetaxel	£9,981	0.31	[redacted]	[redacted]	£58,926

ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year; TTD = time to treatment discontinuation

6.4 Conclusions of the cost effectiveness section

The company base-case ICER (probabilistic) of nivolumab (with PAS) compared with docetaxel was £36,255 per QALY gained. The ERG has incorporated various adjustments to the company base-case. The ERG base-case resulted in an ICER range (probabilistic) of £54,348 to £61,293 per QALY gained for nivolumab (with PAS) versus docetaxel. The most influential adjustments/corrections made by the ERG were 1) using a generalised gamma distribution for estimating TTD; 2) using treatment independent utilities for PFS and PD health states; 3) including treatment waning of nivolumab OS benefit after year 5 and; 4) excluding the two-year stopping rule. Additionally, the company explored using SACT data to estimate TTD (i.e. nivolumab treatment duration) in scenario analyses. Compared with the CheckMate 141 trial, the SACT data provides real-world data that might better reflect UK

clinical practice. The higher TTD observed in the SACT data resulted in a substantially increased ICER (+£14,198 compared to the CS base-case) highlighting the importance of the TTD assumptions in the model.

The equivalence assumptions between docetaxel and methotrexate as well as between the nivolumab flat dose and weight-based nivolumab can be questioned. Unfortunately, the company did not provide analyses based on the docetaxel subgroup (requested during the clarification phase), nor evidence to support the equivalence assumption for nivolumab (flat dose versus weight-based). An additional area of uncertainty is the extrapolation of the nivolumab quality-of-life benefit over time. Although the company implemented utility decrements related to the time to death, the ERG believes that the committee's concern (i.e. emphasizing that quality-of-life benefit cannot be assumed to remain constant over time) is not appropriately addressed. Therefore, the ERG base-case is presented as a range conditional on treatment dependent and treatment independent utilities to address the uncertainty related to the nivolumab utility benefit over time.

The subgroup analyses (based on PD-L1 status) performed by the ERG resulted in ICERs that ranged between £53,152 and £62,895 per QALY gained. It should however be noted that these subgroup analyses did not incorporate any additional costs related to PD-L1 which would be required if PD-L1 testing is not part of UK clinical practice.

In conclusion, the ERG base-case ICERs are estimated to be in the range between £54,348 and £61,293 per QALY gained, reflecting the uncertainty related to nivolumab quality-of-life benefits over time. Uncertainty that was not captured in this range included the equivalence assumptions between docetaxel and methotrexate as well as between the nivolumab flat dose and weight-based nivolumab. Additionally, if the nivolumab treatment duration from the SACT is believed to better reflect UK clinical practice (than TTD from CheckMate 141), this would substantially increase the estimated ICERs.

7. END OF LIFE

The ToE stated that nivolumab meets the end-of-life criteria, i.e. "*the treatment is indicated for patients with a short life expectancy, normally less than 24 months and there is sufficient evidence to indicate that the treatment has the prospect of offering an extension to life, normally of a mean value of at least an additional 3 months, compared with current NHS treatment*".[1, 16] The ERG can confirm that there is no change in OS, however measured, that would suggest that they are not still fulfilled.

8. REFERENCES

[1] National Institute for Health and Care Excellence. *Nivolumab for treating squamous cell carcinoma of the head and neck after platinum-based chemotherapy (TA490): terms of engagement for CDF review*. London: National Institute for Health and Care Excellence, 2018

[2] Bristol-Myers Squibb Pharmaceuticals Ltd. *Nivolumab for treating squamous cell carcinoma of the head and neck after platinum-based chemotherapy (CDF review of TA490) [ID1585]. Cancer Drugs Fund Review of TA490. Company evidence submission for committee*: Bristol-Myers Squibb Pharmaceuticals Ltd., 2020

[3] European Medicines Agency. *Opdivo. Procedural steps taken and scientific information after the authorisation [Internet]*. Amsterdam: European Medicines Agency, 2020 [accessed 21.1.20] Available from: https://www.ema.europa.eu/en/documents/procedural-steps-after/opdivo-epar-procedural-steps-taken-scientific-information-after-authorisation_en.pdf

