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

ERRATUM

This document contains errata in respect of the ERG report in response to the company's factual accuracy check. The table below lists the page to be replaced in the original document and the nature of the change:

Page nr:	**Change:**
36	Text changed: "Therefore it appears that not all patients are required to have had at least one line of platinum therapy. This is indicated further by Table 6 of the CS which indicates that a maximum of 60.2% of patients received prior systemic therapies. Therefore, the subgroup of patients from CheckMate 032 used in the CS appears not in accordance with the population defined in the scope. However, this is contradicted by the CSR, which shows 96.2% receipt in any setting."
13	Text changed: "switched to ipilimumab" changed to "switched to ipilimumab plus nivolumab" 275 changed to 270 "BIRC" changed to "investigator assessed"
33	Text changed: "switched to ipilimumab" changed to "switched to ipilimumab plus nivolumab"
16	275 changed to 270
14	"ORR data for BSC was not identified." was changed to "ORR for BSC from one trial (n=85) was found in zero patients." "PFS data from other comparators were not available." changed to "Paclitaxel (one trial, n=65) had a median PFS of 4.1 months (80% CI: 3.0 to 5.6)." "gemcitabine and cisplatin (one trial, n=65) had a median OS of 10.5 months (95% CI: 3 to 22.9)," changed to "gemcitabine and cisplatin (one trial, n=33) had a mean OS of 10.5 months (95% CI: 3 to 22.9), CIC highlighting added.
83 to 86	Tables 4.15 to 4.17 updated with the most recent CheckMate 275 trial results.
114	Removed statement: "Lastly, any adjustment for background mortality should be applied to responder and non-responder groups separately, if response-based analysis is used. However, the company applied it to the combined responder and non-responder groups, which, due to the different prognoses in both groups, is inappropriate. This issue becomes redundant with a conventional, not response-based analysis."
132	Text changed: "in the ERG base-case a missed dose was only assumed in case the length exceeded seven days; resulting in a proportion of unadministered drug doses of 6.6% (CS dose intensity) × 63.4*% (the proportion of dose delays that exceeded 7 days*; averaged for CheckMate 275 and CheckMate 032) = 2.4% (i.e. dose intensity of 95.8%)." Further text changed: "With the PAS, nivolumab treatment resulted in incremental cost effectiveness ratios (ICERs) of £37,646, £44,960 and £38,164 per QALY gained versus paclitaxel, docetaxel and BSC respectively (Table 5.17)."
21	Text changed: "With the PAS, nivolumab treatment resulted in deterministic incremental cost effectiveness ratios (ICERs) of £37,646, £44,960, £38,164, and £71,608 per QALY gained versus paclitaxel, docetaxel, BSC and cisplatin plus gemcitabine respectively."
24	Text changed: "The ERG's concerns on validation include the lack of internal and cross validity efforts as well as sparse use of expert opinion; external validation efforts for nivolumab that are based on a lung cancer study only and therefore questionable in terms of their relevance; the use of only CheckMate 275 for validating model predictions; as well as transparency issues with the model."
26	Text changed to reflect corrected ERG base-case ICERs: "This resulted in ICERs (probabilistic) of £86,030, £67,205 and £68,348 for nivolumab (with PAS) versus docetaxel, paclitaxel and BSC, respectively."
27-28	Table 1.1 updated with corrected probabilistic ERG base-case and exploratory analyses
142	Text changed to reflect corrected ERG base-case ICERs: "This resulted in ICERs (probabilistic) of £86,030, £67,205 and £68,348 for nivolumab (with PAS) versus docetaxel, paclitaxel and BSC, respectively (Table 5.22)."

143	ICERs in text changed to: “For comparison, the deterministic ERG base-case ICERs were £82,028, £64,298 and £66,161 per QALY gained, with cisplatin plus gemcitabine dominating Table 5.22 updated with corrected probabilistic ERG base-case.
145	ICERs in text changed to: “The ERG base-case resulted in ICERs (probabilistic) of £86,030, £67,205 and £68,348 per QALY gained for nivolumab (with PAS) versus docetaxel, paclitaxel and BSC respectively. In the ERG base-case, cisplatin plus gemcitabine dominated nivolumab, with a larger QALY gain and lower costs. For comparison, the deterministic ERG base-case ICERs were £82,028, £64,298 and £66,161 per QALY gained, with cisplatin plus gemcitabine dominating nivolumab.”
146	ICERs and text changed to: “In exploratory analysis, the ERG found that using the naïve comparison resulted in pronounced increases in the ICERs (£90,465, £63,548, dominated, £64,429 per QALY gained when comparing nivolumab against docetaxel, paclitaxel, cisplatin plus gemcitabine and BSC respectively). These further increased in an extreme scenario where no relative treatment effect was assumed for nivolumab. The use of time-independent HRs also had a significant effect on ICERs, with some ICERs increasing and others decreasing compared to the ERG base-case ICERs (£70,452, £94,067, £74,858, £54,707 per QALY gained when comparing nivolumab against docetaxel, paclitaxel, cisplatin plus gemcitabine and BSC respectively). The use of alternative parametric time-to-event models for OS (lognormal) and PFS (log-logistic) in the conventional approach produced further increases in ICERs (£95,759, £78,505, dominated, £77,739 per QALY gained when comparing nivolumab against docetaxel, paclitaxel, cisplatin plus gemcitabine and BSC respectively). Using the response-based analysis with alternative time-to-event models for OS and PFS, but not for TTD, also resulted in a marked increase in ICERs compared with the response-based company’s base-case (£122,716, £96,836, dominated, £94,964 per QALY gained when comparing nivolumab against docetaxel, paclitaxel, cisplatin plus gemcitabine and BSC respectively). Lastly, the alternative landmark drove the company’s base-case ICERs up (£77,167, £73,309, £93,439, £62,903 per QALY gained when comparing nivolumab against docetaxel, paclitaxel, cisplatin plus gemcitabine and BSC respectively).”
147 to 151	Tables 6.1 and 6.2 updated with corrected probabilistic ERG base-case and exploratory analyses
155	ICERs in text changed to: “This resulted in ICERs (probabilistic) of £86,030, £67,205 and £68,348 for nivolumab (with PAS) versus docetaxel, paclitaxel and BSC, respectively.

1. SUMMARY

1.1 Critique of the decision problem in the company's submission

The patient population described in the final scope issued by the National Institute for Health and Care Excellence (NICE) was '*Adults with metastatic or unresectable urothelial cancer whose disease has progressed after platinum-based chemotherapy'*. Nivolumab was to be compared to retreatment with first-line platinum-based chemotherapy (only for people whose disease has had an adequate response), paclitaxel, docetaxel or best supportive care. Outcomes included overall survival (OS), progression free survival (PFS), response rates, adverse effects of treatment and health-related quality of life (HRQoL).

There were several deviations between the decision problem addressed by the company submission and that of the final scope issued by NICE. For the population, the company submission (CS) was in agreement with the scope, although only one of the two pivotal nivolumab trials included patients from the UK. Both nivolumab studies were small (270 and 78 patients for CheckMate 275 and CheckMate 032 respectively); only six patients were from the UK. For the intervention, the CheckMate 275 trial was in line with the scope, but in the CheckMate 032 trial 23% patients switched to ipilimumab plus nivolumab (referred to throughout this document as 'switched to ipilimumab'). For the comparator, both nivolumab trials were single arm studies and therefore no direct or indirect comparators were included. Simulated treatment comparisons (STC) were performed for comparisons of nivolumab to paclitaxel, docetaxel and best supportive care (BSC). Comparisons of nivolumab to cisplatin plus gemcitabine were included only as part of a scenario analysis. The ERG would have considered cisplatin and gemcitabine suitable for inclusion in the STC, especially given the limitations in the quantity and quality of evidence for nivolumab and all other comparator trials. For the outcomes, comparative data in the form of an STC was only provided for OS, PFS and objective response rate (ORR). There were no comparative analyses for adverse events or quality of life.

1.2 Summary of clinical effectiveness evidence submitted by the company

1.2.1 Direct evidence

The company conducted a systematic literature review (SLR) to inform the submission. The aim of the SLR was '*to understand the relative efficacy and safety of nivolumab compared to alternative therapies for adult patients with metastatic or unresectable urothelial cancer whose disease has progressed after platinum-based chemotherapy'*.

The company did not identify any randomised controlled trials (RCTs) for nivolumab. Two ongoing phase I/II single arm studies for nivolumab were identified (CheckMate 275 and CheckMate 032). Therefore no studies were found that directly compared nivolumab with any specified comparator.

Single arm data for nivolumab

Data from the individual trials indicated that for Check Mate 275 (n=270) nivolumab led to a confirmed ORR (BIRC) in 54 (20.0%) patients (95% CI: 15.4 to 25.3). In CheckMate 032 (n=78) nivolumab led to a confirmed ORR (investigator assessed) in 19 (24.4%) patients (95% CI: 15.3–35.4).

