NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Nivolumab for treating metastatic or unresectable urothelial cancer after platinum-based chemotherapy [ID995]

Committee Papers

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Nivolumab for treating metastatic or unresectable urothelial cancer after platinum-based chemotherapy [ID995]

Contents:

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

Contains AIC ,CIC

Pre-meeting briefing

Nivolumab for treating metastatic or unresectable urothelial cancer after platinum-based chemotherapy [ID995]

This slide set is the pre-meeting briefing for this appraisal. It has been prepared by the technical team with input from the committee lead team and the committee chair. It is sent to the appraisal committee before the committee meeting as part of the committee papers. It summarises:

- the key evidence and views submitted by the company, the consultees and their nominated clinical experts and patient experts and
- the Evidence Review Group (ERG) report

It highlights key issues for discussion at the first appraisal committee meeting and should be read with the full supporting documents for this appraisal

Please note that this document includes information from the ERG before the company has checked the ERG report for factual inaccuracies

The lead team may use, or amend, some of these slides for their presentation at the Committee meeting

1

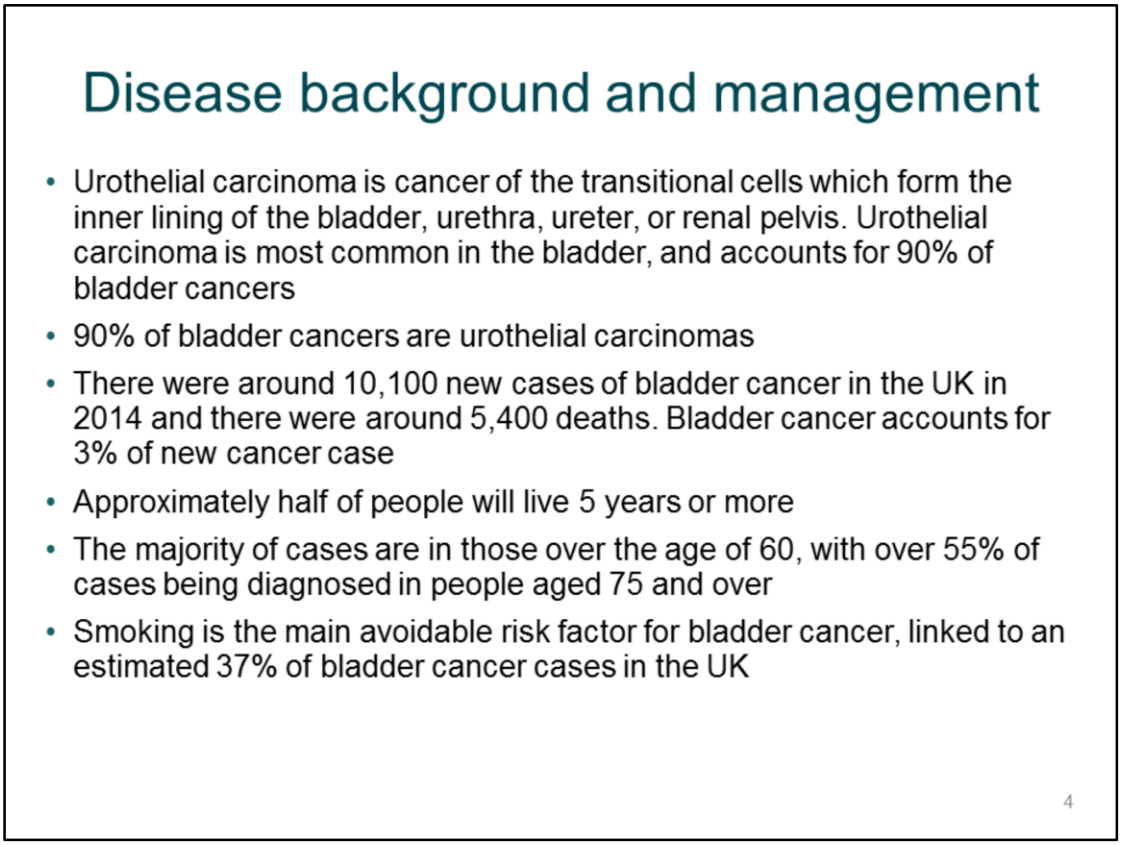
Key points for consideration

- Quality of evidence
  - no comparative nivolumab trial data
  - are the nivoumab studies generalisable to UK practice?
  - how reliable is the simulated treatment comparison? Does the company account for all of the important prognostic factors?
  - how reliable is the network meta-analysis? Are the included studies sufficiently homogeneous?
- How effective is nivolumab?
- Is there enough evidence to make recommendations for PD-L1 subgroups?
- The company excluded gemcitabine and cisplatin from its base case. Is this appropriate?
- Company used a response-based analyses to model survival. ERG preferred conventional approach. Which approach is most appropriate?
- What is the most plausible ICER?
- Can nivolumab be considered innovative? Does end of life apply?

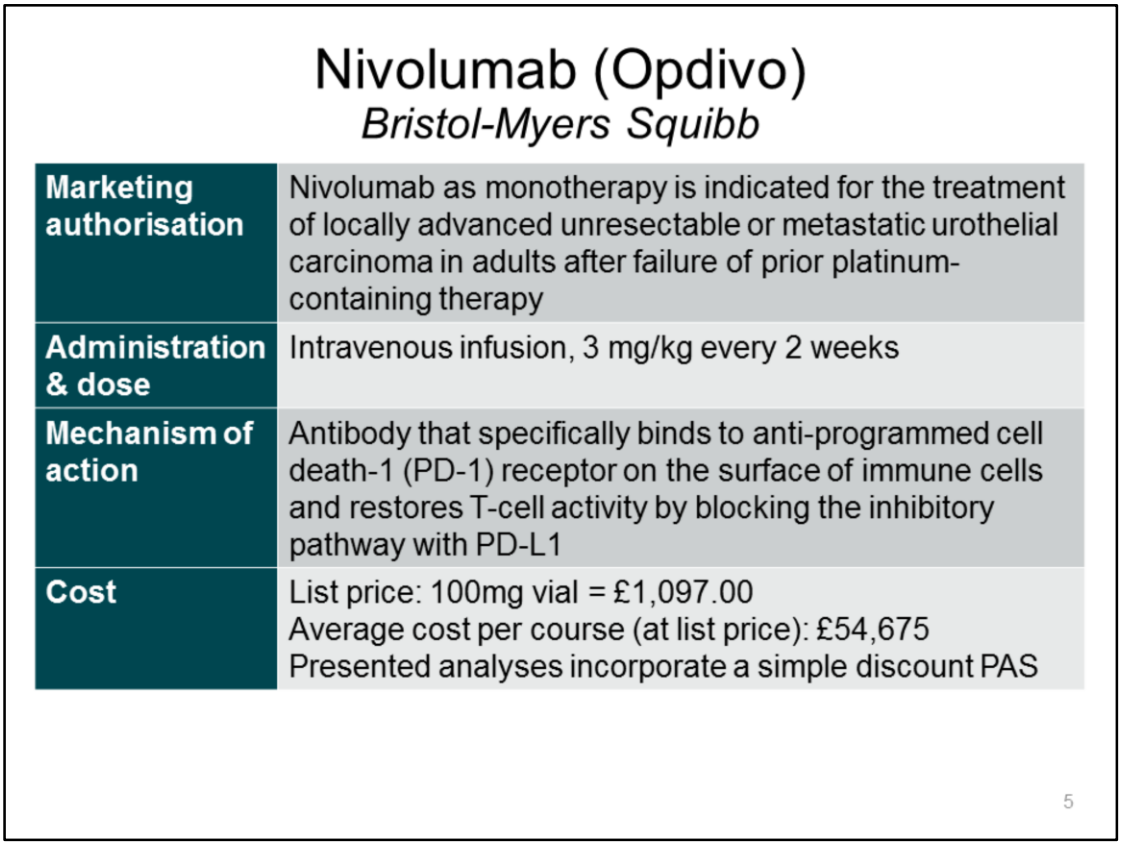
2

Common abbreviations			
AE	Adverse event		
AIC	Akaike information criterion	mRECIST	modified RECIST
ASaT	All subjects as treated	NMA	Network meta-analysis
BIC	Bayesian information criterion	NR	Not reported
BICR	Blinded independent central review	ORR	Objective response rate
CDF	Cancer Drugs Fund	OS	Overall survival
CHMP	Committee for Medicinal Products for Human Use	PAS	Patient access agreement
CI	Confidence Interval	PD	Progressed disease
CPS	Combined proportion score	PD-L1	Programmed death-ligand 1
CR	Complete response	PFS	Progression-free survival
CS	Company submission	PH	Proportional hazards
CSR	Clinical study report	PR	Partial response
DCR	Disease control rate	PSA	Probabilistic sensitivity analysis
EAMS	Early Access to Medicines Scheme	PSS	Personal and Social Services
ECOG	Eastern Cooperative Oncology Group	Q3W	Every 3 weeks
EMA	European Medicines Agency	QALY	Quality adjusted life year
EORTC	European Organisation for the Treatment of Cancer	QLQ	Quality of life questionnaire
EQ-5D	European Quality of Life - 5 Dimensions Questionnaire	RCT	Randomised controlled trial
ERG	Evidence Review Group	RECIST	Response Evaluation Criteria In Solid Tumors
HR	Hazard ratio	RPSFT	Rank preserving structural failure time
HRQoL	Health-related quality of life	RR	Response rate
IA1	First interim analysis	SAE	Serious adverse event
IA2	Second interim analysis	sd	Standard deviation
ICER	Incremental cost effectiveness ratio	SD	Stable disease
Incr.	Incremental	SmPC	Summary of product characteristics
IPCW	Inverse Probability of Censoring Weighting	SOC	Standard of care
ITT	Intention-to-treat	TCC	transitional cell carcinoma
K-M	Kaplan-Meier	TPS	Tumour proportion score
LS	Least squares	TTR	Time to response
LY	Life Years	UK SOC	UK standard of care (i.e. paclitaxel and docetaxel)

3



Updated figures compared to those reported in the scope



Nivolumab (Opdivo)

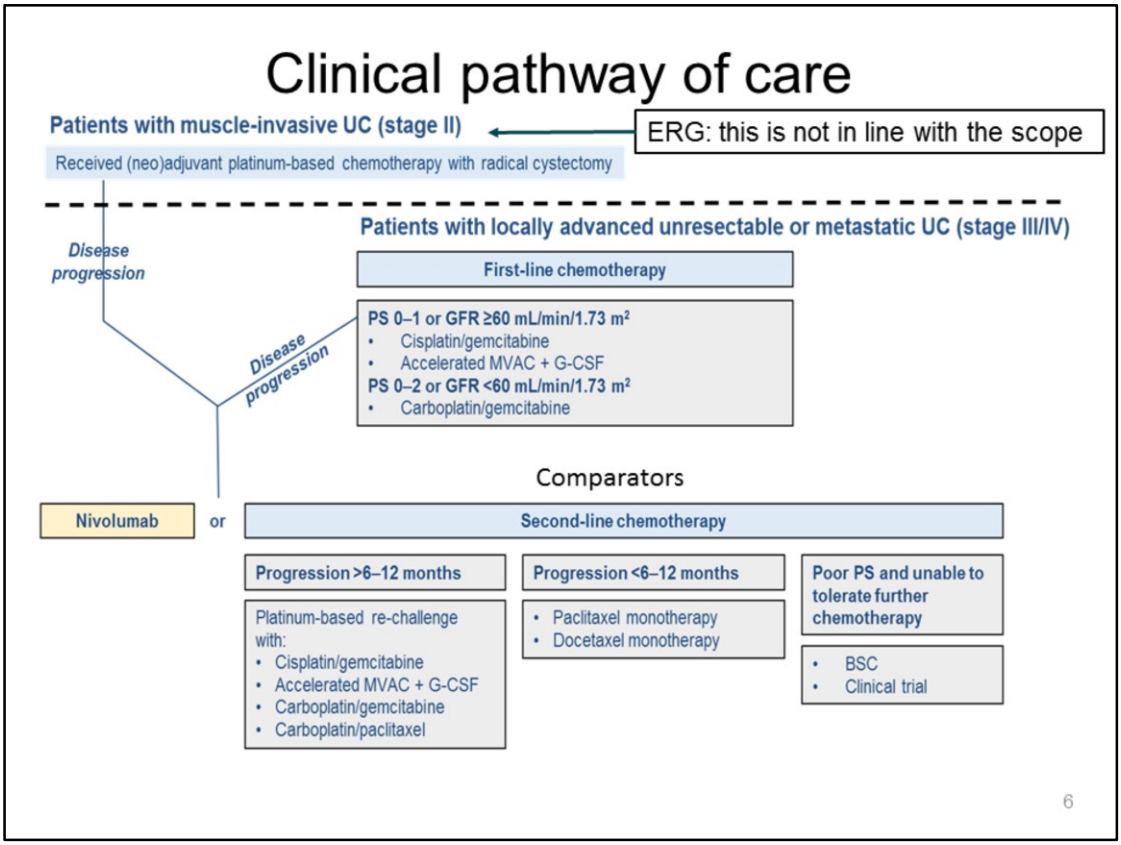
Bristol-Myers Squibb

Marketing authorisation	Nivolumab as monotherapy is indicated for the treatment of locally advanced unresectable or metastatic urothelial carcinoma in adults after failure of prior platinum-containing therapy
Administration & dose	Intravenous infusion, 3 mg/kg every 2 weeks
Mechanism of action	Antibody that specifically binds to anti-programmed cell death-1 (PD-1) receptor on the surface of immune cells and restores T-cell activity by blocking the inhibitory pathway with PD-L1
Cost	List price: 100mg vial = £1,097.00 Average cost per course (at list price): £54,675 Presented analyses incorporate a simple discount PAS

5

- Nivolumab has a marketing authorisation for 5 other indications
  - Melanoma [TA384 – recommended in adults]
  - Non-Small Cell Lung Cancer (NSCLC) [NICE TA ID811 (after chemo) – not yet published]
  - Renal Cell Carcinoma (RCC) [NICE TA417 – recommended in previously treated adults]
  - Classical Hodgkin lymphoma (cHL) [TA426 – recommended in adults with relapsed or refractory cHL]
  - Squamous Cell Cancer of the Head and Neck (SCCHN) [NICE TA ID971 – in appraisal, publication expected November 2017]

- Other PD-L1/PD-1 inhibitors are being appraised for this indication:
  - 'Atezolizumab for treating metastatic urothelial bladder cancer after platinum-based chemotherapy' (ID939) – Committee D
  - 'Pembrolizumab for previously treated advanced or metastatic urothelial cancer'



Source: Figure 7, page 23, company submission

Based on the above treatment pathway, the treatment options representing potentially relevant comparators to nivolumab in the context of this submission are as follows:

- Paclitaxel monotherapy (*standard of care)*
- Docetaxel monotherapy
- BSC
- Retreatment with platinum-based chemotherapy *(<10% of patients)* – company estimate
    - Cisplatin plus gemcitabine
    - Accelerated MVAC plus G-CSF
    - Carboplatin plus gemcitabine
    - Carboplatin plus paclitaxel.

Company estimates the eligible population to be **894 patients**. Full details are in section B.4.1 of the company submission.

Patient perspectives

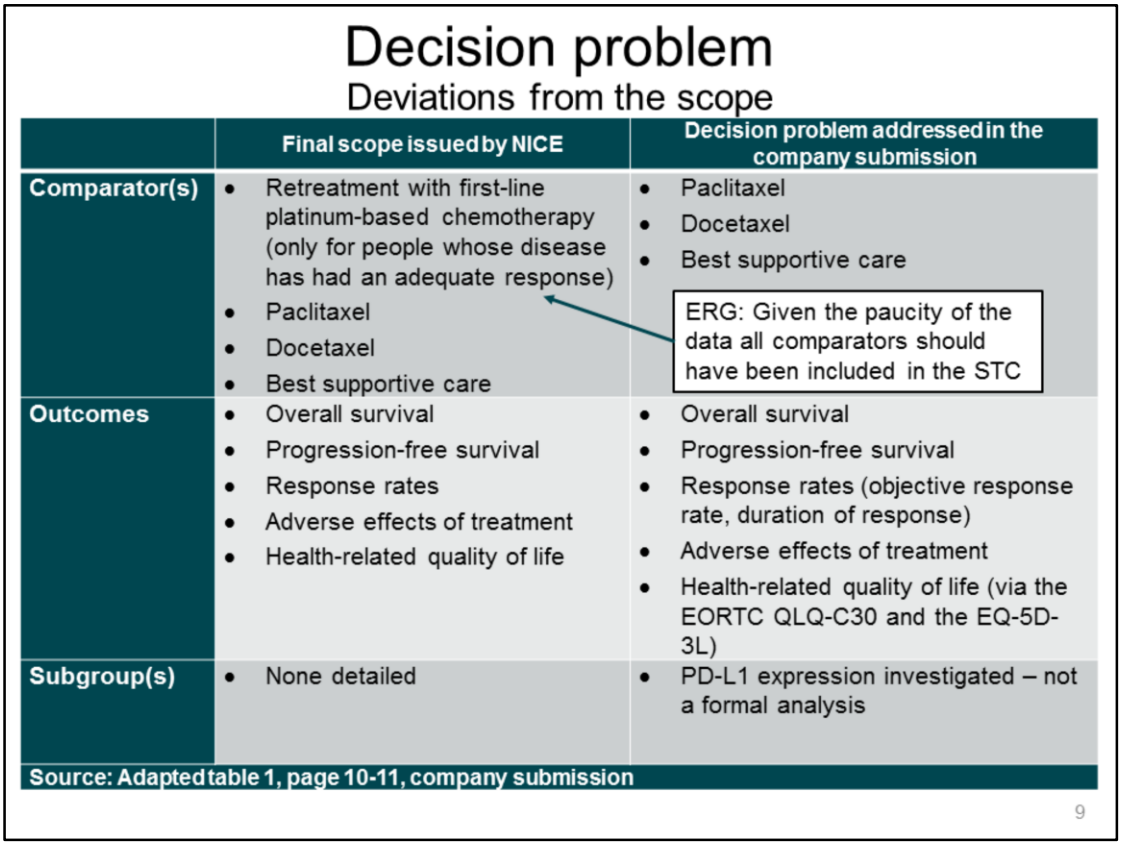
- "Bladder cancer has a very poor prognosis"
- After platinum chemo, few options, "survival rates…exceptionally poor"
- "Many are unable to tolerate the preferred cisplatin chemo… huge unmet need…patients generally overlooked"
- Nivolumab:
  - "Trials show treatment prolongs life, and for 20% of patients the effects are enduring"
  - "Side effects for the majority are minor and tolerable"
  - "Innovative breakthrough treatment"
  - "For a cancer with so few advance in decades, this gives hope to many"

7

Clinician perspectives

- Main treatment aims: "to palliate symptoms, improve quality of life and delay time to further progression of disease and improve survival"
- NHS second-line treatment:
  - Paclitaxel commonest regimen
  - with around 10% response rate, many patients decline further chemo
  - Many centres used PD1/PD-L1 inhibitors last year instead of chemo
- Nivolumab:
  - "For use in good performance status patients"
  - Should increase overall survival and health-related quality of life
  - Acceptable side-effect profile
  - Would be administered in specialist clinics in secondary care
  - Facilities and equipment already in place, some training required (e.g. on side-effects)

8



Decision problem

Deviations from the scope

	Final scope issued by NICE	Decision problem addressed in the company submission
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 ERG: Given the paucity of the data all comparators should have been included in the STC
Outcomes	• Overall survival • Progression-free survival • Response rates • Adverse effects of treatment • Health-related quality of life	• Overall survival • Progression-free survival • Response rates (objective response rate, duration of response) • Adverse effects of treatment • Health-related quality of life (via the EORTC QLQ-C30 and the EQ-5D-3L)
Subgroup(s)	• None detailed	• PD-L1 expression investigated – not a formal analysis

Source: Adapted table 1, page 10-11, company submission

9

Population: 0 UK patients in CheckMate275, 6 (7.7%) UK patients in CheckMate032

Intervention: in Checkmate032, 23% switched to nivolumab in combination with ipilimumab.

Comparator(s): According to the company the evidence for cisplatin+gemcitabine was not relevant to this population as they had M-VAC in the first-line. Also, as platinum chemo re-challenge is limited to <10% of cases, the company excluded this from their base-case. The ERG noted that, platinum re-challenge should be included regardless of the quality of the information as there is lots of uncertainty in the STC anyway.

Outcomes: CM275: ORR was based on BIRC assessment using recist v1.1 in all patients regardless of PD-L1 expression. CM032: ORR was based on investigator assessed, defined as the number of patients with best overall response of a CR or PR using recist v1.1, divided by the total number of patients. Disease progression is normally measured by CT, however RECIST is an accepted tool.

Subgroup(s): The MA is not restricted based on PD-L1 expression status, however the EPAR mentions different outcomes between these groups. See EPAR notes on the clinical effectiveness conclusions slide.

Clinical effectiveness

Evidence overview

- No RCTs directly comparing the efficacy and safety of nivolumab in the patient population versus any relevant comparators or placebo were identified
- Two trials provide evidence for the efficacy and safety of patients after at least one line of platinum-based chemotherapy
  - CheckMate 275
  - CheckMate 032
- CheckMate 275 is a study of nivolumab in patients from the target population
- CheckMate 032 is a study of *nivolumab monotherapy* or *nivolumab in combination with ipilimumab* in patients with one of the following tumour types: UC, triple-negative breast cancer, gastric cancer, pancreatic adenocarcinoma, small cell lung cancer, and ovarian cancer
  - Only a subgroup from this study is relevant to this submission

10

CheckMate 275

Study design	Multicentre, open-label, single-arm phase II study
Population (N=270)	Patients with locally advanced unresectable or metastatic UC who had progressed or recurred after at least one previous line of platinum-containing chemotherapy
Location	63 sites across 11 countries in North America (USA), Europe, Australia and Asia
Intervention(s)	Nivolumab (IV 3 mg/kg Q2W)
Comparator(s)	N/A (single-arm)
Reported outcomes specified in the decision problem	• ORR (primary outcome – BIRC assessed) • OS • PFS • HRQoL via the European Organisation for Research and Treatment of Cancer (EORTC) general cancer module (QLQ-C30) and the EuroQoL-5 dimensions-3 levels (EQ-5D-3L) questionnaires • Adverse events (AEs)
All other reported outcomes	• Duration of response and additional safety outcomes

Source: Table 4, page 26, company submission

BIRC – blinded independent review committee

11

- Patients with histologically confirmed metastatic or surgically unresectable UC with disease progression or recurrence after at least one platinum-based chemotherapy were enrolled and assigned to a cohort according to tumor PD-L1 expression status (PD-L1 ≥5%, PD-L1 < 5%, or indeterminate). Enrollment in the trial continued until approximately 70 subjects with confirmed PD-L1 expression of ≥5% were treated.

- Treatment beyond initial investigator-assessed RECIST v1.1-defined progression was permitted if the subject had an investigator-assessed clinical benefit, did not have rapid disease progression, and was tolerating the study drug.

- The primary endpoint was ORR based on BIRC assessment using RECIST v1.1 in the all-treated population, in patients with PD-L1 expression ≥1%, and in patients with PD-L1 expression ≥5%
    - ORR was defined as the proportion of people with complete response (CR) or partial response (PR), as determined by a BIRC
- Time to response and duration of response were estimated in patients with a confirmed CR or PR
    - Responses were confirmed at the second scan at least 4 weeks after criteria for objective response were met

CheckMate 032

Study design	Multicentre, open-label, two-stage, multi-arm, phase I/II[a]
Population (N=78)	Patients with locally advanced unresectable or metastatic UC who had progressed or recurred after treatment with at least one platinum-containing chemotherapy regimen
Location	International: 16 sites in 5 countries: Finland, Germany, Spain, UK and USA
Intervention(s)	Nivolumab (IV 3 mg/kg Q2W)
Comparator(s)	N/A[a]
Reported outcomes specified in the decision problem	• ORR (investigator assessed) • OS • PFS • EQ-5D-3L • AEs
All other reported outcomes	• Duration of response and additional safety outcomes

[a]CheckMate 032 investigated nivolumab or nivolumab combined with ipilimumab in patients with UC, triple-negative breast cancer, gastric cancer, pancreatic adenocarcinoma, small cell lung cancer, and ovarian cancer. Here, presentation of CheckMate 032 refers only to the nivolumab monotherapy UC cohort (n=86) of relevance to this submission.
Source: Table 4, page 26, company submission

12

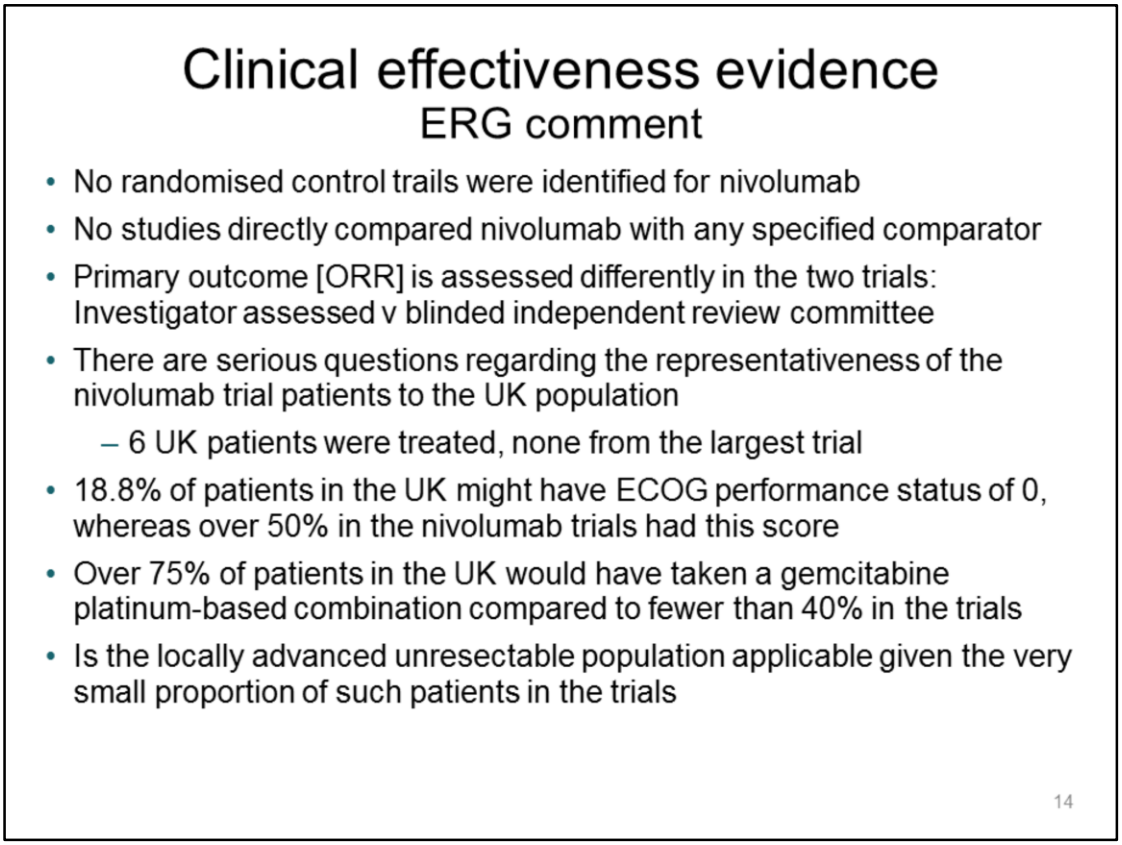
- A study investigating the efficacy and safety of nivolumab or nivolumab combined with ipilimumab in patients in a variety of tumour types as well as UC.
- Patients receiving nivolumab monotherapy 18 (23%) of the 78 patients switched to nivolumab plus ipilimumab . Switching was only allowed if they met pre-specified criteria:
    - Patient had confirmed radiologic disease progression (investigator-assessed RECIST 1.1–defined progression confirmed at least 4 weeks after the initial tumour assessment showing progression) in the absence of clinical deterioration. For patients with clear evidence of new or progressing brain metastases, a confirmation was not required. These patients may proceed with brain radiation therapy, and after having completed the radiation therapy, a switch to the nivolumab-ipilimumab regimen could be considered.
    - Patient had not experienced nivolumab-related adverse events leading to permanent discontinuation.
    - Patient was not continuing to derive any clinical benefit from nivolumab single agent therapy as assessed by the investigator which would allow continuation of nivolumab monotherapy.

Key baseline characteristics

Characteristic	CheckMate 275	CheckMate 032
Mean age years (range)	66 (38–90)	66 (31–85)
% ECOG PS: 0/1/3	53.7 / 45.9 / 0.3	53.8 / 46.2 / 0
Male, n (%)	211 (78.1)	54 (69.2)
% smoking: current+former / never / unknown	71.9 / 24.8 / 3.3	61.5 / 37.2 / 1.3
% PD-L1: <1% / ≥1%	54.1 / 45.9	53.8 / 31.8
% PD-L1: <5% / ≥5%	69.3 / 30.7	67.9 / 17.9
% metastases at baseline visceral / liver / lymph node	84.1 / 27.8 / 15.9	78.2 / 25.6 / 16.7
% disease setting: metastatic / locally-unresectable	96.7 / 3.3	91.0 / 9.0
% previous therapies: 0 / 1 / 2 / ≥3 & 0 / 1 / 2-3 / >3	28.5 / 42.2 / 21.2 / 8.1	n/a / 33.3 / 53.8 / 12.8
% subsequent therapy any / radiotherapy / surgery	19.6 / 9.4 / 3.0	29.5 /11.5 / 6.4
% UK / Non-UK	0 / 100	7.7 / 92.3

Source: Adapted from table 6 (page 35), company submission

13



During clarification the company stated
"… clinical expert attendees at the advisory board stated that the CheckMate 275 and CheckMate 032 trial populations could be considered generally representative of the UK patient population."

"… it should be noted that ECOG performance status was adjusted for as a prognostic factor in the prediction model for the simulated ITC. As such, any differences in ECOG performance status between the patient populations of the nivolumab and comparator trials, are accounted for in the relative effectiveness estimates"

"It is difficult to determine what proportion of the scope population in UK practice might have locally unresectable non-metastatic disease as opposed to metastatic disease. The two groups are classified together for the purposes of treatment decision-making"

CONFIDENTIAL

CheckMate 275: Latest efficacy results

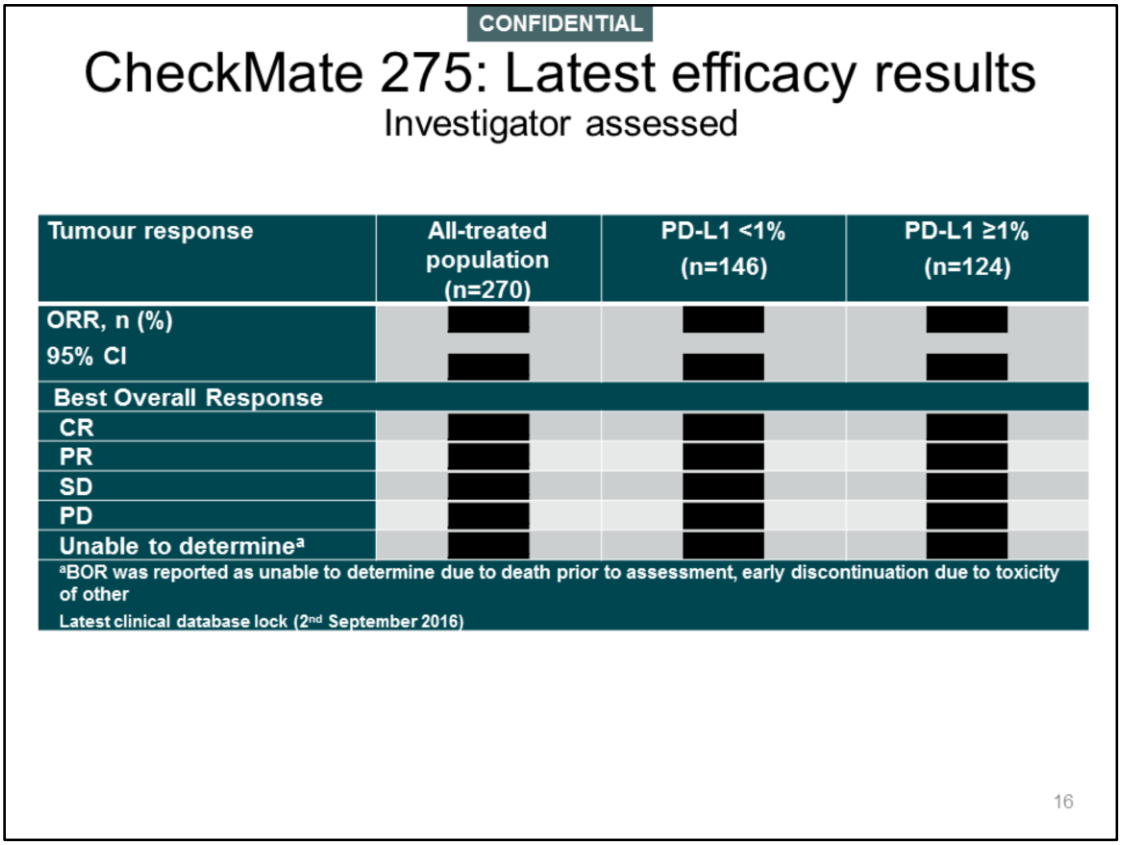
Blinded independent review committee

Tumour response	All-treated population (n=270)	PD-L1 <1% (n=146)	PD-L1 ≥1% (n=124)
ORR, n (%) 95% CI	54 (20.0) 95% CI: 15.4–25.3	23 (15.8) 95% CI: 10.3–22.7)	31 (25.0) 95% CI: 17.7–33.6
CR	8 (3.0)	[redacted]	[redacted]
PR	46 (17.0)	[redacted]	[redacted]
SD	60 (22.2)	[redacted]	[redacted]
PD	[redacted]	[redacted]	[redacted]
Unable to determine[a]	[redacted]	[redacted]	[redacted]
Median Time to response [TTR] (n=54), months IQR	1.94 IQR: 1.84–2.50	1.97 IQR: 1.87–3.48	[redacted] [redacted]
Median duration of response [DOR] (n=54), months 95% CI	10.35 95% CI: 7.52–NR	10.35 95% CI: 7.43–NR	NR 95% CI: 7.52–NR

[a]BOR was reported as unable to determine in 51 patients (18.5%); main reason was death prior to assessment.
Latest clinical database lock (2nd September 2016)

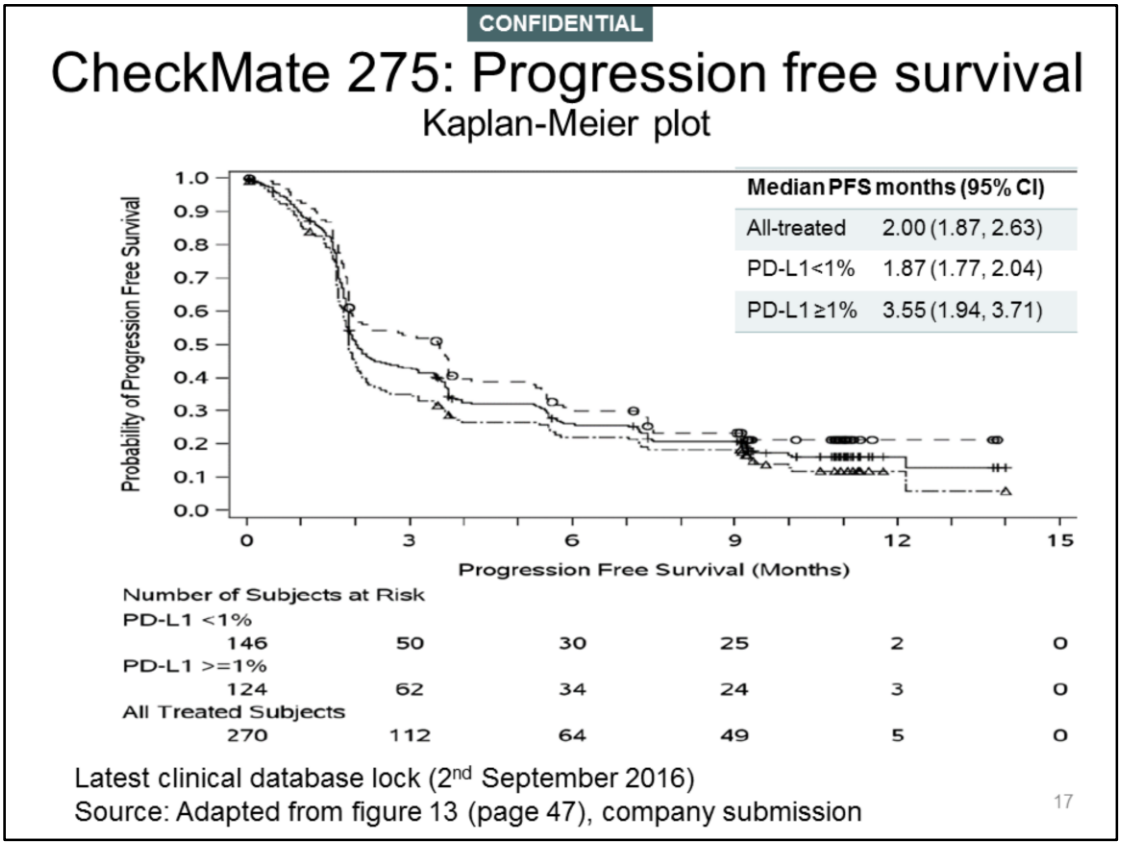
15

Source: Table 14 (page 47), company submission



Source: Table 5 (page 24), company response to clarification

Results for investigator-assessed ORR were investigated as a secondary outcome and the results were consistent with BIRC-assessed ORR



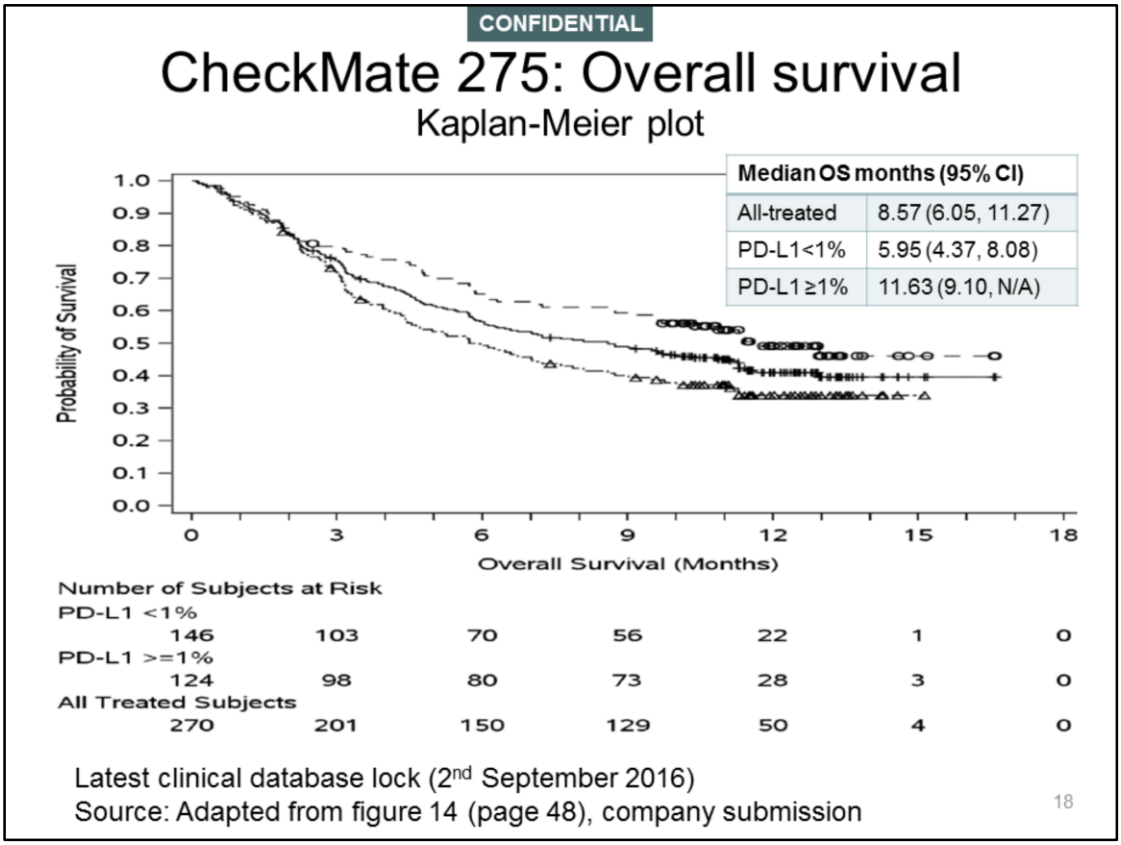
Source: Adapted from figure 13 (page 47), company submission

Graph shows BIRC assessed data

Investigator assessed median PFS from the latest database lock was provided by the company in their response to clarification:

- Median PFS (95% ci), Months:
  - All treated population = [redacted]
  - PD-L1 <1% = [redacted]
  - PD-L1 >1% = [redacted]

These results are mostly consistent with the figures in the graph above.



Source: Adapted from figure 14 (page 48), company submission

Results from the second database lock of CheckMate 275 (2 September 2016) were consistent with those from the primary analysis database lock in terms of ORR, PFS and OS.

There continues to be a statistically signification difference in median OS between PD-L1 <1% and PD-L1 >= 1% (5.95 months (95% CI: 4.37 to 8.08), and in the PD-L1 <1%, median PFS was 11.63 months (95% CI: 9.10 to NA).

CONFIDENTIAL

CheckMate 032 results

Investigator assessed

Tumour response	Nivolumab (n=78)	PD-L1<1% (n=42)	PD-L1 ≥1% (n=25)
ORR, n (%)	19 (24.4) [95% CI 15.3–35.4]	11 (26.2)	6 (24.0)
BOR, n (%)			
CR	5 (6.4)	1 (2.4)	4 (16.0)
PR	14 (17.9)	10 (23.8)	2 (8.0)
SD	22 (28.2)	11 (26.2)	8 (32.0)
PD	30 (38.5)	18 (42.9)	8 (32.0)
Unable to determine	7 (9.0)	2 (4.8)	3 (12.0)
Median TTR, months (IQR)	1.48 (1.25–4.14)	[redacted]	[redacted]
Median DOR, months (95% CI)	NR (9.92–NR)	[redacted]	[redacted]

Primary clinical database lock (24th March 2016)

19

Source: Adapted from table 14 (page 47) company submission & table 56 (page 149) company appendix E

There is a smaller difference in results according to PD-L1 expression status compared to the differences observed in CheckMate 275

*TTR and DoR data from CheckMate 032 are yet to be published – anticipated to be published in Q1 2018

CONFIDENTIAL

CheckMate 032: Progression free survival

Kaplan-Meier plot

Figure redacted AIC

Median PFS months (95% CI)	
No PD-L1*	2.89 (1.05, 6.51)
PD-L1<1%	2.76 (1.41, 6.51)
PD-L1 ≥1%	5.45 (1.41, 11.71)

Primary clinical database lock (24th March 2016); *No quantifiable PD-L1
Source: Figure 27 (page 149), company appendix E

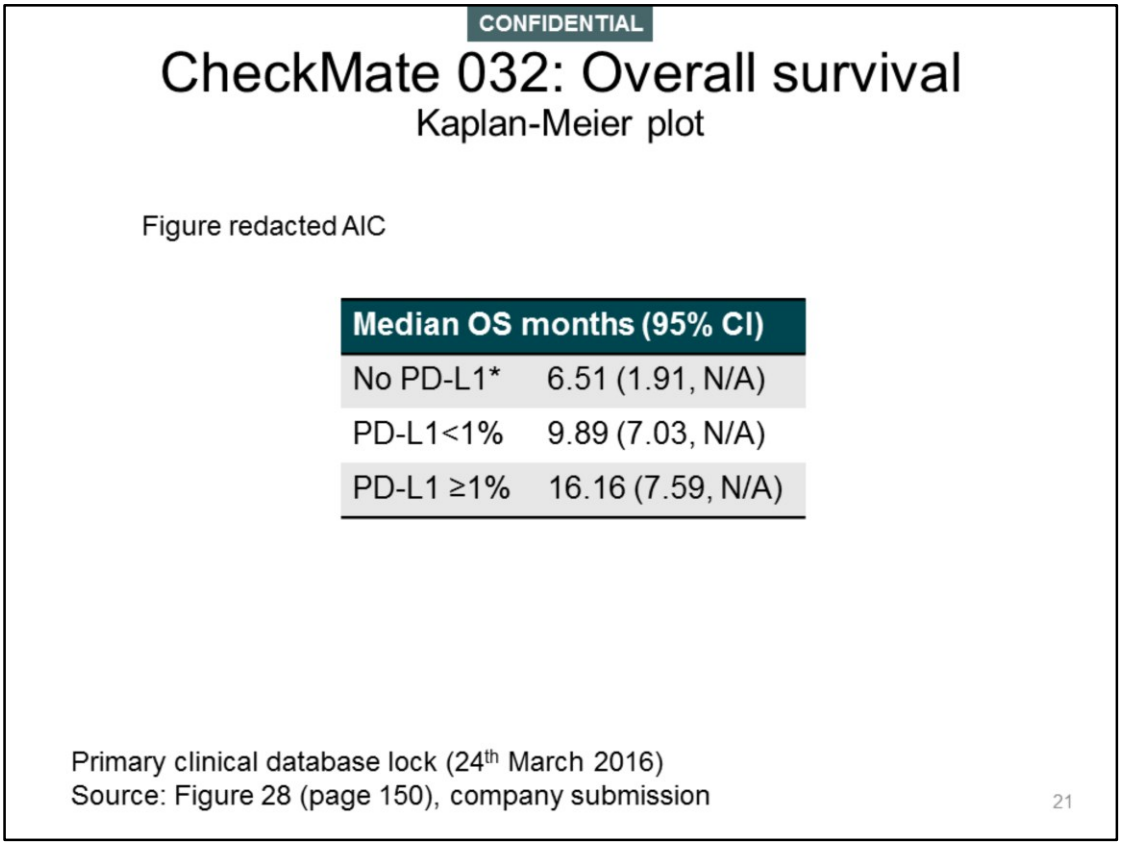
20

Source: Figure 27 (page 149), company appendix E

*No quantifiable PD-L1
Of 18 (23.1%) censored patients, [redacted] had their PFS time censored on either the date of last on-study tumour assessment or date of last assessment prior to subsequent anti-cancer therapy. The most common reason for censoring among these patients was [redacted]. PFS rates (95% CI) were [redacted] at three months, [redacted] at six months and 20.8% (12.3 to 30.9) at 12 months.

Median PFS for patients in the PD-L1 ≥1% cohort was longer than in the all-treated population

Certain PFS data from CheckMate 032 are yet to be published – anticipated to be published in Q1 2018



Source: Figure 28 (page 150), company submission

Median OS was 9.7 months (95% CI 7.3 to 16.2) and 46 (59%) of 78 patients had died at the time of data cut-off. OS rates (95% CI) were [redacted] at three months, [redacted] at six months, and 45.6% (34.2 to 56.3) at 12 months. Median follow-up for OS (time between dose date and last known date alive or death) for all nivolumab monotherapy treated UC patients was 9.69 months (range: 0.7 to 20.7 months).

Median OS for patients in the PD-L1 ≥1% cohort was longer than in the all-treated population.

*Certain OS data from CheckMate 032 are yet to be published – anticipated to be published in Q1 2018

Overview of clinical effectiveness

Outcome	CheckMate 275		CheckMate 032
	Initial database lock: 30 May 2016 n=265[c]	Latest database lock: 2 Sep 2016 n=270[c]	n=78
ORR, n (%), [95% CI]	52 (19.6), [15.0–24.9]	54 (20.0), [15.4–25.3][b]	19 (24.4) [15.3–35.4]
TTR, median (IQR), months	1.87 (1.81–1.97)[a]	1.94 (1.84–2.50)[b]	1.48 (1.25–4.14)
DOR, median (95% CI), months	NR (7.43–NR)[a]	10.35 (7.52–NR)[b]	NR (9.92–NR)
PFS, median (95% CI), months	2.00 (1.87–2.63)[a]	2.00 (1.87–2.63)[b]	2.78 (1.45–5.85)
OS, median (95% CI), months	8.74 (6.05–NR)[a]	8.57 (6.05–11.27)[b]	9.72 (7.26–16.16)

Source: Table 11 (page 43), company submission

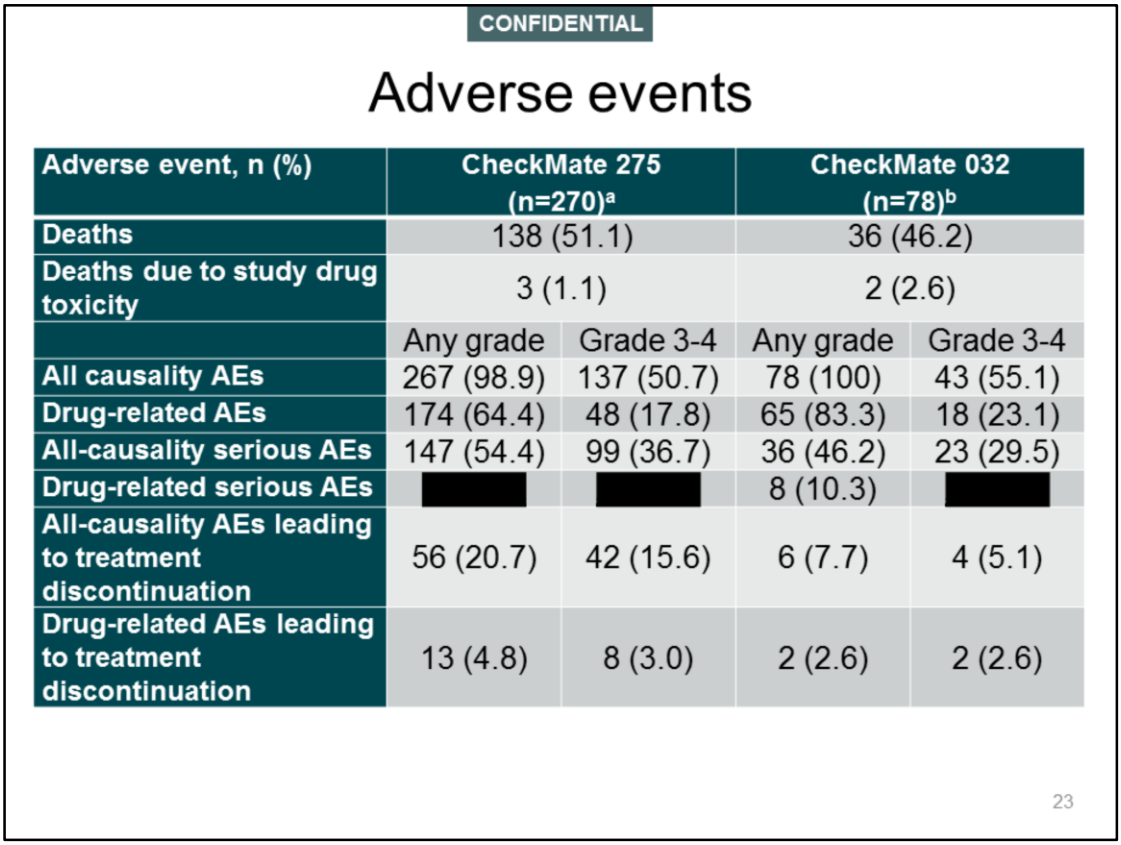
[a]Minimum follow-up of 6 months from the date of first dose. [b]Minimum follow-up of 8.3 months. [c]Follow-up for the latest database lock was sufficient to include 5 patients from Japan who were not included in efficacy analyses in the initial database lock.

CI = confidence intervals; DOR = duration of response; NR = not reached.ORR = objective response rate; OS = overall survival; PFS =progression free survival; TTR = time to response

22

Results from the second database lock of CheckMate 275 (2 September 2016) were consistent with those from the primary analysis database lock in terms of ORR, PFS and OS

Clinical results in key characteristic subgroups can be seen in company appendix E, table 55 and table 57



Adverse event, n (%)	CheckMate 275 (n=270)[a]		CheckMate 032 (n=78)[b]	
Deaths	138 (51.1)		36 (46.2)	
Deaths due to study drug toxicity	3 (1.1)		2 (2.6)	
	Any grade	Grade 3-4	Any grade	Grade 3-4
All causality AEs	267 (98.9)	137 (50.7)	78 (100)	43 (55.1)
Drug-related AEs	174 (64.4)	48 (17.8)	65 (83.3)	18 (23.1)
All-causality serious AEs	147 (54.4)	99 (36.7)	36 (46.2)	23 (29.5)
Drug-related serious AEs	[redacted]	[redacted]	8 (10.3)	[redacted]
All-causality AEs leading to treatment discontinuation	56 (20.7)	42 (15.6)	6 (7.7)	4 (5.1)
Drug-related AEs leading to treatment discontinuation	13 (4.8)	8 (3.0)	2 (2.6)	2 (2.6)

Source: Table 23 (page 72/73), company submission

* Company “It is not anticipated that certain outcomes of the overall safety analysis in CheckMate 275 and CheckMate 032 will be published. These unpublished data are commercially important to Bristol-Myers Squibb”

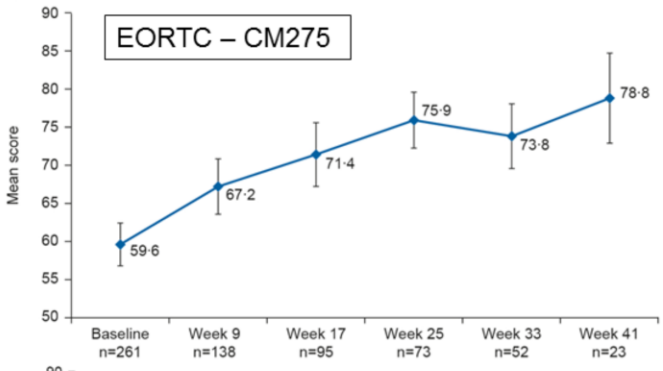
The majority of treated patients experienced at least one AE regardless of causality, during treatment with nivolumab or within 30 days of the last nivolumab dose. As of their respective clinical database locks, a total of 138 (51.5%) patients and 36 (46.2%) patients in the CheckMate 275 and CheckMate 032 trials had died, respectively. The proportion of deaths due to study drug toxicity was 1.1% and 3%, respectively. All-cause AEs leading to treatment discontinuation were reported in 20.7% and 7.7% of patients in CheckMate 275 and CheckMate 032, respectively.

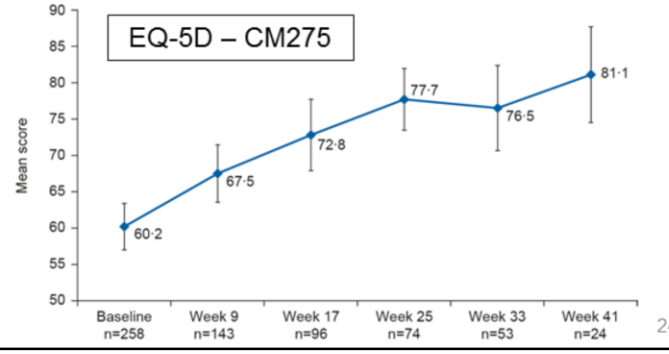
Striking difference in terms of deaths due to study drug toxicity.

Health-related quality of life (HRQoL)

CheckMate 275 and CheckMate 032

- HRQoL was assessed via the EQ-5D-3L in both CheckMate 275 and 032
- CheckMate 275 also collected responses to the EORTC QLQ-C30
  - Commonly used in oncology trials
- At the latest database lock (CM275); 4/9 symptom scales in the EORTC showed improvements, and EQ-5D results were consistent with the initial database lock
- CheckMate 032 reported improvements in EQ-5D-3L results over time





Source EORTC graph: Figure 20 (p54), CS
Source EQ-5D graph: Figure 22 (p56), CS

24

The model used the EQ-5D data from CheckMate275. This is covered in the cost-effectiveness section.

CheckMate275: Patient-reported outcomes data for the measurement of HRQoL was assessed via the EORTC QLQ-C30 questionnaire and the EQ-5D-3L questionnaire in CheckMate 275. Due to the limited study follow-up, interpretations of EORTC QLQ-30 results are limited to the first 41 weeks of follow-up for the all-treated population. Overall, patient HRQoL continued to increase or was maintained throughout the trial from baseline to Week 41.

CheckMate032: Patient-reported outcomes data for the measurement of HRQoL was assessed via the EQ-5D-3L . A total of 73 (93.5%) UC patients treated completed the EQ-5D VAS questionnaire at baseline and the mean baseline EQ-5D VAS score was 72.4 (SD 24.5). Overall, the mean EQ-5D VAS score increased over time. By Week 19, clinically meaningful improvements (>7-point change from baseline) were reported and the average EQ-5D VAS score was >80 points. The EQ-5D VAS continued to improve through Week 61. After week 61, the sample size was too small to interpret (<10).

CheckMate results

ERG comment

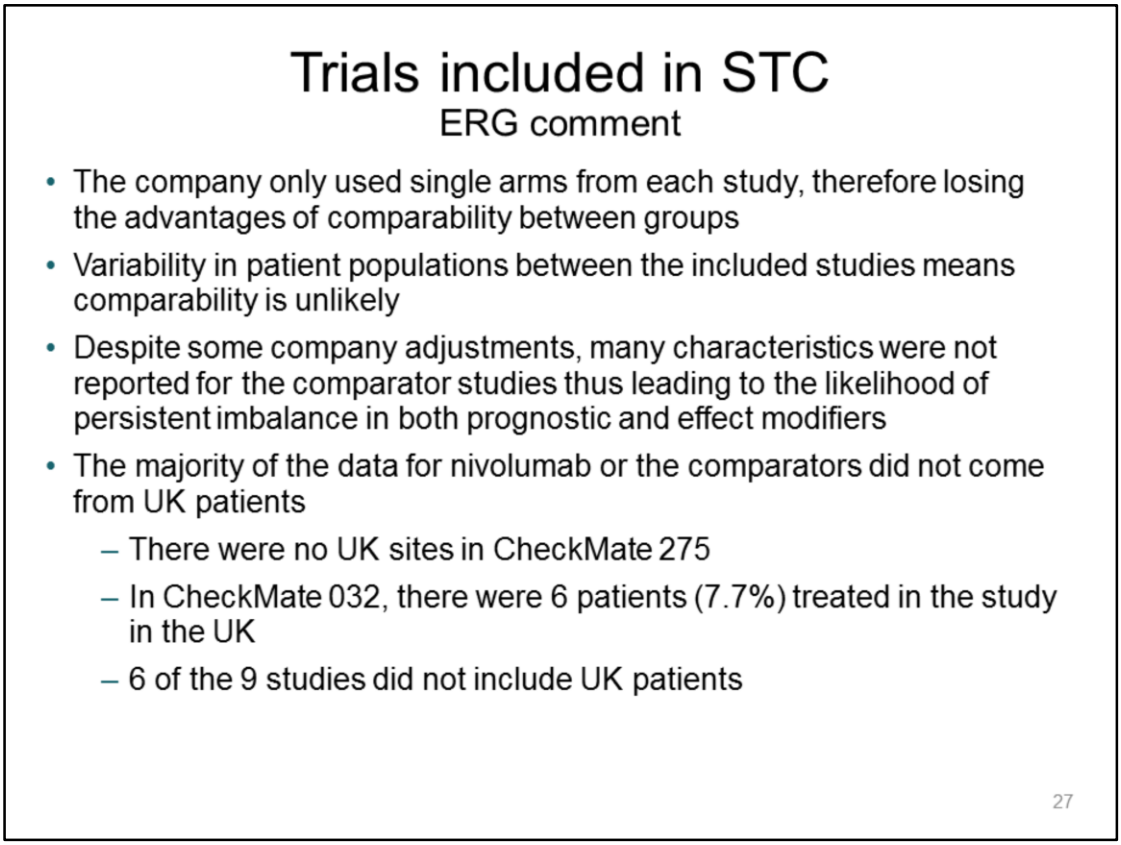
- The outcomes for nivolumab in CheckMate 275 are generally worse than in the CheckMate 032 trial
  - Given the low sample sizes, this could be explained by sampling error
- In CheckMate 032, 23% switched from nivolumab upon disease progression to combination treatment with ipilimumab
- There appeared to be little change between the database locks
- The company did not provide the recent data, as requested by the ERG
- There was a statistically significant difference in OS between the PD-L1 < 1% and PD-L1 >= 1% subgroups
  - The company did not perform a indirect treatment comparison for these subgroups, citing unavailability of PD-L1 status evidence in the comparator studies
  - PD-L1 status is unimportant for the comparators given their mode of action, therefore indirect treatment comparison should have been undertaken
  - Lack of information on other baseline characteristics did not preclude their inclusion in the prediction model for the STC, since missing data was imputed

25

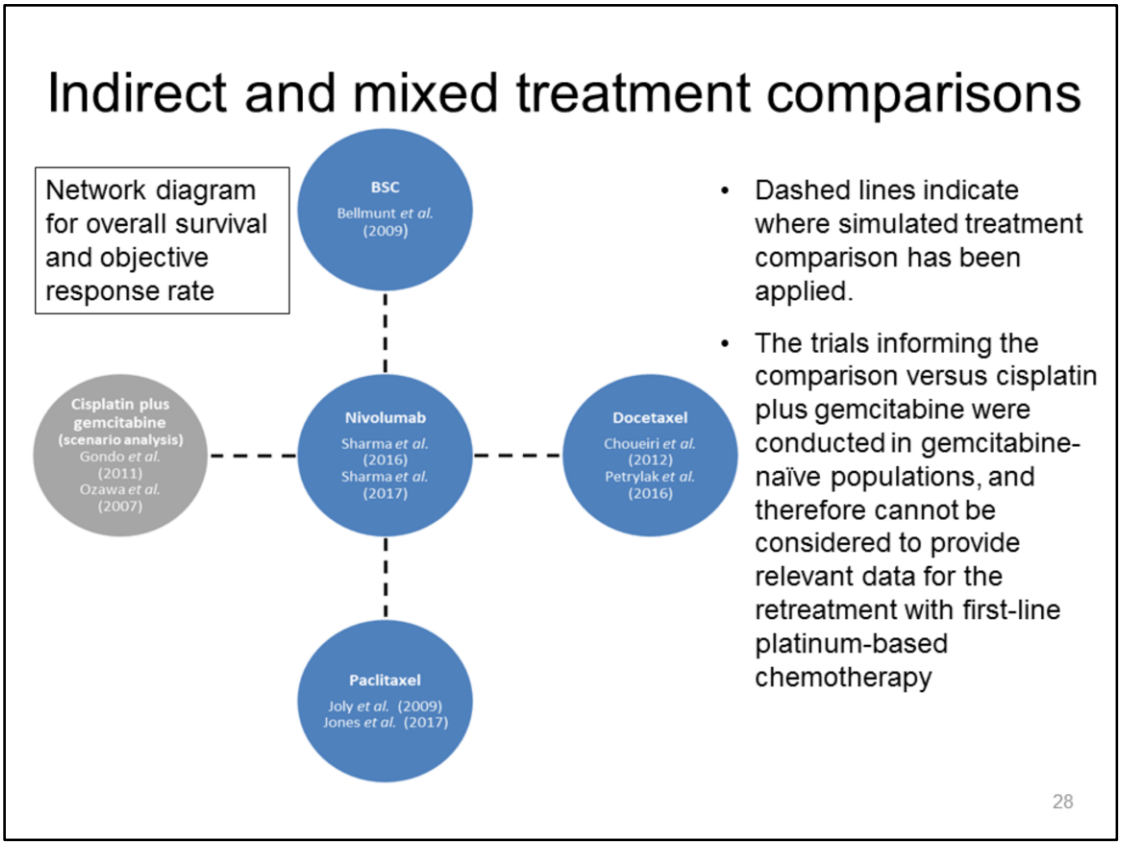
Indirect and mixed treatment comparisons

- The systematic literature review (SLR) identified no randomised controlled trials (RCTs) directly comparing nivolumab with the relevant comparators or placebo
- 12 eligible trials were identified for paclitaxel, docetaxel and BSC, 3 of which were excluded because they did not reflect UK practice
- No relevant trials were identified for retreatment with first-line platinum-based chemotherapy, relevant to the population in this appraisal
  - Two trials were identified for cisplatin plus gemcitabine, however all patients had received MVAC in first-line treatment
  - Results from this comparison were presented in a scenario analysis
- The network for nivolumab and its comparators is disconnected; there are no direct links between nivolmab and it's comparators
  - No direct or indirect links between the nivolumab and comparator trials were identified
  - Hence the indirect comparison was conducted using simulated treatment comparison (STC) methodology

26

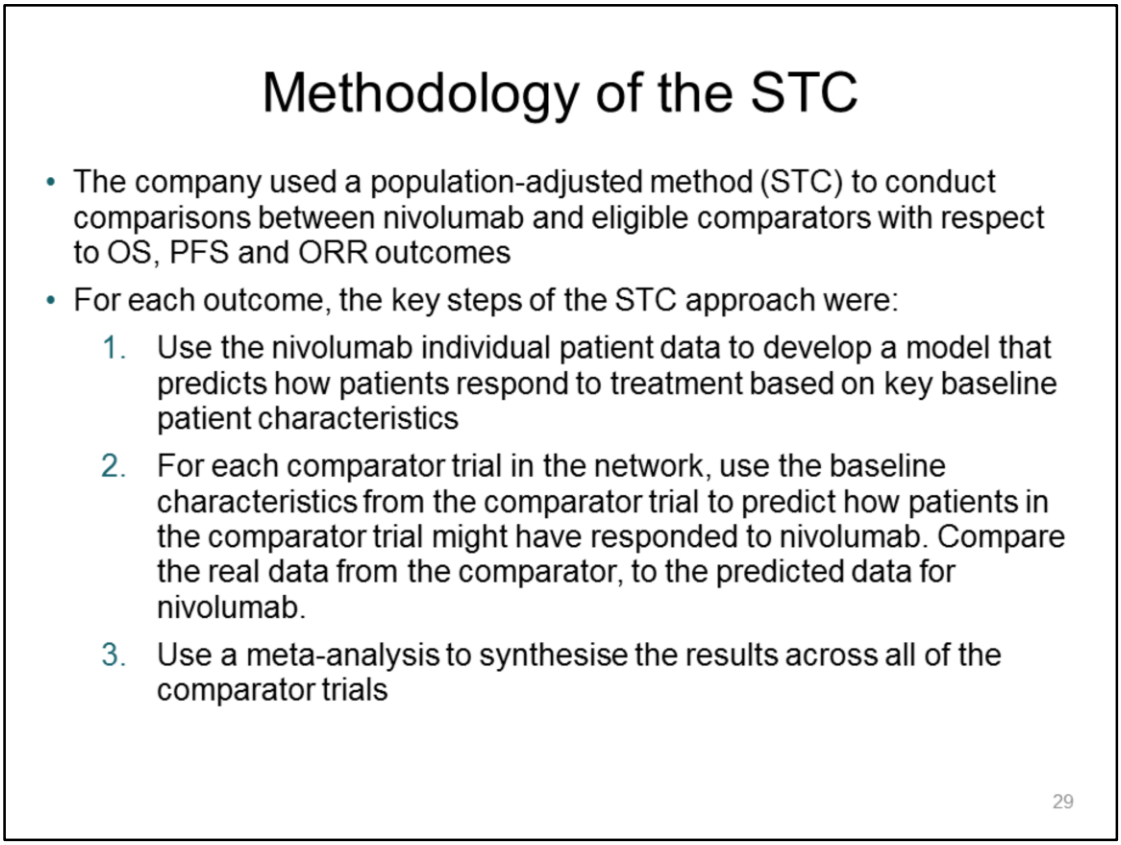


Only one of the studies was conducted exclusively in the UK (Jones et al 2017), one study included some patients from the UK (CheckMate 032: six out of 78), one study was conducted in multiple countries, but it was unclear whether this included the UK (Bellmunt et al 2009) and the remaining six studies did not include UK patients

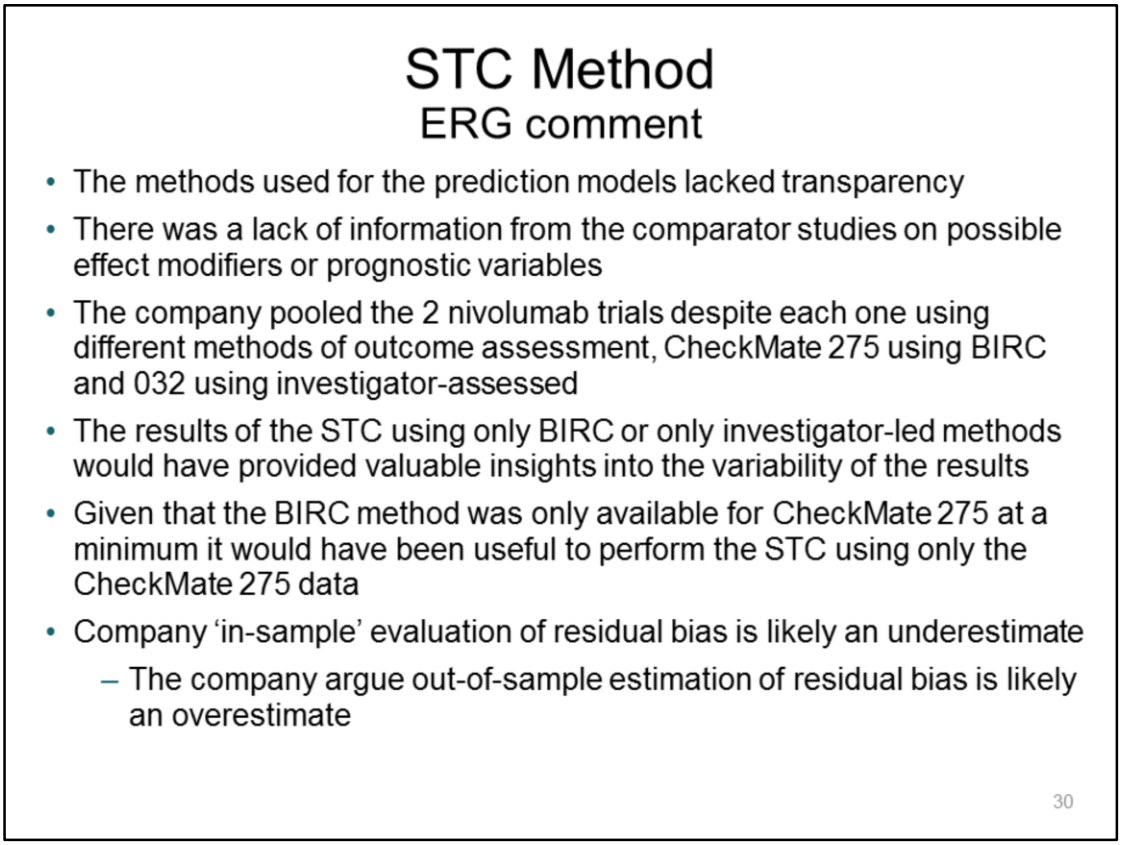


Source: Figure 24 (page 60), company submission

For each comparator trial, and each outcome, the response to nivolumab was estimated by applying the final prediction model to the baseline characteristics in the trial in order to produce adjusted values of the outcome.



ERG: Ideally, for each outcome, the STC should adjust for all the effect modifiers and prognostic variables. However, this is rarely possible, as some effect modifiers and prognostic variables may not be reported by all of the trials or may not be known (for example, as yet undiscovered genetic markers). The company followed the recommendations in the NICE DSU TSD 18. However, we reiterate an unanchored STC '…effectively assumes that absolute outcomes can be predicted from the covariates; that is, it assumes that all effect modifiers and prognostic factors are accounted for. This assumption is very strong, and largely considered impossible to meet. Failure of this assumption leads to an unknown amount of bias in the unanchored estimate".

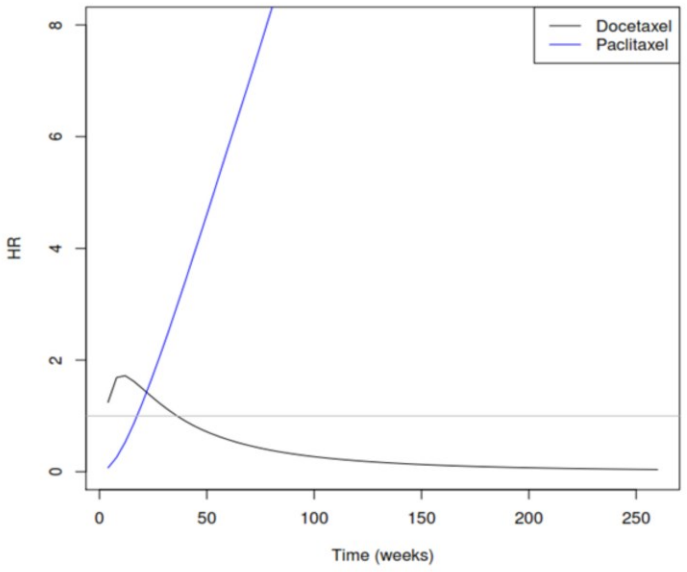


The company state the out-of-sample method of estimating residual bias would not provide a good estimate because:

1. The method described NICE DSU TSD 18 involves a comparison of the between-study variability in the observed and predicted data. However, in this case, there was very limited data to estimate the between-study variability
2. In this case, the 'out-of-sample' method is likely to overestimate the amount of residual bias for the survival outcomes

Progression free survival

Network meta-analysis



Source: Figure 28 (page 66), company submission

- PFS was evaluated using a fractional polynomial
- Second order (P1=0, P2=0) fixed effect model was used because it had clinical plausibility and the lowest DIC
- The network for PFS does not include BSC or platinum-based chemotherapy
- HRs >1 favour nivolumab
- For docetaxel, the HR is initially greater than 1, indicating that patients receiving docetaxel have a higher hazard, but over time the HR decreases

31

As well as in the CheckMate 032 and CheckMate 275 trials, PFS was reported by three comparators studies, for docetaxel and paclitaxel. Jones et al. (2017) did not report a definition for PFS. The median PFS ranged from 1.58 months in response to docetaxel and placebo to 4.1 months in response to paclitaxel.

In all three studies evaluating PFS (Choueri et al. (2012), Jones et al. (2017) (paclitaxel) and Petrylak et al. (2016)) patients would have had a better response to nivolumab than patients in the nivolumab trials.

HRs and credible intervals for each comparator at any given time interval can be seen in Table 20 (p66) of the company submission

Progression free survival

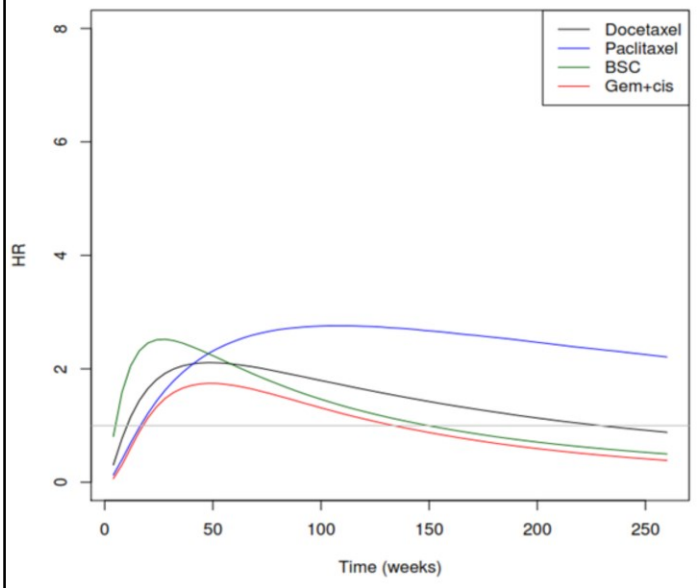
Data included in the STC

Trial ID	Treatment arm	N	PFS definition	Median PFS months (CI)
Sharma et al. (2017) CheckMate 275	Nivolumab	265	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) 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)
Choueiri et al. (2012)	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)	Paclitaxel	65	NR	**4.1** (80% CI 3 to 5.6)
Petrylak et al. (2016)	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)

Source: Table 25 of CS Appendix D

Overall survival

Network meta-analysis



- A fractional polynomial NMA was favoured over a proportion hazards approach because of different mechanisms of action for the treatments
- A second order (P1=0, P2=0) fixed effect model was used because it provided the most clinically plausible extrapolations
- It estimates hazard ratios (HRs) over time for each pairwise treatment comparison
- HRs > 1 favour nivolumab

Source: Figure 26 (page 63), company submission

33

OS was reported by seven studies, including five for the four comparators with two for docetaxel. All of the studies except Bellmunt et al. (2009) reported a definition of OS. Median survival was reported in all of the studies except Gondo et al. (2011), which reported a mean OS of 10.5 months. Median OS ranged from 4.6 months in response to BSC to 9.7 months in response to nivolumab.

In terms of OS, these data suggested that patients in Choueri et al. (2012) (docetaxel and placebo), Petrylak et al. (2016) (docetaxel) and Gondo et al. (2011) (Gemcitabine and cisplatin) would have had on average a better response to nivolumab than patients in the nivolumab trials. However patients in Bellmunt et al. (2009) (BSC) and Jones et al. (2017) (paclitaxel) would have had on average a poorer response.

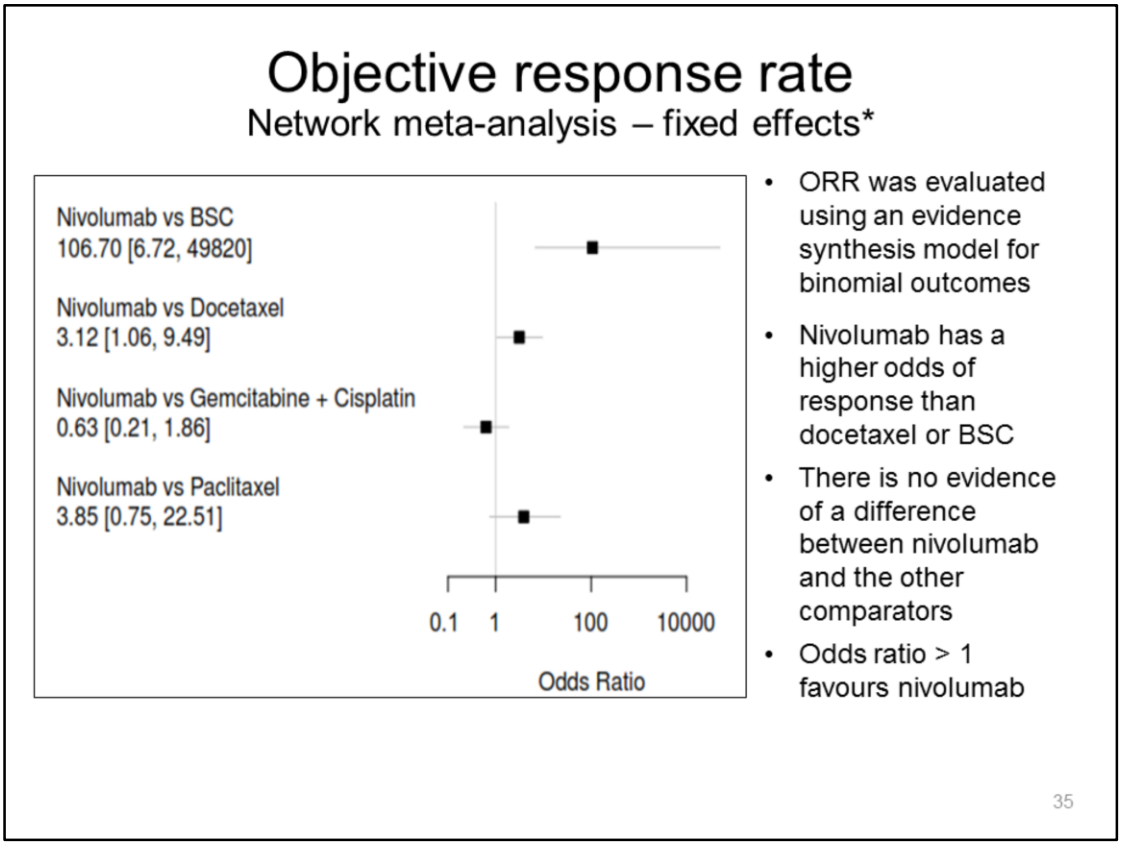
HRs and credible intervals for each comparator at any given time interval can be seen in Table 18 (p63/64) of the company submission

Overall survival

Data included in the STC

Trial ID	Treatment arm	N	Survival definition	Median OS months (CI)
Sharma et al. (2017) CheckMate 275	Nivolumab	265	From first dose and last known date alive or death	8.74 (95%CI 6.05 to NR)
Sharma et al. (2016) 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)	BSC	117	NR	4.6 (95% CI 4.1 to 6.6)
Choueiri et al. (2012)	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)	Paclitaxel	65	From the date of randomisation	8 (80% CI 6.9 to 9.7)
Petrylak et al. (2016)	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)	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)

Source: Tables 24 and 27 of CS Appendix D



Source: Figure 30 (page 68), company submission

*Fixed effect model for ORR was used in the company base case. A random effect model was also presented in figure 19 of the company appendix.

The deviance information criterion (DIC) was used to evaluate model fit and guide the best choice of model. The Fixed effect model had the best fit and was used for ORR in the main analysis.

Eight studies reported ORR, including six for the four comparators. Only one study of paclitaxel by Jones et al. (2017) did not. Four comparator studies did not report a definition of ORR. The ORR ranged from 0% in response to BSC to 40% in response to gemcitabine and cisplatin.

STC results
ERG comment

Overall Response Rate

- Main analysis using the fixed effect model presented finds that nivolumab is significantly better than BSC and docetaxel
- No significant differences were found for nivolumab paclitaxel and gemcitabine
- In the random effects model nivolumab is only superior to BSC*

OS and PFS

- The results of the STC show that nivolumab is superior to all comparators at most time points
  - The credible intervals for the HRs are wide, crossing 1 in many cases
- No formal comparison was made of AEs between the comparators
  - However, it appears rates for nivolumab are lower or comparable to those for the comparators
- Naïve indirect comparison results not reported

36

*statistically significantly different

Clinical effectiveness conclusions

- No randomised control trials, or disconnected network of evidence
- Two single arm studies of nivolumab, included in a STC together with the single arms from some RCTs
- The methods of the STC largely followed DSU TSD18, however:

"unless all baseline characteristics that might be prognostic variables and effect modifiers are incorporated in any model to adjust for bias, it is unclear what the size of any bias might be" (DSU TSD18)

- Comparison with gemcitabine plus cisplatin excluded from the base-case

ERG Comment:

- It's not clear how the fit of the prediction model was tested
- To compound the uncertainty, the numbers of actual patients are small for all comparisons and not all studies provided data for all outcomes
- The survival data are not fully mature in the nivolumab trials
- Not all study outcomes are based on independent review
- The polynomial fraction model appears valid and flexible for estimating HRs, however few functional forms were presented leaving doubt as to the most appropriate

37

- EPAR - OS outcomes for nivolumab and chemo were similar in those with PD-L1 <1%.