[4] European Medicines Agency. *Opdivo 10 mg/mL concentrate for solution for infusion: EPAR - Product Information. Annex I. Summary of product characteristics [Internet]*. Amsterdam: European Medicines Agency, 2015 [accessed 21.1.20] Available from: https://www.ema.europa.eu/en/documents/product-information/opdivo-epar-product-information_en.pdf

[5] Bristol-Myers Squibb Pharmaceuticals Ltd. *Nivolumab for treating squamous cell carcinoma of the head and neck after platinum-based chemotherapy (CDF review of TA490) [ID1585]. Cancer Drugs Fund Review of TA490. Response to request for clarification from the ERG*: Bristol-Myers Squibb Pharmaceuticals Ltd., 2020

[6] National Institute for Health and Care Excellence. *Nivolumab for treating recurrent or metastatic squamous-cell carcinoma of the head and neck after platinum-based chemotherapy (Single Technology Assessment): Appraisal consultation committee papers [Internet]*. London: NICE, 25th April 2017 [accessed 24.3.20]. 713p. Available from: https://www.nice.org.uk/guidance/ta490/documents/committee-papers-2

[7] Ferris RL, Blumenschein GR, Fayette J, Guigay J, Colevas AD, Licitra LF, et al. Further evaluations of nivolumab (nivo) versus investigator's choice (IC) chemotherapy for recurrent or metastatic (R/M) squamous cell carcinoma of the head and neck (SCCHN): CheckMate 141. *J Clin Oncol* 2016;34(15 Suppl):6009.

[8] Gillison ML, Blumenschein G, Fayette J, Guigay J, Colevas AD, Licitra L, et al. Nivolumab (nivo) vs investigator's choice (IC) for recurrent or metastatic (R/M) head and neck squamous cell carcinoma (HNSCC): CheckMate-141. Paper presented at the 107th Annual Meeting of the American Association for Cancer Research; 2016 Apr 16-20; New Orleans: United States. *Cancer Res* 2016;76(14 Suppl):CT099.

[9] Public Health England. *Nivolumab for treating recurrent or metastatic squamous-cell carcinoma of the head and neck – data review. Report for the NICE Appraisal Committee - Review of TA490. Commissioned by NHS England and NHS Improvement*. London: Public Health England, 2020

[10] Bristol-Myers Squibb. CheckMate 141: clinical study report for study CA209141 (7th June 2016). 2016.

[11] Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer* 2009;45(2):228-47.

[12] Rupniewska E. *RE: ID1585 Nivolumab in SCCHN cancer (CDF rev TA490): clarification questions [Personal email communication to Nigel Armstrong]*, 12th March 2020 [accessed 26.3.20]

[13] Bristol-Myers Squibb Pharmaceuticals Ltd. *Nivolumab for treating squamous cell carcinoma of the head and neck after platinum-based chemotherapy (CDF review of TA490) [ID1585]. Cancer Drugs Fund Review of TA490. Company evidence submission. Appendices*: Bristol-Myers Squibb Pharmaceuticals Ltd., 2020

[14] Armstrong N, Ramaekers BLT, Pouwels X, Zaim R, Wolff RF, Riemsma RR, et al. *Nivolumab for treating recurrent or metastatic squamous-cell carcinoma of the head and neck after platinum-based chemotherapy: a Single Technology Assessment [Word document]*. York: Kleijnen Systematic Reviews Ltd, 2016. 68p.

[15] National Institute for Health and Care Excellence. *Nivolumab for treating recurrent or metastatic squamous-cell carcinoma of the head and neck after platinum-based chemotherapy [ID971]: Final appraisal determination committee papers [Internet]*. London: NICE, 13th October 2017 [accessed 24.3.20]. 109p. Available from: https://www.nice.org.uk/guidance/ta490/documents/committee-papers

[16] National Institute for Health and Care Excellence. *PMG19 Addendum A - Final amendments to the NICE technology appraisal processes and methods guides to support the proposed new Cancer Drugs Fund arrangements. Technology Appraisal Processes - CDF [Internet]*. London: NICE, 2016 [accessed 12.9.16]. 11p. Available from: https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisals/process-and-methods-guide-addendum.pdf