For CheckMate 275, at the latest database lock of 2 September 2016 (n=270 analysed), nivolumab led to a median OS of 8.57 months (95% CI: 6.05–11.27) and for CheckMate 032 (n=78) nivolumab led to a median OS of 9.72 months (95% CI: 7.26–16.16).

For CheckMate 275, at the latest database lock of 2 September 2016 (n=270 analysed), nivolumab led to a median PFS of 2.0 months (95% CI: 1.87–2.63) and for CheckMate 032 (n=78) nivolumab led to a median PFS of 2.78 months (95% CI: 1.45–5.85).

Health related-quality of life (HRQoL) data was limited either by the currently available follow-up data or patient numbers.

For CheckMate 275 (May 2016 database lock) 75.6% of patients discontinued treatment with nivolumab (disease progression, 53.3%; adverse events (AEs) unrelated to nivolumab, 12.6%; nivolumab toxicity, 5.2%). For CheckMate 032 (March 2016 database lock) 76.9% of patients discontinued study treatment (disease progression, 64.1%; nivolumab toxicity, 2.6%).

In the CheckMate 275 trial 51.1% of patients died (1.1% attributed to nivolumab toxicity), whilst in CheckMate 032 trial 46.2% of patients died (2.6% attributed to nivolumab toxicity). In the CheckMate 275 trial 64.4% of patients had a drug related AE ([redacted] serious drug related AE), whilst in CheckMate 032 trial 83.3% of patients had a drug related AE (10.3% serious drug related AE).

Data for the CheckMate trials were pooled for the STC but the pooled results or method were not provided, despite a request in the clarification letter.

1.2.2 Indirect evidence

The identification of two single arm studies for nivolumab precluded any conventional mixed treatment comparison (MTC) or indirect meta-analysis. There were no studies that could provide a common comparator to support any indirect comparison or MTC. As a consequence the company decided to perform an unanchored (no common comparator) stimulated treatment comparison (STC).

Single arm data for comparators

Single arm data is provided as an alternative to the STC to allow naive comparisons to the single arm data of nivolumab. Data from the comparator trials indicated that paclitaxel (one trial, n=45) led to overall ORR (definition not reported) in four (9.0%) patients (95% CI: 2 to 21), gemcitabine and cisplatin (two trials, n=53) led to ORR (not defined) in 13 (39.4%) to eight (40.0%) patients (95% CI: NR), docetaxel and placebo (one trial, n=72) led to confirmed ORR (overall PR or CR) in eight (7.1%) patients (95% CI: NR) and docetaxel (one trial, n=45) led to ORR (best overall PR or CR) in four (8.9%) patients (95% CI: 2.5 to 21.2). ORR for BSC from one trial (n=85) was found in zero patients.

BSC (one trial, n= 117) had a median OS of 4.6 months (95% CI: 4.1 to 6.6), paclitaxel (one trial, n=65) had a median OS of eight months (80% CI: 6.9 to 9.7), gemcitabine and cisplatin (one trial, n=65) had a mean OS of 10.5 months (95% CI: 3 to 22.9), docetaxel and placebo (one trial, n=72) had a median OS of 7.03 months (95% CI: 5.19 to 10.41) and docetaxel (one trial, n=45) had a median OS of 9.2 months (95% CI: 5.7 to 11.7).

Docetaxel and placebo (one trial, n=72) had a median PFS of 1.58 months (95% CI: 1.48 to 3.09) and docetaxel (one trial, n=45) had a median PFS of 2.8 months (95% CI: 1.9 to 3.6). PFS data from other comparators were not available.

Simulated treatment comparison The STC approach uses nivolumab IPD to attempt to model how patients might respond to treatment if they were more like those in a comparator trial based on key baseline characteristics. A prediction model is intended to adjust the difference in outcomes observed between the nivolumab and comparator studies given the high risk of bias that must exist in comparing observational data. The

The analysis based on the STC and using a fixed effect FP model of PFS with P1=0 AND P2=0 was only possible for nivolumab compared to paclitaxel or compared to docetaxel. For PFS nivolumab was statistically superior to: paclitaxel at time points between 20 to 72 weeks (HR 7.26, 95% CrI 1.40 to 28.85, 68 to 72 weeks); docetaxel at time points between 8 to 12 weeks only (HR 1.72, 95% CrI 1.18 to 2.49).

The STC analysis of ORR using a fixed effect model found that nivolumab is significantly better than BSC (OR 106.70, 95% CrI 6.72 to 49820) or docetaxel (OR 3.12, 95% CrI 1.06 to 9.49), although the uncertainty was large. No significant differences were found for nivolumab compared to paclitaxel or gemcitabine plus cisplatin. In the random effects model nivolumab was only statistically superior to BSC (OR 108.1, 95% CrI 4.17 to 52240).

No formal comparison of AEs including no evidence synthesis was performed. However, the rate of neutropaenia was generally lower than for most comparators, the exception being BSC, and much lower than for cisplatin and gemcitabine. The rate for anaemia was a little lower except for being much lower than BSC and even lower again in comparison to cisplatin and gemcitabine. For leaukopaenia the rate was comparable i.e. 0% between all comparators where it was reported except against cisplatin plus gemcitabine. The rate of asthaenia was also lower than all comparators except cisplatin plus gemcitabine.

1.3 Summary of the ERG's critique of clinical effectiveness evidence submitted

The company submission (CS) and response to clarification provided sufficient details for the ERG to appraise the searches for eligible trials. Searches were carried out in accordance with the NICE guide to the methods of technology appraisal sections 5.2.2 and 5.2.4 using a good range of databases. Additional searches of conference proceedings were reported, along with trials registers and the checking of reference lists of existing systematic reviews and health technology assessments (HTAs). The systematic review was performed to a good standard.

The ideal scenario to determine the relative benefits of nivolumab and its comparators would be a series of RCTs comparing nivolumab to its comparators. Failing this, a network meta-analysis of RCTs using a set of common comparators would be the preferred approach. However the submission relies on two single arm studies of nivolumab, which are entered into a STC together with the single arms of comparator studies. Single arm studies are basically observational studies and are considered low order for study quality. The methods used by the company to conduct the STC largely follow those described in NICE DSU TSD 18, but, as stated in the same TSD, given no comparative data (unanchored analysis) the results obtained should be treated with caution. The ERG found the following limitations in the STC analysis:

1. There was no STC analysis for AEs or HRQoL. Therefore the value of any potential extension to life cannot be judged in relation to any changes to the patients' quality of life.
2. The analysis relies on two small single arm nivolumab studies, one includes 78 patients and the other included 270. Therefore any statistical analyses have increased uncertainty due to the small sample size.
3. The numbers of patients are small for all comparator studies (33 to 117) and not all studies provided data for all outcomes.
4. There were no common comparators; therefore an unanchored STC had to be performed.
5. The company pooled the two nivolumab trials despite each one using different methods of outcome assessment, CheckMate 275 using BIRC and CheckMate 032 using investigator-assessed. The results of this pooling (and its variability) were not reported.

2. CRITIQUE OF COMPANY'S DEFINITION OF DECISION PROBLEM

Table 3.1: Statement of the decision problem (as presented by the company)

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG comment
Population (s)	Adults with metastatic or unresectable urothelial cancer whose disease has progressed after platinum-based chemotherapy	Adults with metastatic or unresectable urothelial cancer whose disease has progressed after platinum-based chemotherapy	NA	CheckMate 275 was in line with the scope of the decision problem, but no patients were included from the UK. CheckMate 032 included a small proportion of patients who had not received platinum-based chemotherapy; only 8% patients were from the UK.
Intervention	Nivolumab	Nivolumab	NA	CheckMate 275 investigated nivolumab, however CheckMate 032 investigated nivolumab monotherapy, but 23% switched to ipilimumab plus nivolumab
Comparator (s)	Retreatment with first-line platinum-based chemotherapy (only for people whose disease has had an adequate response) Paclitaxel Docetaxel Best supportive care	Paclitaxel Docetaxel Best supportive care	No data on retreatment with first-line platinum-based chemotherapy was identified in the clinical systematic literature review (SLR). However, the use of retreatment is limited to <10% of patients and is not a primary comparator for nivolumab in UC after platinum-based chemotherapy. Data from a trial involving cisplatin plus gemcitabine after the failure of MVAC (methotrexate, vinblastine,	Both included trials were single arm studies and therefore no direct or indirect comparators were included. Given the paucity of data generally the ERG believes evidence for all specified NICE comparators should have been included in the STC.