[in the PD-L1 <1% group] "..the 12-month survival rate decreased to 33.5%, which appeared similar to those described in larger trials with single-agent chemotherapy (25%-30%)"

- Further mention of worse outcomes for those with PD-L1<1%

"…for patients with tumour PD-L1 <1% a shorter median OS was observed in the vast majority of subsets." & "The SmPC has been updated to reflect that results from post-hoc, exploratory analyses indicate that in patients with low (e.g. <5%) to no tumour PD-L1 expression […] might contribute to the clinical outcome"

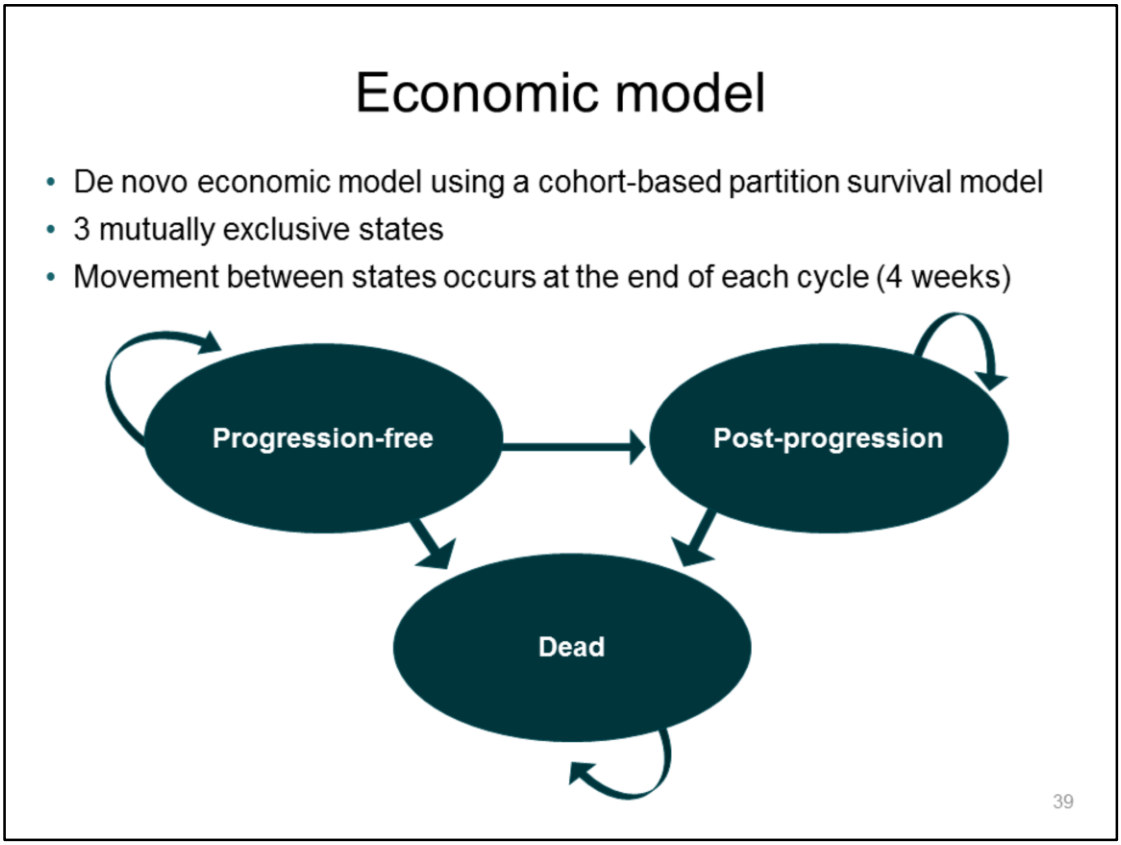
- Obligation for the company to complete these post-authorisation measures (by 30/06/2018)

1. The value of biomarkers to predict the efficacy of nivolumab and/or nivolumab + ipilimumab combination therapy should be further explored, specifically to further investigate the value of biomarkers other than PD-L1 expression status at tumour cell membrane level by IHC […] as predictive of nivolumab and/or nivolumab + ipilimumab combination therapy efficacy.
2. To further explore in UC patients the early identification of those who do/do not respond to treatment with nivolumab, as well as to evaluate the association between improved clinical outcomes to nivolumab and the presence of mutational and neoantigen load, and PD -L1 expression on tumour- and tumour associated immune cells …

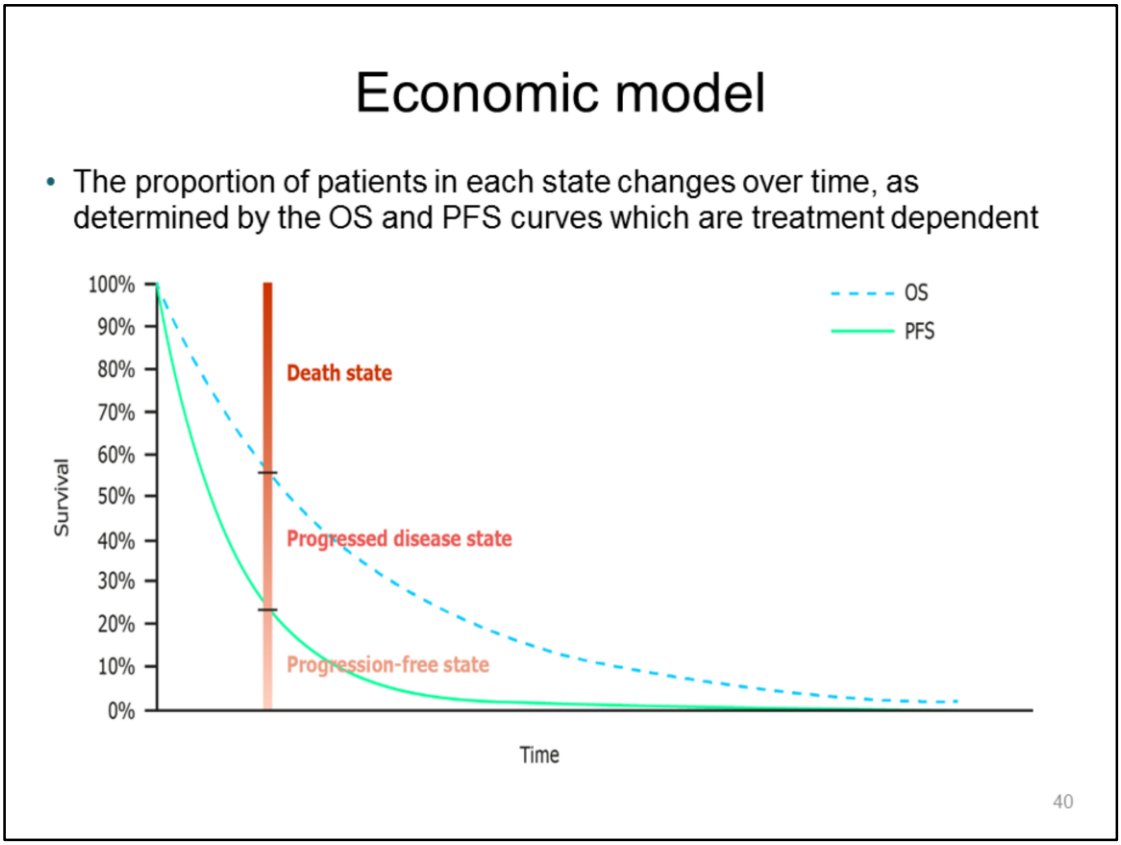
Cost-effectiveness

Company submission section B.3

38



Consistent with TA272 model structure



Source: figure 33 of the company submission

This choice of model structure was made to capture the progressive nature of UC disease and is consistent with previous submissions to NICE relating to metastatic cancers, including the only previous submission in this specific indication (TA272, 2013).

Model inputs

	Company model	ERG comment
Population	• Consistent with the CheckMate 275 & 032 trials • Age, gender, weight and body surface area (BSA) are all included in the model	• Consistent with the final scope • Characteristics included are relevant for calculating background mortality
Comparators	• Paclitaxel: 80mg/m^2 Q3W of a four-week cycle • Docetaxel: 75mg/m2 Q3W • Best supportive care (BSC)	Cisplatin plus gemcitabine was not included in the base-case
Perspective	NHS+PSS (England and Wales)	Appropriate
Time horizon	Life time horizon	Appropriate
Cycle length	Four weeks to account for length of treatment cycles	Appropriate
Discounting	3.5% per year for cost and utilities	Appropriate
Stopping rule	• None (base-case) • 75% of those still on treatment discontinue after 2 years (scenario)	No comment
Utilities source	CheckMate275	Utilities should have been pooled with CheckMate032

41

Population:
- Weight and BSA influence the calculation of dose

Comparator
- Company scenario analysis (not base-case), in which cisplatin + gemcitabine was added as a comparator. The company said it wasn't suitable for the base-case as the population in the Gondo (2011) study differed from the UK population in that the study population received MVAC in first line instead of cisplatin plus gemcitabine.

Stopping rule
- For this analysis the stopping rule was applied to 75% of patients who were yet to discontinue. It was assumed that 25% remained on treatment to reflect a potential minority of patients and/or their clinician who chose to remain on treatment for a longer time period.
- Assumed treatment benefit = life time / no treatment waning

Economic model

ERG comment

- Partitioned survival model has limitations
  - Survival functions are modelled independently even though there are dependencies such as that progression is a prognostic factor for mortality
  - Transition probabilities are not estimated for each possible transition between health states
- The use of response-based analysis without reflecting responder and non-responder states in the model structure is inappropriate
  - The proportions of responders and non-responders remain the same throughout the model time horizon, an unrealistic assumption
  - The proportion of responders will increase over time as they are likely to survive longer

42

Limitations in using partitioned survival model can lead to inappropriate extrapolation

Survival analysis

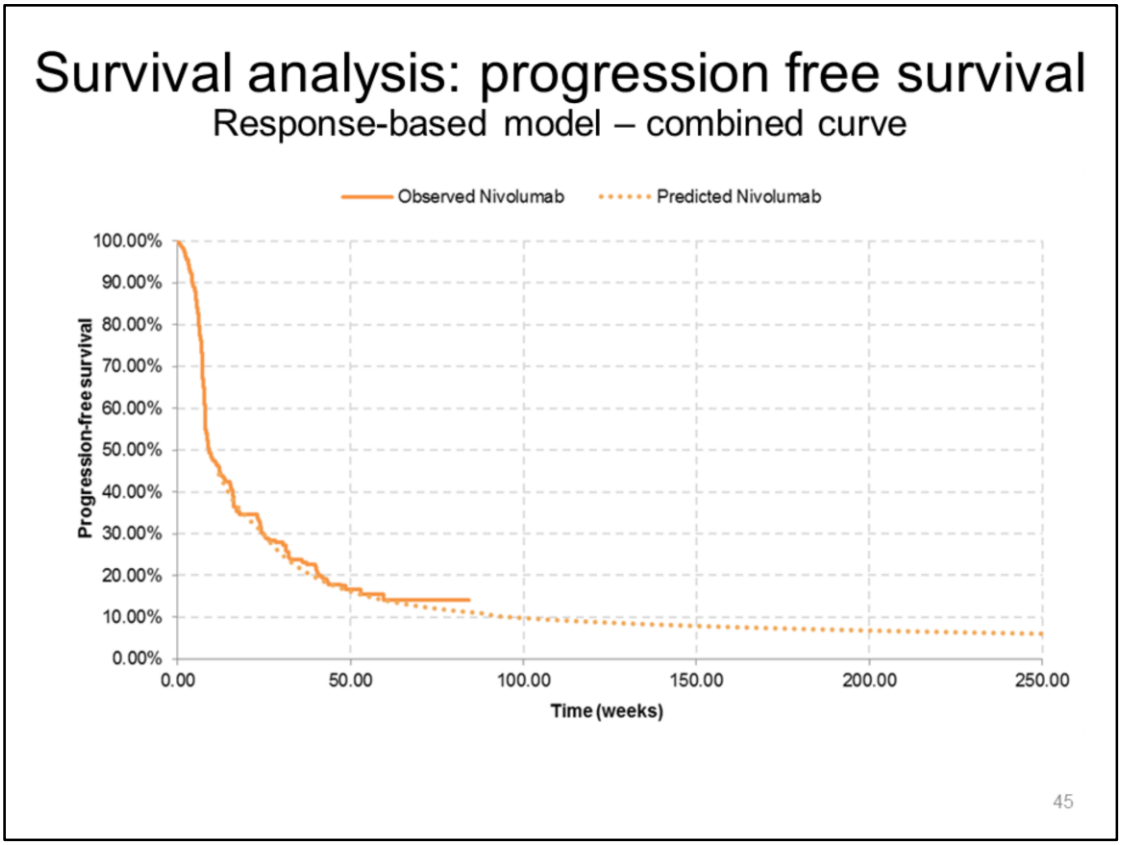
- Parametric time-to-event survival curves were plotted to estimate progression free survival (PFS), overall survival (OS), time to treatment discontinuation (TTD)
- Standard survival modelling approaches may not accurately reflect the mechanism of action of immunotherapy
- A response-based modelling approach was adopted
  - Fit parametric survival curves to the responders and non-responders separately to more accurately characterise the hazard and survival curve in these two groups
- There can be a risk of immortal time bias in response-based models
  - This occurs when responder and non-responder curves are plotted immediately following the start of treatment
  - In the model responders cannot progress or die until their response whereas non-responders can do so at any point
  - The curve for responders may overestimate long-term survival

43

Survival analysis

- To overcome immortal time bias, landmark analysis was undertaken
  - OS and PFS of responders and non-responders is estimated together until a specified landmark point at after which different survival curves are fitted for each group
  - The base-case landmark point coincides with the median time to response in the CheckMate trials – 8 weeks
  - Kaplan-Meier estimates for the whole group are used until the landmark point, then parametric time-to-event models are fitted after
- For the combined curves, responder and non-responder curves were weighted based on patients measured as being progression-free and alive at 8-weeks, these weights were then held constant
  - Assuming constant weights is a conservative assumption, as weighting would be expected to increase in favour of responders
- Generalised gamma was selected for OS and PFS simultaneously for both responder and non-responder groups because it provided the best overall fit

44



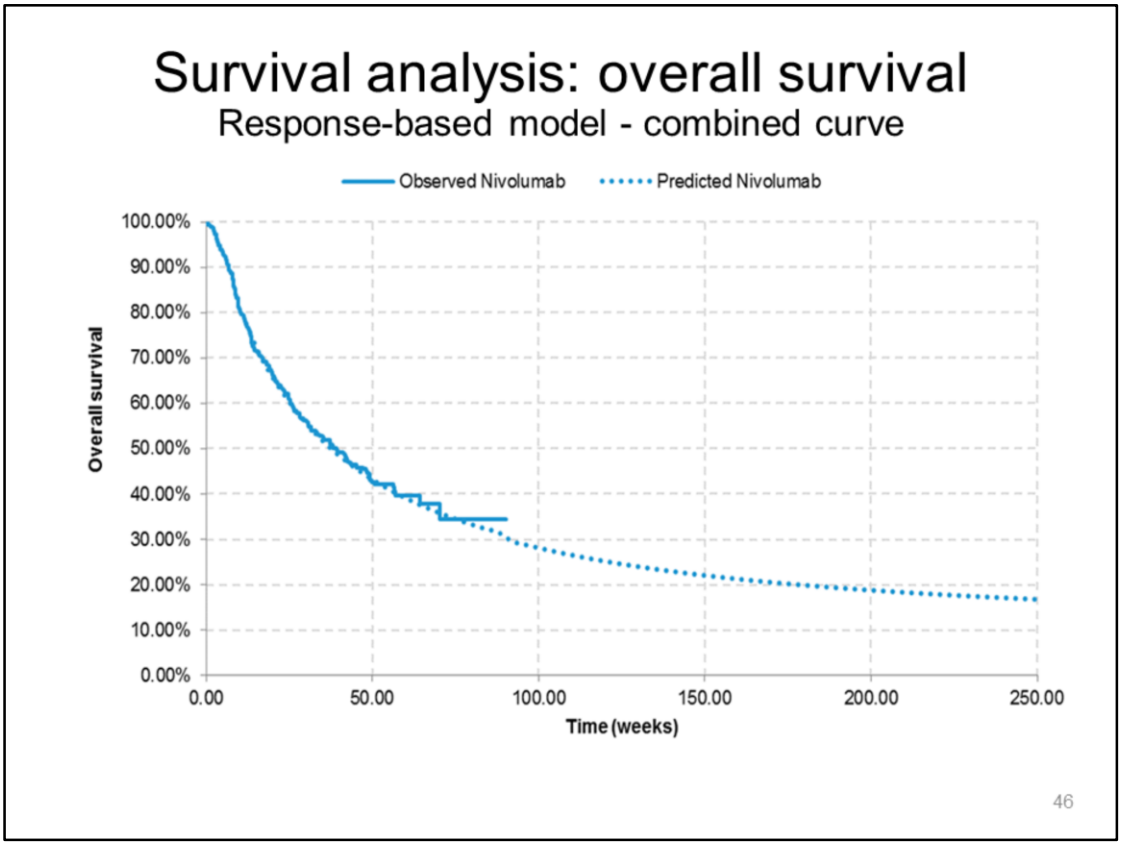
Source: Figure 36 (page 93), company submission

To make the PFS and OS curves suitable for the structure of the economic model, and the application of relative treatment effects, it was necessary to combine the separate responder and non-responder curves. Separate response-based PFS curves can be seen in figure 35 of the company submission.

The final PFS and OS curves for nivolumab are composites of the pre-landmark pooled Kaplan-Meier data from CheckMate 275 and CheckMate 032 and a weighted average of the responder and non-responder curves from the landmark point onwards

The PFS and OS curves were adjusted to account for general population mortality using age-adjusted annual mortality rates based on life tables for England and Wales.

Comparator PFS and OS curves can be seen in figures 40-43



Source: Figure 37 (page 94), company submission

Separated response-based OS curves can be seen in figure 34 of the company submission.

Comparator PFS and OS curves can be seen in figures 40-43

CONFIDENTIAL

Overall survival extrapolation

Data source	Curve	Proportion alive, % 1y	1.5y	2y	3y	4 y	5y	Data source	Curve	Proportion alive, % 1y	1.5y	2y	3y	4y	5y
Nivolumab								**Paclitaxel**							
Model estimates for OS	Gen. Gamma	42.3	33.8	27.5	21.7	18.5	16.6	Model estimates for OS	Gen. Gamma	31.4	17.4	10.6	5.7	3.9	3.2
CheckMate 275	KM	[redacted]	[redacted]	-	-	-	-	Jones et al. 2017	KM	31.6	15.1				
CheckMate 003 (NSCLC)	-	42	-	24	18	-	16	Sideris et al. 2016	KM (Bytescout)	19	8	6	-	-	-
Docetaxel								**BSC**							
Model estimates for OS	Gen. Gamma	25.0	15.7	11.1	7.7	6.4	5.7	Model estimates for OS	Gen. Gamma	14.0	9.0	6.6	5.0	4.4	4.1
Choueiri et al. (2012)	KM	24.3	13.0	-	-	-	-	Bellmunt et al. 2013	KM	21.3	10.7	7.4	1.4	-	-
Sideris et al. 2016	KM	19	8	6	-	-	-								

Abbreviations: BSC: best supportive care; KM: Kaplan-Meier data; NSCLC: non-small cell lung cancer; OS: overall survival.

Source: Adapted from table 55 in the company submission

Survival analysis

ERG comment

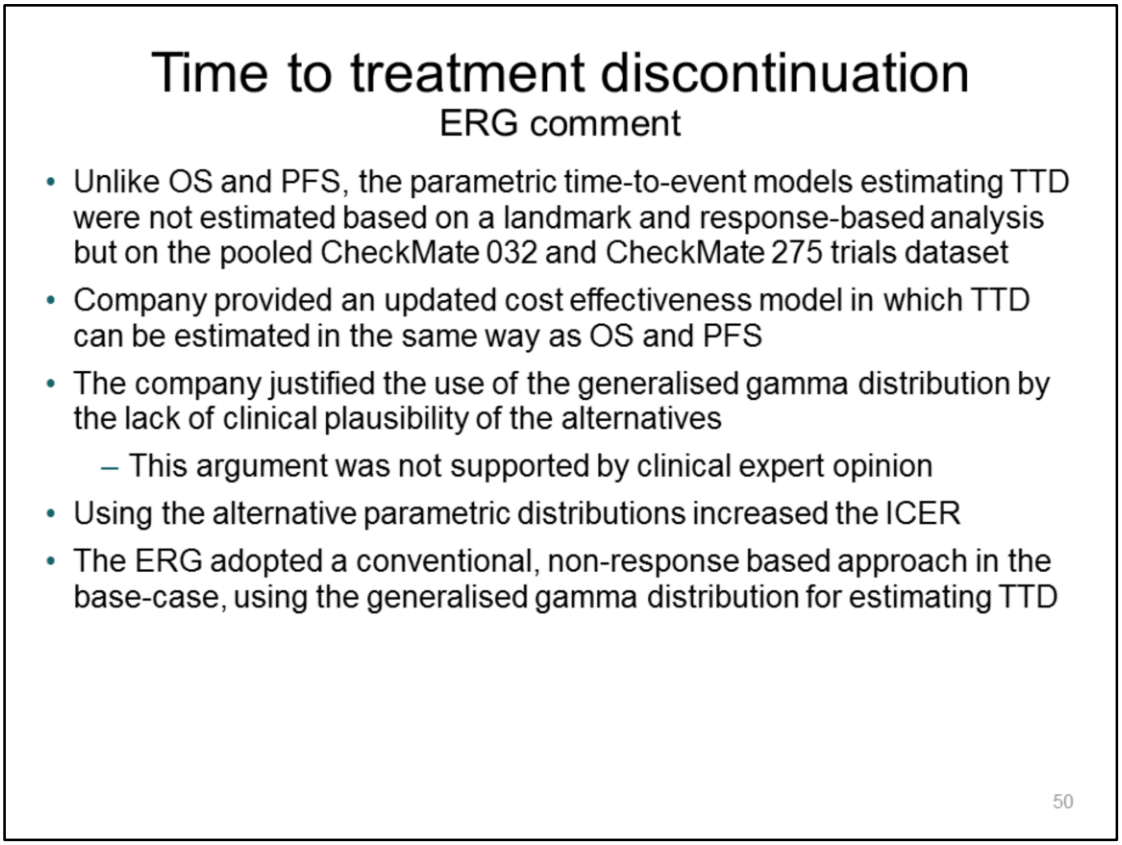
- It was not demonstrated by the company that a response-based model was superior to alternative and flexible modelling approaches
- Standard models provided a good fit for OS and a reasonable fit for PFS
- Responders and non-responders are combined for the indirect comparison, reducing the benefit achieved with a response-based model
- ERG base-case uses conventional parametric time-to-event modelling
- Prefer to use parametric time-to-event model to the estimate survival to the landmark point to avoid the problem of overfitting
- Additional assumptions in response-based model add uncertainty
  - Choice of landmark point has an unpredictable effect on results
  - Only data after the landmark point is used
- Response-based and conventional approaches give different results
  - 2.8 life years (response-based) v 1.84 life years (conventional)
- Limited expert consultation in the choice and validation of the model
- Unrealistic to assume a constant weighting of responder groups

48

Time to treatment discontinuation

- Nivolumab should be administered as long as clinical benefit is observed or until treatment is no longer tolerated by the patient
  - Discontinuation is not based solely on progression
- Time-to-treatment discontinuation (TTD) was estimated through a parametric time-to-event model
  - Generalised gamma was used in the base-case
  - Gompertz and log-logistic showed better statistical fit, but had implausibly long tails, therefore lacking clinical validity
- TTD of the comparators was based on PFS because it's assumed that treatment continues until disease progression or unacceptable toxicity
- Treatment with paclitaxel was assumed to stop after 6 (model) cycles
  - Representing the clinical use of paclitaxel in the UK
- It was assumed that all BSC patients receive this treatment until death
- Scenario analysis in which a proportion of patients (25%) remain on nivolumab after 2 years

49



In the updated cost effectiveness model in which TTD can be estimated in the same way as OS and PFS, the ERG noticed that the company calculated the proportion of responders and non-responders based on the sum of patients in the OS and PFS health states, thereby double-counting patients. The ERG considered it more appropriate to use all responders alive for the calculation of proportion of responders.

TTD curve for nivolumab is presented in the company submission, figure 38, page 96

Relative effectiveness

- HRs for BSC were not available for PFS – it's impact was assumed
- For cisplatin plus gemcitabine, no relevant PFS data were identified in the SLR and therefore this was not included in the STC for PFS
  - Assumed a similar PFS profile as paclitaxel v nivolumab by using the same HR – assumption limits generalisability
- Predicted survival curves for the comparators often underestimate survival when compared with the available trial data, because the STC accounts for differences in characteristics between studies

ERG comment

- There is a discrepancy between the a priori population on which the relative effectiveness is based on the a posteriori population in which the HRs are applied
- HRs used to estimate PFS of BSC and cisplatin plus gemcitabine were based on assumptions, and not supported by clinical evidence
- Not all combinations of parameter values used in the fractional polynomial model result in plausible survival estimates, e.g >100%
  - ERG adjusted the model to prevent this occurring

51

Background mortality

- Company approach:
  - PFS and OS curves were adjusted to account for general population mortality using age-adjusted annual mortality rates based on life tables for England and Wales
  - Due to differences in the rate of mortality between males and females, the annual rates were weighted by gender based on the ratio of males to females (78:22) reported in the CheckMate 275 trial
  - To avoid double counting, background mortality was only applied from week 88 onwards in the model, which is the end of the follow-up period in the CheckMate trials
- ERG comment:
  - Mortality rates implemented did not match ONS life table figures
  - Unconventional company approach, slightly higher mortality
  - Background mortality was applied to the combined responder and non-responder groups which is inappropriate given different prognoses

52

Adverse events

Company approach:

- The rates of adverse events were taken from the clinical trials that inform the PFS and OS curves in the model
- Any all-cause Grade 3 or 4 AEs were included if the incidence was ≥5% and the impact on costs and utilities were incorporated in the first cycle of the model only

ERG comment:

- For nivolumab, CheckMate275 was the only source informing adverse event rates, whereas CheckMate 032 was used for clinical effectiveness
- The company did not justify the selection of the source used to estimate AE rates of the comparator
- Nausea/vomiting, diarrhoea, and ALT increase have an incidence <5% for all treatments included in the cost effectiveness model. Hence it is inconsistent to include these AEs in the cost effectiveness model. The ERG removed these adverse events from its analyses.

53

Health-related quality of life (HRQoL)

- A total of 9 records (6 unique studies) were included in the SLR
- The majority of these studies did not provide enough descriptive data to be deemed consistent with the NICE reference case
- The EQ-5D-3L data from CheckMate 275 were used to inform utility estimates in the health economic model
  - No mapping was required
- Grade 3 and 4 AEs are included in the model in order to represent those AEs that are more likely to have a substantial effect on quality of life
  - Disutilities were sourced from the literature and included in the model
- The number of missing responses for patients still on treatment during the trial was 188 (20.4%)
- Progression and treatment status are statistically significant predictors of missing data, therefore data may not be missing at random
  - The imputed dataset was used to ensure no potential for bias

54

In order to impute these missing values, the multiple imputation by chained equations (MICE) procedure was conducted. MICE is an extension of the multiple imputation method, and uses regression analysis to simultaneously impute values for all the variables in the dataset in one procedure.

Health-related quality of life (HRQoL)

ERG comment

- Inconsistencies in the number of reported observations
  - Interpolated, imputed and valid observations don't sum to the total
- The exclusion of CheckMate 032 utilities, is inconsistent with the pooling of other outcomes
- The imputation of immature trial data is inappropriate as none of the immature observations will be censored due to death of patients
- There was no justification for using multiple imputation in favour of a mixed model to adjust for missing data
- Lack of justification for not using time-dependent utilities
- Dis-utilities for adverse events were inconsistent with those used for a previous nivolumab appraisal, they were derived from the literature
  - It was unclear how the studies were selected – not from the SLR
- Maintained the company's pre- and post-progression utility values, as opposed to on- and off-treatment

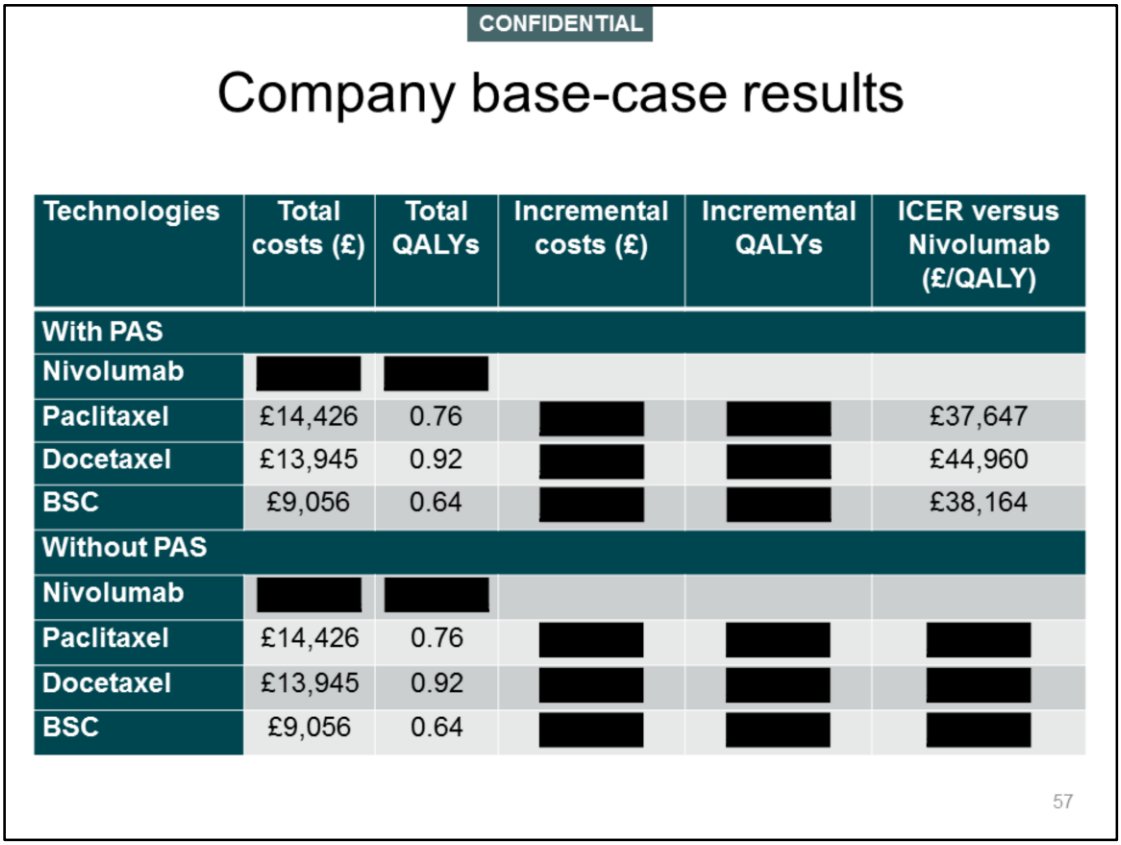
55

Utility values for cost-effectiveness

State	Utility/disutility value: mean (standard error)	95% CI	Source
Pre-progression	Imputed value: 0.718 (0.016) Observed value: 0.713 (0.017)	Imputed value: 0.686 to 0.75 Observed value: 0.679 to 0.747	Imputed from Checkmate 275
Change in utility – pre-progression to post-progression	Imputed value: -0.115 Observed value: -0.061	Imputed value: -0.143 to -0.087 Observed value: -0.123 to -0.055	Imputed from Checkmate 275
Post-progression	Imputed value 0.603 (N/A) Observed value: 0.623 (N/A)	N/A	Checkmate 275
Neutropenia	-0.18	NR	Attard et al. (2014)
Anaemia	-0.09	-0.13, -0.06	Beusterien et al. (2010)
Thrombocytopenia	-0.18	NR	Attard et al. (2014)
Asthenia/Fatigue	-0.12	NR	Attard et al. (2014)
Nausea/vomiting	-0.05	-0.08,-0.02	Nafees et al. (2008)
Diarrhoea	-0.29	NR	Attard et al. (2014)
ALT increase	-0.05	-0.07, -0.03	NICE TA347 (2015)
Leukopenia	-0.09	NR	Frederix et al. (2013)

56

Source: Table 35 company evidence submission



Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER versus Nivolumab (£/QALY)
With PAS					
Nivolumab	[redacted]	[redacted]			
Paclitaxel	£14,426	0.76	[redacted]	[redacted]	£37,647
Docetaxel	£13,945	0.92	[redacted]	[redacted]	£44,960
BSC	£9,056	0.64	[redacted]	[redacted]	£38,164
Without PAS					
Nivolumab	[redacted]	[redacted]			
Paclitaxel	£14,426	0.76	[redacted]	[redacted]	[redacted]
Docetaxel	£13,945	0.92	[redacted]	[redacted]	[redacted]
BSC	£9,056	0.64	[redacted]	[redacted]	[redacted]

Source: Adapted from tables 44 and 45 from the company evidence submission

Base case ICERs in response to clarification – updating the economic model and fixing minor errors
ICERs for Nivolumab v:

- Paclitaxel £37,643
- Docetacel £44,996
- BSC £38,302

CONFIDENTIAL

Probabilistic sensitivity analysis

- Patient age, weight and BSA, costs, resource use, utilities, TTD, PFS and OS were varied
- Incremental costs increased and incremental QALYs decreased compared to the deterministic results

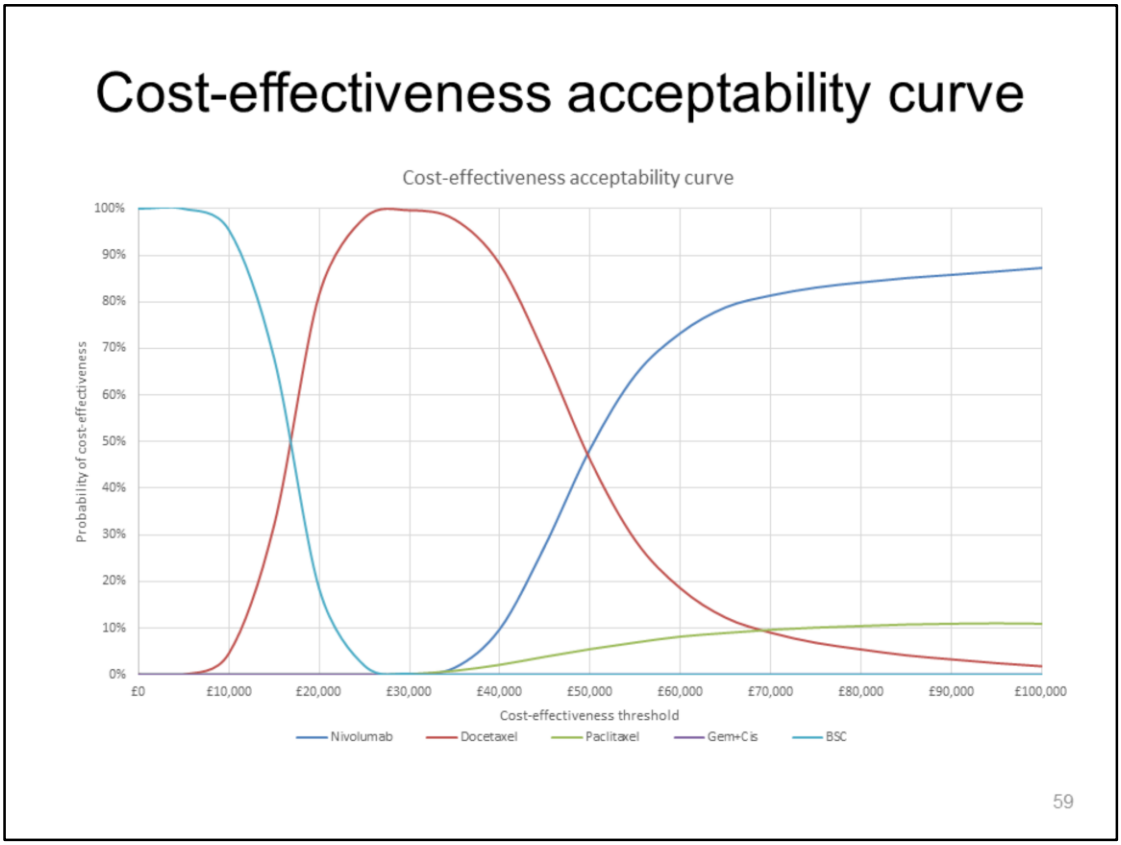
Technologies	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)	Probability of cost effectiveness[a]
Paclitaxel	[redacted]	[redacted]	£46,209	72.10%
Docetaxel	[redacted]	[redacted]	£54,220	49.00%
BSC	[redacted]	[redacted]	£44,698	76.30%
Company scenario				
Cis+gem	[redacted]	[redacted]	£103,568	6.9%

[a]The probability of nivolumab being cost-effective vs the stated comparator at a CE threshold of £50,000/QALY.
Abbreviations: Cis+gem: cisplatin plus gemcitabine; BSC: best supportive care, ICER: incremental cost-effectiveness ratio; QALYs: quality-adjusted life years

58

Sources: Table 46 (page 116) company submission and table 5.18 (page 125), ERG report

The probabilistic results generated during the PSA were similar to the base case results, with a slight increase in the probabilistic ICERs compared with the deterministic analysis

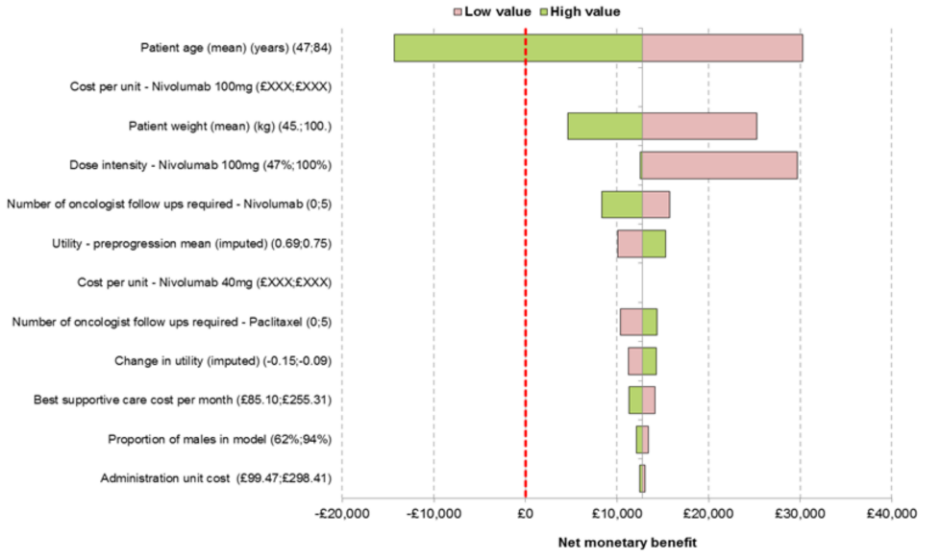


Source: Company model received in response to clarification

CONFIDENTIAL

Deterministic sensitivity analysis

Nivolumab v paclitaxel



60

The DSA results show that the model results are robust to changes to the majority of parameters with only three parameters causing the direction of the ICER to change. Therefore, the key drivers of ICER uncertainty either related to the cost per cycle of nivolumab (e.g. unit price, patient weight) or patient age.

CONFIDENTIAL

Deterministic sensitivity analysis

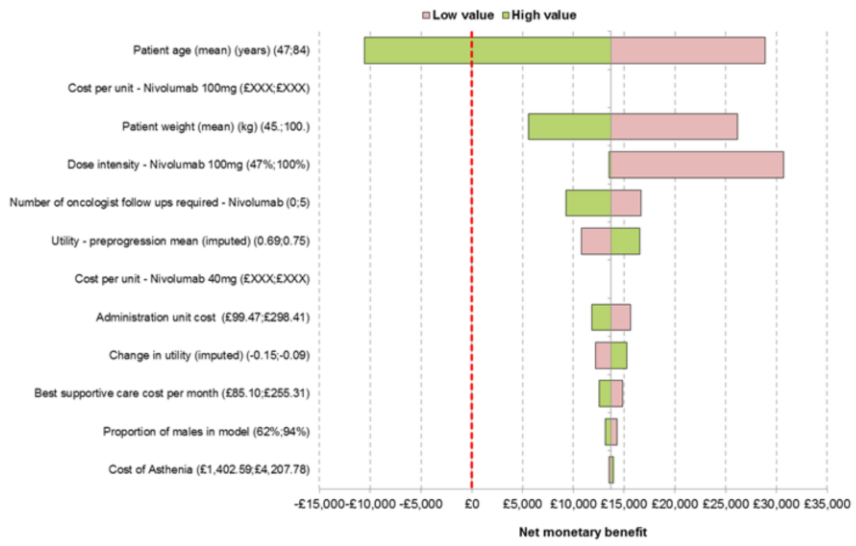
Nivolumab v docetaxel

61

CONFIDENTIAL

Deterministic sensitivity analysis

Nivolumab v BSC



62

Deterministic scenario analysis

Scenario	Scenario info	ICER vs. Paclitaxel	ICER vs. Docetaxel	ICER vs. BSC
Base case	Gen. gamma	£37,647	£44,960	£38,164
1 Survival curves	Landmark 8 weeks			
	Weibull	£101,994	£114,823	£91,372
	Gompertz	£49,010	£59,858	£50,201
	Lognormal	£52,900	£72,044	£53,634
	Log-logistic	£58,279	£78,063	£59,695
	Exponential	£57,998	£70,582	£59,564
	Landmark 26 weeks			
	Gen. Gamma	£34,541	£40,246	£34,774
	Weibull	£50,060	£62,866	£51,378
	Gompertz	£35,655	£41,933	£35,269
	Lognormal	£38,834	£48,610	£38,192
	Log-logistic	£42,475	£54,235	£43,097
	Exponential	£60,279	£76,786	£61,389

63

Sources: Tables 48 – 54 company submission

Deterministic scenario analysis

Scenario	Scenario info	ICER vs. Paclitaxel	ICER vs. Docetaxel	ICER vs. BSC
Base case	Generalised gamma	£37,647	£44,960	£38,164
2 Fractional polynomial model[a]	p1=1, p2=1	£56,073	£59,504	£43,554
3 Exponential piecewise model	Piecewise exponential at 8 weeks	£53,616	£65,450	£55,597
	Piecewise exponential at 26 weeks	£55,681	£71,147	£57,293
4 Vial sharing	Inclusion of vial sharing	£35,651	£42,630	£36,333
5 Stopping rule[b]	Stopping rule included	£31,561	£37,781	£32,743
6 Alternative TTD parametric curves	Weibull	£33,562	£40,141	£34,525
	Gompertz	£183,467	£216,984	£168,053
	Lognormal	£61,810	£73,465	£59,688
	Log-logistic	£61,994	£73,683	£59,851
	Exponential	£28,331	£33,971	£29,866

[a] Second-best fitted fractional polynomial model
[b] Stopping rule applied where after 2 years treatment, 75% of patients still receiving treatment will discontinue treatment

64

Sources: Tables 48 – 54 company submission

Company PSA and DSA
ERG comment

- A number of parameters were excluded from the DSA
  - Survival curves were not explored
  - HRs were not varied
- A number of parameters were excluded from the PSA
  - Excluded HRs and Kaplan-Meier estimates used to estimate nivolumab survival before the landmark, and erroneously included patient characteristics
- The number of iterations (1000) used in the PSA was too small
  - Upped to 10,000 in response to clarification
  - Discrepancies in results between runs remain at 20,000 iterations
- Nivolumab OS and PFS is lower in the PSA than the deterministic analysis
  - Results were more consistent using the conventional model
- Fully incremental analysis did not include all comparators

65

ERG base-case

#	Amendment from company analysis
Fixing errors	
1	Error in the use of UK life tables and conversion of background mortality rate to probability
2	Apply dose intensity after calculating the number of vials per weight category, instead of before
Fixing violations	
3	Added cisplatin plus gemcitabine to the base-case and fully incremental analysis in the PSA
4	Used OS to calculate the responder and non-responder proportions used for response-based TTD - avoiding double counting of patients
5	The ERG removed adverse events with an incidence <5% from the analysis
6	Used the pooled utility estimates from CheckMate275 and CheckMate032
7	Used the pooled weight from CheckMate 275 and 032
8	Removed patient characteristics and comparator treatment costs from the PSA
Matters of judgement	
9	Using conventional survival analysis, not response-based analysis
10	Assumed only doses delayed by 7 days or more to be missed doses, not all delayed doses

66

Generated from information on p142 of the ERG report – see for more in-depth descriptions of the amendments

CONFIDENTIAL

ERG base-case

Amendment	Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Nivolumab ICER (£/QALY)
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 CheckMate studies (6)[b]	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£12,803	0.84	[redacted]	[redacted]	£49,613
	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

Source: Table 6.1 ERG report

(b) Conditional on the fixing errors adjustment (1) and (2)

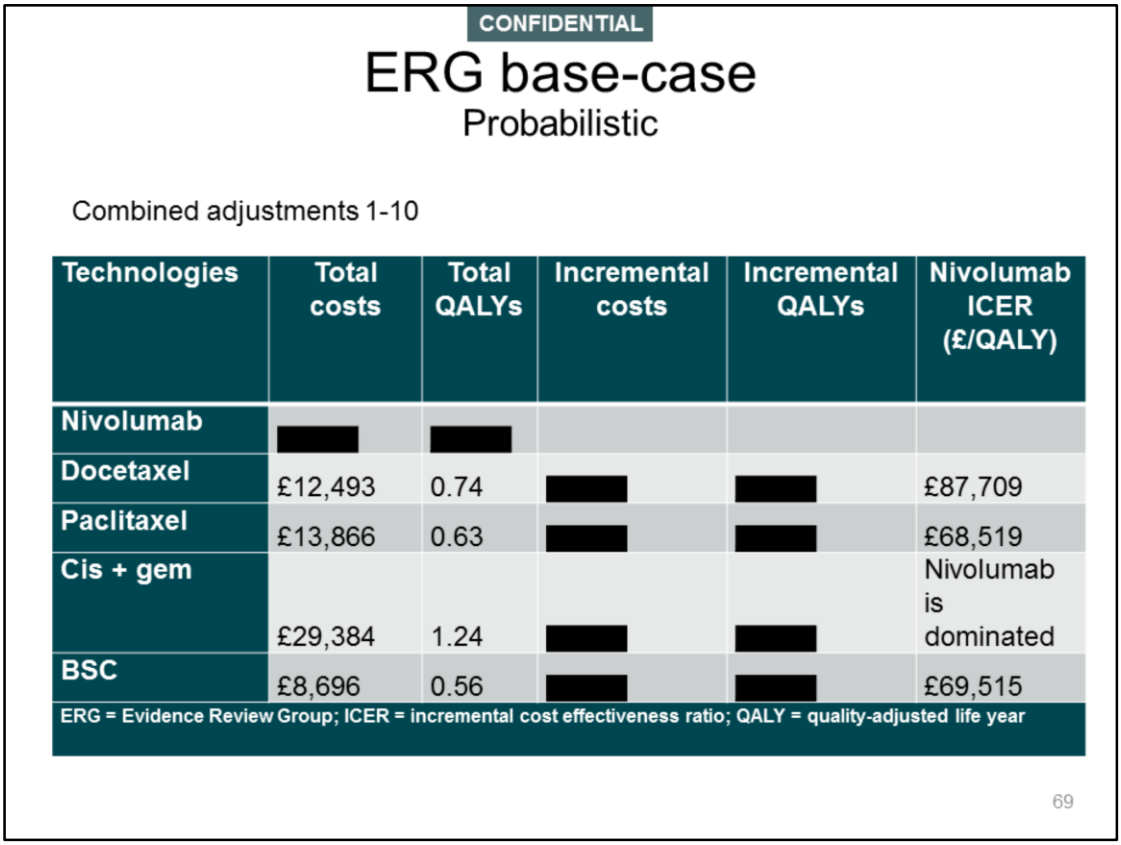
CONFIDENTIAL

ERG base-case

Amendment	Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Nivolumab ICER (£/QALY)
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	£12,803	0.82	[redacted]	[redacted]	£52,858
	Paclitaxel	£14,198	0.71	[redacted]	[redacted]	£44,330
	Cis+gem	£30,315	1.35	[redacted]	[redacted]	£97,665
	BSC	£8,835	0.58	[redacted]	[redacted]	£43,958

Source: Table 6.1 ERG report

(b) Conditional on the fixing errors adjustment (1) and (2)



CONFIDENTIAL

ERG base-case

Probabilistic

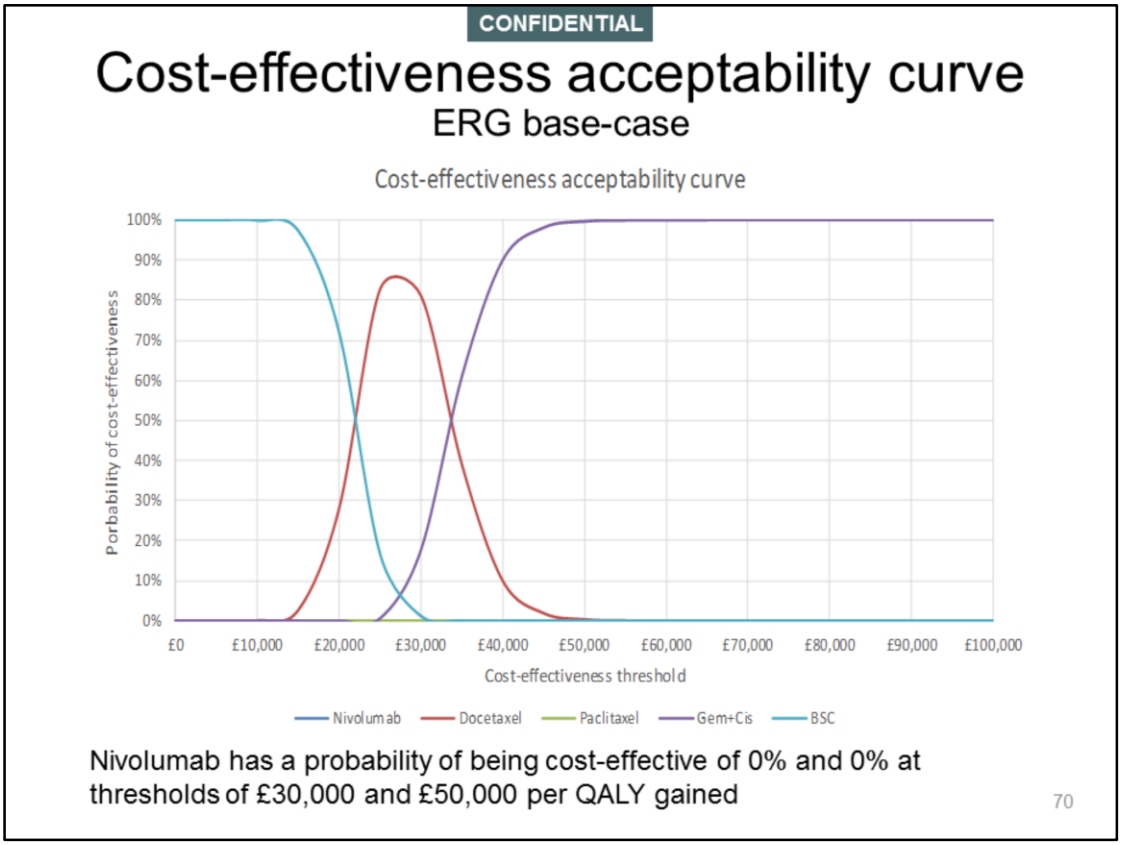
Combined adjustments 1-10

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Nivolumab ICER (£/QALY)
Nivolumab	[redacted]	[redacted]			
Docetaxel	£12,493	0.74	[redacted]	[redacted]	£87,709
Paclitaxel	£13,866	0.63	[redacted]	[redacted]	£68,519
Cis + gem	£29,384	1.24	[redacted]	[redacted]	Nivolumab is dominated
BSC	£8,696	0.56	[redacted]	[redacted]	£69,515

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

69

Source: Table 5.22 ERG report



Source: Figure 5.13 ERG report

ERG exploratory analysis

#	Additional exploratory analysis based on the ERG base-case
A. Exploratory analyses using the ERG base-case	
1	Use of the lognormal distribution for OS and log-logistic for PFS
2	Alternative fractional polynomial model spec, 'mini-PSA' across different P1/P2 values
3	Use of naïve comparison performed by the ERG to derive HRs for OS and PFS
4	Use of time-independent HRs for OS and PFS
5	Use of HRs for OS as proxy for HR for PFS for BSC and cisplatin plus gemcitabine
6	Use of adverse event disutilities and resource use from technology appraisal ID971
7	Use of the UK dosage schedule for cisplatin plus gemcitabine
8	Extreme scenario of assuming no treatment effect of nivolumab vs comparators
B. Exploratory analyses on ERG base-case using response-based model for OS, PFS, TTD	
1	Maintaining the company's base-case choice of parametric time-to-event models
2	Use of parametric time-to-event models with the best fit for OS and PFS (AIC/BIC based)
3	Use of parametric time-to-event models with the best fit
4	Use of 26-week landmark instead of 8-week landmark

71

Generated with information from P144 of the ERG report – more in-depth descriptions can be seen in that section of the report

CONFIDENTIAL

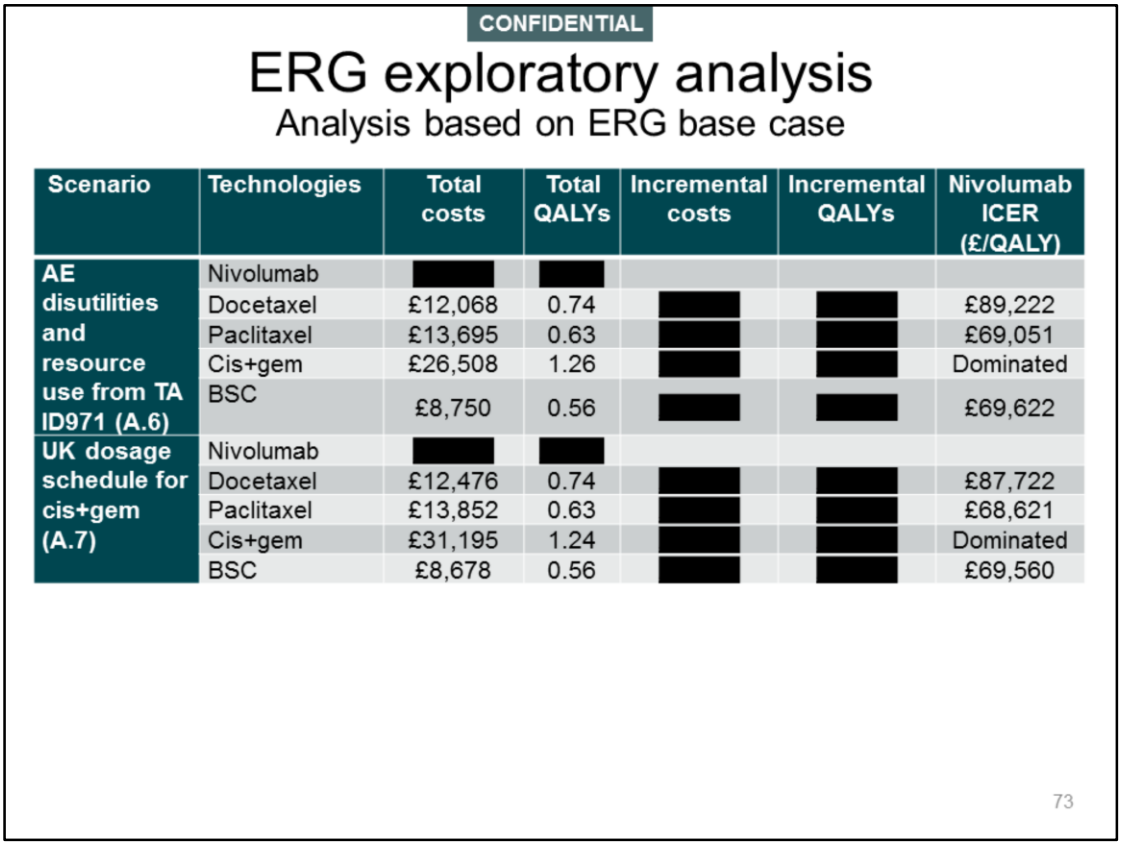
ERG exploratory analysis

Analysis based on ERG base case

Amendment	Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Nivolumab ICER (£/QALY)
Alternative parametric Time-to-event models (A.1)	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£13,173	1.01	[redacted]	[redacted]	£45,721
	Paclitaxel	£14,654	0.89	[redacted]	[redacted]	£39,286
	Cis+gem	£29,736	1.58	[redacted]	[redacted]	£72,732
	BSC	£9,235	0.72	[redacted]	[redacted]	£38,147
Naïve comparison data instead of STC results (A.3)	Nivolumab	[redacted]	[redacted]	[redacted]	[redacted]	
	Docetaxel	£13,005	0.77	[redacted]	[redacted]	£92,335
	Paclitaxel	£13,914	0.60	[redacted]	[redacted]	£64,914
	Cis+gem	£30,910	1.56	[redacted]	[redacted]	Dominated
	BSC	£8,630	0.52	[redacted]	[redacted]	£65,593
Time-independent HRs (A.4)	Nivolumab	[redacted]	[redacted]	[redacted]	[redacted]	
	Docetaxel	£10,213	0.60	[redacted]	[redacted]	£71,639
	Paclitaxel	£13,081	0.78	[redacted]	[redacted]	£95,775
	Cis+gem	£26,584	0.86	[redacted]	[redacted]	£76,576
	BSC	£8,173	0.40	[redacted]	[redacted]	£55,577
Alternative PFS HRs for BSC and cis+gem (A.5)	Nivolumab	[redacted]	[redacted]	[redacted]	[redacted]	
	Docetaxel	£12,507	0.74	[redacted]	[redacted]	£87,863
	Paclitaxel	£13,858	0.63	[redacted]	[redacted]	£68,679
	Cis+gem	£34.999	1.26	[redacted]	[redacted]	Dominated
	BSC	£8,698	0.55	[redacted]	[redacted]	£68,369

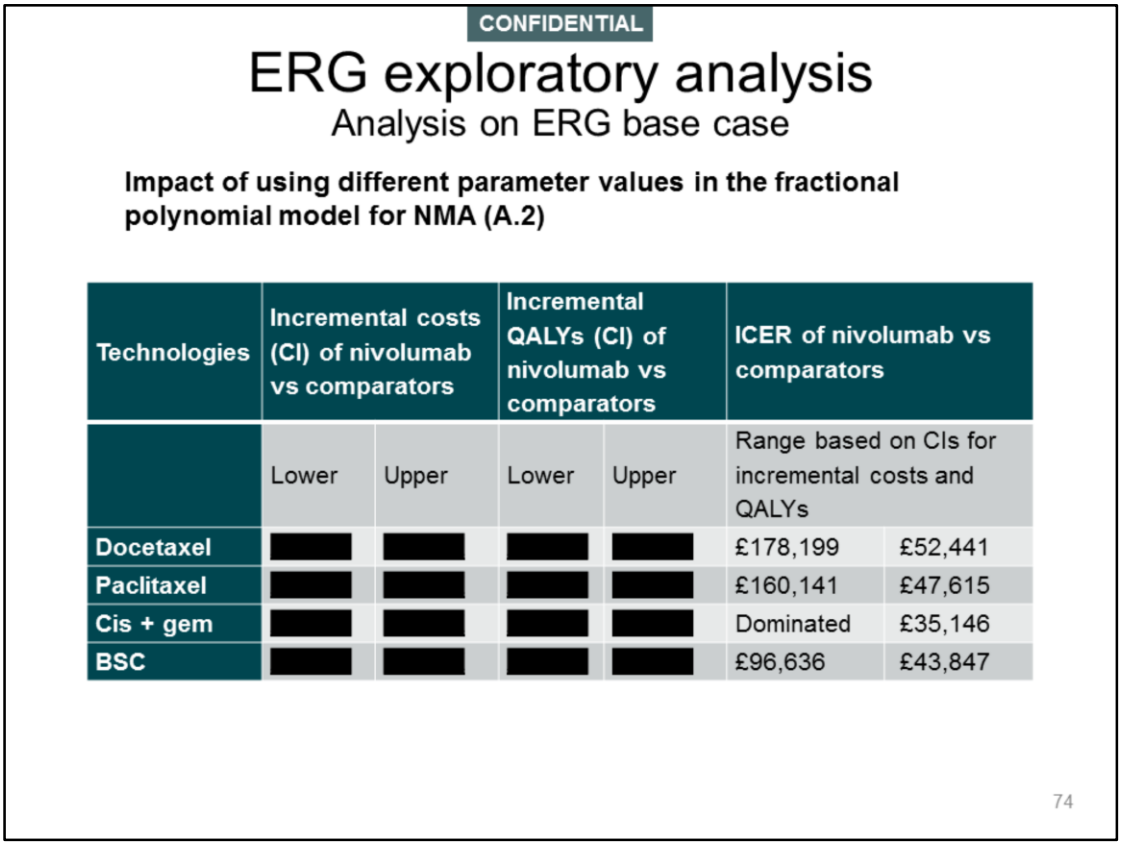
72

Source: Table 6.2 ERG report



Scenario	Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Nivolumab ICER (£/QALY)
AE disutilities and resource use from TA ID971 (A.6)	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£12,068	0.74	[redacted]	[redacted]	£89,222
	Paclitaxel	£13,695	0.63	[redacted]	[redacted]	£69,051
	Cis+gem	£26,508	1.26	[redacted]	[redacted]	Dominated
	BSC	£8,750	0.56	[redacted]	[redacted]	£69,622
UK dosage schedule for cis+gem (A.7)	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£12,476	0.74	[redacted]	[redacted]	£87,722
	Paclitaxel	£13,852	0.63	[redacted]	[redacted]	£68,621
	Cis+gem	£31,195	1.24	[redacted]	[redacted]	Dominated
	BSC	£8,678	0.56	[redacted]	[redacted]	£69,560

Source: Table 6.2 ERG report
Extreme ‘no treatment effect’ scenario not presented here



CONFIDENTIAL

ERG exploratory analysis

Analysis on ERG base case

Impact of using different parameter values in the fractional polynomial model for NMA (A.2)

Technologies	Incremental costs (CI) of nivolumab vs comparators		Incremental QALYs (CI) of nivolumab vs comparators		ICER of nivolumab vs comparators	
	Lower	Upper	Lower	Upper	Range based on CIs for incremental costs and QALYs	
Docetaxel	[redacted]	[redacted]	[redacted]	[redacted]	£178,199	£52,441
Paclitaxel	[redacted]	[redacted]	[redacted]	[redacted]	£160,141	£47,615
Cis + gem	[redacted]	[redacted]	[redacted]	[redacted]	Dominated	£35,146
BSC	[redacted]	[redacted]	[redacted]	[redacted]	£96,636	£43,847

74

Source: Table 6.3 ERG report

CONFIDENTIAL

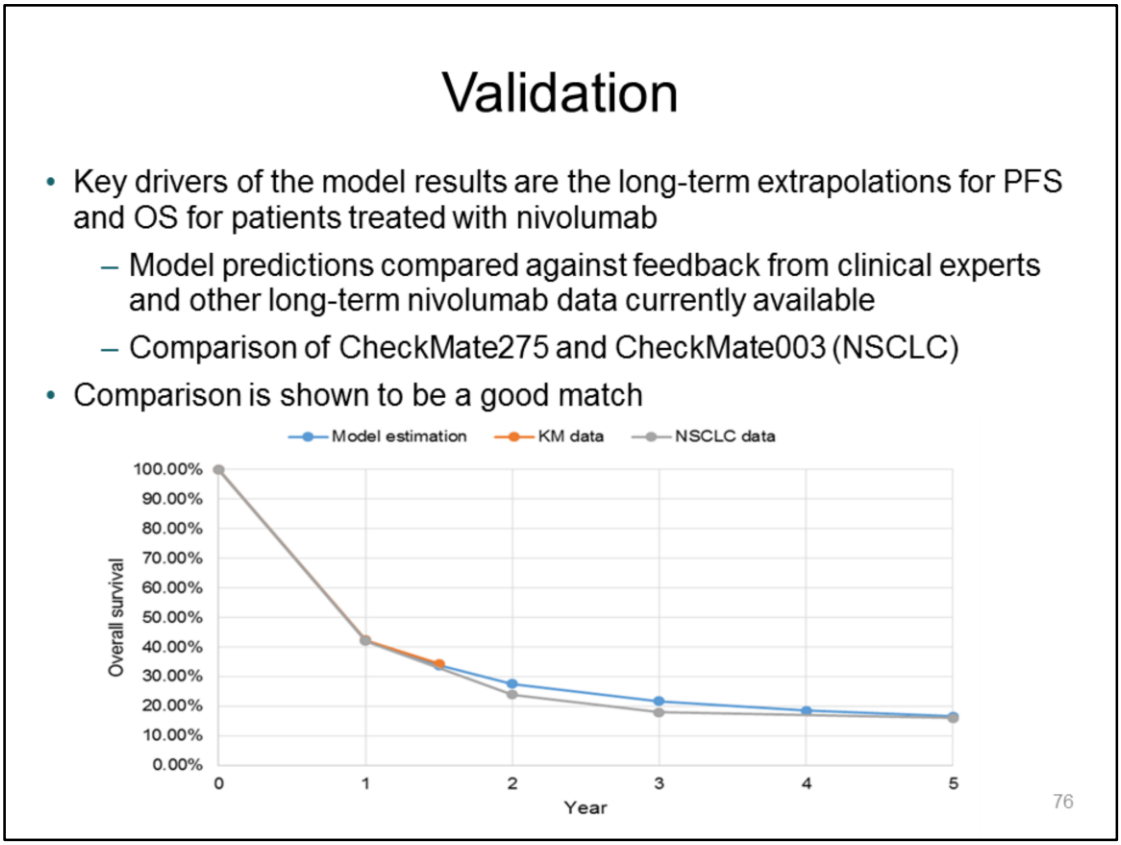
ERG exploratory analysis

Analysis on ERG base case using response-based model

Scenario	Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Nivolumab ICER (£/QALY)
Response-based analysis using ERG base-case (B.1)	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£12,783	0.84	[redacted]	[redacted]	£53,273
	Paclitaxel	£14,163	0.73	[redacted]	[redacted]	£44,877
	Cis+gem	£30,310	1.39	[redacted]	[redacted]	£103,186
	BSC	£8,811	0.59	[redacted]	[redacted]	£44,183
Response-based analysis alternative OS and PFS (B.2)	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£12,475	0.77	[redacted]	[redacted]	£78,795
	Paclitaxel	£13,983	0.68	[redacted]	[redacted]	£68,594
	Cis+gem	£29,893	1.25	[redacted]	[redacted]	£146,721
	BSC	£8,678	0.55	[redacted]	[redacted]	£65,249
Response-based analysis alternative OS, PFS and TTD (B.3)	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£12,452	0.77	[redacted]	[redacted]	£77,597
	Paclitaxel	£13,948	0.67	[redacted]	[redacted]	£67,608
	Cis+gem	£29,880	1.25	[redacted]	[redacted]	£143,923
	BSC	£8,662	0.55	[redacted]	[redacted]	£64,282
Response-based analysis using 26-week landmark (B.4)	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£10,849	0.51	[redacted]	[redacted]	£75,094
	Paclitaxel	£13,689	0.52	[redacted]	[redacted]	£71,255
	Cis+gem	£28,678	0.79	[redacted]	[redacted]	£87,022
	BSC	£8,035	0.35	[redacted]	[redacted]	£61,647

75

Source: Table 6.2 ERG report



Source: Figure 49 company evidence submission

The ERG's concerns include:

1. the lack of internal and cross validity efforts as well as sparse use of expert opinion,
    - There is no description of face validity checks or cross validity checks
    - Clinical experts did not provide feedback on the distributions used for estimating OS and PFS in the company's base-case response-based approach, consulted prior to model development
2. external validation efforts that are based on a lung cancer study,
    - Questionable whether lung cancer really is similar enough to bladder cancer to enable data from the CheckMate 003 trial to be used for external validation of model predictions in bladder cancer
3. the use of only CheckMate 275 for validating model predictions,
    - not using the pooled estimates from CheckMate 275 and 032, impairs the credibility of this validation effort
4. transparency issues with the model.
    - unnecessary difficulties in validating and amending the model

CONFIDENTIAL

Conclusions

- Compared nivolumab with paclitaxel, the current UK standard of care, nivolumab is associated with ICERs (deterministic) that are below the cost-effectiveness threshold for an end of life medicine at £37,647
- 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 ERG base-case resulted in ICERs (probabilistic) of £87,709, £68,519 and £69,515 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

77

Limitations

- Lack of RCT data
  - Uncertainty about the relative treatment effectiveness estimates, which were entirely derived from single-arm studies
- Immaturity of the OS data
  - Due to this the choice of parametric distribution to predict long-term outcomes has a large impact on the final ICERs
- Largely reliant on data from CheckMate275
  - No UK patients
- Uncertainties with a response-based survival analysis approach
  - Unnecessary assumptions
- The exclusion of cisplatin plus gemcitabine as a base-line comparator
- Residual bias could not be quantified in the company's analysis
- Model validation comparison with NSCLC data may not be valid
- The uncertainty introduced from using time-varying HRs was unfortunately not assessed within the PSA

78

End of life

Criterion	Data available
Short life expectancy, less than 24 months	• No studies identified in the systematic literature review provided evidence of OS estimates for this patient population that approached 24 months • Highest median modelled OS of any of the comparators was 10.5 months (Gemcitabine+Cisplatin) (95% CI 3 to 22.9)
Treatment offers an extension to life, normally of at least an additional 3 months, compared with current NHS treatment	• The economic analysis predicted mean life years per patient with nivolumab of 2.78 years (33.36 months) • In comparison, predicted mean life years per patient with comparator therapies were 1.19 years (14.28 months) with paclitaxel, 1.40 years (16.80 months) with docetaxel and 1.01 years (12.12 months) with BSC. Nivolumab was therefore predicted to offer an extension to life of considerably greater than 3 months versus each of these comparators. Furthermore, in the context of the average survival of patients receiving paclitaxel, docetaxel or BSC, the survival gains offered by nivolumab represent a significant extension to life.

ERG comment:
This argument is partly based on lack of evidence to argue that there is no evidence of life expectancy over 24 months, and partly on very weak evidence from the economic model based on a comparison of single arm studies to show an extension to life of at least 3 months

79

Key points for consideration

- Quality of evidence
  - no comparative nivolumab trial data
  - are the nivoumab studies generalisable to UK practice?
  - how reliable is the simulated treatment comparison? Does the company account for all of the important prognostic factors?
  - how reliable is the network meta-analysis? Are the included studies sufficiently homogeneous?
- How effective is nivolumab?
- Is there enough evidence to make recommendations for PD-L1 subgroups?
- The company excluded gemcitabine and cisplatin from its base case. Is this appropriate?
- Company used a response-based analyses to model survival. ERG preferred conventional approach. Which approach is most appropriate?
- What is the most plausible ICER?
- Can nivolumab be considered innovative? Does end of life apply?

80

Authors

- **Thomas Paling**
  Technical Lead
- **Christian Griffiths**
  Technical Adviser
- with input from the Lead Team:
  - Matt Bradley (cost lead)
  - William Turner (clinical lead)
  - Malcolm Oswald (lay lead)

81

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Nivolumab for treating metastatic or unresectable urothelial cancer after platinum-based chemotherapy [ID995]

Document A
Company evidence submission summary for committee

Bristol-Myers Squibb Pharmaceuticals Limited confirm that all information in the submission summary is an accurate summary or replication of evidence in the main submission and accompanying appendices and that wherever possible a cross reference to the original source is provided.

June 2017

File name	Version	Contains confidential information	Date
ID995_Nivolumab for urothelial cancer_Document A_FINAL ACIC	**1.0**	**Yes**	**26/06/2017**

Instructions for companies

This is the template you should use to summarise your evidence submission to the National Institute for Health and Care Excellence (NICE) as part of the single technology appraisal (STA) process. This document will provide the appraisal committee with an overview of the important aspects of your submission for decision-making.

This submission summary must not be longer than 25 pages, excluding the pages covered by this template. If it is too long it will not be accepted. Please submit a draft summary with your main evidence submission. The NICE technical team may request changes later.

When cross referring to evidence in the main submission or appendices, please use the following format: Document, heading, subheading (page X).

For all figures and tables in this summary that have been replicated, cross refer to the evidence from the main submission or appendices in the caption in the following format: Table/figure name – document, heading, subheading (page X).

Companies making evidence submissions to NICE should also refer to the NICE guide to the methods of technology appraisal and the NICE guide to the processes of technology appraisal.

Highlighting in the template (excluding the contents list)

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

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

Grey highlighted text in the footer does not work as an automatic form field, but serves the same purpose – as prompt text to show where you need to fill in relevant details. Replace the text highlighted in [grey] in the header and footer with appropriate text. (To change the header and footer, double click over the header or footer text. Double click back in the main body text when you have finished.)

Contents

Tables and figures

Submission summary

This submission addresses the clinical and cost-effectiveness of nivolumab within its full marketing authorisation for the treatment of locally advanced unresectable or metastatic urothelial carcinoma (UC) in adults after the failure of prior platinum-containing therapy.

A.1 Health condition

UC (or transitional cell carcinoma) is a cancer that originates in the urothelium, the lining of the urinary tract, and accounts for approximately 90% of all bladder cancers.[1]

Depending on how far the tumour has grown and invaded the muscle layers of the bladder wall, UC can be described as either non-muscle-invasive or muscle-invasive. Locally advanced and metastatic disease refers to tumours that have grown through the bladder wall and/or have spread to lymph nodes or other distant sites of the body.[2]

The most common symptoms of UC are haematuria (blood in the urine) and a variety of other irritative and obstructive urinary symptoms such as dysuria, frequency, urgency, feeling of incomplete voiding, and straining.[3] UC has a considerable impact on urinary, bowel and sexual functions and therefore impacts on daily life and sleeping patterns. These symptoms and disruption to normal bodily function can cause considerable impairment to patient health-related quality of life (HRQoL).

Progression of UC to an advanced or metastatic stage is associated with further worsening of HRQoL, with patients in late stages of the disease potentially suffering significant limitations to their mobility. In addition, symptoms related to metastases may include abdominal or pelvic pain, anorexia, renal failure or respiratory symptoms.[3]

The prognosis of patients with locally advanced unresectable or metastatic UC whose disease has progressed following prior platinum-containing chemotherapy is between 5 to 9 months with standard chemotherapy options and thus patients are considered to be at an end-of-life disease stage.[4-7] Furthermore, only 10% of patients typically respond to second-line single-agent chemotherapy regimens hence there is a significant unmet need for effective and tolerable treatment options in this patient population.[8, 9] Further details of the health condition are presented in section B.1.3.1 of the main submission document.

A.2 Clinical pathway of care

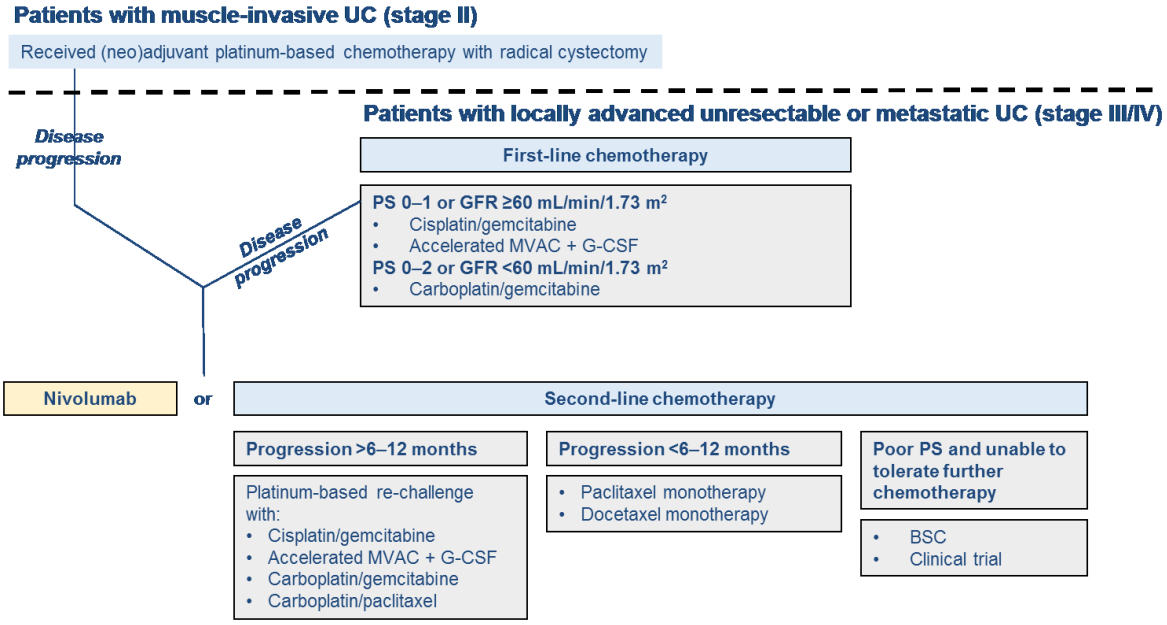
The patient population considered within this submission are patients with locally advanced unresectable or metastatic UC after failure of prior platinum-containing therapy. The clinical pathway of care preceding this and the treatment options for patients who reach this point in the clinical pathway are detailed below.

Patients with newly diagnosed muscle-invasive UC in the UK are typically treated with (neo)adjuvant chemotherapy using a cisplatin-containing combination regimen with either a radical cystectomy or radiotherapy.[10] Patients who are diagnosed with locally advanced or metastatic disease, or have progressed after (neo)adjuvant therapy, will receive cisplatin plus gemcitabine, the standard of care in the first-line setting in the UK. Some patients may receive accelerated (high-dose) MVAC (methotrexate, vinblastine, doxorubicin and cisplatin) in combination with granulocyte-colony stimulating factor [G-CSF]) or a carboplatin-based regimen, if cisplatin is not tolerated.[2, 10-12]

For patients who progress on or after first-line platinum chemotherapy, effective and tolerated treatment options in the second-line setting are severely limited; the vast majority of these patients typically receive paclitaxel monotherapy (or docetaxel monotherapy in some centres), or best supportive care (BSC). A small proportion of patients (less than 10%) who are considered fit enough and have been progression-free for normally at least 9–12 months may receive retreatment with the same first-line platinum-based chemotherapy regimens (see Figure 1).[2, 10-12]

Further details in relation to the clinical pathway of care for UC in the UK are provided in section B.1.3.2 of the main submission document.

Figure 1: Current treatment pathway for patients with urothelial carcinoma in the UK



Abbreviations: BSC: best supportive care; G-CSF: granulocyte-colony stimulating factor; GFR: glomerular filtration rate; MVAC: methotrexate, vinblastine, doxorubicin and cisplatin; PS: performance status; UC: urothelial carcinoma.
Source: Adapted from NICE and EAU/ESMO guidelines and expert clinician feedback.[2, 11, 10]

Based on the above treatment pathway, and as listed in the NICE final scope, the treatment options representing potentially relevant comparators to nivolumab for the treatment of locally advanced unresectable or metastatic UC after failure of prior platinum-containing therapy are as follows:

- Paclitaxel monotherapy
- Docetaxel monotherapy
- BSC
- Retreatment with platinum-based chemotherapy

No data on retreatment with first-line platinum-based chemotherapy were identified in the clinical systematic literature review (SLR). However, as mentioned, the use of retreatment is limited to less than 10% of patients and so this treatment option is not considered to be standard UK clinical practice for the vast majority of patients. Data from a trial involving cisplatin plus gemcitabine after the failure of MVAC were included as a scenario analysis, in the absence of any data on retreatment with first-line platinum-based chemotherapy.

A.3 Equality considerations

The therapies available in current UK clinical practice for patients with unresectable or metastatic UC whose disease has progressed following platinum-based chemotherapy comprise further chemotherapy agents, many of which are associated with high toxicity. UC patients are typically older patients (54% of cases in the UK each year are diagnosed in patients aged 75 and over),[13] who in many cases cannot tolerate retreatment with chemotherapy and may only be suitable or wish for palliative therapy with BSC.

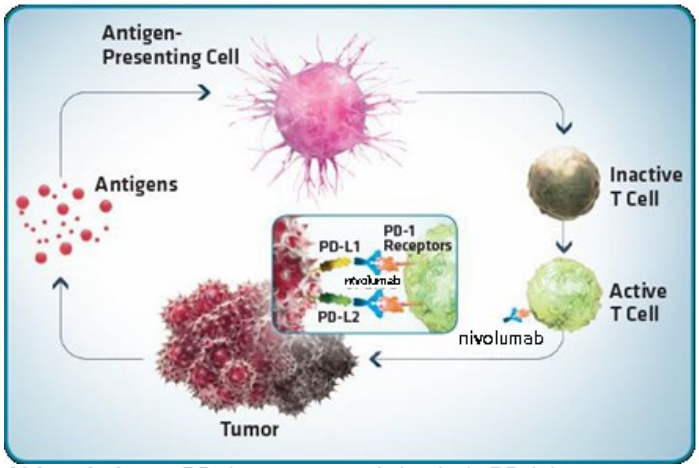
Due to the lack of well-tolerated, effective treatment options after the failure of prior platinum-based chemotherapy, some patients could instead be enrolled into clinical trials. The reliance on clinical trials presents a potential equity issue, given that trial centres may not have an equitable geographic distribution and enrolment criteria and numbers for trials are restricted. The

availability of a nationally funded treatment option on the NHS for patients whose best, or only, option for receiving active treatment is entry into a clinical trial,would help to move towards addressing this equity issue.

A.4 The technology

A summary of the mechanism of action, marketing authorisation status, costs and administration requirements associated with nivolumab is presented in Table 1.

Table 1: Technology being appraised – see Section B.1.2 (page 13)

UK approved name and brand name	Nivolumab (Opdivo®)
Mechanism of action	Nivolumab (Opdivo®) is a human, monoclonal immunoglobulin G4 (IgG4) antibody that acts as a programmed death 1 (PD-1) inhibitor, blocking the interaction of PD-1 with programmed death-ligand 1 and 2 (PD-L1 and PD-L2) (see Figure 2). The programmed death 1 (PD-1) receptor is a negative regulator of T-cell activity and is expressed on activated T-cells. Interaction of PD-1 with its ligands (programmed death-ligand 1, PD-L1, and programmed death-ligand 2, PD-L2) results in the inhibition of T-cell activation and subsequent T-cell death. PD-L1 and PD-L2 are expressed on antigen-presenting cells (such as dendritic cells), and may also be expressed by tumours or other cells in the tumour microenvironment (see Figure 2).[14, 15] **Figure 2: Nivolumab stimulation of immune-mediated destruction**  **Abbreviations:** PD-1 programmed death 1; PD-L1: programmed death ligand 1; PD-L2: programmed death ligand 2. There is increasing evidence that implicates the PD-1 signalling pathway in UC tumour evasion.[16] By preventing inactivation of T-cells, nivolumab effectively restores T-cell activity against tumour cells, i.e. nivolumab harnesses the patient's own immune system to directly fight cancer cells (in the same way that it would any other "foreign" antigen), resulting in destruction of the tumour. Furthermore, the potential of targeting immune inhibitory pathways to treat UC is indicated by the effectiveness in some patients of bacillus calmette-guerin therapy (BCG). This immunotherapy treatment has been used for over 40 years in patients with high-grade non-muscle-invasive UC following surgical resection. Administered intravesically, BCG induces the secretion of cytokines from urothelial cells and the attraction of vast numbers of neutrophils and monocytes to the tumour site, leading to an immune response against tumour cells.[17, 18] There is also evidence in studies of patients with localised UC that the use of ipilimumab, an

	immune checkpoint inhibitor that blocks CTLA-4, enhances immune responses and tumour regression.[19, 20] As such, this evidence provides a compelling biological rationale for the effectiveness of nivolumab and the blocking of PD-1 as a therapeutic target in UC.[21, 22]
Marketing authorisation/CE mark status	The licensed indication for nivolumab as a treatment for UC is detailed below: "Nivolumab (Opdivo®) is indicated for the treatment of locally advanced unresectable or metastatic urothelial carcinoma in adults after failure of prior platinum-containing chemotherapy" An application for a marketing authorisation in this indication was submitted to the European Medicines Agency (EMA) on the 25th August 2016 and a positive opinion from the Committee for Medicinal Products for Human Use (CHMP) was received on the 21st April 2017.[23] Full marketing authorisation was received from the EMA on Monday 5th June 2017. Nivolumab has already been granted a marketing authorisation by the EMA for the following indications, as detailed in the SmPC:[24] • As monotherapy or in combination with ipilimumab, for the treatment of advanced (unresectable or metastatic) melanoma in adults • For the treatment of locally advanced or metastatic non-small cell lung cancer (NSCLC) after prior chemotherapy in adults • As monotherapy for the treatment of advanced renal cell carcinoma after prior therapy in adults • For the treatment of adult patients with relapsed or refractory classical Hodgkin lymphoma after autologous stem cell transplant and treatment with brentuximab vedotin • For the treatment of squamous cell cancer of the head and neck in adults progressing on or after platinum-based therapy
Indications and any restriction(s) as described in the SmPC	There are no restrictions associated with the licensed indication for nivolumab in locally advanced unresectable or metastatic UC.
Method of administration and dosage	The recommended dose and schedule of nivolumab monotherapy in UC is 3 mg/kg administered as an intravenous (IV) infusion over 60 minutes every 2 weeks (Q2W), consistent with the existing approved dose and schedule of nivolumab monotherapy in adults in other indications. Dose escalation or reduction is not recommended; dosing delay or discontinuation may be required based on individual safety and tolerability.
Additional tests or investigations	As detailed in the SmPC, nivolumab treatment must be initiated and supervised by physicians experienced in the treatment of cancer.[24] Hospital oncology units already have the staffing and infrastructure needed for the administration of IV oncology therapies. Administration of nivolumab is therefore not expected to require any additional NHS infrastructure, as the majority of the comparators included in the final scope for this appraisal are also intravenously administered. The only expected source of differential resource use to the NHS for nivolumab relative to current clinical comparators is in the management of immune-related AEs. AEs observed with immunotherapies, such as nivolumab, may differ from those observed with non-immunotherapies that are currently used in clinical practice. The immune-based mechanism of action of nivolumab means that many of its treatment-related AEs are

	immunological in origin. Patients treated with nivolumab are advised to be vigilant and report any changes whilst on treatment to help ensure quick resolution of potential AEs. Immune-related AEs associated with nivolumab, including severe AEs, are well characterised and are generally manageable with topical and/or systemic immunosuppressants.[24] They are often resolved following initiation of appropriate medical therapy, for example corticosteroids, and/or withdrawal of nivolumab.[24] A full list of AEs and guidelines for the discontinuation or withholding of doses in response to immune-related AEs is provided in the SmPC.[24] As detailed in the SmPC for nivolumab, adequate evaluation should be performed to confirm the aetiology or exclude other causes for suspected immune-related AEs.[24]
List price and average cost of a course of treatment	**40 mg vial** / **100 mg vial** **List price:** £439.00 / £1,097.00 **PAS price:** [redacted] / [redacted] Treatment with nivolumab should be continue as long as clinical benefit is observed or until treatment is no longer tolerated by the patient. Based on the economic model developed for this submission, the average cost of treating a patient with nivolumab in this indication is estimated to be: **List price:** £54,675 **PAS price:** [redacted]
Patient access scheme (if applicable)	A PAS is already in place with the Department of Health for inclusion in this technology appraisal, representing a simple discount of [redacted] on the list price of nivolumab.

Abbreviations: Ab: antibody; CD28: cluster of differentiation 28; CHMP: Committee for Medicinal Products for Human Use; EMA: European Medicines Agency; IV: intravenous; MHC: major histocompatibility complex; NHS: National Health Service; NSCLC: non-small cell lung cancer; PAS: patient access scheme; PD-1: programmed death 1; PD-L1: programmed death-ligand 1; PD-L2: programmed death-ligand 2; Q2W: twice weekly; SmPC: summary of product characteristics; UC: urothelial carcinoma.

A.5 Decision problem and NICE reference case

This submission covers the full marketing authorisation for nivolumab in this indication. The company submission is consistent with the final NICE scope and NICE reference case, with minor discrepancies noted in Table 2 on the next page.

Table 2. The decision problem – see Section B.1.1 (page 10)

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
Population	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	N/A
Intervention	Nivolumab	Nivolumab	N/A
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, 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 effects 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 effects of treatment • health-related quality of life (via the EORTC QLQ-C30 and the EQ-5D-3L)	N/A

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. Costs were considered from an NHS and Personal Social Services perspective.	N/A
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.
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.