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG comment
			doxorubicin and cisplatin) was identified and included as a scenario analysis, in the absence of clinical data to inform a comparison of nivolumab versus retreatment.	
Outcomes	The outcome measures to be considered include: overall survival progression-free survival response rates adverse events of treatment health-related quality of life	The outcome measures considered include: overall survival progression-free survival response rates (objective response rate, duration of response) adverse events of treatment health-related quality of life (via the EORTC QLQ-C30 and the EQ-5D-3L)	N/A	The ERG notes that comparative data in the form of an STC was only provided for overall survival, progression free survival and objective response rate. There was no formal comparison for adverse events or quality of life.
Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. Costs will be considered from an NHS and Personal Social Services perspective.	The cost effectiveness of treatments are expressed in terms of incremental cost per quality-adjusted life year. A lifetime time horizon was adopted to capture all relevant costs and health-related utilities. All costs and utilities were discounted at a rate of 3.5% per year in alignment with the NICE guide to the methods of technology appraisal.	N/A	N/A

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG comment
		Costs were considered from an NHS and Personal Social Services perspective.		
Subgroups to be considered	If appropriate, the appraisal should include consideration of the costs and implications of additional testing for biological markers, but will not make recommendations on specific diagnostic tests or devices. Guidance will only be issued in accordance with the marketing authorisation. Where the wording of the therapeutic indication does not include specific treatment combinations, guidance will be issued only in the context of the evidence that has underpinned the marketing authorisation granted by the regulator.	No subgroup analysis was undertaken.	The effect of nivolumab in relation to baseline tumour PD-L1 expression status was investigated as part of the pivotal clinical trials informing the clinical evidence base for nivolumab within this submission. However, the link between baseline tumour PD-L1 expression status and the efficacy of PD-1/PD-L1 targeting agents is yet to be fully established and the testing methodologies of PD-L1 expression status are yet to be fully validated; as such, no formal subgroup analyses have been presented within this submission. This is in line with the marketing authorisation for nivolumab which is not restricted based on PD-L1 expression status.	The company was requested in the clarification letter to perform these subgroup analyses in the STC, but declined to do so arguing that data on PD-L1 expression was not available in the comparator trials.[7]
Special considerations including issues related to equity or equality	None detailed.	Treatment access being available only via clinical trials currently represents an inequality for some patients.	The availability of a nationally funded treatment option on the NHS would help to move towards addressing this equity issue.	No comment.

Source: CS, Table 1, page 11-13.

CR = complete response; N.A.= not applicable; ORR = objective response rate; PR = partial response; PD-L1: programmed death-ligand 1; STC simulated treatment comparison

3.1 Population

The population defined in the scope is: 'Adults with metastatic or unresectable urothelial cancer whose disease has progressed after platinum-based chemotherapy'.[6]
The licensed indication for nivolumab is: *'Nivolumab (Opdivo®) is indicated for the treatment of locally advanced unresectable or metastatic urothelial carcinoma in adults after failure of prior platinum-containing chemotherapy'* (CS, page 16).'[2]

The submission relies on two single arm studies, the CheckMate 275 trial[8] and the CheckMate 032 trial.[9] Examination of the inclusion criteria for these trials indicated that the CheckMate 275 trial included patients with metastatic or surgically unresectable transitional cell carcinoma of the urothelium (bladder, urethra, ureter, or renal pelvis). Patients have progression or recurrence after treatment with at least one platinum-containing chemotherapy regimen or within 12 months of peri-operative treatment with a platinum agent in the setting of cystectomy for localized muscle-invasive UC. Patients must have an ECOG performance status of 0 or 1.[10] Therefore the ERG considers this a good match with regards to the final scope. However, none of the patients included in this trial were from the UK.

CheckMate 032 included patients with histologically confirmed locally advanced or metastatic disease of one of the following tumour types: triple negative breast cancer, gastric cancer, pancreatic cancer, small cell lung cancer, bladder cancer, ovarian cancer. Patients must have an ECOG performance status of 0 or 1.[11] Prior chemotherapy was not stipulated as an inclusion criterion and reading Appendix 3.8 of the Checkmate 032 CSR indicated that a proportion of patients did not previously receive a platinum-based chemotherapy. For the purposes of the CS *'a subgroup of the enrolled population in this trial is of relevance to this submission: the cohort of patients enrolled to receive nivolumab monotherapy for the treatment of locally advanced unresectable or metastatic UC who had progressed after at least one previous line of platinum-containing chemotherapy (n=86).'*(CS section B.2.2)[2] In Table 5 of the CS, previous platinum based therapies are found in two of three inclusion criteria for progression or recurrence, the third criteria states 'refusal of standard treatment with chemotherapy'. Therefore it appears that not all patients are required to have had at least one line of platinum therapy. This is indicated further by Table 6 of the CS which indicates that a maximum of 60.2% of patients received prior systemic therapies. Therefore, the subgroup of patients from CheckMate 032 used in the CS appears not in accordance with the population defined in the scope. However, this is contradicted by the CSR, which shows 96.2% receipt in any setting. In addition, only 6/78 (8%) of bladder cancer patients in CheckMate 032 were from the UK.

3.2 Intervention

The intervention is in line with the scope. The intervention described in the scope is 'Nivolumab'. The CS describes the recommended dose and schedule of nivolumab monotherapy in urothelial carcinoma as follows: '3 mg/kg administered as IV infusion over 60 minutes every 2 weeks (Q2W), which is consistent with the existing approved dose and schedule of nivolumab monotherapy in adults in other indications.' (CS, page 17).[2] Dose escalation or reduction is not recommended; dosing delay or discontinuation may be required based on individual safety and tolerability.

A marketing authorisation application for nivolumab was submitted to the European Medicines Agency (EMA) on the 25 August 2016. A positive opinion from the Committee for Medicinal Products for Human Use (CHMP) was received on the 21 April 2017. Full marketing authorisation was received from the EMA on Monday 5 June 2017.[12]

clarification letter.[2, 7]The results for the individual nivolumab trials were added to tables 4.15 to 4.17 to provide a comparison, in the absence of the pooled data.

Table 4.2: Overall survival in studies included in the simulated treatment comparison

Trial ID	Treatment arm	Population assessed (n)	Survival definition	Survival median (CI)
Sharma et al. (2017)[8] CheckMate 275	Nivolumab	270	From first dose and last known date alive or death	8.57 (6.05–11.27)
Sharma et al. (2016)[9] CheckMate 032	Nivolumab	78	From first dose and last known date alive or death	9.7 (95% CI 7.3 to 16.2)
Bellmunt et al. (2009)[26]	BSC	117	NR	4.6 (95% CI 4.1 to 6.6)
Choueiri et al. (2012)[27]	Docetaxel and placebo	72	From date of random assignment until date of death	7.03 (95% CI 5.19 to 10.41)
Jones et al. (2017)[15]	Paclitaxel	65	From the date of randomisation	8 (80% CI 6.9 to 9.7)
Petrylak et al. (2016)[16]	Docetaxel	45	The time from random assignment to death resulting from any cause	9.2 (95% CI 5.7 to 11.7)
Gondo et al. (2011)[13]	Gemcitabine and cisplatin	33	OS was measured from the start of the gemcitabine-cisplatin regimen until the date of death or the last follow-up.	10.5 (95% CI 3 to 22.9)
Joly et al. (2009)[28]	Paclitaxel	Outcome not reported		
Ozawa et al. (2007)[14]	Gemcitabine and cisplatin	Outcome not reported		

Source: Tables 24 and 27 of CS Appendix D
BSC = best supportive care; CI = confidence interval; NR = not reported; OS = overall survival

Table 4.3: Progression-free survival in studies included in the simulated treatment comparison

Trial ID	Treatment arm	Population assessed (n)	PFS definition	PFS median (CI)
Sharma et al. (2017)[8] CheckMate 275	Nivolumab	270	Time from first dosing date to the date of the first documented tumour progression, based on BIRC assessments (per RECIST 1.1), or death due to any cause.	2.00 (95% CI 1.87 to 2.63)
Sharma et al. (2016)[9] CheckMate 032	Nivolumab	78	Time from treatment assignment to the date of the first documented tumour progression, as determined by the investigator (per RECIST 1.1), or death due to any cause.	2.78 (95% CI 1.45 to 5.85)
Bellmunt et al. (2009)[26]	BSC	Outcome not reported		
Choueiri et al. (2012)[27]	Docetaxel and placebo	72	Time between random assignment and documented progression per RECIST criteria or death.	1.58 (95% CI 1.48 to 3.09)
Jones et al. (2017)[15]	Paclitaxel	65	NR	4.1 (80% CI 3 to 5.6)
Petrylak et al. (2016)[16]	Docetaxel	45	The time from random assignment until the first radiographic documentation of objective progression defined by RECIST v1.1 or death resulting from any cause	2.8 (95% CI 1.9 to 3.6)
Gondo et al. (2011)[13]	Gemcitabine and cisplatin	Outcome not reported		
Joly et al. (2009)[28]	Paclitaxel	Outcome not reported		
Ozawa et al. (2007)[14]	Gemcitabine and cisplatin	Outcome not reported		

Source: Table 25 of CS Appendix D
BSC = best supportive care; CI = confidence interval; NR = not reported; PFS = survival

Table 4.4: Objective response rate in studies included in the simulated treatment comparison