Abbreviations: N/A: not applicable; PD-L1: programmed death-ligand 1. **Source:** NICE final scope [ID995] – issue date: April 2017.[25]

A.6 Clinical effectiveness evidence

An SLR was conducted to identify relevant clinical evidence on the efficacy and safety of nivolumab for the treatment of unresectable or metastatic UC. Full details of the SLR search strategy, study selection process and results are presented in Appendix D.

A total of 18 publications reporting on 12 trials investigating comparators listed in the NICE final scope were identified, including two clinical trials providing evidence for the efficacy and safety of nivolumab in locally advanced unresectable or metastatic UC after failure of prior platinum-containing chemotherapy: CheckMate 275 and CheckMate 032.[26, 27]

No RCTs directly comparing the efficacy and safety of nivolumab in the patient population of interest versus any of the comparators relevant to this submission or placebo were identified.

A summary of CheckMate 275 and CheckMate 032 is provided in Table 3.

Table 3: Clinical effectiveness evidence

Study title	CheckMate 275 (N=270)	CheckMate 032 (N=78)
Study design	Multicentre, open-label, single-arm phase II trial	Multicentre, open-label, two-stage, single-arm, phase I/II[a] trial
Population	Patients with locally advanced unresectable or metastatic UC who had progressed or recurred after at least one previous line of platinum-containing chemotherapy	Patients with locally advanced unresectable or metastatic UC who had progressed or recurred after treatment with at least one platinum-containing chemotherapy regimen[a]
Intervention(s)	Nivolumab (IV 3 mg/kg Q2W)	Nivolumab (IV 3 mg/kg Q2W)
Comparator(s)	N/A (single-arm)	N/A[a]
Outcomes specified in the decision problem	• ORR • OS • PFS • HRQoL via the European Organisation for Research and Treatment of Cancer (EORTC) general cancer module (QLQ-C30) and the EuroQoL-5 dimensions-3 levels (EQ-5D-3L) questionnaires • AEs	• ORR • OS • PFS • EQ-5D-3L • AEs
Reference to section in submission	See Section B.2.3.1 (page 27), B.2.6 (page 42) and Appendix M (page 227)	See Section B.2.3.2 (page 28), B.2.6.5 (page 50) and Appendix M (page 227)

[a]CheckMate 032 investigated nivolumab or nivolumab combined with ipilimumab in patients with UC, triple-negative breast cancer, gastric cancer, pancreatic adenocarcinoma, small cell lung cancer, and ovarian cancer. Here, presentation of CheckMate 032 refers only to the nivolumab monotherapy UC cohort (n=86) of relevance to this submission.
Abbreviations: EORTC QLQ-C30: European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30; EQ-5D-3L: 3-level EuroQoL 5-Dimensions; HRQoL: health-related quality of life; IV: intravenous; N/A: not applicable; ORR: objective response rate; OS: overall survival; PFS: progression-free survival; Q2W: every two weeks; UC: urothelial carcinoma.

A.7 Key results of the clinical effectiveness evidence

An overview of the clinical effectiveness results from CheckMate 275 and CheckMate 032 is presented in Table 4.

Table 4: Clinical effectiveness results from CheckMate 275 and CheckMate 032

Outcome	CheckMate 275		CheckMate 032
	Initial database lock: 30th May 2016 n=265[c]	Latest database lock: 2nd Sep 2016 n=270[c]	n=78
ORR, n (%), [95% CI]	52 (19.6), [15.0–24.9]	54 (20.0), [15.4–25.3][c]	19 (24.4), [15.3–35.4]
TTR, median (IQR), months	1.87 (1.81–1.97)[b]	1.94 (1.84–2.50)[c]	1.48 (1.25–4.14)
DOR, median (95% CI), months	NR (7.43–NR)[b]	10.35 (7.52–NR)[c]	NR (9.92–NR)
PFS, median (95% CI), months	2.00 (1.87–2.63)[b]	2.00 (1.87–2.63)[c]	2.78 (1.45–5.85)
OS, median (95% CI), months	8.74 (6.05–NR)[b]	8.57 (6.05–11.27)[c]	9.72 (7.26–16.16)

[a]Follow-up for the latest database lock was sufficient to include 5 patients from Japan who were not included in efficacy analyses in the initial database lock.[b]Minimum follow-up of 6 months from the date of first dose. [c]Minimum follow-up of 8.3 months.
Abbreviations: CI confidence interval; DOR: duration of response; IQR: interquartile range; PFS: progression-free survival; ORR: objective response rate; OS: overall survival; TTR: time to response; NR: not reached.
Source: Sharma *et al.* (2016),[26] Sharma *et al.* (2017)[27]; CheckMate 275 CSR[28] and CheckMate 275 CSR Addendum (25 October 2016).[29]

A.7.1 Objective response rate

At the primary database lock of CheckMate 275 (30th May 2016 [n=265]), treatment with nivolumab led to a clinically meaningful confirmed objective response per blinded independent review committee (BIRC) (primary efficacy endpoint) in a total of 52 (19.6%) patients (95% CI: 15.0–24.9) with 6 (2.3%) patients achieving a CR. At a minimum follow-up of 6 months (primary database lock), median duration of response (DOR) had not yet been reached; [redacted] of responders were continuing in response and nearly all patients ([redacted]) had experienced a DOR of at least 3 months. At the second database lock (2nd September 2016 [n=270]), after a minimum follow-up of 8.3 months, median DOR was 10.35 months.

Results for the primary efficacy endpoint of CheckMate 032 were consistent with those from CheckMate 275: treatment with nivolumab led to a confirmed investigator-assessed objective response in 19 (24.4%) patients (95% CI 15.3–35.4) (n=78). At a minimum follow-up of 9 months, median DOR had not yet been reached; [redacted]%) of responders had experienced a DOR of at least 3 months, and [redacted]%) of responders had a DOR of at least 6 months.

ORR results were consistent across all PD-L1 subgroups in both trials, with clinically meaningful ORRs observed even for patients with low to no PD-L1 expression (PD-L1<1%). In CheckMate 275, patients in the PD-L1≥1% cohort achieved an ORR of 23.8% (95% CI: 16.5–32.3) and patients with <1% PD-L1 expression had a confirmed ORR of 16.1% (15.8% at the second database lock). Further details are provided in Section B.2.6.2 of the full submission.

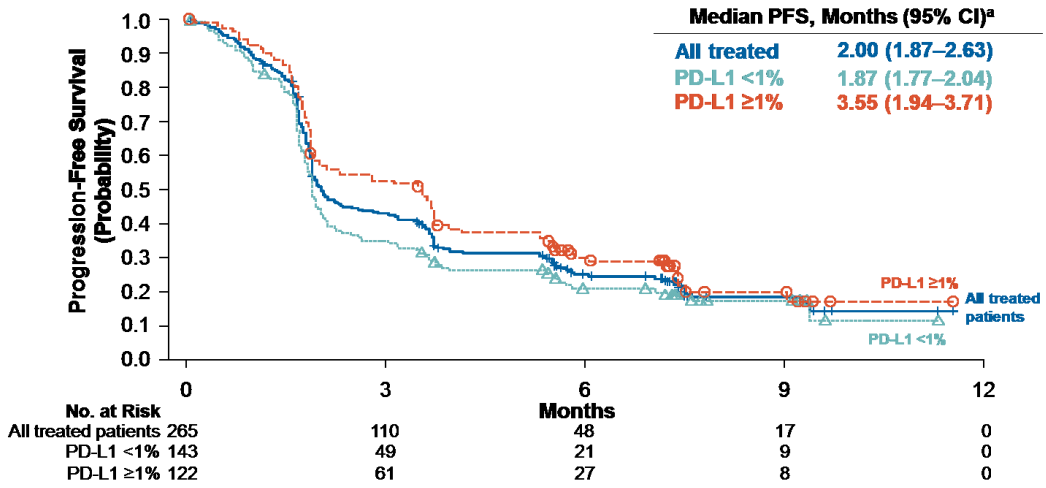
A.7.2 Progression-free survival

At the primary database lock of CheckMate 275, median PFS (by BIRC) in the efficacy-treated population was 2.00 months (95% CI: 1.87–2.63), with a quarter (25.2% [95% CI: 20.0–30.8]) of patients remaining progression-free 6 months after initiation of therapy (Figure 3). At the second database lock, the proportion of patients remaining progression-free was 26.1% (95% CI: 20.9–31.5) at 6 months, and 16.1% (95% CI: 11.7–21.1) at 12 months.

Median PFS for patients in the PD-L1 ≥1% cohort was slightly longer than in the efficacy-treated population at 3.55 months (95% CI: 1.94–3.71), and in the PD-L1 <1%, median PFS was 1.87 months (95% CI: 1.77–2.04) (Figure 3).

In CheckMate 032, median PFS (investigator-assessed) was consistent with that in CheckMate 275 at 2.78 months (95% CI: 1.45–5.85). Further details are presented in Sections B.2.6.3, B.2.6.4, B.2.6.5 and Appendix E of the main submission document.

Figure 3: Kaplan-Meier plot for progression-free survival in CheckMate 275



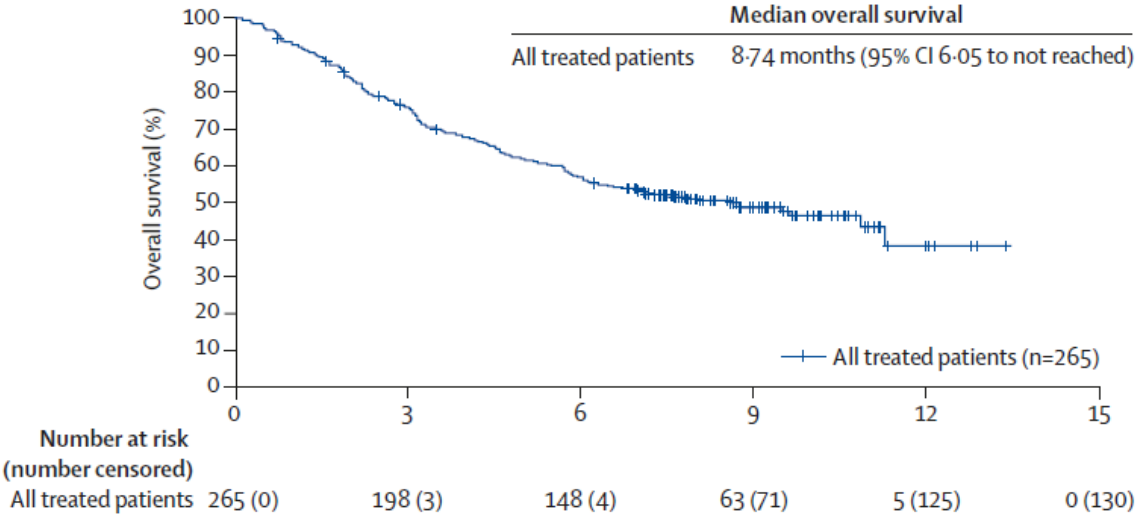
Abbreviations: CI: confidence interval; PD-L1: programmed death ligand 1; PFS: progression-free survival.
Source: Galsky *et al.* (2016).[30]

A.7.3 Overall survival

At the primary database lock of CheckMate 275, 138 patients (51.1%) had died with median OS estimated at 8.74 months (95% CI: 6.05–NR) for the efficacy-treated population. The 3-month and 6-month OS rates were 75.8% (95% CI: 70.2–80.5) and 57.0% (95% CI: 50.7–62.7). At the secondary database lock, median OS was 8.57 months (95% CI: 6.05 –11.27), and the proportion of patients still alive at 12 months was 41.0% (95% CI: 34.8–47.1).

In CheckMate 032, OS results were consistent with CheckMate 275 with a median OS of 9.72 months (95% CI: 7.26–16.16) at a median follow-up for OS of 9.69 months. Further details are presented in Sections B.2.6.3, B.2.6.4, B.2.6.5 and Appendix E. At 12 months, OS was 45.6% (95% CI: 34.2, 56.3).

Figure 4: Kaplan-Meier plot for overall survival in CheckMate 275



Abbreviations: CI: confidence interval; OS: overall survival; PD-L1: programmed death ligand 1.
Source: Sharma *et al.* (2017).[27]

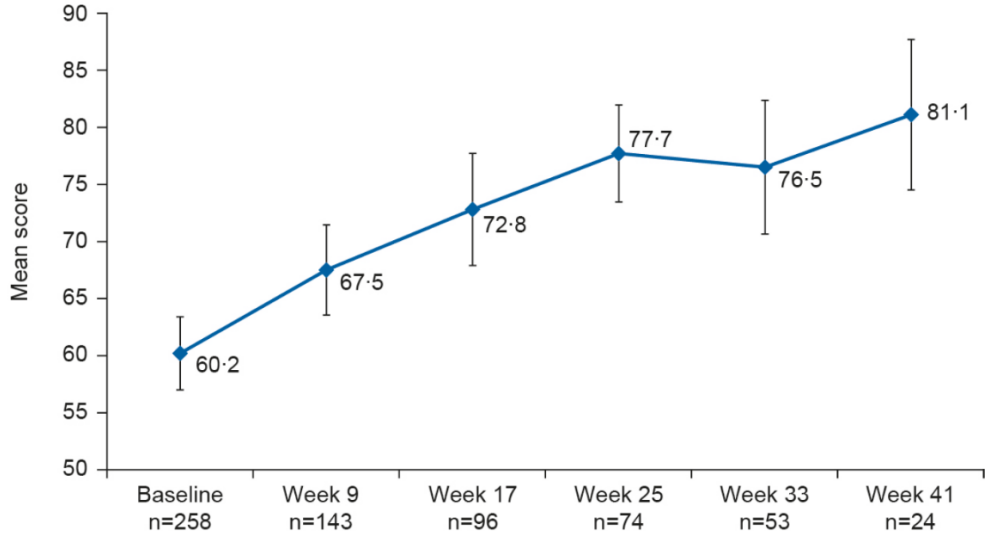
A.7.4 Patient-reported outcomes

Patient-reported outcomes for the measurement of HRQoL were assessed via the EORTC QLQ-C30 questionnaire in CheckMate 275, and the EQ-5D-3L questionnaire in both CheckMate 275 and CheckMate 032.

In CheckMate 275, HRQoL measured via the EORTC QLQ-30 questionnaire demonstrated that nivolumab increased or maintained patient HRQoL from baseline to Week 41, and a meaningful improvement was observed for the dyspnoea, insomnia and financial difficulties domains.

During post baseline follow-up, the percentage of patients reporting health problems decreased by 10% for all dimensions of the EQ-5D: mobility at Week 9, self-care at Week 33, usual activities at Week 17, pain/discomfort at Week 9, and anxiety/depression at Week 17. The EQ-5D visual analogue scale was completed by 259 patients (96%) at baseline, and scores showed clinically relevant improvements in HRQoL by week 9, with continued improvement to the end of week 41, demonstrating the positive impact of nivolumab on patient HRQoL (Figure 5).

Figure 5. Mean EQ-5D-3L VAS score in all-treated population in CheckMate 275



Abbreviations: EQ-5D: EuroQoL-5 dimensions 3-levels questionnaire.
Source: Sharma *et al*. (2017).[27]

The mean EQ-5D VAS score in CheckMate 032 also increased over time. By Week 19, clinically meaningful improvements (>7-point change from baseline) were reported, and the average EQ-5D VAS score was >80 points. Measured via the EQ-5D-3L in the same population, an improvement of ≥10% from baseline in pain/discomfort and anxiety/depression was seen at Week 5; and for mobility and usual activities at Week 19. The proportion of patients with health problems continued to decrease over time for these 4 dimensions. Detailed results of patients-reported outcomes are presented in Section B.2.6.6 of the main submission document.

A.7.5 Adverse reactions

The safety and tolerability of nivolumab for patients with locally advanced unresectable or metastatic UC was evaluated as an exploratory endpoint in CheckMate 275 and as a secondary endpoint in CheckMate 032. The safety profile of nivolumab across both trials was consistent and no new safety signals were raised.

Median duration of therapy was [redacted] months (95% CI: [redacted]) and [redacted] months (95% CI: [redacted]) in CheckMate 275 and CheckMate 032, respectively. As of their respective primary database locks, the proportion of deaths due to study drug toxicity was extremely low (1.1% and 2.6%, respectively). All-cause AEs leading to treatment discontinuation were reported in 20.7% and 7.7% of patients in CheckMate 275 and CheckMate 032, respectively (Table 5).

The vast majority of drug-related AEs were grade 1 or 2 and the frequency of drug-related grade 3 or 4 AEs was low (see Table 5); the most commonly-reported AEs of any grade across both trials were fatigue, nausea and decreased appetite.

Predicted select immune-related AEs did occur, but were mostly grade 1 or 2 and were manageable using the recommended treatment guidelines. The safety data demonstrate that nivolumab in the treatment of locally advanced unresectable and metastatic UC is well tolerated, and the safety profile is manageable and consistent with expectations based on prior data in multiple other tumour types. Full details of the safety analysis are presented in Section B.2.10 of the full submission.

Table 5: Summary of safety analysis in CheckMate 275 and CheckMate 032

Adverse event, n (%)	CheckMate 275 (n=270)[a]		CheckMate 032 (n=78)[b]	
Deaths	138 (51.1)		36 (46.2)	
Deaths due to study drug toxicity	3 (1.1)[c]		2 (2.6)[d]	
	Any grade	**Grade 3-4**	**Any grade**	**Grade 3-4**
All causality AEs	267 (98.9)	137 (50.7)	78 (100)	43 (55.1)
Drug-related AEs	174 (64.4)	48 (17.8)	65 (83.3)	18 (23.1)
All-causality serious AEs	147 (54.4)	99 (36.7)	36 (46.2)	23 (29.5)
Drug-related serious AEs	[redacted]	[redacted]	8 (10.3)	[redacted]
All-causality AEs leading to treatment discontinuation	56 (20.7)	42 (15.6)	6 (7.7)	4 (5.1)
Drug-related AEs leading to treatment discontinuation	13 (4.8)	8 (3.0)	2 (2.6)	2 (2.6)

[a]AEs were coded using the MedDRA version 19.0 and were graded for severity according to the NCI CTCAE version 4.0. [b]AEs were coded using the MedDRA version 18.1 and were graded for severity according to the NCI CTCAE version 4.0. [c]Three deaths (Grade 5 pneumonitis, Grade 5 acute respiratory failure, and Grade 5 cardiovascular failure) were judged as study drug-related. [d]Two deaths (Grade 4 pneumonitis and Grade 4 thrombocytopenia) were assessed as study drug-related.
Abbreviations: AEs: adverse events; MedDRA: Medical Dictionary for Regulatory Activities; NCI CTCAE: National Cancer Institute Common Terminology Criteria for Adverse Events; SAEs: serious adverse events.
Source: Sharma *et al* (2017),[27] CheckMate 275 CSR,[28] Galsky *et al.* (2016),[30] Sharma *et al* (2016)[26] CheckMate 032 CSR.[31]

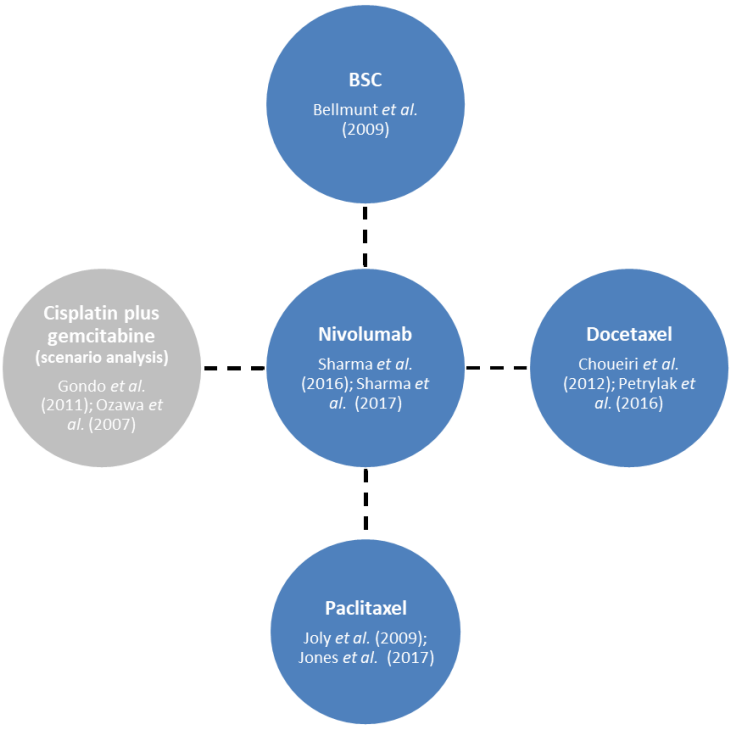
A.8 Evidence synthesis

An SLR identified no RCTs directly comparing the efficacy and safety of nivolumab in the patient population of interest versus any of the comparators relevant to this submission or placebo. As such, the feasibility of conducting an indirect treatment comparison (ITC) was assessed between the two nivolumab trials and ten comparator trials ultimately included in the evidence network.

Eligible trials were identified for paclitaxel, docetaxel and BSC but none which investigated retreatment with first-line platinum-based chemotherapy. Two trials were identified for cisplatin plus gemcitabine but were limited in their generalisability to the decision problem as one trial (Gondo *et al.* (2011)[32]) used a different first-line treatment (MVAC) and so could not be classified as retreatment with first-line platinum-based chemotherapy, and the Ozawa *et al.* (2007)[33] trial included chemotherapy-naïve patients in addition to patients who had previously undergone (unspecified) first-line treatment. The two identified trials also employed dosing regimens different to that used in UK clinical practice. As the trials were conducted in gemcitabine-naïve populations, and therefore cannot be considered to provide relevant data for retreatment with first-line platinum-based chemotherapy, a comparison to cisplatin plus gemcitabine is included as a scenario analysis only and the results versus this comparator should be treated with caution.

An overall network diagram is illustrated in Figure 6. No direct or indirect links were identified between the nivolumab and comparator trials, hence a population-adjusted approach (simulated treatment comparison [STC]) was conducted using individual patient-level data from the nivolumab trials and summary data from the comparator trials where available, to estimate how patients in each of the comparator trials would have responded to nivolumab in terms of OS, PFS and ORR. Further details of this method are provided in Section B.2.9 (page 59) and Appendix D (page 64) and the analysis was conducted in accordance with the new technical support document (TSD) from the NICE Decision Support Unit (DSU) (TSD18).[34]

Figure 6: Network diagram – B.2.9 (page 59)

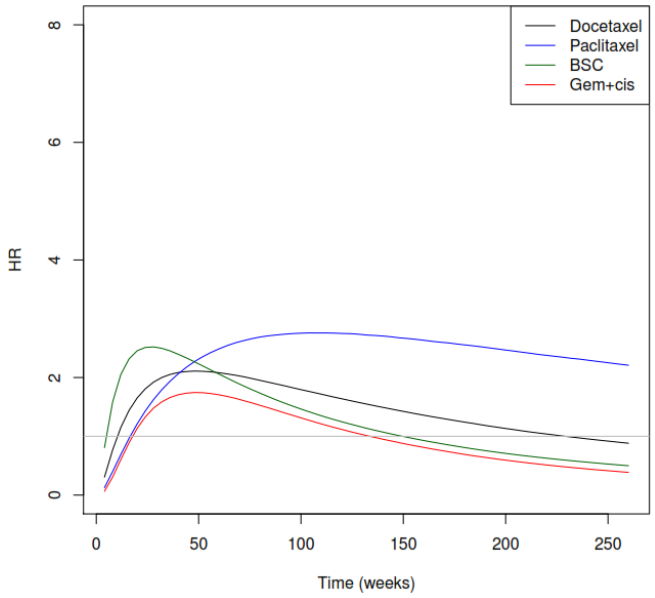


Dashed lines indicate where simulated treatment comparison has been applied.
Abbreviations: BSC: best supportive care.

OS and PFS were evaluated using a fractional polynomial NMA approach, which estimates hazard ratios (HRs) over time.[35] ORR was evaluated using a NMA model for binomial outcomes.[36] For all outcomes, both fixed effect and random effects models were applied. For the survival outcomes, different types of fractional polynomial model were also explored. The deviance information criterion (DIC) was used to evaluate model fit and guide the best choice of model. For the survival outcomes, clinical plausibility of the extrapolated HRs was also considered based on expert clinical feedback elicited via an advisory board and further clinician interviews.[12, 37]

For OS, the second order (P1=0, P2=0) fixed effect model was used in the base case analysis because it provided the most clinically plausible extrapolations out of the three models with the lowest DIC. Figure 7 illustrates the HRs for each of the comparators versus nivolumab over time. HRs greater than 1 favour nivolumab. Further details are provided in Section B.2.9.2 (page 61) of the main submission document.

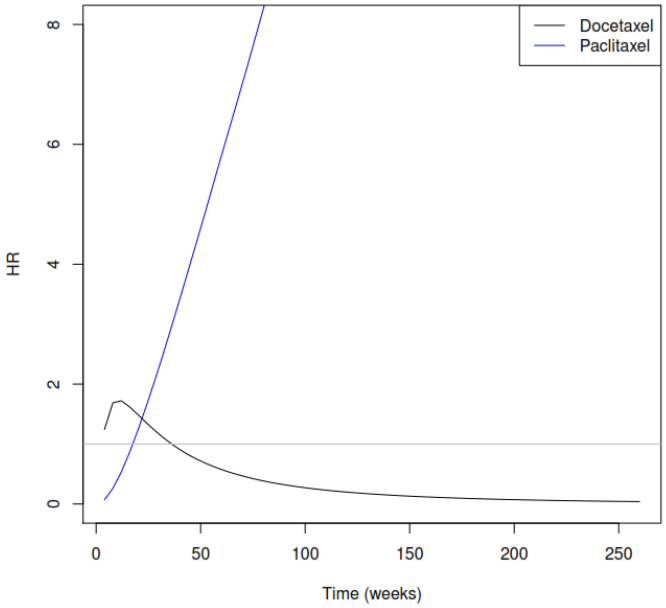
Figure 7: Overall survival: network meta-analysis results (second order (P1=0, P2=0) fixed effect model): HRs for each of the comparators versus nivolumab – B.2.9.2 (page 63)



HRs greater than 1 favour nivolumab.
Abbreviations: BSC: best supportive care; cis: cisplatin; gem: gemcitabine; HR: hazard ratio.

For PFS, the second order (P1=0, P2=0) fixed effect model was taken forward for the base case analysis in the cost-effectiveness model because it had clinical plausibility and the lowest DIC. Figure 8 illustrates the HRs for each of the comparators versus nivolumab over time; HRs greater than 1 favour nivolumab. No PFS data were available for cisplatin plus gemcitabine or BSC. Further details are provided in Section B.2.9.3 (page 65) of the main submission document.

Figure 8: Progression-free survival: network meta-analysis results (fixed effect second order (P1=0, P2=0) model): HRs for each of the comparators versus nivolumab – B.2.9.3 (page 66)

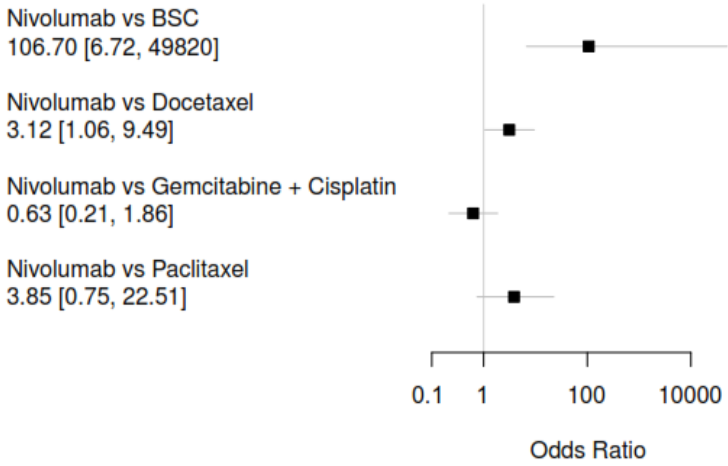


HRs greater than 1 favour nivolumab.
Abbreviations: HR: hazard ratio.

For ORR, the fixed effect model was taken forward for the base case analysis because it had had the lowest DIC. Figure 9 illustrates the ORR odds ratios for nivolumab versus each of the

comparators. The results suggest that patients who receive nivolumab have higher odds of response than patients who receive BSC or docetaxel. Further details are provided in B.2.9.4, page 64.

Figure 9: Objective response rate: NMA results (fixed effect model): Odds ratios for nivolumab versus each of the comparators – B.2.9.4 (page 66)



Abbreviations: BSC: best supportive care.

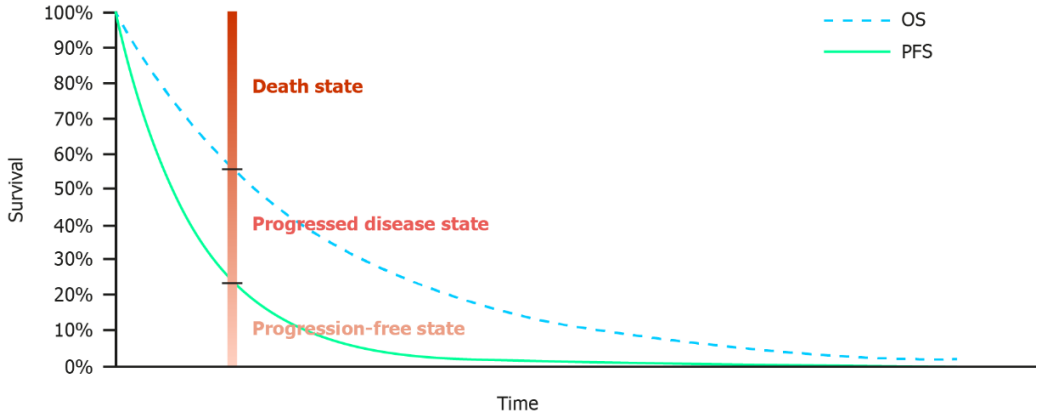
A.9 Key clinical issues

- Clinical evidence for nivolumab consists of single-arm trials; there is no RCT evidence for nivolumab. An STC was therefore conducted to generate indirect estimates of relative effectiveness of nivolumab versus the key comparators including the UK standard of care, paclitaxel. The STC adheres to the recommendations outlined in the recent NICE DSU TSD (TSD18) and has been informed by clinical expert opinion to ensure that all relevant treatment effect modifiers were included. The results presented are consistent with the clinical input elicited and expected outcomes for nivolumab and the comparators in second-line locally advanced unresectable or metastatic UC.
- The OS evidence provided by CheckMate 275 is relatively immature: only 51.6% of patients had died at the time of database lock for the primary analysis (57.0% at the secondary database lock) and, at the secondary database lock, 41.0% of patients were still alive at 1 year. However, validation of the survival outcomes has been undertaken for both nivolumab and the comparators using additional clinical data (from other trials and real-world practice) and clinical expert opinion.

A.10 Overview of the economic analysis

An economic SLR identified no previous economic evaluations for nivolumab as a treatment of locally advanced unresectable or metastatic UC hence a de novo cost-utility model was constructed for the purposes of this appraisal. A cohort-based partitioned survival model was developed that included three mutually exclusive health states: progression-free (PF), post-progression (PP) and death (Figure 10). This choice of model structure was made to capture the progressive nature of UC disease and is consistent with previous submissions to NICE relating to metastatic cancers, including the only previous submission in this specific indication (TA272, 2013).[38]

Figure 10: Schematic representation of the partitioned survival method – see Section B.3.2 (page 88)



Note: The model uses 4-week cycles and a lifetime horizon to capture all relevant costs and patient outcomes.

A summary of the features of the economic analysis is presented in Table 6.

Table 6: Features of the economic analysis

	Previous appraisal	Current appraisal	
Factor	**TA272[38]**	**Chosen values**	**Justification**
Time horizon	5 years	Lifetime	To capture all relevant health consequences and costs
Treatment waning effect?	Not included	Not included	Other immunotherapy trials provide evidence that a continued treatment benefit is observed for some patients up to 10 years[39]
Source of utilities	• Pre-progression utility values were based on trial data (EORTC QLQ-C30 questionnaire) • Post-progression utility values were taken from a study reporting EQ-5D values in terminally ill patients with painful bone metastases or poor prognosis non-small-cell lung cancer • Disutility values associated with treatment-related adverse events were not included	• EQ-5D-3L data from the CheckMate 275[27] trial were used and adjusted using multiple imputation and a mixed model regression • Disutility values associated with treatment-related adverse events were taken from the literature	• The CheckMate 275 trial was deemed the best source following the literature review, which identified only one previous study to use the EQ-5D in this indication.[40] This study was not deemed appropriate due to the use of US population weights, which do not match the NICE reference case • Multiple imputation was used to impute missing values and a mixed model regression analysis used to account for autocorrelation

Source of costs	NHS reference costs, literature and expert opinion.	• Therapy costs were taken from the BNF and eMit • Administration and resource use costs were taken from NHS reference costs and supplemented with evidence from the literature and an advisory board where necessary	• Unit costs were taken from recognised national databases • Clinicians provided advice on resource use for a number of parameters due to a paucity of relevant data in the wider literature[12]

Abbreviations: BNF: British National Formulary; eMIT: electronic market information tool; EORTC QLQ-C30: European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30; EQ-5D: EuroQoL 5-Dimensions; NHS: National Health Service; NICE: National Institute for Health and Care Excellence; US: United States.

A.11 Incorporating clinical evidence into the model

For the base case analysis, a response-based modelling approach was adopted in order to characterise the mechanism of action of nivolumab whereby a subset of patients receive a long and durable response to treatment and survival. The response-based approach works by fitting parametric survival curves to the responding and non-responding patients separately to more accurately characterise the hazard and survival curve in these two groups. With this approach, there can be a risk of immortal time bias, which occurs when the responder and non-responder curves are plotted from treatment initiation. As such, landmark analysis was undertaken. With this approach, PFS and OS for nivolumab were based on the full cohort of patients (i.e. not separated by response), using the pooled Kaplan-Meier data from CheckMate 275 and CheckMate 032 up until a designated landmark point. After this landmark, separate responder and non-responder curves are plotted for the remaining time horizon.[41] The landmark point was chosen to be 8 weeks to reflect the median time to response in the CheckMate trials (1.87 months and 1.48 months, based on RECIST v1.1 criteria). The impact of using longer 26-week landmark, to ensure all patients had responded to treatment, was examined in a scenario analysis. The Kaplan-Meier trial data is used to model PFS and OS up until the landmark point.

Following the advice of the NICE DSU, six parametric distributions were plotted for both responders and non-responders in order to predict long term PFS and OS.[42] These were: exponential, Weibull, Gompertz, lognormal, log-logistic and generalised gamma. In the base case analysis, the generalised gamma distribution was chosen for both responders and non-responders, based on the statistical goodness-of-fit and clinical plausibility of the extrapolation. The impact of this choice was examined during scenario analysis.

To make the PFS and OS curves suitable for the structure of the economic model, and the application of relative treatment effects, it was necessary to combine the separate responder and non-responder curves for all nivolumab-treated patients. To generate combined curves, the separate responder and non-responder curves were weighted based on the number of patients measured as being progression-free and alive at the landmark point in the CheckMate 275 and CheckMate 032. This weighting was assumed to remain constant for the remaining time horizon in each parametric model, which is a conservative assumption given that the weighting would be expected to shift to responding patients over time. The combined PFS and OS curves were further adjusted to account for general population mortality using age-adjusted annual mortality rates based on life tables for England and Wales.[43] Further details are presented in section B.3.3.1 and Appendix L of the main submission document.

For each of the comparators, time-varying HRs were generated based on the STC detailed above in Section A.8 and applied to the combined nivolumab curves for PFS and OS. It was necessary to generate time-varying HRs as the proportional hazard assumption did not hold for these comparisons given the unique mechanism of action for nivolumab. Within the economic model, a separate HR was applied for each cycle in the model.

For BSC, no relevant PFS data were identified during the SLR and so an assumption was made that applying the HR of vinflunine versus BSC for second-line UC patients (1.47) from Bellmunt *et al.* (2009)[44] to one of the other chemotherapy PFS curves provided a reasonable estimate. This HR was assumed to remain fixed for the timeframe of the analysis. Further details are presented in section B.3.3.2 and Appendix L of the main submission document.

A.12 Key model assumptions and inputs

A summary of the key model assumptions and inputs is presented in Table 7.

Table 7: Key model assumptions and inputs

Model input and cross reference	Source/assumption	Justification
Progression-free survival Section B.3.3	Nivolumab: pooled KM data from CheckMate 275 and 032 then parametric survival curve using a generalised gamma distribution. Comparators: STC used to estimate time-varying HRs	See Section A.11
Overall survival Section B.3.3	Nivolumab: pooled KM data from CheckMate 275 and 032 then parametric survival curve using a generalised gamma distribution. Comparators: STC used to estimate time-varying HRs	See Section A.11
Treatment duration Section B.3.3	Treatment duration for nivolumab was derived from CheckMate 275 and extrapolated using a generalised gamma distribution. For the comparators, it was assumed that treatment continued until progression.	The licence for nivolumab indicates that the treatment should be administered as long as clinical benefit is observed or until treatment is no longer tolerated by the patient.[24] Therefore, time on treatment was plotted using individual patient-level data from CheckMate275 and extrapolated. A generalised gamma distribution was chosen based on statistical fit and clinical plausibility from all considered distributions. For the comparators, in clinical practice all therapies are administered until disease progression or unacceptable toxicity.
Utility Section B.3.4	Utility values recorded in CheckMate 275 were used and adjusted using multiple imputation and mixed model regression analysis. Utility was assumed to remain fixed for the full model time horizon. The same utility values were used for nivolumab and chemotherapy with disutilities applied for AEs.	The CheckMate 275 trial was deemed the best source following the literature review, which identified only one previous study in this indication [40]. This study was not deemed appropriate due to the use of US population weights, which do not match the NICE reference case. Multiple imputation was used to impute missing values from CheckMate 275 and a mixed model regression analysis used to account for autocorrelation. Utility remained fixed as it was judged that no data were available to allow

		clinically plausible estimations of utility change over time in this indication.
Treatment costs Section B.3.5	Nivolumab unit costs were taken from the British National Formulary (BNF) with a PAS discount of [redacted] applied. The cost of each comparator was taken from the electronic Market information tool (eMit).	BNF and eMit are standard sources of unit costs for drugs in England and Wales [45, 46].
Administration costs Section B.3.5	The cost of drug administration was derived from the NHS reference cost schedule 2015-16 and applied dependent on doses required per 4-week cycle.[47]	All included drugs are administered intravenously and, therefore, the same cost per event was applied.
Monitoring costs Section B.3.5	Monitoring consisted of regular follow-up visits with an oncologist and a series of ongoing diagnostic tests whilst patients remain on treatment. The type and frequency of visits/tests was based on the cycle length of each treatment regimen. This was based on clinical advice.	Clinical advice was sought due to a lack of published evidence of monitoring costs in this indication.
Best supportive care costs Section B.3.5	The resources provided to patients receiving BSC, including the frequency per month, were informed by clinical advice. It is also assumed that patients on each of the other treatment options receive BSC from the period in which the original treatment stopped until death. This is also based on clinical advice.	Clinical advice was sought due to a lack of published evidence of BSC costs in this indication.

Abbreviations: BNF: British National Formulary; eMIT: electronic market information tool; EORTC QLQ-C30: European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30; EQ-5D: EuroQoL 5-Dimensions; KM: Kaplan-Meier; NHS: National Health Service; NICE: National Institute for Health and Care Excellence; US: United States.

A.13 Base case ICER (deterministic)

The results of the base case analysis including the PAS for nivolumab are summarised in Table 8. Nivolumab was associated with higher costs but also higher quality-adjusted life year (QALY) gains compared with docetaxel, paclitaxel and BSC.

Table 8: Base case results (deterministic) with PAS – see Section B.3.7 (page 114)

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	Incremental ICER (£/QALY)
Nivolumab	[redacted]	2.78	[redacted]				
Paclitaxel	£14,426	1.19	0.76	[redacted]	1.60	[redacted]	£37,647
Docetaxel	£13,945	1.40	0.92	[redacted]	1.38	[redacted]	£44,960
BSC	£9,056	1.01	0.64	[redacted]	1.77	[redacted]	£38,164

Abbreviations: BSC: best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life years gained; PAS: patient access scheme; QALYs, quality-adjusted life years.

A.14 Probabilistic sensitivity analysis

Probabilistic sensitivity analysis (PSA) was undertaken to assess the impact of uncertainty on the model results. The results of the PSA with 1,000 iterations are shown in Table 9 and are similar to the deterministic outputs but with slightly higher ICERs. This increase is due to lower estimates for OS and PFS in the probabilistic analysis caused by the sampling approach for this analysis. These results of the PSA are also presented graphically via scatterplots in Figure 11.

Table 9: Base case results (probabilistic) with PAS – see Section B.3.8 (page 115)

Technologies	Incremental costs (£)	Incremental QALYs	Incremental ICER (£/QALY)
Paclitaxel	[redacted]	[redacted]	£46,209
Docetaxel	[redacted]	[redacted]	£54,220
BSC	[redacted]	[redacted]	£44,698

Abbreviations: BSC: best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life years gained; PAS: patient access scheme; QALYs, quality-adjusted life years.

Figure 11: Scatterplot of probabilistic results – see Section B.3.8 (page 120)

Paclitaxel

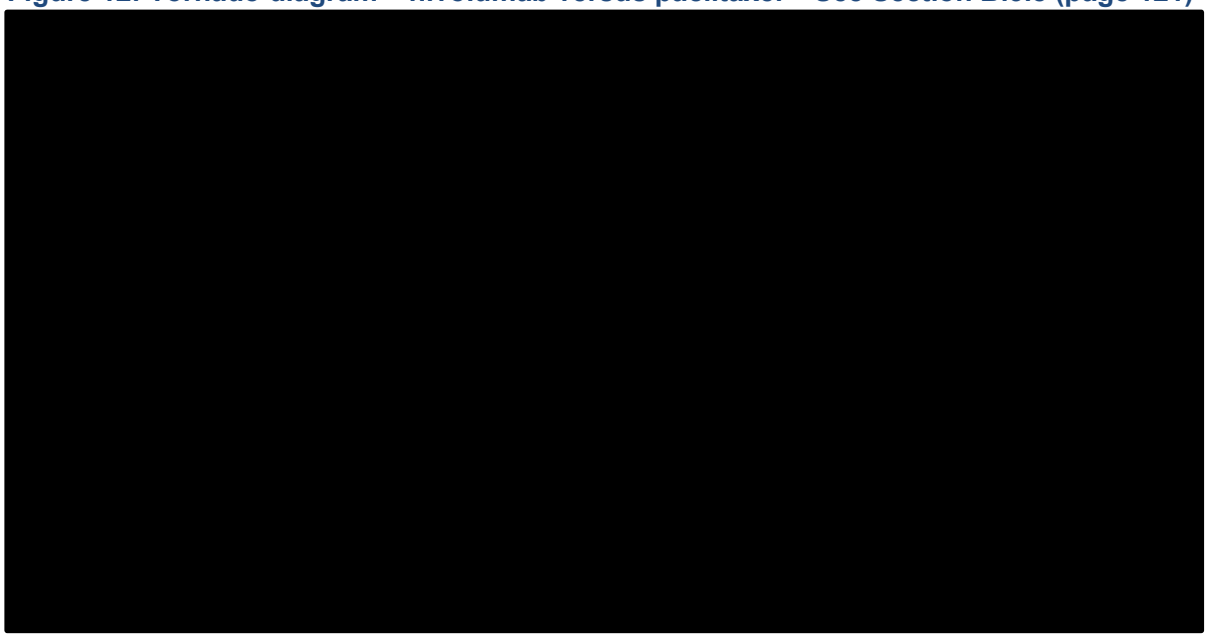
Docetaxel

BSC

A.15 Key sensitivity and scenario analyses

Several model parameters were varied individually in one-way sensitivity analysis with the exception of the survival curves for all treatments. All parameters were varied within the 95% CI, or where this was not possible from available data, within a range of +/-50%. These results are summarised as tornado diagrams in Figure 12 to Figure 14, using net monetary benefit as the metric of cost-effectiveness. Further details are presented in section B.3.8.2 (page 118).

Figure 12: Tornado diagram – nivolumab versus paclitaxel – See Section B.3.8 (page 121)

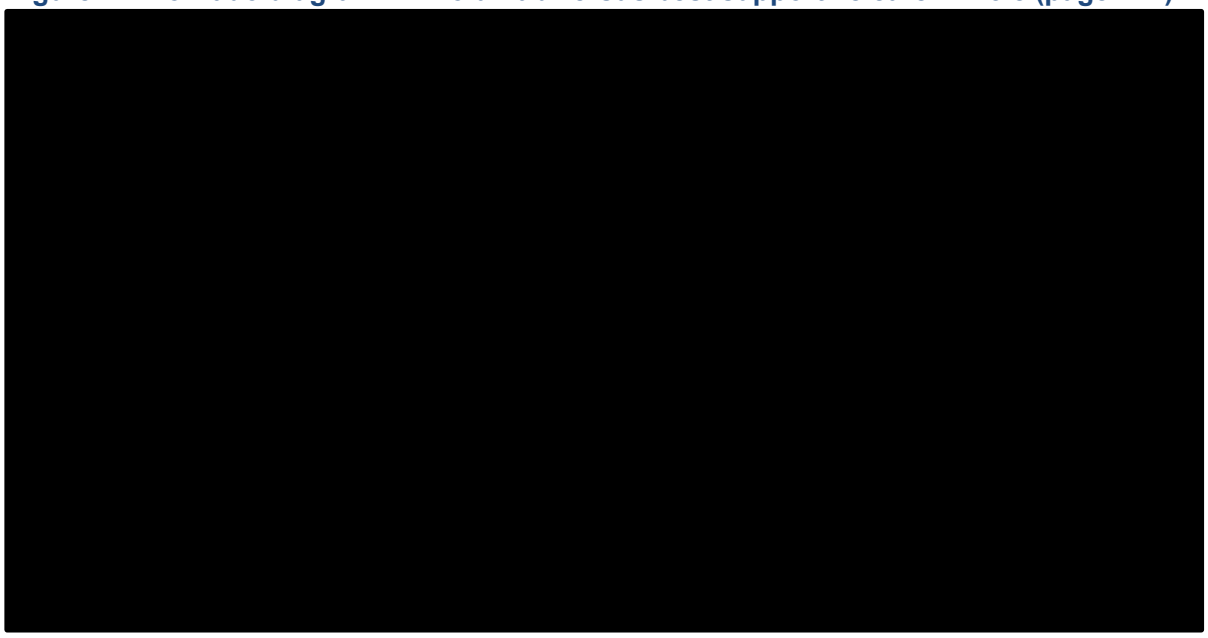


Abbreviations: SD: standard deviation.

Figure 13: Tornado diagram – nivolumab versus docetaxel – B.3.8 (page 122)

Abbreviations: SD: standard deviation.

Figure 14: Tornado diagram – nivolumab versus best supportive care – B.3.8 (page 122)



Abbreviations: SD: standard deviation.

A series of scenario analyses were completed to explore uncertainty regarding key structural assumptions of the analysis. These scenarios, and the key results from each scenario as measured by the ICER (per QALY gained) are presented in Section B.3.8.3 of the main submission document. Table 10 shows the results of the four scenarios that explore the uncertainty associated with the survival modelling and the STC. Three other scenarios were also conducted (inclusion of vial sharing, inclusion of a 2-year stopping rule for nivolumab and alternative parametric curves for time to discontinuation) and within most other scenarios there was a small reduction in all ICERs.

Table 10: Key scenario analysis

Scenario and cross reference	Scenario detail	Brief rationale	Impact on base case ICER
Base case			**Paclitaxel: £37,647** **Docetaxel: £44,960** **BSC: £38,164**
Alternative parametric distributions for PFS and OS curves with 8-week landmark analysis [B.3.8.3 (p 124)]	Weibull Gompertz Lognormal Log-logistic Exponential	The ICER is sensitive to the choice of parametric distribution for PFS and OS.	Paclitaxel: £49,010 to £101,994 Docetaxel: £59,858 to £114,823 BSC: £50,201 to £91,372
All parametric distributions for PFS and OS curves with 26 week	Generalised gamma Weibull	The ICER is sensitive to the choice of parametric distribution for PFS and OS.	Paclitaxel: £34,541 to £60,279 Docetaxel: £40,246 to £76,786 BSC: £34,774 to £61,389

landmark analysis [B.3.8.3 (p 124)]	Gompertz Lognormal Log-logistic Exponential		
Alternative fractional polynomial model [B.3.8.3 (p 125)]	p1=1, p2=1 fractional polynomial model used to estimate OS HRs for comparators.	A number of fractional polynomial models were tested. For OS, two models had almost identical statistical goodness of fit. P1=0, p2=0 model was chosen for the base case as it generated results, which were more clinically plausible in the long term extrapolations of survival but the alternative was model tested here.	Paclitaxel: £56,073 Docetaxel: £59,504 BSC: £43,554
Conservative exponential piecewise modelling [B.3.8.3 (p 125)]	KM data plus exponential distribution from 8 weeks	Piecewise exponential survival modelling was requested in a previous nivolumab appraisal.	Paclitaxel: £53,616 Docetaxel: £65,450 BSC: £55,597
	KM data plus exponential distribution from 26 weeks		Paclitaxel: £55,681 Docetaxel: £71,147 BSC: £57,293

Abbreviations: BSC: best supportive care; KM: Kaplan-Meier; OS: overall survival; PFS: progression-free survival.

A.16 Innovation

For patients with locally advanced unresectable or metastatic UC who have progressed following platinum-based chemotherapy, tolerable treatment options with proven survival benefits are extremely limited. Patients treated with current chemotherapy regimens have an estimated life expectancy of only 5–9 months and are thus considered to be at an end-of-life disease stage. Only 10% of patients typically respond to second-line single-agent chemotherapy regimens, and complete responses are rare and short-lived.[8, 9] Furthermore, many of the chemotherapy agents are associated with high toxicity and many patients instead choose to receive BSC only or clinicians will seek to enrol their patients in a clinical trial. Therefore, there is a critical unmet need for novel, effective and tolerable treatment options, offering durable survival benefit for patients at this stage of disease.[2]

As detailed in Section B.1.2 of the main submission document, rather than relying on the indiscriminate cytotoxic effects of chemotherapy, nivolumab harnesses the body's own immune system to destroy cancer cells via the restoration of anti-tumour T-cell activity and represents a highly innovative mechanism of action. The awarding of a Breakthrough Therapy Designation by the FDA is recognition of the innovative nature of nivolumab.[48] Acting via this novel mechanism of action, nivolumab has demonstrated a predictable and manageable safety profile in UC, consistent with that demonstrated across several previous indications, thus offering improvements in tolerability compared to the cytotoxic chemotherapies currently available to these patients.

The introduction of nivolumab as a highly-innovative and well-tolerated therapy with demonstrable and durable tumour response rates and survival outcomes represents a step-change in the management of patients with locally advanced unresectable or metastatic UC after

the failure of prior platinum-containing chemotherapy. These patients currently have limited effective, tolerable treatment options available and nivolumab has the potential to help address the considerable unmet medical need for these patients at an end-of-life stage.

A.17 End-of-life criteria

It is considered that nivolumab as a treatment for locally advanced unresectable or metastatic UC whose disease has progressed following platinum-containing chemotherapy meets NICE's end-of-life criteria and a summary of this justification is provided in Table 11 below.

Table 11: End-of-life criteria

Criterion	Data available
The treatment is indicated for patients with a short life expectancy, normally less than 24 months	• No studies identified in the SLR reported in Appendix D provided evidence of OS estimates for this patient population that approached 24 months
There is sufficient evidence to indicate that the treatment offers an extension to life, normally of at least an additional 3 months, compared with current NHS treatment	• The economic analysis predicted mean life years per patient with nivolumab of 2.78 years (33.36 months) • In comparison, predicted mean life years per patient with comparator therapies were 1.19 years (14.28 months) with paclitaxel, 1.40 years (16.80 months) with docetaxel and 1.01 years (12.12 months) with BSC. Nivolumab was therefore predicted to offer an extension to life of considerably greater than 3 months versus each of these comparators. Furthermore, in the context of the average survival of patients receiving paclitaxel, docetaxel or BSC, the survival gains offered by nivolumab represent a significant extension to life.

Abbreviations: BSC: best supportive care; OS: overall survival; SLR: systematic literature review.

A.18 Budget impact

Based on available data from Cancer Research UK and expert clinician feedback, the number of patients in England and Wales eligible for treatment with nivolumab, as per the licensed indication for locally advanced unresectable or metastatic UC whose disease has progressed following platinum-containing chemotherapy, is estimated to be 894 patients. Full details regarding the calculation for this eligible patient population are presented in Section B.4 of the full submission document.

The annual net budget impact on the NHS in England and Wales associated with the introduction of nivolumab as a treatment for locally advanced unresectable or metastatic UC is presented in Section B.6 of the main submission document; by Year 5, the annual net budget impact of introducing nivolumab is estimated to be [redacted] (with PAS).

A.19 Interpretation and conclusions of the evidence

Patients with locally advanced unresectable or metastatic UC after failure of platinum-based chemotherapy face a poor prognosis of only 5–9 months with current treatment options; only 10% of patients typically respond to second-line single-agent chemotherapy regimens and complete responses are short, hence many patients in practice choose to be treated with BSC or enter clinical trials in search of effective treatment options.

Two international, multicentre studies have demonstrated the clinical effects of nivolumab in generating durable tumour responses and delaying progression in patients with locally advanced unresectable or metastatic UC after failure of platinum-based chemotherapy. At the latest database lock of CheckMate 275 and CheckMate 032, nivolumab provided median OS estimates of 8.57 months and 9.72 months, with 41.0% and 45.6% of patients alive at 1 year, respectively.

In several other indications, a plateau effect in the survival curve for nivolumab has been demonstrated and the same may be seen in UC with further follow-up. Nivolumab provided stable HRQoL outcomes and a tolerable safety profile consistent with that observed in other indications.

Due to the lack of RCT data, novel techniques were required to generate an STC versus scope-defined comparators for which data were available. This demonstrated clinical benefit of nivolumab versus comparator therapies and provided robust time-varying HRs for input into the economic model.

The economic model utilised a standardised structure consistent with that previously presented to NICE for oncology products, including the only submission in this indication. The results of the economic evaluation indicate that nivolumab is cost-effective for second-line UC patients after failure of prior platinum-containing chemotherapy when compared with the treatment options used in these patients in the UK, especially when compared to the most commonly used treatment, paclitaxel. This conclusion was found to be robust to changes in key assumptions and modelling choices in scenario analyses, and when accounting for combined uncertainty in the model in the PSA.

A.20 References

1. Pasin E, Josephson DY, Mitra AP, et al. Superficial Bladder Cancer: An Update on Etiology, Molecular Development, Classification, and Natural History. Reviews in Urology 2008;10:31-43.
2. Bellmunt J, Orsola A, Leow JJ, et al. Bladder cancer: ESMO Practice Guidelines for diagnosis, treatment and follow-up. Annals of Oncology 2014;25:iii40-iii48.
3. Sharma S, Ksheersagar P, Sharma P. Diagnosis and treatment of bladder cancer. American family physician 2009;80:717-23.
4. Bellmunt J, Theodore C, Demkov T, et al. Phase III trial of vinflunine plus best supportive care compared with best supportive care alone after a platinum-containing regimen in patients with advanced transitional cell carcinoma of the urothelial tract. Journal of Clinical Oncology 2009;27:4454-4461.
5. Choueiri TK, Ross RW, Jacobus S, et al. Double-blind, randomized trial of docetaxel plus vandetanib versus docetaxel plus placebo in platinum-pretreated metastatic urothelial cancer. Journal of Clinical Oncology 2012;30:507-12.
6. Jones R, Hussain S, Protheroe A, et al. Randomized Phase II Study Investigating Pazopanib Versus Weekly Paclitaxel in Relapsed or Progressive Urothelial Cancer. Journal of Clinical Oncology 2017;0:JCO.2016.70.7828.
7. Petrylak DP, Tagawa ST, Kohli M, et al. Docetaxel As Monotherapy or Combined With Ramucirumab or Icrucumab in Second-Line Treatment for Locally Advanced or Metastatic Urothelial Carcinoma: An Open-Label, Three-Arm, Randomized Controlled Phase II Trial. Journal of Clinical Oncology 2016;34:1500-9.
8. Ortmann CA, Mazhar D. Second-line systemic therapy for metastatic urothelial carcinoma of the bladder. Future Oncology 2013;9:1637-1651.
9. Raggi D, Miceli R, Sonpavde G, et al. Second-line single-agent versus doublet chemotherapy as salvage therapy for metastatic urothelial cancer: a systematic review and meta-analysis. Annals of Oncology 2015:mdv509.
10. National Institute for Health and Care Excellence (NICE). [NG2]: Bladder cancer: diagnosis and management. Available: https://www.nice.org.uk/guidance/NG2 [Accessed 3 Feb 2017].
11. European Association of Urology. Muscle-invasive and Metastatic Bladder Cancer Guidelines. Available at: http://uroweb.org/guideline/bladder-cancer-muscle-invasive-and-metastatic/#1 [Accessed 3 Feb 2017], 2016.
12. Bristol-Myers Squibb. Meeting minutes, Clinical Advisory Board Meeting: 06 March 2017, 0930-1600 [Not in the public domain]. 30 Euston Square, London, UK: Bristol-Myers Squibb, 2017.
13. Cancer Research UK. Bladder cancer incidence statistics. Available at: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bladder-cancer/incidence#heading-Zero [Accessed 3 Feb 2017].
14. Brahmer JR, Drake CG, Wollner I, et al. Phase I study of single-agent anti–programmed death-1 (MDX-1106) in refractory solid tumors: safety, clinical activity, pharmacodynamics, and immunologic correlates. Journal of Clinical Oncology 2010;28:3167-3175.
15. Chen DS, Irving BA, Hodi FS. Molecular pathways: next-generation immunotherapy—inhibiting programmed death-ligand 1 and programmed death-1. Clinical cancer research 2012;18:6580-6587.
16. Zandberg DP, Strome SE. The role of the PD-L1: PD-1 pathway in squamous cell carcinoma of the head and neck. Oral oncology 2014;50:627-632.
17. Kamat AM, Flaig TW, Grossman HB, et al. Expert consensus document: consensus statement on best practice management regarding the use of intravesical immunotherapy with BCG for bladder cancer. Nature Reviews Urology 2015;12:225-235.
18. Gontero P, Bohle A, Malmstrom P-U, et al. The role of bacillus Calmette-Guérin in the treatment of non–muscle-invasive bladder cancer. European urology 2010;57:410-429.
19. Liakou CI, Kamat A, Tang DN, et al. CTLA-4 blockade increases IFNγ-producing CD4+ICOShi cells to shift the ratio of effector to regulatory T cells in cancer patients. Proceedings of the National Academy of Sciences 2008;105:14987-14992.

20. Carthon B, Wolchok JD, Yuan J, et al. Pre-operative CTLA-4 blockade: tolerability and immune monitoring in the setting of a pre-surgical clinical trial. Clinical cancer research : an official journal of the American Association for Cancer Research 2010;16:2861-2871.
21. Inman BA, Sebo TJ, Frigola X, et al. PD-L1 (B7-H1) expression by urothelial carcinoma of the bladder and BCG-induced granulomata. Cancer 2007;109:1499-1505.
22. Boorjian SA, Sheinin Y, Crispen PL, et al. T-cell coregulatory molecule expression in urothelial cell carcinoma: clinicopathologic correlations and association with survival. Clinical cancer research 2008;14:4800-4808.
23. European Medicines Agency. Committee for Medicinal Products for Human Use (CHMP) Opinion: Opdivo (nivolumab) EMA/CHMP/251204/2017. Available at http://www.ema.europa.eu/docs/en_GB/document_library/Summary_of_opinion/human/003985/WC500226181.pdf [Accessed 28 Apr 2017]. 2017.
24. Bristol-Myers Squibb Pharmaceuticals Ltd. Nivolumab (OPDIVO) 10 mg/mL concentrate for solution for infusion. Summary of Product Characteristics. 2017. Available at: https://www.medicines.org.uk/emc/medicine/30476 [Accessed 20 June 17].
25. National Institute for Health and Care Excellence (NICE). Final Scope: Nivolumab for treating metastatic or unresectable urothelial cancer - Issued April 2017.
26. Sharma P, Callahan MK, Bono P, et al. Nivolumab monotherapy in recurrent metastatic urothelial carcinoma (CheckMate 032): a multicentre, open-label, two-stage, multi-arm, phase 1/2 trial. The Lancet Oncology 2016;17:1590-1598.
27. Sharma P, Retz M, Siefker-Radtke A, et al. Nivolumab in metastatic urothelial carcinoma after platinum therapy (CheckMate 275): a multicentre, single-arm, phase 2 trial. The Lancet Oncology 2017.
28. Bristol-Myers Squibb. CheckMate 275: Clinical Study Report for Study CA209275 (25th July 2016).
29. Bristol-Myers Squibb. CheckMate 275 Addendum: Clinical Study Report for Study CA209275 (25th October 2016).
30. Galsky M, Retz M, Siefker-Radtke A, et al. Efficacy and safety of nivolumab monotherapy in patients with metastatic urothelial cancer (mUC) who have received prior treatment. Results from the phase II CheckMate 275 study. Presented at the European Society for Medical Oncology (ESMO) Copenhagen 2016. 2016.
31. Bristol-Myers Squibb. CheckMate 032: Clinical Study Report for Study CA209032 (29th June 2016).
32. Gondo T, Ohori M, Hamada R, et al. The efficacy and safety of gemcitabine plus cisplatin regimen for patients with advanced urothelial carcinoma after failure of M-VAC regimen. International Journal of Clinical Oncology 2011;16:345-51.
33. Ozawa A, Tanji N, Ochi T, et al. Gemcitabine and cisplatin for advanced urothelial carcinomas: the Ehime University Hospital experience. International Journal of Clinical Oncology 2007;12:279-83.
34. Phillippo DM, Ades AE, Dias S, et al. NICE DSU Technical Support Document 18: Methods for population-adjusted indirect comparisons in submissions to NICE. Volume 2016. Sheffield: NICE Decision Support Unit, 2016.
35. Jansen JP. Network meta-analysis of survival data with fractional polynomials. BMC medical research methodology 2011;11:61.
36. Dias S, Welton NJ, Sutton AJ, et al. NICE DSU Technical Support Document 2: A generalised linear modelling framework for pairwise and network meta-analysis of randomised controlled trials, 2011; last updated 2016.
37. Bristol-Myers Squibb. Data on File: Expert Clinician Opinion.
38. National Institute for Health and Care Excellence (NICE). [TA272]: Vinflunine for the treatment of advanced or metastatic transitional cell carcinoma of the urothelial tract. Available: https://www.nice.org.uk/guidance/ta272 [Accessed 3 Feb 2017].
39. Schadendorf D, Hodi FS, Robert C, et al. Pooled Analysis of Long-Term Survival Data From Phase II and Phase III Trials of Ipilimumab in Unresectable or Metastatic Melanoma. Journal of Clinical Oncology 2015;33:1889-1894.
40. Pickard AS, Jiang R, Lin HW, et al. Using Patient-reported Outcomes to Compare Relative Burden of Cancer: EQ-5D and Functional Assessment of Cancer Therapy-General in Eleven Types of Cancer. Clinical Therapeutics 2016;38:769-777.

41. Anderson JR, Cain KC, Gelber RD. Analysis of survival by tumor response and other comparisons of time-to-event by outcome variables. J Clin Oncol 2008;26:3913-5.
42. Latimer N. NICE DSU Technical Support Document 14: Survival analysis for economic evaluations alongside clinical trials - Extrapolation with patient-level data. In: Decision Support Unit, ScHARR, Sheffield. Uo, eds, 2011.
43. Office for National Statistics. National life tables, 2013-2015.
44. Bellmunt J, Fougeray R, Rosenberg JE, et al. Long-term survival results of a randomized phase III trial of vinflunine plus best supportive care versus best supportive care alone in advanced urothelial carcinoma patients after failure of platinum-based chemotherapy. Ann Oncol 2013;24:1466-72.
45. British National Formulary. 2017.
46. Department of Health. Drugs and pharmaceutical electronic market information (eMit). In: (CMU) CMU, ed, 2016.
47. Department of Health. NHS reference costs 2015 to 2016. NHS reference costs
48. Bristol-Myers Squibb. Bristol-Myers Squibb's Opdivo (nivolumab) Receives Breakthrough Therapy Designation from U.S. Food and Drug Administration for Advanced Form of Bladder Cancer. Press Release. http://investor.bms.com/investors/news-and-events/press-releases/press-release-details/2016/Bristol-Myers-Squibbs-Opdivo-nivolumab-Receives-Breakthrough-Therapy-Designation-from-US-Food-and-Drug-Administration-for-Advanced-Form-of-Bladder-Cancer/default.aspx [Accessed 17 Feb 2017].

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Nivolumab for treating metastatic or unresectable urothelial cancer after platinum-based chemotherapy

[ID995]

Document B

Company evidence submission

Bristol-Myers Squibb Pharmaceuticals Limited

June 2017

File name	Version	Contains confidential information	Date
ID995_Nivolumab for urothelial cancer_Document B_FINAL ACIC	1.0	Yes	26/07/2017

Instructions for companies

This is the template for submission of evidence to the National Institute for Health and Care Excellence (NICE) as part of the single technology appraisal (STA) process. Please note that the information requirements for submissions are summarised in this template; full details of the requirements for pharmaceuticals and devices are in the user guide.

This submission must not be longer than 150 pages, excluding appendices and the pages covered by this template. If it is too long it will not be accepted.

Companies making evidence submissions to NICE should also refer to the NICE guide to the methods of technology appraisal and the NICE guide to the processes of technology appraisal.

In this template any information that should be provided in an appendix is listed in a box.

Highlighting in the template (excluding the contents list)

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

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

Grey highlighted text in the footer does not work as an automatic form field, but serves the same purpose – as prompt text to show where you need to fill in relevant details. Replace the text highlighted in grey in the header and footer with appropriate text. (To change the header and footer, double click over the header or footer text. Double click back in the main body text when you have finished.)

Contents

Tables and figures

List of abbreviations

Acronym	Definition
Ab	Antibody
AE	Adverse event
AIC	Akaike information criterion
ALT	Alanine transaminase
BIC	Bayesian information criterion
BIRC	Blinded independent review committee
BNF	British National Formulary
BOR	Best overall response
BSA	Body surface area
BSC	Best supportive care
CD28	Cluster of differentiation 28
CEAC	Cost-effectiveness acceptability curve
CHMP	Committee for Medicinal Products for Human Use
CI	Confidence interval
CIS	Carcinoma in situ
Cis	Cisplatin
CR	Complete response
CrI	Credible interval
CSR	Clinical study report
CT	Computer tomography
CTCAE	Common Terminology Criteria for Adverse Events
CTLA-4	Cytotoxic T-lymphocyte-associated protein 4
$\bar{D}_{res}$	Residual deviance
DIC	Deviance information criteria
DOR	Duration of response
EAU	European Association of Urology
ECOG	Eastern Cooperative Oncology Group
EMA	European Medicines Agency
eMIT	Electronic market information tool
EORTC QLQ-C30	European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30
EQ-5D	EuroQoL 5-Dimensions
EQ-5D-3L	EuroQoL 5-Dimensions 3-Levels
ESMO	European Society for Medical Oncology
FDA	Food and Drug Administration
FP	Fractional polynomial
G-CSF	Granulocyte-colony stimulating factor
GCP	Good Clinical Practice
Gem	Gemcitabine
GFR	Glomerular filtration rate

GP	General practitioner
HR	Hazard ratio
HRQoL	Health-related quality of life
ICER	Incremental cost-effectiveness ratio
IFNγR	Interferon gamma receptor
IPD	Individual patient data
IQR	Interquartile range
ITC	Indirect treatment comparison
IV	Intravenous
LPFT	Last patient first treatment
LYG	Life years gained
LYs	Life years
MedDRA	Medical Dictionary for Regulatory Activities
MHC	Major histocompatibility complex
MICE	Multiple imputation by chained equations
MRI	Magnetic resonance imaging
MVAC	Methotrexate, vinblastine, doxorubicin and cisplatin
N/A	Not applicable
NCI	National Cancer Institute
NF-κB	Nuclear transcription factor-κB
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NMA	Network meta-analysis
NMB	Net monetary benefit
NR	Not reached/not reported
NSCLC	Non-small cell lung cancer
OR	Odds ratio
ORR	Objective response rate
OS	Overall survival
PAS	Patient access scheme
PD	Progressive disease
pD	Number of effective parameters
PD-1	Programmed death 1
PD-L1	Programmed death ligand 1
PD-L2	Programmed death ligand 2
PF	Progression-free
PFS	Progression-free survival
PI3K	Phosphoinositide 3-kinase
PP	Post-progression
PR	Partial response
PROs	Patient-reported outcomes
PS	Performance status
PSA	Probabilistic sensitivity analysis

PSS	Personal and Social Services
PSSRU HCHS	Personal and Social Services Research Unit Hospital and Community Health Services
Q2W	Every 2 weeks
Q3W	Every 3 weeks
QALY	Quality-adjusted life year
RECIST	Response Evaluation Criteria In Solid Tumours
SAE	Serious adverse event
SD	Stable disease/standard deviation
SE	Standard error
Shp-2	Src homology 2 domain-containing protein tyrosine phosphatase 2
SLR	Systematic literature review
SmPC	Summary of product characteristics
STA	Single technology appraisal
STC	Simulated treatment comparison
TNM	Tumour-node-metastasis
TTD	Time to treatment discontinuation
TTR	Time to response
TURBT	Transurethral resection of the bladder tumour
UC	Urothelial carcinoma
UICC	Union for International Cancer Control
UK	United Kingdom
US	United States
USA	United States of America
VAS	Visual analogue scale

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

B.1.1 Decision problem

This submission addresses the clinical and cost-effectiveness of nivolumab within its full marketing authorisation for the treatment of locally advanced unresectable or metastatic urothelial carcinoma (UC) in adults after failure of prior platinum-containing chemotherapy.

The decision problem addressed within this submission is consistent with the NICE final scope for this appraisal as outlined in Table 1, with minor discrepancies noted.

Table 1: The decision problem

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
Population	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	N/A
Intervention	Nivolumab	Nivolumab	N/A
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, 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 effects 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 effects of treatment • health-related quality of life (via the EORTC QLQ-C30 and the EQ-5D-3L)	N/A
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. Costs were considered from an NHS and Personal Social Services perspective.	N/A
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.
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.

Abbreviations: N/A: not applicable; PD-L1: programmed death-ligand 1.
Source: NICE final scope [ID995] – issue date: April 2017.[1]

B.1.2 Description of the technology being appraised

A summary of the mechanism of action, marketing authorisation status, costs and administration requirements associated with nivolumab as a treatment for locally advanced unresectable or metastatic UC is presented in Table 2.

Table 2: Technology being appraised

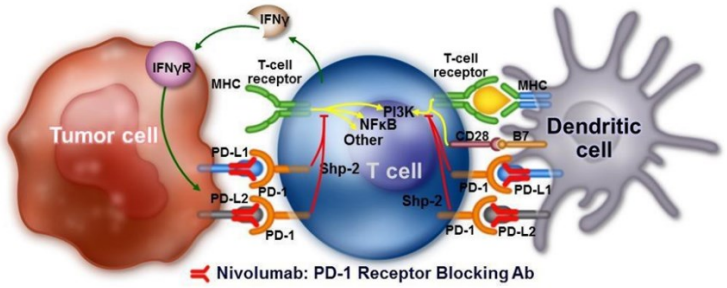
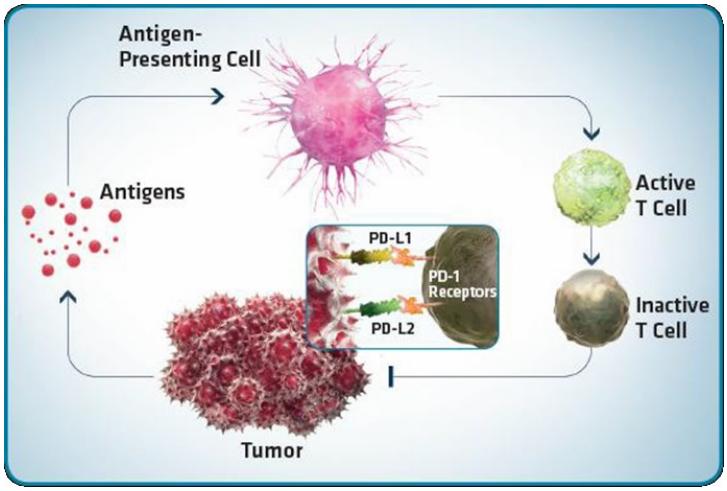
UK approved name and brand name	Nivolumab (Opdivo®)
Mechanism of action	A major part of the immune response to foreign antigens or cells is the activation of T-cells that can destroy them. Activation and de-activation of T-cells is regulated through a complex balance of positive and negative signals via receptors on the T-cell surface (see Figure 1). Cancer cells can exploit these pathways by stimulating inhibitory receptors and in doing so can avoid destruction and facilitate tumour development.[2] Antibodies designed to bind to and block these inhibitor receptors can prevent tumour-driven T-cell suppression and allow restoration of T-cell activity, as depicted in Figure 1. **Figure 1: Regulation of the T-cell immune response**  **Abbreviations:** Ab: antibody; CD28: cluster of differentiation 28; IFNγ: interferon gamma; IFNγR: interferon gamma receptor; MHC: major histocompatibility complex; NF-κB: nuclear transcription factor-κB; PD-1: programmed death 1; PD-L1: programmed death-ligand 1; PD-L2: programmed death-ligand 2; PI3K: phosphoinositide 3-kinase; Shp-2: Src homology 2 domain-containing protein tyrosine phosphatase 2. The programmed death 1 (PD-1) receptor is a negative regulator of T-cell activity and is expressed on activated T-cells. Interaction of PD-1 with its ligands (programmed death-ligand 1, PD-L1, and programmed death-ligand 2, PD-L2) results in the inhibition of T-cell activation and subsequent T-cell death. PD-L1 and PD-L2 are expressed on antigen-presenting cells (such as dendritic cells), and may also be expressed by tumours or other cells in the tumour microenvironment (see Figure 2).[3, 4]

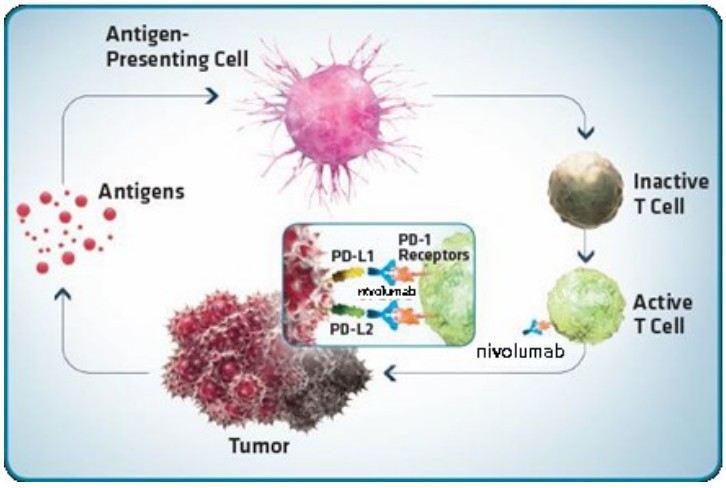
Figure 2: Tumour immune evasion



Abbreviations: PD-1 programmed death 1; PD-L1: programmed death ligand 1; PD-L2: programmed death ligand 2.

Nivolumab (Opdivo®) is a human, monoclonal immunoglobulin G4 (IgG4) antibody that acts as a PD-1 inhibitor, blocking the interaction of PD-1 with PD-L1 and PD-L2 (see Figure 3). By preventing inactivation of T-cells, nivolumab effectively restores T-cell activity against tumour cells, i.e. nivolumab harnesses the patient's own immune system to directly fight cancer cells (in the same way that it would any other "foreign" antigen), resulting in destruction of the tumour.

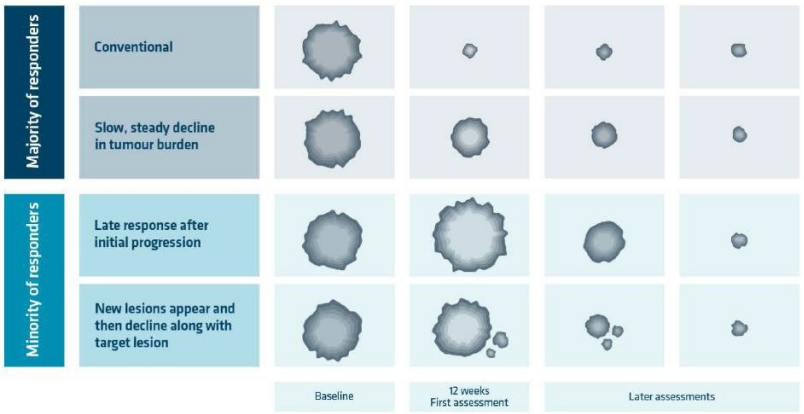
Figure 3: Nivolumab stimulation of immune-mediated destruction



Abbreviations: PD-1 programmed death 1; PD-L1: programmed death ligand 1; PD-L2: programmed death ligand 2.

There is increasing evidence that implicates the PD-1 signalling pathway in UC tumour evasion.[5] By preventing inactivation of T-cells, nivolumab effectively restores T-cell activity against tumour cells, i.e. nivolumab harnesses the patient's own immune system to directly fight cancer cells (in the same way that it would any other "foreign" antigen), resulting in destruction of the tumour.

Furthermore, the potential of targeting immune inhibitory pathways to treat UC is indicated by the effectiveness in some patients of bacillus calmette-guerin therapy (BCG). This immunotherapy

	treatment has been used for over 40 years in patients with high-grade non-muscle-invasive UC following surgical resection. Administered intravesically, BCG induces the secretion of cytokines from urothelial cells and the attraction of vast numbers of neutrophils and monocytes to the tumour site, leading to an immune response against tumour cells.[6, 7] There is also evidence in studies of patients with localised UC that the use of ipilimumab, an immune checkpoint inhibitor that blocks CTLA-4, enhances immune responses and tumour regression.[8, 9] As such, this evidence provides a compelling biological rationale for the effectiveness of nivolumab and the blocking of PD-1 as a therapeutic target in UC.[10, 11] Contrary to conventional anti-cancer therapies, where response to treatment is observed as an immediate shrinkage of the tumour, immune-mediated tumour destruction results in varying patterns of response. Rapid responses are seen, but in other cases immune-checkpoint inhibitors can have an initial effect of making the tumour appear bigger – this is thought to be due to the proliferation of activated T-cells infiltrating the tumour to destroy it. This is commonly referred to as an "unconventional immune-related response" and can result in "pseudo-progression," where patients who ultimately achieve a positive clinical outcome may appear to have progressed when assessed in the early stages of treatment. Typical patterns of response observed with immunotherapies are presented in Figure 4. **Figure 4: Typical patterns of response observed with immune checkpoint inhibitors** 
Marketing authorisation/CE mark status	The licensed indication for nivolumab as a treatment for UC is detailed below: *"Nivolumab (Opdivo®) is indicated for the treatment of locally advanced unresectable or metastatic urothelial carcinoma in adults after failure of prior platinum-containing chemotherapy"* An application for a marketing authorisation in this indication was submitted to the European Medicines Agency (EMA) on the 25th August 2016. A positive opinion from the Committee for Medicinal Products for Human Use (CHMP) was received on the 21st April 2017.[12] Full marketing authorisation was received from the EMA on Monday 5th June 2017. The Summary of Product Characteristics (SmPC) for nivolumab, which details the licensed indication for nivolumab in UC, is provided in the reference pack accompanying this submission.[13]

	In the United States (US), nivolumab received an accelerated approval for patients with locally advanced or metastatic urothelial carcinoma who have disease progression during or following platinum-containing chemotherapy or have disease progression within 12 months of neoadjuvant or adjuvant treatment with a platinum-containing chemotherapy on the 2nd February 2017. This followed the earlier granting of a Breakthrough Therapy Designation by the US Food and Drug Administration (FDA).[14] The Breakthrough Therapy Designation and subsequent accelerated approval reflect the innovative nature and potential benefit of nivolumab to address an unmet medical need.[15] Nivolumab has already been granted a marketing authorisation by the EMA for the following indications, as detailed in the SmPC:[13] • As monotherapy or in combination with ipilimumab, for the treatment of advanced (unresectable or metastatic) melanoma in adults • For the treatment of locally advanced or metastatic non-small cell lung cancer (NSCLC) after prior chemotherapy in adults • As monotherapy for the treatment of advanced renal cell carcinoma after prior therapy in adults • For the treatment of adult patients with relapsed or refractory classical Hodgkin lymphoma after autologous stem cell transplant and treatment with brentuximab vedotin • For the treatment of squamous cell cancer of the head and neck in adults progressing on or after platinum-based therapy
Indications and any restriction(s) as described in the summary of product characteristics (SmPC)	The licensed indication for nivolumab as a treatment for UC is detailed below: *"Nivolumab (Opdivo®) is indicated for the treatment of locally advanced unresectable or metastatic urothelial cancer in adults after failure of prior platinum-containing chemotherapy"*
Method of administration and dosage	The recommended dose and schedule of nivolumab monotherapy in UC is 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. Dose escalation or reduction is not recommended; dosing delay or discontinuation may be required based on individual safety and tolerability.
Additional tests or investigations	As detailed in the SmPC, nivolumab treatment must be initiated and supervised by physicians experienced in the treatment of cancer.[13] Hospital oncology units already have the staffing and infrastructure needed for the administration of IV oncology therapies. Administration of nivolumab is therefore not expected to require any additional NHS infrastructure, as the majority of the comparators included in the final scope for this appraisal (with the exception of best supportive care) are also intravenously administered. The only expected source of differential resource use to the NHS for nivolumab relative to current clinical comparators is in the management of immune-related AEs. In patients with locally advanced unresectable or metastatic UC, nivolumab was found to be well tolerated with a favourable safety profile and a low rate of treatment-related grade 3–4 adverse events (AEs), as detailed in Section B.2.10. However, AEs observed with immunotherapies, such as nivolumab, may differ from those observed with non-

	immunotherapies that are currently used in clinical practice. The immune-based mechanism of action of nivolumab means that many of its treatment-related AEs are immunological in origin. Patients treated with nivolumab are advised to be vigilant and report any changes whilst on treatment to help ensure quick resolution of potential AEs. Immune-related AEs associated with nivolumab, including severe AEs, are well characterised and are generally manageable with topical and/or systemic immunosuppressants.[13] They are often resolved following initiation of appropriate medical therapy, for example corticosteroids, and/or withdrawal of nivolumab.[13] A full list of AEs and guidelines for the discontinuation or withholding of doses in response to immune-related AEs is provided in the SmPC.[13] As detailed in the SmPC for nivolumab, adequate evaluation should be performed to confirm the aetiology or exclude other causes for suspected immune-related AEs.[13]
List price and average cost of a course of treatment	**Concentrate for solution for infusion (sterile concentrate)** — **40 mg vial** — **100 mg vial** **List price:** £439.00 £1,097.00 **PAS price:** [redacted] [redacted] Treatment with nivolumab should be continued as long as clinical benefit is observed or until treatment is no longer tolerated by the patient. Based on the economic model developed for this submission, the average cost of treating a patient with nivolumab in this indication is estimated to be: **List price:** £54,675 **PAS price:** [redacted]
Patient access scheme (if applicable)	A PAS is already in place with the Department of Health for inclusion in this technology appraisal, representing a simple discount of [redacted] on the list price of nivolumab.

Abbreviations: Ab: antibody; CD28: cluster of differentiation 28; CHMP: Committee for Medicinal Products for Human Use; EMA: European Medicines Agency; FDA: Food and Drug Administration; IFNγ: interferon gamma; IFNγR: interferon gamma receptor; IV: intravenous; MHC: major histocompatibility complex; NF-κB: nuclear transcription factor-κB; NHS: National Health Service; NSCLC: non-small cell lung cancer; PAS: patient access scheme; PD-1: programmed death 1; PD-L1: programmed death-ligand 1; PD-L2: programmed death-ligand 2; PI3K: phosphoinositide 3-kinase; Q2W: twice weekly; Shp-2: Src homology 2 domain-containing protein tyrosine phosphatase 2; SmPC: summary of product characteristics; UC: urothelial carcinoma.

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

Overview of the disease

- Urothelial carcinoma (UC) (or transitional cell carcinoma) originates in the lining of the urinary tract and can be characterised based on the extent to which the tumour has invaded the bladder muscle wall as either non-muscle-invasive or muscle-invasive
- Locally advanced and metastatic disease refers to tumours that have grown through the bladder wall and/or have spread to lymph nodes or other distant sites
- UC is the 10th most common cancer in the UK; in 2014, there were 9,021 patients newly diagnosed with UC in England and Wales, of which 7,307 (73%) were in males and 2,756 (27%) were in females
- Due to the anatomical location of the disease and the importance of urinary, bowel and sexual functions to everyday life, the symptoms of UC can have a significant detrimental impact on patient HRQoL

Clinical pathway of care

- Clinical guidelines for the management of UC are available from NICE, the European Society for Medical Oncology (ESMO), and the European Association of Urology (EAU)
- Patients who are diagnosed with locally advanced unresectable or metastatic UC, or have progressed after (neo)adjuvant therapy, will receive cisplatin plus gemcitabine, standard of care in the first-line setting in the UK. Some patients may receive accelerated (high-dose) MVAC (methotrexate, vinblastine, doxorubicin and cisplatin) in combination with granulocyte-colony stimulating factor [G-CSF]) or a carboplatin-based regimen if cisplatin is not tolerated
- For patients who progress on or after first0line platinum-based chemotherapy, effective and tolerated treatment options are limited; expert clinician feedback is that the vast majority of these patients are treated with a taxane-based monotherapy regimen (paclitaxel) or best supportive care (BSC), and the few patients who are fit enough and have a progression-free window of more than 9–12 months following prior platinum-based therapy may receive retreatment with the same first-line platinum-based chemotherapy regimens
- The prognosis of patients with locally advanced unresectable or metastatic UC whose disease has progressed following prior platinum-containing chemotherapy is poor and there is a significant unmet need for licensed, effective and tolerable therapies in this patient population

Estimated UK nivolumab eligible population

- The number of patients in England and Wales potentially eligible for treatment with nivolumab, as per the licensed indication for locally advanced unresectable or metastatic UC whose disease has progressed following platinum-containing chemotherapy, is estimated to be 894 patients

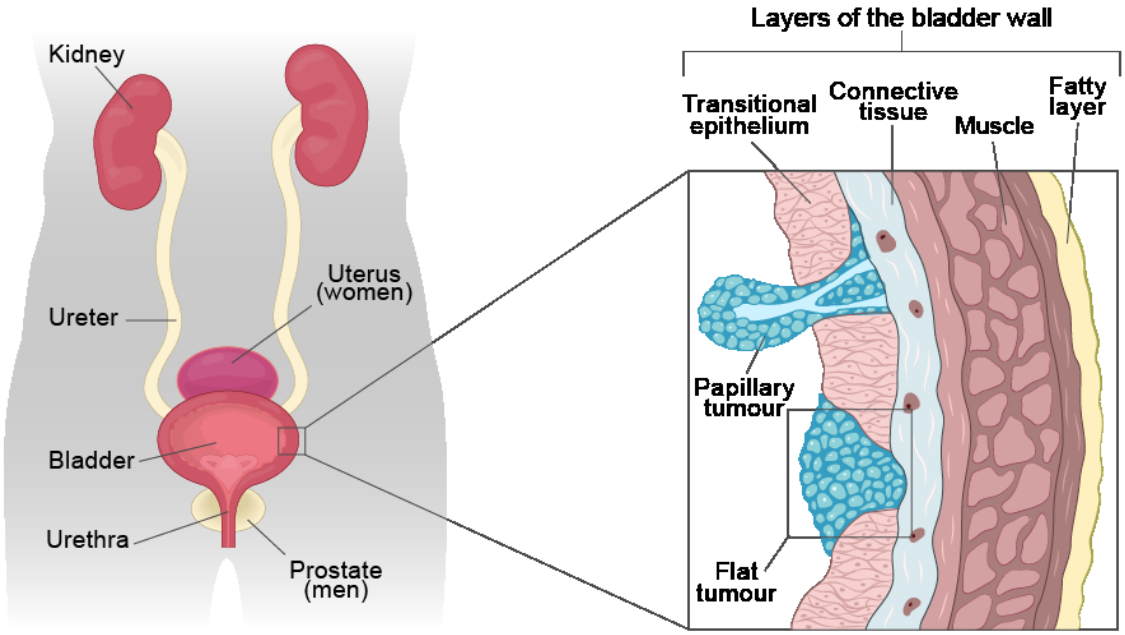
B.1.3.1 Disease overview

UC (or transitional cell carcinoma) is a cancer that originates in the urothelium, the transitional epithelial tissue lining the inner surface of the urinary tract from the renal pelvis (in the kidneys) to the ureter, bladder and proximal two-thirds of the urethra (Figure 5), and accounts for approximately 90% of all bladder cancers.[16] Approximately 10% of bladder tumours may originate in cells other than the transitional epithelium, including squamous cell carcinoma (~5%)

and adenocarcinoma (~1–2%);[16, 17] these rarer, non-transitional forms of bladder cancer tend to be associated with more aggressive disease.[18]

Depending on how far the tumour has invaded the bladder wall, UC can be described as either non-muscle-invasive (early, superficial UC or carcinoma in situ [CIS]) or muscle-invasive UC. Non-muscle-invasive UC can further be divided into two types, papillary or flat, based on how the tumour has grown. With muscle-invasive UC, the tumour has grown and invaded the muscle layers of the bladder wall (Figure 5). Locally advanced and metastatic disease refers to tumours that have grown through the bladder wall and/or have spread to lymph nodes or other distant sites.[19]

Figure 5: Anatomical location of urothelial carcinoma



Source: Adapted from the American Cancer Society.[20]

Presentation, diagnosis and staging

The most common presenting symptom of UC is painless haematuria (blood in the urine), which is typically seen in >80% of patients.[19, 21] In addition to haematuria, patients can present with a variety of other irritative and obstructive urinary symptoms such as dysuria, frequency, urgency, feeling of incomplete voiding, and straining.[22] UC has a considerable impact on urinary, bowel and sexual functions and therefore impacts on daily life and sleeping patterns. These symptoms and disruption to normal bodily function can cause considerable impairment to patient health-related quality of life (HRQoL).

Progression of bladder cancer to an advanced or metastatic stage is associated with further worsening of HRQoL, with patients in late stages of the disease potentially suffering significant limitations to their mobility. Patients with metastatic UC may also present with signs and symptoms of metastatic disease, such as abdominal, bone or pelvic pain, anorexia, cachexia (wasting), or pallor.[22]

Pathological diagnosis of UC is typically made from a biopsy obtained via transurethral resection of the bladder tumour (TURBT). Upon diagnosis, UC tumours are typically staged using the

Union for International Cancer Control (UICC) tumour-node-metastasis (TNM) classification of malignant tumours, a system that describes the anatomical extent of disease based on an assessment of the extent of the primary tumour, the absence or presence and extent of regional lymph node metastasis, and the absence or presence of distant metastasis. Details of the TNM staging system for UC and the progression of the disease are presented in Table 3 and Figure 6 below.

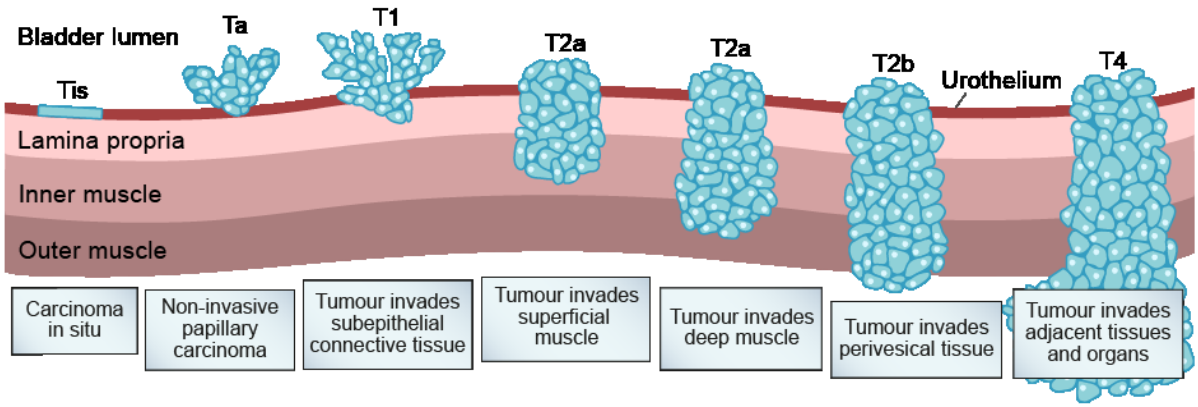
The staging of UC at diagnosis is critical to determining the rate of recurrence and progression of the disease and its associated prognosis, in addition to the appropriate treatment pathway for the patient (see Section B.1.3.2).

Table 3: TNM tumour staging system in urothelial carcinoma

Stage	T	N	M	Description
Stage 0a	Ta	N0	M0	Non-invasive papillary carcinoma. Cancer cells found only on surface of bladder inner lining and can often be easily removed. Tumour has not invaded the muscle or connective tissue of the bladder wall.
Stage 0is	Tis	N0	M0	Flat or carcinoma in situ. Cancer cells found only on the inner lining of the bladder. Tumour has not grown in toward the hollow part of the bladder or spread to the thick layer of muscle or connective tissue of the bladder
Stage I	T1	N0	M0	Tumour has grown through the inner lining of the bladder into the lamina propria but not spread to the thick layer of muscle in the bladder wall or to lymph nodes or other organs
Stage II	T2a–T2b	N0	M0	Muscle-invasive cancer. Tumour has spread into the thick muscle wall of the bladder but has not reached the perivesical tissue (fatty tissue surrounding the bladder) or spread to the lymph nodes or other organs
Stage III	T3a–T3b, T4a	N0	M0	Tumour has spread throughout the muscle wall to the perivesical tissue surrounding the bladder. Tumour may also have spread to the prostate in a man or the uterus and vagina in a woman but not spread to the lymph nodes or other organs
Stage IV	T4b	N0	M0	Tumour has spread to the pelvic wall or the abdominal wall but not to the lymph nodes or other parts of the body
	Any T	N1–N3	M0	Tumour has spread to one or more regional lymph nodes but not to other parts of the body
	Any T	Any N	M1	Tumour may or may not have spread to the lymph nodes but has spread to other parts of the body e.g. bones, liver or lungs

Source: Cancer Research UK.[23]

Figure 6: Progression of urothelial carcinoma



Source: Adapted from Knowles and Hurst (2015).[24]

B.1.3.2 Clinical pathway of care

Clinical guidelines for the management of UC are available from NICE (NICE Guideline 2 – "Bladder cancer: diagnosis and management"), the European Society for Medical Oncology (ESMO), and the European Association of Urology (EAU).[19, 25, 26]

The only published technology appraisal guidance in locally advanced unresectable or metastatic UC is for vinflunine, which was issued a negative recommendation from NICE in 2013 for the treatment of advanced or metastatic transitional cell cancer of the urothelial tract that has been treated previously with platinum-containing chemotherapy.[27]

Details of the current treatment pathway for patients in the UK are presented below, based on the recommendations from these guidelines and feedback from UK expert clinicians experienced in the management of UC.[19, 26, 28, 29]

Muscle-invasive urothelial carcinoma

Patients with newly diagnosed muscle-invasive UC are typically treated with neoadjuvant chemotherapy using a cisplatin-containing combination regimen given prior to either radical cystectomy (surgical removal of the bladder) or radical radiotherapy.[25] Patients for whom neoadjuvant chemotherapy is not suitable (because muscle invasion was not shown before the cystectomy) may be treated with adjuvant cisplatin-containing combination therapy given after radical cystectomy.[25]

Locally advanced or metastatic bladder cancer

First-line chemotherapy

For patients with locally advanced or metastatic UC who are physically fit (Eastern Cooperative Oncology Group [ECOG] performance status [PS] 0 or 1) and have adequate renal function (glomerular filtration rate [GFR] ≥60 mL/min/1.73 m^2), the current standard of care in the first-line setting is cisplatin plus gemcitabine. Some patients may receive accelerated (high-dose) MVAC (methotrexate, vinblastine, doxorubicin and cisplatin) in combination with granulocyte-colony stimulating factor [G-CSF]).[19, 25, 26, 29] Patients with ECOG PS 0-2 and GFR<60 mL/min/1.73m^2 or who are unsuitable for a cisplatin-based regimen may be treated with carboplatin in combination with gemcitabine.[25]

Second-line chemotherapy

For patients who progress on or after first-line platinum-based chemotherapy, effective and tolerated treatment options in the second-line setting are severely limited. Feedback from expert clinicians was that in UK clinical practice, the vast majority of patients with locally advanced unresectable or metastatic UC following prior platinum-based chemotherapy would be treated with paclitaxel monotherapy, with docetaxel monotherapy also used in some centres.[29]

Of those patients considered fit enough to be offered second-line treatment with paclitaxel monotherapy, approximately one third to one half of these patients would typically refuse further chemotherapy treatment, and this figure may be even higher in some smaller centres.[29] These patients would therefore currently opt for best supportive care (BSC), which may include painkillers, steroids and blood transfusions. Some patients would also be unsuitable for chemotherapy altogether, and would therefore be offered BSC instead of taxane-based chemotherapy.[29]

For patients with locally advanced unresectable or metastatic UC whose condition has progressed after first-line therapy and who are physically fit [ECOG PS 0 or 1] with adequate renal function [GFR 60 ml/min/1.73 m^2], NICE recommends retreatment with cisplatin in combination with gemcitabine, or accelerated (high-dose) MVAC in combination with G-CSF. Feedback from expert clinicians was that in UK clinical practice, they would only consider retreatment with platinum-based chemotherapy for patients they considered fit enough and who had been progression-free for at least 9–12 months (or 6 months in some centres) following prior platinum-based chemotherapy; as such, this would very much be the minority of patients, representing only 5–10% of cases in the second-line setting.[29] Patients for whom cisplatin-based chemotherapy is unsuitable (i.e. GFR <60 ml/min/1.73 m^2) may be treated with carboplatin plus paclitaxel in this setting.[25]

The NICE guideline also recommends the use of gemcitabine plus paclitaxel in this second-line setting; however, expert clinician feedback, elicited at an advisory board, was that this regimen is used rarely in few centres across the UK and only for patients who have progressed quickly following first-line platinum chemotherapy and are very symptomatic (e.g. lymph node pain).[29] The ESMO guidelines also refer to the use of gemcitabine monotherapy; however, expert clinician opinion was that this is not used in UK clinical practice and as such, gemcitabine monotherapy is not considered a relevant treatment option in relation to this submission.[28, 29] Furthermore, although both gemcitabine monotherapy and gemcitabine plus paclitaxel are discussed here for transparency, neither is included in the final NICE scope, which further reflects that they do not constitute part of routine UK clinical practice.

Finally, feedback from expert clinicians was that many patients in practice are also entered into clinical trials following progression on their first-line chemotherapy regimen, highlighting the distinct lack of therapeutic options in the second-line setting for patients with this disease.

The prognosis of patients with locally advanced unresectable or metastatic UC whose disease has progressed following prior platinum-containing chemotherapy is between 5 to 9 months with standard chemotherapy options and thus patients are considered to be at an end-of-life disease stage.[30-33] Furthermore, only 10% of patients typically respond to second-line single-agent chemotherapy regimens hence there is a significant unmet need for effective and tolerable treatment options in this patient population.[34, 35]

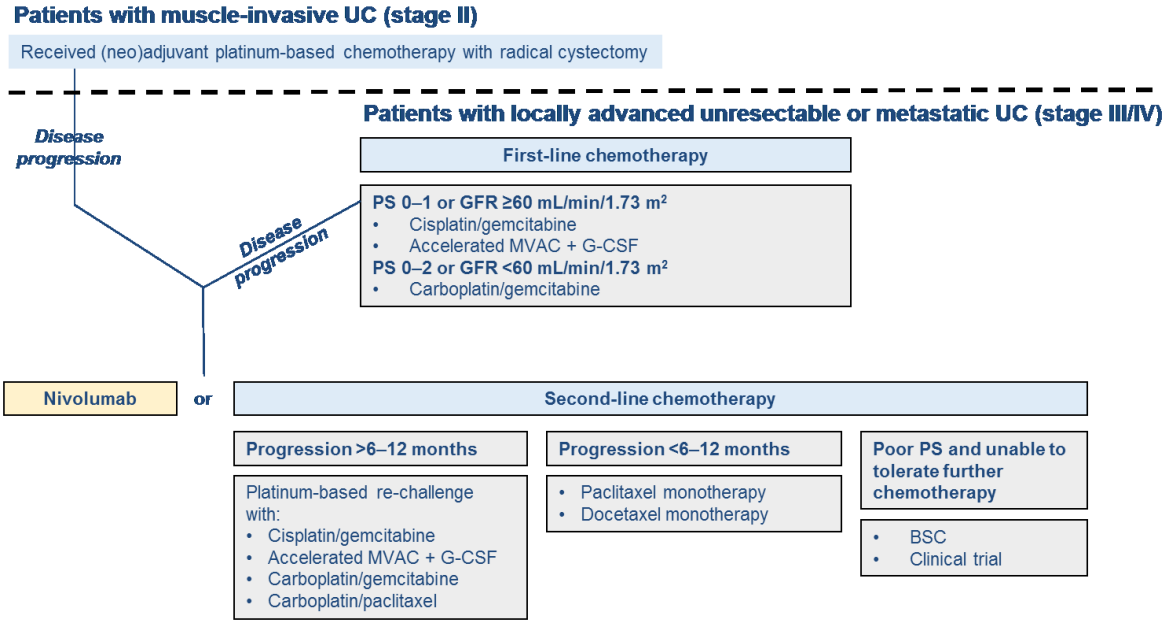
B.1.3.3 Positioning of nivolumab

The patient population for which nivolumab is considered in this submission is in line with its licensed indication, for locally advanced or metastatic UC after failure of prior platinum-containing chemotherapy.[13] As such, there are two potential positions for nivolumab in the treatment pathway of UC:

1. In first-line locally advanced unresectable or metastatic disease, following disease progression after prior platinum-containing therapy received as (neo)adjuvant therapy with radical cystectomy in the muscle-invasive disease stage
2. In second-line unresectable or metastatic disease, following disease progression after prior platinum-containing therapy received in the locally advanced unresectable or metastatic disease stage.

The current treatment pathway for patients with UC and the expected positioning of nivolumab in the UK is presented in Figure 7.

Figure 7: Current treatment pathway for patients with UC



Abbreviations: BSC: best supportive care; G-CSF: Granulocyte-colony stimulating factor; GFR: glomerular filtration rate; MVAC: methotrexate, vinblastine, doxorubicin and cisplatin; PS: performance status; UC: urothelial carcinoma
Source: Adapted from NICE and EAU/ESMO guidelines and expert clinician feedback[19, 25, 26, 29]

Based on the above treatment pathway, the treatment options representing potentially relevant comparators to nivolumab in the context of this submission are as follows:

- Paclitaxel monotherapy (*standard of care)*
- Docetaxel monotherapy
- BSC
- Retreatment with platinum-based chemotherapy *(<10% of patients)*:
  - Cisplatin plus gemcitabine
  - Accelerated MVAC plus G-CSF
  - Carboplatin plus gemcitabine
  - Carboplatin plus paclitaxel.

B.1.3.4 Estimated nivolumab-eligible population

UC is the 10th most common cancer in the UK, and is 3–4 times more commonly found in males than females.[36] In 2014, there were 9,021 patients newly diagnosed with UC in England and Wales, of which 7,307 (73%) were in males and 2,756 (27%) were in females.[36] The disease is also more common in older adults, with more than half (54%) of UC cases in the UK each year diagnosed in patients aged 75 and over.[36]

The majority of patients with UC are diagnosed in stages I and II (62%), with approximately 20% diagnosed at the advanced, metastatic stage.[36] Based on available data from Cancer Research UK and expert clinician feedback, the number of patients in England and Wales eligible for treatment with nivolumab, as per the licensed indication for locally advanced unresectable or metastatic UC whose disease has progressed following platinum-containing chemotherapy, is estimated to be 894 patients. Full details regarding the calculation for this eligible patient population are provided in Section B.4.1.

B.1.4 Equality considerations

The therapies available in current UK clinical practice for patients with locally advanced unresectable or metastatic UC whose disease has progressed following platinum-based chemotherapy comprise further chemotherapy agents, many of which are associated with high toxicity. UC patients are typically older patients, who in many cases cannot tolerate retreatment with chemotherapy and may only be suitable or wish for palliative therapy with BSC.[36]

Due to the lack of well-tolerated, effective treatment options after the failure of platinum-based therapy, some patients could instead be enrolled into clinical trials. The reliance on clinical trials presents a potential equity issue, given that trial centres may not have an equitable geographic distribution and enrolment criteria and numbers for trials are restricted. The availability of a nationally funded treatment option on the NHS for patients whose best, or only, option for receiving active treatment is entry into a clinical trial would help to move towards addressing this equity issue.

B.2 Clinical effectiveness

Summary of the clinical evidence

- A systematic literature review (SLR) identified two clinical trials of nivolumab relevant to the decision problem of this submission: CheckMate 275 (phase II) and CheckMate 032 (phase I/II). No randomised controlled trials for nivolumab versus any of the relevant comparators were identified
- In total, 270 patients (CheckMate 275) and 78 patients (CheckMate 032) with locally advanced unresectable or metastatic UC whose disease had progressed or recurred after treatment with at least one platinum-containing chemotherapy regimen were treated with IV nivolumab 3 mg/kg every 2 weeks until treatment progression (based on blinded independent review committee [BIRC] assessment [CheckMate 275] or investigator assessment [CheckMate 032] of response according to the Response Evaluation Criteria in Solid Tumors (RECIST) v1.1) or unacceptable toxicity
- At the primary analysis database lock of CheckMate 275 (30th May 2016 [n=265]), treatment with nivolumab led to a confirmed objective response per BIRC in 52 patients (19.6%; 95% CI: 15.0–24.9), with 6 (2.3%) patients achieving a complete response (CR)
- Equivalent efficacy was observed in CheckMate 032 (database lock 24th March 2016), with 19 patients (24.4%; 95% CI: 15.3–35.4) achieving a confirmed investigator-assessed objective response, and 5 patients (6.4%) achieving a CR
- At a minimum follow-up of 6 months (primary database lock), median duration of response (DOR) had not yet been reached in CheckMate 275; 76.9% of responders were continuing in response and nearly all patients ([redacted]%) had experienced a DOR of at least 3 months. DOR was also not reached at the time of the database lock in CheckMate 032 and the majority of responders ([redacted]) were still continuing in response, with most responders ([redacted]%) having a DOR of at least 6 months, and [redacted] with a DOR of at least 12 months
- Median PFS was 2.0 months (95% CI: 1.87–2.63) in CheckMate 275 (as per BIRC) and 2.78 months (95% CI: 1.45–5.85) in CheckMate 032 (investigator-assessed). Median OS was 8.74 months (95% CI: 6.05–N/A) in CheckMate 275 and 9.72 months (95% CI: 7.26–16.16) in CheckMate 032. At 12 months, the proportion of patients still alive was 41.0% (95% CI: 34.8–47.1) in CheckMate 275 and 45.6% (34.2–56.3) in CheckMate 032
- Across both trials, consistent results (ORR, PFS and OS) were observed regardless of baseline tumour PD-L1 expression status, including those with PD-L1 expression <1%
- A second database lock of CheckMate 275 (2nd September 2016) provided consistent results with the primary database lock for ORR, PFS and OS; 54 patients (20.0%) achieved an objective response (95% CI: 15.4–25.3), and 2 more patients achieved a CR
- Assessment of HRQoL via the EQ-5D-3L demonstrated that nivolumab increased or maintained patient HRQoL from baseline throughout both trials

Summary of the results of the indirect treatment comparison

- In the absence of RCT data for nivolumab versus the relevant comparators to this submission, an indirect treatment comparison (ITC) using simulated treatment comparison (STC) techniques was conducted. The STC estimated time-varying hazard ratios (HR) for PFS and OS and odd ratios (OR) for ORR for nivolumab versus each of the relevant comparators for which data were available
- Results of the STC found that patients who receive nivolumab have a higher odds of response than patients who receive paclitaxel, docetaxel or BSC (OR: 3.85, 3.12 and 106.7, respectively)
- In terms of OS, the HR for death was greater than 1 (favouring nivolumab) for the majority of time points to week 96, for nivolumab versus paclitaxel, docetaxel and BSC

B.2.1 Identification and selection of relevant studies

An SLR was conducted to identify relevant clinical evidence on the efficacy and safety of nivolumab for the treatment of unresectable or metastatic UC. Full details of the SLR search strategy, study selection process and results can be found in Appendix D.

The searches were conducted on 3 March 2017 and 6 March 2017 and identified 10,866 records. A total of 35 publications reporting on 29 unique trials were ultimately included in the SLR. Of these, 18 publications reported on 12 trials investigating comparators listed in the NICE final scope and considered relevant to this submission, and 5 publications reported on 2 trials investigating nivolumab.

B.2.2 List of relevant clinical effectiveness evidence

Two clinical trials were identified in the SLR that provide evidence for the efficacy and safety of nivolumab in patients with locally advanced unresectable or metastatic UC following the failure of at least one previous line of platinum-based chemotherapy: CheckMate 275 and CheckMate 032.[37, 38]

No RCTs directly comparing the efficacy and safety of nivolumab in the patient population of interest versus any of the comparators relevant to this submission or placebo were identified.

CheckMate 275 is an ongoing, phase II single-arm clinical trial investigating the efficacy and safety of nivolumab in patients with locally advanced unresectable or metastatic UC who had failed at least one previous line of therapy.[38]

CheckMate 032 is an ongoing phase I/II multi-arm trial investigating the efficacy and safety of nivolumab monotherapy or nivolumab in combination with ipilimumab in patients with one of the following tumour types: UC, triple-negative breast cancer, gastric cancer, pancreatic adenocarcinoma, small cell lung cancer, and ovarian cancer.[37] Therefore, only 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). From this point onwards in this submission, reference to CheckMate 032 will refer only to this subgroup of UC patients.[37]

An overview of CheckMate 275 and CheckMate 032 is provided in Table 5.

Table 4: Clinical effectiveness evidence

Study	**CheckMate 275 (NCT02387996)**	**CheckMate 032 (NCT01928394)**
Publications (primary reference in bold)	**Sharma *et al.* (2017)[38]** Clinical study report (plus addendum)[39,40] Galsky *et al.* (2016)[41,a] Sharma *et al.* (2016)[42,a]	**Sharma *et al.* (2016)[37]** Clinical study report[43]
Study design	Multicentre, open-label, single-arm phase II study	Multicentre, open-label, two-stage, multi-arm, phase I/II[b]
Population	Patients with locally advanced unresectable or metastatic UC who had progressed or recurred after at	Patients with locally advanced unresectable or metastatic UC who had progressed or recurred after

	least one previous line of platinum-containing chemotherapy			treatment with at least one platinum-containing chemotherapy regimen		
Intervention(s)	Nivolumab (IV 3 mg/kg Q2W)			Nivolumab (IV 3 mg/kg Q2W)		
Comparator(s)	N/A (single-arm)			N/A[b]		
Indicate if trial supports application for marketing authorisation	Yes	**Indicate if trial used in the economic model**	Yes	Yes	**Indicate if trial used in the economic model**	Yes
Reported outcomes specified in the decision problem	• ORR • OS • PFS • HRQoL via the European Organisation for Research and Treatment of Cancer (EORTC) general cancer module (QLQ-C30) and the EuroQoL-5 dimensions-3 levels (EQ-5D-3L) questionnaires • Adverse events (AEs)			• ORR • OS • PFS • EQ-5D-3L • AEs		
All other reported outcomes	• Duration of response and additional safety outcomes			• Duration of response and additional safety outcomes		

[a]Note that these records were identified by the SLR but subsequently excluded as per the pre-specified protocol. Please see appendix D for further information. [b]CheckMate 032 investigated nivolumab or nivolumab combined with ipilimumab in patients with UC, triple-negative breast cancer, gastric cancer, pancreatic adenocarcinoma, small cell lung cancer, and ovarian cancer. Here, presentation of CheckMate 032 refers only to the nivolumab monotherapy UC cohort (n=86) of relevance to this submission.
Abbreviations: BIRC: blinded independent review committee; CSR: clinical study report; EORTC QLQ-C30: European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30; EQ-5D-3L: 3-level EuroQoL 5-Dimensions; HRQoL: health-related quality of life; IV: intravenous; N/A: not applicable; ORR: objective response rate; OS: overall survival; PD-L1: programmed death ligand 1; PFS: progression-free survival; Q2W: every two weeks; UC: urothelial carcinoma.
Source: Sharma *et al.* (2017),[38] CheckMate 275 CSR,[39] Sharma *et al.* (2016)[37] and CheckMate 032 CSR.[43]

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

B.2.3.1 CheckMate 275

CheckMate 275 is a phase II trial of nivolumab in patients with locally advanced unresectable or metastatic UC with disease progression or recurrence following treatment with at least one prior-platinum containing agent.

Patients with histologically confirmed metastatic or surgically unresectable UC with disease progression or recurrence after at least one platinum-based chemotherapy were enrolled and assigned to a cohort according to tumor PD-L1 expression status (PD-L1 ≥5%, PD-L1 < 5%, or indeterminate). Enrollment in the trial continued until approximately 70 subjects with confirmed PD-L1 expression of ≥5% were treated. Enrollment continued further in Japan until approximately 25 Japanese subjects were treated, or until November 2015, whichever occurred sooner.

Enrolled patients were treated with IV nivolumab 3mg/kg Q2W until documented disease progression (based on RECIST v1.1 criteria) and clinical deterioration, unacceptable toxicity, or other protocol-defined reasons. Treatment beyond initial investigator-assessed RECIST v1.1-

defined progression was permitted if the subject had an investigator-assessed clinical benefit, did not have rapid disease progression, and was tolerating the study drug.

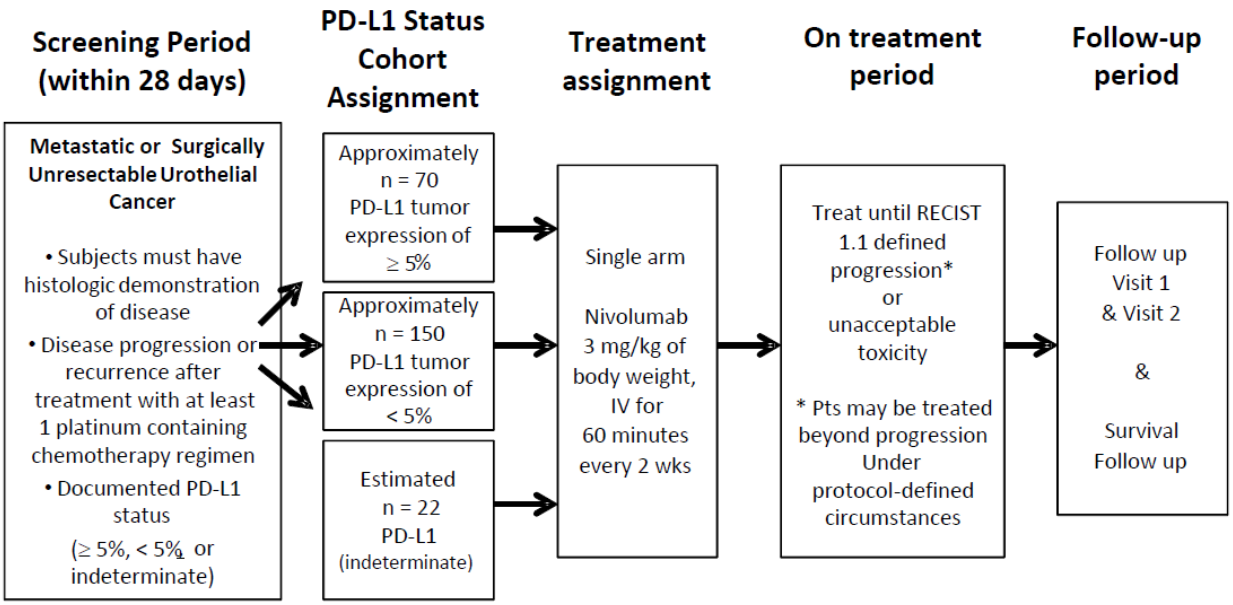
The primary endpoint of CheckMate 275 was objective response rate (ORR) based on Blinded Independent Review Committee (BIRC) assessment using RECIST v1.1 in the all-treated population, in patients with PD-L1 expression ≥1%, and in patients with PD-L1 expression ≥5%. Objective response was defined as the proportion of patients with best overall response of confirmed complete response (CR) or partial response (PR) assessed by the BIRC. Time to response and duration of response were estimated in patients with a confirmed CR or PR. Responses were confirmed at the second scan at least 4 weeks after criteria for objective response were met.

The trial consisted of 3 phases: screening, treatment, and follow-up. Treated subjects were evaluated for response according to the RECIST v1.1 guidelines beginning 8 weeks (±1 week) after the first dose of nivolumab and then every 8 weeks (±1 week) thereafter up to 48 weeks, then every 12 weeks (±1 week) until disease progression (investigator-assessed RECIST v1.1-defined progression) or treatment discontinuation, whichever occurred later. Subjects were followed for OS every 3 months until death, lost to follow-up, or withdrawal of study consent.

A schematic of the CheckMate 275 study design is presented in Figure 8, and a summary of the methodology of CheckMate 275 is presented in Table 6.

Further details of the methodology of CheckMate 275, including the full eligibility criteria, can be found in Appendix M.

Figure 8: CheckMate 275 trial design



Abbreviations: IV: intravenous; PD-L1: programmed cell death ligand 1; RECIST: Response Evaluation Criteria In Solid Tumors; wks: weeks.
Source: CheckMate 275 CSR.[39]

B.2.3.2 CheckMate 032

CheckMate 032 is a multicentre, open-label, two-stage, multi-arm, phase I/II study investigating the efficacy and safety of nivolumab or nivolumab combined with ipilimumab in patients with one of the following tumour types: UC, triple-negative breast cancer, gastric cancer, pancreatic adenocarcinoma, small-cell lung cancer, and ovarian cancer.[37]

Presentation of the CheckMate 032 study refers only to the nivolumab monotherapy UC cohort of relevance to this submission. Eligible patients with histologically or cytologically confirmed carcinoma of the renal pelvis, ureter, bladder, or urethra and disease progression after at least one previous platinum-based chemotherapy treatment were treated with IV nivolumab 3 mg/kg Q2W until documented disease progression (based on RECIST v1.1 criteria), unacceptable toxicity, or other protocol-defined reasons.

A total of 86 patients were enrolled in the trial, of which 78 patients received at least one dose of nivolumab. The primary endpoint of CheckMate 032 was the proportion of patients with a confirmed investigator-assessed objective response, defined as the number of patients with a best overall response of a CR or PR as per the RECIST v1.1 criteria divided by the number of treated patients. Patients were evaluated for response at baseline, 6 weeks after the first dose of nivolumab, continuing every 6 weeks for the first 24 weeks, and then every 12 weeks until disease progression or treatment discontinuation, whichever occurred later. Patients receiving nivolumab monotherapy could switch to nivolumab plus ipilimumab (nivolumab 1 mg/kg and ipilimumab 3 mg/kg or nivolumab 3 mg/kg and ipilimumab 1 mg/kg intravenously, every 3 weeks for four cycles) following disease progression if they met prespecified criteria.

For a CR or PR to be judged to be a best overall response, the assessment needed to be confirmed by a second scan no less than 4 weeks after the criteria for response was first met. Patients who did not meet response-evaluable criteria (i.e. at least one target lesion at baseline and at least one on-study assessment) were judged to be not assessable. Treatment beyond initial investigator-assessed RECIST v1.1-defined progression was permitted if the subject had an investigator-assessed clinical benefit and was tolerating the study drug.

A summary of the methodology and trial design of CheckMate 032 is presented in Table 5. Further details of the methodology of CheckMate 032, including the full eligibility criteria can be found in Appendix M.

Table 5: Summary of CheckMate 275 and CheckMate 032 study methodology

Trial name	CheckMate 275 (n=270)	CheckMate 032 (n=78)
Location	International: 63 sites across 11 countries in North America (USA), Europe, Australia and Asia	International: 16 sites in 5 countries: Finland, Germany, Spain, UK and USA
Trial design	Multicentre, open-label, single-arm phase II study	Multicentre, open-label, multi-arm, phase I/II study[b]
Eligibility criteria for participants	**Key inclusion criteria** • Males and females ≥18 years of age with an ECOG PS 0 or 1 • Histologically or cytologically confirmed metastatic or surgically unresectable transitional cell carcinoma of the urothelium involving the bladder, urethra,	**Key inclusion criteria** • Males and females ≥18 years of age with an ECOG PS 0 or 1 • Measurable disease by CT or MRI per RECIST v1.1 criteria • Locally advanced or metastatic urothelial cell carcinoma • Progression or recurrence

	ureter, or renal pelvis
	• Measurable disease by CT or MRI per RECIST v1.1 criteria
	• Progression or recurrence after treatment
	  ○ With at least 1 platinum-containing chemotherapy regimen for metastatic or surgically unresectable locally advanced urothelial cancer, or
	  ○ Within 12 months of peri-operative (neo-adjuvant or adjuvant) treatment with platinum agent in the setting of cystectomy for localised muscle-invasive urothelial cancer
	• Patients that had received more than 2 prior lines of chemotherapy must not have had liver metastases
	• Availability of tumour samples for PD-L1 expression analysis[a]
	• Previous palliative radiotherapy must have been completed at least 2 weeks before administration of the study drug
	Key exclusion criteria
	• Active brain or leptomeningeal metastases
	• Active, known or suspected autoimmune disease
	• Previous malignancy active within the previous 3 years (except locally curable cancers that appeared to have been cured or carcinoma in situ)
	• Any serious or uncontrolled medical disorder
	• Autoimmune disease (vitiligo, type 1 diabetes mellitus, residual hypothyroidism due to an autoimmune condition only requiring hormone replacement, psoriasis not requiring systemic treatment, or conditions not expected to recur in the absence of an external trigger were permitted)
	• Systemic treatment with either corticosteroids (>10 mg daily prednisone equivalents) or other immunosuppressive medications within 14 days of first study drug administration
	• Prior treatment with an anti-PD-1, anti-PD-L1, anti-PD-L2, anti-CTLA-4 antibody, anti-CD137, or any other antibody or drug specifically
		○ After at least 1 previous platinum-containing chemotherapy treatment for metastatic or locally advanced unresectable urothelial cancer, or
		  ○ Recurrence within 1 year of completing previous platinum-based neoadjuvant or adjuvant treatment
		  ○ After previously refusing standard treatment with chemotherapy for the treatment of metastatic (stage IV) or locally advanced disease
		Key exclusion criteria
		• Active brain metastases or leptomeningeal metastases
		• Any serious or uncontrolled medical disorder
		• History of or active, known or suspected autoimmune disease (vitiligo, type 1 diabetes mellitus, residual hypothyroidism caused by auto immune thyroiditis, and disorders not expected to recur in the absence of an external trigger were permitted)
		• Need for immunosuppressive doses of systemic corticosteroids (>10 mg daily prednisone equivalents) for at least 2 weeks before study drug administration
		• Prior treatment with experimental anti-tumour vaccines or any modulator of T-cell function or checkpoint pathway
		A full list of inclusion and exclusion criteria is presented in Appendix M.

	targeting T-cell co-stimulation or immune checkpoint pathways • Treatment with any chemotherapy, radiation therapy, biologics for cancer, or investigational therapy within 28 days of first study drug administration • All toxicities attributed to previous anticancer therapy other than neuropathy, alopecia, and fatigue must have resolved to grade 1 or baseline before administration of study drug. A full list of inclusion and exclusion criteria is presented in Appendix M.	
Settings and locations where the data were collected	• The study was conducted in a secondary care (hospital) setting at 63 sites across 11 countries worldwide • The study was conducted in accordance with Good Clinical Practice guidelines by qualified investigators using a single protocol to promote consistency across sites	• The study was conducted in a secondary care (hospital) setting at 16 sites across 5 countries worldwide • The study was conducted in accordance with Good Clinical Practice guidelines by qualified investigators using a single protocol to promote consistency across sites
Method of study drug administration	• Nivolumab 3mg/kg Q2W via IV infusion over 60 minutes • Treatment was continued until progression, unacceptable toxicity, or withdrawal of consent • Patients were permitted to continue treatment beyond investigator-assessed RECIST v1.1-defined progression if they were experiencing a clinical benefit, as determined by the investigator, and were tolerating the study drug • No dose modifications were allowed, but predefined dose delays were permitted for adverse events	• Nivolumab 3mg/kg Q2W via IV infusion over 60 minutes • Treatment was continued until progression, unacceptable toxicity, or withdrawal of consent. Patients were permitted to continue treatment beyond investigator-assessed RECIST v1.1-defined progression if they were experiencing a clinical benefit, as determined by the investigator, and were tolerating the study drug • Patients could switch to nivolumab plus ipilimumab (nivolumab 1 mg/kg and ipilimumab 3 mg/kg or nivolumab 3 mg/kg and ipilimumab 1 mg/kg intravenously, every 3 weeks for four cycles) after progression if they met pre-specified criteria.
Permitted and disallowed concomitant medication	The following medications were prohibited during the study: • Immunosuppressive agents (except to treat a drug-related adverse events) or systemic corticosteroids (>10 mg daily prednisone equivalent) within 14 days of study drug administration[b] • Any antibody or drug specifically targeting T-cell co-stimulation or checkpoint pathways, or chemotherapy, radiation therapy, biologics for cancer, or	The following medications were prohibited during the study: • Immunosuppressive agents (except to treat a drug-related adverse event) • Systemic corticosteroids >10 mg daily prednisone equivalent[b] • Any concurrent antineoplastic therapy (i.e. surgery, chemotherapy, hormonal therapy, immunotherapy, radiation therapy except for palliative radiation therapy described above or

	investigational therapy within 28 days of first study drug administration	standard or investigational agents for treatment of cancer) Supportive care for disease-related symptoms was permitted to be offered to all patients on the trial. Palliative (limited-field) radiation therapy and palliative surgical resection were permitted if the certain protocol-defined criteria were met.
Primary endpoint	• The primary endpoint of CheckMate 275 was BIRC-assessed ORR (as per RECIST v1.1) in the all-treated population, in patients with PD-L1 expression ≥1%, and in patients with PD-L1 expression ≥5% • ORR was defined as the number of patients with a best overall response (BOR) of confirmed complete response (CR) or partial response (PR) divided by the number of all-treated patients, PD-L1 ≥1% patients or PD-L1 ≥5% subjects, respectively	• The primary endpoint of CheckMate 032 was confirmed investigator-assessed ORR • ORR was defined as the number of patients with a BOR of CR or PR as per RECIST v1.1 divided by the number of treated patients

Secondary and exploratory endpoints	**Secondary endpoints:**
	• BIRC-assessed PFS
	• OS
	• Investigator-assessed ORR
	(in the all-treated population, patients with PD-L1 expression ≥1%, and patients with PD-L1 expression ≥5%)
	
	Exploratory endpoints:
	• Investigator-assessed PFS
	• Safety
	• HRQoL via the EORTC QLQ-C30 questionnaire
	• General health status via the EQ-5D-3L questionnaire
	• Pharmacokinetics and exploration of exposure-response relationships*
	• Immunogenicity*
	• Pharmacodynamic activity in the peripheral blood and tumour tissue as measured by flow cytometry, immunohistochemistry, soluble factor analysis, and gene expression (microarray technology, quantitative RT-PCR)*
	• Association between biomarkers in the peripheral blood and tumour tissue with safety and efficacy*
	**Outcomes not considered relevant to present in this submission*
		Secondary endpoints:
		• Investigator-assessed PFS
		• OS
		• DOR
		• Safety
		
		Exploratory endpoints:
		• Assessed by PD-L1 expression (≥1% and <1%):
		  ○ ORR
		  ○ OS
		  ○ PFS
		• HRQoL via the EQ-5D and EQ-VAS questionnaires

Timing of assessments	• Tumour assessments were scheduled at 8 weeks from the date of first dose (±1 week), then every 8 weeks (±1 week) thereafter up to 48 weeks, then every 12 weeks (±1 week) until documented disease progression or treatment discontinuation (whichever occurred last). Assessments were performed using CT or MRI and included the pelvis, chest, abdomen and all known sites of disease • Survival assessment was scheduled every 3 months until death, lost to follow-up or withdrawal of study consent • AEs were assessed during treatment visits and were included in the safety analyses if they occurred within 30 days from the day of the last dose received • HRQoL and general health status were assessed before each dose at Week 1, then every 8 weeks up to 48 weeks, then every 12 weeks until disease progression or treatment discontinuation (whichever occurred later) Two follow-up visits and subsequent survival follow-up visits were also scheduled for AEs and HRQoL measures[c]	• Treated subjects were evaluated for response by the investigator according to the RECIST v1.1 at baseline and then every 6 weeks (±1 week) from first dose for the first 24 weeks, then every 12 weeks (±1 week) until disease progression or treatment was discontinued (whichever occurred later) • Assessments were performed using CT or MRI and included the pelvis, chest, abdomen and all known sites of disease • AEs were assessed during treatment visits. Safety was defined as the incidence of treatment-related adverse events leading to drug discontinuation within the first 12 weeks of treatment in patients who had at least one dose of study drug • HRQoL was assessed before study drug administration through Week 13, then at the same time of subsequent tumour assessments, during Follow-Up Visit 1 and 2 and survival visits Two follow-up visits and subsequent survival follow-up visits were also scheduled (AEs and HRQoL)[c]
Pre-planned subgroups	• A pre-planned analysis of the primary and secondary endpoints in patients with PD-L1 expression <1% and ≥1% was conducted • Further subgroup analyses were conducted to assess the impact of pre-specified baseline characteristics, site of original tumour origin (bladder, renal pelvis/ureter), number of Bellmunt risk factors, and prior cancer therapy regimens (number of prior regimens in a metastatic setting, time from completion of most recent prior regimen to study treatment) on confirmed ORR per BIRC, PFS and OS	• As part of the exploratory endpoints, ORR, OS and PFS were analysed in subgroups defined by PD-L1 expression (<1% and ≥1%). • In addition, ad-hoc subgroup analyses were conducted to assess the impact several key baseline factors such as ECOG-PS, metastases, or haemoglobin on investigator-assessed ORR
Duration of study and follow-up	The first patient was treated on the 9th March 2015 and the trial is currently ongoing. The last patient last visit date for the primary database lock of the 30th May 2016, data from which are presented in this submission, was the 15th April 2016.	The first patient was treated on the 5th June 2014 and the trial is currently ongoing. The last patient last visit date for the primary database lock of 24th March 2016 was the 11th February 2016, data from which are presented in this submission.

	A further database lock took place on 2nd September 2016 and data from this are also presented in this submission.	

[a]Patients were required to have an evaluable tumour tissue sample for PD-L1 expression testing at screening, but were not excluded based on PD-L1 status. [b]Several advanced or metastatic solid tumour types were studied in CheckMate 032, but only the urothelial carcinoma arm treated with nivolumab monotherapy is presented in this submission. CPatients were followed for at least 100 days after the last dose of study drug. Follow-up Visit 1 was scheduled for 35 days from the last dose ±7 days or coincided with the date of discontinuation (± 7 days) if date of discontinuation was >35 days after last dose. Follow-up Visit 2 was scheduled for 80 days (±7 days) from follow-up Visit 1. Survival follow-up visits were scheduled for every 3 months (± 7 days) from Follow-up Visit 2.
Abbreviations: AEs: adverse events; BIRC: blinded independent review committee; BOR: best overall response; CR: complete response; CT: computer tomography; CTLA-4: cytotoxic T-lymphocyte-associated protein 4; ECOG: Eastern Cooperative Oncology Group; EORTC QLQ-C30: European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30; EQ-5D-3L: 3-level EuroQoL 5-Dimensions; GCP: Good Clinical Practice; HRQoL: health-related quality of life; IV: intravenous; MRI: magnetic resonance imaging; ORR: objective response rate; OS: overall survival; PD-1: programmed death 1; PD-L1: programmed death ligand 1; PD-L2: programmed death ligand 2; PFS: progression-free survival; PR: partial response; PROs: patient-reported outcomes; PS: performance status; RECIST: response evaluation criteria in solid tumours.
Source: Sharma *et al.* (2017),[38] CheckMate 275 CSR,[39] Sharma *et al.* (2016)[37] and CheckMate 032 CSR.[43]

B.2.3.3 Eligibility criteria

The full eligibility criteria for enrolment in CheckMate 275 and CheckMate 032 are provided in Appendix M.

B.2.3.4 Baseline characteristics

Baseline demographics, disease characteristics and a summary of prior therapies of the patients included in CheckMate 275 and CheckMate 032 are presented in Table 6.

In CheckMate 275, median age was 66 years, the majority of patients were white and male, and over 70% were current or former smokers. The vast majority of patients (96.7%) had metastatic disease. Overall 71.5% of patients had received at least one prior regimen in the metastatic disease setting, and 29.3% had received two or more prior regimens for metastatic disease. Prior systemic cancer therapy was less common in the neoadjuvant and adjuvant settings, with 22.2% receiving at least one neoadjuvant regimen and 30.7% of patients receiving prior regimen(s) in the adjuvant setting.

The baseline characteristics of the patients in CheckMate 032 were similar to those in CheckMate 275. The median age of the patient population in CheckMate 032 was 66 years; the majority were white (92.3%) and male (69.2%). The vast majority (91%) of patients had metastatic (stage IV) disease, and 75.6% of patients had at least two disease sites.

Expert clinician feedback was that the patient populations of CheckMate 275 and CheckMate 032 were very similar, and could be considered generally representative of the patient population expected to receive nivolumab in UK clinical practice.[29] Across both trials, expert clinician feedback was that the proportion of patients with PS 0 was perhaps slightly over-representative of the number of patients likely to have PS 0 in this setting, and that the median age of the patients in both trials may be slightly lower than the age of the average UC patient treated in the second-line setting in UK clinical practice. However, a recent chart review conducted in UK clinical practice of patients with locally advanced unresectable or metastatic UC initiating second-line therapy found that the mean patient age was in fact very similar, albeit slightly lower (mean of 62.8 years), than in both CheckMate trials.[44]

Table 6: Baseline characteristics of patients in the all-treated population of CheckMate 275 and CheckMate 032

Characteristic	CheckMate 275 Total (n=270)	CheckMate 032 Total (n=78)
Demographics		
Age, median years (range)	66 (38–90)	66 (31–85)
Age categorisation, n (%)		
<65	122 (45.2)	37 (47.4)
≥65 and <75	110 (40.7)	31 (39.7)
≥75 and <85	35 (13.0)	N/A
≥75	N/A	10 (12.8)
>85	3 (1.1)	N/A
Male, n %	211 (78.1)	54 (69.2)
Race, n %		
White	231 (85.6)	72 (92.3)
Asian	30 (11.1)	1 (1.3)
Black	2 (0.7)	4 (5.1)
Other	3 (1.1)	1 (1.3)
Not reported	4 (1.5)	N/A
Region, n (%)		
US	106 (39.3)	59 (75.6)
Japan	23 (8.5)	0 (0.0)
Rest of world	141 (52.2)	19 (24.4)
Tobacco use, n (%)		
Current/former smoker	194 (71.9)	48 (61.5)
Never smoked	67 (24.8)	29 (37.2)
Unknown	9 (3.3)	1 (1.3)
Disease characteristics		
ECOG PS, n (%)		
0	145 (53.7)	42 (53.8)
1	124 (45.9)	36 (46.2)
3	1 (0.3)	0
Bellmunt risk factors, n (%)		
0	98 (36.3)	27 (34.6)
1	111 (41.1)	39 (50.0)
2	46 (17.0)	8 (10.3)
3	15 (5.6)	4 (5.1)
Site of primary tumour, n (%)		
Urinary bladder	197 (73.0)	NR
Renal pelvis	46 (17.0)	NR
Ureter	19 (7.0)	NR
Urethra	8 (3.0)	NR
Disease setting, n (%)		
Metastatic	261 (96.7)	71 (91.0)
Locally unresectable/non-metastatic	9 (3.3)	7 (9.0)
Baseline metastases, n (%)		
Any visceral involvement	227 (84.1)	61 (78.2)
Liver	75 (27.8)	20 (25.6)

Lymph node only	43 (15.9)	13 (16.7)
PD-L1 expression, n (%)		
Assessable	N/A	67 (85.9)
<1%	N/A	42 (53.8)
≥1%	124 (45.9)	25 (31.8)
<5%	N/A	53 (67.9)
≥5%	83 (30.7)	14 (17.9)
Number of sites with ≥1 lesion, n (%)		
1	85 (31.5)	19 (24.4)
2	94 (34.8)	30 (38.5)
3	51 (18.9)	24 (30.8)
4	29 (10.7)	3 (3.8)
≥5	11 (4.1)	2 (2.6)
Prior therapy		
Prior systemic therapy regimen setting, n (%)		
Metastatic	193 (71.5)	N/A
Adjuvant	83 (30.7)	33 (42.3)
Neo-adjuvant	60 (22.2)	14 (17.9)
Previous therapies in metastatic setting, n (%)		
0	77 (28.5)	N/A
1	114 (42.2)	26 (33.3)
2	57 (21.2)	N/A
2-3	N/A	42 (53.8)
>3	N/A	10 (12.8)
≥3	22 (8.1)	N/A
Prior surgery related to cancer, n (%)	250 (92.6)	71 (91.0)
Prior radiotherapy, n (%)	85 (31.5)	25 (32.1)

Abbreviations: ECOG PS: Eastern Cooperative Oncology Group performance status; N/A: not applicable; NR: not reported; PD-L1: programmed death ligand 1.
Source: Sharma *et al.* (2017)[38] Sharma *et* al. (2016),[37] CheckMate 275 CSR[39] and CheckMate 032 CSR.[43]

B.2.3.5 Subsequent therapies

Details of the subsequent therapies received by patients in CheckMate 275 and CheckMate 032 following discontinuation of nivolumab are provided in Table 7 below. Expert clinician feedback was that the proportion of patients receiving subsequent therapy in both trials might be considered slightly lower than would typically be seen in clinical practice, where a larger proportion of patients would likely go on to try further treatment.[29] This is likely due to the fact that patients in clinical trials are kept on treatment for longer (with some patients treated beyond progression in some cases).

Table 7: Summary of subsequent anti-cancer therapies received in CheckMate 275 and CheckMate 032

Subsequent therapy, n (%)	CheckMate 275 (n=265)	CheckMate 032 (n=78)
Patients with any subsequent therapy[a]	52 (19.6)	23 (29.5)
Patients who received subsequent radiotherapy	25 (9.4)	9 (11.5)

Subsequent therapy, n (%)	CheckMate 275 (n=265)	CheckMate 032 (n=78)
Patients who received subsequent surgery	8 (3.0)	5 (6.4)
Patients who received subsequent systemic therapy	26 (9.8)	14 (17.9)
Other systemic cancer therapy – chemotherapy	25 (9.4)	11 (14.1)
Antineoplastic	1 (0.4)	1 (1.3)
Bevacizumab	1 (0.4)	0 (0)
Carboplatin	5 (1.9)	3 (3.8)
Cisplatin	5 (1.9)	3 (3.8)
Cyclophosphamide	1 (0.4)	0 (0)
Docetaxel	4 (1.5)	1 (1.3)
Doxorubicin	1 (0.4)	3 (3.8)
Everolimus	1 (0.4)	1 (1.3)
Gemcitabine	10 (3.8)	7 (9.0)
Ifosfamide	0 (0)	1 (1.3)
Lapatinib	0 (0)	1 (1.3)
Methotrexate	2 (0.8)	0 (0)
Nivolumab	2 (0.8)	0 (0)
Paclitaxel	5 (1.9)	4 (5.1)
Pemetrexed	2 (0,8)	0 (0)
Trametinib	1 (0.4)	0 (0)
Vinblastine	1 (0.4)	1 (1.3)
Vincristine	1 (0.4)	0 (0)
Vinflunine	4 (1.5)	0 (0)
Other systemic cancer therapy - experimental drugs	3 (1.1)	3 (3.8)
Investigational antineoplastic drug	3 (1.1)	3 (3.8)

[a] Patient may have received more than one type of subsequent therapy. Subsequent therapy was defined as therapy started on or after first dosing date.
Abbreviations: PD-L1: programmed death ligand 1.
Source: CheckMate 275 CSR.[39]

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

The statistical analyses used for the primary and secondary endpoints alongside sample size calculations and methods for handling missing data are presented in Table 8.

Table 8: Statistical methods for the primary analysis of CheckMate 275 and CheckMate 032

Trial name	CheckMate 275	CheckMate 032
Hypothesis objective	Treatment with nivolumab monotherapy would lead to clinical benefit in patients with metastatic or	Treatment with nivolumab monotherapy will have clinical activity

	surgically unresectable UC who have progressed post platinum treatment as demonstrated by a clinically meaningful ORR	in subjects with advanced or metastatic tumours
Statistical analysis	• ORRs (both BIRC- and investigator-assessed) were summarised by a binomial response rate and their corresponding two-sided 95% exact CIs using the Clopper-Pearson method.[45] BOR was summarised by response category • Median values of DOR were calculated along with two-sided 95% CI using Brookmeyer and Crowley method.[46] TTR was summarised using descriptive summary statistics for the responders • Time-to-event distributions were estimated using Kaplan-Meier techniques. This was done for PFS, OS and DOR (note that time to response was analysed using summary statistics such as mean, SD, median, min, max). ○ Median survival time along with 95% CIs were constructed based on a log-log transformed CI for the survivor function S(t)[46, 47] ○ Rates at fixed time points were derived from the Kaplan-Meier estimate and corresponding confidence interval were derived based on Greenwood formula[48] for variance derivation and on log-log transformation applied on the survivor function S(t)[49]	• ORR was summarised by a binomial response rate and corresponding two-sided 95% exact CI using the Clopper-Pearson method. • Time-to-event distributions (DOR, PFS and OS) were estimated using Kaplan-Meier techniques ○ When appropriate, the median along with 95% CI was provided using Brookmeyer and Crowley methodology (using the log-log transformation for construction of CIs). ○ Rates at fixed time points (e.g. OS at 12 months) were derived from the Kaplan Meier estimate along with their corresponding log-log transformed 95% CIs.
Sample size, power calculation	• The primary objective was to estimate ORR as per BIRC assessment for: ○ All treated patients ○ Patients with PD-L1 expression ≥1% ○ Patients with PD-L1 expression ≥5% • For all treated patients, a sample size of 242 would provide 90% power to reject the null hypothesis that ORR was 10% at a two-sided 5% type I error if the true ORR in this population was 16.9%. • Assuming ORR is 30%, 70 treated patients with PD-L1 expression ≥5% would provide 99.1% power at 5% type 1 error to reject the null hypothesis of a two-sided test that	• The primary objective was to estimate investigator-assessed ORR • An ORR of 10% or less was considered not of clinical value, and an ORR of 25% or greater was considered of strong clinical interest • A sample size of 60–100 treated subjects would provide 90% to 97% power to reject the null hypothesis of 10% response rate if the true response rate was 25% with a two-sided Type I error rate of 5%

	the true ORR was 10%, based on historical control data for single-agent chemotherapy,[34, 35, 50] a threshold below which was considered not clinically meaningful in this population, and 90% power at 5% type I error to reject the null hypothesis of a two-sided test that the true ORR was 14.7%. • Under the assumption of 32% prevalence rate of PD-L1 ≥5% among all PD-L1 evaluable patients, approximately up to 220 PD-L1 evaluable patients would be treated. Assuming an additional 10% of treated patients with PD-L1 indeterminate status, the total sample size was expected to be approximately 242. • Under the assumption of 50% prevalence rate of PD-L1 ≥1% among all PD-L1 evaluable patients, approximately up to 110 patients with PD-L1 expression ≥1% would be treated. This would provide 90% power to reject the null hypothesis that ORR was 10% at a two-sided 5% type 1 error if the true ORR in this population was 20.6%.	
Data management, patient withdrawals	• The final analysis of the primary endpoint ORR (based on BIRC assessments) was to be performed six months after approximately 70 patients with PD-L1 expression of ≥5% had been treated (i.e. six months after last patient first treatment)	• All 78 patients who received at least one dose of nivolumab were included in the safety and efficacy analyses

Abbreviations: BOR: best overall response; CI: confidence interval; ORR: overall response rate; PD-L1: programmed death ligand 1; TTR: time to response.
Source: CheckMate 275 CSR[39] and CheckMate 032 CSR.[43]

B.2.4.1 Definitions of study groups

CheckMate 275

A total of 386 patients were initially enrolled in CheckMate 275, of which 270 patients went on to receive at least one dose of nivolumab. Reasons for non-receipt of study drug amongst these 116 enrolled patients are provided in Appendix D. At the primary database lock (30th May 2016), five patients from Japan who were enrolled and first treated after the closure of global enrolment were excluded from the primary efficacy analyses for having less than 6 months of follow-up time, giving rise to an efficacy-treated population of 265 patients. These five patients were included in subsequent efficacy analyses from the second database lock (2nd September 2016).

Definitions of the study populations are presented in Table 9 below. Further details regarding study populations, including the participant flow (CONSORT diagram) and the full eligibility criteria of CheckMate 275 are provided in Appendix D and M.

Table 9: Trial populations used in the primary analysis of CheckMate 275

Analysis	Trial population
All-treated population (n=270)	• Population for baseline demographics and disease characteristics, safety, and dosing evaluation • All patients that received at least one dose of nivolumab
Efficacy-treated population (n=265)	• Population for efficacy analysis • All patients that received at least one dose of nivolumab excluding 5 patients in Japan who started treatment after the last patient first treatment (LPFT) date of patients enrolled before closure of global enrolment.

Abbreviations: LPFT: last patient first treatment.
Source: CheckMate 275 CSR.[39]

CheckMate 032

A total of 86 patients were enrolled in the nivolumab monotherapy treatment arm of CheckMate 032, of whom 78 patients received at least one dose of nivolumab. Reasons for non-receipt of study drug among those eight patients are summarised in Appendix D. All 78 patients who received at least one dose of nivolumab were included in the safety and efficacy analyses.

B.2.4.2 Participant flow

Full details of the participant flow (CONSORT diagrams) for CheckMate 275 and CheckMate 032 can be found in Appendix D.

B.2.5 Quality assessment of the relevant clinical effectiveness evidence

The risk of bias assessments for CheckMate 275 and CheckMate 032 were conducted using the CRD cohort study checklist and is summarised in Table 10.[51] The checklist was adapted to remove the three criteria that referred to comparative studies: *"Were the groups comparable on all important confounding variables?"*, "*Was there adequate adjustment for the effect of these confounding variables?" and "Were drop-out rates and reasons for dropout similar across intervention and unexposed groups?"*

Both CheckMate 275 and CheckMate 032 are considered to be of satisfactory quality based on the CRD cohort study checklist.[51] A summary of the quality assessments is provided below in Table 10, and full details of the quality assessments are reported in Appendix D.

Table 10: Quality assessment of CheckMate 275 and CheckMate 032

Trial	CheckMate 275	CheckMate 032
Is there sufficient description of the groups and the distribution of prognostic factors?	Yes	Yes
Are the groups assembled at a similar point in their disease progression?	Yes	Yes
Are the intervention/treatment reliable ascertained?	Yes	Yes
Was a dose response relationship between intervention and outcome demonstrated?	No	No
Was outcome assessment blind to exposure status?	No	No
Was follow-up long enough for the outcomes to occur?	Yes	Yes

What proportion of the cohort was followed up?	100%	100%

Quality assessment performed using the CRD cohort study checklist.[51]
Source: Sharma *et al.* (2017),[38] CheckMate 275 CSR,[39] CheckMate 275 CRS Addendum,[40] Sharma *et al.* (2016)[37] and CheckMate 032 CSR.[43]

B.2.6 Clinical effectiveness results of the relevant trials

Summary of the clinical effectiveness results of CheckMate 275 and CheckMate 032

- At the primary database lock of CheckMate 275, treatment with nivolumab led to a clinically meaningful confirmed objective response per BIRC (primary efficacy endpoint) in a total of 52 (19.6%) patients (95% CI: 15.0–24.9) and 6 (2.3%) patients achieved a CR (n=270)
    - At a median follow-up of 7 months, median duration of response (DOR) had not yet been reached; 76.9% of responders were continuing in response and nearly all patients ([redacted]) had experienced a DOR of at least 3 months
- Results for the primary efficacy endpoint of CheckMate 032 were consistent with those from CheckMate 275: treatment with nivolumab led to a confirmed investigator-assessed objective response in a total of 19 (24.4%) patients (95% CI 15.3–35.4) (n=78)
- ORR results were consistent across all PD-L1 subgroups in both trials, with clinically meaningful ORRs observed even for patients with low to no PD-L1 expression (PD-L1<1%).
    - In CheckMate 275, patients in the PD-L1≥1% cohort achieved an ORR of 23.8% (95% CI: 16.5–32.3) and patients with <1% PD-L1 expression had a confirmed ORR of 16.1% (15.8% at the second database lock).
- PFS as per BIRC was 2.00 months (95% CI, 1.87–2.63) in CheckMate 275 and 2.78 months (95% CI: 1.45–5.85) in CheckMate 032 (investigator-assessed).
    - In CheckMate 275, patients with PD-L1 expression ≥1% and <1% experienced a PFS period of 3.55 months (95% CI: 1.94–3.71) and 1.87 (95% CI: 1.77–2.04), respectively.
- Median OS in the efficacy-treated population was 8.74 months (95% CI: 6.05–N/A) in CheckMate 275 and 9.72 months (95% CI: 7.26–16.16) in CheckMate 032.
    - In CheckMate 275, 3-month and 6-month OS rates were 75.8% (95% CI: 70.2–80.5) and 57.0% (95% CI: 50.7–62.7).
- Results from the second database lock of CheckMate 275 (2nd September 2016) were consistent with those from the primary analysis database lock in terms of ORR, PFS and OS. In total, 54 patients (20.0%) had achieved an ORR (95% CI: 15.4–25.3), and 2 more patients had achieved a CR. Median DOR was 10.35 months (95% CI: 7.52–NR).
- In CheckMate 275, HRQoL measured via the EORTC QLQ-30 questionnaire demonstrated that nivolumab increased or maintained patient HRQoL from baseline to Week 41, and a meaningful improvement was observed for the dyspnoea, insomnia and financial difficulties domains
- In summary, across both CheckMate 275 and CheckMate 032, nivolumab provided meaningful clinical benefit with a substantial and durable clinical response, irrespective of PD-L1 expression status, for patients with locally advanced unresectable or metastatic UC after failure of prior platinum-containing chemotherapy

B.2.6.1 Overview of the clinical effectiveness results

An overview of the clinical effectiveness results from CheckMate 275 and CheckMate 032 is presented in Table 11. Full results for the primary, secondary and exploratory clinical endpoints are presented in the subsequent sections. Clinical effectiveness results for the PD-L1 <1% and ≥1% subgroups and key baseline characteristics subgroup populations are presented in Appendix E.

Table 11: Overview of clinical effectiveness results from CheckMate 275 and CheckMate 032

Outcome	CheckMate 275		CheckMate 032
	Initial database lock: 30th May 2016 n=265[c]	Latest database lock: 2nd Sep 2016 n=270[c]	n=78
ORR, n (%), [95% CI]	52 (19.6), [15.0–24.9]	54 (20.0), [15.4–25.3][b]	19 (24.4) [15.3–35.4]
TTR, median (IQR), months	1.87 (1.81–1.97)[a]	1.94 (1.84–2.50)[b]	1.48 (1.25–4.14)
DOR, median (95% CI), months	NR (7.43–NR)[a]	10.35 (7.52–NR)[b]	NR (9.92–NR)
PFS, median (95% CI), months	2.00 (1.87–2.63)[a]	2.00 (1.87–2.63)[b]	2.78 (1.45–5.85)
OS, median (95% CI), months	8.74 (6.05–NR)[a]	8.57 (6.05–11.27)[b]	9.72 (7.26–16.16)

[a]Minimum follow-up of 6 months from the date of first dose. [b]Minimum follow-up of 8.3 months. [C]Follow-up for the latest database lock was sufficient to include 5 patients from Japan who were not included in efficacy analyses in the initial database lock.
Abbreviations: CI confidence interval; DOR: duration of response; IQR: interquartile range; PFS: progression-free survival; ORR: objective response rate; OS: overall survival; TTR: time to response; NR: not reached.
Source: Sharma *et al.* (2016),[37] Sharma *et al.*, 2017[38]; CheckMate 275 CSR[39] and CheckMate 275 CSR Addendum (25 October 2016).[40]

B.2.6.2 CheckMate 275

Primary endpoint: objective response rate

Nivolumab demonstrated a clinically meaningful objective response rate in patients with locally advanced unresectable or metastatic urothelial carcinoma

Treatment with nivolumab led to a confirmed objective response per BIRC in a total of 52 (19.6%) patients (95% CI: 15.0–24.9) and 6 (2.3%) patients achieved a CR (Table 12).

ORR results were consistent across all PD-L1 subgroups in both trials, with clinically meaningful ORRs observed even for patients with low to no PD-L1 expression (PD-L1<1%). Patients in the PD-L1≥1% cohort achieved an ORR of 23.8% (95% CI: 16.5–32.3) and patients with <1% PD-L1 expression had a confirmed ORR of 16.1% (15.8% at the second database lock).

These ORR rates of more than 15% achieved across all patients, including those with <1% PD-L1 expression, can be considered clinically meaningful in the context of current therapeutic options for locally advanced unresectable or metastatic UC, where only 10% of patients typically respond to second-line single-agent chemotherapy regimens.[34, 35]

Full results of the PD-L1 <1% and ≥1% subgroup analyses are presented in Appendix E. Results for investigator-assessed ORR were investigated as a secondary outcome and the results were consistent with BIRC-assessed ORR (see Section B.2.6.3).

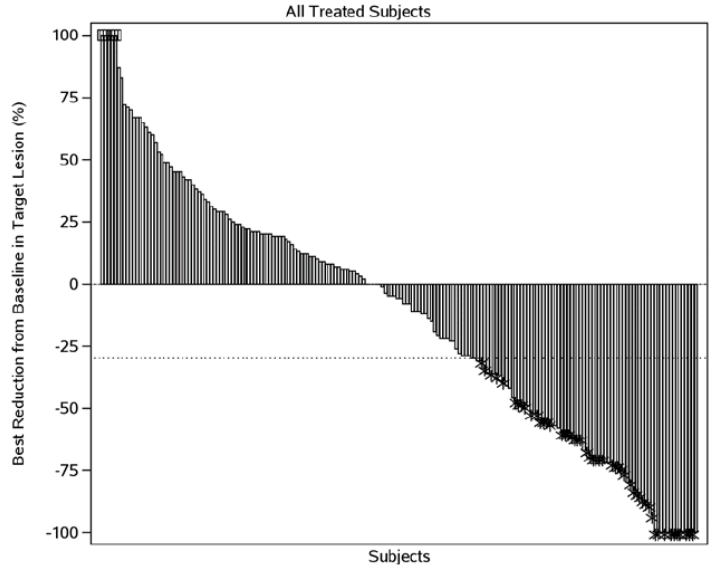
Table 12: Primary efficacy results of CheckMate 275

Tumour response	Efficacy-treated population (n=265)	PD-L1 <1% (n=143)	PD-L1 ≥1% (n=122)
ORR, n (%) **95% CI**	52 (19.6) 95% CI: 15.0–24.9	23 (16.1) 95% CI: 10.5–23.1	29 (23.8) 95% CI: 16.5–32.3
BOR			
CR	6 (2.3)	1 (0.7)	5 (4.1)
PR	46 (17.4)	22 (15.4)	24 (19.7)
SD	60 (22.6)	25 (17.5)	35 (28.7)
PD	104 (39.2)	67 (46.9)	37 (30.3)
Unable to determine[a]	49 (18.5)	28 (19.6)	21 (17.2)
Median TTR (n=52), months; IQR	1.87 IQR: 1.81–1.97	1.94 IQR: 1.81–2.10	1.87 IQR: 1.81–1.97
Median DOR (n=52), months; 95% CI	NR 95% CI: 7.43–NR	NR 95% CI: 7.43–NR	NR 95% CI: 7.52–NR

[a]BOR was reported as unable to determine in 49 patients (18.5%); main reasons were because the patient had died or started subsequent therapy before the first scan visit at Week 8.
Abbreviations: BOR: best overall response; CI: confidence intervals; CR: complete response; DOR: duration of response; IQR: interquartile range; ORR: objective response rate; PD: progressive disease; PR: partial response; SD: stable disease; TTR: time to response NR: not reached.
Source: Sharma *et al.* (2017)[38] and CheckMate 275 CSR.[39]

The magnitude of BIRC-assessed change from baseline in tumour burden for response-evaluable patients is shown in Figure 9. All responders (identified by asterisks in the figure) had a more than 30% reduction in tumour burden consistent with a RECIST v1.1 defined response.

Figure 9: Waterfall plot of best reduction from baseline in sum of diameters of target lesions per BIRC-response evaluable patients



Patients with target lesion at baseline and at least one on-treatment tumour assessment. Negative/positive value means maximum tumour reduction/minimum tumour increase. Best reduction is based on evaluable target lesion measurements up to progression or start of subsequent therapy. Horizontal reference line indicates the 30% reduction consistent with a RECIST v1.1 response.
*Responder per RECIST v1.1 criteria, confirmation of response required. Square symbol represents % change truncated to 100%.
Abbreviations: BIRC: blinded independent review committee: RECIST: Response Evaluation Criteria In Solid Tumours.
Source: CheckMate 275 CSR.[39]

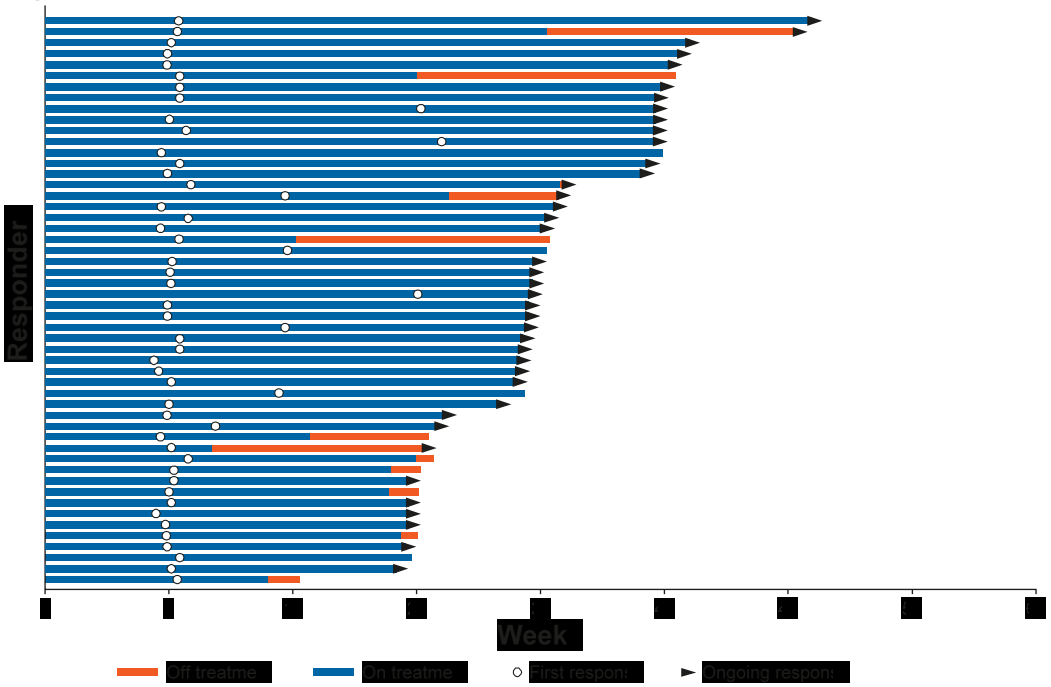
Time to response and duration of response

Objective response rates occurred rapidly and were durable across all PD-L1 cohorts

TTR and DOR were estimated in patients with a confirmed PR or CR. Median TTR as per BIRC was 1.87 months (IQR: 1.81–1.97 months) and the majority of responders achieved their response at the time of first tumour assessment (Week 8).

At the time of the clinical database lock (30th May 2016), median DOR as per BIRC had not been reached in the efficacy-treated population and across the <1% and ≥1% PD-L1 subgroups. The majority of responders (76.9%) were still continuing to respond and almost all patients ([redacted]) had a DOR of at least 3 months (see Figure 10). In a small number of responders, an ongoing response was also seen to continue beyond treatment discontinuation (Figure 10).

Figure 10: Time to and duration of response of responders in CheckMate 275



Bars indicate progression-free survival.
Source: Galsky *et al.* (2016).[41]

B.2.6.3 Secondary efficacy results of CheckMate 275

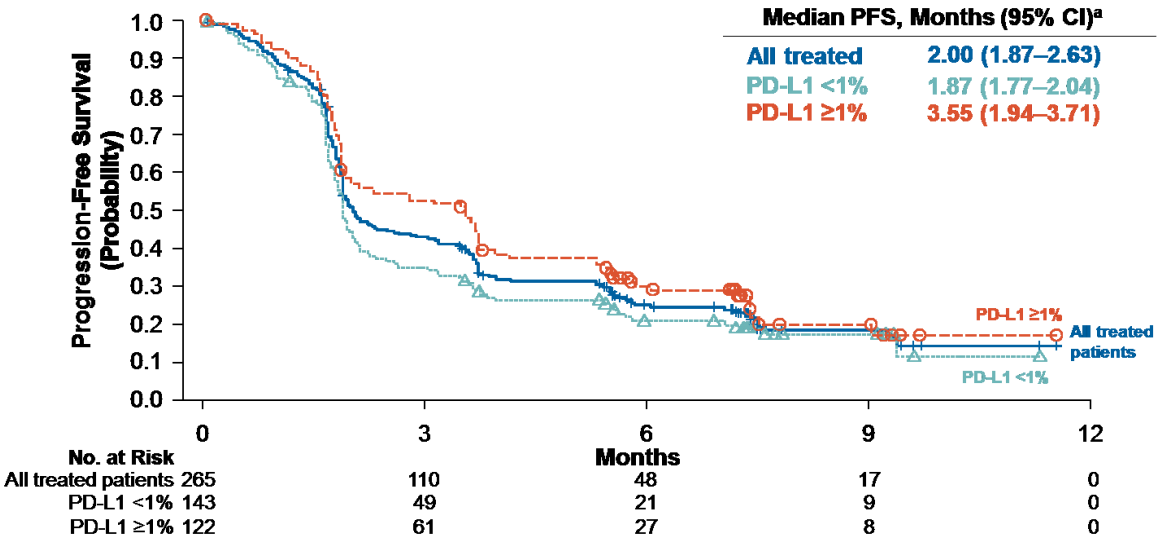
Progression-free survival

At the time of the primary clinical database lock (30th May 2016), 201 patients (75.8%) had experienced a PFS event. Median PFS in the efficacy-treated population was 2.00 months (95% CI: 1.87–2.63), and the PFS rates at 3 and 6 months were 43.1% (95% CI: 37.0–49.1) and 25.2% (95% CI: 20.0–30.8), respectively (Figure 11).

PFS was consistent irrespective of baseline PD-L1 status; median PFS for patients in the PD-L1 ≥1% cohort was slightly longer than in the all-treated population at 3.55 months (95% CI: 1.94–3.71), and in the PD-L1 <1%, median PFS was 1.87 months (95% CI: 1.77–2.04). The Kaplan-Meier plot for PFS is presented in Figure 11.

PFS results based on investigator assessment were evaluated as an exploratory endpoint and were consistent with those based on BIRC assessment. These results are provided in Appendix M.

Figure 11: Kaplan-Meier plot for progression-free survival in CheckMate 275



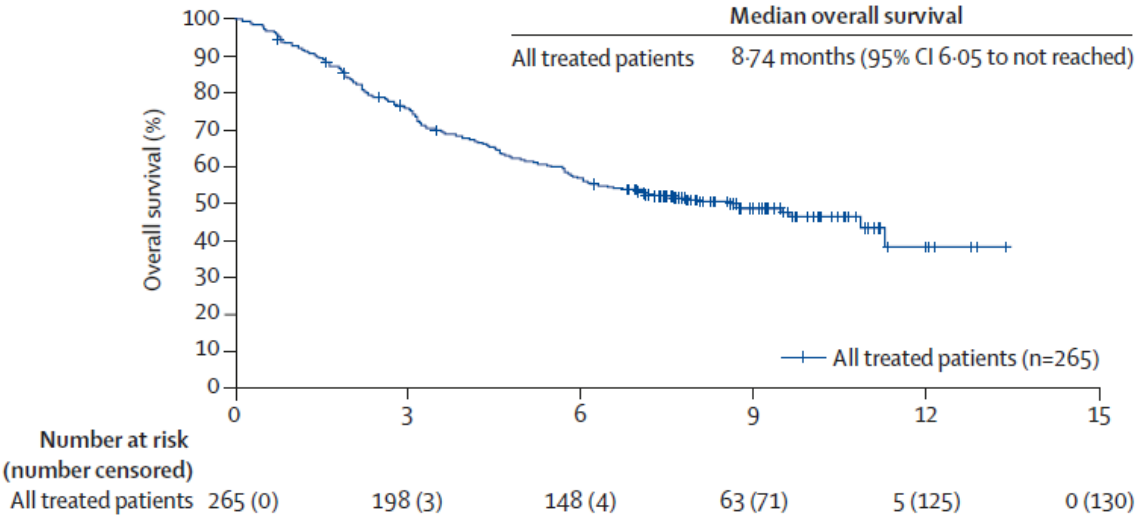
Abbreviations: CI: confidence interval; PD-L1: programmed death ligand 1; PFS: progression-free survival.
Source: Galsky *et al.* (2016).[41]

Overall survival

Median follow-up for OS (time between first dose and last known date alive or death) was 7.00 months (IQR: 2.96–8.77 months). At the primary analysis database lock (30th May 2016), 138 patients (51.1%) had died. Median OS in the efficacy-treated population was 8.74 months (95% CI: 6.05–N/A); 3-month and 6-month OS rates were 75.8% (95% CI: 70.2–80.5) and 57.0% (95% CI: 50.7–62.7).

The Kaplan-Meier plot for OS is presented in Figure 12. Results of the PD-L1 <1% and ≥1% subgroup analyses are presented in full in Appendix E.

Figure 12: Kaplan-Meier plot for overall survival in CheckMate 275



Abbreviations: CI: confidence interval; OS: overall survival.
Source: Sharma *et al.* (2017).[38]

Treatment beyond progression

As of the primary clinical database lock of CheckMate 275 (30th May 2016), a total of 70 patients (26.4%) received at least one dose of nivolumab after initial RECIST v1.1-defined progression. Treatment beyond progression was defined as a last dosing date after a RECIST v1.1 progression date.

Of the 70 patients treated beyond progression, 24 were considered non-conventional responders, defined as patients who had not experienced a BOR of PR/CR prior to initial RECIST v1.1-defined progression, and met at least 1 of the following criteria:

- Criterion 1: Appearance of a new lesion followed by decrease from baseline of at least 10% in the sum of the target lesions (15 patients)
- Criterion 2: Initial increase from nadir ≥20% in the sum of the target lesions followed by reduction from baseline of at least 30% (2 patients)
- Criterion 3: Initial increase from nadir ≥20% in the sum of the target lesions followed by at least 2 tumour assessments showing no further progression defined as a 10% additional increase in sum of target lesions and new lesions (3 patients)
- Criteria 1 and 2 (1 patient)
- Criteria 1 and 3 (3 patients).

The kinetics of tumour burden change over time for patients treated beyond initial RECIST v1.1-defined progression are presented as a subgroup analysis in the PD-L1 <1% and PD-L1 ≥1% cohorts in Appendix E.

Objective response rate as per investigator assessment

ORR as per investigator assessment was investigated as a secondary outcome in CheckMate 275. Rates of confirmed objective response were similar to those reported for BIRC-assessed ORRs and are presented in Table 13. A total of [redacted] patients ([redacted]) achieved an objective response of which [redacted] patients ([redacted]%) achieved a CR.

Table 13: Investigator-assessed ORR in CheckMate 275

Tumour response	Efficacy-treated population (n=265)
ORR, n (%) **95% CI**	[redacted]
BOR	
CR	[redacted]
PR	[redacted]
SD	[redacted]
PD	[redacted]
Unable to determine[a]	[redacted]

[a]BOR was reported as unable to determine in 39 patients (14.7%) due to death prior to disease assessment or early discontinuation due to toxicity.
Abbreviations: BOR: best overall response; CI: confidence intervals; CR: complete response; ORR: objective response rate; PD: progressive disease; PR: partial response; SD: stable disease; NR: not reached.
Source: CheckMate 275 CSR.[43]

B.2.6.4 Latest efficacy results of CheckMate 275: database lock (2nd September 2016)

Objective response rate (latest database lock)

At the time of the latest database lock of CheckMate 275 (2nd September 2016), the ORR was 54 (20.0%; 95% CI: 15.4, 25.3). A CR had been achieved by 8 (3.0%) and a PR by 46 (17.0%) patients. Median TTR was 1.94 and median DOR was 10.35 months. Detailed ORR results are presented in Table 14.

Table 14: Latest database lock efficacy results of CheckMate 275

Tumour response	All-treated population (n=270)	PD-L1 <1% (n=146)	PD-L1 ≥1% (n=124)
ORR, n (%) **95% CI**	54 (20.0) 95% CI: 15.4–25.3	23 (15.8) 95% CI: 10.3–22.7)	31 (25.0) 95% CI: 17.7–33.6
CR	8 (3.0)	[redacted]	[redacted]
PR	46 (17.0)	[redacted]	[redacted]
SD	60 (22.2)	[redacted]	[redacted]
PD	[redacted]	[redacted]	[redacted]
Unable to determine[a]	[redacted]	[redacted]	[redacted]
Median TTR (n=54), months **IQR**	1.94 IQR: 1.84–2.50	1.97 IQR: 1.87–3.48	[redacted]
Median DOR (n=54), months **95% CI**	10.35 95% CI: 7.52–NR	10.35 95% CI: 7.43–NR	NR 95% CI: 7.52–NR

[a]BOR was reported as unable to determine in 51 patients (18.5%); main reason was death prior to assessment.
Abbreviations: BOR: best overall response; CI: confidence intervals; CR: complete response; DOR: duration of response; IQR: interquartile range; ORR: objective response rate; PD: progressive disease; PR: partial response; SD: stable disease; TTR: time to response NR: not reached.
Source: CheckMate 275 CSR Addendum (25 October 2016).[40]

Progression-free survival (latest database lock)

At the time of the latest clinical database lock (2nd September 2016), an additional 5 PFS events (1.9%) had occurred since the initial database lock. Median PFS in the all-treated population remained unchanged at 2.00 months (95% CI: 1.87–2.63), and the PFS rates at 3 and 6 months were also relatively similar, at [redacted] and 26.1% (95% CI: 20.9–31.5) respectively. PFS rates at 9 and 12 months were also reported in the latest database lock, as [redacted] and [redacted], respectively.

Patients in the <1% and ≥1% PD-L1 cohorts experienced a similar duration of PFS as in the primary analysis, with a median PFS of 1.87 months (95% CI: 1.77–2.04) and 3.55 months (95% CI: 1.94–3.71), respectively.

The Kaplan-Meier plot for PFS in the latest database lock for the all-treated population and the PD-L1 <1% and PD-L1 ≥1% subgroups is presented in Figure 13.