Trial ID	Treatment arm	Population assessed (n)	OR definition	Observed cases, n (%) (CI)
Sharma et al. (2017)[8] CheckMate 275	Nivolumab	270	The best response designation, as determined by BIRC, recorded between the date of first dose and the date of objectively documented progression per RECIST v1.1 or the date of subsequent therapy.	54 (20.0) (95% CI 15.4 to 25.3)
Sharma et al. (2016)[9] CheckMate 032	Nivolumab	78	Best overall response (BOR) of complete response (CR) or partial response (PR) divided by the number of treated subjects, as determined by the investigator. Assessment of ORR in accordance with RECIST 1.1. Recorded between the date of treatment assignment and documented progression or the start date of subsequent anti-cancer therapy.	19 (24.2) (95% CI 15.3 to 35.4)
Bellmunt et al. (2009)[26]	BSC	85	NR	0 (NR)
Choueiri et al. (2012)[27]	Docetaxel and placebo	72	The percentage of participants who achieved a confirmed overall PR or CR using RECIST criteria on treatment. Patients without measurable disease only at baseline are included, based on status of non-target lesions.	8 (7.1) (NR)
Jones et al. (2017)[15]	Paclitaxel	Outcome not reported		
Petrylak et al. (2016)[16]	Docetaxel	45	Objective response: defined as the proportion of patients with a best overall response of complete or partial.	4 (8.9) (95% CI 2.5 to 21.2)
Gondo et al. (2011)[13]	Gemcitabine and cisplatin	33	NR	13 (39.4) (NR)

Trial ID	Treatment arm	Population assessed (n)	OR definition	Observed cases, n (%) (CI)
Joly et al. (2009)[28]	Paclitaxel	45	Overall ORR – not further defined	4 (9) (95% CI 2 to 21)
Ozawa et al. (2007)[14]	Gemcitabine and cisplatin	20	Objective response – not further defined	8 (40) (NR)

Source: Tables 24 and 27 of CS Appendix D
BSC = best supportive care; CI = confidence interval; CR = complete response; NR = not reported; ORR = objective response rate; PR = partial response

Background mortality

After 88 weeks, general population mortality estimates were used to adjust OS and PFS estimations. This was implemented in order to '*appropriately characterise the relationship between age and increasing risk of death.*'[2] To avoid double-counting, general population mortality estimates were applied from the 88th week onwards, which represented the end of the CheckMate 032 and CheckMate 275 studies' follow-up. This adjustment was implemented by multiplying the survival estimates obtained from the parametric time-to-event model estimating OS (described in previous sections) by the probability of being alive according to age-adjusted UK life tables.

ERG comment: The ERG's comments relate to (1) an error in the calculation of background mortality, (2) the use of an age distribution to calculate background mortality, and (3) the implementation of adjusting OS and PFS by background mortality.

(1) When reviewing the cost effectiveness model, the ERG noted that the mortality rates implemented in the model did not match the values reported by the Office of National Statistics UK life tables. The ERG therefore used the correct age-adjusted background mortality rates and fixed the conversion of the background mortality rate into a probability.

(2) Not in line with conventional methods of incorporating background mortality in parametric survival models, the company used a distribution of age instead of a fixed mean age, to reflect patient heterogeneity. This resulted in slightly higher background mortality compared to standard background mortality estimates. Despite this being unconventional in cohort models, the ERG considers that it is appropriate to reflect patient heterogeneity in the calculation of background mortality.

(3) The conventional approach seen in many technology appraisals is to implement a maximum function to incorporate general UK population mortality data in the cost effectiveness model, to ensure that the probability of dying does not become lower than the probability of dying based on the age-adjusted UK life tables. However, the company's approach of implementing this background mortality by multiplying OS by the probability of being alive based on the age-adjusted UK life tables, was viewed as appropriate.

5.2.6.2 Relative effectiveness of nivolumab

The relative effectiveness of nivolumab versus the comparators was modelled through time-varying hazard ratios (HRs) because the '*proportional hazard assumption did not hold for these comparisons given the unique mechanism of action for nivolumab*'.[2] No evidence was provided to support the violation of the proportional hazard assumption. A STC was performed to obtain these time-varying HRs. More detail about this methodology is provided in Section 4.4.1. The STC was performed based on the pooled CheckMate 032 and CheckMate 275 trials dataset, in which response status was not taken into account. The HRs obtained from the STC were then applied to the combined parametric time-to-event models of nivolumab which took response status into account. Figures 5.8 to 5.9 present the survival curves estimating OS and PFS of each comparator, obtained by applying the time-varying HRs to the combined survival curves of nivolumab (Figures 5.10 and 5.11), compared to the Kaplan-Meier estimates observed in the comparator studies. The company explained that the predicted OS and PFS of the comparators were mostly lower than the observed OS and PFS, especially for

Given the lack of clarity and justification for the AE unit costs reported in CS Table 41, the alternatively calculated AE unit costs, based on ID971, were used in the ERG exploratory analyses.

(6) In the CS it is stated that '*In UK clinical practice, cisplatin plus gemcitabine is given in the first-line setting as gemcitabine (1250mg/m2) plus cisplatin (70mg/m2) on days 1 and 8 of a 21 day cycle (cisplatin on day 1 only)*'.[2] However, in response to clarification question B17.E[7] the company responded that, in the economic model, it assumed the administration regimen with gemcitabine on days 1, 8 and 15 and cisplatin on days 1 and 2. This was based on the administration regimen from the Gondo (2011) study[13] and justified by stating that this study was the key source for efficacy data. The ERG performed scenario analyses incorporating the cisplatin + gemcitabine administration scheme that is likely applicable to UK clinical practice.

(7) In response to clarification question B17.B[7] the company stated that dose delays that exceed the duration of a nivolumab treatment cycle (i.e. 14 days) can reasonably be assumed to be missed. Hence, the company assumed that all delayed doses were missed doses. This seems reasonable to the ERG if all dose delays exceed the duration of a nivolumab treatment cycle. However, it is highly questionable whether this is applicable to all dose delays. Particularly given that the length of dose delays was less than one week in 34.6% and 38.5% of all delayed doses for CheckMate 275 and CheckMate 032 and the large majority of dose delays (71.7% and 80.8% respectively) does not exceed the duration of a nivolumab treatment cycle[10, 11]. Therefore, in the ERG base-case a missed dose was only assumed in case the length exceeded seven days; resulting in a proportion of unadministered drug doses of 6.6% (CS dose intensity) × 63.4*% (the proportion of dose delays that exceeded 7 days*; averaged for CheckMate 275 and CheckMate 032) = 2.4% (i.e. dose intensity of 95.8%).

(8) The calculated dose intensity of 93.4% for nivolumab was assumed to be applicable for the comparators; assuming that 6.6% of the doses would be missed. In response to clarification question B17.C[7], the company stated that this was assumed in absence of evidence. In addition, the company stated that assuming no dose intensity for the comparators would induce bias in favour of nivolumab.[7] However, the ERG questions whether the current approach (assuming a dose intensity of 93.4% for all comparators) does not induce bias in favour of nivolumab as well. Particularly considering the AE occurrence that was used for the comparators (Table 5.7), it is not unlikely that that the number of missed doses is higher for (some of) the comparators than for nivolumab. Hence the drug costs for the comparators might be overestimated.

5.2.10 Cost effectiveness results

In the deterministic base-case analysis, nivolumab was associated with larger QALY and LY gains than docetaxel, paclitaxel and BSC (Table 5.15). The main benefit of nivolumab versus these comparators stemmed from QALY gains post-progression ([redacted], [redacted] and [redacted] of incremental QALYs in post-progression health state for the comparisons with docetaxel, paclitaxel and BSC respectively). Compared with cisplatin plus gemcitabine, nivolumab's incremental QALYs were increased in pre-progression and decreased in post-progression.

Nivolumab also induced larger life time costs than docetaxel, paclitaxel and BSC. Incremental costs mainly stemmed from higher treatment costs ([redacted]), which reflect the technology costs of nivolumab, and to a minor degree stemmed from higher costs in the post-progression health state ([redacted]) (Table 5.16). With the PAS, nivolumab treatment resulted in incremental cost effectiveness ratios (ICERs) of £37,646, £44,960 and £38,164 per QALY gained versus paclitaxel, docetaxel and BSC respectively (Table 5.17).

In the deterministic base-case analysis, nivolumab was associated with larger QALY and LY gains and costs than docetaxel, paclitaxel, and BSC. With the PAS, nivolumab treatment resulted in deterministic incremental cost effectiveness ratios (ICERs) of £37,646, £44,960, £38,164, and £71,608 per QALY gained versus paclitaxel, docetaxel, BSC and cisplatin plus gemcitabine respectively.