Figure 13: Kaplan-Meier plot for progression-free survival in CheckMate 275 (latest database lock)

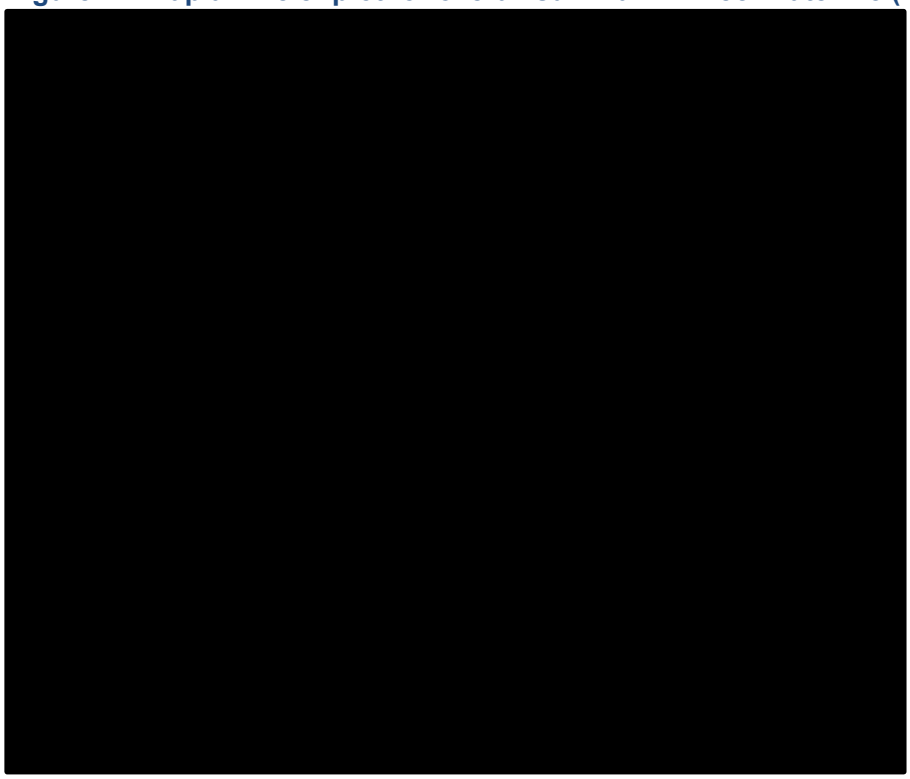
Overall survival (latest database lock)

Median follow-up time for OS was 11.5 months (range: 8.3–15.7 months), about 3 months longer than the initial database lock. At the latest database lock, 154 patients (57%) had died, which is an additional 16 deaths since the initial database lock. Median OS was 8.57 months (95% CI: 6.05–11.27) in all-treated patients, 11.63 months (95% CI: 9.10–NR) in patients with PD-L1 ≥1%, and [redacted] months (95% CI: [redacted]) in patients with PD-L1 ≥5%.

OS rates were [redacted] at 3 months, [redacted] at 6 months, [redacted] at 9 months and 41.0% (95% CI: 34.8–47.1) at 12 months.

The Kaplan-Meier plot for OS in the latest database lock for the all-treated population and the PD-L1 <1% and PD-L1 ≥1% subgroups is presented in Figure 14.

Figure 14: Kaplan-Meier plot for overall survival in CheckMate 275 (latest database lock)



B.2.6.5 CheckMate 032

The following section presents the clinical efficacy results from CheckMate 032 (primary database lock: 24th March 2016).

Objective response rate

An overview of the primary efficacy results from the UC cohort of CheckMate 032 is presented in Table 15. A confirmed investigator-assessed objective response was achieved in 19 (24.4%) patients (95% CI: 15.3–35.4) of 78 treated patients, with five patients (6%) achieving a CR and 14 patients (18%) achieving a PR.

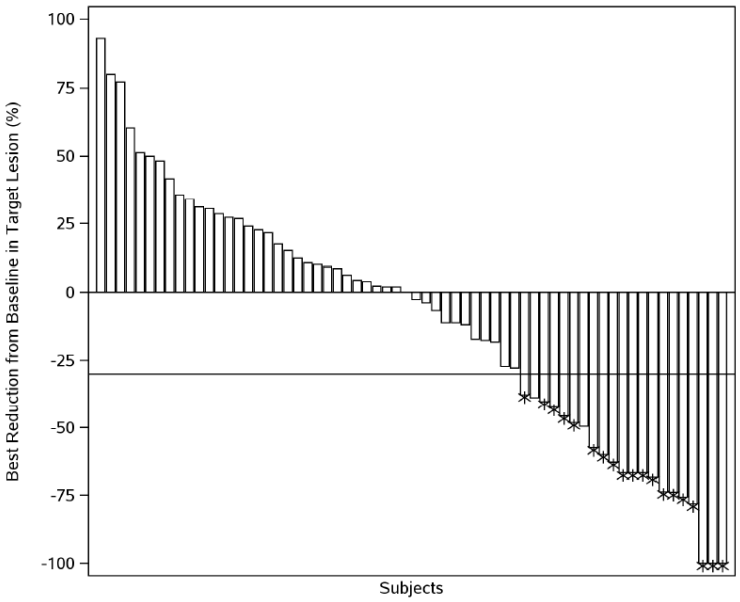
Table 15: Overview of clinical effectiveness results from CheckMate 032

Tumour response	Nivolumab (n=78)
ORR, n (%)	19 (24.4) [95% CI 15.3–35.4]
BOR, n (%)	
CR	5 (6.4)
PR	14 (17.9)
SD	22 (28.2)
PD	30 (38.5)
Unable to determine	7 (9.0)
Median TTR, months (IQR)	1.48 (1.25–4.14)
Median DOR, months (95% CI)	NR (9.92–NR)

Abbreviations: BOR: best overall response; CI: confidence intervals; CR: complete response; DOR: duration of response; IQR: interquartile range; ORR: objective response rate; PD: progressive disease; PR: partial response; SD: stable disease; TTR: time to response NR: not reached.
Source: Sharma *et al.* (2016)[37 37 38 33 33 3434] and CheckMate 032 CSR.[43]

The magnitude of best change in tumour burden in target lesions relative to baseline is shown in Figure 15. All responders had a >30% reduction in tumour burden consistent with a RECIST v1.1-defined response.

Figure 15: Waterfall plot of best change in target lesion per investigator-assessed objective response

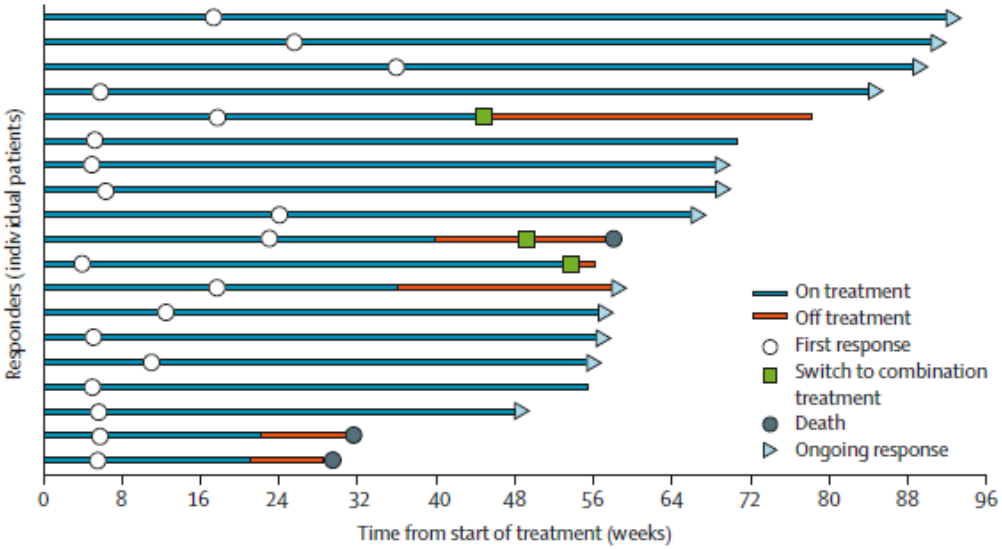


Patients with target lesion at baseline and at least one on-treatment tumour assessment. Negative/positive value means maximum tumour reduction/minimum tumour increase. Horizontal reference line indicates the 30% reduction consistent with a RECIST 1.1 response. Best reduction is based on evaluable target lesion measurements up to progression or start of subsequent therapy.
Asterisk indicates responders. Crossover patients are truncated at crossover date. Symbol square represents % change truncated to 100%.
Source: CheckMate 032 CSR.[43]

Time to response and duration of response

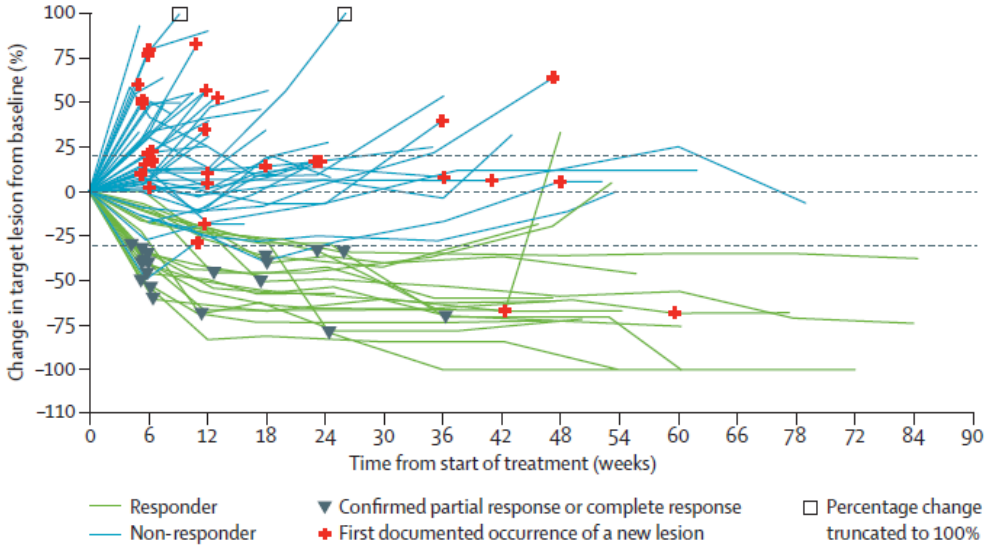
The median TTR was 1.48 months with the majority of responders achieving their response at the time of first tumour assessment (week 6). At the time of the clinical database lock, median DOR was not reached and the majority of responders ([redacted]) were still continuing in response. Most responders ([redacted]) had a DOR of at least 6 months, and [redacted] had a DOR of at least 12 months.

Figure 16: Time to and duration of response in CheckMate 032



Source: Sharma et al. (2016).[37]

Figure 17: Best change in target lesion observed in CheckMate 032



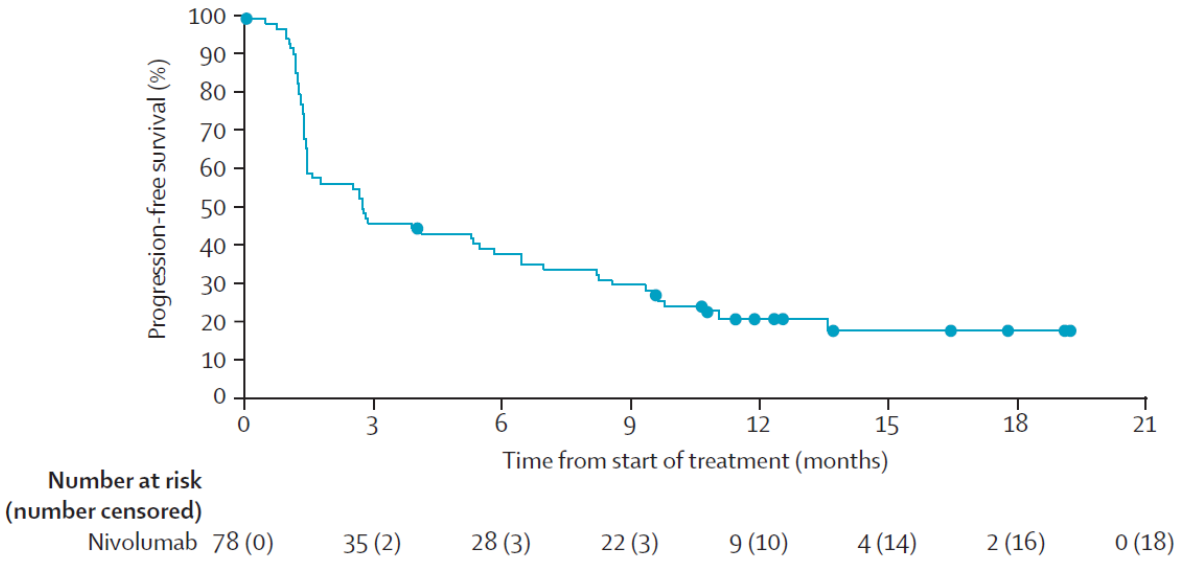
Source: Sharma et al. (2016).[37]

Progression-free survival

The Kaplan-Meier plots for PFS and OS in CheckMate 032 are presented in Figure 18 and Figure 19.

Median PFS was 2.78 months (95% CI 1.45–5.85) and 60 (77%) of 78 patients had disease progression or died by data cut-off. Of 18 (23.1%) censored patients, [redacted] had their PFS time censored on either the date of last on-study tumour assessment or date of last assessment prior to subsequent anti-cancer therapy. The most common reason for censoring among these patients was [redacted]. PFS rates (95% CI) were [redacted] at 3 months, [redacted] at 6 months and 20.8% (12.3–30.9) at 12 months.

Figure 18: Kaplan-Meier plot for progression-free survival in CheckMate 032

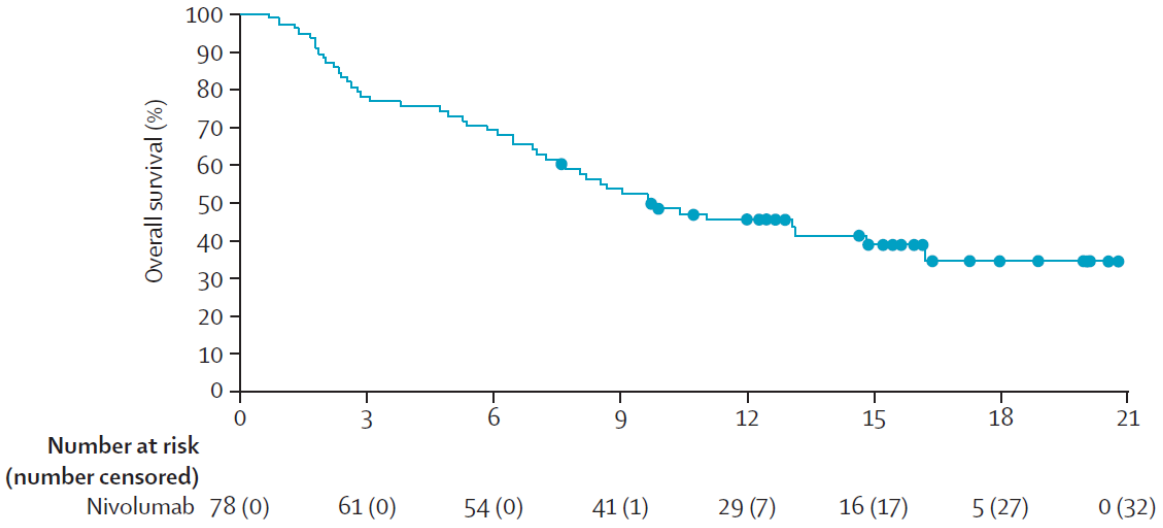


Source: Sharma *et al.* (2016).[37]

Overall survival

Median OS was 9.7 months (95% CI 7.3–16.2) and 46 (59%) of 78 patients had died at the time of data cut-off. OS rates (95% CI) were [redacted] at 3 months, [redacted] at 6 months, and 45.6% (34.2–56.3) at 12 months. Median follow-up for OS (time between dose date and last known date alive or death) for all nivolumab monotherapy treated UC patients was 9.69 months (range: 0.7–20.7 months).

Figure 19: Kaplan-Meier plot for overall survival in CheckMate 032



Source: Sharma *et al.* (2016).[37]

B.2.6.6 Patient-reported outcomes

Patient-reported outcomes data for the measurement of HRQoL was assessed via the EORTC QLQ-C30 questionnaire in CheckMate 275, and the EQ-5D-3L questionnaire, collected in both CheckMate 275 and CheckMate 032.

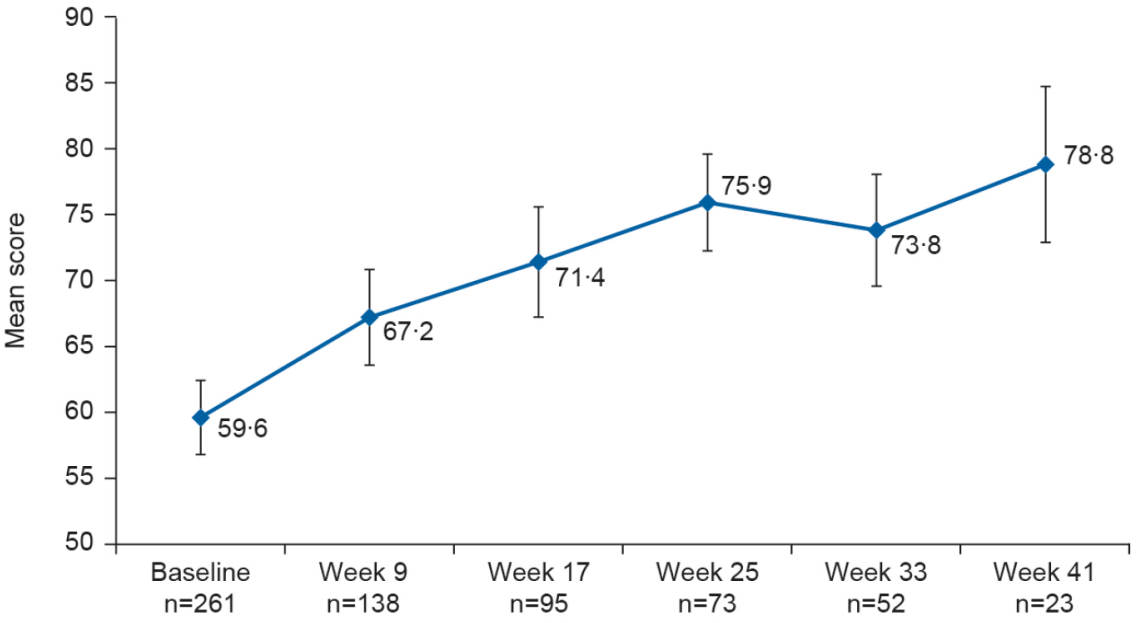
CheckMate 275: EORTC QLQ-30

The EORTC QLQ-C30 is the most commonly used quality-of-life instrument in oncology trials that includes five functional scales (physical, role, cognitive, emotional, and social), nine symptom scales (fatigue, pain, nausea/vomiting, dyspnoea, insomnia, appetite loss, constipation, diarrhoea, and financial difficulties), and a global health/quality of life scale.[52]

A total of 262 patients (97.0%) in the all-treated population completed the EORTC QLQ-30 at baseline. Calculated as a percentage of patients on study, completion rates for treated patients met or exceeded 75% at all assessments through the first 49 weeks of on-treatment visits, after which no patients were eligible for on-treatment patient-reported outcomes assessment.

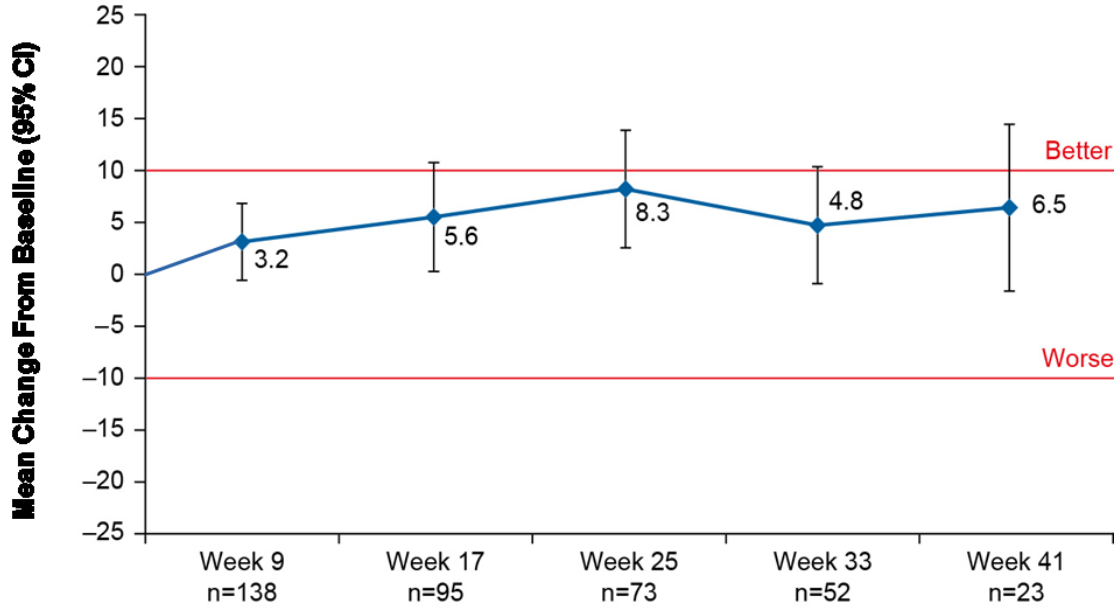
Due to the limited study follow-up, interpretations of EORTC QLQ-30 results are limited to the first 41 weeks of follow-up for the all-treated population. Overall, patient HRQoL continued to increase or was maintained throughout the trial from baseline to Week 41 (see Figure 20 and Figure 21).

Figure 20. Mean score in EORTC QLQ-30 global health status in CheckMate 275



Abbreviations: EORTC QLQ-30: European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30.
Source: Sharma *et al.* (2017).[38]

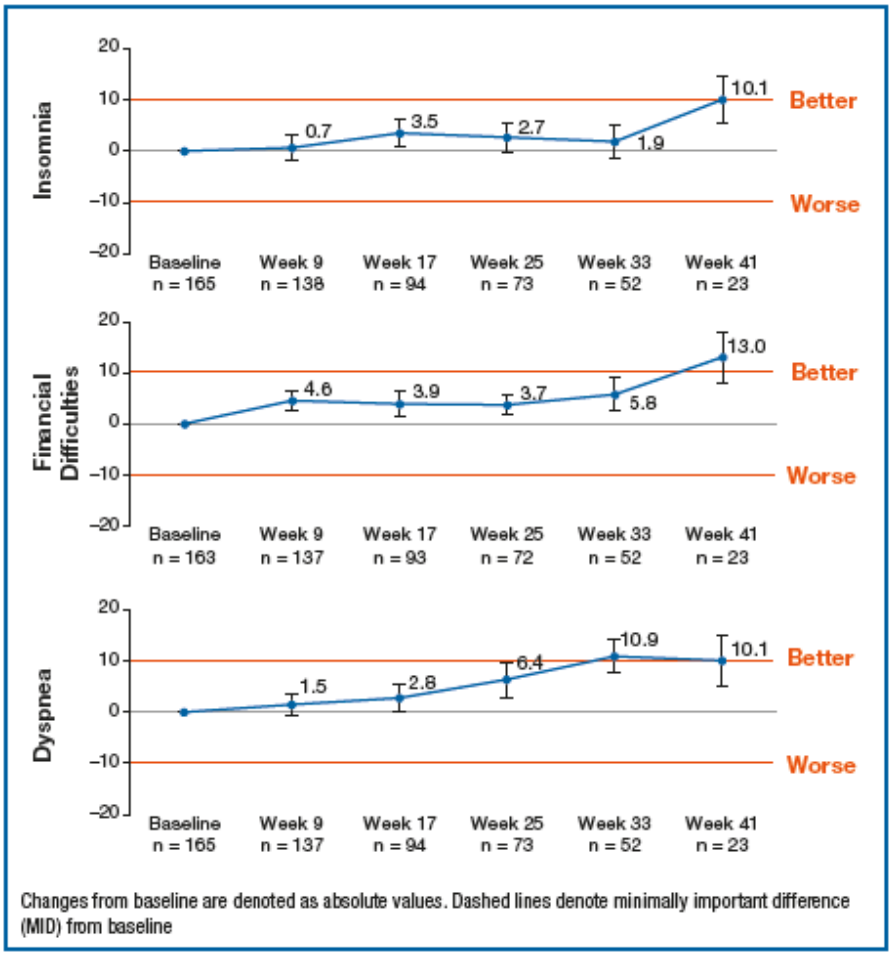
Figure 21: Mean change from baseline in EORTC QLQ-30 global health status score in CheckMate 275



Abbreviations: CI: confidence interval; EORTC QLQ-30: European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30.
Source: Sharma *et al.* (2017).[38]

A meaningful improvement (defined as a ≥10-point increase from baseline score)[53] was experienced for the following domains: dyspnoea (mean change: 10.9 points) at Week 33 with continued improvement through to Week 41; insomnia (mean change: 10.1 points) at Week 41; and financial difficulties (mean change: 13.0 points) at Week 41 (see Figure 22)

Figure 22: EORTC QLQ-C30 mean score change from baseline for insomnia, financial difficulties and dyspnoea



Abbreviations: EORTC QLQ-30: European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30.
Source: Necchi *et al.* (2017).[54]

CheckMate 275: EQ-5D-3L

The EQ-5D-3L is a generic multi-attribute health-state classification system by which health is described in 5 dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression, over 3 levels: no problems, some problems, and severe problems.[55] In addition, the EQ-5D includes a VAS, allowing patients to rate their health on a scale from 0–100, with a clinically meaningful change in EQ-5D VAS score regarded as 7 points.[56]

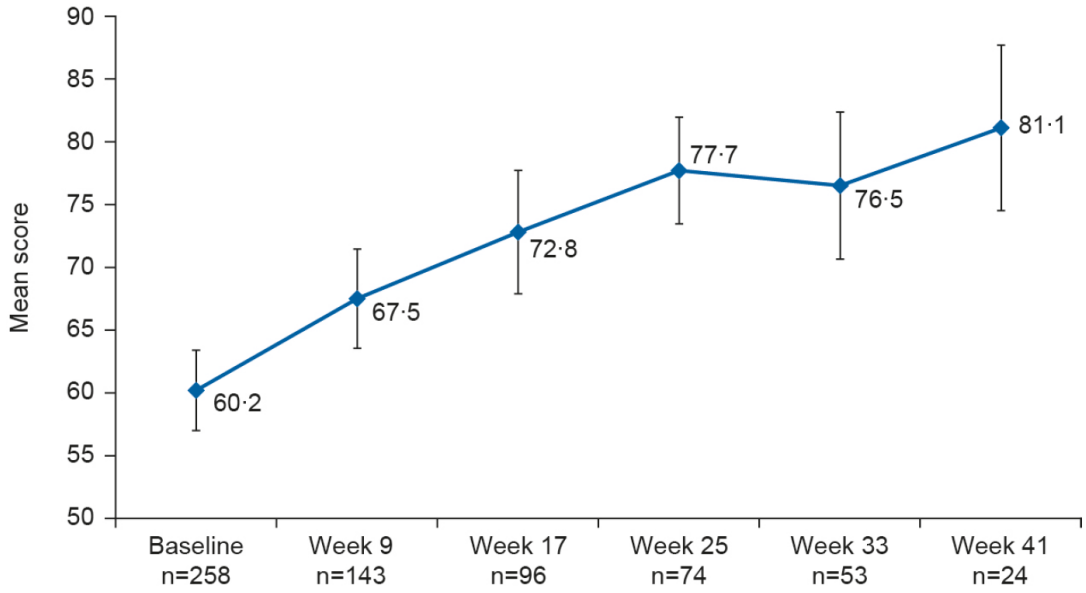
Baseline completion rates for the 5 items included in the EQ-5D descriptive system ranged from 95.9% (mobility) to 96.7% (self-care, usual activities, and anxiety/depression), while 95.6% (258/270) of treated subjects completed the EQ-5D VAS.

During post baseline follow-up, the percentage of patients reporting health problems decreased by 10% for all dimensions of the EQ-5D: mobility at Week 9, self-care at Week 33, usual activities at Week 17, pain/discomfort at Week 9, and anxiety/depression at Week 17. The proportion of patients reporting no health problems continued to increase or remain stable from baseline through to Week 41 of treatment for all dimensions.

The mean baseline EQ-5D VAS score was 60.2 (Figure 23), and mean scores were higher at Week 9 on treatment (67.5). By Week 41, the average EQ-5D VAS was more than 80 points,

which is in alignment with that of the US general population (the country with the largest representation in the study).[57]

Figure 23. Mean EQ-5D-3L score in all-treated population in CheckMate 275



Abbreviations: EQ-5D: EuroQoL-5 dimensions 3-levels questionnaire.
Source: Sharma *et al*. (2017).[38]

CheckMate 275: Updated EQ-5D-3L data (database lock 2nd September 2016)

With the additional 3 months of follow-up for the latest database lock, descriptive interpretations of results are provided to the first 49 weeks of follow-up for treated subjects. Four EORTC QLQ-C30 scales showed new improvements versus baseline compared with the initial database lock (social functioning, global health status, appetite loss, and pain). Results of EQ-5D or EQ-5D VAS were consistent with the those obtained at the initial database lock; the proportion of subjects reporting no health problems continued to increase or remain stable through Week 49 of treatment for all dimensions and the mean scores remained higher at Week 9 on treatment (67.5).

CheckMate 032: EQ-5D-3L

HRQoL data were collected via the EQ-5D-3L questionnaire in CheckMate 032. At baseline, 76 (97.4%) patients completed the questionnaire for each of the 5 items included in the EQ-5D-3L descriptive system. Improvement of ≥10% from baseline was reported for pain/discomfort and anxiety/depression at Week 5; and for mobility and usual activities at Week 19. The proportion of patients with health problems continued to decrease over time for these 4 dimensions. 72 (94.7%) patients reported no problems (Level 1) at baseline for the remaining health dimension, self-care, and remained stable over time.

A total of 73 (93.5%) UC patients treated completed the EQ-5D VAS questionnaire at baseline and the mean baseline EQ-5D VAS score was 72.4 (SD 24.5). Overall, the mean EQ-5D VAS score increased over time. By Week 19, clinically meaningful improvements (>7-point change from baseline) were reported and the average EQ-5D VAS score was >80 points. The EQ-5D VAS continued to improve through Week 61. After week 61, the sample size was too small to interpret (<10).

B.2.7 Subgroup analysis

CheckMate 275

Pre-planned subgroup analyses were performed for patients with baseline PD-L1 status <1% and ≥1%; results of these analyses for the primary outcome ORR as per BIRC, and the secondary outcomes PFS and OS are presented in Appendix E.

Subgroup analyses were also conducted to assess the impact of several key baseline patient characteristics including age, gender, race, baseline ECOG-PS, baseline metastases (liver, visceral, lymph nodes only), baseline haemoglobin, site of original tumour origin (bladder, renal pelvis/ureter), and prior cancer therapy regimens (number of prior regimens in a metastatic setting, time from completion of most recent prior regimen to study treatment) on the primary endpoint of confirmed ORR as per BIRC. The ORRs were consistent across the vast majority of these predefined subgroups and are presented in Appendix E.

CheckMate 032

As part of the exploratory endpoints, ORR, OS and PFS were analysed in subgroups defined by PD-L1 expression (<1% and ≥1%). Results of these analyses are presented in Appendix E.

In addition, ad-hoc subgroup analyses were conducted to assess the impact several key baseline factors such as ECOG-PS, metastases, or haemoglobin on investigator-assessed ORR. Details of these analyses are presented in Appendix E.

Summary of the clinical effectiveness evidence from CheckMate 275 and CheckMate 032

- Across both CheckMate 275 and CheckMate 032, a total of 348 patients with locally advanced unresectable or metastatic UC whose disease had progressed or recurred after treatment with at least one platinum-containing chemotherapy regimen were treated with IV nivolumab 3 mg/kg every 2 weeks until disease progression or unacceptable toxicity
- In CheckMate 275, treatment with nivolumab led to a clinically meaningful confirmed objective response per BIRC (primary efficacy endpoint) in a total of 52 (19.6%) patients (95% CI: 15.0–24.9) and 6 (2.3%) patients achieved a CR. Results for the primary efficacy endpoint of CheckMate 032 were consistent with those from CheckMate 275: treatment with nivolumab led to a confirmed investigator-assessed objective response in a total of 19 (24.4%) patients (95% CI 15.3–35.4).
    - ORR results were consistent across all PD-L1 subgroups in both trials, even for patients with low to no PD-L1 expression
    - PFS as per BIRC was 2.0 months (95% CI, 1.87–2.63) in CheckMate 275 and 2.78 months (95% CI: 1.45–5.85) in CheckMate 032 (investigator-assessed); median OS in the efficacy-treated population was 8.74 months (95% CI: 6.05–N/A) in CheckMate 275 and 9.72 months (95% CI: 7.26–16.16) in CheckMate 032.
- Nivolumab provided meaningful clinical benefit with a substantial and durable clinical response, irrespective of PD-L1 expression status, for patients with locally advanced unresectable or metastatic UC after failure of prior platinum-containing chemotherapy

B.2.8 Meta-analysis

Data from CheckMate 275 and CheckMate 032 were pooled in the context of the ITC presented in Section B.2.9 and Appendix D.

B.2.9 Indirect and mixed treatment comparisons

Summary of the indirect treatment comparison

- No RCTs directly comparing the efficacy and safety of nivolumab in the patient population of interest versus any of the comparators relevant to this submission or placebo were identified in the SLR
- As such, the feasibility of conducting an ITC was assessed between the two nivolumab trials and the comparator trials identified from the SLR. Eligible trials were identified for paclitaxel, docetaxel and BSC; no relevant trials were identified for retreatment with first-line platinum-based chemotherapy
- No direct or indirect links between the nivolumab and comparator trials were identified hence a population-adjusted approach (simulated treatment comparison [STC]) was conducted using individual patient level data from the nivolumab trials and summary data from the comparator trials, to estimate how patients in each of the comparator trials would have responded to nivolumab
- OS and PFS were evaluated using a fractional polynomial approach as exploratory analyses indicated that the proportional hazards assumption was not appropriate for comparisons between nivolumab and its comparators; ORR was evaluated using an NMA model for binomial outcomes
- Time-varying hazard ratios (HR) for PFS and OS and odd ratios (OR) for ORR were then estimated for nivolumab versus each of the relevant comparators with available data
- The results for OS demonstrated that the HR for death was greater than 1 (favouring nivolumab) at the majority of time points through to week 96 for paclitaxel, docetaxel and BSC
- The results for ORR suggest that patients who receive nivolumab have a higher odds of response than patients who receive BSC or docetaxel
- A comparison versus cisplatin plus gemcitabine was conducted as a scenario analysis only due to the non-generalisability of the trial data for this data to UK clinical practice: all patients were gemcitabine-naïve and the dosing regimens used in the trials did not match those used in UK clinical practice

B.2.9.1 Methodology

The SLR identified no RCTs directly comparing the efficacy and safety of nivolumab in the patient population of interest versus any of the comparators relevant to this submission or placebo.

Indirect treatment comparison (ITC) was used to evaluate the relative efficacy of nivolumab and its comparators with respect to OS, PFS and ORR. This section provides a summary of the available data and the results. Appendix D provides full details of the methodology and additional information about the results.

Only trials identified for comparators listed in the NICE final scope and considered relevant to this submission were taken forward for consideration for inclusion within the ITC. These included 18 publications reporting on 12 unique trials (see list of 12 trials in Appendix D).

Eligible trials were identified for paclitaxel, docetaxel and BSC. Three trials (Kim *et al.* [2016],[58] McCaffrey *et al.* [1997][59] and Vaughn *et al.* [2002][60]) were excluded from the ITC because they investigated doses and/or treatments that did not correlate with current UK clinical practice.[29] Appendix D provides further details on these trials.

No relevant trials were identified for retreatment with first-line platinum-based chemotherapy. Two trials were identified for cisplatin plus gemcitabine (Gondo *et al.* [2011] and Ozawa *et al.* [2007]).[61, 62] However, these trials were limited in their generalisability to the decision problem; all patients in Gondo *et al.* (2011)[61] had received MVAC in first-line treatment and are therefore not considered to be directly comparable to those receiving cisplatin plus gemcitabine retreatment in current UK clinical practice, as they are gemcitabine naïve.[29] The Ozawa *et al.* (2007)[62] trial included chemotherapy-naïve patients in addition to patients who had previously undergone first-line treatment. Although outcome data are reported separately for these two populations, patient baseline characteristic data are reported for the two populations combined. Therefore, it is not possible to determine baseline characteristics for patients who had only received first-line treatment, precluding a comparison with patients in other studies included in this analysis. Additionally, the two trials did not use the standard dosing regimen typically used for cisplatin plus gemcitabine in the UK. As such, these trials are non-generalisable to UK clinical practice where cisplatin plus gemcitabine is the standard of care in the first-line setting for locally advanced unresectable or metastatic UC and therefore cannot be considered to provide relevant data for retreatment with first-line platinum-based chemotherapy. Furthermore, the study by Gondo *et al.* (2011) provided no PFS data, and the study by Ozawa *et al.* (2007) provided neither OS not PFS data. As the only identified evidence for cisplatin plus gemcitabine, these trials were taken forwards for the ITC, but the comparison between nivolumab and cisplatin plus gemcitabine was conducted for the purposes of a scenario analysis only, and the results versus this comparator should be treated with caution.

Table 16 provides a summary of the nine trials included in the ITC and identifies which of the outcomes of interest were available for each trial.

Table 16: Summary of the trials used to carry out the ITC

	Outcomes			Interventions				
References of trial	**OS**	**PFS**	**ORR**	**Nivolumab**	**Paclitaxel**	**Docetaxel**	**BSC**	**Cisplatin plus gemcitabine**
CheckMate 032[37]	Yes	Yes	Yes	Yes				
CheckMate 275[38]	Yes	Yes	Yes	Yes				
Bellmunt *et al.* (2009)[33]	Yes		Yes				Yes	
Choueiri *et al.* (2012)[31]	Yes	Yes	Yes			Yes		
Gondo *et al.* (2011)[61]	Yes		Yes					Yes
Joly *et al.* (2009)[63]			Yes		Yes			

Jones *et al.* (2017) [32]	Yes	Yes			Yes			
Ozawa *et al.* (2007)[62]			Yes					Yes
Petrylak *et al.* (2016)[30]	Yes	Yes	Yes			Yes		

Abbreviations: BSC: best supportive care; ORR: objective response rate; OS: overall survival; PFS: progression-free survival.

An overall network diagram is illustrated in Figure 24. The diagram indicates that the network is disconnected; there are no direct or indirect links between nivolumab and any of the comparators. Hence, it was necessary to use a population-adjusted method (simulated treatment comparison [STC]) to conduct the ITC.[64] The STC used individual patient data from the two nivolumab trials, CheckMate 275 and CheckMate 032,[37, 38] along with baseline characteristics from the comparator trials, to estimate how patients in each of the comparator trials would have responded to nivolumab. Further details of this method are provided in Appendix D. Using this approach allowed the generation of pseudo-trials that include real data for the comparator and simulated data for nivolumab. NMA could then be used to establish a network (with a shared comparator of nivolumab) of these pseudo-trials in order to generate relative effectiveness estimates across all treatments.

Figure 24: Network diagram

Dashed lines indicate where simulated treatment comparison has been applied. The trials informing the comparison versus cisplatin plus gemcitabine were conducted in gemcitabine-naïve populations, and therefore cannot be considered to provide relevant data for the retreatment with first-line platinum-based chemotherapy. This comparison is included as a scenario analysis only and results should be interpreted with caution.
Abbreviations: BSC: best supportive care.

B.2.9.2 Overall survival

OS was evaluated using a fractional polynomial network meta-analysis (NMA) approach.[65] This modelling approach was selected because exploratory analyses indicated that the proportional hazards assumption was not appropriate for comparisons between nivolumab and its comparators. This decision was supported by clinical advisory board input – proportional hazards are not expected to hold because of the different mechanisms of action of the treatments.[29] The fractional polynomial NMA approach estimates hazard ratios (HRs) over time for each pairwise treatment comparison (standard NMA models for survival estimate fixed HRs for each comparison). The network diagram for OS is provided in Figure 25 below.

Figure 25: Network diagram for overall survival

Dashed lines indicate where simulated treatment comparison has been applied. The trials informing the comparison versus cisplatin plus gemcitabine were conducted in gemcitabine-naïve populations, and therefore cannot be considered to provide relevant data for the retreatment with first-line platinum-based chemotherapy. This comparison is included as a scenario analysis only and results should be interpreted with caution.
Abbreviations: BSC: best supportive care.

Both first order and second order fractional polynomial models, and fixed and random effects models, were evaluated. Table 17 lists the models that were evaluated and provides a summary of the model fit statistics. The table indicates that the three best fitting models had very similar deviance information criteria (DIC) (the second order (P1=1, P2=1) fixed and random effects models and the second order (P1=0, P2=0) fixed effect model). The DICs for these models are all within 0.3 of each other. DIC differences of less than 3 are generally regarded as unimportant.[66]

As such, the second order (P1=0, P2=0) fixed effect model was used as the base case in the cost-effectiveness model because it provided the most clinically plausible extrapolations out of the three best fitting models. Specifically, the model (P1=1, P2=1) generated very long flat tails for the comparator therapies, due to the estimated HR at later time points. This led to clinically implausible results for the comparator therapies (i.e. 5% of patients alive at 15 years for docetaxel, 4% of patients alive at 15 years for paclitaxel) which was wholly inconsistent with the available clinical evidence and expert opinion.[67] The DIC for the equivalent random effects model was slightly higher, suggesting that there is minimal between-study heterogeneity. It was not possible to evaluate inconsistency because the network does not include any comparisons informed by both direct and indirect evidence.

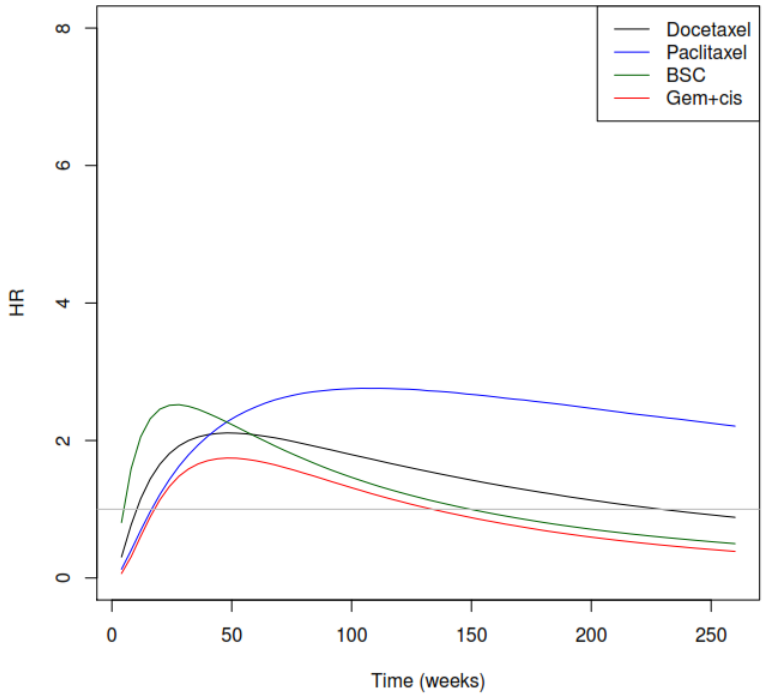
The results of the second order (P1=0, P2=0) fixed effect model are provided below. Further results are provided in Appendix D.

Table 17: Model fit statistics for overall survival

	Fractional polynomial model type	Fractional polynomial model	$\bar{D}_{res}$	pD	DIC
Fixed effect models	First order	P1=0	292.8	14.0	306.8
		P2=1	301.2	13.9	315.2
	Second order	P1=0, P2=1	287.1	18.9	305.9
		P1=0, P2=0	283.6	18.4	302.1
		P1=1, P2=1	282.9	19.2	302.0
Random effects models	First order	P1=0	292.5	14.8	307.3
		P2=1	301.1	14.7	315.8
	Second order	P1=0, P2=1	287.0	19.5	306.5
		P1=0, P2=0	284.3	19.6	303.9
		P1=1, P2=1	282.5	19.3	301.8

Abbreviations: $\bar{D}_{res}$: residual deviance; DIC: deviance information criterion; pD: number of effective parameters.

Figure 26: Overall survival: network meta-analysis results (second order (P1=0, P2=0) fixed effect model): HRs for each of the comparators versus nivolumab



HRs greater than 1 favour nivolumab.
Abbreviations: BSC: best supportive care; cis: cisplatin; gem: gemcitabine; HR: hazard ratio.

Figure 26 illustrates the HRs for each of the comparators versus nivolumab over time, with HRs greater than 1 favouring nivolumab. Table 18 provides estimates of the HRs and their 95% credible intervals for specific time intervals.

Table 18: Overall survival: network meta-analysis results (second order (P1=0, P2=0) fixed effect model): HRs and 95% credible intervals for each of the comparators versus nivolumab for selected time intervals

Comparison	Time Interval (weeks)	HR (95% CrI)
Paclitaxel versus nivolumab	0-4	0.13 (0.02–0.64)
	8-12	0.69 (0.36–1.26)
	20-24	1.43 (0.86–2.31)
	44-48	2.27 (1.41–3.56)
	68-72	2.63 (1.17–5.52)
	92-96	2.75 (0.82–8.52)
Docetaxel versus nivolumab	0-4	0.31 (0.09–0.84)
	8-12	1.15 (0.75–1.72)
	20-24	1.81 (1.25–2.62)
	44-48	2.11 (1.46–3.00)
	68-72	2.01 (1.14–3.37)
	92-96	1.83 (0.8–3.87)

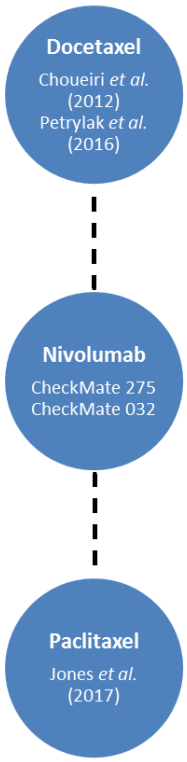
BSC versus nivolumab	0-4	0.81 (0.33–1.79)
	8-12	2.05 (1.36–3.08)
	20-24	2.51 (1.69–3.72)
	44-48	2.27 (1.57–3.25)
	68-72	1.86 (1.17–2.85)
	92-96	1.51 (0.82–2.66)
Cisplatin plus gemcitabine versus nivolumab (scenario analysis only)[a]	0-4	0.06 (0.00–0.70)
	8-12	0.61 (0.21–1.37)
	20-24	1.33 (0.66–2.49)
	44-48	1.75 (0.96–2.99)
	68-72	1.61 (0.68–3.31)
	92-96	1.36 (0.37–4.05)

[a]The trials informing the comparison versus cisplatin plus gemcitabine were conducted in gemcitabine-naïve populations, and therefore cannot be considered to provide relevant data for the retreatment with first-line platinum-based chemotherapy. This comparison is included as a scenario analysis only and results should be interpreted with caution.
Abbreviations: BSC: best supportive care; CrI: credible interval; HR: hazard ratio.

B.2.9.3 Progression-free survival

As per OS, PFS was evaluated using a fractional polynomial NMA approach.[65] The network diagram for PFS is provided in Figure 27 below.

Figure 27: Network diagram for progression-free survival



Dashed lines indicate where simulated treatment comparison has been applied.

Both first order and second order fractional polynomial models, and fixed and random effects models, were evaluated. Table 19 lists the models that were evaluated and provides a summary of the model fit statistics. The table indicates that the second order (P1=0, P2=0) model had a much lower DIC than the other models. The DIC values indicate that the fixed effect and random effects second order (P1=0, P2=0) models have similar fits; the DIC for the fixed effect model is slightly lower. This suggests that there is minimal between-study heterogeneity. It was not possible to evaluate inconsistency because the network does not include any comparisons informed by both direct and indirect evidence.

The results of the fixed effect second order (P1=0, P2=0) model are provided below. Further results are provided in Appendix D.

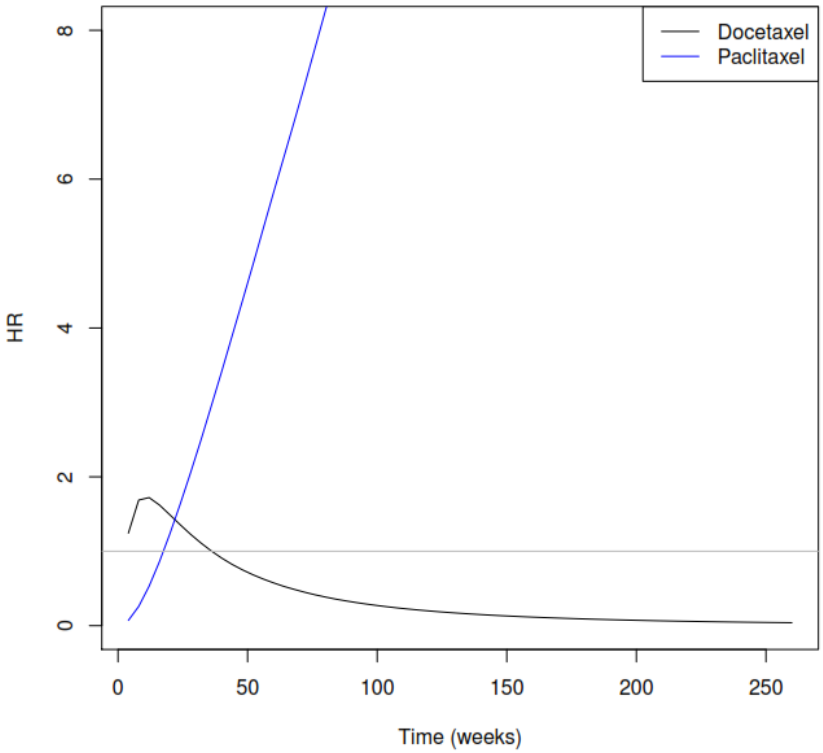
Table 19: Model fit statistics for progression-free survival

	Fractional polynomial model type	Fractional polynomial model	$\bar{D}_{res}$	pD	DIC
Fixed effect models	First order	P1=0	177.3	8.0	185.3
		P2=1	171.1	8.0	179.0
	Second order	P1=0, P2=1	143.9	10.8	154.7
		P1=0, P2=0	132.9	10.8	143.7
		P1=1, P2=1	153.5	10.8	164.3
Random effects models	First order	P1=0	176.4	8.9	185.3
		P2=1	170.4	8.9	179.3
	Second order	P1=0, P2=1	143.6	11.6	155.2
		P1=0, P2=0	132.5	11.6	144.1
		P1=1, P2=1	153.1	11.5	164.6

Abbreviations: $\bar{D}_{res}$: residual deviance; DIC: deviance information criterion; pD: number of effective parameters.

Figure 28 illustrates the HRs for each of the comparators versus nivolumab over time. HRs greater than 1 favour nivolumab. Table 20 provides estimates of the HRs and their 95% credible intervals for specific time intervals. Initially the HR for paclitaxel versus nivolumab is less than 1, indicating that patients receiving paclitaxel have a lower hazard, but over time the HR increases above 1. For docetaxel, the HR is initially greater than 1, indicating that patients receiving docetaxel have a higher hazard, but over time the HR decreases.

Figure 28: Progression-free survival: network meta-analysis results (fixed effect second order (P1=0, P2=0) model): HRs for each of the comparators versus nivolumab



HRs greater than 1 favour nivolumab.
Abbreviations: HR: hazard ratio.

Table 20: Progression-free survival: network meta-analysis results (fixed effect second order (P1=0, P2=0) model): HRs and 95% credible intervals for each of the comparators versus nivolumab for selected time intervals

Comparison	Time Interval (weeks)	HR (95% CrI)
Paclitaxel versus nivolumab	0-4	0.07 (0.01, 0.36)
	8-12	0.53 (0.30, 0.90)
	20-24	1.63 (1.04, 2.52)
	44-48	4.36 (1.84, 9.08)
	68-72	7.26 (1.40, 28.85)
	92-96	10.21 (0.91, 76.04)
Docetaxel versus nivolumab	0-4	1.24 (0.61, 2.42)
	8-12	1.72 (1.18, 2.49)
	20-24	1.36 (0.78, 2.20)
	44-48	0.75 (0.16, 3.19)
	68-72	0.45 (0.04, 4.82)
	92-96	0.29 (0.01, 6.93)

Abbreviations: CrI: credible interval; HR: hazard ratio.

B.2.9.4 Objective response rate

ORR was evaluated using an NMA model for binomial outcomes.[68] The network diagram for ORR is provided in Figure 29 below.

Figure 29: Network diagram for objective response rate

Dashed lines indicate where simulated treatment comparison has been applied. The trials informing the comparison versus cisplatin plus gemcitabine were conducted in gemcitabine-naïve populations, and therefore cannot be considered to provide relevant data for the retreatment with first-line platinum-based chemotherapy. This comparison is included as a scenario analysis only and results should be interpreted with caution.
Abbreviations: BSC: best supportive care.

Both fixed and random effects models were evaluated. Table 21 provides a summary of the model fit statistics. The table indicates that the fixed effect model had the lowest DIC. This suggests that there is minimal between-study heterogeneity. It was not possible to evaluate inconsistency because the network does not include any comparisons informed by both direct and indirect evidence. The results of the fixed effect model are shown below. The results of the random effects model are provided in Appendix D.

Table 21: Model fit statistics for objective response rate

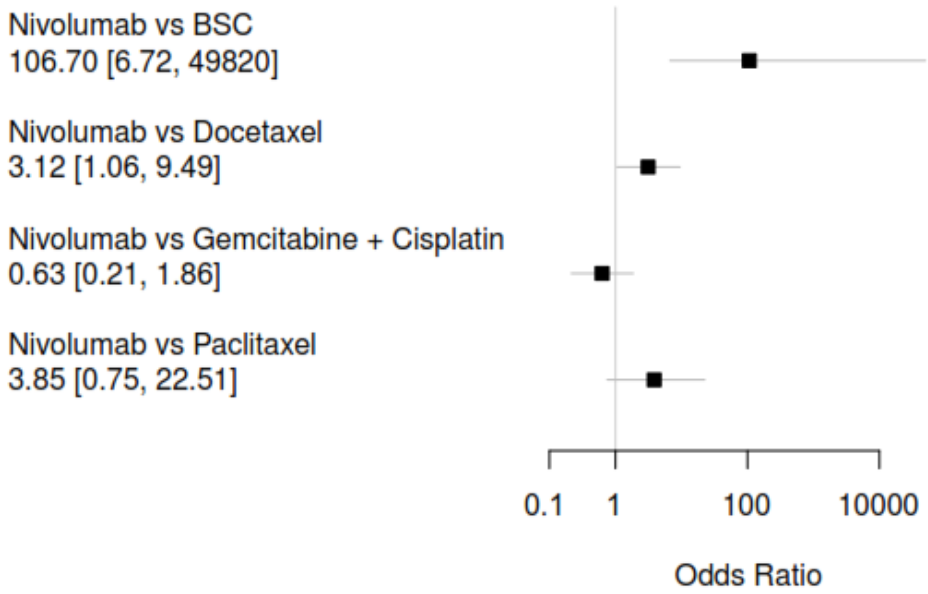
	$\bar{D}_{res}$	pD	DIC
Fixed effect model	10.5	9.7	20.3
Random effects model	11.2	10.5	21.6

Abbreviations: $\bar{D}_{res}$: residual deviance; DIC: deviance information criterion; pD: number of effective parameters.

Figure 30 illustrates the ORR odds ratios for nivolumab versus each of the comparators and Table 22 provides the estimates of the odds ratios and their 95% credible intervals. The results suggest that patients who receive nivolumab have higher odds of response than patients who

receive BSC or docetaxel. There is no evidence of a difference between nivolumab and the other comparators. Note that the comparisons with BSC are very uncertain. This is because no patients responded to BSC in the only trial of this treatment (Bellmunt *et al.* [2009]).[33]

Figure 30: Objective response rate: network meta-analysis results (fixed effect model): Odds ratios for nivolumab versus each of the comparators



The trials informing the comparison versus cisplatin plus gemcitabine were conducted in gemcitabine-naïve populations, and therefore cannot be considered to provide relevant data for the retreatment with first-line platinum-based chemotherapy. This comparison is included as a scenario analysis only and results should be interpreted with caution.
Abbreviations: BSC: best supportive care.

Table 22: Objective response rate: network meta-analysis results (fixed effect model): Odds ratios and 95% credible intervals for each pairwise comparison

	Nivolumab	BSC	Docetaxel	Cisplatin plus gemcitabine
BSC	106.7 (6.72, 49820)			
Docetaxel	3.12 (1.06, 9.49)	0.03 (0.00, 0.59)		
Paclitaxel	3.85 (0.75, 22.5)	0.03 (0.00, 1.00)	1.23 (0.17, 9.74)	6.15 (0.87, 48.4)
Cisplatin plus gemcitabine	0.63 (0.21, 1.86)	0.01 (0.00, 0.12)	0.20 (0.04, 0.93)	

ORs greater than 1 favour the column treatment. The trials informing the comparison versus cisplatin plus gemcitabine were conducted in gemcitabine-naïve populations, and therefore cannot be considered to provide relevant data for the retreatment with first-line platinum-based chemotherapy. This comparison is included as a scenario analysis only and results should be interpreted with caution.
Abbreviations: BSC: best supportive care; OR: odds ratio.

B.2.9.5 Uncertainties in the indirect and mixed treatment comparisons

One source of the uncertainties in the indirect and mixed treatment comparisons is the inclusion of single-arm trial data both for the nivolumab data and some of the comparator data. The inclusion of this evidence needs to be considered in the overall results of this analysis.

Efforts have been made to only include trials that have treatment regimens indicative of UK clinical practice. However, as highlighted above, it needs to be noted that the only trials where patients have been treated with cisplatin plus gemcitabine are Gondo *et al.* (2011)[61] and Ozawa *et al.* (2007),[62] which cannot be considered representative in terms of the doses of gemcitabine and cisplatin used. Additionally, all patients in Gondo *et al.* (2011)[61] had received MVAC in first-line treatment and are therefore not considered to be directly comparable to those receiving gemcitabine and cisplatin re-challenge in current UK clinical practice, as they are gemcitabine naïve.[29] Also the Ozawa *et al.* (2007)[62] trial included chemotherapy-naïve patients in addition to patients who had previously undergone first-line treatment. Although outcome data are reported separately for these two populations, patient baseline characteristic data are reported for the two populations combined. Therefore, it is not possible to determine baseline characteristics for patients who had only received first-line treatment, precluding a comparison with patients in other studies included in this analysis. Due to these serious limitations and the limited use of platinum-based re-challenge in clinical practice, this comparison is presented only as a scenario analysis for the cost-effectiveness analysis.

The network for nivolumab and its comparators is disconnected. Hence the indirect comparison was conducted using STC methodology. Ideally, for each outcome, the STC should adjust for all the effect modifiers and prognostic variables. However, this is rarely possible, as some effect modifiers and prognostic variables may not be reported by all of the trials or may not be known (for example, as yet undiscovered genetic markers).[64] In order to explore the potential error due to missing effect modifiers or prognostic variables we have followed the recommendations in the NICE DSU TSD 18[64] and estimated the residual bias (see Appendix D).

B.2.10 Adverse reactions

CheckMate 275 and CheckMate 032 safety analysis

- The safety and tolerability of nivolumab for patients with locally advanced unresectable or metastatic UC was evaluated as an exploratory endpoint in CheckMate 275 and as a secondary endpoint in CheckMate 032
- The safety profile of nivolumab across both trials was consistent and no new safety signals were raised
- Median duration of therapy was [redacted] months (95% CI: [redacted]) and [redacted] months (95% CI: [redacted]) in CheckMate 275 and CheckMate 032, respectively
- The majority of drug-related AEs were grade 1 or 2 and the frequency of drug-related grade 3 or 4 AEs was low; the most commonly-reported AEs of any grade across both trials were fatigue, nausea and decreased appetite
- Deaths due to study drug toxicity occurred in 3 (1.1%) and 2 (2.6%) of patients in CheckMate 275 and CheckMate 032, respectively
- Predicted select immune-related AEs did occur, but were mostly grade 1 or 2 and were manageable using the recommended treatment guidelines
- Overall, nivolumab in the treatment of locally advanced unresectable and metastatic UC is well tolerated and the safety profile is manageable and consistent with expectations based on prior data in multiple other tumour types

B.2.10.1 Overview

The safety and tolerability of nivolumab for patients with locally advanced unresectable or metastatic UC were evaluated as an exploratory endpoint in the phase II CheckMate 275 trial and as a secondary endpoint in the phase I/II CheckMate 032 trial. The safety data from both trials are presented together in this section of the submission.

In both trials, the safety population included all patients who had received at least one dose of nivolumab (CheckMate 275 all-treated population, n=270; CheckMate 032 all-treated population, n=78). Safety was analysed through the incidence of deaths, AEs, serious AEs, AEs leading to discontinuation, AEs leading to dose delay, select AEs, immune-related AEs (IMAEs) and specific laboratory abnormalities (worst grade). Select AE analyses included incidence, time-to-onset, and time-to-resolution. AEs and laboratory abnormalities were graded using the National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE) version 4.0. AEs were coded using the MedDRA Version 19.0 (CheckMate 275) or 18.1 (CheckMate 032).

B.2.10.2 Treatment duration

CheckMate 275

A total of [redacted] of patients received ≥90% of the planned nivolumab dose intensity, and the median number of doses received was [redacted] (range: [redacted]). The median duration of therapy was [redacted] months.

The Kaplan-Meier plot for duration of therapy for the all-treated population, patients with PD-L1 ≥1% and patients with PD-L1 ≥5% is presented in Figure 31. Patients in the PD-L1 ≥1% and PD-L1 ≥5% cohorts had a longer median duration of therapy ([redacted] and [redacted] months, respectively) than those in the PD-L1 <1% and PD-L1 <5% cohorts ([redacted] and [redacted] months, respectively).

At the time of the 30th May 2016 database lock, 75.6% of patients had discontinued treatment with nivolumab. The most common reasons for discontinuation were disease progression (53.3%), AEs unrelated to nivolumab (12.6%), and nivolumab toxicity (5.2%).

Figure 31: Kaplan-Meier plot of time on treatment for CheckMate 275

CheckMate 032

In CheckMate 032, the majority ( ) of patients received ≥90% of the planned nivolumab dose intensity; the median number of nivolumab doses received was 8.5 with % receiving >4 doses. The median duration of therapy was months (95% CI:).

At the time of the 24th March 2016 database lock, 76.9% of patients in the UC cohort of CheckMate 032 had discontinued study treatment; the most common reason was disease progression (64.1%). Two (2.6%) patients discontinued due to study drug toxicity. The Kaplan-Meier plot for duration of therapy in CheckMate 032 is presented in Figure 32.

Figure 32: Kaplan-Meier plot of time on treatment for CheckMate 032

B.2.10.3 Safety analysis in CheckMate 275 and CheckMate 032

A summary of the safety results from CheckMate 275 and CheckMate 032 is presented in Table 23 below. The majority of treated patients experienced at least one AE regardless of causality, during treatment with nivolumab or within 30 days of the last nivolumab dose.

As of their respective clinical database locks, a total of 138 (51.5%) patients and 36 (46.2%) patients in the CheckMate 275 and CheckMate 032 trials had died, respectively. The proportion of deaths due to study drug toxicity was extremely low (1.1% and 3%, respectively).

All-cause AEs leading to treatment discontinuation were reported in 20.7% and 7.7% of patients in CheckMate 275 and CheckMate 032, respectively.

Table 23: Summary of safety analysis in CheckMate 275 and CheckMate 032

Adverse event, n (%)	CheckMate 275 (n=270)[a]		CheckMate 032 (n=78)[b]	
Deaths	138 (51.1)		36 (46.2)	
Deaths due to study drug toxicity	3 (1.1)[c]		2 (2.6)[d]	
	Any grade	**Grade 3-4**	**Any grade**	**Grade 3-4**
All causality AEs	267 (98.9)	137 (50.7)	78 (100)	43 (55.1)
Drug-related AEs	174 (64.4)	48 (17.8)	65 (83.3)	18 (23.1)
All-causality serious AEs	147 (54.4)	99 (36.7)	36 (46.2)	23 (29.5)
Drug-related serious AEs	[redacted]	[redacted]	8 (10.3)	[redacted]
All-causality AEs leading to treatment discontinuation	56 (20.7)	42 (15.6)	6 (7.7)	4 (5.1)

Drug-related AEs leading to treatment discontinuation	13 (4.8)	8 (3.0)	2 (2.6)	2 (2.6)

[a] AEs were coded using the MedDRA version 19.0 and were graded for severity according to the NCI CTCAE version 4.0. [b] AEs were coded using the MedDRA version 18.1 and were graded for severity according to the NCI CTCAE version 4.0. C Three deaths (Grade 5 pneumonitis, Grade 5 acute respiratory failure, and Grade 5 cardiovascular failure) were judged as study drug-related. [d] Two deaths (Grade 4 pneumonitis and Grade 4 thrombocytopenia) were assessed as study drug-related.
Abbreviations: AEs: adverse events; MedDRA: Medical Dictionary for Regulatory Activities; NCI CTCAE: National Cancer Institute Common Terminology Criteria for Adverse Events; SAEs: serious adverse events.
Source: Sharma *et al* (2017),[38] CheckMate 275 CSR,[39] Galsky *et al.* (2016),[41] Sharma *et al* (2016)[37] CheckMate 032 CSR.[43]

All-cause and drug-related AEs

AEs of any cause that occurred in at least 10% of patients are presented in Table 24. The most commonly reported AEs of any grade across both trials were fatigue (32.2% and 53.8% in CheckMate 275 and CheckMate 032, respectively), nausea (22.2% and 29.5%, respectively), and decreased appetite (21.9% and 14.1%, respectively).

Table 24: All-cause adverse events in ≥10% patients in CheckMate 275 and CheckMate 032

Adverse event	CheckMate 275 (n=270)		CheckMate 032 (n=78)	
	Any grade	Grade 3-4	Any grade	Grade 3-4
Total patients with an event	267 (98.9)	137 (50.7)[a]	78 (100.0)	43 (55.1)[f]
General disorders and administration site conditions	177 (65.6)	31 (11.5)[b]	54 (69.2)	7 (9.0)
Fatigue	87 (32.2)	7 (2.6)	42 (53.8)	3 (3.8)
Pyrexia	47 (17.4)	1 (0.4)	9 (11.5)	1 (1.3)
Asthenia	38 (14.1)	11 (4.1)	N/A	N/A
Oedema peripheral	30 (11.1)	0 (0.0)	11 (14.1)	0 (0.0)
Pain	N/A	N/A	12 (15.4)	0 (0.0)
Gastrointestinal disorders	151 (55.9)	30 (11.1)	46 (59.0)	9 (11.5)
Nausea	60 (22.2)	2 (0.7)	23 (29.5)	1 (1.3)
Diarrhoea	47 (17.4)	7 (2.6)	13 (16.7)	1 (1.3)
Constipation	42 (15.6)	1 (0.4)	13 (16.7)	1 (1.3)
Vomiting	32 (11.9)	5 (1.9)	13 (16.7)	0 (0.0)
Abdominal pain	29 (10.7)	4 (1.5)	14 (17.9)	2 (2.6)
Dry mouth	N/A	N/A	8 (10.3)	0 (0.0)
Musculoskeletal and connective tissue disorders	114 (42.2)	11 (4.1)	44 (56.4)	4 (5.1)
Back pain	32 (11.9)	3 (1.1)	12 (15.4)	1 (1.3)
Arthralgia	N/A	N/A	18 (23.1)	1 (1.3)
Myalgia	N/A	N/A	8 (10.3)	0 (0.0)
Infections and infestations	103 (38.1)	41 (15.2)	30 (38.5)	10 (12.8)
Urinary tract infection	45 (16.7)	17 (6.3)	10 (12.8)	3 (3.8)
Metabolism and nutrition disorders	103 (38.1)	25 (9.3)	33 (42.3)	7 (9.0)
Decreased appetite	59 (21.9)	6 (2.2)	11 (14.1)	0 (0.0)

Hyperglycaemia	N/A	N/A	15 (19.2)	4 (5.1)
Respiratory, thoracic and mediastinal disorders	101 (37.4)	18 (6.7)[c]	41 (52.6)	8 (10.3)
Cough	45 (16.7)	0 (0.0)	17 (21.8)	0 (0.0)
Dyspnoea	35 (13.0)	9 (3.3)	17 (21.8)	4 (5.1)
Skin and subcutaneous tissue disorders	93 (34.4)	6 (2.2)	40 (51.3)	3 (3.8)
Pruritus	32 (11.9)	0 (0.0)	24 (30.8)	0 (0.0)
Rash	28 (10.4)	3 (1.1)	N/A	N/A
Rash maculo-papular	N/A	N/A	16 (20.5)	2 (2.6)
Blood and lymphatic system disorders	61 (22.6)	23 (8.5)	28 (35.9)	7 (9.0)
Anaemia	46 (17.0)	18 (6.7)	24 (30.8)	6 (7.7)
Investigations	N/A	N/A	38 (48.7)	13(16.7)
Blood creatinine increased	N/A	N/A	14 (17.9)	0 (0.0)
Lipase increased	N/A	N/A	11 (14.1)	4 (5.1)
Nervous system disorders	N/A	N/A	29 (37.2)	2 (2.6)
Headache	N/A	N/A	10 (12.8)	0 (0.0)
Peripheral sensory	N/A	N/A	8 (10.3)	0 (0.0)
Renal and urinary disorders	N/A	N/A	29 (37.2)	11 (14.1)
Haematuria	N/A	N/A	14 (17.9)	4 (5.1)
Acute kidney injury	N/A	N/A	8 (10.30	4 (5.1)
Psychiatric disorders	N/A	N/A	13 (16.7)	1 (1.3)
Vascular disorders	N/A	N/A	13 (16.7)	2 (2.6)
Neoplasms benign, malignant and unspecified (including cysts and polyps)	57 (21.1)	29 (10.7)[d]	10 (12.8)	1 (1.3)[g]
Malignant neoplasm progression	35 (13.0)	14 (5.2)[e]	N/A	N/A
Endocrine disorders	N/A	N/A	9 (11.5)	0 (0.0)

[a] 31 (11.5%) Grade 5 all-cause AEs. [b] 3 (1.1%) Grade 5 all-cause AEs. C 3 (1.1%) Grade 5 all-cause AEs. [d] 21 (7.8%) Grade 5 all-cause AEs. [e] 20(7.4%) Grade 5 all-cause AEs. [f] 7 (9.0%) Grade 5 all-cause AEs. [g] 6 (7.7%) Grade 5 all-cause AEs.
Abbreviations: AEs: adverse events; N/A: not applicable.
Source: CheckMate 275 CSR[39] and CheckMate 032 CSR.[43]

Table 25: Drug-related adverse events in ≥5% patients in CheckMate 275 and CheckMate 032

Adverse event	CheckMate 275 (n=270)		CheckMate 032 (n=78)	
	Any grade	Grade 3-4	Any grade	Grade 3-4
Total patients with an event	174 (64.4)	48 (17.8)[a]	65 (83.3)	18 (23.1)[b]
General disorders and administration site conditions	80 (29.6)	10 (3.7)	29 (37.2)	2 (2.6)
Fatigue	45 (16.7)	5 (1.9)	28 (35.9)	2 (2.6)
Asthenia	16 (5.9)	4 (1.5)	N/A	N/A
Pyrexia	15 (5.6)	0 (0.0)	N/A	N/A

Gastrointestinal disorders	54 (20.0)	7 (2.6)	24 (30.8)	2 (2.6)
Diarrhoea	24 (8.9)	5 (1.9)	7 (9.0)	0 (0.0)
Nausea	19 (7.0)	1 (0.4)	10 (12.8)	1 (1.3)
Skin and subcutaneous tissue disorders	54 (20.0)	6 (2.2)	34 (43.6)	3 (3.8)
Pruritus	25 (9.3)	0 (0.0)	23 (29.5)	0 (0.0)
Rash	16 (5.9)	3 (1.1)	5 (6.4)	0 (0.0)
Rash maculo-papular	N/A	N/A	14 (7.9)	2 (2.6)
Dry skin	N/A	N/A	5 (6.4)	0 (0.0)
Investigations	N/A	N/A	26 (33.3)	8 (10.3)
Lipase increased	N/A	N/A	11 (14.1)	4 (5.1)
Amylase increased	N/A	N/A	7 (9.0)	3 (3.8)
Lymphocyte count decreased	N/A	N/A	5 (6.4)	2 (2.6)
Blood creatinine increased	N/A	N/A	4 (5.1)	0 (0.0)
Endocrine disorders	31 (11.5)	1 (0.4)	6 (7.7)	0 (0.0)
Hypothyroidism	21 (7.8)	0	4 (5.1)	0 (0.0)
Musculoskeletal and connective tissue disorders	N/A	N/A	13 (16.7)	1 (1.3)
Arthralgia	N/A	N/A	9 (11.5)	0 (0.0)
Metabolism and nutrition	27 (10.0)	3 (1.1)	10 (12.8)	2 (2.6)
Decreased appetite	22 (8.1)	0	5 (6.4)	0 (0.0)
Hyperglycaemia	N/A	N/A	5 (6.4)	1 (1.3)
Blood and lymphatic system disorders	N/A	N/A	11 (14.1)	1 (1.3)
Anaemia	N/A	N/A	8 (10.3)	0 (0.0)
Respiratory, thoracic and mediastinal disorders	N/A	N/A	11 (14.1)	1 (1.3)[b]
Dyspnoea	N/A	N/A	6 (7.7)	2 (2.6)
Nervous system disorders	N/A	N/A	7 (9.0)	0 (0.0)

[a]Grade 5 events reported in 3 (1.1%) patients (1 death due to pneumonitis, 1 death due to acute respiratory failure, 1 death due to cardiovascular failure). [b] 1 (1.3%) Grade 5 drug-related AE (pneumonitis).
Abbreviations: AEs: adverse events; N/A: not applicable.
Source: Sharma *et al.* (2017),[38] Sharma *et al.* (2016),[37] CheckMate 275 CSR[39] and CheckMate 032 CSR.[43]

Select AEs

Select AEs were defined as AEs of special clinical interest that are potentially associated with the use of nivolumab, and were identified based on the following principles:

- AEs that may differ in type, frequency, or severity from AEs caused by non-immunotherapies
- AEs that may require immunosuppression (e.g. Corticosteroids) as part of their management
- AEs whose early recognition and management may mitigate severe toxicity
- AEs for which multiple event terms may be used to describe a single type of AE, thereby necessitating the pooling of terms for full characterisation.

Considering the AEs already observed across other studies of nivolumab therapy, the AEs considered as select AEs were endocrinopathies, diarrhoea/colitis, hepatitis, pneumonitis,

interstitial nephritis, rash and hypersensitivity/infusion reactions. Hypersensitivity/infusion reactions were analysed along with select AEs because multiple event terms may be used to describe such events and pooling of terms was therefore necessary for full characterisation; they would not otherwise meet the criteria to be considered select AEs.

Multiple event terms that may describe each of these AEs were grouped into endocrine, GI, hepatic, pulmonary, renal, skin and hypersensitivity/infusion reactions select AE categories, respectively.

The majority of select AEs were grade 1 or 2, with very few higher-grade hepatic and pulmonary events reported: 1 subject with a grade 4 hepatic select AE and 2 patients with grade 5 pulmonary select AEs. Most select AEs were considered drug-related by the investigator, with the exception of hepatic and renal events, where a lower proportion of select AEs were deemed to be drug-related. The most frequently reported any-grade drug-related select AE categories were skin (17.4%) and endocrine (14.4%) – see Table 26 below.

Overall, across all select AE categories, the majority of events were manageable, with resolution occurring when immune-modulating medications (mostly systemic corticosteroids) were administered.

Table 26: Drug-related select adverse events in CheckMate 275 and CheckMate 032

Select adverse event, n (%)	CheckMate 275		CheckMate 032	
	Any grade	Grade 3–4	Any grade	Grade 3–4
Total patients with an event, by category				
Skin	47 (17.4)	4 (1.5)	33 (42.3)	2 (2.6)
Endocrine	39 (14.4)	1 (0.4)	6 (7.7)	0 (0.0)
Gastrointestinal	25 (9.3)	6 (2.2)	8 (10.3)	1 (1.3)
Hepatic	10 (3.7)	5 (1.9)	4 (5.1)	1 (1.3)
Pulmonary	11 (4.1)	3 (1.1)	2 (2.6)	0 (0.0)
Renal	3 (1.1)	1 (0.4)	7 (9.0)	1 (1.3)
Hypersensitivity/infusion reactions	3 (1.1)	1 (0.4)	2 (2.6)	0 (0.0)
Drug-related 'select' AEs, by category				
Skin				
Pruritis	25 (9.3)	0 (0.0)	23 (29.5)	0 (0.0)
Rash	16 (5.9)	3 (1.1)	5 (6.4)	0 (0.0)
Rash maculo-papular	4 (1.5)	1 (0.4)	14 (17.9)	2 (2.6)
Erythema	2 (0.7)	0 (0.0)	N/A	N/A
Pruritis generalised	2 (0.7)	1 (0.4)	N/A	N/A
Rash macular	2 (0.7)	0 (0.0)	N/A	N/A
Rash pruritic	2 (0.7)	1 (0.4)	N/A	N/A
Rash erythematous	N/A	N/A	2 (2.6)	0 (0.0)
Rash papular	N/A	N/A	1 (1.3)	0 (0.0)
Palmar-plantar erythrodysaesthesia syndrome	N/A	N/A	1 (1.3)	0 (0.0)
Blister	1 (0.4)	0 (0.0)	N/A	N/A
Dermatitis	1 (0.4)	0 (0.0)	N/A	N/A

Select adverse event, n (%)	CheckMate 275		CheckMate 032	
	Any grade	Grade 3–4	Any grade	Grade 3–4
Eczema	1 (0.4)	0 (0.0)	N/A	N/A
Rash generalised	1 (0.4)	0 (0.0)	N/A	N/A
Skin exfoliation	1 (0.4)	0 (0.0)	N/A	N/A
Skin irritation	N/A	N/A	1 (1.3)	0 (0.0)
Urticaria	1 (0.4)	0 (0.0)	N/A	N/A
Endocrine				
Thyroid disorder	35 (13.0)	0 (0.0)	6 (7.7)	0 (0.0)
Hypothyroidism	21 (7.8)	0 (0.0	4 (5.1)	0 (0.0)
Hyperthyroidism	11 (4.1)	0 (0.0	3 (3.8)	0 (0.0)
Blood thyroid stimulating hormone increased	10 (3.7)	0 (0.0	1 (1.3)	0 (0.0)
Blood thyroid stimulating hormone decreased	5 (1.9)	0 (0.0	N/A	N/A
Thyroiditis	2 (0.7)	0 (0.0	N/A	N/A
Drug-related 'select' AEs, by category - continued				
Thyroxine increased	2 (0.7)	0 (0.0	N/A	N/A
Autoimmune thyroiditis	1 (0.4)	0 (0.0	N/A	N/A
Thyroxine decreased	1 (0.4)	0 (0.0	N/A	N/A
Thyroxine free increased	1 (0.4)	0 (0.0	N/A	N/A
Adrenal disorder	2 (0.7)	0 (0.0)	N/A	N/A
Adrenal insufficiency	2 (0.7)	0 (0.0)	N/A	N/A
Pituitary disorder	2 (0.7)	1 (0.4)	N/A	N/A
Hypophysitis	2 (0.7)	1 (0.4)	N/A	N/A
Diabetes	1 (0.4)	0 (0.0)	N/A	N/A
Type I diabetes mellitus	1 (0.4)	0 (0.0)	N/A	N/A
Gastrointestinal				
Diarrhoea	24 (8.9)	5 (1.9)	7 (9.0)	0 (0.0)
Colitis	2 (0.7)	1 (0.4)	1 (1.3)	1 (1.3)
Hepatic				
Alanine aminotransferase increased	8 (3.0)	2 (0.7)	3 (3.8)	0 (0.0)
Aspartate aminotransferase increased	6 (2.2)	3 (1.1)	1 (1.3)	1 (1.3)
Blood alkaline phosphatase increased	3 (1.1)	2 (0.7)	1 (1.3)	0 (0.0)
Blood bilirubin increased	2 (0.7)	1 (0.4)	1 (1.3)	0 (0.0)
Liver function test increased	2 (0.7)	1 (0.4)	N/A	N/A
Transaminases increased	2 (0.7)	0 (0.0)	N/A	N/A
Hyperbilirubinaemia	1 (0.4)	0 (0.0)	N/A	N/A
Pulmonary				
Pneumonitis	10 (3.7)	2 (0.7)	2 (2.6)	0 (0.0)
Interstitial lung disease	1 (0.4)	1 (0.4)	N/A	N/A
Renal				

Select adverse event, n (%)	CheckMate 275		CheckMate 032	
	Any grade	Grade 3–4	Any grade	Grade 3–4
Acute kidney injury	1 (0.4)	0 (0.0)	1 (1.3)	1 (1.3)
Blood creatinine increased	1 (0.4)	1 (0.4)	4 (5.1)	0 (0.0)
Renal failure	1 (0.4)	0 (0.0)	N/A	N/A
Blood urea increased	N/A	N/A	3 (3.8)	0 (0.0)
Hypersensitivity/infusion reactions				
Infusion related reaction	2 (0.7)	1 (0.4)	1 (1.3)	0 (0.0)
Hypersensitivity	1 (0.4)	0 (0.0)	1 (1.3)	0 (0.0)

Includes events reported between first dose and 30 days after last dose of study therapy.
Abbreviations: AEs: adverse events; N/A: not applicable.
Source: Sharma *et al.* (2017),[38] CheckMate 275 CSR[39] and CheckMate 032 CSR.[43]

B.2.10.4 Safety conclusions

As described in Section B.1.3.2, currently available therapies for locally advanced unresectable or metastatic UC comprise chemotherapy options, many of which are associated with high toxicity. There is a critical unmet need for well-tolerated treatment options for patients at this stage of disease. Nivolumab represents a novel treatment option with an innovative immunological mechanism of action. The safety profile of nivolumab is favourable and has been consistently demonstrated in two large clinical trials of patients with locally advanced unresectable or metastatic UC.

'Select' AEs that represent AEs of particular interest for patients treated with nivolumab did occur in CheckMate 275 and CheckMate 032; however, these were mainly grade 1–2 in severity, and the majority of events were resolved and generally manageable using recommended treatment guidelines. No new safety concerns with nivolumab were identified across the two trials, and the demonstrated safety profile is consistent with the safety/tolerability profile observed with nivolumab in trials for multiple other tumour types.[69]

B.2.11 Ongoing studies

Both CheckMate 275 and CheckMate 032 are still ongoing and interim analyses are planned following the next database locks for CheckMate 275 and CheckMate 032 in [redacted] and [redacted], respectively. No further trials are currently ongoing or planned for nivolumab in locally advanced unresectable or metastatic UC.

B.2.12 Innovation

For patients with locally advanced unresectable or metastatic UC who have progressed following platinum-based chemotherapy, tolerable treatment options with proven survival benefits are extremely limited. Patients treated with current chemotherapy regimens have an estimated life expectancy of only 5–9 months and are thus considered to be at an end-of-life disease stage.[30-33] Only 10% of patients typically respond to second-line single-agent chemotherapy regimens, and complete responses are rare and short-lived.[34, 35] Furthermore, many of the chemotherapy agents are associated with high toxicity and many patients instead choose to receive BSC only or clinicians will seek to enrol their patients in a clinical trial. Therefore, there is a critical unmet need for novel, effective and tolerable treatment options, offering durable survival benefit for patients at this stage of disease.[19]

As detailed in Section B.1.2, rather than relying on the indiscriminate cytotoxic effects of chemotherapy, nivolumab harnesses the body's own immune system to destroy cancer cells via the restoration of anti-tumour T-cell activity, representing a highly innovative mechanism of action. The use of immunotherapy for the treatment of UC has been ongoing for over 40 years and the potential of targeting immune inhibitory pathways to treat UC is indicated by the effectiveness in some patients of BCG, an immunotherapy treatment and the standard of care for patients with high-grade non-muscle-invasive UC following surgical resection. Given intravesically, BCG induces an immune response against tumour cells, leading to the secretion of cytokines from urothelial cells and the attraction of vast numbers of neutrophils and monocytes.[6, 7] There is also evidence in studies of patients with localised UC that the use of ipilimumab, an immune checkpoint inhibitor that blocks CTLA-4, enhances immune responses and tumour regression.[8, 9] As such, this evidence provides a compelling biological rationale for the effectiveness of nivolumab and the blocking of PD-1 as a therapeutic target in UC.[10, 11] The awarding of a Breakthrough Therapy Designation by the FDA is recognition of the innovative nature of nivolumab.[14]

With this innovative mechanism of action, nivolumab has demonstrated clinically meaningful improvements in tumour response rates in CheckMate 275, a large phase II trial of patients with unresectable or metastatic UC following progression on prior platinum-containing chemotherapy, with 52 patients (19.6%) achieving an objective response, including a complete response in 6 (2.3%) of patients (see Section B.2.3.1).[41] These response rates are supported by data from CheckMate 032, a smaller phase I/II trial, in which a total of 19 (24.4%) patients achieved an objective response. In addition, nivolumab demonstrated promising survival for these patients: at the latest database lock of CheckMate 275, median OS was 8.57 months and 41.0% of patients were still alive at 1 year.[41] This efficacy was observed regardless of baseline tumour PD-L1 expression status, including those with PD-L1 expression <1%. Furthermore, long-term survival benefits with nivolumab have been observed in the other cancer indications that have been investigated, such as advanced NSCLC, advanced renal cell carcinoma, and advanced melanoma, and for which data from longer follow-up are available.[70-72]

In addition, nivolumab has demonstrated a predictable and generally manageable safety profile across both CheckMate 275 and CheckMate 032 in UC, consistent with that demonstrated across several previous indications, illustrating that nivolumab may offer improvements in tolerability compared to the cytotoxic chemotherapies that represent the currently available therapies for these patients.