Probabilistic sensitivity analysis (PSA) and deterministic sensitivity analysis (DSA) were undertaken and presented by the company. Patient age, weight and BSA, costs, resource use, utilities, TTD, PFS and OS were varied but relative effectiveness estimates were not included in these analyses. The PSA with 1,000 iterations resulted in ICERs of £54,220, £46,209, £44,698 and £103,568 per QALY gained for nivolumab versus docetaxel, paclitaxel, BSC and cisplatin plus gemcitabine The company reasoned that the PSA ICER increases were mainly driven by a reduction in PFS and OS in the PSA (compared with the deterministic analysis), but did not provide further insights into the mechanism by which this occurred.

1.5 Summary of the ERG's critique of cost effectiveness evidence submitted

Systematic literature review

The cost effectiveness searches in the CS were well documented and easily reproducible, and were carried out in line with the NICE guide to the methods of technology appraisal, using a good range of databases. Additional searches of conference proceedings and organisational websites were reported, along with the checking of reference lists of existing systematic reviews, meta-analyses and health technology assessments.

Model structure and main modelling decisions

The choice of partitioned survival analysis for this decision problem is in line with other appraisals in metastatic cancer, but it should be noted that the recent NICE DSU TSD 19 advocates for alternative model structures that can more accurately reflect interdependent survival functions and use transition probabilities for each possible transition between health states. Another criticism relates to the company's response-based analysis, which if deemed appropriate, should have been incorporated in the model via separate responder and non-responder health states. The ERG considers the adopted perspective, time horizon and discounting to be appropriate for this appraisal.

The patient population used in the model was deemed consistent with the population of the CheckMate 275 and CheckMate 032 studies, as well as the final scope issued by NICE for this appraisal. The company did not provide the comparison of nivolumab with cisplatin plus gemcitabine in the base-case, despite it being in the scope and despite ERG request. The company justified this by citing expert opinion that the population in the only available cisplatin plus gemcitabine study differed from the UK population in that the study population received MVAC in first line instead of cisplatin plus gemcitabine. The ERG considered this to be challengeable in that patients in the cited study would have had exposure to platinum-based therapy and that the precise combination of first-line treatment or naivety to gemcitabine might therefore be irrelevant. Furthermore, a relevant comparator should not be excluded based on issues with the data.

Treatment effectiveness, relative effectiveness and TTD

One of the main issues was that it was unclear whether pooling both CheckMate 032 and CheckMate 275 trials was appropriate and how this was done. The company failed to provide further details upon the ERG's request.

is a subtype of leukopenia. There was an inconsistency in that not all included adverse events matched the inclusion criteria of having an incidence of ≥5%.

Health-related quality of life

The ERG identified several inconsistencies and choices lacking justification in the handling of health-related quality of life estimates. The main issues include inconsistencies in reported observations, the use of utilities derived only from CheckMate 275, the imputation of immature data, the use of multiple imputation instead of the mixed model to adjust for missing data, and inconsistencies in disutilities for adverse events with those used for a previous nivolumab appraisal.

Resource use and costs

Estimation of resource use and costs included a technical error in calculating the dose intensity; inconsistencies in using the average weight and BSA from CheckMate 275 (not using CheckMate 032) and in using the subsequent treatment proportions from CheckMate 275 (not using CheckMate 032). Further inconsistencies related to not using cost and resource use data from TA272 (identified in the SLR), and using different AE unit costs compared with a previous nivolumab appraisal. Some assumptions lacked justification, such as the assumption of an administration scheme that is inconsistent with UK clinical practice for cisplatin plus gemcitabine, the assumption that all delayed doses are missed doses for calculating nivolumab dose intensity, and assuming that the dose intensity for the comparators is equal to that of nivolumab.

Cost effectiveness results

Cost effectiveness results were not presented for one comparator identified in the scope (cisplatin plus gemcitabine) in the base-case. In their sensitivity analyses, the company did not explore important parameters regarding relative effectiveness. The number of iterations (1,000) used in the PSA was shown to not yield stable results. The company subsequently provided a PSA with 10,000 simulations, but this still did not achieve stability. Furthermore, there were marked differences between the deterministic and probabilistic results in the company's base-case, which the company did not provide explanation for. These differences were largely resolved by removing response-based analysis. The PSA did not include relative effectiveness estimates, but it did include inappropriate parameters, such as patient characteristics (age, weight) and comparator treatment costs. The company justified the exclusion of hazard ratios from the PSA by stating that sampling the time-dependent hazard ratios in each period independently would yield counter-intuitive results. However, it is possible to circumvent this problem, for example, by using a fixed set of random numbers. Because relative effectiveness estimates are by far the largest contributor to decision uncertainty, the PSA was deemed to be insufficient.

The ERG's concerns on validation include the lack of internal and cross validity efforts as well as sparse use of expert opinion; external validation efforts for nivolumab that are based on a lung cancer study only and therefore questionable in terms of their relevance; the use of only CheckMate 275 for validating model predictions; as well as transparency issues with the model.

1.6 ERG commentary on the robustness of evidence submitted by the company

1.6.1 Strengths

The searches for eligible studies in the CS were well documented and easily reproducible. Searches were carried out on a good range of databases. Supplementary searches of conference proceedings, and clinical trials registers, and the checking of references lists were undertaken by the company in order to identify additional studies not retrieved by the main searches.

effectiveness; the use of Kaplan-Meier estimates for the period up to the landmark instead of fitting a parametric curve until then, which may result in overfitting; increased uncertainty resulting from fitting parametric models due to decreased sample size; and the combination of responder and non-responder groups using a weighted average, with the weight being the proportion of responders at the landmark, which was held constant. If a response-based analysis is used, this should translate into separate responder and non-responder health states in the model, with differential estimation of relative effectiveness, TTD, HRQoL and resource use and costs. There is therefore an inconstancy in using such an analysis without including these health states. Furthermore, alternative methods to the employed landmark analysis are recommended in NICE DSU TSD 14, but these were not considered by the company.

With respect to the relative effectiveness, the company ruled out proportionality of hazards between responders and non-responders without sufficient justification. OS and PFS estimates derived using the pooled CheckMate studies and response-based analysis were not validated by clinical experts, posing a non-adherence to TSD 14 recommendations. This is of even greater concern because (1) best statistical fit was not the only criterion used for selecting the parametric time-to-event models and (2) model predictions using the response-based approach were significantly different from model predictions using the conventional approach. The application of hazard ratios to an artificially created a posteriori mixed responder and non-responder population while these were derived from the a priori Checkmate matched population poses an inconsistency. The use of time-dependent HRs was not appropriately justified and potentially caused over-parameterisation. Assumptions around the relative effectiveness of nivolumab versus cisplatin plus gemcitabine and BSC in terms of PFS were not supported by clinical evidence. The parameterisation of the fractional polynomial model contributed significant uncertainty, which was not sufficiently explored.

There were inconsistencies in resource use, costs and disutilities associated with adverse events compared with a previous nivolumab appraisal.

Uncertainty caused by the many modelling assumptions was not appropriately explored in deterministic and probabilistic sensitivity analyses. The PSA did not include the, perhaps, most influential and uncertain relative effectiveness parameters.

1.7 Summary of exploratory and sensitivity analyses undertaken by the ERG

A number of issues were identified by the ERG. The ERG was able to adjust/correct some of these in its base-case. This resulted in ICERs (probabilistic) of £86,030, £67,205 and £68,348 for nivolumab (with PAS) versus docetaxel, paclitaxel and BSC, respectively. Cisplatin plus gemcitabine dominated nivolumab.

Additional sensitivity analyses were performed to examine the potential impact of alternative assumptions on the cost effectiveness estimates. These included two scenario analyses: a) exploratory analyses performed using the ERG base-case, and b) exploratory analyses performed using the ERG base-case, except that a response-based approach was used.

The company's and ERG base-case results as well as those scenario analyses with the largest influence on the ICERs are shown in Table 1.1. The uncertainty about the treatment and relative effectiveness evidence is characterised by scenarios A.3 (using a naïve treatment comparison), which increases the ICERs. Using alternative parametric time-to-event models within the ERG base-case can decrease the ICERs significantly (A.1). Finally, using the response-based (B.1) approach significantly decreases the ICER, but these ICERs can increase significantly with the use of best-fitting parametric

time-to-event models (B.3). In addition to these exploratory analyses, the ERG also demonstrated that alternative parameter values informing the fractional polynomial model for the NMA could have a vast impact on the ICERs.

Table 1.5: Scenario analyses with significant impact on ICERs

	Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Nivolumab ICER (£/QALY)
Probabilistic Company base-case[a]	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£12,748	0.82	[redacted]	[redacted]	£54,131
	Paclitaxel	£14,186	0.71	[redacted]	[redacted]	£45,482
	Cis+gem	£30,443	1.34	[redacted]	[redacted]	£100,417
	BSC	£8,811	0.57	[redacted]	[redacted]	£44,873
ERG base-case	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£12,540	0.74	[redacted]	[redacted]	£86,030
	Paclitaxel	£13,905	0.63	[redacted]	[redacted]	£67,205
	Cis+gem	£29,284	1.24	[redacted]	[redacted]	Dominated
	BSC	£8,741	0.56	[redacted]	[redacted]	£68,348
Alternative parametric TTE models (lognormal for OS, log-logistic for PFS) (A.1)[b]	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£11,696	0.66	[redacted]	[redacted]	£95,759
	Paclitaxel	£13,688	0.59	[redacted]	[redacted]	£78,505
	Cis+gem	£28,094	1.10	[redacted]	[redacted]	Dominated
	BSC	£8,611	0.52	[redacted]	[redacted]	£77,739
Naïve comparison data instead of STC results (A.3)[b]	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£12,959	0.77	[redacted]	[redacted]	£90,465
	Paclitaxel	£13,850	0.60	[redacted]	[redacted]	£63,548
	Cis+gem	£30,716	1.56	[redacted]	[redacted]	Dominated
	BSC	£8,588	0.52	[redacted]	[redacted]	£64,429
Response-based analysis (B.1)[c]	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£12,919	0.85	[redacted]	[redacted]	£53,937
	Paclitaxel	£14,198	0.73	[redacted]	[redacted]	£45,466
	Cis+gem	£31,662	1.40	[redacted]	[redacted]	£108,156
	BSC	£8,838	0.60	[redacted]	[redacted]	£44,600

	Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Nivolumab ICER (£/QALY)
Response-based analysis using alternative TTE models for OS, PFS and TTD (B.3)[c]	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£12,507	0.77	[redacted]	[redacted]	£75,916
	Paclitaxel	£13,978	0.68	[redacted]	[redacted]	£66,008
	Cis+gem	£29,779	1.25	[redacted]	[redacted]	£140,296
	BSC	£8,699	0.55	[redacted]	[redacted]	£62,998

Note: [a] results have been reproduced by the ERG, based on the economic model submitted by the company in their clarification response; [b] using the ERG base-case ; [c] using ERG base-case except the change to conventional, not response-based approach
ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year

Additionally, exploratory sensitivity analyses were performed by the ERG to examine the potential impact of alternative assumptions on the cost effectiveness estimates.

The ERG's base-case:

Fixing errors

1. Error in the use of UK life tables and conversion of background mortality rate to probability

   The ERG corrected the error.
2. Error in calculating dose intensity
   The ERG corrected the error by applying dose intensity after calculating the number of vials per weight category, instead of before.

Fixing violations

3. Exclusion of cisplatin plus gemcitabine from base-case and fully incremental analysis in PSA.

   The ERG added cisplatin plus gemcitabine to the base-case and fully incremental analysis in the PSA.
4. Calculation of responder and non-responder proportions for response-based TTD analysis based on OS and PFS, thereby double-counting patients.
   The ERG used only OS to calculate the responder and non-responder proportions used for response-based TTD analysis.
5. Adverse events with an incidence <5% were included in the model, despite the company stating that these should be excluded.
   The ERG removed adverse events with an incidence <5% from the analysis.
6. Use of utilities from CheckMate 275 only.
   The ERG employed the pooled utility estimates from both CheckMate 275 and 032 studies.
7. Use of BSA and weight from CheckMate 275 only.
   The ERG employed the pooled weight from CheckMate 275 and 032, but, due to BSA data not being available from CheckMate 032, kept the BSA estimate from CheckMate 275 only. It should be noted that the re-calculation of weight categories was based on the pooled mean only, the standard deviation was unchanged.
8. Inappropriate parameters in PSA: Patient characteristics were included in the PSA, although they are considered first order uncertainty and typically not reflected in cohort model PSAs. Comparator treatment costs were included in the PSA, but are not typically included.
   The ERG removed patient characteristics and comparator treatment costs from the PSA.

Matters of judgment

9. Use of response-based analysis, without sufficient justification and despite it introducing additional uncertainty.

   The ERG used a not response-based, conventional, survival analysis in its base-case, making redundant the choice of a landmark and retaining the same parametric time-to-event models as chosen by the company (goodness-of fit suggests it is second for OS and first or second for PFS).
10. The assumption that all delayed doses are missed doses.
    The ERG assumed only doses delayed by 7 days or more to be missed doses.

5.3.1 Probabilistic ERG base-case

The ERG performed a PSA to obtain the ERG base-case incorporating all abovementioned adjustments. This resulted in ICERs (probabilistic) of £86,030, £67,205 and £68,348 for nivolumab (with PAS) versus docetaxel, paclitaxel and BSC, respectively (Table 5.22). Cisplatin plus

gemcitabine dominated nivolumab. The individual effects of each change on costs, QALYs and ICERs are presented in Section 6, Table 6.1. For comparison, the deterministic ERG base-case ICERs were £82,028, £64,298 and £66,161 per QALY gained, with cisplatin plus gemcitabine dominating nivolumab.

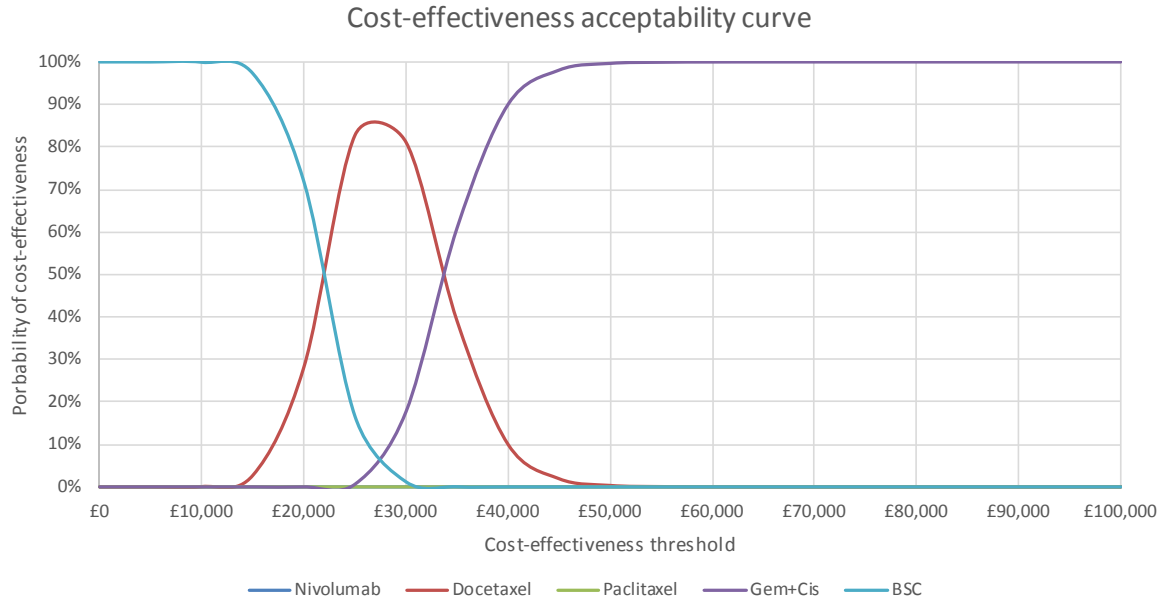
Table 5.6: ERG base-case (probabilistic)

	Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Nivolumab ICER (£/QALY)
ERG base-case	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£12,540	0.74	[redacted]	[redacted]	£86,030
	Paclitaxel	£13,905	0.63	[redacted]	[redacted]	£67,205
	Cis + gem	£29,284	1.24	[redacted]	[redacted]	Dominated
	BSC	£8,741	0.56	[redacted]	[redacted]	£68,348

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

The CEACs based on the ERG base-case (Figure 5.13) show that nivolumab has a probability of being cost effective of 0% and 0% at thresholds of £30,000 and £50,000 per QALY gained, respectively.

Figure 5.1: Cost effectiveness acceptability curve for ERG base-case



The ERG wishes to reiterate that the probabilistic model results are different from the deterministic results. This difference was more pronounced using the company's base-case (with fixed errors) than when using the ERG base-case. The difference is explained by using the response-based approach. However, it is not clear what in the response-based approach causes the probabilistic results to deviate as much from the deterministic results. The ERG considers it to be related to a) the increased uncertainty introduced by the response-based approach, b) the skew of the parametric models used and c) potentially the significant quantitative difference in OS and PFS caused by the response-based compared to the conventional approach.

Reviewing the overall evidence, the ERG confirmed that there was no existing cost effectiveness model for nivolumab for the current indication, and thus that development of a de novo model was necessary. The economic model described in the CS is considered by the ERG to meet the NICE reference case, with the notable exceptions of a) the exclusion of a comparator that was identified in the scope, and b) a PSA that excludes crucial parameters, includes parameters usually not included in the PSA (such as patient characteristics), and yields results significantly different from the deterministic results. The company model follows a logical structure with respect to the nature of the disease. The economic model was primarily informed by the CheckMate 275 and CheckMate 032 studies, both single-arm studies. Relative treatment effectiveness were informed based on a simulated treatment comparison using studies that were identified through the systematic literature review on the comparators docetaxel, paclitaxel, cisplatin plus gemcitabine and BSC.