The introduction of nivolumab as a highly-innovative and well-tolerated therapy with demonstrable and durable tumour response rates and survival outcomes represents a step-change in the management of patients with locally advanced unresectable or metastatic UC after the failure of prior platinum-containing chemotherapy. These patients currently have limited effective, tolerable treatment options available and nivolumab has the potential to help address the considerable unmet medical need for these patients at an end-of-life stage.

B.2.13 Interpretation of clinical effectiveness and safety evidence

Patients with locally advanced unresectable or metastatic UC who have progressed after prior platinum-containing chemotherapy currently face an unmet need for effective and tolerable treatment options. Current treatment options are associated with limited efficacy and are restricted to relatively toxic chemotherapy drugs that may not be tolerated by many patients given the advanced age of many individuals with this condition. For patients unable or unwilling to tolerate chemotherapy, remaining treatment options are restricted to palliative BSC or enrolment in clinical trials. Access to clinical studies is limited by the availability of a recruiting clinical trial in an appropriate geography for a given patient and by the inclusion/exclusion criteria governing entry to such trials. There is therefore a clear need for national availability on the NHS of a licensed treatment that presents a tolerable and effective therapeutic option for this patient group.

Evidence for the efficacy and safety of nivolumab, a novel PD-1 inhibitor-based immunotherapy, is provided by the pivotal CheckMate 275 study, an ongoing phase II single-arm study in patients with locally advanced unresectable or metastatic UC whose disease has progressed or recurred after treatment with at least one platinum-containing chemotherapy regimen. CheckMate 275 investigated nivolumab monotherapy at the licensed dose (IV 3 mg/kg Q2W) in a patient population that matches the final scope of this appraisal and provides evidence on outcomes relating to tumour response rates, disease progression, OS, patient-reported outcomes and safety profile. Evidence for the efficacy and safety is also supported by a cohort of patients within the CheckMate 032 study who were treated with nivolumab monotherapy at the licensed dose for locally advanced unresectable or metastatic UC who had progressed after at least one previous line of platinum-containing chemotherapy. This evidence base forms the basis of the positive EMA approval on nivolumab for this indication, as well as the granting of an initial Breakthrough Therapy Designation and subsequent Accelerated Approval by the FDA in the USA (see Section B.1.2).

In the absence of RCT data, evidence for the comparative efficacy of nivolumab versus the relevant comparators to this submission is provided from an STC (see Section B.2.9).

B.2.13.1 Principal findings from the clinical evidence base

Nivolumab provided clinically meaningful tumour responses and survival outcomes

At the primary database lock, BIRC-assessed ORR in the CheckMate 275 study was 19.6% (95% CI: 15.0–24.0), with 6 (2.3%) patients achieving a best overall response of CR. Median TTR (as per BIRC) was 1.87 months (range: 1.6–5.9), with the majority of objective responders seen to achieve their response within the first 8 weeks. The vast majority of responders (96.2%) had a response lasting at least 3 months, and in a number of patients tumour responses were observed to continue after treatment discontinuation. Tumour responses were therefore seen to occur early following initiation of treatment and to be durable once established in the majority of cases. At the latest database lock, ORR was consistent at 20.0% (95% CI: 15.4-25.3).

Consistent results were also achieved in CheckMate 032, with 19 (24.4%) patients (95% CI: 15.3–35.4) of 78 treated patients achieving a confirmed investigator-assessed objective response, five patients (6%) achieving a CR and 14 patients (18%) achieving a PR.

Median PFS (by BIRC) in the efficacy-treated population was 2.00 months (95% CI: 1.87–2.63) at the primary database lock, with a quarter (25.2%) of patients being progression-free 6 months after initiation of therapy. Results by investigator assessment provided consistent evidence of median PFS on nivolumab. The Kaplan-Meier plot of PFS highlights a potential plateauing of the PFS curve, with more gradual decline in the proportion of patients who remain progression-free from approximately 4 months. This may indicate the potential for a proportion of patients to achieve more prolonged progression-free status, which would be consistent with the observation that a majority of those patients that did respond exhibited durable response (see above). At the latest clinical database lock an additional 5 PFS events had occurred, with median PFS remaining unchanged at 2.00 months (95% CI: 1.87-2.63). Similar results were observed in CheckMate 032, with median PFS 2.78 months (95% CI 1.5–5.9)

In interpreting results relating to tumour response rates and disease progression it should be remembered that these were determined based on the RECIST v1.1 criteria. These criteria provide a well-established measure frequently used in clinical trials of anti-cancer therapies; however, they may have limitations as a method of evaluating clinical benefit in terms of response or progression with immune-checkpoint inhibitors. This is because some patients who ultimately derive clinical benefit from immunotherapy may initially progress by RECIST criteria before exhibiting a response (see Section B.1.2, Figure 4, for further information). Consistent with this notion, a notable proportion of patients in the CheckMate 275 study (24.6%) were treated with nivolumab beyond RECIST v1.1-defined progression. Given that treatment beyond progression was permitted where the study investigator determined that patients were achieving clinical benefit and tolerating the study drug, this may indicate that progression as defined by the RECIST criteria did not fully align with investigator opinion regarding continuing patient clinical benefit on nivolumab therapy.

At the primary analysis database lock (30th May 2016), 51.1% of patients in the CheckMate 275 study had died, with median OS estimated at 8.74 months (95% CI: 6.05–NR) for the efficacy-treated population. Median OS was therefore considerably higher than median PFS observed in the study, implying prolonged post-progression survival; perhaps again reflecting the limitations of the RECIST criteria in assessing clinically relevant disease progression for immunotherapies. At 3 months and 6 months following treatment initiation the proportion of patients remaining alive in the study was 75.8% (95% CI: 70.2–80.5) and 57.0% (50.7–62.7), respectively. The latest database lock provides updated OS data with a median follow-up time for OS of 11.5 months, which is approximately 3 months later than the primary database lock. At this point, 57% of patients had died (16 additional deaths versus the primary database lock) and the median OS estimate remained consistent at 8,57 months (95% CI: 6.05-11.27). Furthermore, 45.6% of patients were still alive at 1 year, demonstrating the durability of response seen with nivolumab. Consistent OS results were observed in CheckMate 032: median OS was 9.7 months (95% CI 7.3–16.2), with 32 (41%) of 78 patients still alive at the time of data cut-off.

Consistent clinical benefits were observed with nivolumab regardless of level of PD-L1 expression

In both CheckMate 275 and CheckMate 032, subgroup analyses were performed on a number of outcomes to investigate treatment efficacy in patients with differential PD-L1 expression. The investigation of any relationship between level of PD-L1 expression and treatment efficacy is an

important consideration for therapies such as nivolumab that specifically target the PD-1/PD-L1 signalling pathway (see Section B.1.2).

Subgroup analyses found nivolumab to demonstrate efficacy regardless of the level of PD-L1 expression on the tumour. Although higher PD-L1 expression was associated with numerically higher ORR, PFS and OS, in CheckMate 275, patients with low to no PD-L1 expression (<1%) had an ORR which exceeded 15%, median PFS of 1.87 months and median OS of almost 6 months, and consistent results were observed in CheckMate 032. Efficacy of nivolumab was therefore established irrespective of PD-L1 expression levels.

Simulated treatment comparison demonstrated a superior clinical benefit with nivolumab versus current treatment options

As no RCTs directly comparing the efficacy and safety of nivolumab in the patient population of interest versus any of the comparators relevant to this submission or placebo were identified in the SLR, a population-adjusted approach using STC techniques was conducted using individual patient level data from the nivolumab trials and summary data from the comparator trials, to estimate how patients in each of the comparator trials would have responded to nivolumab. The STC was conducted in accordance with the recently published technical support document on population-adjusted indirect comparisons (MAIC and STC) in NICE submissions (TSD18).[64] The STC was informed by studies identified through a robust clinical SLR that provided an evidence base for the key comparators of paclitaxel monotherapy, docetaxel monotherapy and BSC. It should be noted, however, that no studies were identified for retreatment with first-line platinum-based chemotherapy. The SLR did identify two trials involving cisplatin plus gemcitabine after the failure of MVAC-regimen which could be included in the STC, and these have been included as a scenario analysis in the absence of other clinical data. However, as these trials were conducted in a gemcitabine-naïve patient population, the results should be treated with a great deal of caution. Additionally, the trials used a dosing regimen which is different to that used in UK clinical practice, further limiting their generalisability to UK clinical practice.

Results from this STC demonstrated the clinical efficacy of nivolumab versus the relevant comparators. Time-varying HRs for OS demonstrated that the HR for death with paclitaxel, docetaxel, BSC and cisplatin plus gemcitabine were all greater than 1 (increased risk of death versus nivolumab) from week 4 throughout week 96. Furthermore, the relative ORR between nivolumab and paclitaxel or docetaxel was estimated using odds ratios, and results from the STC demonstrated that patients who receive nivolumab have a statistically significantly higher odds of response than patients who receive BSC or docetaxel (odds ratio: 106.70 and 3.12 for nivolumab versus BSC and docetaxel, respectively). The odds ratio versus paclitaxel was 3.85, though this was not significant.

Despite the absence of RCT data, these results demonstrate the superior efficacy of nivolumab versus the therapies that are currently used in clinical practice.

Nivolumab provided stable health-related quality of life outcomes for patients and a tolerable safety profile consistent with that observed in other indications

A major concern with currently available active therapies for the treatment of patients with locally advanced unresectable or metastatic UC who have progressed after prior platinum-containing chemotherapy is their tolerability and the detrimental impact such treatment can therefore have on patient HRQoL. For patients unable to tolerate chemotherapy the only remaining treatment option is BSC. As BSC essentially represents palliative care, it would not be expected to induce tumour responses that lead to reductions in tumour bulk and the alleviation of symptoms

associated with this. Key goals of new treatments for these patients are therefore achievement of control of disease (i.e. tumour growth) and maintenance of HRQoL.

Assessment of EORTC QLQ-C30, a commonly used quality of life measure in oncology trials, found that HRQoL with nivolumab increased or was maintained throughout the trial from baseline to Week 41, with clinically meaningful improvements observed for domains of dyspnoea, insomnia and financial difficulties at isolated time points. Response rates for the EORTC QLC-C30 were high, meeting or exceeding 75% at all assessments for which patients were eligible for on-treatment patient-reported outcomes assessment. Using the EQ-5D VAS measure, the mean baseline score of 60.2 was seen to have increased to 67.5 by Week 9 and more than 80 points by Week 41. A score of over 80 points is in alignment with that of the US general population (the country with the largest representation in the study). Taken together, these patient-reported outcome results provide compelling evidence for the maintenance and potential improvement, in some domains, of HRQoL on nivolumab. In the context of the toxicity and negative HRQoL impact of currently available therapies in clinical practice, these results indicate the potential for nivolumab to help address the unmet need for a tolerable treatment option on which patients can continue to enjoy reasonable quality of life.

B.2.13.2 End-of-life criteria

The systematic literature review presented in Appendix D identified a number of studies providing estimates of OS, for therapies both relevant and not relevant to clinical practice in the UK specifically. However, across all studies identified from the SLR, no study provided evidence of OS estimates for this patient population that approached the 24 months that represents the threshold for NICE's end of life criteria.

Table 27: End-of-life criteria

Criterion	Data available
The treatment is indicated for patients with a short life expectancy, normally less than 24 months	• No studies identified in the SLR reported in Appendix D provided evidence of OS estimates for this patient population that approached 24 months
There is sufficient evidence to indicate that the treatment offers an extension to life, normally of at least an additional 3 months, compared with current NHS treatment	• The economic analysis predicted mean life years per patient with nivolumab of 2.78 years (33.36 months) • In comparison, predicted mean life years per patient with comparator therapies were 1.19 years (14.28 months) with paclitaxel, 1.40 years (16.80 months) with docetaxel and 1.01 years (12.12 months) with BSC. Nivolumab was therefore predicted to offer an extension to life of considerably greater than 3 months versus each of these comparators. Furthermore, in the context of the average survival of patients receiving paclitaxel, docetaxel or BSC, the survival gains offered by nivolumab represent a significant extension to life.

Abbreviations: BSC: best supportive care; OS: overall survival; SLR: systematic literature review.

B.2.13.3 Strengths of the clinical evidence base

Both the CheckMate 275 and CheckMate 032 studies match the decision problem outlined in the final scope for this appraisal; patients with advanced metastatic or unresectable UC who have progressed after prior platinum-containing chemotherapy. The baseline characteristics of the patients recruited to both CheckMate 275 and CheckMate 032 are aligned to that which would be expected in UK clinical practice, with patients generally being older (median age of 66), male

(78.1%) and current/former smokers (71.9%), which can be considered generally in line with the expected profile of a 'typical' bladder cancer patient.[29] Overall, both CheckMate 275 and CheckMate 032 can therefore be considered to provide evidence on the efficacy and safety of nivolumab in a patient population relevant to both the scope of this appraisal and to the expected patient population in clinical practice.

Although a single-arm trial, which presents acknowledged limitations for the evidence base as outlined below, the CheckMate 275 study was assessed to be of satisfactory quality, using appropriate methods for data collection (see Section B.2.5). Furthermore, the outcomes assessed in both CheckMate 275 and CheckMate 032 represent standard outcome measures for the assessment of anticancer therapies. Control of tumour response, delayed progression, maintenance of quality of life and ultimately extended length of life are key goals in the treatment of cancer generally and for bladder cancer specifically. The outcome measures evaluated are therefore also relevant to patients and clinicians in clinical practice. The use of RECIST criteria in cancer trials is recommended by the EMA and provides an objective measure of tumour response and PFS.[28] Nonetheless, it should be noted that in clinical practice response to therapy will most likely be assessed based on clinical judgement rather than radiological assessments and that RECIST may have limitations as a method of evaluating clinical benefit in terms of response or progression with immune-checkpoint inhibitors (see Section B.2.13.1).

B.2.13.4 Limitations of the evidence base

The key limitation of the evidence base is the lack of randomised controlled trials to inform relative efficacy estimates with nivolumab. The single-arm nature of the CheckMate 275 (and CheckMate 032) trials means that any 'placebo effect' resulting from the receipt of an active intervention (irrespective of the biological activity of that agent) cannot be adequately accounted for, reducing reliability of study results as a true estimation of treatment effect. Single-arm studies are more susceptible to selection and assessment bias, which may further reduce confidence in study results.

However, whilst randomised controlled trials represent the current 'gold standard' of trial design, it is not always possible or appropriate to conduct such a trial and single-arm studies may in some cases be the most appropriate form of study design. A single-arm study design was chosen for CheckMate 275 on the basis that there was no standard available therapy for patients with metastatic or unresectable UC who have progressed on prior platinum chemotherapy. Furthermore, currently available chemotherapy regimens are unlicensed, and associated with high rates of toxicity. Recommendations for the use of current chemotherapy regimens for UC are based on small, phase II studies that vary in terms of eligibility criteria and definitions of second-line treatment, hence there is a distinct lack of evidence of proven clinical benefit to warrant their appropriate inclusion within a comparative study. In such cases where there is no clear, effective, standard of care and, due to the end-of-life nature of the indication, it would be inappropriate and unethical to randomise patients to placebo, hence a single-arm trial is the appropriate choice of trial design.

In the absence of RCT data for nivolumab versus the relevant comparators to this submission, a robust STC was conducted using data available from a clinical SLR. The STC adheres to the recommendations outlined in the recent technical support document (TSD18)[64] and has been informed by clinical expert opinion, to ensure all relevant treatment effect modifiers were included.[29, 67] The results presented are consistent with elicited clinical input and expected outcomes for nivolumab and the comparators in second-line locally advanced unresectable or metastatic UC.

A limitation of the evidence provided by CheckMate 275 and CheckMate 032 is the relative immaturity of the OS outcome, with 41.0% and 45.6% of patients still alive at 1 year, respectively (latest database lock). However, validation of the survival outcomes has been undertaken for both nivolumab and the comparators, using additional clinical data (from other trials and real world practice) and clinical expert opinion.

It should also be noted that there is an extensive body of evidence on survival outcomes with nivolumab from longer-term follow-up in other indications. Specifically, data from longer follow-up are available from trials of nivolumab in advanced NSCLC, advanced renal cell carcinoma (RCC) and advanced melanoma that demonstrate long-term survival benefits with nivolumab.[70-72] Increasing evidence suggests that immune-checkpoint inhibitors (including those targeting PD-1 and cytotoxic T-lymphocyte-associated protein 4 [CTLA-4]) are characterised by survival curves with a long, plateauing tail for a subset of patients, and that marked differences in the shape of survival curves (OS and PFS) may be observed compared to standard cytotoxic therapies due to differences in mechanism of action.[73] Based on survival patterns observed in longer-term data for nivolumab in these other cancer indications there is evidence that the highly innovative mechanism of action of nivolumab as an immune-checkpoint inhibitor may offer some patients a long-term, durable survival.[71, 72] Additionally, as highlighted previously, there is evidence from other therapies on the potential of immunotherapy in UC. BCG has been used for over 40 years in patients with high-grade non-muscle-invasive UC following surgical resection; given intravesically, BCG induces an immune response against tumour cells, leading the secretion of cytokines from urothelial cells and the attraction of vast numbers of neutrophils and monocytes.[6, 7] There is also evidence in studies of patients with localised UC that the use of ipilimumab, an immune checkpoint inhibitor that blocks CTLA-4, enhances immune responses and tumour regression.[8, 9] As such, this evidence provides a compelling biological rationale for the effectiveness of nivolumab and the blocking of PD-1 as a therapeutic target in UC.[10, 11]

B.3 Cost effectiveness

Summary of the cost-effectiveness analysis

- An economic SLR identified no previous economic evaluations for nivolumab as a treatment for locally advanced or metastatic UC hence a de novo cost-utility model was constructed for the purposes of this appraisal
- The model used a partitioned survival approach and included three health states: *progression-free, progressed disease*, and *death*, consistent with previous submissions to NICE in metastatic cancers, including the only previous submission in this specific indication
- Nivolumab was compared to paclitaxel, docetaxel and BSC with clinical data derived from an ITC.
- OS and PFS estimates for nivolumab were extrapolated from pooled CheckMate 275 and CheckMate 032 trial data using appropriate survival analyses; TTD was also used to determine the duration of treatment for nivolumab and comparators in the model
- A response-based modelling approach using landmark analysis was used; from week 8, separate curves were fitted to the responding and non-responding patients, respectively
- Health-state utilities for the *progression-free* and *progressed disease* states were derived from EQ-5D-3L data collected from patients in the CheckMate 275 trial; disutilities for AEs were also included
- Resource use and costs included in the model were based on information from CheckMate 275, published sources identified in the SLR and expert clinician feedback.
- Feedback from UK clinicians and health economists was sought in order to validate assumptions and inputs included in the model

Base case cost-effectiveness results

- Nivolumab was found to be associated with higher costs but also higher QALYs than paclitaxel, docetaxel or BSC
- In the base case analysis, nivolumab was associated with ICERs of between £37,647 and £44,960 per QALY gained when nivolumab was provided with the confidential PAS; these ICERs are well below the cost-effectiveness threshold of £50,000 per QALY considered for therapies meeting end-of-life criteria.

Sensitivity analyses

- ICER estimates obtained from probabilistic sensitivity analysis to take account of combined uncertainty in the model were similar to the base case deterministic ICERs
- Of parameters explored in deterministic sensitivity analysis, the cost of nivolumab and patient age were found to be the most influential parameter on the ICERs.
- Scenario analyses were conducted to explore the impact of different time horizons and alternative parametric distributions for OS, PFS and TTD, amongst other sensitivity analyses. The analyses indicate that the choice of parametric distribution is a key driver of the overall results. In the majority of the other scenarios a reduction in the ICER was identified.

B.3.1 Published cost-effectiveness studies

An SLR was conducted to identify evidence to support the development of a cost-effectiveness model for nivolumab as a treatment for locally unresectable or metastatic UC. A single review was performed to identify relevant studies in UC that included published economic evaluations, studies reporting cost/resource use data, and studies reporting utility values.

Full details of the search strategy and results of the economic SLR are presented in Appendix G. A total of 576 articles were identified in the SLR, of which 9 records were ultimately included, comprising 3 unique economic evaluations, 6 unique utilities studies, and 2 unique cost/resource studies.

The three economic evaluations identified in patients with locally advanced unresectable or metastatic UC included a cost-utility analysis of cisplatin plus gemcitabine versus MVAC, and the economic evaluations submitted as part of the NICE and SMC appraisals for vinflunine versus BSC. Whilst some of these economic evaluations were performed in populations that match the final scope of this appraisal, they do not consider the cost-effectiveness of nivolumab and therefore a de novo health economic analysis was conducted for the purposes of this appraisal.

Full details of the 3 economic evaluations included in the SLR and the quality assessments of these economic evaluations can be found in Appendix G.

B.3.2 Economic analysis

B.3.2.1 Patient population

The de novo economic model considers patients with metastatic or unresectable UC who have progressed following first-line platinum-based chemotherapy. This patient group is consistent with the population of the CheckMate 275 and CheckMate 032 trials, as well as the final scope issued by NICE for this appraisal.

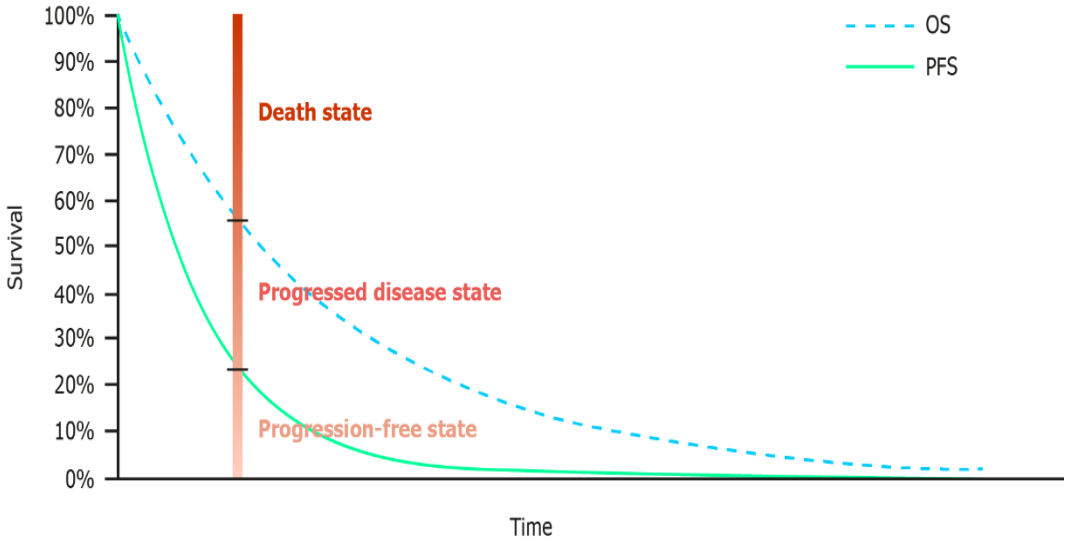
B.3.2.2 Model structure

A cohort-based partitioned survival model was developed that included three mutually exclusive health states: progression-free (PF), post-progression (PP) and death. The model structure is presented in Figure 33. Patients enter the model in the PF state and receive one of the five treatment options included in the model. Patients remain in the PF state until either disease progression or death. Movement between states occurs at the end of each cycle; however, to reflect the fact that patients may, in fact, progress or die at any point, a half cycle correction has been applied (half-cycle correction not used when estimating treatment costs as the majority of costs incurred at the start of each cycle). The death state is absorbing such that patients cannot leave it once they have entered. The proportion of patients in each state therefore changes over time, as determined by the PFS and OS curves, which are treatment dependent. The PFS curve determines the number of patients in the PF state, the OS curve determines the number of patients in the death state and the difference between the two curves determines the number of patients in the PP state (Figure 33). More details on the survival analysis are provided in Section B.3.3.

This choice of model structure was made to capture the progressive nature of UC disease and is consistent with previous submissions to NICE relating to metastatic cancers, including the only previous submission in this specific indication (TA272, 2013).[27]

The model was constructed from the perspective of the NHS and Personal and Social Services (PSS) in England and Wales. Four-week cycles were adopted to account for the length of treatment cycles and 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. The key features of the economic model are summarised in Table 28 below.

Figure 33: Schematic representation of the partitioned survival method



Abbreviations: OS: overall survival; PFS: progression-free survival.

B.3.2.3 Intervention technology and comparators

Nivolumab was included in the analysis as per the licensed indication for second-line UC (i.e. 3 mg/kg Q2W).[13] Four comparators were included in the analysis: docetaxel, paclitaxel, cisplatin plus gemcitabine and BSC. Docetaxel, paclitaxel and BSC were specifically named in the NICE final scope for this appraisal, with retreatment with first-line platinum-based chemotherapy also listed.

As highlighted in Section B.2.9 and Appendix D, there is limited evidence for retreatment with first-line platinum-based chemotherapy regimens for patients with locally advanced unresectable or metastatic UC and no relevant trials for this comparator were identified in the clinical SLR. As described in Section B.2.9.1, data for one of the first-line platinum-based regimens used in re-challenge, cisplatin plus gemcitabine, were identified in the SLR and were included within the ITC and the economic model as a scenario analysis. This comparison however suffers from a number of limitations in relation to the generalisability of the trials to the decision problem of this submission. The trials were conducted in a gemcitabine-naïve patient population, making them non-generalisable to UK clinical practice where cisplatin plus gemcitabine is the standard of care in the first-line setting for locally advanced unresectable or metastatic UC.[61] Therefore, given the lack of generalisability and relevance of this comparison, the results are presented only briefly in Appendix O.

The following treatment regimens were implemented in the economic model based on their anticipated use in clinical practice in England and Wales:

- Paclitaxel: 80mg/m^2 Q3W of a four week cycle
- Docetaxel: 75mg/m^2 Q3W

For the scenario analysis versus cisplatin plus gemcitabine, data for were only available from the Gondo *et al*. (2011) trial where it was administered at a dose not reflective of clinical practice in the UK.[61] The dose administered was gemcitabine 1,000 mg/m^2 on days 1, 8, and 15, plus cisplatin 35 mg/m^2 on days 1 and 2. In UK clinical practice, cisplatin plus gemcitabine is given in the first-line setting as gemcitabine (1250mg/m^2) plus cisplatin (70mg/m^2) on days 1 and 8 of a

21 day cycle (cisplatin on day 1 only).[29] This, combined with the fact the patient population in Gondo *et al.* (2011) are gemcitabine-naïve, severely limits the applicability of this data and comparison to the decision problem being considered.[61]

A PAS involving a [redacted] discount to the unit list price of nivolumab was also applied in the analysis.

Table 28: Features of the economic analysis

	Previous appraisal	Current appraisal	
Factor	**TA272[27]**	**Chosen values**	**Justification**
Time horizon	5 years	Lifetime	To capture all relevant health consequences and costs
Treatment waning effect?	Not included	Not included	Evidence from immunotherapy trials provide evidence that a continued treatment benefit is observed for some patients up to 10 years[74]
Source of utilities	• Pre-progression utility values were based on trial data (EORTC QLQ-C30 questionnaire) • Post-progression utility values were taken from a study reporting EQ-5D values in terminally ill patients with painful bone metastases or poor prognosis non-small-cell lung cancer • Disutility values associated with treatment-related adverse events were not included	• EQ-5D data from the CheckMate 275[38] trial were used and adjusted using multiple imputation and a mixed model regression • Disutility values associated with treatment-related adverse events were taken from the literature	• The CheckMate 275 trial was deemed the best source following the literature review, which identified only one previous study to use the EQ-5D in this indication.[75] This study was not deemed appropriate due to the use of US population weights, which do not match the NICE reference case • Multiple imputation was used to impute missing values and a mixed model regression analysis used to account for autocorrelation
Source of costs	NHS reference costs, literature and expert opinion.	• Therapy costs were taken from the BNF and eMit • Administration and resource use costs were taken from NHS reference costs and supplemented with evidence from the literature and an advisory board where necessary	• Unit costs were taken from recognised national databases • Clinicians provided advice on resource use for a number of parameters due to a paucity of relevant data in the wider literature[29]

Abbreviations: BNF: British National Formulary; eMIT: electronic market information tool; EORTC QLQ-C30: European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30; EQ-5D: EuroQoL 5-Dimensions; NHS: National Health Service; NICE: National Institute for Health and Care Excellence; US: United States.

B.3.3 Clinical parameters and variables

Parametric time-to-event survival curves were plotted to estimate the long-term outcomes (i.e. PFS and OS) with each treatment option. In accordance with guidance from the NICE Decision Support Unit (TSD 14)[76] the following six distributions were plotted for both PFS and OS for each treatment included in the model:

- Exponential
- Weibull
- Gompertz
- Lognormal
- Log-logistic
- Generalised gamma.

B.3.3.1 Survival analysis: nivolumab

Previous appraisals of nivolumab have highlighted that due to the innovative mechanism of action of immunotherapy, the standard survival modelling approaches outlined in the TSD14 may not accurately reflect the mechanism of action of immunotherapy as the advice was published before immunotherapy drugs were available.[77] Therefore, novel approaches to modelling survival curves with the characteristics of immunotherapy treatment were explored. For the base case analysis, a response-based modelling approach was adopted in order to characterise the mechanism of action of nivolumab whereby a subset of patients exhibit a long and durable response to treatment and survival. Due to the nature of extrapolation of data via parametric models, such as those listed above, simply modelling all nivolumab patients as one may fail to fully characterise the additional benefits received by these responders. This is because the shape of the nivolumab survival curve changes over time as the hazard changes and standard parametric models are unlikely to be flexible enough to characterise this change. The response-based approach works by fitting parametric survival curves to the responding and non-responding patients separately to more accurately characterise the hazard and survival curve in these two groups.

When using a response-based modelling approach, there can be a risk of immortal time bias, which occurs when the responder and non-responder curves are plotted immediately following the start of treatment (i.e. month 0 in the survival analysis). This is because response to treatment does, in fact, not occur instantaneously and often takes a number of months. For responders, this means that they are unable to progress or die in the preceding time. Alternatively, non-responders risk progression or death at any point in this period. Therefore, if separate curves based on response are plotted from treatment initiation, then the difference in PFS and OS is likely to be exaggerated between the two curves thereby generating implausible results (i.e. given that responders artificially have no risk of death until response, the curve for responders may overestimate long-term survival).

To overcome immortal time bias, landmark analysis was undertaken. For this approach, PFS and OS for nivolumab is based on the full cohort of patients (i.e. not separated by response), using pooled Kaplan-Meier data from CheckMate 032 and CheckMate 275, up until a designated landmark point. After this landmark, separate responder and non-responder curves are plotted

for the remaining time horizon.[78] Two landmark points were included in the analysis: 8 weeks and 26 weeks. The 8-week landmark point was selected to reflect the median time to response (1.87 months and 1.48 months, based on RECIST v1.1 criteria) as measured in the CheckMate 275 and CheckMate 032 trials.[37, 38] Twenty-six weeks was chosen as a time point by which all patients had responded while leaving a sufficiently long observational period for further extrapolation. The use of a 52-week time point was also investigated, but a very small number of events occurred after this point making extrapolation difficult so this was excluded from any further analysis.

For the base case analysis, the 8-week landmark was chosen to better capture the benefits of those patients who received an active response to nivolumab. The impact of using the 26-week landmark was examined in a scenario analysis, and further details can be found in Appendix L.

Table 29 compares the model fit statistics for different survival distributions with the 8-week landmark analysis. Lower values of the Akaike information criterion (AIC) and Bayesian information criterion (BIC) indicate better fits. In total, there are eight criteria and two distributions accounted for the lowest scores for 7 out of 8 criterion thereby indicating they provide the best fit. These were the Weibull (lowest score for 4 criteria) and Generalised Gamma (lowest score for 3 criteria) distributions. However, the Weibull distribution was seen to be a particularly poor fit for responders whilst the magnitude of difference between the Weibull and Generalised Gamma distributions for non-responders was much smaller. Therefore, altogether the Generalised Gamma was deemed to provide the best fit so this was chosen for the base case analysis.

A visual inspection of the distributions also indicate that the Generalised Gamma distribution provides a close match to the observed OS data from CheckMate -275 and -032. The long-term extrapolation for OS with this distribution also closely matches the longer-term data that is available for nivolumab in another indication whereby patients were treated with nivolumab in the second-line setting after prior platinum therapy (i.e. NSCLC). More information is provided on this validation in Section B.3.10.

Nivolumab PFS and OS with the Generalised Gamma distribution are presented in Figure 34 and Figure 35, respectively. The alternative distributions are presented in Appendix L. Appendix L also contains further details of the methods applied for the landmark analysis, along with the outputs from the 26-week landmark analysis.

Table 29: Week 8 landmark model fit measures

Endpoint	Distribution	Responders		Non-responders	
		AIC	BIC	AIC	BIC
OS	Exponential	90.06	**92.35**	1402.65	1406.09
	Weibull	91.13	95.71	**1393.29**	**1400.18**
	Gompertz	91.86	96.44	1395.42	1402.30
	Lognormal	90.43	95.01	1397.41	1404.30
	Log-logistic	91.04	95.62	1394.43	1401.32
	Generalised gamma	**87.94**	94.81	1394.51	1404.84
PFS	Exponential	276.86	279.15	787.77	790.62
	Weibull	266.93	271.51	**763.44**	**769.15**

	Gompertz	273.11	277.69	780.65	786.35
	Lognormal	262.40	266.98	773.05	778.76
	Log-logistic	264.58	269.16	776.65	782.35
	Generalised gamma	**256.62**	**263.49**	764.96	773.51

Abbreviations: AIC: Akaike information criterion; BIC: Bayesian information criterion; OS: overall survival; PFS: progression-free survival.

Figure 34: Week 8 landmark - overall survival with generalised gamma

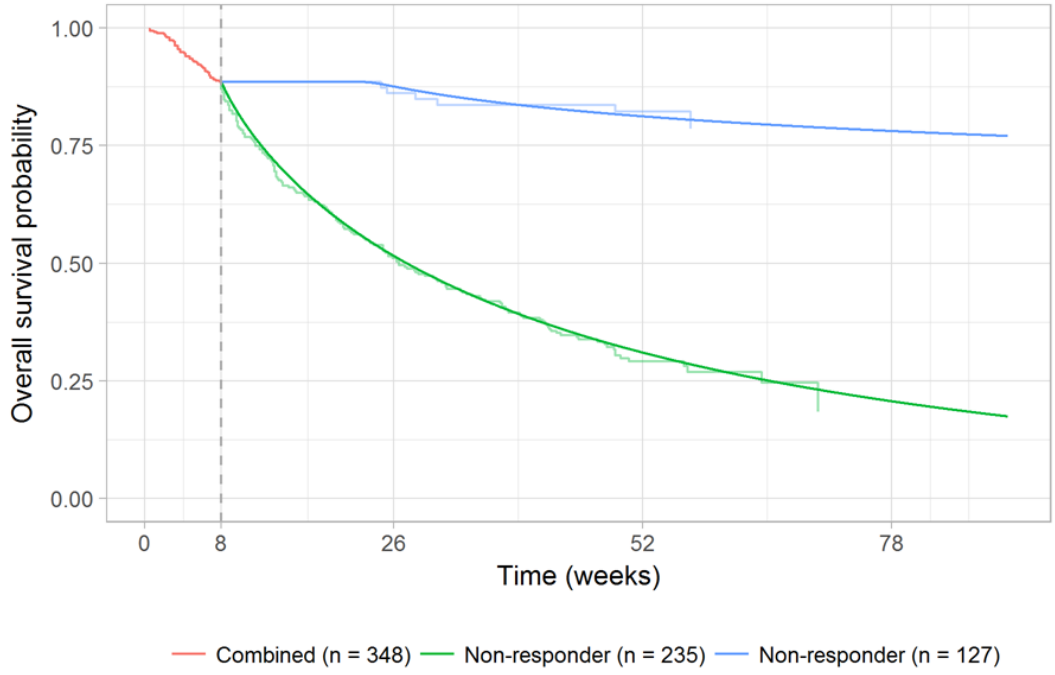
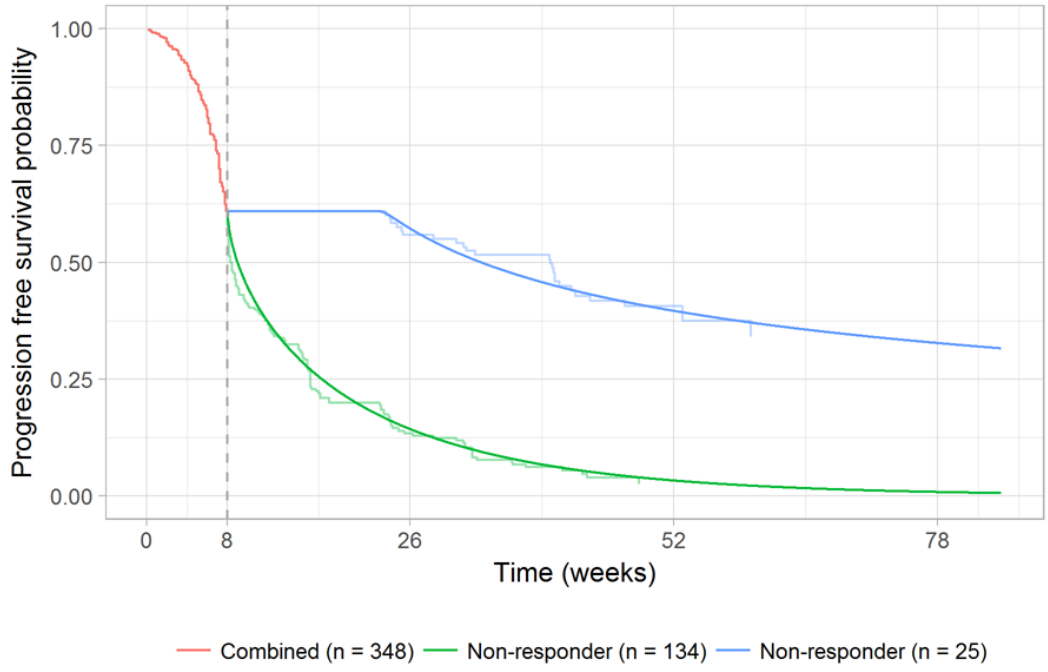


Figure 35: Week 8 landmark – progression-free survival with generalised gamma



To make the PFS and OS curves suitable for the structure of the economic model, and the application of relative treatment effects, it was necessary to combine the separate responder and non-responder curves. This generates a combined curve that can used to estimate PFS and OS for all nivolumab-treated patients and the comparator treatments. To generate combined curves, the separate responder and non-responder curves were weighted based on the number of patients measured as being progression-free and alive at the 8-week landmark point in the CheckMate 275 and CheckMate 032 trials. This weighting was assumed to remain constant for the remaining time horizon in each parametric model. This is likely to be a conservative assumption as the weighting would be expected to increase in favour of the responding patients across time, who die at a much slower rate than the non-responding patients. Therefore, the final PFS and OS curves for nivolumab are composites of the pre-landmark pooled Kaplan-Meier data from CheckMate 275 and CheckMate 032 and a weighted average of the responder and non-responder curves from the landmark point onwards.

One important consideration when modelling survival with short-term data over a long-term period is ensuring that curves appropriately characterise the relationship between age and increasing risk of death. To account for this, the PFS and OS curves were adjusted to account for general population mortality using age-adjusted annual mortality rates based on life tables for England and Wales.[79] Due to differences in the rate of mortality between males and females, the annual rates were weighted by gender based on the ratio of males to females (78:22) reported in the CheckMate 275 trial.[38] To avoid double counting, background mortality was only applied from week 88 onwards in the model, which is the end of the follow-up period in the CheckMate trials.

The final survival curves for nivolumab, using the Generalised Gamma distribution for the longer-term extrapolation, are presented in Figure 36 and Figure 37. The survival curves for all other distributions are presented in Appendix L.

Figure 36: Progression-free survival with nivolumab – observed and predicted values with the generalised gamma distribution

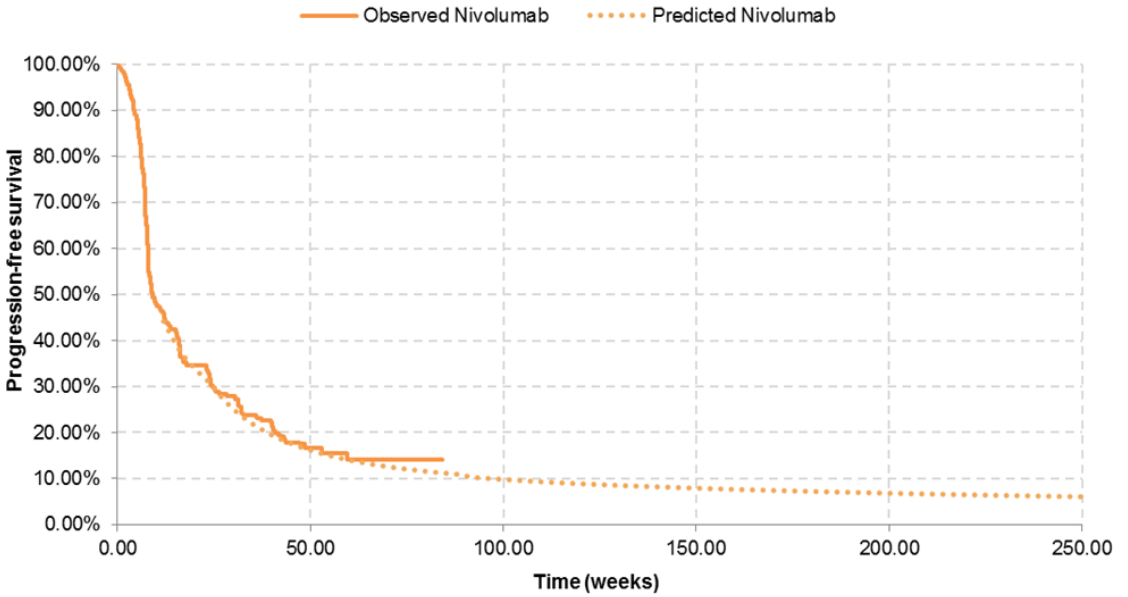
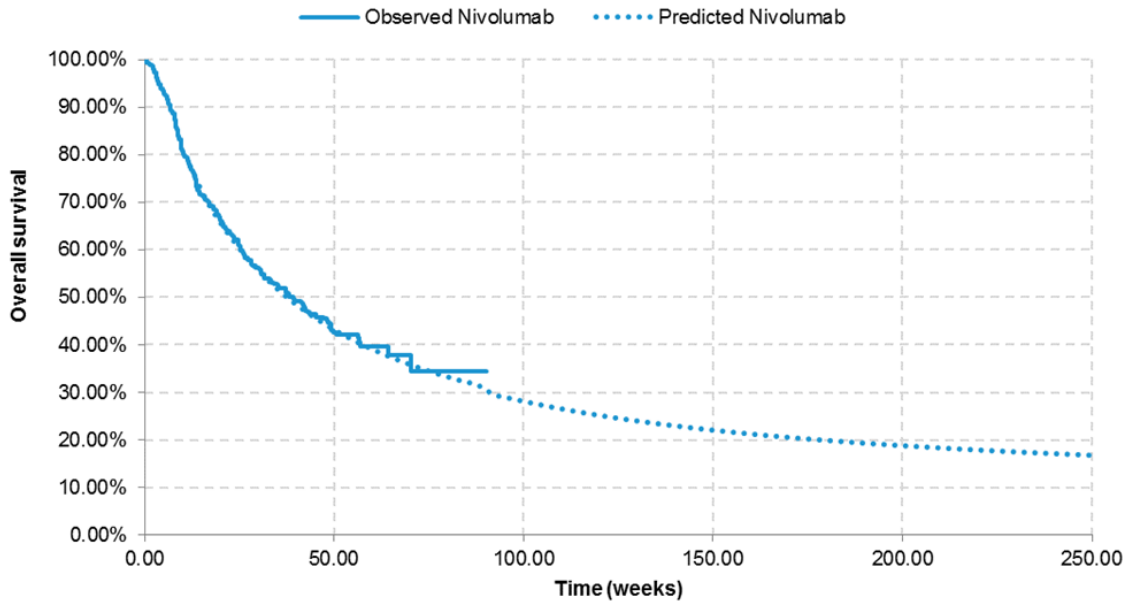


Figure 37: Overall survival with nivolumab – observed and predicted values with the generalised gamma distribution



B.3.3.2 Survival analysis: comparators

As described in Section B.2 and Appendix D, an STC was conducted to estimate time-varying HRs for nivolumab versus each comparator (both PFS and OS). It was necessary to generate time-varying HRs as the proportional hazard assumption did not hold for these comparisons given the unique mechanism of action for nivolumab. Within the economic model, a separate HR was applied for each cycle in the model.

For BSC, no relevant PFS data were identified during the clinical SLR and, therefore this was not included in the STC for PFS. Consequently, HRs for BSC were not available for PFS and it was necessary to make assumptions regarding the impact of BSC on patient PFS. It was determined that it would not be appropriate to apply the HR for either of the chemotherapy agents given the lack of active therapy for patients receiving BSC. However, Bellmunt *et al.* (2009) report the HR of vinflunine versus BSC for second-line UC patients (1.47).[33] Given the expected similarity in terms of outcomes between vinflunine and paclitaxel/docetaxel, the assumption was made that this HR could be applied to the paclitaxel PFS curve in order to estimate PFS with BSC. Paclitaxel was chosen over docetaxel as an analysis of drug usage in the UK indicates that, in the patient population under consideration, paclitaxel is used more commonly then docetaxel.[44] In the absence of a time-varying HR, the HR was assumed to remain fixed for the timeframe of the analysis. This simplifying assumption was required in the absence of alternative data.

In the scenario analysis for cisplatin plus gemcitabine, no relevant PFS data were identified during the clinical SLR and therefore this was not included in the STC for PFS. It was determined that treatment would be expected to have a similar PFS profile to paclitaxel/docetaxel given they are all chemotherapy agents. Therefore, the HR for paclitaxel versus nivolumab was also applied for cisplatin plus gemcitabine. This necessary simplification further limits the generalisability of this comparator (see Appendix O).

The base case survival curves for each comparator are presented in Figure 40 to Figure 43Figure 42. All other survival curves are presented in Appendix L. These figures indicate that the predicted survival curves for the comparators often underestimate survival when compared with the available trial data, particularly for docetaxel. This occurrence is due to the STC in which the predicted curves account for differences in the patient characteristics between the nivolumab studies and the comparator studies; it represents how the patients in the nivolumab studies might have responded to each comparator. The prediction model for PFS is based on ECOG PS, age, visceral metastases and liver metastases (see Appendix D). Overall, based on these characteristics, patients in the nivolumab studies were worse off at baseline than patients in the docetaxel studies. Hence the predicted docetaxel curve, in the nivolumab population, is lower than the observed docetaxel curve.

B.3.3.3 Time to discontinuation

The licence for nivolumab indicates that treatment should be administered as long as clinical benefit is observed or until treatment is no longer tolerated by the patient.[13] Therefore, discontinuation from treatment is not based solely on patient progression (which is common for oncology therapies). To estimate time on treatment, a parametric survival curve for time to treatment discontinuation (TTD) was plotted to predict the proportion of patients receiving treatment at each cycle. This curve was based on individual patient data (IPD) from the CheckMate 275 and CheckMate 032 trials and, in line with the approach to PFS and OS, six distributions were plotted. Table 30 compares the model fit statistics for the six distributions.

The generalised gamma distribution was chosen for the base case. This was to ensure consistency with the choice made for PFS and OS. Furthermore, whilst two distributions produced lower AIC/BIC scores (Gompertz and log-logistic), indicating a better fit, these two distributions also produced very long tails with a percentage of patients on treatment at 5 and even 10 years. This is not in keeping with the expected clinical use of nivolumab and therefore these distributions lack clinical validity.

The resulting survival curve is presented in Figure 38 and also includes the observed Kaplan-Meier data from CheckMate 275 and CheckMate 032. All other distributions are presented in Appendix L.

The impact of adopting alternative distributions was examined in the scenario analysis. As part of this analysis, the impact of incorporating a treatment stopping rule for nivolumab was also examined. This was to reflect the possibility that, due to the unique mechanism of action of immune-checkpoint inhibitors in restoring anti-tumour immunity, it may be feasible to stop treatment with nivolumab for patients who have not yet progressed and exhibit a durable response and maintenance of clinical benefit. Evidence to support the stopping of treatment for patients who are responding to nivolumab is available from the CheckMate 003 trial in which treatment was continued up to 96 weeks.[80] Ongoing responses after treatment cessation were observed in this trial for patients with advanced NSCLC who had completed 96 weeks of therapy with nivolumab (see Figure 39). In addition, clinical stopping rules have been explored as part of other appraisals by NICE for nivolumab.[81, 82] Scenario analyses were therefore conducted to explore the impact of a majority of patients stopping treatment after 2 years for those who were yet to discontinue. This was implemented in the model by discontinuing 75% of patients who were still receiving treatment at the end of 2 years with all other parameters remaining the same. This also allows for the possibility that a minority of patients and/or their clinician want them to continue treatment and this approach has been accepted in other immunotherapy appraisals.[83, 84]

For docetaxel, paclitaxel and cisplatin plus gemcitabine, treatment is administered until disease progression or unacceptable toxicity.[30, 32, 61] Therefore, time on treatment was based on the PFS curves discussed previously. For paclitaxel, UK clinical practice is to stop treatment at 6 cycles (for patients who have not already discontinued) as was implemented in the UK-based study by Jones *et al.* (2017) and was also confirmed by UK expert clinical feedback.[29, 32] Therefore, all paclitaxel patients in the model were assumed to have discontinued treatment by week 24. The PFS curves for the four treatments are provided in Figure 40: Progression-free survival and overall survival with paclitaxel – observed and predicted values with the generalised gamma distributionto Figure 43. It was assumed that all BSC patients receive this treatment until death.

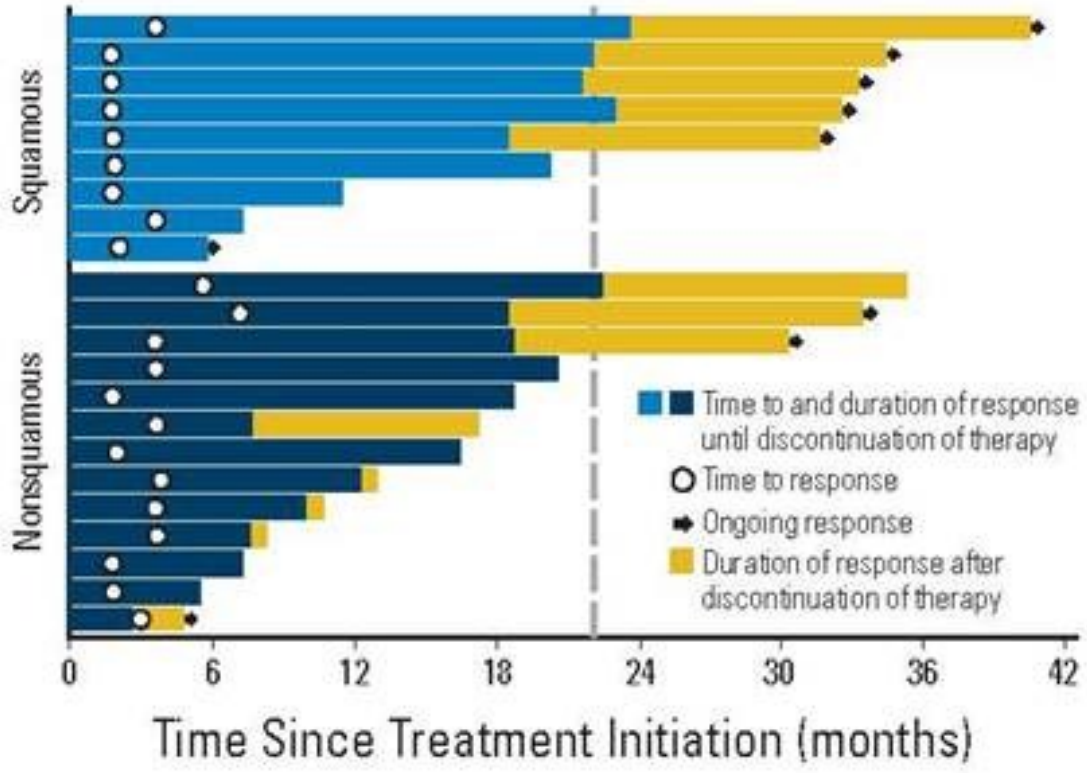
Table 30: Time to discontinuation model fit measures

Endpoint	Distribution	AIC	BIC
Time to discontinuation	Exponential	2381.86	2385.71
	Weibull	2329.96	2337.67
	Gompertz	2318.29	2325.99
	Lognormal	2341.69	2349.40
	Log-logistic	2322.93	2330.63
	Generalised gamma	**2328.48**	**2340.04**

Abbreviations: AIC: Akaike information criterion; BIC: Bayesian information criterion.

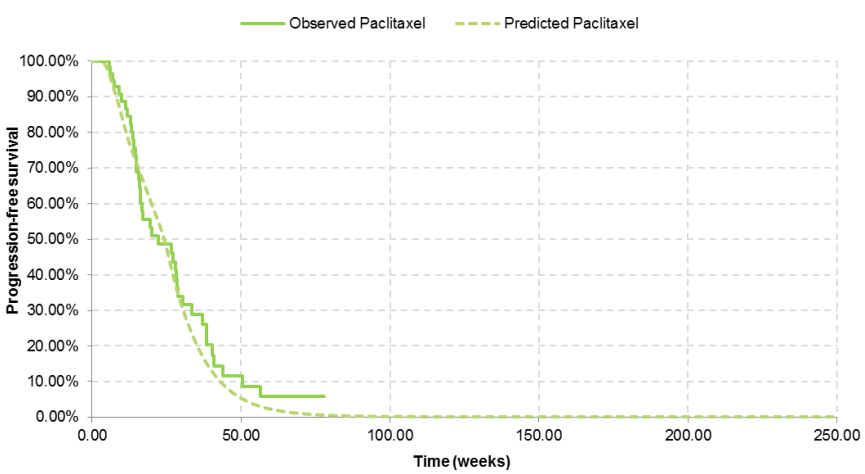
Figure 38: Time to discontinuation for nivolumab

Figure 39: Swimmers plot from CheckMate 003



Squamous NSCLC (n=9) and non-squamous NSCLC (n=13) treated with nivolumab
Vertical dashed line at 22 months indicates maximum planned duration of continuous nivolumab therapy.
Eighteen responders discontinued nivolumab therapy for reasons other than disease progression, including: completion of maximum cycles (n=7), adverse events (n=8), withdrawal of consent (n=2), and other (n=1)
Abbreviations: NSCLC: non-small cell lung cancer.
Source: Gettinger *et al.* (2015).[80]

Figure 40: Progression-free survival and overall survival with paclitaxel – observed and predicted values with the generalised gamma distribution



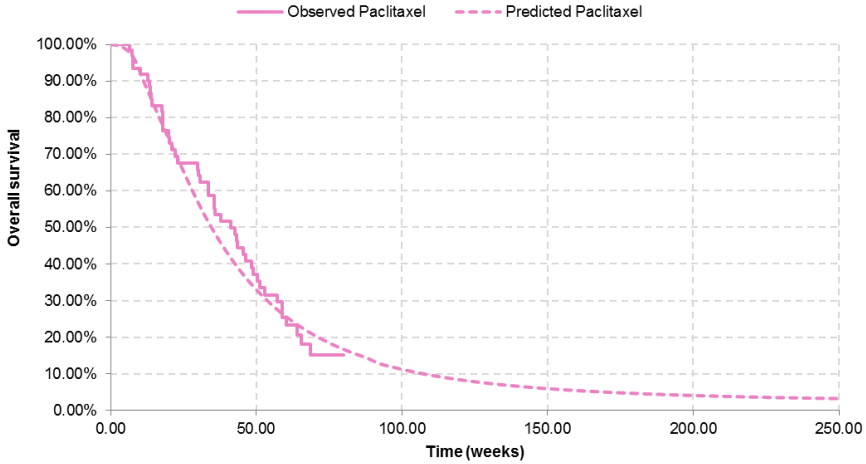
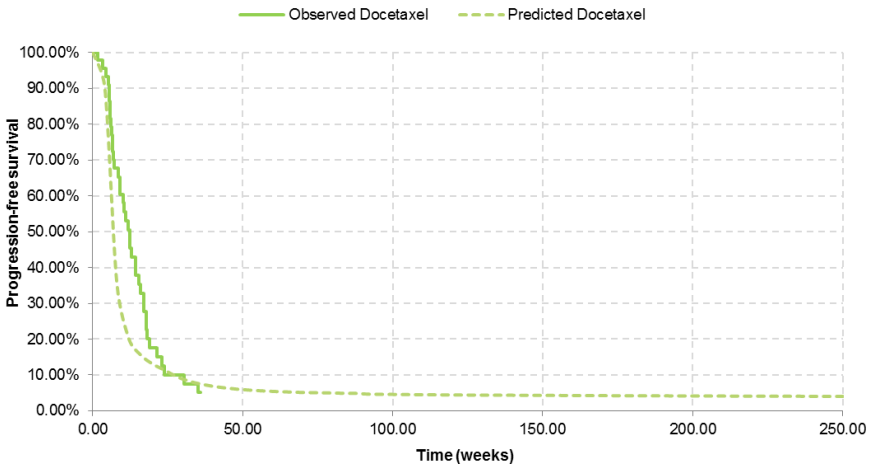


Figure 41: Progression-free survival and overall survival with docetaxel – observed and predicted values with the generalised gamma distribution



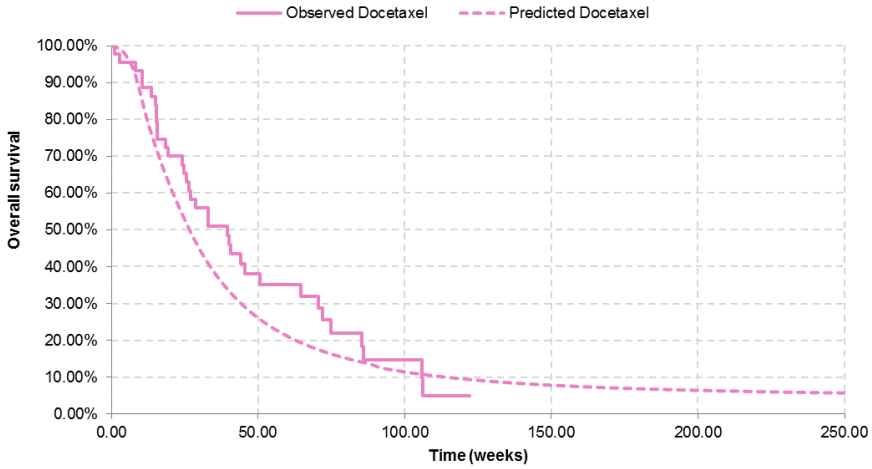
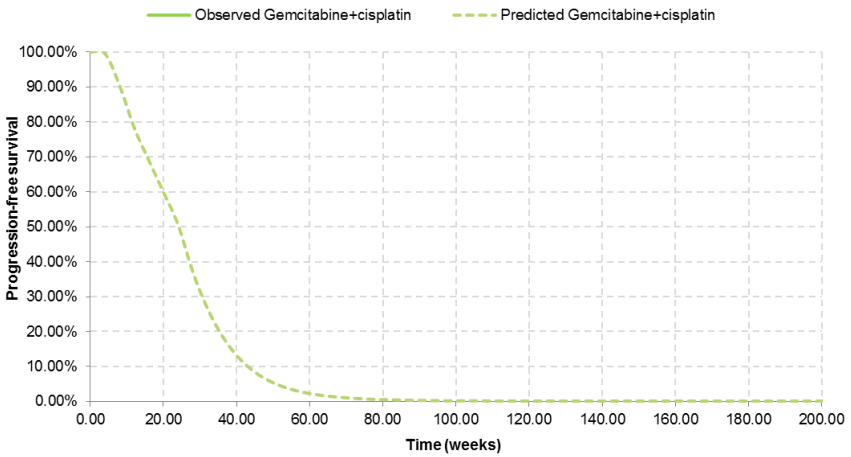
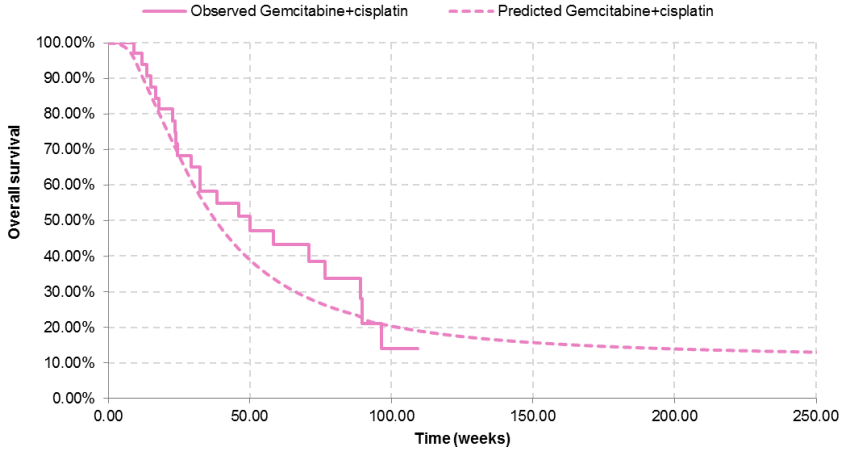


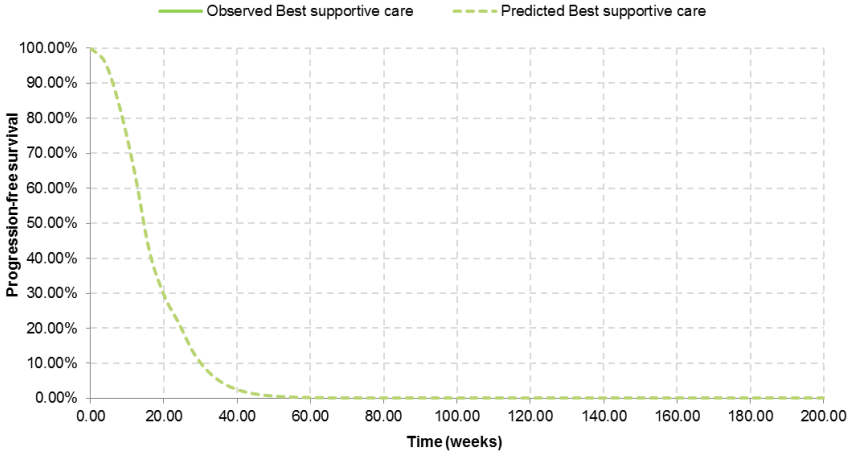
Figure 42: Progression-free survival and overall survival with best supportive care – observed and predicted values with the generalised gamma distribution

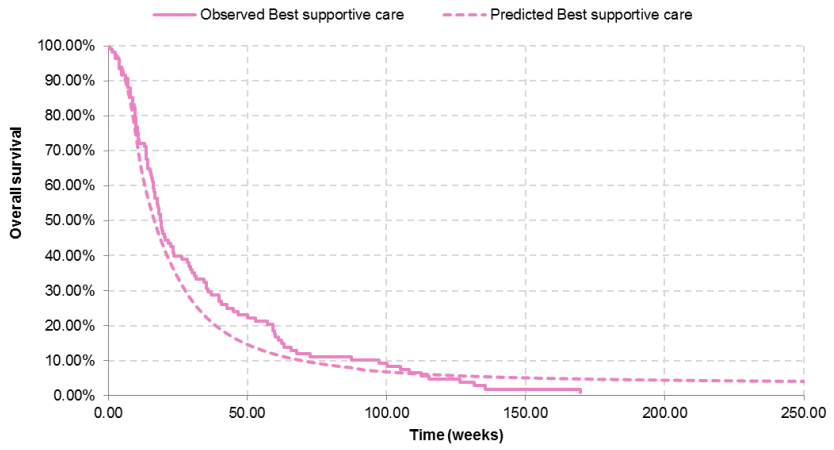




No observed progression-free survival data were identified for best supportive care

Figure 43: Progression-free survival and overall survival with cisplatin plus gemcitabine – observed and predicted values with the generalised gamma distribution[a]





[a]Cisplatin plus gemcitabine treatment was analysed as a scenario analysis. No observed progression-free survival data were identified for cisplatin plus gemcitabine

B.3.3.4 Adverse events

Adverse event rates were included in the model to capture the toxicity of each treatment option. The rates were taken from the clinical trials that inform the PFS and OS curves in the model. Any all-cause Grade 3 or 4 AEs were included if the incidence was ≥5% and the impact on costs and utilities were incorporated in the first cycle of the model only. The adverse event rates are summarised in Table 31.

Table 31: Adverse event rates

Therapy	Neutropenia	Anaemia	Thrombocytopenia	Asthenia	Nausea/vomiting	Diarrhoea	ALT increase	Leukopenia	Source
Nivolumab	1.00%	1.48%	NR	1.48%	0.37%	1.85%	0.74%	0.00%	Checkmate 275[38]
Docetaxel	14.00%	1.00%	NR	6.00%	NR	0.00%	0.00%	0.00%	Choueiri *et al.* (2012)[31]
Paclitaxel	6.00%	0.00%	0.00%	5.00%	0.00%	2.00%	2.00%	0.00%	Jones *et al.* (2017)[32]
BSC	0.90%	8.10%	0.90%	17.90%	0.90%	NR	NR	NR	Bellmunt *et al.* (2009)[33] Bellmunt *et al.* (2013)[85]
Cisplatin plus gemcitabine[a]	66.67%	42.42%	33.33%	0.00%	0.00%	NR	NR	45.45%	Gondo *et al.* (2011)[61]

[a]The trials informing the comparison versus cisplatin plus gemcitabine were conducted in gemcitabine-naïve populations, and therefore cannot be considered to provide relevant data for the retreatment with first-line platinum-based chemotherapy. This comparison has been briefly included in Appendix O as a scenario analysis only and results should be interpreted with caution.
Abbreviations: ALT: alanine transaminase; BSC: best supportive care; NR: not reported

B.3.4 Measurement and valuation of health effects

B.3.4.1 Health-related quality-of-life data from clinical trials

As detailed in Section B.2.3, HRQoL data were collected in CheckMate 275 using the EQ-5D-3L questionnaire. Patients were scheduled to complete the EQ-5D-3L questionnaire before each dose at Week 1, then every 8 weeks up to 48 weeks, then every 12 weeks until disease progression or treatment discontinuation (whichever occurred later). Two follow-up visits also scheduled for HRQoL measures.

The EQ-5D-3L data from CheckMate 275 were therefore used to inform utility estimates in the health economic model. Details on the methods used to derive the utility values from the EQ-5D-3L questionnaire data and take account of any missing data are described below in Section B.3.4.5.

B.3.4.2 Mapping

Utility data in the model were based on EQ-5D data collected in CheckMate 275 using UK preference weights. Therefore, no mapping was required.

B.3.4.3 Health-related quality-of-life studies

As stated previously, EQ-5D-3L data were collected from the CheckMate 275 trial and were therefore used to derive the utility values for the economic model. In line with the NICE guide to the methods of technology appraisal, an SLR to identify relevant utility studies was performed. Full details of the search strategy and results can be found in Appendix G.

A total of 9 records (6 unique studies) were included in the SLR that reported health-state utility values for patients with locally advanced or metastatic UC.[27, 75, 86-92] In the vast majority of these studies, EQ-5D health state descriptions were not used, and full details of the elicitation and valuation methods were not reported. As such, none of the included utility studies were deemed to be consistent with the NICE reference case for consideration for use within the health economic model. Further details of these studies are presented in Appendix H.

B.3.4.4 Adverse reactions

Therapies for advanced cancer can be associated with a number of toxicities that are likely to have a significant impact on patients' quality of life. Grade 3 and 4 AEs are included in the model in order to represent those AEs that are more likely to have a substantial effect on quality of life. Disutilities were sourced from the literature and included in the model. These are shown in Table 35; these were applied as a one-off event at the start of treatment.

B.3.4.5 Health-related quality-of-life data used in the cost-effectiveness analysis

A proportion of patients in the CheckMate 275 trial did not complete the EQ-5D questionnaire at each time point, resulting in an incomplete dataset. Therefore, it was necessary to investigate the extent of, and reasons for, the missing data and take steps to mitigate against potential biases created by it. The analyses undertaken is summarised below.

The majority of the 270 patients in the CheckMate 275 trial receiving nivolumab were required to provide questionnaires at nine scheduled time points: at baseline, then every eight weeks up to

48 weeks, and two further follow-ups at 60 and 72 weeks. However, some patients were followed up at four points, the timings of these varied from patient to patient. To simplify the analysis, the patients with measurements at the four time points were mapped onto the nine-point follow-up schedule using their questionnaire results if the date on which they filled them out fell within three weeks either side of scheduled follow-up. For example, the results of a patient followed up between weeks 29 and 35 are mapped to scheduled follow-up at week 32.

At baseline, 96% of patients completed the EQ-5D questionnaire, as described in Section B.2.6.6, with the completion rate showing a slight decline over time, though remaining over 70% until 49 weeks. The completion rates over time are shown in Table 32.

Table 32: EORTC QLQ-C30 and EQ-5D-3L questionnaire completion rates over time (total enrolled population)

Assessment	EORTC QLQ-C30[a]		EQ-5D-3L[a]	
	n/N	%	n/N	%
Week 1 (baseline)	262/270	97.0	261/270	96.7
Week 9	144/167	86.2	144/167	86.2
Week 17	98/116	84.5	97/116	83.6
Week 25	76/91	83.5	75/91	82.4
Week 33	54/70	77.1	54/70	77.1
Week 41	24/32	75.0	24/32	75.0
Week 49	**6/7[b]**	**85.7**	**6/7**	**85.7**

[a]Completion rates = patients who completed the PRO with ≥1 score at the assessment time point/expected population (total population minus patients who have died or dropped out)
Abbreviations: EORTC QLQ-30: European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30; EQ-5D-3L: EuroQoL 5-Dimensions 3-Levels; PRO: patient reported outcomes.
Source: Necchi *et al.* (2017).[54]

After taking into account that not all patients were alive to complete questionnaires at each follow-up time point, a complete set of utility data would amount to 1,465 completed questionnaires. In total, data was imputed for 677 of observations (46.2%), although it should be noted that this included 204 questionnaires that were due to be administered after the date at which the dataset was finalised (2nd September 2016) and a further 285 relating to patients who had already discontinued treatment for reasons other than death. This means that the number of missing responses for patients still on treatment during the trial was 188 (20.4%).

To ascertain whether these missing data systematically bias estimates of the utility scores, the characteristics of the missing observations were compared with those with complete data. As shown in Table 33, the missing observations are slightly more likely to be female (25.4% versus 21.4% in the complete cases) but do not differ by age or EQ-5D score at baseline. However, the missing group are considerably more likely to be in the progressed disease state (67.2% versus 29.5%) and to have discontinued treatment (61.2% versus 20%). This was corroborated by a logistic regression model that estimates the effects of patient characteristics on the odds of data being missing, which showed that progression status and treatment status are statistically significant predictors of missing data. This is sufficient evidence to suggest that the data, using Rubin's taxonomy, may not be missing completely at random and so the imputed dataset was used to ensure no potential for bias.[93]

Table 33: Characteristics of observations with missing and complete EQ-5D questionnaire data

Variable	Missing	Complete
Age	64.9	65.3
% male	74.6%	78.6%
% progressed	67.0%	29.5%
% on treatment	38.8%	80.0%
Baseline EQ-5D	0.734	0.7312

Abbreviations: EQ-5D: EuroQoL 5-Dimensions.

In order to impute these missing values, the multiple imputation by chained equations (MICE) procedure was conducted. MICE is an extension of the multiple imputation method, and uses regression analysis to simultaneously impute values for all the variables in the dataset in one procedure. If $z_1, z_2, \ldots z_k$ represent all the variables with missing values, then MICE begins by regressing z_1 on all other variables in the dataset, including $z_2, \ldots z_k$. Missing values for z_1 are imputed by randomly drawing from probability distributions around the coefficients from the regression equation to create a predicted value. This process is then repeated for the next variable with missing values, z_2, this time incorporating the predicted values for z_1, and so on. Once this procedure has taken place for all the variables with missing values, it is run m times to create a series of m imputed datasets. It has been recommended in the literature that m should increase with the proportion of data missing; therefore, m is set to 40 in the analysis.[94] As the categorical variables are being imputed in the EQ-5D dimension scores, predictive mean matching (PMM) is used to predict missing values. Whereas the process described above imputes missing values using the random coefficient draws $\boldsymbol{\beta}^*$ directly, such that $\boldsymbol{z_i} = \boldsymbol{\beta}^* \boldsymbol{x_i}$, PMM identifies a set of q observations from the total number of observations with complete data, p, that match $\boldsymbol{\beta}^* \boldsymbol{x_i}$ as closely as possible, such that $|\widehat{\boldsymbol{\beta}} \boldsymbol{x_j} - \boldsymbol{\beta}^* \boldsymbol{x_i}|$ is minimised, where $j \in (1,2 \ldots p)$. Of these q observations, one is selected at random and becomes the imputed value.

Whilst it is possible to extract point estimates and standard errors for the mean utility scores of pre-progressed and progressed patients at this stage, there is an additional risk that these estimates may be biased by the presence of autocorrelation. This relates to the fact that the EQ-5D scores for each patient over time will be correlated with each other. To account for this hierarchical data structure, a linear mixed effects model was used, in which a random effect was assumed for each patient and a fixed linear effect for progression status:

$$h_{is} = \alpha + \beta p_{is} + \pi_s + \varepsilon_{is}$$

Where h_i is the EQ-5D score of observation i and patient s, p_i the progression status, s_i a patient identifier and ε_i the idiosyncratic error. β and π represent the fixed and random effects, respectively. The constant term α therefore provides the EQ-5D score of an individual in the pre-progressed state whilst β is the effect of progression on EQ-5D.

To reflect the uncertainty pertaining to the imputed values, the linear mixed effects model described above was run on each of the imputed datasets. Therefore m estimates of α and β were generated from which pooled estimates $\hat{\alpha}$ and $\hat{\beta}$ are calculated by taking the average across the datasets. For example, $\hat{\beta}$ was calculated by the following:

$$\hat{\beta} = \frac{1}{m} \sum_{i=1}^{m} \hat{\beta}_i$$

When estimating the variance, within-dataset variation is combined with between-dataset variation using Rubin's rules:

$$Var(\hat{\beta}) = \left[\left(\frac{1}{m}\right)\sum_{i=1}^{m} \boldsymbol{W_i}\right] + \left[\left(1+\frac{1}{m}\right)\left(\frac{1}{m-1}\right)\sum_{i=1}^{m}(\hat{\beta}_i - \hat{\beta})^2\right]$$

Where $\boldsymbol{W_i}$ represents the variance-co-variance (VCV) matrices for all $i = (1,2 \ldots m)$. All analyses were conducted in R, with imputation implemented using the MICE package.

Table 34 reports the results of linear mixed effects models when run on both the complete case data and the imputed datasets. Whilst the pre-progression utilities are similar, the imputed data find a larger decrement from being in the progressed disease state, with a coefficient of -0.1148 versus -0.0608 in the complete case analysis. The EQ-5D scores using the imputed data are therefore 0.7182 for the pre-progression state and 0.6038 for post-progression patients. These utility values are assumed to remain constant for the full time horizon for the analysis.

The imputed dataset was also used to examine how the utility values changed over the duration of CheckMate 275, with utility values generated for pre-progression and post-progression patients at each follow-up time point in the trial. This was an attempt to account for potential changes in utility over time. However, the generated utility values for post-progression patients was seen to increase and decrease in a manner that would not be expected in clinical practice. Due to the implausible nature of the values, they were not used within the economic analysis. Nevertheless, the cycle by cycle results are provided in Appendix N.

Table 34: Regression output for the linear mixed models run on the imputed and complete case datasets

	EQ-5D – imputed (SE)	EQ-5D – complete (SE)
Constant	0.7182* (0.0165)	0.7124* (0.0168)
Progressed	-0.1148* (0.0291)	-0.0608* (0.0167)
AIC		-190.102
BIC		-171.47
N	1465	781

*p<0.001. AIC and BIC are not provided when pooling the model results over the imputed datasets
Abbreviations: AIC: Akaike information criterion; BIC: Bayesian information criterion; EQ-5D: EuroQoL 5-Dimensions; SE: standard error.

Table 35: Summary of utility values for cost-effectiveness analysis

State	Utility/disutility value: mean (standard error)	95% CI	Source
Pre-progression	Imputed value: 0.718 (0.016) Observed value: 0.713 (0.017)	Imputed value: 0.686 to 0.75 Observed value: 0.679 to 0.747	Imputed from Checkmate 275[38]
Change in utility – pre-progression to post-progression	Imputed value: -0.115 Observed value: -0.061	Imputed value: -0.143 to -0.087 Observed value: -0.123 to -0.055	Imputed from Checkmate 275[38]
Post-progression	Imputed value	N/A	Checkmate 275[38]

	0.603 (N/A) Observed value: 0.623 (N/A)		
Neutropenia	-0.18	NR	Attard *et al.* (2014)[95]
Anaemia	-0.09	-0.13, -0.06	Beusterien *et al.* (2010)[96]
Thrombocytopenia	-0.18	NR	Attard *et al.* (2014)[95]
Asthenia/Fatigue	-0.12	NR	Attard *et al.* (2014)[95]
Nausea/vomiting	-0.05	-0.08,-0.02	Nafees *et al.* (2008)[97]
Diarrhoea	-0.29	NR	Attard *et al.* (2014)[95]
ALT increase	-0.05	-0.07, -0.03	NICE TA347 (2015)[98]
Leukopenia	-0.09	NR	Frederix *et al.* (2013)[99]

Abbreviations: ALT: alanine transaminase; CI: confidence interval; N/A: not applicable; NR: not reported.

B.3.5 Cost and healthcare resource use identification, measurement and valuation

An SLR to identify relevant cost/resource use data was performed. Full details of the search strategy and results can be found in Appendix G. A total of 4 records reporting 2 unique studies were identified that reported cost/resource use data for the treatment of UC. Further details of these studies are presented in Appendix I.

Resource use and unit costs for the economic model were based on a number of sources, including: data from CheckMate 275, national databases, published sources identified in the SLR and clinical advice. These are described in more detail below. In the absence of any additional sources of evidence, assumptions were made for cost/resource inputs included in the model where necessary and were validated through discussions with clinicians.

B.3.5.1 Intervention and comparator costs and resource use

Drug acquisition costs

The unit cost of nivolumab was taken from the British National Formulary (BNF) with two formulations (40mg and 100mg) included to enable dosing for different patient weights.[100] A PAS was applied in the model. The drug costs for docetaxel, paclitaxel and cisplatin plus gemcitabine were taken from the electronic market information tool.[101] The unit costs for all treatments are presented in Table 38.

Drug dosing

The cost per administration is dependent on the total dose required. For nivolumab, the dosage is based on weight with a dose of 3 mg/kg required, as adopted in the CheckMate trials and in line with the licensed dose. For docetaxel, paclitaxel, and cisplatin plus gemcitabine, the dose is based on total body surface area (BSA). Weight and BSA vary across the population and to account for this variation specific categories were generated for both. The patient population was separated into the categories based on a normal distribution using mean and standard deviation values for weight and BSA as reported in the CheckMate 275 trial. A normal distribution was chosen as BSA has previously been shown to be normally distributed in UK cancer patients[102] and weight is also known to follow this distribution. The distributions are presented in Table 36 and

Table 37.

Table 36: Weight distribution

Weight (kg)	% population in group
40	1.12%
60	13.36%
80	42.08%
100	35.20%
120	7.79%
140	0.44%
160	0.01%
Total	**100%**

Table 37: Body surface area distribution

BSA (m^2)	% population in group
1.50	2.55%
1.75	20.67%
2.00	45.50%
2.50	31.11%
2.75	0.17%
Total	**100%**

Abbreviations: BSA: body surface area.

For each weight category, the total required dose was estimated and used to calculate the total number of vials needed to obtain that dose assuming there was no vial sharing but that pharmacists would use the optimal combination of vial sizes for nivolumab to reduce wastage. The assumption of drug wastage may be conservative as is it possible for pharmacies to introduce vial sharing, particularly if nivolumab is used across multiple indications. Therefore, vial sharing has been incorporated as a scenario analysis.

In the CheckMate trials, a proportion of patients received a delayed or missed dose. The average length of dose delay was 15.1 days and 11.6 days in CheckMate 275 and CheckMate 032, respectively.[44] Given the 14-day cycle length for nivolumab, it can be reasonably assumed that these delayed doses represented a missed dose. Therefore, it was necessary to adjust for dose intensity so these missed doses were not costed in the economic model. In total, there were 211 missed doses across the two CheckMate trials, out of a total of 3,214 doses, which equates to a dose intensity of 93.44%.[39] Therefore, it was assumed in the model that 93.44% of drug doses were received by nivolumab patients with the remaining 6.56% not administered. This approach has been adopted and accepted for other appraisals of nivolumab.[77, 103]

Whilst drug costs were not included for the 6.56% of nivolumab patients with missed doses, it was assumed that the cost of administrating the drug would still be incurred to factor in that the chair time would still have been reserved for the patient. This may be a conservative assumption as it is possible the appointment may be taken up by another patient. More information on administration costs is provided below.

For the comparators, no relevant data on dose intensity was identified following a targeted search of the literature. In the absence of alternative data it was assumed that the dose intensity for docetaxel, paclitaxel, cisplatin plus gemcitabine was equal to that of nivolumab. Again, it was assumed that administration costs were incurred in 100% of patients. The impact on alternative dose intensity values was examined during sensitivity analyses.

Drug administration and monitoring

The cost of drug administration was derived from the NHS reference cost schedule 2015-16[104] and applied dependent on doses required per 4-week cycle. As all included drugs are administered intravenously, the same cost per event was applied. The cost of administration is summarised in Table 38.

Monitoring consists of regular follow-up visits with an oncologist and a series of ongoing diagnostic tests whilst patients remain on treatment. The type and frequency of visits/tests was based on the cycle length of each treatment regimen. Nivolumab patients were also assumed to require two thyroid function and pituitary function tests per cycle with these tests not required for the other treatments. The cost of monitoring is summarised in Table 38.

Computed tomography (CT) scans were also included as part of the monitoring regimen. Clinicians advised that the frequency of CT scans was linked to the underlying treatment regimen (i.e. the frequency of drug administration). Therefore, based on this advice, it was assumed that nivolumab and cisplatin plus gemcitabine patients require a CT scan every 8 weeks whilst docetaxel and paclitaxel patients require a scan every 9 weeks.

Monitoring resource use and unit costs are summarised in Table 38. The impact of these costs on the model results are examined in the sensitivity analyses.

Best supportive care

The resources provided to patients receiving BSC, including the frequency per month, were informed by clinical advice. They advised that BSC patients would receive a combination of supportive therapies (i.e. painkillers, steroids, bisphosphonates and blood transfusions) and GP and nurse visits. The cost of BSC is summarised in Table 39.

Patients receiving BSC at the start of the model time horizon are assumed to remain receiving the treatment until death. It is also assumed that patients on each of the other treatment options receive BSC from the period in which the original treatment is stopped until death.

Discontinuation and terminal care costs

A proportion of patients in the model receive radiotherapy and/or surgery following discontinuation of treatment (i.e. excluding BSC patients). The proportion receiving each resource is based on data from CheckMate 275 (9.3% for radiotherapy and 3.3% for surgery) with the values applied equally for nivolumab, docetaxel, paclitaxel and cisplatin plus gemcitabine. The resources are included as one-off costs that occur at the point of treatment discontinuation and these are presented in Table 40.

A one-off terminal care cost is also applied on entering the death health state with the same cost applied to all treatment groups (see Table 40). This cost is based on the average acute care and

community costs for cancer patients in their last eight weeks of life and is additional to all previous costs incurred by patients.[105]

Table 38: Intervention and comparator costs and resource use

Items	Nivolumab	Reference	Docetaxel (CI)	Reference	Paclitaxel (CI)	Reference	Cisplatin plus gemcitabine (scenario analysis only) (CI)	Reference
Technology cost (price per unit)	[redacted] (40mg) [redacted] (100mg) (with PAS)	BNF 2017	£12.47 (£12.36, £12.58)	eMit 2016	£8.50 (£8.38, £8.62)	eMit 2016	Gemcitabine = £178.56 (£160.58, £196.60) Cisplatin = £6.99 (£6.96, £7.02)	eMit 2016
Unit size (mg/unit)	40; 100	BNF 2017	80	eMit 2016	100	eMit 2016	Gemcitabine = 1000 Cisplatin = 50	eMit 2016
Pack size (no. of units	1	BNF 2017	1	eMit 2016	1	eMit 2016	Gemcitabine = 1 Cisplatin = 5	eMit 2016
Dose required (mg) per kg or m^2	3	CheckMate 275	75	eMit 2016 [101]	80	Jones *et al.* (2017)	Gemcitabine = 1000 Cisplatin = 70	Gondo *et al.* (2011)
Doses per cycle	2	CheckMate 275	1.33	Choueiri *et al.* (2011), Petrylak *et al.* (2016)	3	Jones *et al.* (2017)	Gemcitabine = 3 Cisplatin = 1	Gondo *et al.* (2011)
Mean total cost of treatment per cycle	With PAS: [redacted] Without PAS: [redacted]	Calculated	£38	Calculated	£51	Calculated	£1,262	Calculated
Administration cost per cycle	£397.88	NHS reference costs 2015-16, Deliver simple parental Chemotherapy at First attendance (SB12Z Outpatient)	£265.25	NHS reference costs 2015-16, Deliver simple parental Chemotherapy at First attendance (SB12Z Outpatient)	£596.82	NHS reference costs 2015-16, Deliver simple parental Chemotherapy at First attendance (SB12Z Outpatient)	£596.82	NHS reference costs 2015-16, Deliver simple parental Chemotherapy at First attendance (SB12Z Outpatient)
Monitoring cost – oncologist	£336	NHS reference costs 2015-16, Consultant led,	£221	NHS reference costs 2015-16, Consultant led,	£498	NHS reference costs 2015-16, Consultant led,	£498	NHS reference costs 2015-16, Consultant led,

Items	Nivolumab	Reference	Docetaxel (CI)	Reference	Paclitaxel (CI)	Reference	Cisplatin plus gemcitabine (scenario analysis only) (CI)	Reference
follow-up, cost/cycle	(unit cost £163, 2 per cycle)	non-admitted face to face attendance, follow-up, oncology (WF01A)	(unit cost £163, 1.33 per cycle)	non-admitted face to face attendance, follow-up, oncology (WF01A)	(unit cost £163, 3 per cycle)	non-admitted face to face attendance, follow-up, oncology (WF01A)	(unit cost £163, 3 per cycle)	non-admitted face to face attendance, follow-up, oncology (WF01A)
Monitoring cost – CT scan, cost/2 cycles	Unit cost = £115 Scan every 8 weeks; therefore, applied every other cycle	NHS reference costs 2015-16, diagnostic imaging, computerised tomography scan of two areas, with contrast, outpatient (RD24Z) Advisory board	Unit cost = £115 Scan every 9 weeks; therefore, unit cost adjusted to 8 weeks and applied every other cycle.	NHS reference costs 2015-16, diagnostic imaging, computerised tomography scan of two areas, with contrast, outpatient (RD24Z) Advisory board	Unit cost = £115 Scan every 9 weeks; therefore, unit cost adjusted to 8 weeks and applied every other cycle.	NHS reference costs 2015-16, diagnostic imaging, computerised tomography scan of two areas, with contrast, outpatient (RD24Z) Advisory board	Unit cost = £115 Scan every 8 weeks; therefore, applied every other cycle	NHS reference costs 2015-16, diagnostic imaging, computerised tomography scan of two areas, with contrast, outpatient (RD24Z) Advisory board
Monitoring cost – Blood tests, cost per cycle	Unit cost per blood test[a] = £1 5 tests twice per cycle = £10	NHS reference costs 2015-16, Directly accessed pathology services, Clinical biochemistry (DAPS04)	Unit cost per blood test[b] = £1 3 tests, 1.33 times per cycle = £4	NHS reference costs 2015-16, Directly accessed pathology services, Clinical biochemistry (DAPS04)	Unit cost per blood test[b] = £1 3 tests, three times per cycle = £9	NHS reference costs 2015-16, Directly accessed pathology services, Clinical biochemistry (DAPS04)	Unit cost per blood test[b] = £1 3 tests, three times per cycle = £9	NHS reference costs 2015-16, Directly accessed pathology services, Clinical biochemistry (DAPS04)

[a]Blood tests include full blood count, hepatic function test, renal function test, thyroid function test, pituitary function test. [b]blood tests include full blood count, hepatic function test, renal function test

Abbreviations: BNF: British National Formulary; CI: confidence interval; CT: computer tomography; eMIT: electronic market information tool; NHS: National Health Service; PAS: patient access scheme.

Source: Advisory board,[29] BNF 2017,[100] CheckMate 275,[38] Choueiri *et al.* (2011),[31] eMIT 2016,[101] Gondo *et al.* (2011),[61] Jones *et al.* (2017)[32] and Petrylak *et al.* (2016).[30]

Table 39: Best supportive care cost and resource use

Items	Resource use	Unit cost	Reference in submission
GP home visit	Proportion of patients = 50% Frequency per month = 2	£77.22	Curtis *et al.* (2015);[106] GP home visit duration Curtis *et al.* (2016);[107] Cost per minute
Community nurse specialist visit	Proportion of patients = 50% Frequency per month = 2	£69.20	NHS reference costs 2015-16,[104] Community health services, specialist nursing, cancer related, Adult, face to face (N10AF)
Blood transfusions	Proportion of patients = 10%	£170.14	NICE guideline NG24 Costing guidance, 2015[108]
Prednisolone	Dose (mg per day) = 10 Pack size (mg) = 140	£0.42	eMit 2016[101]
Morphine	Dose (mg per day) = 40 Pack size (mg) = 200	£0.73	eMit 2016[101]
Gabapentin	Dose (mg per day) = 300 Pack size (mg) = 30,000	£2.13	eMit 2016[101]
Alendronic acid	Dose (mg per day) = 10 Pack size (mg) = 280	£1.25	eMit 2016[101]
Total cost per cycle	-	**£170.21**	-

Abbreviations: GP: general practitioner; NHS: National Health Service; NICE: National Institute for Health and Care Excellence.

Table 40: Health-state unit costs and resource use

Health states	Items	Value	Reference in submission
Pre-progressed	Drug, administration and monitoring costs	Varies dependent on treatment.	See Table 38
	Adverse event costs	See Table 41	See Table 41
Post-progressed	Subsequent radiotherapy	Unit cost = £128.22 Proportion of patients = 9.3%	Applied to patients who have discontinued treatment NHS reference costs 2015-16,[104] weighted average of outpatient costs SC22Z and SC47Z CheckMate 275 CSR, Addendum, table S.5.7.[40]
	Subsequent surgery	Unit cost = £3,201.68 Proportion of patients = 3.3%	Applied to patients who have discontinued treatment NHS reference costs 2015-16,[104] Weighted average of LB19C and LB19D CheckMate 275 CSR, Addendum, table S.5.7.[40]
	BSC	£170.21 (per cycle)	Applied to patients who have discontinued treatment See Table 39

Death	Terminal care	£6,153	Addicott and Dewer (2008),[105] inflated to 2015-16 cost using PSSRU HCHS inflation index[107]

Abbreviations: NHS: National Health Service; PSSRU HCHS: Personal and Social Services Research Unit Hospital and Community Health Services.

Table 41: Adverse reaction unit costs and resource use

Adverse reactions	Value	Reference in submission
Neutropenia	£4,111	NHS reference costs 2015-16,[104] Weighted average of NEL PM45B, PM45C, PM45D
Anaemia	£2,971	NHS reference costs 2015-16,[104] Weighted average of NEL SA01G, SA01H, SA01J, SA01K
Leucopoenia	£1,207	Robinson *et al.* (2004).[92] Inflated to 2015-16 cost using PSSRU HCHS inflation index[107]
Nausea and vomiting	£1,907	NHS reference costs 2015-16,[104] Weighted average of NEL PF28C, PF28D, PF28E
Thrombocytopenia	£2,519	NHS reference costs 2015-16,[104] Weighted average of NEL SA12G, SA12H, SA12J, SA12K
Asthenia/Fatigue	£2,805	Brown *et al.* (2013).[91] Inflated to 2015-16 cost using PSSRU HCHS inflation index[107]
Diarrhoea	£490	Brown *et al.* (2013).[91] Inflated to 2015-16 cost using PSSRU HCHS inflation index[107]
ALT increase	£595	NICE ID900 (Nivolumab).[109] Inflated to 2015-16 cost using PSSRU HCHS inflation index[107]

Abbreviations: ALT: alanine transaminase; NICE: National Institute for Health and Care Excellence; PSSRU HCHS: Personal and Social Services Research Unit Hospital and Community Health Services.

B.3.6 Summary of base case analysis inputs and assumptions

B.3.6.1 Summary of base case analysis inputs

A summary of the base case model inputs and assumptions is provided in Table 42 and Table 43.

Table 42: Summary of variables applied in the economic model

Variable	Value (reference to appropriate table or figure in submission)	Measurement of uncertainty and distribution: CI (distribution)	Reference to section in submission
Starting age	65 years	SD 9.38 (normal)	-
Weight (kg)	Mean = 77.3 Normal distribution applied in model	SD 16.34 (normal)	Cost and healthcare resource use identification, measurement and valuation, Section B.3.5
Body surface area (m^2)	Mean = 1.90 Normal distribution applied in model	SD 0.205 (normal)	Cost and healthcare resource use identification, measurement and valuation, Section B.3.5

Proportion male	78%	-	Clinical parameters and variables, Section B.3.3
Days per cycle	28	-	Model structure, Section B.3.4
Overall survival	Varied by treatment		Clinical parameters and variables, Section B.3.3
Progression-free survival	Varied by treatment		Clinical parameters and variables, Section B.3.3
Discount rate – costs and benefits	3.5%	-	NICE reference case
Utilities	Pre-progression: 0.718 Post-progression: 0.603		Summary of utility values for cost-effectiveness analysis, Section B.3.4
Drug costs	Varies by treatment		Intervention and comparators' costs and resource use, Section B.3.5
Resource use	Varies by treatment/health state		Health-state unit costs and resource use, Section B.3.5
Adverse event rates	Varies by treatment		Adverse event rates, Section B.3.3
Adverse event costs	Varies by treatment		Adverse reaction unit costs and resource use, Section B.3.5
Adverse event utilities	Varied by treatment		Summary of utility values for cost-effectiveness analysis, Section B.3.4

Abbreviations: CI, confidence interval; SD, standard deviation.

B.3.6.2 Assumptions

Table 43: Assumptions

Assumption	Justification
Patients cannot return to the progression-free state from the progressive disease state.	Assumption made to minimise the complexity of the model and the available data. Consistent with clinical evidence.
Utility values do not change over time as long as the patient remains in the same health state.	Reflects available utility data.
Patients receiving nivolumab, docetaxel, paclitaxel and cisplatin plus gemcitabine are monitored until treatment discontinuation with the type of resources, and frequency of use, based on clinical advice.	No relevant data identified in the wider literature.
Patients receiving nivolumab, docetaxel, paclitaxel and cisplatin plus gemcitabine move onto BSC following treatment discontinuation, based on clinical advice.	No relevant data identified in the wider literature.
A proportion of patients receiving nivolumab, docetaxel, paclitaxel and cisplatin plus gemcitabine will receive subsequent surgery and/or radiotherapy at the point of discontinuation with the same value applied across treatments.	No evidence of a difference between treatment groups.
Vial sharing did not occur for any treatment option.	This is a conservative assumption to account for the fact that drug wastage may occur in some hospital pharmacies.
Doses being recorded as delayed in the CheckMate 275 and 032 trials would be missed and, therefore, do not incur a cost. This equates to a total of 6.56% of nivolumab doses. The same value was applied to each comparator.	The average length of dose delay was 15.1 days and 11.6 days in CheckMate 275 and CheckMate 032, respectively. Given the 14-day cycle length for nivolumab, it can be reasonably assumed that these delayed doses represented a missed dose. This approach has been adopted and accepted for other appraisals of nivolumab.[77, 103] No data on dose intensity found for the comparators.
Administration costs are incurred in the proportion of patients who do not receive treatment at a particular time point due to a missed dose.	This is a conservative assumption to account for the fact missed doses may still incur the cost of an administration because the appointment has to be booked in advanced and the hospital may not be able to allocate another patient to the appointment.
In the absence of a PFS data for BSC, the HR of vinflunine versus BSC for second-line UC patients from Bellmunt *et al*. (2009) was applied to the paclitaxel PFS curve in order to estimate PFS with BSC.	For BSC, no relevant PFS data were identified during the clinical SLR. It was determined that it would not be appropriate to apply the HR for either of the chemotherapy agents given the lack of active therapy for patients receiving BSC. However, Bellmunt *et al*. (2009) report the HR of vinflunine versus BSC for second-line UC patients (1.47).[57] Given the expected similarity in terms of outcomes between vinflunine and paclitaxel/docetaxel, this HR was applied to the paclitaxel PFS curve in order to estimate PFS with BSC. Paclitaxel was chosen as an analysis

	of drug usage in the UK indicates that, in the patient population under consideration, paclitaxel is used more commonly then docetaxel.[44] In the absence of a time-varying HR, the HR was assumed to remain fixed for the timeframe of the analysis. This simplifying assumption was required in the absence of alternative data.
In the absence of PFS data for cisplatin plus gemcitabine, the HR for paclitaxel versus nivolumab was also applied for cisplatin plus gemcitabine	No relevant PFS data were identified for cisplatin plus gemcitabine during the clinical SLR. It was determined that treatment would be expected to have a similar PFS profile to docetaxel and paclitaxel given they are all chemotherapy agents. An analysis of drug usage in the UK indicates that, in the patient population under consideration, paclitaxel is used more commonly then docetaxel.[44] Therefore, the HR for paclitaxel versus nivolumab was also applied for cisplatin plus gemcitabine. This necessary simplification further limits the generalisability of this comparator.
The separate responder and non-responder curves were weighted based on the number of patients measured as being progression-free and alive at the 8-week landmark point in the CheckMate 275 and CheckMate 032 trials. This weighting was assumed to remain constant for the remaining time horizon in each parametric model.	This is likely to be a conservative assumption as the weighting would be expected to increase in favour of the responding patients across time, who die at a much slower rate than the non-responding patients.

Abbreviations: BSC: best supportive care; cisplatin plus gemcitabine: gemcitabine and cisplatin.

B.3.7 Base case results

B.3.7.1 Base case incremental cost-effectiveness analysis results

The results of the base case analysis with and without the PAS are summarised in Table 44 and Table 45. Nivolumab is more effective than docetaxel, paclitaxel and BSC in terms of quality-adjusted life year (QALY) gains but is associated with higher lifetime costs than all treatments irrespective of whether a PAS was included. All remaining results, including the outputs from the sensitivity and scenario analyses, are presented with the PAS only. Under the end-of-life criteria that should be considered relevant to nivolumab in this appraisal (see Section B.2.13.2), these base case ICERs fall below the cost-effectiveness threshold adopted by NICE of £50,000 per QALY.

Clinical outcomes from the model, which have been included in the cost-effectiveness analysis, have been presented in Appendix J. Appendix J also contains the disaggregated results of base case incremental cost effectiveness analysis.

Table 44: Base case results – with PAS

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus Nivolumab (£/QALY)
Nivolumab	[redacted]	2.78	[redacted]				

Paclitaxel	£14,426	1.19	0.76	[redacted]	1.60	[redacted]	£37,647
Docetaxel	£13,945	1.40	0.92	[redacted]	1.38	[redacted]	£44,960
BSC	£9,056	1.01	0.64	[redacted]	1.77	[redacted]	£38,164

Abbreviations: BSC: best supportive care; ICER: incremental cost-effectiveness ratio; LYG: life years gained; QALYs: quality-adjusted life years.

Table 45: Base case results – without PAS

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus Nivolumab (£/QALY)
Nivolumab	[redacted]	2.78	[redacted]				
Paclitaxel	£14,426	1.19	0.76	[redacted]	1.60	[redacted]	[redacted]
Docetaxel	£13,945	1.40	0.92	[redacted]	1.38	[redacted]	[redacted]
BSC	£9,056	1.01	0.64	[redacted]	1.77	[redacted]	[redacted]

Abbreviations: BSC: best supportive care; ICER: incremental cost-effectiveness ratio; LYG: life years gained; QALYs: quality-adjusted life years.

B.3.8 Sensitivity analyses

B.3.8.1 Probabilistic sensitivity analysis

Probabilistic sensitivity analysis (PSA) was undertaken to assess the impact of uncertainty on the model results. The inputs varied and their parameters are presented in Table 46. The survival curves for TTD, PFS and OS were also varied for nivolumab using Cholesky decomposition to estimate the range of values that could be applied. Whereas all other parameters in the PSA were assumed to be independent, Cholesky decomposition is necessary to account for covariance between the coefficients for each survival curve.[110] The parameters generated from the Cholesky decomposition were varied using a lognormal distribution.

One main set of parameters was not included in the PSA – the time-varying HRs for each of the comparators. This was due to the complexity of varying HRs that change for every cycle (i.e. it would be illogical to generate a HR that increases and decreases at an irregular pattern for the full-time horizon). Also, as the comparator curves are dependent on the underlying survival curves for nivolumab, and these curves are varied during the PSA, the impact of varying the survival curves for each of the comparators should still be captured.

The results of the probabilistic analysis with 1,000 iterations are shown in Table 47. These results are similar to the deterministic outputs but, overall, there is a small increase in the probabilistic ICERs. This appears to be due to a small reduction in PFS and OS within the probabilistic analysis. For example the mean PFS for nivolumab is 15.55 and 12.23 for the deterministic and probabilistic analyses respectively. These reductions occur for all treatments but due to the larger PFS and OS estimates for nivolumab the impact of the reduction is greater for this treatment, hence the reduction in the probabilistic ICER. The results of the PSA are also presented graphically via scatterplots in Figure 44.

Table 46: Probability sensitivity analysis parameters and distributions

Parameter	Mean	Standard error	Distribution
Patient characteristics			
Age	65	0.576	Normal
Weight	77.3	16.340	Normal
Body surface area	2	0.205	Normal
Cost and resource use			
BSC cost per cycle	£170	£42.55	Gamma
Monitoring – Nivolumab	£336	£84.00	Gamma
Monitoring – Docetaxel	£498	£124.50	Gamma
Monitoring - Paclitaxel	£221	£55.33	Gamma
Subsequent radiotherapy	£11.92	£2.98	Gamma
Subsequent surgery	£105.66	£26.41	Gamma
Terminal care	£6,153	£1,538.16	Gamma
Utility			
Pre-progression	0.72	0.02	Beta
Change to post-progression	0.115	0.014	Gamma

Abbreviations: BSC: best supportive care.

Table 47: Probabilistic results

Technologies	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)	Probability of cost-effectiveness[a]
Paclitaxel	[redacted]	[redacted]	£46,209	72.10%
Docetaxel	[redacted]	[redacted]	£54,220	49.00%
BSC	[redacted]	[redacted]	£44,698	76.30%

[a]The probability of nivolumab being cost-effective versus the stated comparator at a cost-effectiveness threshold of £50,000/QALY.
Abbreviations: BSC: best supportive care, ICER: incremental cost-effectiveness ratio; QALYs: quality-adjusted life years.

The PSA results presented above are based on a cost-effectiveness threshold of £50,000 per QALY gained. A £50,000 threshold has been applied as it is expected that nivolumab will meet NICE's end-of-life criteria. This is discussed further in Section B.2.13.2. The impact of adopting an alternative threshold value is shown in the cost-effectiveness acceptability curves (CEACs) presented in Figure 45.

Figure 44: Probabilistic sensitivity analysis scatterplots

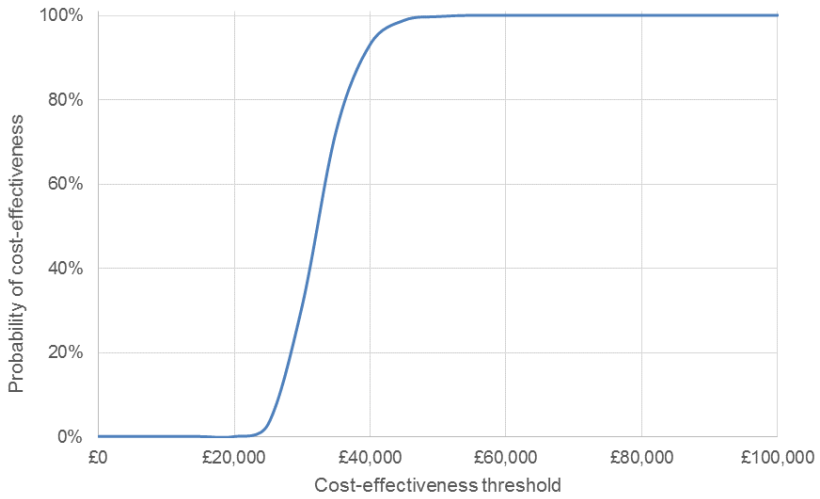
Paclitaxel

Docetaxel

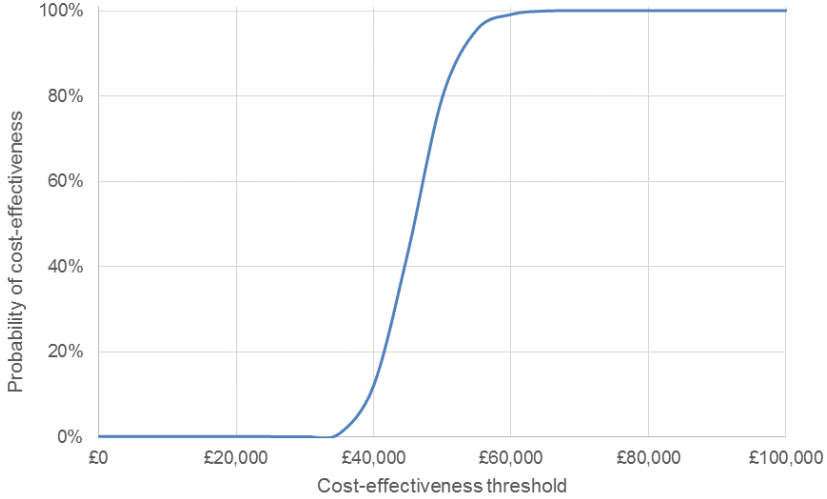
Best supportive care

Figure 45: Cost-effectiveness acceptability curves

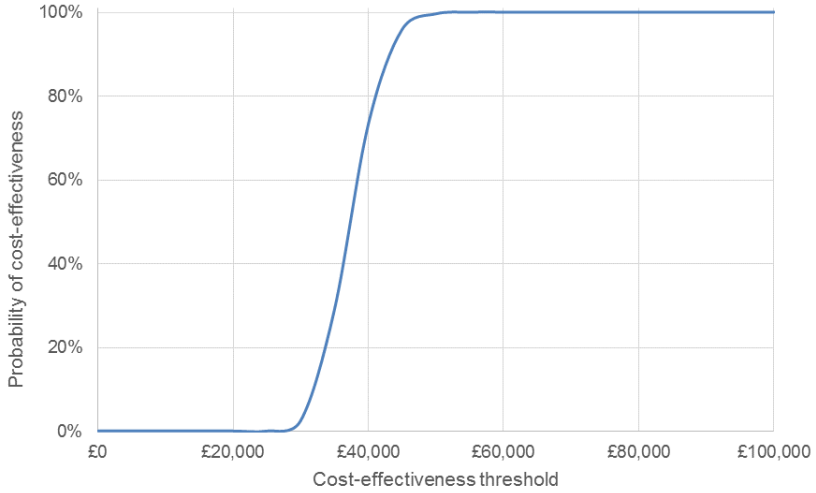
Paclitaxel



Docetaxel



Best supportive care



B.3.8.2 Deterministic sensitivity analysis

All parameters were varied individually during one-way deterministic sensitivity analysis (DSA) with the exception of the survival curves for all treatments as these could not be varied deterministically. All parameters were varied within the 95% CI, or if this was not available from the original data source, or could not be estimated, then the parameter was varied within a range of +/-50%.

The results of the DSA (including PAS) are summarised as tornado diagrams in Figure 46 to

Figure 48. These tornado diagrams use net monetary benefit (NMB), rather than ICERs, as the metric of cost-effectiveness. This is because negative ICERs are difficult to interpret (i.e. it is not clear whether this indicates cost-effectiveness or not at the given threshold value) and negative ICERs could be generated during the DSA. For NMB, a positive value always indicates cost-effectiveness, at the given threshold value (i.e. £50,000/QALY), whilst a negative value indicates the opposite. A large number of parameters had no meaningful impact on the results and, therefore, only the 12 parameters with the largest impact are presented here.

The results indicate that the three parameters with the largest impact on the results are: cost per unit for nivolumab, patient weight and patient age.

Figure 46: Tornado diagram – nivolumab versus paclitaxel

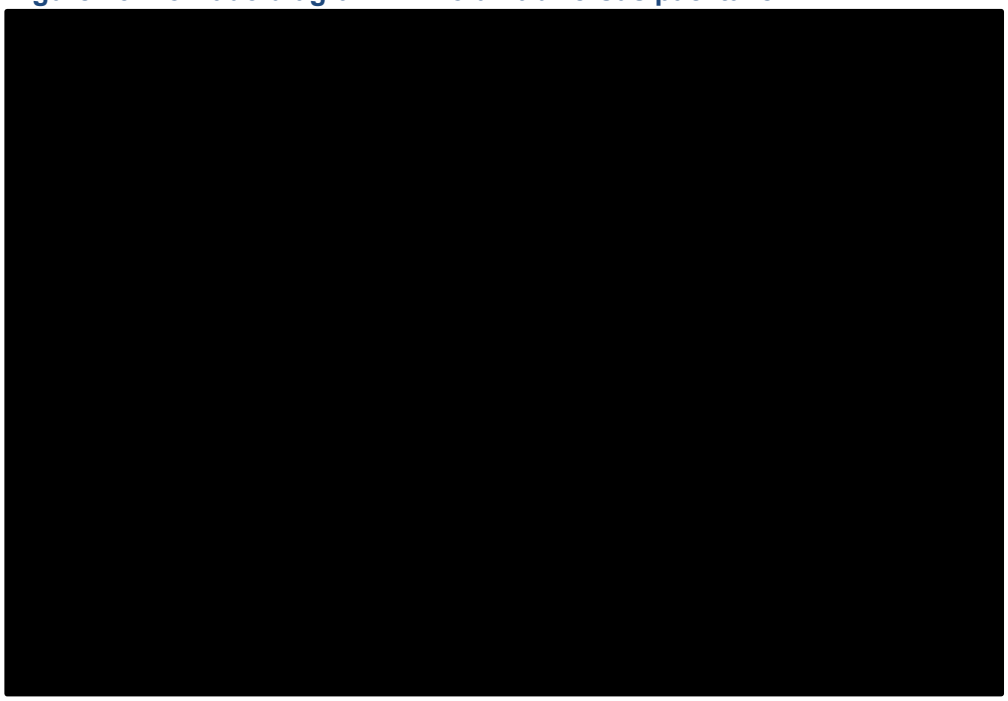


Figure 47: Tornado diagram – nivolumab versus docetaxel

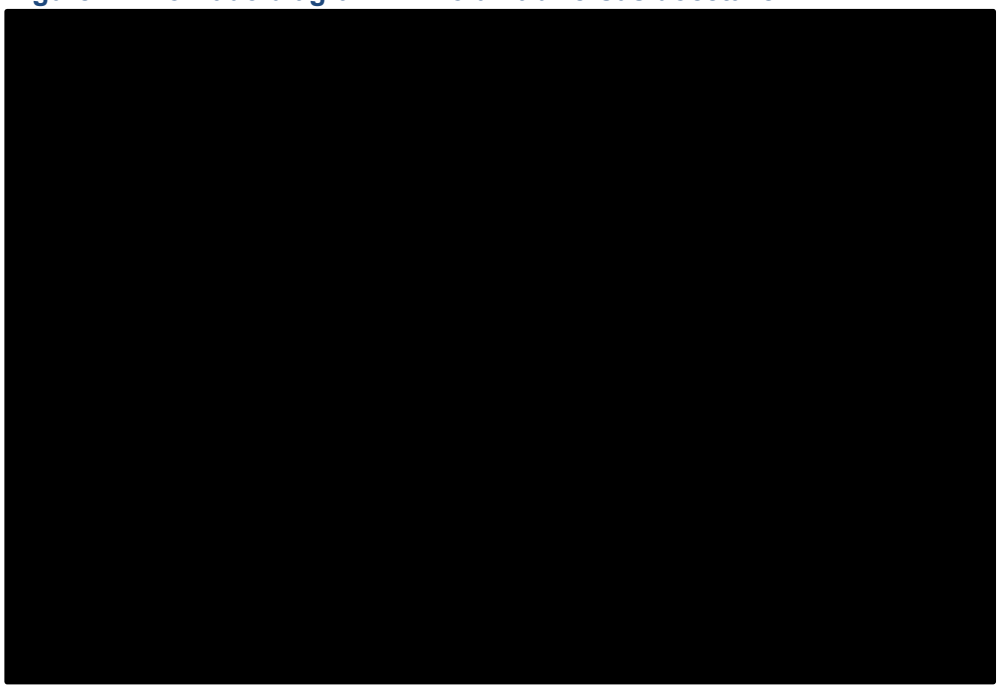
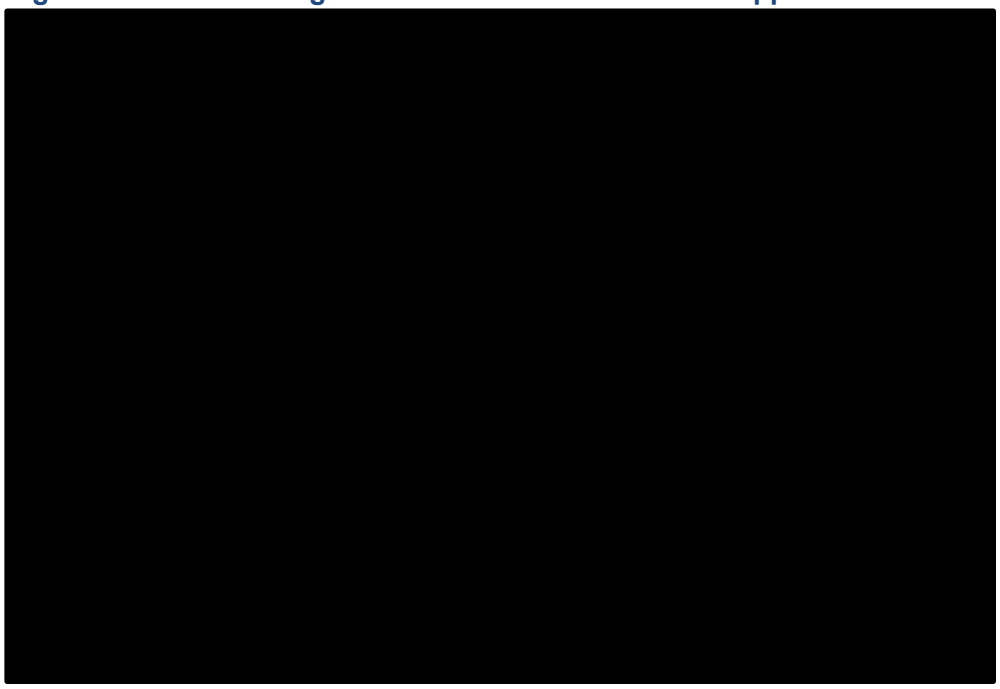


Figure 48: Tornado diagram – nivolumab versus best supportive care



B.3.8.3 Scenario analysis

A series of scenario analyses (including PAS) were completed to explore uncertainty regarding key structural assumptions of the analysis. These scenarios, and the key results from each scenario as measured by the ICER (per QALY gained), are described below.

Scenario 1 – Survival curves

In total, six parametric distributions were examined for the PFS and OS curves with the 8-week and 26-week landmark analysis. The generalised gamma distribution at 8 weeks was applied in Company evidence submission template for ID995.

the base case and the results of the following distributions are reported here: Weibull, Gompertz, lognormal, log-logistic and exponential at 8 weeks and generalised gamma, Weibull, Gompertz, lognormal, log-logistic and exponential at 26 weeks. To avoid an excessive number of scenarios the same distribution was always applied for PFS and OS. The results with each distribution are summarised in Table 48 and Table 49.

The results of this scenario analysis indicate that for the current UK standard of care, paclitaxel, the majority of distributions are associated with ICERs that are below or around the cost-effectiveness threshold for end of life medicines. This is also the same for BSC, another commonly used treatment strategy.

Table 48: Summary of scenario 1 – alternative parametric curves at 8 weeks

Distribution	ICER (per QALY) for nivolumab versus:		
	Paclitaxel	Docetaxel	BSC
Weibull	£101,994	£114,823	£91,372
Gompertz	£49,010	£59,858	£50,201
Lognormal	£52,900	£72,044	£53,634
Log-logistic	£58,279	£78,063	£59,695
Exponential	£57,998	£70,582	£59,564
Generalised gamma (base case)	*£37,647*	*£44,960*	*£38,164*

Abbreviations: BSC: best supportive care: Gen. Gamma: Generalised Gamma; ICER: incremental cost-effectiveness ratio; QALY: quality-adjusted life years.

Table 49: Summary of scenario 1 – alternative parametric curves at 26 weeks

Distribution	ICER (per QALY) for nivolumab versus:		
	Paclitaxel	Docetaxel	BSC
Gen. Gamma	£34,541	£40,246	£34,774
Weibull	£50,060	£62,866	£51,378
Gompertz	£35,655	£41,933	£35,269
Lognormal	£38,834	£48,610	£38,192
Log-logistic	£42,475	£54,235	£43,097
Exponential	£60,279	£76,786	£61,389
8-week analysis (base case)	*£37,647*	*£44,960*	*£38,164*

Abbreviations: BSC: best supportive care: Gen. Gamma: Generalised Gamma; ICER: incremental cost-effectiveness ratio; QALY: quality-adjusted life years.

Scenario 2 – Alternative fractional polynomial model

A number of fractional polynomial (FP) models were tested during the STC (see Appendix L for more details) and for the OS curves two models had almost identical scores for statistical goodness-of-fit. The p1=0, p2=0 model was chosen in the base case as it was seen to generate more clinically plausible long-term extrapolations of survival for each treatment. Specifically, the survival curves estimated using p1=1, p2=1 produces long, flat tails for the comparator treatments, which indicate a significant proportion of patients at later time points (e.g. approximately 5% of docetaxel, paclitaxel and BSC patients remain alive at 15, 10.5 and 5 years respectively). This is not reflective of current clinical experience with docetaxel/paclitaxel in

second line, post-platinum treated patients with advanced and metastatic bladder cancer and therefore represents an implausible scenario.

Table 50 presents the results of the model when alternative FP model (p1=1, p2=1) is chosen to generate the time-varying HRs for the comparators. These results using the p1=1, p2=1 model generate much less favourable results for nivolumab due to the very optimistic and clinically implausible survival curves estimated for comparator treatments.

Table 50: Summary of scenario 2 – alternative fractional polynomial model (p1=1, p2=1)

FP model choice	ICER (per QALY) for nivolumab versus:		
	Paclitaxel	Docetaxel	BSC
p1=1, p2=1	£56,073	£59,504	£43,554
p1=0, p2=0 (base case)	*£37,647*	*£44,960*	*£38,164*

Abbreviations: BSC: best supportive care: FP: fractional polynomial; ICER: incremental cost-effectiveness ratio; QALY: quality-adjusted life years.

Scenario 3 – conservative exponential piecewise modelling

An exponential piecewise approach (using a 26-week cut-off for OS, PFS and TTD) was explored because of a previously stated preference by the appraisal committee for piecewise exponential modelling. As stated in previous appraisals of nivolumab however, the exponential is characterised by a constant hazard, which is not necessarily appropriate for the mechanism of action of nivolumab, whereby there is evidence of a slowly decreasing hazard (see Appendix L). The scenario should therefore be considered to represent a very conservative outcome, which does not account for the plausibility that some patients treated with nivolumab have a lower hazard of progression or death than the entire cohort being considered.

The results of the exponential piecewise approach are shown in Table 51. They indicate that with the highly conservative piecewise exponential approach, nivolumab is just outside the threshold for cost-effectiveness versus paclitaxel, the UK standard of care. However, this scenario should be considered very conservative, given the longer-term data for nivolumab from other indications and clinical feedback from experts consulted.

Table 51: Summary of scenario 3 – conservative piecewise exponential modelling

TTD distribution	ICER (per QALY) for nivolumab versus:		
	Paclitaxel	Docetaxel	BSC
Piecewise exponential at 8 weeks	£53,616	£65,450	£55,597
Piecewise exponential at 26 weeks	£55,681	£71,147	£57,293
Generalised gamma (base case)	*£37,647*	*£44,960*	*£38,164*

Abbreviations: BSC: best supportive care; Gen. Gamma: Generalised Gamma; ICER: incremental cost-effectiveness ratio; QALY: quality-adjusted life years; TTD: time to treatment discontinuation.

Scenario 4 – inclusion of vial sharing

The results when vial sharing is included within the economic model are shown in Table 52. They indicate that when vial sharing is included slightly more favourable results are generated for nivolumab versus docetaxel, paclitaxel and BSC.

Table 52: Summary of scenario 4 – inclusion of vial sharing

Vial sharing	ICER (per QALY) for nivolumab versus:		
	Paclitaxel	Docetaxel	BSC
Vial sharing	£35,651	£42,630	£36,333
No vial sharing (base case)	*£37,647*	*£44,960*	*£38,164*

Abbreviations: BSC: best supportive care; ICER: incremental cost-effectiveness ratio; QALY: quality-adjusted life years.

Scenario 5 – Treatment stopping rule

As noted in Section B.3.3.3 it may be feasible to stop treatment with nivolumab for patients who have not yet progressed and still maintain a clinical benefit due to the unique mechanism of action of nivolumab in restoring anti-tumour activity. Evidence to support the stopping of treatment for patients who are responding to nivolumab is available from the CheckMate 003 trial in which treatment was continued up to 96 weeks.[80] Ongoing responses after treatment cessation were observed in this trial for patients with advanced NSCLC who had completed 96 weeks of therapy with nivolumab (see Figure 39). For this analysis the stopping rule was applied to 75% of patients who were yet to discontinue. It was assumed that 25% remained on treatment to reflect a potential minority of patients and/or their clinician who chose to remain on treatment for a longer time period.

The results when the stopping rule is included at 2 years is shown in Table 53. They indicate that the inclusion of a stopping rule produces more favourable results for nivolumab with all ICERs lower than the £50,000/QALY threshold.

Table 53: Summary of scenario 5 – treatment stopping rule

Nivolumab stopping rule	ICER (per QALY) for nivolumab versus:		
	Paclitaxel	Docetaxel	BSC
Stopping rule included	£31,561	£37,781	£32,743
No stopping rule (base case)	*£37,647*	*£44,960*	*£38,164*

Abbreviations: BSC: best supportive care; ICER: incremental cost-effectiveness ratio; QALY: quality-adjusted life years.

Scenario 6 – Time to discontinuation

In total six parametric distributions were examined for the time to discontinuation (TTD) curves. The Generalised Gamma distribution was applied in the base case and the results of the following distributions are reported here: Weibull, Gompertz, Lognormal, Log-logistic and Exponential. To avoid an excessive number of scenarios the same distribution was always applied for PFS and OS. The results with each distribution are summarised in Table 54.

The results of this scenario analysis indicate the choice of distribution for TTD has a large impact on the ICER as shown by the range of values presented in Table 54. When the exponential or Weibull distribution is selected the ICER is below £50,000/QALY for all comparisons. Alternatively, with the Gompertz, Lognormal and Log-logistic distributions all ICERs were greater than £50,000/QALY. However, as noted previously, the Gompertz and Lognormal distributions produce long tails with a similar effect also shown with the Log-logistic distribution. Therefore, a small proportion of patients are being modelled as remaining on treatment for a number of years (i.e. 5 and 10 years), which is not expected for nivolumab in clinical practice. Therefore, these clinically implausible distributions are likely to overestimate the treatment costs for nivolumab.

Table 54: Summary of scenario 6 – alternative parametric curves for TTD

Distribution	ICER (per QALY) for nivolumab versus:		
	Paclitaxel	Docetaxel	BSC
Weibull	£33,562	£40,141	£34,525
Gompertz	£183,467	£216,984	£168,053
Lognormal	£61,810	£73,465	£59,688
Log-logistic	£61,994	£73,683	£59,851
Exponential	£28,331	£33,971	£29,866
Generalised gamma (base case)	*£37,647*	*£44,960*	*£38,164*

Abbreviations: BSC: best supportive care; Gen. Gamma: Generalised Gamma; ICER: incremental cost-effectiveness ratio; QALY: quality-adjusted life years.

B.3.8.4 Summary of sensitivity analyses results

The probabilistic results generated during the PSA were similar to the base case results, with a slight increase in the probabilistic ICERs compared with the deterministic analysis. Altogether, the PSA indicated that the nivolumab has a probability of cost-effectiveness, at a £50,000 threshold, of 72%, 49% and 76% when compared with paclitaxel, docetaxel and BSC respectively.

The DSA results show that the model results are robust to changes to the majority of parameters with only three parameters causing the direction of the ICER to change. Therefore, the key drivers of ICER uncertainty either related to the cost per cycle of nivolumab (e.g. unit price, patient weight) or patient age.

The results of the scenario analyses indicate that there are three further drivers of the model results: the choice of parametric distribution for the nivolumab PFS and OS curves; the choice of parametric distribution for the nivolumab TTD and the NMA HR estimates. The choice of parametric distributions has a particularly large impact on the overall results with a wide range of ICERs identified. However, the base case distributions have been selected based on the statistical goodness-of-fit and clinical plausibility and, therefore, are deemed to be the most appropriate distributions to use for the analysis.

B.3.9 Subgroup analysis

No subgroup analysis was undertaken, which is in line with the NICE scope.

B.3.10 Validation

B.3.10.1 Validation of cost-effectiveness analysis

Validation of nivolumab estimates

Key drivers of the model results are the long-term extrapolations for PFS and OS for patients treated with nivolumab. Therefore, the predictions from the model have been compared against feedback from clinical experts and other long-term nivolumab data currently available. These are data from the CheckMate 003 study that examine the safety and efficacy of nivolumab in NSCLC and several other solid tumours. Clinical experts at the advisory board indicated that lung cancer would be the most biologically similar to bladder cancer, in relation to the strong link to smoking, the choice of treatment used in clinical practice, and the poor outcomes associated with both diseases without treatment.[29] Survival data with a minimum of five-years follow-up for NSCLC patients was recently presented at an international conference.[111] A comparison of the outcomes with the generalised gamma distribution versus the CheckMate 275 and CheckMate 003 results is presented in Table 55 and illustrated graphically in Figure 49. The comparison shows that the model estimates of OS closely match both the Kaplan-Meier data from CheckMate 275 and the long-term data from CheckMate 003 with no difference in the 5-year estimates.

Validation of comparator estimates

The survival curve extrapolations for each of the comparators has also been validated against available clinical (trial and registry) data and expert opinion. Each treatment has been compared helwith the observed Kaplan-Meier data that was used to inform the STC (Table 55), which is shown to be a good match.

Sideris *et al.* (2016)[112] have also published 2-year follow-up data from Belgium on the efficacy of paclitaxel when used in real-world clinical practice. This data shows that OS for paclitaxel (and docetaxel, if it is assumed this has a similar profile given the mechanism of action) may be higher in the respective clinical trials, than is observed in clinical practice, although the differences are not very significant. Validation with two clinicians with expertise in the treatment of patients with locally advanced unresectable or metastatic UC indicated that they would not expect more than 5% of patients to be alive at two years.[67] This feedback is more closely aligned with the outcomes for paclitaxel, which was informed by the UK trial PLUTO.

Unfortunately further registry data, specific to the UK, could not be located. However, the outcomes estimated by the model are reasonable, given the clinical data and expert feedback provided, with a potential slight overestimate for the comparators. The implications of this, however, is that the cost-effectiveness of nivolumab may be underestimated.

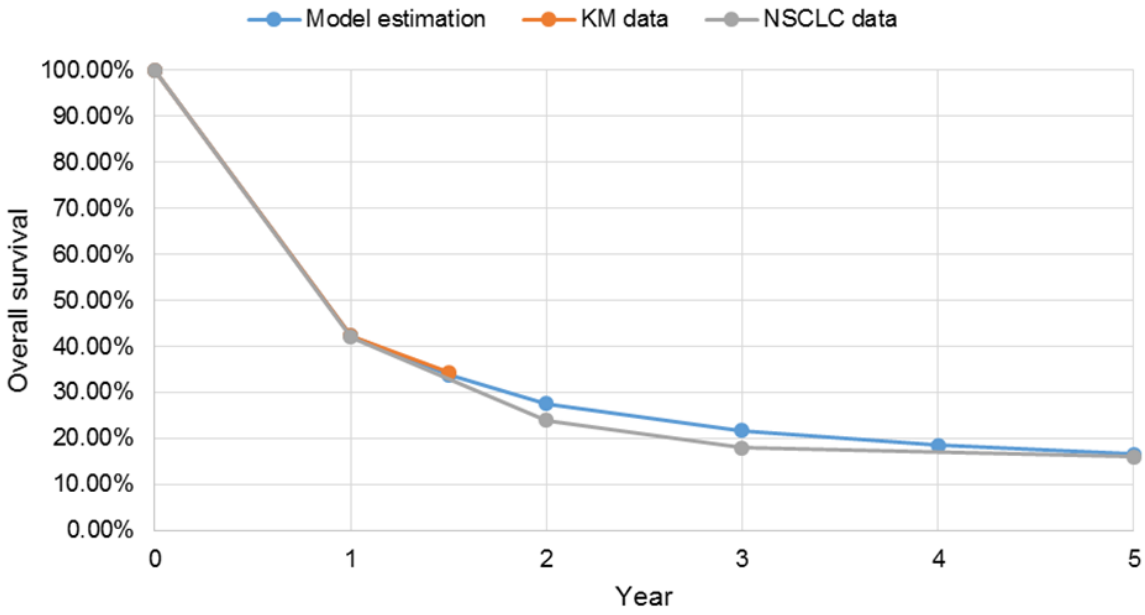
Finally, Bellmunt *et al.* (2013) published three-year follow-up data, which was used to inform the STC.[85] The Kaplan-Meier data from this study indicates that the model may underestimate OS initially, followed by very similar profiles during year 2, and then overestimate survival from the end of year 2 onwards. These comparisons are summarised in Table 55.

Table 55: Comparison of overall survival extrapolation in model against observed data

Data source	Survival curve	Proportion alive, %					
		1 year	1.5 years	2 years	3 years	4 years	5 years
Nivolumab							
Model estimates for OS	Gen. Gamma (Base case)	42.34%	33.82%	27.54%	21.66%	18.51%	16.55%
CheckMate 275	Kaplan-Meier data	[redacted]	[redacted]	-	-	-	-
CheckMate 003 (NSCLC)	-	42%	-	24%	18%	-	16%
Docetaxel							
Model estimates for OS	Gen. Gamma (Base case)	25.01%	15.67%	11.05%	7.67%	6.36%	5.69%
Chouieri *et al.* (2012)[31]	Kaplan-Meier data	24.33%	13.03%	-	-	-	-
Sideris *et al.* (2016)[112]	Kaplan-Meier data (Bytescout)	19%	8%	6%	-	-	-
Paclitaxel							
Model estimates for OS	Gen. Gamma (Base case)	31.41%	17.40%	10.56%	5.66%	3.94%	3.15%
Jones *et al.* (2017)[32]	Kaplan-Meier data	31.58%	15.08%				
Sideris *et al.* (2016)[112]	Kaplan-Meier data (Bytescout)	19%	8%	6%	-	-	-
BSC							
Model estimates for OS	Gen. Gamma (Base case)	14.00%	8.96%	6.64%	5.03%	4.42%	4.09%
Bellmunt *et al.* (2013)[85]	Kaplan-Meier data	21.30%	10.65%	7.41%	1.39%	-	-

Abbreviations: BSC: best supportive care; NSCLC: non-small cell lung cancer; OS: overall survival.

Figure 49: Validation of model predictions of OS with nivolumab



B.3.11 Interpretation and conclusions of economic evidence

The economic evaluation considered patients with locally advanced unresectable or metastatic UC who have progressed following first-line platinum-based chemotherapy. This reflects the population of the CheckMate 275 and CheckMate 032 trials and is extendable to all patients included in the final NICE scope.

It is expected that the results of the economic evaluation are generalisable to clinical practice in England. This is for the following reasons:

- The structure of the economic model is consistent with previous oncology submissions to NICE, including the only submission in this indication (Vinflunine; TA272).
- The population from the CheckMate 275 and 032 trials are considered to be reflective of the patient population in England.
- The economic model uses utility data generated from these trials and using UK preference weights.
- Unit costs have been sourced from relevant, well-established UK sources (e.g. NHS Reference Costs, eMit).
- The approach adopted takes into account feedback from the ERG in previous nivolumab HTA submissions to NICE.
- The model structure and inputs have been validated by UK-based experts, including clinicians and health economists.

The economic evaluation makes use of the best available evidence to estimate the costs and QALYs with each treatment option. This includes the use of IPD from the CheckMate 275 and CheckMate 032 trials, including data on patient quality of life. It was necessary to extrapolate from trial data in order to predict long-term treatment benefit and robust methods were used for these extrapolations. This included the use of an extensive range of distributions, following the NICE Decision Support Unit (DSU) methods guidance. A novel approach (i.e. landmark analysis)

was also adopted to capture the immune-response survival effect of nivolumab in a proportion of patients and, therefore, better estimate the long-term outcomes with this treatment. Overall, it is believed that the model produces clinically plausible estimates of PFS and OS for nivolumab.

The main weaknesses of the evaluation are believed to be the immaturity of the OS data and a lack of RCT evidence to estimate the efficacy of nivolumab. The immature nature of the nivolumab was due to the relatively short follow-up data available at the time of the analysis. Due to the immaturity of the data the choice of parametric distribution to predict long-term outcomes has a large impact on the final ICERs. However, as noted previously, it is believed that clinically plausible estimates of PFS and OS were used in the base case analysis. Further, OS and PFS may be slightly overestimated for the comparators, given clinical feedback. Therefore, the base case ICERs may actually underestimate the cost-effectiveness of nivolumab. No randomised, comparative evidence is available for nivolumab in this indication. However, sophisticated prediction models were generated, following guidelines outlined by the NICE DSU, in order to predict the efficacy of nivolumab versus each comparator. Extensive sensitivity analysis was also undertaken to examine the uncertainty relating to the model effectiveness data.

The outputs from the economic evaluation are largely based on data from the CheckMate 275 trial. The clinical cut-off from this trial was July 2016; however, follow-up continues in patients and additional OS data are expected in [redacted]. These additional data could be used to re-analyse the survival estimates in the economic model and are expected to confirm the current results.

Overall, the results of the economic evaluation indicate that nivolumab is cost-effective for second-line UC patients when compared with the treatment options most commonly used in these patients in the UK (i.e. docetaxel, paclitaxel and BSC). When compared with paclitaxel, the current UK standard of care[67], nivolumab is associated with ICERs that are well below the cost-effectiveness threshold for an end of life medicine at £37,647. Importantly, this comparison is informed by a UK based trial (Jones *et al.* 2017) which provides robust and relevant evidence of the clinical outcomes seen with paclitaxel in UK practice. The results for the clinical and cost-effectiveness of nivolumab versus paclitaxel can therefore be seen as highly generalisable to clinical practice and decision-making in the UK.

B.4 Assessment of factors relevant to the NHS and other parties

B.4.1 Number of patients eligible for treatment in England and Wales

For the analysis of budget impact, the incident number of patients in England and Wales eligible for treatment with nivolumab, as per the licensed indication for patients with locally advanced unresectable or metastatic UC after failure of prior platinum-containing chemotherapy, was estimated to be 894 patients per year. Full details of the derivation of this calculation are presented in Table 11 below.

Table 56: Estimated eligible population for nivolumab in England and Wales

Stage of treatment pathway	Estimate		Source
Newly diagnosed with bladder cancer (all stages) in England and Wales	9,021		Cancer Research UK (2014)[36]
Transitional cell carcinoma histology	90%	8,119	Pasin *et al.* (2008)[16]
Muscle-invasive disease			
Newly diagnosed with muscle invasive bladder cancer (stage II)	23%	1,868	Cancer Research UK (2014)[36]
Receive neoadjuvant/adjuvant cisplatin-based chemotherapy with radical cystectomy or radiotherapy	40-60%	934	Expert clinician feedback
Progress/recur after neoadjuvant/adjuvant cisplatin-based chemotherapy with radical cystectomy or radiotherapy to locally advanced unresectable or metastatic disease	40%	374	Expert clinician feedback[29]
Considered for nivolumab therapy Considered for further chemotherapy	**35%** 65%	**131** 244	Expert clinician feedback[29]
Locally advanced or metastatic disease			
Newly diagnosed with locally advanced muscle invasive bladder cancer or metastatic disease (Stage III/IV)	20%	1,624	Cancer Research UK (2014)[36]
Total newly diagnosed/ progressed to locally advanced unresectable or metastatic disease (Stage III/IV) and considered for chemotherapy	244 + 1,624	1,868	Calculation
Receive first-line platinum-based chemotherapy	50–66%	1,090	Expert clinician feedback[29]
Progress/recur and eligible for second-line treatment	**60–80%**	**763**	**Expert clinician feedback[29]**
Patients eligible to receive nivolumab as per licensed indication in England and Wales	**131 + 763**	**894**	**Calculation**

Note that numbers have been subjected to rounding within each calculation.

B.4.2 Assumptions made about current treatment options and uptake of technologies

All comparators included in the final scope for this appraisal (paclitaxel, docetaxel, BSC and retreatment with platinum-based chemotherapy) have been considered in the budget impact analysis and are assumed to be equally displaced by the introduction of nivolumab. Market share estimates used in the budget impact analysis are presented in Section B.4.3 below.

B.4.3 Assumptions made about market share in England and Wales

The proportion of patients receiving each therapy, based on internal market share estimates, is presented in Table 57 for the scenario without nivolumab and in Table 58 for the scenario with nivolumab. This budget impact analysis was based on a closed cohort; as a result, the total number of patients eligible to receive nivolumab was estimated to be 894 each year over the 5-year time horizon.

As described in Section B.1.3.2, the majority of patients with locally advanced unresectable or metastatic UC after failure of prior platinum-containing chemotherapy are currently expected to receive treatment with paclitaxel monotherapy, with docetaxel monotherapy and BSC used to a lesser extent. Expert clinician feedback was that <10% of patients would be likely to receive retreatment with platinum-containing chemotherapy in this setting. Therefore, in the budget impact analysis the proportion of patients receiving retreatment with platinum-containing chemotherapy was estimated to be 10%, with the remaining 90% of patients assumed to receive paclitaxel and docetaxel or BSC, in line with the relative use of these therapies as reported in a recent chart review conducted by Bristol-Myers Squibb. Based on Bristol-Myers Squibb internal estimates, in the world with nivolumab, nivolumab is expected to have a market share of [redacted]% in Year 1, rising to [redacted]% in subsequent years.

Table 57: Proportion of patients receiving each therapy – NHS without nivolumab

Treatment	Year 1	Year 2	Year 3	Year 4	Year 5
Paclitaxel monotherapy	36.6%	36.6%	36.6%	36.6%	36.6%
Docetaxel monotherapy	26.8%	26.8%	26.8%	26.8%	26.8%
BSC	26.6%	26.6%	26.6%	26.6%	26.6%
Retreatment with platinum-containing chemotherapy	10.0%	10.0%	10.0%	10.0%	10.0%

Abbreviations: BSC: Best supportive care; NHS: National Health Service.

Table 58: Proportion of patients receiving each therapy – NHS with nivolumab

Treatment	Year 1	Year 2	Year 3	Year 4	Year 5
Nivolumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Paclitaxel monotherapy	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Docetaxel monotherapy	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Retreatment with platinum-containing chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: BSC: Best supportive care; NHS: National Health Service.

Table 59: Number of patients receiving each therapy – NHS without nivolumab

Treatment	Year 1	Year 2	Year 3	Year 4	Year 5
Paclitaxel monotherapy	327	327	327	327	327
Docetaxel monotherapy	240	240	240	240	240
BSC	238	238	238	238	238
Retreatment with platinum-containing chemotherapy	89	89	89	89	89

Abbreviations: BSC: Best supportive care; NHS: National Health Service.

Table 60: Number of patients receiving each therapy – NHS with nivolumab

Treatment	Year 1	Year 2	Year 3	Year 4	Year 5
Nivolumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Paclitaxel monotherapy	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Docetaxel monotherapy	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
BSC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Retreatment with platinum-containing chemotherapy	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: BSC: Best supportive care; NHS: National Health Service.

B.4.4 Cost inputs

Costs associated with drug acquisition and administration were included in the budget impact analysis. The unit costs for these are consistent with those used in the cost-effectiveness analysis, described in Section B.3.5. Based on expert clinician feedback, it was assumed that patients receiving chemotherapy (paclitaxel, docetaxel or retreatment with platinum-containing chemotherapy) would receive a maximum of 6 treatment cycles (21-day or 28-day as appropriate). Patients receiving nivolumab were assumed to receive a total of [redacted] doses, based on the median number of doses received in the CheckMate 275 trial. Patients receiving BSC were assumed to receive 13 4-week cycles of BSC costs per year. For simplification, retreatment with platinum-based chemotherapy was based on cisplatin plus gemcitabine costs only.

Table 61: Treatment costs included within the budget impact analysis

Therapy	Total cost per dose (inc. admin)	Total cost per cycle	Number of cycles per year	Number of doses per year	Total cost per year

Nivolumab (with PAS)	[redacted]	[redacted]	-	[redacted]	[redacted]
Paclitaxel monotherapy	£215.95	£647.86	6	-	£3,887.17
Docetaxel monotherapy	£227.78	£227.78	6	-	£1,366.68
BSC	-	£6.77	13	-	£88.07
Retreatment with platinum-based chemotherapy	£833.83	£2,057.66	6	-	£12,345.93

Abbreviations: BSC: Best supportive care.

B.4.5 Estimates of resource savings

There are no estimates of resource savings although nivolumab is associated with fewer adverse events for patients versus the standard of care.

B.4.6 Estimated annual budget impact on the NHS in England and Wales

The budget impact analysis compares total costs over a 5-year time horizon between scenarios with and without nivolumab, with Year 1 coinciding with the year of introduction of nivolumab in the former scenario. The annual net budget impact associated with the introduction of nivolumab is presented in Table 62 (with PAS); by Year 5, the annual net budget impact of introducing nivolumab is estimated to be [redacted].

Results of these analyses are limited by the accuracy of market share predictions. Furthermore, by only modelling a closed cohort, the analysis does not include patients who may continue to receive treatment across the 5-year time horizon.

Table 62: Estimated annual budget impact to NHS England and Wales of introducing nivolumab – over the first 5 years (with PAS for nivolumab)

Treatment	Year 1	Year 2	Year 3	Year 4	Year 5
NHS without nivolumab					
Paclitaxel monotherapy	£1,271,899	£1,271,899	£1,271,899	£1,271,899	£1,271,899
Docetaxel monotherapy	£327,446	£327,446	£327,446	£327,446	£327,446
BSC	£20,943	£20,943	£20,943	£20,943	£20,943
Retreatment with platinum-based chemotherapy	£1,103,726	£1,103,726	£1,103,726	£1,103,726	£1,103,726
Total cost	£2,724,014	£2,724,014	£2,724,014	£2,724,014	£2,724,014
NHS with nivolumab					
Nivolumab (with PAS)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Paclitaxel monotherapy	£1,017,519	£890,329	£890,329	£890,329	£890,329
Docetaxel monotherapy	£261,957	£229,212	£229,212	£229,212	£229,212

BSC	£16,754	£14,660	£14,660	£14,660	£14,660
Retreatment with platinum-based chemotherapy	£882,981	£772,608	£772,608	£772,608	£772,608
Total cost	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Net budget impact	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Cumulative net budget impact					[redacted]

Abbreviations: NHS: National Health Service; PAS: Patient Access Scheme.

B.5 References

1. National Institute for Health and Care Excellence (NICE). Final Scope: Nivolumab for treating metastatic or unresectable urothelial cancer - Issued April 2017.

2. Mellman I, Coukos G, Dranoff G. Cancer immunotherapy comes of age. Nature 2011;480:480-489.

3. Brahmer JR, Drake CG, Wollner I, et al. Phase I study of single-agent anti–programmed death-1 (MDX-1106) in refractory solid tumors: safety, clinical activity, pharmacodynamics, and immunologic correlates. Journal of Clinical Oncology 2010;28:3167-3175.

4. Chen DS, Irving BA, Hodi FS. Molecular pathways: next-generation immunotherapy—inhibiting programmed death-ligand 1 and programmed death-1. Clinical cancer research 2012;18:6580-6587.

5. Zandberg DP, Strome SE. The role of the PD-L1: PD-1 pathway in squamous cell carcinoma of the head and neck. Oral oncology 2014;50:627-632.

6. Kamat AM, Flaig TW, Grossman HB, et al. Expert consensus document: consensus statement on best practice management regarding the use of intravesical immunotherapy with BCG for bladder cancer. Nature Reviews Urology 2015;12:225-235.

7. Gontero P, Bohle A, Malmstrom P-U, et al. The role of bacillus Calmette-Guérin in the treatment of non–muscle-invasive bladder cancer. European urology 2010;57:410-429.

8. Liakou CI, Kamat A, Tang DN, et al. CTLA-4 blockade increases IFNγ-producing CD4+ICOShi cells to shift the ratio of effector to regulatory T cells in cancer patients. Proceedings of the National Academy of Sciences 2008;105:14987-14992.

9. Carthon B, Wolchok JD, Yuan J, et al. Pre-operative CTLA-4 blockade: tolerability and immune monitoring in the setting of a pre-surgical clinical trial. Clinical cancer research : an official journal of the American Association for Cancer Research 2010;16:2861-2871.

10. Inman BA, Sebo TJ, Frigola X, et al. PD-L1 (B7-H1) expression by urothelial carcinoma of the bladder and BCG-induced granulomata. Cancer 2007;109:1499-1505.

11. Boorjian SA, Sheinin Y, Crispen PL, et al. T-cell coregulatory molecule expression in urothelial cell carcinoma: clinicopathologic correlations and association with survival. Clinical cancer research 2008;14:4800-4808.

12. European Medicines Agency. Committee for Medicinal Products for Human Use (CHMP) Opinion: Opdivo (nivolumab) EMA/CHMP/251204/2017. Available at http://www.ema.europa.eu/docs/en_GB/document_library/Summary_of_opinion/human/003985/WC500226181.pdf [Accessed 28 Apr 2017]. 2017.

13. Bristol-Myers Squibb Pharmaceuticals Ltd. Nivolumab (OPDIVO) 10 mg/mL concentrate for solution for infusion. Summary of Product Characteristics. 2017. Available at: https://www.medicines.org.uk/emc/medicine/30476 [Accessed 20 June 17].

14. Bristol-Myers Squibb. Bristol-Myers Squibb's Opdivo (nivolumab) Receives Breakthrough Therapy Designation from U.S. Food and Drug Administration for Advanced Form of Bladder Cancer. Press Release. http://investor.bms.com/investors/news-and-events/press-releases/press-release-details/2016/Bristol-Myers-Squibbs-Opdivo-nivolumab-Receives-Breakthrough-Therapy-Designation-from-US-Food-and-Drug-Administration-for-Advanced-Form-of-Bladder-Cancer/default.aspx [Accessed 17 Feb 2017].

15. U.S. Food and Drug Administration. Fact Sheet: Breakthrough Therapies. Available at: http://www.fda.gov/RegulatoryInformation/Legislation/SignificantAmendmentstotheFDCAct/FDASIA/ucm329491.htm [Accessed 17 Feb 2017].

16. Pasin E, Josephson DY, Mitra AP, et al. Superficial Bladder Cancer: An Update on Etiology, Molecular Development, Classification, and Natural History. Reviews in Urology 2008;10:31-43.

17. Amin MB, McKenney JK, Paner GP, et al. ICUD-EAU international consultation on bladder cancer 2012: pathology. European urology 2013;63:16-35.

18. Action Bladder Cancer UK. Types of Bladder Cancer. available at: http://actionbladdercanceruk.org/types-of-bladder-cancer/ [Accessed 3 Feb 2017].

19. Bellmunt J, Orsola A, Leow JJ, et al. Bladder cancer: ESMO Practice Guidelines for diagnosis, treatment and follow-up. Annals of Oncology 2014;25:iii40-iii48.

20. The American Cancer Society. What Is Bladder Cancer? Revised 23 May 2016. Available at: https://www.cancer.org/cancer/bladder-cancer/about/what-is-bladder-cancer.html [Accessed 3 Feb 2017].

21. Razmaria AA. Bladder Cancer. Jama 2015;314:1886-1886.

22. Sharma S, Ksheersagar P, Sharma P. Diagnosis and treatment of bladder cancer. American family physician 2009;80:717-23.

23. Cancer Research UK. Bladder Cancer. Available at: http://www.cancerresearchuk.org/about-cancer/bladder-cancer/types-stages-grades/stages [Accessed 3 Feb 2017].

24. Knowles MA, Hurst CD. Molecular biology of bladder cancer: new insights into pathogenesis and clinical diversity. Nature Reviews Cancer 2015;15:25-41.

25. National Institute for Health and Care Excellence (NICE). [NG2]: Bladder cancer: diagnosis and management. Available: https://www.nice.org.uk/guidance/NG2 [Accessed 3 Feb 2017].

26. European Association of Urology. Muscle-invasive and Metastatic Bladder Cancer Guidelines. Available at: http://uroweb.org/guideline/bladder-cancer-muscle-invasive-and-metastatic/#1 [Accessed 3 Feb 2017], 2016.

27. National Institute for Health and Care Excellence (NICE). [TA272]: Vinflunine for the treatment of advanced or metastatic transitional cell carcinoma of the urothelial tract. Available: https://www.nice.org.uk/guidance/ta272 [Accessed 3 Feb 2017].

28. European Medicines Agency. Guideline on evaluation of anticancer medicinal products in man (Draft). Available at: http://www.ema.europa.eu/ema/doc_index.jsp?curl=pages/includes/document/document_detail.jsp?webContentId=WC500203320&murl=menus/document_library/document_library.jsp&mid=0b01ac058009a3dc Accessed: 23rd May 2016., 2016.

29. Bristol-Myers Squibb. Meeting minutes, Clinical Advisory Board Meeting: 06 March 2017, 0930-1600 [Not in the public domain]. 30 Euston Square, London, UK: Bristol-Myers Squibb, 2017.

30. Petrylak DP, Tagawa ST, Kohli M, et al. Docetaxel As Monotherapy or Combined With Ramucirumab or Icrucumab in Second-Line Treatment for Locally Advanced or Metastatic Urothelial Carcinoma: An Open-Label, Three-Arm, Randomized Controlled Phase II Trial. Journal of Clinical Oncology 2016;34:1500-9.

31. Choueiri TK, Ross RW, Jacobus S, et al. Double-blind, randomized trial of docetaxel plus vandetanib versus docetaxel plus placebo in platinum-pretreated metastatic urothelial cancer. Journal of Clinical Oncology 2012;30:507-12.

32. Jones R, Hussain S, Protheroe A, et al. Randomized Phase II Study Investigating Pazopanib Versus Weekly Paclitaxel in Relapsed or Progressive Urothelial Cancer. Journal of Clinical Oncology 2017;0:JCO.2016.70.7828.

33. Bellmunt J, Theodore C, Demkov T, et al. Phase III trial of vinflunine plus best supportive care compared with best supportive care alone after a platinum-containing regimen in patients with advanced transitional cell carcinoma of the urothelial tract. Journal of Clinical Oncology 2009;27:4454-4461.

34. Ortmann CA, Mazhar D. Second-line systemic therapy for metastatic urothelial carcinoma of the bladder. Future Oncology 2013;9:1637-1651.

35. Raggi D, Miceli R, Sonpavde G, et al. Second-line single-agent versus doublet chemotherapy as salvage therapy for metastatic urothelial cancer: a systematic review and meta-analysis. Annals of Oncology 2015:mdv509.

36. Cancer Research UK. Bladder cancer incidence statistics. Available at: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bladder-cancer/incidence#heading-Zero [Accessed 3 Feb 2017].

37. Sharma P, Callahan MK, Bono P, et al. Nivolumab monotherapy in recurrent metastatic urothelial carcinoma (CheckMate 032): a multicentre, open-label, two-stage, multi-arm, phase 1/2 trial. The Lancet Oncology 2016;17:1590-1598.

38. Sharma P, Retz M, Siefker-Radtke A, et al. Nivolumab in metastatic urothelial carcinoma after platinum therapy (CheckMate 275): a multicentre, single-arm, phase 2 trial. The Lancet Oncology 2017.

39. Bristol-Myers Squibb. CheckMate 275: Clinical Study Report for Study CA209275 (25th July 2016).

40. Bristol-Myers Squibb. CheckMate 275 Addendum: Clinical Study Report for Study CA209275 (25th October 2016).

41. Galsky M, Retz M, Siefker-Radtke A, et al. Efficacy and safety of nivolumab monotherapy in patients with metastatic urothelial cancer (mUC) who have received prior treatment. Results from the phase II CheckMate 275 study. Presented at the European Society for Medical Oncology (ESMO) Copenhagen 2016. 2016.

42. Sharma P, Bono P, Kim J, et al. Efficacy and safety of nivolumab monotherapy in metastatic urothelial cancer (mUC): Results from the phase I/II CheckMate 032 study. 2016 ASCO Annual Meeting 2016.

43. Bristol-Myers Squibb. CheckMate 032: Clinical Study Report for Study CA209032 (29th June 2016).

44. Bristol-Myers Squibb. CA209-9GH mUC Report UK, 2016.

45. Clopper CJ, Pearson ES. The Use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial. Biometrika 1934;26:404-413.

46. Brookmeyer R, Crowley J. A Confidence Interval for the Median Survival Time. Biometrics 1982;38:29-41.

47. Klein JP, Moschberger ML. Survival Analysis: Techniques for Censored and Truncated Data. 2003.

48. Greenwood M. The errors of sampling of the survivorship tables. Reports on Public Health and Statistical Subjects, 33, Appendix 1: HMSO, London, 1926.

49. Kalbfleisch J, Prentice R. The statistical analysis of failure time data. 2002. Hoboken: John Wiley & Sons 1980;2.

50. Oing C, Rink M, Oechsle K, et al. Second Line Chemotherapy for Advanced and Metastatic Urothelial Carcinoma: Vinflunine and Beyond—A Comprehensive Review of the Current Literature. The Journal of urology 2016;195:254-263.

51. Centre for Reviews and Dissemination. Quality Assessment: Cohort Studies (Box 5.9). CRD Report 4: Undertaking systematic reviews of research on effectiveness. York: University of York, 2001.

52. Aaronson NK, Ahmedzai S, Bergman B, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. J Natl Cancer Inst 1993;85:365-76.

53. Osoba D, Rodrigues G, Myles J, et al. Interpreting the significance of changes in health-related quality-of-life scores. Journal of Clinical Oncology 1998;16:139-144.

54. Necchi A, Grimm M, Retz M, et al. Health-Related Quality of Life as a Marker of Treatment Benefit With Nivolumab in Platinum-Refractory Patients With Metastatic or Unresectable Urothelial Carcinoma From CheckMate 275. Poster presented at the American Society of Clinical Oncology (ASCO) Annual Meeting 2017; June 2–6, 2017; Chicago, IL, USA.

55. The EuroQol Group. EuroQol-a new facility for the measurement of health-related quality of life. Health policy 1990;16:199-208.

56. Pickard AS, Neary MP, Cella D. Estimation of minimally important differences in EQ-5D utility and VAS scores in cancer. Health and Quality of Life Outcomes 2007;5:70-70.

57. Bharmal M, Thomas J. Comparing the EQ-5D and the SF-6D descriptive systems to assess their ceiling effects in the US general population. Value in Health 2006;9:262-271.

58. Kim YS, Lee SI, Park SH, et al. A phase II study of weekly docetaxel as second-Llne chemotherapy in patients with metastatic urothelial carcinoma. Clinical Genitourinary Cancer 2016;14:76-81.

59. McCaffrey JA, Hilton S, Mazumdar M, et al. Phase II trial of docetaxel in patients with advanced or metastatic transitional-cell carcinoma. Journal of Clinical Oncology 1997;15:1853-7.

60. Vaughn DJ, Broome CM, Hussain M, et al. Phase II trial of weekly paclitaxel in patients with previously treated advanced urothelial cancer. Journal of Clinical Oncology 2002;20:937-40.

61. Gondo T, Ohori M, Hamada R, et al. The efficacy and safety of gemcitabine plus cisplatin regimen for patients with advanced urothelial carcinoma after failure of M-VAC regimen. International Journal of Clinical Oncology 2011;16:345-51.

62. Ozawa A, Tanji N, Ochi T, et al. Gemcitabine and cisplatin for advanced urothelial carcinomas: the Ehime University Hospital experience. International Journal of Clinical Oncology 2007;12:279-83.

63. Joly F, Houede N, Noal S, et al. Do patients with advanced urothelial carcinoma benefit from weekly paclitaxel chemotherapy? A GETUG phase II study. Clinical Genitourinary Cancer 2009;7:E28-33.

64. Phillippo DM, Ades AE, Dias S, et al. NICE DSU Technical Support Document 18: Methods for population-adjusted indirect comparisons in submissions to NICE. Volume 2016. Sheffield: NICE Decision Support Unit, 2016.

65. Jansen JP. Network meta-analysis of survival data with fractional polynomials. BMC medical research methodology 2011;11:61.

66. Sutton AJ, Welton NJ, Cooper N, et al. Evidence synthesis for decision making in healthcare: John Wiley & Sons, 2012.

67. Bristol-Myers Squibb. Data on File: Expert Clinician Opinion.

68. Dias S, Welton NJ, Sutton AJ, et al. NICE DSU Technical Support Document 2: A generalised linear modelling framework for pairwise and network meta-analysis of randomised controlled trials, 2011; last updated 2016.

69. Bristol-Myers Squibb. CheckMate 141: Clinical Study Report for Study CA209141 (7th June 2016).

70. McDermott D, Motzer R, Atkins M, et al. Long-term overall survival (OS) with nivolumab in previously treated patients with advanced renal cell carcinoma (aRCC) from phase I and II studies. J Clin Oncol 2016;34:abstr 4507.

71. Hodi SF, Kluger H, Sznol M, et al. Durable, long-term survival in previously treated patients with advanced melanoma (MEL) who received nivolumab (NIVO) monotherapy in a phase I trial. Cancer Research 2016;76:Suppl. Abstract CT001.

72. Ramalingam S, Lena H, Rizvi NA, et al. 137O - Nivolumab in patients (pts) with advanced refractory squamous (SQ) non-small cell lung cancer (NSCLC): 2-year follow-up from CheckMate 063 and exploratory cytokine profiling analyses. Presented at the European Lung Cancer Conference - Geneva, Switzerland 2016. Abstract number 137O. 2016.

73. Ascierto PA, Marincola FM. What have we learned from cancer immunotherapy in the last 3 years? J Transl Med 2014;12:141.

74. Schadendorf D, Hodi FS, Robert C, et al. Pooled Analysis of Long-Term Survival Data From Phase II and Phase III Trials of Ipilimumab in Unresectable or Metastatic Melanoma. Journal of Clinical Oncology 2015;33:1889-1894.

75. Pickard AS, Jiang R, Lin HW, et al. Using Patient-reported Outcomes to Compare Relative Burden of Cancer: EQ-5D and Functional Assessment of Cancer Therapy-General in Eleven Types of Cancer. Clinical Therapeutics 2016;38:769-777.

76. Latimer N. NICE DSU Technical Support Document 14: Undertaking survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data., 2011.

77. National Institute for Health and Care Excellence (NICE). [ID971]: Nivolumab for treating recurrent or metastatic squamous-cell carcinoma of the head and neck after platinum-based chemotherapy. Available: https://www.nice.org.uk/guidance/indevelopment/gid-ta10080 [Accessed 2 Feb 2017].

78. Anderson JR, Cain KC, Gelber RD. Analysis of survival by tumor response and other comparisons of time-to-event by outcome variables. J Clin Oncol 2008;26:3913-5.

79. Office for National Statistics. National life tables, 2013-2015.

80. Gettinger SN, Horn L, Gandhi L, et al. Overall Survival and Long-Term Safety of Nivolumab (Anti–Programmed Death 1 Antibody, BMS-936558, ONO-4538) in Patients With Previously Treated Advanced Non–Small-Cell Lung Cancer. Journal of Clinical Oncology 2015;33:2004-2012.

81. National Institute for Health and Care Excellence (NICE). [TA384]: Nivolumab for treating advanced (unresectable or metastatic) melanoma (2016). Available: https://www.nice.org.uk/guidance/ta384. Accessed: 9th June 2016.

82. National Institute for Health and Care Excellence (NICE). [TA400]: Nivolumab in combination with ipilimumab for treating advanced melanoma (2016). Available: https://www.nice.org.uk/guidance/ta400 [Accessed 2 Feb 2017].

83. National Institute for Health and Care Excellence (NICE). [TA428]: Pembrolizumab for treating PD-L1-positive non-small-cell lung cancer after chemotherapy (2017). Available: https://www.nice.org.uk/guidance/ta428 [Accessed: 22 June 2016].

84. National Institute for Health and Care Excellence (NICE). [ID990]: Pembrolizumab for untreated PD-L1 positive metastatic non-small-cell lung cancer (2017). Available: https://www.nice.org.uk/guidance/indevelopment/gid-ta10092 [Accessed: 22 June 2016].

85. Bellmunt J, Fougeray R, Rosenberg JE, et al. Long-term survival results of a randomized phase III trial of vinflunine plus best supportive care versus best supportive care alone in advanced urothelial carcinoma patients after failure of platinum-based chemotherapy. Annals of Oncology 2013;24:1466-1472.

86. Berdeaux G, Roskell N, Hemstock M. Influence of the method of analysis on estimates of QALY treatment difference: Phase iii trial of vinflunine versus best supportive care in patients with TCCU. Value in Health 2015;18 (7):A467.

87. Scottish Medicines Consortium. Vinflunine ditartrate 25mg/ml concentrate for solution for infusion (Javlor®). SMC No. (686/11). 2011. Available online at https://www.scottishmedicines.org.uk/files/advice/vinflunine_ditartrate_Javlor_FINAL_February_2011.doc_for_website.pdf. [Last accessed 27/02/2017].

88. Scottish Medicines Consortium. Vinflunine (as ditartrate), 25mg/mL, concentrate for solution for infusion (Javlor®) SMC No. (686/11). 2015. Available online at https://www.scottishmedicines.org.uk/files/advice/vinflunine_Javlor_Resubmission_FINAL_June_2015_for_website.pdf. [Last accessed 27/02/2017].

89. von der Maase H, Hansen SW, Roberts JT, et al. Gemcitabine and cisplatin versus methotrexate, vinblastine, doxorubicin, and cisplatin in advanced or metastatic bladder cancer: results of a large, randomized, multinational, multicenter, phase III study. J Clin Oncol 2000;18:3068-77.

90. Davey P, Rajan N, De Souza P, et al. Examining preferences, utility values and cost-effectiveness for gemcitabine plus cisplatin (GEM/cis) for the treatment of bladder cancer — A discrete choice conjoint analysis conducted in Australia. European Journal of Cancer 2001;37:S163.

91. Brown A, Aristides M, FitzGerald P, et al. Examining Preferences and Time-Trade-Off Utility for Gemcitabine Plus Cisplatin in the Treatment of Advanced Bladder Bancer: A Survey Using Discrete Choice Conjoint Analysis. . Value in Health 2002;5:543-544.

92. Robinson P, Maase H, Bhalla S, et al. Cost-utility analysis of the GC versus MVAC regimens for the treatment of locally advanced or metastatic bladder cancer. Expert Rev Pharmacoecon Outcomes Res 2004;4:27-38.

93. Little R, Rubin D. Statistical Analysis with Missing Data. Hoboken, New Jersey: Wiley-Interscience, 2002.

94. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. Stat Med 2011;30:377-99.

95. Attard CL, Brown S, Alloul K, et al. Cost-effectiveness of folfirinox for first-line treatment of metastatic pancreatic cancer. Curr Oncol 2014;21:e41-51.

96. Beusterien KM, Davies J, Leach M, et al. Population preference values for treatment outcomes in chronic lymphocytic leukaemia: a cross-sectional utility study. Health Qual Life Outcomes 2010;8:50.

97. Nafees B, Stafford M, Gavriel S, et al. Health state utilities for non small cell lung cancer. Health Qual Life Outcomes 2008;6:84.

98. National Institute for Health and Care Excellence. Nintedanib for previously treated locally advanced, metastatic, or locally recurrent nonsmallcell lung cancer [TA347], 2015.

99. Frederix GW, Quadri N, Hovels AM, et al. Utility and work productivity data for economic evaluation of breast cancer therapies in the Netherlands and Sweden. Clin Ther 2013;35:e1-7.

100. British National Formulary. 2017.

101. Department of Health. Drugs and pharmaceutical electronic market information (eMit). In: (CMU) CMU, ed, 2016.

102. Sacco JJ, Botten J, Macbeth F, et al. The average body surface area of adult cancer patients in the UK: a multicentre retrospective study. PLoS One 2010;5:e8933.

103. National Institute for Health and Care Excellence (NICE). [TA417]: Nivolumab for previously treated advanced renal cell carcinoma (2016). Available: https://www.nice.org.uk/guidance/ta417 [Accessed 2 Feb 2017].

104. Department of Health. NHS reference costs 2015 to 2016. NHS reference costs

105. Addicott R, Dewar S. Improving choice at end of life: a descriptive analysis of the impact and costs of the Marie Curie delivering choice programme in Lincolnshire: London: Kings Fund, 2008.

106. Personal Social Services Research Unit. Unit Costs of Health and Social Care 2015, 2015.

107. Personal Social Services Research Unit. Unit Costs of Health and Social Care 2015, 2016.

108. National Institute for Health and Care Excellence. Costing statement: Blood transfusion, 2015.

109. National Institute for Health and Care Excellence (NICE). [ID900]: Lung cancer (non-small-cell, non-squamous, metastatic, after treatment) - nivolumab. Available: https://www.nice.org.uk/guidance/indevelopment/gid-tag524 [Accessed 2 Feb 2017].

110. Briggs A, Claxton K, Sculpher M. Decision Modelling for Health Economic Evaluation. Oxford: Oxford University Press, 2008.

111. Brahmer JR, Horn L, Jackman D, et al. Five-Year Follow-up From the CA209-003 Study of Nivolumab in Previously Treated Advanced Non-Small Cell Lung Cancer: Clinical Characteristics of Long-term Survivors. Presented at the American Association for Cancer Research (AACR) 2017 Annual Meeting, Washington, DC, USA. 2017.

112. Sideris S, Aoun F, Zanaty M, et al. Efficacy of weekly paclitaxel treatment as a single agent chemotherapy following first-line cisplatin treatment in urothelial bladder cancer. Mol Clin Oncol 2016;4:1063-1067.

B.6 Appendices

Appendix C: Summary of product characteristics (SmPC) and European public assessment report (EPAR)
Appendix D: Identification, selection and synthesis of clinical evidence
Appendix E: Subgroup analysis
Appendix F: Adverse reactions
Appendix G: Published cost-effectiveness studies
Appendix H: Health-related quality-of-life studies
Appendix I: Cost and healthcare resource identification, measurement and valuation
Appendix J: Clinical outcomes and disaggregated results from the model
Appendix K: Checklist of confidential information
Appendix L: Survival analysis: supplementary information
Appendix M: Full eligibility criteria for CheckMate 275 and CheckMate 032
Appendix N: Additional utility data
Appendix O: Comparison with cisplatin plus gemcitabine

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Single technology appraisal

Nivolumab for treating metastatic or unresectable urothelial cancer after platinum-based chemotherapy [ID995]

Dear Company,

The Evidence Review Group, Kleijnen Systematic Reviews Ltd, and the technical team at NICE have looked at the submission received on 26 June 2017 from Bristol-Myers Squibb. In general they felt that it is well presented and clear. However, the ERG and the NICE technical team would like further clarification on the clinical and cost effectiveness data (see questions listed at end of letter).

The ERG and the technical team at NICE will be addressing these issues in their reports.

Please provide your written response to the clarification questions by **5pm on Wednesday 02 August 2017.** Your response and any supporting documents should be uploaded to NICE Docs.

Two versions of your written response should be submitted; one with academic/commercial-in-confidence information clearly marked and one with this information removed.

Please underline all confidential information, and separately highlight information that is submitted as commercial in confidence in turquoise, and all information submitted as academic in confidence in yellow.

If you present data that are not already referenced in the main body of your submission and that are academic/commercial in confidence, please complete the attached checklist for confidential information.

Please do not embed documents (PDFs or spreadsheets) in your response because this may result in them being lost or unreadable.

If you have any queries on the technical issues raised in this letter, please contact Thomas Paling, Technical Lead (Thomas.paling@nice.org.uk). Any procedural questions should be addressed to Kate Moore, Project Manager (kate.moore@nice.org.uk).

Yours sincerely

Helen Knight
Associate Director – Appraisals
Centre for Health Technology Evaluation

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Encl. checklist for confidential information

Section A: Clarification on effectiveness data

Literature searches

A1. **Priority question**: Please provide search strategies for the following databases listed in G.1 of Appendix G: EconLit, NHS EED, HTAD.

A2. Please clarify the results found in Appendix D, D.1.1, Table 4, search line #9. This search line appears to retrieve 0 records, however the ERG found 3687 records when replicating the search.

A3. Please provide details of the search terms used for the PubMed search listed in Appendix D, D1.1.

A4. Please provide details of the search terms used for the conference handsearching listed in G.1 of Appendix G.

A5. Please provide details of the search terms used for the following resources listed in G.1 of Appendix G: NICE, SMC, NCPE, CEA Registry, ScHARRHUD, EQ-5D Publications Database.

Nivolumab studies:

A6. **Priority question**: Please list the number of patients from each country in the two Checkmate studies, including numbers from the UK in each study.

A7. Both CheckMate studies are still ongoing. Please list any planned analyses after those reported in the company submission for each study. Are any further data available apart from those reported in the company submission?

A8. Please add details of the CheckMate studies to the following tables reported in Appendix D of the company submission: Table 17-Trial Design, page 69; Table 19-Trial methods, page 77; Table 21-patients' characteristics, page 86; and Table 23-statistical analysis, page 94.

A9. A. Please confirm that results for ORR and PFS from the latest database lock for CheckMate 275 (company submission-B, page 47) are based on BIRC assessment.

B. Please report investigator-assessed results for ORR and PFS from the latest database lock or CheckMate 275 as well (or BIRC results if it was not BIRC in the company submission).

A10. Please provide results for ORR, TTD, DOR and PFS for CheckMate 032 based on BIRC assessment.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

A11. **Priority question:** Can the company explain the differences in effectiveness of nivolumab in the CheckMate 275 and 032 studies? Nivolumab seems to be more effective in CheckMate 032. Although the difference is not statistically significant, it is consistent across all outcome measures.

A12. On page 58 (company submission, section B.2.8) it is mentioned that data from the CheckMate studies were pooled. Please provide details of the statistical method(s) used for pooling the data from Checkmate 275 and 032 and please explain which data were used (BIRC or investigator-assessed)**.** Please conduct all analyses using data from each method separately.