The company base-case ICERs (probabilistic) of nivolumab (with PAS) compared with docetaxel, paclitaxel, cisplatin plus gemcitabine and BSC were £54,220, £46,209, £103,568 and £44,698 per QALY gained respectively. The cost effectiveness results were not robust to scenario and one-way sensitivity analyses conducted by the company. Scenario analyses indicated that the choice of nivolumab parametric OS, PFS and TTD curves, the position of the landmark, as well as the choice of the fractional polynomial model used for the NMA were major drivers of model results, mostly resulting in less favourable cost effectiveness estimates for nivolumab versus its comparators.
The ERG incorporated various adjustments to the company's base-case. The ERG base-case resulted in ICERs (probabilistic) of £86,030, £67,205 and £68,348 per QALY gained for nivolumab (with PAS) versus docetaxel, paclitaxel and BSC respectively. In the ERG base-case, cisplatin plus gemcitabine dominated nivolumab, with a larger QALY gain and lower costs. For comparison, the deterministic ERG base-case ICERs were £82,028, £64,298 and £66,161 per QALY gained, with cisplatin plus gemcitabine dominating nivolumab. The single most influential adjustment made by the ERG in its base-case was the use of conventional survival analysis instead of adopting the company's preferred response-based approach.

The ERG identified substantial issues and uncertainties that affected the cost effectiveness analysis. The main issues with the analysis include the use of a response-based survival analysis approach, which was not appropriately and sufficiently justified, necessitated a number of additional assumptions and therefore caused additional uncertainty. These additional assumptions included the choice of a landmark; the use of KM estimates up to the chosen landmark; assumptions surrounding the proportionality of hazards between responders and non-responders; increased uncertainty surrounding the choice of parametric time-to-event models for OS, PFS and TTD; the a posteriori combination of responder and non-responder groups; and the application of HRs in this artificial a posteriori population, which is not the same as the one that HRs were derived from. The ERG deemed the introduction of these additional uncertainties, some of which were shown to have a substantial effect on the ICERs in the ERG's exploratory analysis, as unjustified, given that the need for response-based analysis and its improvement over conventional analysis was not demonstrated. Further issues related to the exclusion of cisplatin plus gemcitabine as a comparator, inconsistencies in the source for nivolumab-related effectiveness, resource use, utilities and adverse event data (use of CheckMate 275 and CheckMate 032 for effectiveness, use of CheckMate 275 only for the others), the inclusion of adverse events with incidence smaller than 5%, the calculation of dose intensity, and the exclusion of important parameters from, and inclusion of inappropriate parameters in, the PSA.

There is substantial uncertainty about the relative treatment effectiveness estimates, which were entirely derived from single-arm studies, using a simulated treatment comparison that aimed at correcting for differences in the study populations. The residual bias could not be quantified in the

company's analysis, and cost effectiveness results should therefore be interpreted with extreme caution. Model estimates for nivolumab were not externally validated, apart from the comparison with NSCLC data, which may not be appropriate. The uncertainty introduced by the derived time-varying HRs was unfortunately not assessed within the PSA. In exploratory analysis, the ERG attempted to give a measure of parts of this uncertainty by using a naive comparison as opposed to the STC, and time-fixed HRs as opposed to time-varying HRs.

In exploratory analysis, the ERG found that using the naïve comparison resulted in pronounced increases in the ICERs (£90,465, £63,548, dominated, £64,429 per QALY gained when comparing nivolumab against docetaxel, paclitaxel, cisplatin plus gemcitabine and BSC respectively). These further increased in an extreme scenario where no relative treatment effect was assumed for nivolumab. The use of time-independent HRs also had a significant effect on ICERs, with some ICERs increasing and others decreasing compared to the ERG base-case ICERs (£70,452, £94,067, £74,858, £54,707 per QALY gained when comparing nivolumab against docetaxel, paclitaxel, cisplatin plus gemcitabine and BSC respectively). The use of alternative parametric time-to-event models for OS (lognormal) and PFS (log-logistic) in the conventional approach produced further increases in ICERs (£95,759, £78,505, dominated, £77,739 per QALY gained when comparing nivolumab against docetaxel, paclitaxel, cisplatin plus gemcitabine and BSC respectively). Using the response-based analysis with alternative time-to-event models for OS and PFS, but not for TTD, also resulted in a marked increase in ICERs compared with the response-based company's base-case (£122,716, £96,836, dominated, £94,964 per QALY gained when comparing nivolumab against docetaxel, paclitaxel, cisplatin plus gemcitabine and BSC respectively). Lastly, the alternative landmark drove the company's base-case ICERs up (£77,167, £73,309, £93,439, £62,903 per QALY gained when comparing nivolumab against docetaxel, paclitaxel, cisplatin plus gemcitabine and BSC respectively). The ERG also found that the use of different parameter values for the fractional polynomial model alone resulted in large variation in absolute costs and QALYs (Table 6.3). These findings illustrate how uncertain the presented cost effectiveness results are.

In conclusion, given the ERG base-case ICERs are estimated to be substantially above £60,000 per QALY gained, and the large uncertainty regarding comparative treatment effectiveness in combination with the lack of appropriate validation, uncertainty around the cost effectiveness of nivolumab remains substantial.

6. IMPACT ON THE ICER OF ADDITIONAL CLINICAL AND ECONOMIC ANALYSES UNDERTAKEN BY THE ERG

In Section 5.3 the ERG's base-case was presented, which was based on various changes compared to the company's base-case. Table 6.1 shows how each individual change impacts the ICER plus the combined effect of all changes simultaneously. The analyses numbers in Table 6.1 correspond to the analyses numbers reported in Section 5.3. Also, the exploratory analysis is presented in Table 6.2 (conditional on the ERG base-case). Finally, the threshold analyses are discussed in Section 5.3.2. Appendix 1 contains technical details on the analyses performed by the ERG.

Table 6.7: ERG base-case (probabilistic), nivolumab with PAS

	Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Nivolumab ICER (£/QALY)
Probabilistic Company base-case[a]	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£12,748	0.82	[redacted]	[redacted]	£54,131
	Paclitaxel	£14,186	0.71	[redacted]	[redacted]	£45,482
	Cis+gem	£30,443	1.34	[redacted]	[redacted]	£100,417
	BSC	£8,811	0.57	[redacted]	[redacted]	£44,873
Fixing errors (1) and (2)	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£12,744	0.82	[redacted]	[redacted]	£50,974
	Paclitaxel	£14,155	0.71	[redacted]	[redacted]	£42,715
	Cis+gem	£29,969	1.34	[redacted]	[redacted]	£91,773
	BSC	£8,813	0.58	[redacted]	[redacted]	£42,532
Proportions of responders based on OS for TTD (4)[b]	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£12,779	0.82	[redacted]	[redacted]	£50,889
	Paclitaxel	£14,162	0.71	[redacted]	[redacted]	£42,644
	Cis+gem	£29,960	1.35	[redacted]	[redacted]	£92,606
	BSC	£8,819	0.58	[redacted]	[redacted]	£42,435
Removing AEs with incidence < 5% (5)[b]	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£12,810	0.82	[redacted]	[redacted]	£51,023
	Paclitaxel	£14,205	0.71	[redacted]	[redacted]	£42,870
	Cis+gem	£29,982	1.34	[redacted]	[redacted]	£92,433
	BSC	£8,858	0.58	[redacted]	[redacted]	£42,566
Utilities from pooled	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£12,803	0.84	[redacted]	[redacted]	£49,613

	Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Nivolumab ICER (£/QALY)
CheckMate studies (6)[b]	Paclitaxel	£14,204	0.73	[redacted]	[redacted]	£41,605
	Cis+gem	£29,994	1.39	[redacted]	[redacted]	£91,388
	BSC	£8,849	0.59	[redacted]	[redacted]	£41,406
Weight from pooled CheckMate studies (7)[b]	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£12,763	0.82	[redacted]	[redacted]	£52,682
	Paclitaxel	£14,165	0.71	[redacted]	[redacted]	£44,199
	Cis+gem	£29,975	1.34	[redacted]	[redacted]	£98,529
	BSC	£8,819	0.58	[redacted]	[redacted]	£43,780
Excluding parameters from PSA (8)[b]	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£12,763	0.82	[redacted]	[redacted]	£51,149
	Paclitaxel	£14,178	0.71	[redacted]	[redacted]	£42,868
	Cis+gem	£29,960	1.34	[redacted]	[redacted]	£92,876
	BSC	£8,829	0.57	[redacted]	[redacted]	£42,632
Conventional instead of response-based analysis (9)[b]	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£12,507	0.72	[redacted]	[redacted]	£84,193
	Paclitaxel	£13,894	0.61	[redacted]	[redacted]	£65,302
	Cis+gem	£29,082	1.20	[redacted]	[redacted]	Dominated
	BSC	£8,736	0.55	[redacted]	[redacted]	£66,951
Missed doses when delayed > 7days (10)[b]	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£12,894	0.82	[redacted]	[redacted]	£54,053
	Paclitaxel	£14,197	0.71	[redacted]	[redacted]	£45,372
	Cis+gem	£31,620	1.35	[redacted]	[redacted]	£105,278
	BSC	£8,844	0.58	[redacted]	[redacted]	£44,704
ERG base-case (combining adjustments 1-10)	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£12,540	0.74	[redacted]	[redacted]	£86,030
	Paclitaxel	£13,905	0.63	[redacted]	[redacted]	£67,205
	Cis+gem	£29,284	1.24	[redacted]	[redacted]	Dominated
	BSC	£8,741	0.56	[redacted]	[redacted]	£68,348

Table 6.8: Exploratory analyses; nivolumab with PAS

	Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Nivolumab ICER (£/QALY)
Probabilistic Company base-case[a]	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£12,748	0.82	[redacted]	[redacted]	£54,131
	Paclitaxel	£14,186	0.71	[redacted]	[redacted]	£45,482
	Cis+gem	£30,443	1.34	[redacted]	[redacted]	£100,417
	BSC	£8,811	0.57	[redacted]	[redacted]	£44,873
ERG base-case	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£12,540	0.74	[redacted]	[redacted]	£86,030
	Paclitaxel	£13,905	0.63	[redacted]	[redacted]	£67,205
	Cis+gem	£29,284	1.24	[redacted]	[redacted]	Dominated
	BSC	£8,741	0.56	[redacted]	[redacted]	£68,348
Alternative parametric TTE models (lognormal for OS, log-logistic for PFS) (A.1)	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£11,696	0.66	[redacted]	[redacted]	£95,759
	Paclitaxel	£13,688	0.59	[redacted]	[redacted]	£78,505
	Cis+gem	£28,094	1.10	[redacted]	[redacted]	Dominated
	BSC	£8,611	0.52	[redacted]	[redacted]	£77,739
Naïve comparison data instead of STC results (A.3)	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£12,959	0.77	[redacted]	[redacted]	£90,465
	Paclitaxel	£13,850	0.60	[redacted]	[redacted]	£63,548
	Cis+gem	£30,716	1.56	[redacted]	[redacted]	Dominated
	BSC	£8,588	0.52	[redacted]	[redacted]	£64,429
Time-independent HRs (A.4)	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£10,172	0.60	[redacted]	[redacted]	£70,452
	Paclitaxel	£13,035	0.78	[redacted]	[redacted]	£94,067
	Cis+gem	£26,435	0.86	[redacted]	[redacted]	£74,858
	BSC	£8,135	0.39	[redacted]	[redacted]	£54,707

	Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Nivolumab ICER (£/QALY)
Alternative assumptions for PFS HRs for BSC and cis+gem (A.5)	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£12,500	0.74	[redacted]	[redacted]	£86,455
	Paclitaxel	£13,882	0.63	[redacted]	[redacted]	£67,486
	Cis+gem	£34,843	1.26	[redacted]	[redacted]	Dominated
	BSC	£8,710	0.55	[redacted]	[redacted]	£67,346
AE disutilities and resource use from TA ID971 (A.6)	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£12,083	0.74	[redacted]	[redacted]	£87,485
	Paclitaxel	£13,680	0.63	[redacted]	[redacted]	£67,677
	Cis+gem	£26,381	1.27	[redacted]	[redacted]	Dominated
	BSC	£8,753	0.57	[redacted]	[redacted]	£68,428
UK dosage schedule for cis+gem (A.7)	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£12,539	0.74	[redacted]	[redacted]	£85,743
	Paclitaxel	£13,900	0.63	[redacted]	[redacted]	£66,966
	Cis+gem	£31,088	1.24	[redacted]	[redacted]	Dominated
	BSC	£8,738	0.56	[redacted]	[redacted]	£68,131
No treatment effect of nivolumab vs comparators (A.8)	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£13,753	1.19	[redacted]	[redacted]	£5,634,843
	Paclitaxel	£14,298	1.20	[redacted]	[redacted]	£11,163,091
	Cis+gem	£31,907	1.15	[redacted]	[redacted]	£404,845
	BSC	£10,670	1.16	[redacted]	[redacted]	£1,153,670
Response-based analysis using ERG base-case (B.1)	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£12,919	0.85	[redacted]	[redacted]	£53,937
	Paclitaxel	£14,198	0.73	[redacted]	[redacted]	£45,466
	Cis+gem	£31,662	1.40	[redacted]	[redacted]	£108,156
	BSC	£8,838	0.60	[redacted]	[redacted]	£44,600
Response-based	Nivolumab	[redacted]	[redacted]			

	Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Nivolumab ICER (£/QALY)
analysis using alternative TTE models for OS, PFS, but not TTD (B.2)	Docetaxel	£12,516	0.74	[redacted]	[redacted]	£122,716
	Paclitaxel	£13,891	0.63	[redacted]	[redacted]	£96,836
	Cis+gem	£29,271	1.24	[redacted]	[redacted]	Dominated
	BSC	£8,718	0.56	[redacted]	[redacted]	£94,964
Response-based analysis using alternative TTE models for OS, PFS and TTD (B.3)	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£12,507	0.77	[redacted]	[redacted]	£75,916
	Paclitaxel	£13,978	0.68	[redacted]	[redacted]	£66,008
	Cis+gem	£29,779	1.25	[redacted]	[redacted]	£140,296
	BSC	£8,699	0.55	[redacted]	[redacted]	£62,998
Response-based analysis using 26-week landmark (B.4)	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£10,711	0.50	[redacted]	[redacted]	£77,167
	Paclitaxel	£13,681	0.52	[redacted]	[redacted]	£73,309
	Cis+gem	£28,436	0.78	[redacted]	[redacted]	£93,439
	BSC	£8,043	0.35	[redacted]	[redacted]	£62,903

that systematic error has been eliminated. Hoaglin,[72, 73] in a series of letters critiquing an unanchored comparison by Di Lorenzo et al.[78] based upon a matching approach similar to MAIC, remarked that, without providing evidence that the adjustment compensates for the missing common comparator arms and the resulting systematic error, the ensuing results "are not worthy of consideration".'[1]

No formal comparison of AEs including no evidence synthesis was performed, although it might be reasonable to conclude, based on few data from the comparators, that the rate of key AEs was generally similar to or lower than the comparators.

In conclusion, it is difficult to be sure what the effectiveness of nivolumab is in comparison to the comparators in the scope. Evidence from directly examining the single arms of the trial data indicates little difference between the outcomes measured from the nivolumab and comparator studies. Such a naive comparison carries a high risk of bias. STC analysis was used to try and reduce this bias, but there is also no clear evidence that risk of bias was reduced by the STC analysis. Multiple limitations in the STC were identified and the test of validity recommended by TSD 18, the 'out-of-sample' method either lack of success in reducing the bias if it is applicable at all given the lack of data and PF model. The ERG was able to estimate the unadjusted hazards for nivolumab, but not with estimates of uncertainty. The effect of an analysis based on different combinations of covariates in the prediction model used to make the adjustment remains unknown.

With regards to the health economic model submitted by the company, the ERG demonstrated that there was large uncertainty surrounding the ICERs and that a number of alternative assumptions could change the ICERs significantly. Most crucially, the ERG questioned the need for the company's response-based approach to survival analysis, which was deemed insufficiently justified. If a response-based approach was indeed deemed necessary, then other, more established methods, should be explored (spline-based or mixture cure models, as recommended in TSD 14).[38] However, it should also be noted, that the company's approach to implementing the response-based approach necessitated additional model assumptions and increased uncertainty. The resulting model predictions were different from those obtained using a conventional approach to an extent that might be implausible; the lack of validation by experts further made the ERG question the plausibility of the company's base-case. Furthermore, the exclusion of cisplatin plus gemcitabine from the base-case stood in contrast to the scope and was inappropriately justified.

Apart from this, numerous issues were identified by the ERG. The ERG was able to adjust/correct some of these in its base-case. This resulted in ICERs (probabilistic) of £86,030, £67,205 and £68,348 for nivolumab (with PAS) versus docetaxel, paclitaxel and BSC, respectively. Cisplatin plus gemcitabine dominated nivolumab.

Additional sensitivity analyses were performed to examine the potential impact of alternative assumptions on the cost effectiveness estimates. These included two scenarios in which changes were implemented: a) exploratory analyses performed using the ERG base-case, and b) exploratory analyses performed using the ERG base-case, except that a response-based approach was used. Scenarios exploring the uncertainty about the treatment and relative effectiveness evidence significantly increased the ICERs. Using one example set of alternative parametric time-to-event models within the ERG base-case decreased the ICERs significantly. Finally, using the response-based approach significantly decreased the ICER, but these ICERs were shown to increase significantly with the use of best-fitting parametric time-to-event models. In addition, alternative parameter values informing the fractional polynomial model for the NMA showed that this model feature alone could have a vast impact on the ICERs.