A13. A. Could the company discuss the generalisability of the CheckMate 275 and CheckMate 032 studies to the UK population, given that more than 50% in both studies had an ECOG performance status of 0?

B. How well do the Checkmate trials fit the UK population in terms of prior treatments received (type and setting of prior systemic therapy)?

C. A very small number of patients in the Checkmate trials have locally unresectable non-metastatic disease. Does thi s reflect the UK population and can the data from these patients be applied to the patients in the scope?

Comparator studies:

A14. Adverse events and Health Related Quality of Life (HRQoL) have been presented for nivolumab, but not for the comparators.

A. Please provide adverse events for all comparators in the same way as reported in Table 23 to 26 of the main submission (Section B.2.10.3, pages 72-78).

B. Please provide HRQoL data for all comparators.

Indirect comparisons

A15. Cisplatin + gemcitabine should be a comparator according to the scope. The company argues that the generalisability of the cisplatin + gemcitabine study is limited because patients were gemcitabine naive. However, they could still be considered as undergoing retreatment with a platinum-based chemotherapy even if the precise combination was different, as stated in the Comparators section of the scope: “Retreatment with 1st line platinum-based chemotherapy (only for people whose disease has had an adequate response)” Could the company explain why cisplatin + gemcitabine cannot be a comparator for patients who have had exposure to cisplatin?

A16. Please provide further details of the three trials excluded from the indirect comparison/mixed treatment comparison and why the doses/treatment regimens

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

were not considered to be in line with current UK clinical practice (See Appendix D.2.3 page 71).

A17. A. Please provide further details of the fractional polynomial network meta-analysis method, and how you judged whether the proportional hazards assumption did not hold, particularly when the Checkmate trials were single-arm only and could not be used to assess the proportional hazards assumption for nivolumab (See Section B.2.9.2 page 62)?

B. Please discuss methods other than the fractional polynomial for conducting the network meta-analysis including their pros and cons.

A18. A. **Priority question**: Please quantify the possible extent of any residual systematic error resulting from unobserved prognostic variables and effect modifiers, using the 'out-of-sample' method described in NICE DSU TSD 18.

B. It is argued on page 103 in the company submission that this method may not provide an accurate estimate of the residual bias. Please explain why and in which direction it differs from an accurate estimate.

A19. Pooled nivolumab data were simulated to match characteristics in the comparator studies. Therefore, it is important that inclusion criteria and population characteristics of comparator studies match the population described in the scope. Please discuss each of the comparator studies and describe whether they reflect the UK population described in the scope.

A20. Studies such as pazopanib vs. docetaxel, or docetaxel vs. BSC, or docetaxel+ ramucurimab vs. vinflunine could have been used to provide indirect comparisons in the meta-analysis conducted by the company i.e. the so called ITC. Were such studies searched for in the systematic literature review? Is it possible that such studies exist, but not found through the systematic literature review?

A21. **Priority question:** It is clear across all outcomes (including ORR, PFS and OS) that patients with PD-L1 < 1% expression do less well with nivolumab than those with PD-L1 >=1% expression.

A. Could the company please provide a justification as to why the 'Indirect treatment comparison' for this subgroup only was not performed?

B. Could the company please perform the 'Indirect treatment comparison' for this subgroup only?

A22. **Priority question**: The code for the 'indirect or mixed treatment comparison' is shown in Appendix D.2.7. Could the company also provide all of the data necessary for running these models so that the ERG can validate the results?

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

A23. Please provide evidence based on the effectiveness analyses that nivolumab provides an extension of life of at least three months compared to the comparators in order to fulfill the end-of-life criteria.

Section B: Clarification on cost-effectiveness data

Treatment effectiveness

B1. To derive nivolumab treatment effectiveness, the CheckMate 275 and 032 studies were pooled. This is inconsistent with how utilities, resource use and adverse event rates were derived (from the CheckMate 275 study only). Please justify why the treatment effectiveness data were derived from the pooled CheckMate 275 and 032 studies, but utilities, resource use and adverse event rates were derived from CheckMate 275 only.

B2. The company states that a response-based modelling approach was adopted in order to reflect the mechanism of action of nivolumab and to reflect that the nivolumab survival curve changes over time as the hazard changes. The company furthermore claims that standard parametric models are unlikely to be flexible enough to characterise this change in the hazard. However, most parametric distributions (except the exponential) can be used to incorporate changing hazards over time. Additionally, standard models (e.g. log-logistic, log-normal and generalised gamma distributions) even include a hazard function that is non-monotonic with respect to time (initially an increasing hazard, followed by a decreasing hazard).[1] Moreover, the NICE technical support document on survival analysis suggests spline-based models as useful, more flexible alternatives.[1] Please provide further justification for the response-based approach, and why landmark analysis was performed, in particular:

A. Please provide justification for why a response-based approach was necessary, including whether standard parametric curves (as described in the NICE technical support document on survival analysis) were tested and why they were deemed to not appropriately reflect nivolumab survival.[1]

B. Were other methods, such as spline-based models (see also TSD 14), or mixture cure models, considered?[1, 2] If so, why was a landmark analysis preferred? If not, please consider the advantages and disadvantages of these methods compared to the landmark analysis and consider implementation of the most suitable approach.

C. Please provide justification for the choice of the selected 8-week landmark using clinical expert opinion.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

D. Please analyse the impact of using different landmarks, by providing scenario analyses results (disaggregated) of alternative landmarks: 12 and 20 weeks.

E. Please provide justification for why no parametric curve was fitted to the Kaplan–Meier estimates prior to the 8-week landmark point. Please also provide the results of an analysis where a parametric curve is fitted to the data before the landmark point.

F. Please provide a scenario analysis where nivolumab patients are not analysed separately by response (i.e. OS and PFS curves fitted to all patients regardless of response status).

B3. **Priority question**: For the analysis of responders versus non-responders, proportionality of hazards was discarded, even though no analysis was presented to justify this. Furthermore, the responders' and non-responders' curves were combined using an average weighted by the 8-week responder proportion, thus artificially over-estimating the weight of non-responders in later periods.

A. Please explore whether proportionality of hazards is violated between responders and non-responders, using log cumulative hazard plots.

B. Please provide justification for, and describe the methods used for, combining the responders and non-responders' curves instead of modelling them separately by using additional health states in the model, and provide comment on the impact of this approach.

B4. **Priority question**: The time-varying hazard ratios are calculated by predicting survival of patients from the comparator studies if they would receive nivolumab based on a prediction model estimated on the pooled data from the CheckMate studies (i.e. not divided into the groups of responders vs non-responders). The hazard ratios obtained are then applied to the newly calculated survival curves that combined responders and non-responders. The model parameterisation of the fractional polynomial approach (i.e. which polynomials are chosen) has a large impact on model outcomes.

A. Please discuss the potential bias induced by deriving hazard ratios from one survival curve (fitted to all patients irrespective of response status) and then applying it to a different survival curve (the one that was derived from combining the responders and non-responders curves using a weighted average). Please provide justification for this approach.

B. Please provide hazard ratios derived for responders and non-responders separately.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

C. Please provide a scenario analysis incorporating hazard ratios that are estimated independent of time (i.e. fixed over time).

D. Please provide scenario analysis using further alternative model specifications, with polynomials other than p1=0, p2=0 and p1=1, p2=1. Please also describe how and justify why these particular polynomials are chosen.

B5. It is not clear why it is necessary to use the same survival model (generalised gamma distribution) for responders and non-responders.

A. Please provide justification for why it was deemed necessary to use the same survival model for responders and non-responders.

B. Were clinical experts consulted to support the choice of survival model? If so, please provide the methods for eliciting expert opinion including the number of experts and questions asked as well as the results.

C. Please provide an implementation in the model by which it is possible to use differential curves for responders and non-responders. Please also provide a scenario analysis, in which the best fitting curves are chosen separately for responders and non-responders, e.g. using the Weibull for non-responders' OS and PFS and, in two separate scenarios, the exponential and the generalised gamma for responders' OS and PFS.

D. PFS and OS curves were adjusted to account for general population mortality using age-adjusted annual mortality rates. Please discuss the method to implement this and provide justification for this approach. Please also justify why both, PFS and OS, had to be adjusted instead of just OS. Please also discuss whether this has any impact on the plausibility of the OS estimates.

B6. Time to treatment discontinuation (TTD) was estimated irrespective of response status (inconsistent with OS and PFS). However, treatment is discontinued when patients no longer benefit from it. It could therefore be suspected that TTD differs significantly for responders versus non-responders.

A. **Priority question**: Please implement survival models for TTD using the same response-based survival analysis as for PFS and OS (currently landmark analysis with 8-week landmark) in the cost effectiveness model and provide the results of this in a scenario analysis.

B. Please provide justification for the survival model choice for TTD, with description of the clinical expert opinion and methods to elicit this.

B7. The ERG noticed several inconsistencies between the Checkmate (032 and 275) trials, the company submission and the cost effectiveness model concerning the

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

number of responders used for the OS landmark analysis. Objective response was achieved in 19 and 52 patients in CheckMate 032 and 275, respectively, totalling 71 responders.[3, 4] However, the number of responders at the 8-week landmark for OS estimation, cells DD10 and DE10 of the cost effectiveness model, is 73 patients. In addition, the numbers of responders and non-responders provided in Figure 35 of the company submission for the PFS landmark analysis do not correspond to the number of responders used in the cost effectiveness model. Please clarify which figures are correct for the PFS and OS landmark analyses, amend the cost effectiveness model if necessary and provide the cost effectiveness results using the correct number of responders.

B8. **Priority question**: Please provide the comparison of nivolumab against cisplatin + gemcitabine in the base-case (see also question A15).

B9. The company assumes that the hazard ratio of BSC versus vinflunine can be applied to the paclitaxel PFS curve in order to calculate PFS for the BSC comparator. The company justifies this assumption by stating that the outcomes between vinflunine and paclitaxel/docetaxel are expected to be similar and therefore that this hazard ratio can be applied to the paclitaxel PFS curve in order to obtain PFS estimations for BSC. However, no evidence is provided to support this assumption. Please provide clinical evidence to support this assumption.

Model structure

B10. **Priority question**: Please provide an implementation of the model, in which there are separate health states for responders and non-responders (instead of using PFS and OS based on weighted averages).

A. Please add an implementation with differential hazard ratios for OS and PFS for responders and non-responders.

B. Please discuss the plausibility of applying differential utility values and resource use for responders and non-responders, and apply these if applicable.

B11. The company uses a partitioned survival model approach. Could the company provide additional justification for this approach, other than that it has been used in previous appraisals on metastatic cancers and TA272, especially in the light of criticism of partitioned survival models compared with state transition models according to TSD19, which includes that endpoints are treated as independent and that intermediate health states are not reflected?[5]

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Adverse events

B12. The company uses adverse event rates for the comparators from sources that are only named briefly in the company submission, without explanation. Please provide an overview of and justification for the chosen sources.

Probabilistic sensitivity analysis

B13. The probabilistic sensitivity analysis (PSA) produces errors, does not include all input parameters, uses a small number of simulations, produces results that are different from the deterministic analysis and does not appear to be reproducible. Full incremental results are not provided.

A. **Priority question**: Please provide full incremental analysis with all comparators included in the PSA simultaneously, showing incremental costs and QALYs of nivolumab and all comparators.

B. In addition to the incremental costs and QALYs provided in Table 47 of the company submission, please also provide absolute costs and QALYs resulting from the PSA.

C. Please include the Kaplan–Meier estimates and hazard ratios in the PSA.

D. After re-running the PSA, the PSA results in #NUM errors in both nivolumab and comparator costs and QALYs. Please provide a corrected PSA, which does not produce errors.

E. Please comment on the reasons for which OS and PFS may be lower in the probabilistic compared with the deterministic analysis.

F. The ERG did not obtain probabilistic results similar to those reported by the company. Could the company ensure the reproducibility of the PSA and provide a version of the model with identical PSA results as provided in the company submission?

G. Please increase the number of PSA simulations to at least 10,000 PSA simulations (or more if needed to provide reproducible PSA results).

Health-related quality of life

B14. The company explains that 204 observations of utilities were missing because of the immaturity of the dataset. The dataset was finalised 2nd September 2016.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

A. Please justify the choice to impute future observations (i.e. questionnaires that were due to be administered after the data-cut of 2nd September 2016) and discuss the implications.

B. Please provide the analyses presented in the company submission using a more recent data-cut.

B15. Utility estimates are derived from CheckMate 275 only, disregarding CheckMate 032 and NICE TA 272.

A. Please comment on the reasons for disregarding the utility estimates used in NICE TA 272, discuss how the utility estimates derived from CheckMate 275 differ, and discuss the implications for model outcomes.

B. Please comment on the reasons for disregarding the utility estimates from CheckMate 032, discuss how the utility estimates derived from CheckMate 275 differ, and discuss the implication for model outcomes.

C. Please provide an analysis using the utility estimates from both CheckMate studies.

B16. Please provide justification for, and discuss the suitability of, the approach used to obtain utility values.

A. Please report the deviation in time for the interpolated observations (i.e. number of cases, mean, median, standard deviation, minimum and maximum).

B. Please specify why it was deemed necessary to impute data despite the use of a mixed model (which has methods to deal with missing data).

C. Please discuss the differences in methods and results of observed and imputed values in Table 35 in the company submission.

D. Please justify why predictive mean matching was chosen as the imputation method and the limitations associated with this method in the context of a large amount of missing data. Additionally, please provide additional details regarding the imputation methods (e.g. which variables were included) and discuss the plausibility of the imputed data.

E. Please justify why a mixed model was used and provide diagnostics of the mixed model.

F. Please explore adding a variable for a patient being on treatment or not to the mixed model and adding a variable for response status to the mixed model, provide results and discuss the impact on model outcomes.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

G. The company submission states that 'the generated utility values for post-progression patients was seen to increase and decrease in a manner that would not be expected in clinical practice.' Please explain how it was determined that the time-dependent utilities obtained were not in line with clinical practice.

Resource use and costs

B17. With regard to the calculation of drug and administration costs:

A. Please comment on the reasons for disregarding resource use (e.g. for calculating drug and administration costs) from CheckMate 032, discuss how the resource use derived from CheckMate 275 differs, and discuss the implications of this.

B. Please provide justification for classing delayed doses as missed doses and discuss the impact of this assumption.

C. Please provide justification (other than absence of evidence) for assuming that the dose intensity for docetaxel, paclitaxel, gemcitabine plus cisplatin was equal to that of nivolumab.

D. Please provide justification that the weight and body surface area from CheckMate 275, used to calculate drug costs, is applicable to patients in the UK setting.

E. In the company submission it states 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)". Please provide justification why 3 gemcitabine administrations were assumed per 4 weeks instead of 2.67 administrations per 4 weeks (=2 × 28 / 21).

F. Please provide justification for why administration costs for cisplatin were incorporated in addition to the gemcitabine administration costs for the cisplatin plus gemcitabine regime (as cisplatin and gemcitabine are both administered on day 1).

B18. TA272 (the only other NICE submission in this indication) was identified in the systematic review for cost-effectiveness evidence.

A. Please provide justification for why TA272 was not used to inform costs and resource use.

B. Please provide explanations for discrepancies with TA272 in terms of monitoring costs in the present assessment that range from £272.44 to £555.50 per 4 weeks

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

while in TA 272 this is £3.18 per treatment cycle of 21 days (see company submission of TA272 Table B37).

C. Please provide explanations for discrepancies with TA272 in terms of BSC costs of £170.21 per 4 weeks in the present assessment while in TA 272 this is £580 and £1,253 per month for pre progression and post progression respectively.

Model validation

B19. The company states that expert opinion has been used to validate OS and PFS predictions of the model for nivolumab and the comparator. However, the company does not provide any information on the number and identification of experts, or the methods used. Could the company please provide the number of experts that were consulted, the methods used, and questions asked to elicit opinion about OS and PFS predictions for nivolumab and the comparators, and an overview of each expert's opinion/statement.

Subgroup analysis

B20. Referring to Question A21, please provide a subgroup analysis using PFS and OS for patients with PD-L1 < 1% and those with PD-L1 >=1% expression along with the corresponding probabilistic results (expected ICER and cost-effectiveness acceptability curves.

Section C: Textual clarifications and additional points

C1. The (blue and green) curves presented in Figures 34 and 35 of the company submission and Figures 32 to 41 of Appendix L are all labelled 'non-responder'. Please correct the labels such that they correspond to their respective subgroups.

[1] Latimer N. NICE DSU Technical Support Document 14: Undertaking survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data. Sheffield: Decision Support Unit, ScHARR, 2017. 52p. Available from: http://scharr.dept.shef.ac.uk/nicedsu/wp-content/uploads/sites/7/2016/03/NICE-DSU-TSD-Survival-analysis.updated-March-2013.v2.pdf

[2] Othus M, Bansal A, Koepl L, Wagner S, Ramsey S. Accounting for cured patients in cost-effectiveness analysis. Value Health 2017;20(4):705-709.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

[3] Bristol-Myers Squibb Pharmaceuticals Ltd. CheckMate 032: Clinical Study Report for Study CA209032 (29th June 2016). 2016

[4] Bristol-Myers Squibb Pharmaceuticals Ltd. CheckMate 275: Clinical Study Report for Study CA209275 (25th July 2016), 2016

[5] Woods B, Sideris E, Palmer S, Latimer N, Soares M. NICE DSU Technical Support Document 19: Partitioned survival analysis for decision modelling in health care: a critical review. Sheffield: Decision Support Unit, ScHARR, 2017. 72p. Available from: http://scharr.dept.shef.ac.uk/nicedsu/wp-content/uploads/sites/7/2017/06/Partitioned-Survival-Analysis-final-report.pdf

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Single technology appraisal

Nivolumab for treating metastatic or unresectable urothelial cancer after platinum-based chemotherapy [ID995]

Dear Helen,

Thank you for the opportunity to respond to the clarification questions from the Evidence Review Group, Kleijnen Systematic Reviews, and the technical team at NICE. We thank the team for their general comments on the submission and hope that our responses to the individual questions in turn below provide clarity for our approach in the submission and the necessary additional information where this has been possible.

As requested, we have uploaded to NICE Docs two versions of this response letter: one with academic/commercial-in-confidence information clearly marked and one with this information removed. Accompanying these response letters is also a zipped folder data package, containing the code and supportive data referred to within this response.

Please do not hesitate to get in touch should you have any questions regarding our response.

Kind regards,

Sarah Breen

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Section A: Clarification on effectiveness data

Literature searches

A1. **Priority question**: Please provide search strategies for the following databases listed in G.1 of Appendix G: EconLit, NHS EED, HTAD.

The search strategy employed for the searches in EconLit as part of the systematic literature review for economic studies is provided below.

Term groups	ID	Search strings	Hits (02/12/16)
Disease area: advanced, metastatic or unresectable bladder cancer	1	"bladder cancer"	12

The Health Technology Assessment Database (HTAD) and the NHS Economic Evaluation Database (NHS-EED) were searched as part of the systematic literature review for economic studies via the Cochrane Library Wiley Online platform. The search strategy employed is provided below.

Health Technology Assessment Database: Issue 4 of 4, October 2016
NHS Economic Evaluation Database: Issue 2 of 4, April 2015

Term groups	ID	Search	Hits (02/12/16)
Disease area: advanced, metastatic or unresectable bladder cancer	#1	[mh "urinary bladder neoplasms"] or [mh "carcinoma, transitional cell"] or [mh "ureteral neoplasms"] or [mh ^"bladder neoplasms"] or [mh ^"urethral neoplasms"]	1227
	#2	((bladder* or urethra* or ureter* or urin* or urotheli* or renal pelvis or calice*) near/3 (cancer* or carcinoma* or adenoma* or adenocarcinoma* or squamous* or neoplas* or tum?r* or malignan*)):ti,ab,kw	2768
	#3	#1 or #2	2786
	#4	[mh "Neoplasm metastasis"] or (metastat* or metastas* or advanced or stage III or "stage 3" or stage IIIa or stage 3a or stage IIIb or stage 3b or stage IIIc or stage 3c or stage IV or "stage 4" or unresectable or non-resectable or nonresectable or inoperable or progressive)	67376
	#5	#3 and #4	682
Total	#6	#5 in Technology Assessments	4
	#7	#5 in Economic Evaluations	3

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

A2. Please clarify the results found in Appendix D, D.1.1, Table 4, search line #9. This search line appears to retrieve 0 records, however the ERG found 3687 records when replicating the search.

Thank you for your comment. We note that there was a typographical error in search line #9 Table 4, Table 5, Table 6 and Table 7. In each of these lines the search line included a full stop and this prevented the retrieval of any records. We have re-run these searches without the full stop and retrieved the following record numbers for search line #9 for each database.

- 621 records for Central (Total 495 records identified in original March searches, 126 new records retrieved in July)
- 13 records for Dare (Total 11 records identified in original March searches, 2 new records retrieved in July)
- 4 records for HTA (Total of 4 records identified in original March searches, no new records retrieved in July)
- 2 records for NHS EED (Total of 2 records identified in original March searches, no new records retrieved in July)

We reviewed the 128 new records excluding those that were published since the original March searches were conducted. No new relevant records were identified following the correction to search line #9.

A3. Please provide details of the search terms used for the PubMed search listed in Appendix D, D1.1.

Thank you for your comment. This was a typographical error. PubMed was not searched for this review. It was not listed as a database for searching in the protocol and therefore should not have been listed in Appendix D, D.1.1.

A4. Please provide details of the search terms used for the conference handsearching listed in G.1 of Appendix G.

The search terms used for the conference handsearching for the systematic literature review of economic studies are provided below.

Conference	Link	Search Strategy
American Society of Clinical Oncology (ASCO):	meetinglibrary.asco.org/abstracts (Keywords)	The website was searched for:

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

• 2014 ASCO Annual Meeting • 2015 ASCO Annual Meeting • 2016 ASCO Annual Meeting		• Bladder • Transitional cell • Urothelial
European Association of Urology (EAU): • 2014 Annual EAU Congress • 2015 Annual EAU Congress • 2016 Annual EAU Congress	2014: http://www.sciencedirect.com/science/journal/15699056/13/1 (Vol 13 Issue 1) 2015: http://www.sciencedirect.com/science/journal/15699056/14/2 (Vol 14 Issue 2) 2016: http://www.sciencedirect.com/science/journal/15699056/15/3 (Vol 15 Issue 3)	The websites/abstract books were searched for: Economic evaluations and cost/resource use studies: • Cost • Resource Utilities studies: • EQ-5D • EuroQoL • Utilit • Quality of life • HRQoL • QoL
European Multidisciplinary Meeting on Urological Cancers (EMUC) • EMUC 2014 • EMUC 2015 • EMUC 2016	2014: http://emuc2014.uroweb.org/uploads/emuc2014.uroweb.org/eau_paragraph_downloads/8/file/EMUC2014_abstract_&_programme__book.compressed.pdf 2015: http://emuc15.uroweb.org/uploads/emuc2015.uroweb.org/eau_paragraph_downloads/13/file/EMUC15_Abstract-Programme_Book_FINAL_VERSION_lr.pdf 2016: http://www.sciencedirect.com/science/journal/15699056/15/13	The websites were searched for: Economic evaluations and cost/resource use studies: • Cost • Resource Utilities studies: • EQ-5D • EuroQoL • Utilit • Quality of life • HRQoL • QoL
European Society for Medical Oncology (ESMO) Annual Meeting • ESMO 2014 • ECC 2015 • ESMO 2016	ESMO 2014: https://cslide.ctimeetingtech.com/library/esmo/browse/search/7Ca ECC 2015: (Abstract Body) www.eccocongress.org/Vienna2015/Scientific-Programme/Abstract-search ESMO Congress 2016: https://academic.oup.com/annonc/issue/27/suppl_6	The websites/abstract books were searched for: • Bladder • Transitional cell • Urothelial
ISPOR Annual International Meeting and Annual European Congress	www.ispor.org/RESEARCH_STUDY_DIGEST/research_index.asp	Each meeting in the "Meeting" drop-down menu was selected and the

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

• 2014 • 2015 • 2016		following terms were searched for in "abstract" • Bladder • Transitional cell • Urothelial

A5. Please provide details of the search terms used for the following resources listed in G.1 of Appendix G: NICE, SMC, NCPE, CEA Registry, ScHARRHUD, EQ-5D Publications Database.

The search terms used for the searching of the NICE, SMC, NCPE websites, the CEA Registry, ScHARRHUD and the EQ-5D publications database for the systematic literature review of economic studies are provided below:

Conference	Link	Search Strategy
The National Institute for Health and Care Excellence (NICE)	https://www.nice.org.uk/	The website was searched for: • Bladder • Transitional cell • Urothelial
Scottish Medicines Consortium (SMC)	https://www.scottishmedicines.org.uk/	The website was searched for: • Bladder • Transitional cell • Urothelial
National Centre for Pharmacoeconomics (NCPE)	http://www.ncpe.ie/	The website was searched for: • Bladder • Transitional cell • Urothelial
The Cost-effectiveness Analysis (CEA) Registry, managed by Tufts Medical Center	healtheconomics.tuftsmedicalcenter.org/cear4/SearchingtheCEARegistry/SearchtheCEARegistry.aspx	The website was searched for: • Bladder • Transitional cell • Urothelial
The University of Sheffield Health Utilities Database	www.scharrhud.org/	The website was searched for: • Bladder • Transitional cell • Urothelial
The EQ-5D Publications Database	www.euroqol.org/eq-5d-publications/search.html	The website was searched for:

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

		• Bladder • Transitional cell • Urothelial

Nivolumab studies:

A6. **Priority question**: Please list the number of patients from each country in the two Checkmate studies, including numbers from the UK in each study.

The number of patients treated in each country in CheckMate 275 and CheckMate 032 are provided below. There were no UK sites in CheckMate 275. In CheckMate 032, there were 6 patients (7.7%) treated in the study in the UK.

CheckMate 275

Country	Number treated (%) (Total N=270)
Australia	6 (2.2)
Belgium	7 (2.6)
Czech Republic	3 (1.1)
Finland	3 (1.1)
Germany	45 (16.7)
Italy	34 (12.6)
Japan	23 (8.5)
Poland	11 (4.1)
Spain	27 (10.0)
Sweden	5 (1.9)
United States	106 (39.3)

CheckMate 032

Country	Number treated (%) (Total N=78)
United Kingdom	6 (7.7)
Finland	2 (2.6)
Germany	3 (3.8)
Spain	8 (10.3)
United States	59 (75.6)

A7. Both CheckMate studies are still ongoing. Please list any planned analyses after those reported in the company submission for each study. Are any further data available apart from those reported in the company submission?

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

As highlighted in the submission, both CheckMate 275 and CheckMate 032 are still ongoing and interim analyses are currently planned following the next database locks for CheckMate 275 and CheckMate 032 in [redacted] and [redacted], respectively. Further database locks are not currently formally planned, but may be scheduled in the future as required.

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

A8. Please add details of the CheckMate studies to the following tables reported in Appendix D of the company submission: Table 17-Trial Design, page 69; Table 19-Trial methods, page 77; Table 21-patients' characteristics, page 86; and Table 23- statistical analysis, page 94.

Table 1: Clinical effectiveness: single-arm trials

Trial ID	Population	Intervention	Reported outcomes specified in the decision problem	All other reported outcomes
Included in indirect treatment comparison				
Sharma *et al.* (2017) [CheckMate 275][1]	Patients with locally advanced unresectable or metastatic UC who had progressed or recurred after at least one previous line of platinum-containing chemotherapy	Nivolumab (IV 3 mg/kg Q2W)	ORR, OS, PFS, HRQoL, adverse events	Duration of response (DoR) and additional safety outcomes
Sharma *et al.* (2016) [CheckMate 032][2]	Patients with locally advanced unresectable or metastatic UC who had progressed or recurred after treatment with at least one platinum-containing chemotherapy regimen	Nivolumab (IV 3 mg/kg Q2W)	ORR, OS, PFS, HRQoL, adverse events	DoR and additional safety outcomes
Gondo *et al.* (2011)[3]	Patients with histologically confirmed advanced and metastatic UC. All patients had evidence of disease progression, relapse or no response after MVAC chemotherapy as first-line treatment.	Gemcitabine (1,000 mg/m^2; D1, D8, D15); Cisplatin (35 mg/m2; D1, D2); 28 day-cycle;	OS, ORR	Toxicity
Joly *et al.* (2009)[4]	Patients had urothelial carcinoma of the bladder, or urothelial tract, with a progressive measurable disease after previous line of chemotherapy for advanced disease (neoadjuvant, adjuvant, or metastatic therapy), life expectancy ≥3 months, WHO performance status of 0-2	Paclitaxel (80mg/m^2 IV over 1 hour, D1, D8, D15); 28 day course;	ORR	CR, PR, SD
Ozawa *et al.* (2007)[5]	Patients had histological or cytological proof of UC, at least one bi-dimensionally measurable lesion according to WHO criteria, and a WHO performance status <2	Gemcitabine (1000mg/m^2 D1, D8, D15) Cisplatin (70mg/m^2	ORR	Toxicity

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Trial ID	Population	Intervention	Reported outcomes specified in the decision problem	All other reported outcomes
		D2); Every 28 days		
Not included in indirect treatment comparison				
Kim *et al.* (2016)[6]	Patients has pathologic proof of advanced or metastatic TCC or the urothelial tract, and were refractory to or relapsed after no more than 1 prior cisplatin-containing treatment. All patients were required to have at least 60% Karnofsky performance status and a at keast one measurable indicator lesion not irradiated and ≥2 cm	Docetaxel (100 mg/m^2 IV over 1 hour); Every 21 days	OS	PR, DoR, TTR, toxicity
McCaffrey *et al.* (1997)[7]	Patients with histologically confirmed UC, measurable lesions, ECOG PS 0 or 1, with documented progression after ≥1 previous platinum-based chemotherapy for advanced of metastatic disease (adjuvant if progressed within 6 months of last dose)	Docetaxel (30 mg/m^2 IV over 1 hour, D1, D8); Every 21 days	OS, PFS,	PR, toxicity
Vaughn *et al.* (2002)[8]	Patients with histologically confirmed bidimensionally measurable carcinoma of the urothelium, evidence of progressing regional or metastatic disease and ECOG performance status 0-2. Patients received at least one prior treatment for advanced UC.	Paclitaxel (80 mg/m^2 IV over 1 hour); Every 1 week for 4 weeks	ORR, PFS, OS	PR, toxicity

Abbreviations: CR: complete response; D: day; DoR: duration of response; ECOG: Eastern Cooperative Oncology Group; HRQoL: health-related quality of life; IV: intravenous; IV: intravenous; MVAC: methotrexate, vinblastine; adriamycin (doxorubicin) and cisplatin; NA: not applicable; ORR: objective response rate; OS: overall survival; PFS: progression-free survival; PR: partial response; SD: stable disease; TCC: transitional cell carcinoma; TTR: time to response; UC: urothelial carcinoma; WHO: World Health Organization.

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Table 2: Trial methods: single-arm trials

Trial ID	Location (number of centres)	Eligibility criteria for participants	Trial drugs	Treatment (n)	Permitted and disallowed concomitant medication	Primary outcomes	Other outcomes used in the economic model / specified in the scope	Pre-planned subgroups
Sharma *et al.* (2017) [CheckMate 275][1]	Australia, Belgium, Czech Republic, Finland, Germany, Italy, Japan, Poland, Spain, Sweden, USA (63)	**Key inclusion criteria** • Males and females ≥18 years of age with an ECOG PS 0 or 1, • Histologically or cytologically confirmed metastatic or surgically unresectable transitional cell carcinoma of the urothelium involving the bladder, urethra, ureter, or renal pelvis. • Measurable disease by CT or MRI per RECIST v1.1 criteria, • Progression or recurrence after treatment either: ○ With at least 1 platinum-containing chemotherapy regimen for metastatic or surgically unresectable locally advanced urothelial cancer, or ○ Within 12 months of peri-operative (neo-adjuvant or adjuvant) treatment with platinum agent in the setting of cystectomy for localised muscle-invasive urothelial cancer. • Patients that had received more than 2	Nivolumab 3 mg/kg Q2W	270	Disallowed: Immunosuppressive agents (except to treat a drug-related adverse events) or systemic corticosteroids (>10 mg daily prednisone equivalent) within 14 days of study drug administration, any antibody or drug specifically targeting T-cell co-stimulation or checkpoint pathways, or chemotherapy, radiation therapy, biologics for cancer, or investigational	BIRC-assessed ORR (RECIST v1.1)	BIRC-assessed PFS, OS and investigator-assessed ORR, PFS, safety, HRQoL (EORTC QLQ-C30 and EQ-5D-3L)	Patients with PD-L1 expression <1% and ≥1%

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Trial ID	Location (number of centres)	Eligibility criteria for participants	Trial drugs	Treatment (n)	Permitted and disallowed concomitant medication	Primary outcomes	Other outcomes used in the economic model / specified in the scope	Pre-planned subgroups
		prior lines of chemotherapy must not have had liver metastases. • Availability of tumour samples for PD-L1 expression analysis • Previous palliative radiotherapy must have been completed at least 2 weeks before administration of the study drug **Key exclusion criteria** • Active brain or leptomeningeal metastases • Active, known or suspected autoimmune disease • Previous malignancy active within the previous 3 years (except locally curable cancers that appeared to have been cured or carcinoma in situ) • Any serious or uncontrolled medical disorder • Autoimmune disease (vitiligo, type 1 diabetes mellitus, residual hypothyroidism due to an autoimmune condition only requiring hormone replacement, psoriasis not requiring systemic treatment, or conditions not expected to recur in the			therapy within 28 days of first study drug administration			

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Trial ID	Location (number of centres)	Eligibility criteria for participants	Trial drugs	Treatment (n)	Permitted and disallowed concomitant medication	Primary outcomes	Other outcomes used in the economic model / specified in the scope	Pre-planned subgroups
		absence of an external trigger were permitted) • Systemic treatment with either corticosteroids (>10 mg daily prednisone equivalents) or other immunosuppressive medications within 14 days of first study drug administration • Prior treatment with an anti-PD-1, anti-PD-L1, anti-PD-L2, anti-CTLA-4 antibody, anti-CD137, or any other antibody or drug specifically targeting T-cell co-stimulation or immune checkpoint pathways • Treatment with any chemotherapy, radiation therapy, biologics for cancer, or investigational therapy within 28 days of first study drug administration • All toxicities attributed to previous anticancer therapy other than neuropathy, alopecia, and fatigue must have resolved to grade 1 or baseline before administration of study drug.						

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Trial ID	Location (number of centres)	Eligibility criteria for participants	Trial drugs	Treatment (n)	Permitted and disallowed concomitant medication	Primary outcomes	Other outcomes used in the economic model / specified in the scope	Pre-planned subgroups
Sharma *et al.* (2016) [CheckMate 032][2]	Finland, Germany, Spain, UK and USA (16)	**Key inclusion criteria** • Males and females ≥18 years of age with an ECOG PS 0 or 1 • Measurable disease by CT or MRI per RECIST v1.1 criteria • Locally advanced or metastatic urothelial cell carcinoma • Progression or recurrence either: ○ After at least 1 previous platinum-containing chemotherapy treatment for metastatic or locally advanced unresectable urothelial cancer, or ○ Recurrence within 1 year of completing previous platinum-based neoadjuvant or adjuvant treatment ○ After previously refusing standard treatment with chemotherapy for the treatment of metastatic (stage IV) or locally advanced disease **Key exclusion criteria** • Active brain metastases or leptomeningeal metastases • Any serious or uncontrolled medical	Nivolumab 3 mg/kg Q2W	78	Disallowed: Immunosuppressive agents (except to treat a drug-related adverse event), systemic corticosteroids >10 mg daily prednisone equivalent, any concurrent antineoplastic therapy Permitted: Supportive care for disease-related symptoms, palliative (limited-field) radiation therapy and palliative surgical resection permitted if the certain protocol-defined criteria were met	Confirmed investigator-assessed ORR (RECIST 1.1)	Investigator-assessed PFS, OS, DOR, safety, HRQoL (EQ-5D)	ORR, OS and PFS analysed in subgroups defined by PD-L1 expression (<1% and ≥1%)

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Trial ID	Location (number of centres)	Eligibility criteria for participants	Trial drugs	Treatment (n)	Permitted and disallowed concomitant medication	Primary outcomes	Other outcomes used in the economic model / specified in the scope	Pre-planned subgroups
		disorder • History of or active, known or suspected autoimmune disease (vitiligo, type 1 diabetes mellitus, residual hypothyroidism caused by auto immune thyroiditis, and disorders not expected to recur in the absence of an external trigger were permitted) • Need for immunosuppressive doses of systemic corticosteroids (>10 mg daily prednisone equivalents) for at least 2 weeks before study drug administration • Prior treatment with experimental anti-tumour vaccines or any modulator of T-cell function or checkpoint pathway						
Gondo *et al.* (2011)[3]	Japan (1)	Histologically confirmed advanced and metastatic UC. Evidence of disease progression, relapse or no response after MVAC chemotherapy as first-line treatment. As MVAC chemotherapy, methotrexate was given at a dose of 30 mg/m^2 on day 1, vinblastine was given at a dose of 3 mg/m^2 on day 2, adriamycin was given at a dose of 30 mg/m^2 also on day 2, and cisplatin was given at a	Gemcitabine (1,000 mg/m^2; D1, D8, D15); Cisplatin (35 mg/m^2; D1, D2); 28 day-cycle;	33	Supportive care, including anti-emetics, analgesics, blood transfusions, and antibiotics, were administered as appropriate. G-CSF was not used	OS, ORR, survival, toxicity. RECIST 1.1; 1 or 2 cycles	NR	Baseline prognostic factors were explored

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Trial ID	Location (number of centres)	Eligibility criteria for participants	Trial drugs	Treatment (n)	Permitted and disallowed concomitant medication	Primary outcomes	Other outcomes used in the economic model / specified in the scope	Pre-planned subgroups
		dose of 35 mg/m^2 on day 2 and 3. ECOG PS ≤1, an adequate bone marrow reserve, that is, WBC count 3.5 x 10^9/l, platelets ≥100 x 10^9/l, and haemoglobin ≥8.0 g/dl, and no signs of CNS metastasis			routinely, but it was administered when granulocytes measured less than 500/µl. No other antineoplastic therapy was permitted during the study.			
Joly *et al.* (2009)[4]	France (NR)	Urothelial carcinoma of the bladder, or urothelial tract, with a progressive measurable disease after previous line of chemotherapy for advanced disease (neoadjuvant, adjuvant, or metastatic therapy), life expectancy ≥3 months, WHO PS of 0-2, normal baseline hematologic parameters, serum bilirubin level ≤ 1.5 normal limits, and transaminases and ALP <2.5 normal limits. Received taxanes in a 3-week schedule in first-line regimen	Paclitaxel (80mg/m^2 IV over 1 hour, D1, D8, D15); 28-day course	45	Dexamethasone, Dexchlorpheniramine, and Ranitidine premedication, given IV 30 minutes before paclitaxel.	ORR (complete response, partial response and stable disease). RECIST; Every 8 weeks (every 2 cycles)	ORR	NR
Ozawa *et al.* (2007)[5]	Japan (1)	Histological or cytological proof of UC, at least one bi-dimensionally measurable lesion according to WHO criteria, and a WHO PS of less than 2	Gemcitabine (1000mg/m^2 D1, D8,	30	NR	ORR, toxicity; WHO (1979);	NR	Patients who had not received

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Trial ID	Location (number of centres)	Eligibility criteria for participants	Trial drugs	Treatment (n)	Permitted and disallowed concomitant medication	Primary outcomes	Other outcomes used in the economic model / specified in the scope	Pre-planned subgroups
			D15) Cisplatin (70mg/m^2 D2); Every 28 days			Unclear		surgery due to metastatic disease

Abbreviations: ALP: alkaline phosphatase; BIRC: blinded independent review committee; D: day; ECOG: Eastern Cooperative Oncology Group; EORTC QLQ-C30: European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30;EQ-5D-3L: EuroQoL 5-Dimensions 3-Levels; IV: intravenous; MVAC: methotrexate, vinblastine, adriamycin (doxorubicin) and cisplatin; NR: not reported; ORR: objective response rate; OS: overall survival; PD-L1: programmed death ligand 1; PFS: progression-free survival; PS: performance status; RECIST: Response Evaluation Criteria in Solid Tumours; UC: urothelial carcinoma; WBC: white blood cell; WHO: World Health Organization.

Table 3: Patients' characteristics: single-arm trials

Trial, treatment arm and population size	Age (years)	Males n (%)	ECOG status n (%)	Location of urothelial cancer n (%)	Presence and location of metastasis n (%)	Prior neoadjuvant or adjuvant treatment n (%)	Prior type of chemotherapy n (%)	Prior radiotherapy n (%)	Prior surgery n (%)	Response to prior chemotherapy n (%)
Sharma *et al.* (2017) [CheckMate 275][1]	Median 66 (38-90)	211 (78.1)	0: 145 (53.7) 1: 124 (45.9)	Urinary bladder: 197 (73.0)	Visceral: 227 (84.1) Liver: 75 (27.8)	Adjuvant: 83 (30.7) Neo-adjuvant: 60 (22.2)	Cisplatin and gemcitabine: 87 (32.2)	85 (31.5)	250 (92.6)	CR: 23 (8.6) PR: 44 (16.4) SD: 51 (19.0) PD: 88 (32.7)

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Trial, treatment arm and population size	Age (years)	Males n (%)	ECOG status n (%)	Location of urothelial cancer n (%)	Presence and location of metastasis n (%)	Prior neoadjuvant or adjuvant treatment n (%)	Prior type of chemotherapy n (%)	Prior radiotherapy n (%)	Prior surgery n (%)	Response to prior chemotherapy n (%)
			3: 1 (0.3)	Renal pelvis: 46 (17.0) Ureter: 19 (7.0) Urethra: 8 (3.0)	Lymph node only: 43 (15.9)		Carboplatin and gemcitabine: 54 (20.0) MVAC: 16 (5.9) Vinflunine 20 (7.4) Paclitaxel 18 (6.7) *Therapies used in ≥5% patients in metastatic setting listed*			N/A, UtD, NR: 63 (23.3)[a] *Percentage based on prior platinum containing regiment associated with recurrence/regression (n=72)*
Sharma *et al.* (2016) [CheckMate 032][2]	Median 66 (31-85)	54 (69.2)	0: 42 (53.8) 1: 36 (46.2)	NR	Visceral: 61 (78.2) Liver: 20 (25.6) Lymph node only: 13 (16.7)	Adjuvant: 33 (42.3) Neo-adjuvant: 14 (17.9)	Cisplatin and gemcitabine: 23 (29.5) Carboplatin and gemcitabine: 15 (19.2) MVAC: 7 (9.0) Carboplatin and paclitaxel: 5 (6.4) Vinflunine: 4 (5.1) *Therapies used in ≥5% patients*	25 (32.1)	71 (91.0)	CR: 2 (2.8) PR: 15 (20.8) SD: (19 (26.4) PD: 24 (33.3) N/A, UtD: 12 (16.7)[a] *Percentage based on prior platinum containing regiment associated with*

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Trial, treatment arm and population size	Age (years)	Males n (%)	ECOG status n (%)	Location of urothelial cancer n (%)	Presence and location of metastasis n (%)	Prior neoadjuvant or adjuvant treatment n (%)	Prior type of chemotherapy n (%)	Prior radiotherapy n (%)	Prior surgery n (%)	Response to prior chemotherapy n (%)
							in metastatic setting listed			*recurrence/regression (n=72)*
Gondo *et al.* (2011)[3] Gemcitabine and Cisplatin n=33	Median 66 (40-82)	26 (78.8)	Inclusion criteria: ECOG PS <1 n: NR	Bladder alone: 19 (57.6); Ureter: 7 (21.2); Renal pelvis: 7 (21.2)	Bone: 5 (15.2); Bone only: 1 (3) Lymph nodes only: 10 (30.3); Lymph nodes and lung: 5 (15.2); Lymph nodes and local recurrence: 4 (12.1); Lymph nodes and liver: 2 (6.1); Lymph nodes and bone: 1 (3.0); Evaluable lymph nodes: 24 (72.7) Lung only: 3 (9.1);	Adjuvant: 14 (42)	MVAC. Number of courses: 1: 2 (6.1); 2: 10 (30.3); 3: 10 (30.3); 4: 14 (12.1); ≥5: 7 (21.2)	NR	32 (97)	NR

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Trial, treatment arm and population size	Age (years)	Males n (%)	ECOG status n (%)	Location of urothelial cancer n (%)	Presence and location of metastasis n (%)	Prior neoadjuvant or adjuvant treatment n (%)	Prior type of chemotherapy n (%)	Prior radiotherapy n (%)	Prior surgery n (%)	Response to prior chemotherapy n (%)
					Evaluable lung: 11 (33.3); Lung and local recurrence: 2 (6.1) Liver: 5 (15.2); Liver and peritoneum: 1 (3.0); Visceral lesions: 23; Other: 10 (30.3)					
Joly *et al.* (2009)[4] Paclitaxel n=45	Mean 64 (47-79)	36 (80[a])	NR	Bladder alone: 38 (84); Non-bladder cancer reported as other: 7 (16[a])	Bone: 14 (33); Visceral: 26 (58); Nodes: 23 (55); Pulmonary: 22 (52); Liver: 16 (38); Other: 11	Adjuvant: 32 (71)	Gemcitabine and Cisplatin: 40(89) MVAC: 5(11) Paclitaxel with cisplatin: 1; Paclitaxel with cisplatin and gemcitabine: 1 first-line adjuvant: 32 (71) first-line for	16 (36)	Total: 39 (87); Radical surgery: 28 (NR); Transurethral resection of the bladder: 7 (NR)	NR (62)

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Trial, treatment arm and population size	Age (years)	Males n (%)	ECOG status n (%)	Location of urothelial cancer n (%)	Presence and location of metastasis n (%)	Prior neoadjuvant or adjuvant treatment n (%)	Prior type of chemotherapy n (%)	Prior radiotherapy n (%)	Prior surgery n (%)	Response to prior chemotherapy n (%)
							metastasis: 13 (29)			
Ozawa *et al.* (2007)[5] Gemcitabine n=55	Median 71 (32-84)	44 (80)	NR	Bladder alone: 28 (50.9); Ureter: 16 (29.1); Renal pelvis: 11 (20)	Lymph nodes: 23; Lymph node and lung: 6; Lymph node and liver: 3; Lymph node and bone: 4; Lymph node, lung and liver: 1; Lymph node, lung, liver and bone: 1; Lung: 5; Lung and liver: 1; Lung and bone: 1; Lung, liver and bone: 2	NR	20/47 patients with metastatic disease received prior chemo MVAC: 14 (25[a]); MEC: 5 (9[a]); Low dose cisplatin: 1 (2[a])	NR	NR	NR

[a]Reviewer-calculated value.

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Abbreviations: CR: complete response; ECOG: Eastern Cooperative Oncology Group; MEC: methotrexate, epirubicin and cisplatin; MVAC: methotrexate, vinblastine, adriamycin (doxorubicin) and cisplatin; N/A: not applicable; NR: not reported; PD: progressive disease; PR: partial response; SD: stable disease; UtD: unable to determine.

Table 4: Statistical analysis: single-arm trials

Trial ID	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
Sharma *et al.* (2017) [CheckMate 275][1]	To evaluate whether treatment with nivolumab monotherapy would lead to clinical benefit in patients with metastatic or surgically unresectable UC who have progressed post platinum treatment as demonstrated by a clinically meaningful ORR	• ORRs (both BIRC- and investigator-assessed) were summarised by a binomial response rate and their corresponding two-sided 95% exact CIs using the Clopper-Pearson method.[9] BOR was summarised by response category • Median values of DOR were calculated along with two-sided 95% CI using Brookmeyer and Crowley method.[10] TTR was summarised using descriptive summary statistics for the responders • Time-to-event distributions were estimated using Kaplan-Meier techniques	For all treated patients, a sample size of 242 would provide 90% power to reject the null hypothesis that ORR was 10% at a two-sided 5% type I error if the true ORR in this population was 16.9%	The final analysis of the primary endpoint ORR (based on BIRC assessments) was to be performed six months after approximately 70 patients with PD-L1 expression of ≥5% had been treated (i.e. six months after last patient first treatment)

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Trial ID	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
Sharma *et al.* (2016) [CheckMate 032][2]	To evaluate whether treatment with nivolumab monotherapy will have clinical activity in subjects with advanced or metastatic tumours	• ORR was summarised by a binomial response rate and corresponding two-sided 95% exact CI using the Clopper-Pearson method. • Time-to-event distributions (DOR, PFS and OS) were estimated using Kaplan-Meier techniques	• An ORR of 10% or less was considered not of clinical value, and an ORR of 25% or greater was considered of strong clinical interest • A sample size of 60–100 treated subjects would provide 90% to 97% power to reject the null hypothesis of 10% response rate if the true response rate was 25% with a two-sided Type I error rate of 5%	All 78 patients who received at least one dose of nivolumab were included in the safety and efficacy analyses
Gondo *et al.* (2011)[3]	To study the efficacy and safety of combination chemotherapy with gemcitabine plus cisplatin for patients with advanced urothelial carcinoma after failure of methotrexate, vinblastine, adriamycin, and cisplatin chemotherapy	Time-to-event endpoints were calculated using the KM method, and compared with a log-rank test. The effect of pre-specified baseline prognostic factors were examined using Cox's proportional hazards models	NR	27/30 (90%) patients were available for evaluation of response
Joly *et al.* (2009)[4]	To evaluate the response rate, clinical benefit, and effect on QoL of a second-line chemotherapy with weekly paclitaxel	ITT analysis. TTP and OS estimated using the KM method. 95% CI of survival rate was estimated using the Rothman and Boice method (1982)	NR	Efficacy and adverse event outcomes were reported for all 24 patients in the study

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Trial ID	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
Ozawa *et al.* (2007)[5]	To determine the ORR and toxicity of gemcitabine and cisplatin	Survival distributions were estimated using the KM method. Two-sided P values of less than 0.05 were regarded as statistically significant	NR	All patients who completed at least two therapy cycles were analysed for chemotherapeutic efficacy. All 55 patients received at least two courses of gemcitabine and cisplatin without any discontinuation due to toxicities; therefore, these patients were evaluated for response and toxicity

Abbreviations: BIRC: blinded independent review committee; BOR: best overall response; CI: confidence interval; DOR: duration of response; ITT: intention to treat; KM: Kaplan Meier; NR: not reported; OS: overall survival; PD-L1: programmed death ligand 1; QoL: quality of life; TTP: time to tumour progression

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

A9. A. Please confirm that results for ORR and PFS from the latest database lock for CheckMate 275 (company submission-B, page 47) are based on BIRC assessment.

Yes, both the ORR and PFS results reported from the latest database lock (2nd September 2016) of CheckMate 275 on page 47 of the manufacturer submission are based on blinded independent review committee (BIRC) assessment.

B. Please report investigator-assessed results for ORR and PFS from the latest database lock or CheckMate 275 as well (or BIRC results if it was not BIRC in the company submission).

The investigator-assessed results for ORR and for PFS from the latest database lock of CheckMate 275 are provided below in Table 5and Table 6, respectively.

Table 5: Investigator-assessed ORR results from the latest database lock of CheckMate 275

Tumour response	All-treated population (n=270)	PD-L1 <1% (n=146)	PD-L1 ≥1% (n=124)
ORR, n (%) **95% CI**	[redacted]	[redacted]	[redacted]
Best overall response			
CR	[redacted]	[redacted]	[redacted]
PR	[redacted]	[redacted]	[redacted]
SD	[redacted]	[redacted]	[redacted]
PD	[redacted]	[redacted]	[redacted]
Unable to determine[a]	[redacted]	[redacted]	[redacted]

[a]BOR was reported as unable to determine due to death prior to assessment, early discontinuation due to toxicity of other.
Abbreviations: BOR: best overall response; CI: confidence intervals; CR: complete response; ORR: objective response rate; PD: progressive disease; PR: partial response; SD: stable disease.
Source: CheckMate 275 CSR Addendum (25 October 2016).[11]

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Table 6: Investigator-assessed PFS results from the latest database lock of CheckMate 275

PFS	All-treated population (n=270)	PD-L1 <1% (n=146)	PD-L1 ≥1% (n=124)
No. events/No. subjects (%)	[redacted]	[redacted]	[redacted]
Median PFS (95% CI), months	[redacted]	[redacted]	[redacted]

Abbreviations: CI: confidence intervals; PFS: progression-free survival

A10. Please provide results for ORR, TTD, DOR and PFS for CheckMate 032 based on BIRC assessment.

Measurement of ORR, TTD, DOR and PFS by BIRC assessment was not part of the protocol of the trial for the urothelial carcinoma patient cohort; these outcomes were therefore not measured by BIRC assessment.

A11. **Priority question:** Can the company explain the differences in effectiveness of nivolumab in the CheckMate 275 and 032 studies? Nivolumab seems to be more effective in CheckMate 032. Although the difference is not statistically significant, it is consistent across all outcome measures.

Firstly, it is worthwhile noting that CheckMate 275 enrolled more patients than CheckMate 032. The smaller sample size in CheckMate 032 means there is more uncertainty around the point estimates for the outcome measures in CheckMate 032, as demonstrated by the wider 95% confidence intervals around the reported results of CheckMate 032 compared to CheckMate 275. As such, it may only be by chance that nivolumab appears more effective in CheckMate 032. As noted by the ERG, the difference was not found to be statistically significant, providing evidence of no difference.

The European Public Assessment Report for nivolumab in urothelial carcinoma notes that overall the CheckMate 275 population seemed to have a poor prognosis, and that the population of CheckMate 032 represents a similar population with regards to baseline characteristics.[12] Reviewing the baseline characteristics of the CheckMate 275 and CheckMate 032 studies in detail, there are some small differences in patient populations that might explain any differences in the observed effectiveness of nivolumab in the two studies, should such differences not represent a spurious finding. The CheckMate 275 study population appears to be marginally less healthy at the outset of the trial. Compared with

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

CheckMate 032, CheckMate 275 enrolled a higher proportion of current or former smokers (71.9% vs 61.5%) and had a higher proportion of patients with ≥4 lesions (14.8% vs 6.4%) or with ≥2 Bellmunt risk factors (22.6% vs 15.4%).

It should be noted that the ITC presented in our submission controlled for baseline characteristics that reflect health status at treatment initiation where these were considered to be prognostic (ECOG performance status, visceral and liver metastases or haemoglobin level). This should alleviate concerns that any differences between trial populations in terms of prognostic factors are accounted for in the relative effectiveness estimates.

A12. On page 58 (company submission, section B.2.8) it is mentioned that data from the CheckMate studies were pooled. Please provide details of the statistical method(s) used for pooling the data from Checkmate 275 and 032 and please explain which data were used (BIRC or investigator-assessed). Please conduct all analyses using data from each method separately.

The sentence in Section B.2.8 regarding pooling of data refers specifically to the simulated treatment comparison (STC). In the STC, we predict how patients in each of the comparator trials would have responded to nivolumab. These predictions are based on data from both CheckMate 032 and CheckMate 275.

For each outcome, we first evaluated whether it was appropriate to combine data from the two studies. For OS and PFS, the Wald test was used to evaluate if there was a difference between the two studies. For objective response, a chi-squared test was used to compare a logistic regression model of objective response with study as the only predictor variable to the equivalent model without any predictor variables. In all cases, there was no evidence of a difference between studies (OS: p=0.42, PFS: p=0.28 and ORR: p=0.41) (see Appendix D.1.6). Hence, in each case, the prediction models were based on a dataset of 348 patients (including the 78 patients from CheckMate 032 and the 270 patients from CheckMate 275). For PFS and objective response the STCs were based on the primary definitions of the outcomes in each study. Thus, for CheckMate 032, the STC is based on investigator assessments of PFS and objective response, and for CheckMate 275, the STC is based on blinded independent review committee (BIRC) assessments of these outcomes. High concordance between BIRC-assessed and investigator-assessed response rates in CheckMate 275, as shown in Table 7, supports the pooling of both studies despite differences in primary endpoint definition. As agreed with the ERG on the preliminary teleconference to discuss the clarification questions, analyses using each method separately have not been provided.

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Table 7: Concordance in objective response rate between BIRC- and investigator-assessed objective response rates

Number of subjects, n (%)	BIRC assessment		
Investigator assessment	**Responders**	**Non-responders**	**Unable to determine**
Responders	48 (17.8)	14 (5.2)	0
Non-responders	6 (2.2)	147 (54.4)	14 (5.2)
Unable to determine	0	4 (1.5)	37 (13.7)
Concordance rate (responders)	92.6%		

Abbreviations: BIRC: blinded independent review committee.
Source: CheckMate 275 CSR Addendum.[11]

A13. A. Could the company discuss the generalisability of the CheckMate 275 and CheckMate 032 studies to the UK population, given that more than 50% in both studies had an ECOG performance status of 0?

Feedback from the advisory board acknowledged that there were fewer patients with an ECOG performance status of 0 in the UK clinical practice than in the CheckMate trials. This is consistent with findings of a chart review study conducted by Bristol-Myers Squibb in 2017, which suggested that a lower proportion (18.8%) of patients in UK practice would be ECOG performance status 0. However, clinical expert attendees at the advisory board stated that the CheckMate 275 and CheckMate 032 trial populations could be considered generally representative of the UK patient population.

Importantly, it should be noted that ECOG performance status was adjusted for as a prognostic factor in the prediction model for the simulated ITC. As such, any differences in ECOG performance status between the patient populations of the nivolumab and comparator trials, are accounted for in the relative effectiveness estimates. Therefore any differences in ECOG performance status should not be a concern for the estimates of relative effectiveness that feed into the economic analysis, and hence for the cost-effectiveness results.

B. How well do the Checkmate trials fit the UK population in terms of prior treatments received (type and setting of prior systemic therapy)?

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Based on a recent chart review study conducted for Bristol-Myers Squibb in 2017 in metastatic UC patients in the UK,[13] prior therapies received by patients in the CheckMate trials are similar to those received by patients in UK clinical practice.

Table 8 presents the prior therapy regimens received by more than 5% of patients in either of the CheckMate trials or in the chart review study across the (neo)adjuvant and metastatic settings separately. Across both CheckMate trials the main therapies received as prior treatment were gemcitabine plus carboplatin/cisplatin, and this can be seen to align closely to the main prior therapies reported in the chart review. Furthermore, these prior treatments are also in line with NICE, ESMO and EAU clinical guidelines,[14-16] as reflected in the treatment pathway presented in Figure 7 (Section B.1.3.3) of the original company submission.

Table 8: Prior treatments received by ≥5% of UC patients in CheckMate studies or UK chart review

Treatment	CheckMate 275, n (%)	CheckMate 032, n (%)	Chart review, n (%)
Neoadjuvant or adjuvant setting			
Gemcitabine + cisplatin	76 (28.1)	17 (21.8)	9 (64.3)[a]
Gemcitabine + carboplatin	24 (8.9)	6 (7.7)	2 (14.3)[a]
MVAC (methotrexate, vinblastine, doxorubicin, cisplatin)	26 (9.6)	3 (3.8)	2 (14.3)[a]
BCG vaccine	1 (0.4)	10 (12.8)	NR
Clinical trial	1 (0.4)	NR	1 (7.1%)
Metastatic setting			
Gemcitabine + cisplatin	87 (32.2)	23 (29.5)	106 (45.3)
Gemcitabine + carboplatin	54 (20.0)	15 (19.2)	80 (34.2)
MVAC	16 (5.9)	7 (9.0)	12 (5.1)
Paclitaxel	18 (6.7)	2 (2.6)	12 (5.1)
Carboplatin + paclitaxel	10 (3.7)	5 (6.4)	NR
Vinflunine	20 (7.4)	5 (5.1)	NR

Prior treatments were reported separately for the adjuvant and neoadjuvant settings in CheckMate 275 and CheckMate 032. These have been combined in this table for the purposes of clarity and alignment with reporting in the chart review study

[a]Data on adjuvant/neoadjuvant treatment were only available for 20 of 234 patients in the study.

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

C. A very small number of patients in the Checkmate trials have locally unresectable non-metastatic disease. Does this reflect the UK population and can the data from these patients be applied to the patients in the scope?

The marketing authorisation for nivolumab in bladder cancer is for the treatment of '*locally advanced unresectable or metastatic urothelial carcinoma in adults after failure of prior platinum-containing chemotherapy'*. The population defined in the final scope of this NICE appraisal is '*adults with metastatic or unresectable urothelial cancer whose disease has progressed after platinum-based chemotherapy*'. Both population descriptions encompass not only metastatic patients, but also patients who have progressed to locally advanced, unresectable disease following prior platinum-based therapy but whose disease has not metastasised. As such, these patients are relevant to the population for which nivolumab is licensed and to the scope of this NICE appraisal.

It is difficult to determine what proportion of the scope population in UK practice might have locally unresectable non-metastatic disease as opposed to metastatic disease. The two groups are classified together for the purposes of treatment decision-making: for example, both the NICE and ESMO guidelines on bladder cancer group locally advanced unresectable disease together with metastatic disease for the purposes of management and treatment recommendations.[14, 16] This highlights that from a clinical perspective the distinction between the two groups of patients is not considered sufficiently meaningful to warrant differing treatment or management recommendations. This is supported by the fact that at the advisory board reported in our original submission, clinicians did not indicate that patient status with regards to locally unresectable non-metastatic disease versus metastatic disease was an important prognostic factor. The location of metastases (e.g. liver metastases, visceral metastases) was noted as a determinant of patient prognosis, but non-metastatic versus metastatic disease in itself was not.

Comparator studies:

A14. Adverse events and Health Related Quality of Life (HRQoL) have been presented for nivolumab, but not for the comparators.

A. Please provide adverse events for all comparators in the same way as reported in Table 23 to 26 of the main submission (Section B.2.10.3, pages 72-78).

We are limited by the data presented by the individual papers for the comparator studies as we do not have access to the clinical study reports for any study except for the CheckMate

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

trials. Please see Table 9 and Table 10 below for the adverse event data reported in the included RCTs and single-arm trials respectively.

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Table 9: Adverse event data reported in the included randomised controlled trials of competitors

Study	Treatment	Safety population n	Neutropenia n (%)	Febrile Neutropenia n (%)	Anaemia n (%)	Thrombocytopenia n (%)	Asthenia n (%)	Nausea n (%)	Vomiting n (%)	Diarrhoea n (%)	Pruritus n (%)	Pneumonia n (%)	Lung infiltration n (%)	ALT increase n (%)	Hepatitis n (%)	Abdominal pain with hospitalisation n (%)	Fever n (%)	Leukopenia n (%)	Constipation n (%)
Bellmunt *et al.* (2009)[17]	Vinflunine and BSC	NR 248 at baseline	123 (50)	15 (6)	47 (19.1)	14 (5.7)	48 (19.3)	6 (2.4)	7 (2.8)	NR	NR	NR	NR	NR	NR	NR	NR	NR	40 (16.1)
	BSC	NR 117 at baseline	1 (0.9)	0 (0)	9 (8.1)	1 (0.9)	21 (17.9)	1 (0.9)	0 (0)	NR	NR	NR	NR	NR	NR	NR	NR	NR	1 (0.9)
Choueiri *et al.* (2012)[18]	Docetaxel and Vandetanib	142	10 (14)	NR	1 (1)	NR	4 (6)	NR	NR	0 (0)	NR	NR	NR	NR	NR	NR	NR	NR	NR

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Jones *et al.* (2017)[19]	Paclitaxel	129	Grade 3>: (6)	NR	NR	Grade 3≥ 0 (0)	Grade 3≥: NR (5)	Grade 3≥: 0 (0)	NR	Grade 3≥: NR (2)	NR	NR	NR	Grade 3>: NR (2)	NR	NR	NR	NR	NR
Petrylak *et al.* (2016)[20]	Docetaxel	140	Grade 3>: 16 (36)	Grade 3≥: 6 (13)	Grade 3≥: 3 (6.7)	Grade 3≥: 0	Grade 3≥: 6 (13)	Grade 3≥: 0 (0)	Grade 3≥: 0 (0)	Grade 3>: 1 (2.2)	NR	Grade 3>: 4 (8.9)	NR	NR	NR	NR	NR	Grade 3>: 6 (13)	NR

Abbreviations: BSC: best supportive care; NR: not reported.

Table 10: Adverse events reported in the included single-arm studies of competitors

Study/Adverse event	Treatment	Safety population n	Neutropenia n (%)	Febrile Neutropenia n (%)	Anaemia n (%)	Thrombocytopenia n (%)	Asthenia n (%)	Nausea n (%)	Vomiting n (%)	Diarrhoea n (%)	Pruritus n (%)	Pneumonia n (%)	Lung infiltration n (%)	ALT increase n (%)	Hepatitis n (%)	Abdominal pain with hospitalisation n (%)	Fever n (%)	Leukopenia n (%)	Constipation n (%)
Gondo *et al.* (2011)[3]	Gemcitabine and cisplatin	33	Grade 3: 19	NR	Grade 3: 12 (36.4);	Grade 3: 5 (15.2)	Grade 3: 0 (0);	Grade 3: 0 (0);	Grade 3: 0 (0);	NR	NR	NR	NR	NR	NR	NR	Grade 3: 0 (0);	Grade 3: 14 (42.4);	NR

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Study/Adverse event	Treatment	Safety population n	Neutropenia n (%)	Febrile Neutropenia n (%)	Anaemia n (%)	Thrombocytopenia n (%)	Asthenia n (%)	Nausea n (%)	Vomiting n (%)	Diarrhoea n (%)	Pruritus n (%)	Pneumonia n (%)	Lung infiltration n (%)	ALT increase n (%)	Hepatitis n (%)	Abdominal pain with hospitalisation n (%)	Fever n (%)	Leukopenia n (%)	Constipation n (%)
			(57.6); Grade 4: 3 (9.1)		Grade 4: 2 (6.1)	; Grade 4: 6 (18.2)	Grade 4: 0 (0)	Grade 4: 0 (0)	Grade 4: 0 (0)								Grade 4 0 (0)	Grade 4: 1 (3).	
Joly *et al*. (2009)[21]	Paclitaxel	44	Grade 3: 1 (2); Grade 4: 2 (4)	NR	Grade 3: 3 (7); Grade 4: 2 (4).	NR	Grade 3: 6 (14); Grade 4: 0 (0)	Grade 3: 1; Grade 4: 0.	Grade 3: 1 (2); Grade 4: 0 (0).	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Ozawa *et al*. (2007)[5]	Gemcitabine and cisplatin	55	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR

Abbreviations: NR: not reported.

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

B. Please provide HRQoL data for all comparators.

Limited information was reported in relation to HRQoL in the comparator trials. Two of the comparator studies (Joly *et al.* (2009) and Jones *et al.* (2017)) reported the use of the Functional Assessment of Cancer Therapy (FACT-G) with the bladder module (FACT-B1).[19, 21] These were the only comparator studies to present data on HRQoL.

In the Joly *et al.* (2009) study, authors report that there was no decrease in the scores of the different quality of life (QoL) domains during the chemotherapy. Paclitaxel did not induce toxicity with negative effect on QoL based on taxane subscales of the FACT-Taxane module. Six of 35 patients (17%) had improved QoL in at least 1 domain (absolute change of score ≥+5). Among the 21 patients with objective response or stabilization, 10% (2 of 21) displayed QoL improvement and 14% (3 of 21) decreased their analgesic consumption.

In the Jones *et al*. (2017) study, FACT-BI trial outcome index is significantly reduced in the pazopanib arm (drug outside of NICE scope) (baseline adjusted standardised area under the curve (AUC) median -2.7; IQR: -10.3 to 0.0) compared to paclitaxel (baseline adjusted standardised AUC median 0.0; IQR: -4.9 to 2.0); 2-sided p=0.0028 (FDR adjusted p=0.0034). Similarly, FACT-BI total score is also significantly reduced with pazopanib (baseline adjusted standardised AUC median -3.8; IQR: -9.8 to 0.0) compared to paclitaxel (baseline adjusted standardised AUC median 0.0; IQR: -5.2 to 0.8); 2-sided p=0.0034 (false discovery rate adjusted p=0.0034).

Indirect comparisons

A15. Cisplatin + gemcitabine should be a comparator according to the scope. The company argues that the generalisability of the cisplatin + gemcitabine study is limited because patients were gemcitabine naive. However, they could still be considered as undergoing retreatment with a platinum-based chemotherapy even if the precise combination was different, as stated in the Comparators section of the scope: "Retreatment with 1st line platinum-based chemotherapy (only for people whose disease has had an adequate response)" Could the company explain why cisplatin + gemcitabine cannot be a comparator for patients who have had exposure to cisplatin?

Clinical expert opinion has stated that patients treated with gemcitabine in the second-line setting who have not received gemcitabine in the first-line setting are not reflected of UK clinical practice. The minutes from an advisory board of six UK clinicians treating bladder cancer are provided in the reference pack and clearly state the following:

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

"The regimen used in the Gondo 2011 study was gemcitabine (1000 mg/m^2)/cisplatin (35 mg/m^2), but it was highlighted that these patients all received MVAC in the first-line setting, and would therefore not be a similar patient population to the population that would currently receive gemcitabine/cisplatin in second-line in current UK clinical practice"

The results of a comparison of nivolumab versus cisplatin + gemcitabine can be found in Appendix O though as highlighted above, the clinical data informing this comparison is not generalisable to UK clinical practice. The use of re-challenge is very limited, given the poor patient prognosis currently.

A16. Please provide further details of the three trials excluded from the indirect comparison/mixed treatment comparison and why the doses/treatment regimens were not considered to be in line with current UK clinical practice (See Appendix D.2.3 page 71).

Please see section D.2.2 of the original company submission, which provides the discussion of the exclusion of Kim *et al.* (2016), McCaffrey *et al.* (1997) and Vaughn *et al.* (2002) from the ITC.

In Kim *et al.* (2016), the dose was 30 mg/m^2 over a 1 hour infusion on days 1 and 8 as part of a 21-day cycle. In McCaffrey *et al.* (1997) 100mg/m^2 was administered intravenous (IV) over 1 hour every 21 days. The dose of docetaxel used in the UK is 75 mg/m^2 and therefore BMS concluded that the Kim *et al.* (2016) and McCaffrey *et al.* (1997) studies would not be eligible for consideration in the NMA as the doses administered in these studies were not comparable with UK clinical practice.

Vaughn *et al.* (2002) administered 80 mg/m^2 of paclitaxel over a 1 hour infusion, but this was infused once weekly for four weeks. Expert clinician feedback confirmed that the dose of paclitaxel used in clinical practice is a weekly dose of 80 mg/m^2 administered once weekly but for the first 3 weeks only, as part of a 28-day treatment course. The difference in dosing regimen between Vaughn *et al.* (2002) and the dose of paclitaxel used in clinical practice was sufficient to exclude Vaughn *et al.* (2002) from further analysis.

A17. A. Please provide further details of the fractional polynomial network meta-analysis method, and how you judged whether the proportional hazards assumption did not hold, particularly when the Checkmate trials were single-arm only and could not be

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

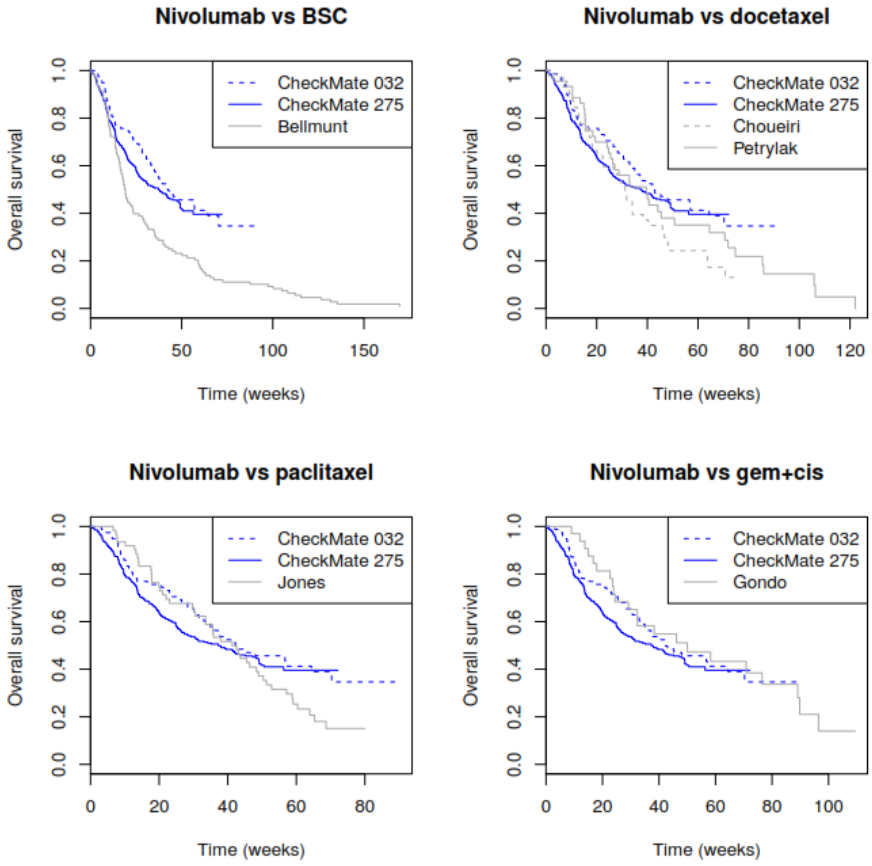
used to assess the proportional hazards assumption for nivolumab (See Section B.2.9.2 page 62)?

Ideally it would be best to assess the proportional hazards assumption for nivolumab versus its comparators using data from within randomized controlled trials. In the absence of such data, we evaluated the proportional hazards assumptions by a) comparing the OS and PFS data across studies and b) by discussing the issue with the clinical advisory board. Figure 1 to Figure 4 show Kaplan-Meier plots and log-cumulative hazard plots for OS and PFS. The nivolumab data comes from the CheckMate 032 and CheckMate 275 studies. For the comparator treatments, the data has been re-constructed from the published Kaplan-Meier plots (see Appendix D.2.5.1). For both OS and PFS, the plots indicate that the proportional hazards assumption does not hold – for the log-cumulative hazard plots, the curves are not parallel.
This observation is supported by the clinical advisory board input – proportional hazards are not expected to hold because of the different mechanisms of action of the treatments.

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Figure 1: Overall survival: Kaplan-Meier plots for nivolumab and its comparators

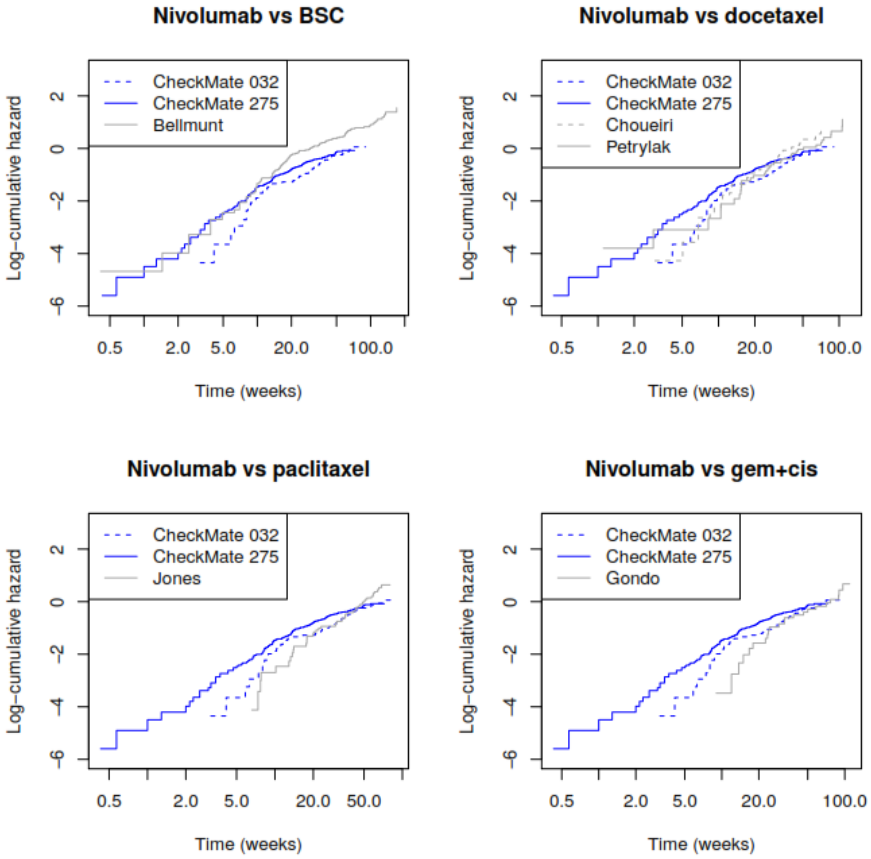


Abbreviations: BSC: best supportive care; gem+cis: gemcitabine and cisplatin.

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

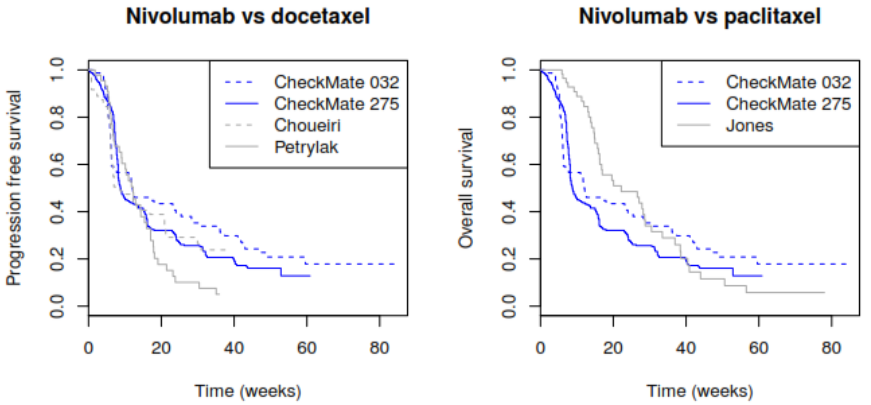
+44 (0)300 323 0140

Figure 2: Overall survival: Log-cumulative hazard plots for nivolumab and its comparators



Abbreviations: BSC: best supportive care; gem+cis: gemcitabine and cisplatin.

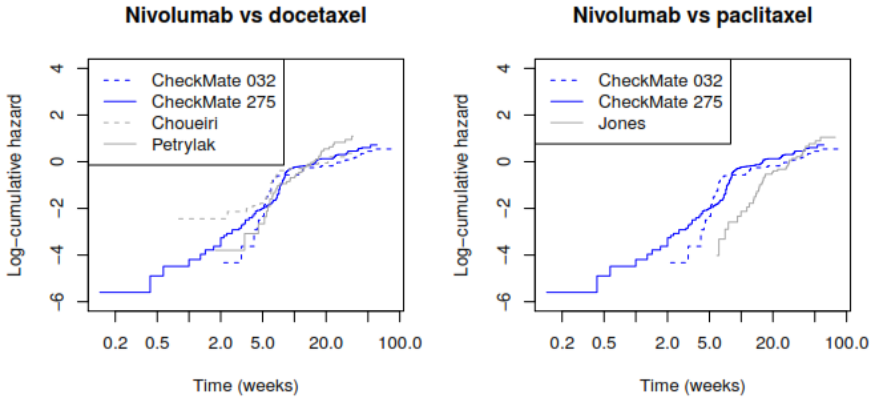
Figure 3: Progression-free survival: Kaplan-Meier plots for nivolumab and its comparators



Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Figure 4: Progression free survival: Log-cumulative hazard plots for nivolumab and its comparators



Appendix D.2.5.5 provided a description of the fractional polynomial NMA method. Further description is provided below.

The fractional polynomial NMA method is based on that proposed in Jansen (2011).[22] The method uses a fractional polynomial model to describe the log hazard rate over time. The time period is divided into equally sized intervals. Both first order and second order fractional polynomial models were fitted.

The first order fractional polynomial NMA model is:

$$\log(h_{jkt}) = \beta_{0jk} + \beta_{1jk}t^p$$

$$\begin{pmatrix}\beta_{0jk}\\ \beta_{1jk}\end{pmatrix} = \begin{cases} \begin{pmatrix}\mu_{0j}\\ \mu_{1j}\end{pmatrix} & \text{if k is the baseline treatment for study j} \\ \begin{pmatrix}\mu_{0j}\\ \mu_{1j}\end{pmatrix} + \begin{pmatrix}\partial_{0jkb}\\ \partial_{1jkb}\end{pmatrix} & \text{if k is not the baseline treatment for study j} \end{cases}$$

Where:

- h_{jkt} is the hazard rate for treatment k in study j in time interval t
- $\log$ is the natural log
- β_{0jk} is the first parameter of the fractional polynomial model for treatment k in study j
- β_{1jk} is the second parameter of the fractional polynomial model for treatment k in study j
- p is the power of the fractional polynomial model. If $p = 0$, then the term t^p is set to $\log t$
- μ_{0j} and μ_{1j} are the fractional polynomial model parameters for the baseline treatment in study j

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

- ∂_{0jkb} and ∂_{1jkb} are the differences in the fractional polynomial model parameters for treatment k, relative to baseline treatment b in study j

The second order fractional polynomial NMA model is:

$$\log(h_{jkt}) = \beta_{0jk} + \beta_{1jk}t^{p_1} + \beta_{2jk}t^{p_2}$$

$$\begin{pmatrix}\beta_{0jk}\\ \beta_{1jk}\\ \beta_{2jk}\end{pmatrix} = \begin{cases} \begin{pmatrix}\mu_{0j}\\ \mu_{1j}\\ \mu_{2j}\end{pmatrix} & \text{if k is the baseline treatment for study j} \\ \begin{pmatrix}\mu_{0j}\\ \mu_{1j}\\ \mu_{2j}\end{pmatrix} + \begin{pmatrix}\partial_{0jkb}\\ \partial_{1jkb}\\ \partial_{2jkb}\end{pmatrix} & \text{if k is not the baseline treatment for study j} \end{cases}$$

Where the parameters are as per the first order model, with the addition of:

- p_1 and p_2: the powers of the fractional polynomial model. If the power is 0, then the term is set to $\log t$. If $p_1 = p_2 = p$, then the model becomes a repeated powers model: $\log(h_{jkt}) = \beta_{0jk} + \beta_{1jk}t^{p} + \beta_{2jk}t^{p}\log t$
- β_{2jk}: the third parameter of the fractional polynomial model for treatment k in study j. μ_{2j} and ∂_{2jkb} are the associated baseline and difference parameters.

For the random effects model, we only allowed for heterogeneity in the ∂_{0jkb} parameter. This means that the between-study variability of the log hazard ratios is constant over time. We felt that this was a reasonable assumption. Models that also allow for heterogeneity in the ∂_{1jkb} and ∂_{2jkb} parameters are more flexible, but less stable.

Thus for the random effects models:

- ∂_{0jkb}~Normal(d_{0kb}, τ^2),
- $\partial_{1jkb} = d_{1kb}$, and
- $\partial_{2jkb} = d_{2kb}$.

For the fixed effect models:

- $\partial_{0jkb} = d_{0kb}$,
- $\partial_{1jkb} = d_{1kb}$, and
- $\partial_{2jkb} = d_{2kb}$.

The parameters d_{0kb}, d_{1kb} and d_{2kb} are the differences in the fractional polynomial model parameters for treatment k, relative to treatment b, and τ^2 is the between-study variance. As for all NMA models the differences between treatments follow the consistency equations such that:

- $d_{0kb} = d_{0k1} - d_{0b1}$,
- $d_{1kb} = d_{1k1} - d_{1b1}$,

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

- $d_{2kb} = d_{2k1} - d_{2b1}$,

Where 1 is the overall reference treatment for the network.
Note that each study only provided data for a single comparator treatment. Hence it was not necessary to include any adjustments for multi-arm trials.
For the comparator treatments, the data is modelled as per Jansen 2011:
$r_{jkt} \sim \text{binomial}(p_{jkt}, n_{jkt})$
Where r_{jkt} is the observed number of events in the interval $[t, t + \Delta t]$ for treatment k in study j, n_{jkt} is the number at risk at the start of the interval and p_{jkt} is the probability of an event in the interval.
The hazard rate is then assumed to be constant within the time interval, such that:
$h_{jkt} = -\log(1 - p_{jkt})/\Delta t$
For each comparator study, the log hazard rates for nivolumab were simulated for a set of 10,000 patients. The mean and variance of the log hazard rates were then included in the model as follows:

$$\text{y}_{jkt} \sim \text{Normal}(\log(h_{jkt}), \sigma_{jkt}^2)$$

Where y_{jkt} is the mean log hazard rate for nivolumab in study j and σ_{jkt}^2 is the variance of the log hazard rate for nivolumab in study j. In our model, nivolumab is the baseline treatment in each study, and overall, so in the above equation $k = 1$. We know that the nivolumab predictions must follow the proportional hazards property (since they are simulated from a proportional hazards model). Hence, in our model, the second and third fractional polynomial parameters for the baseline treatment (i.e. nivolumab) are fixed: $\mu_{1j} = \mu_1$ and : $\mu_{2j} = \mu_2$.
The baseline (μ) and difference parameters (d) were assigned vague normal priors: Normal(0,100^2). For the random effects models, the between study standard deviation (τ) was assigned a Uniform(0,2) prior.

B. Please discuss methods other than the fractional polynomial for conducting the network meta-analysis including their pros and cons.

An alternative to the fractional polynomial NMA model, is a standard NMA model for HR data. Standard NMA models for HR data:

- Synthesise log HRs from each of the studies (the WinBUGS code in Example 7 of the NICE DSU TSD 2 can be used), and
- Lead to HRs that are constant over time

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

An advantage of the standard NMA model for HR data is that it is easier to interpret, however it is only appropriate if the proportional hazards assumption holds. As per our response to part A of this question, we do not believe a standard NMA model for HR data is appropriate, in this case, because the proportional hazards assumption does not hold. However, we have provided the results of standard NMA models for HR data in response to Clarification Question B4.C.

A18. A. **Priority question**: Please quantify the possible extent of any residual systematic error resulting from unobserved prognostic variables and effect modifiers, using the 'out-of-sample' method described in NICE DSU TSD 18.

B. It is argued on page 103 in the company submission that this method may not provide an accurate estimate of the residual bias. Please explain why and in which direction it differs from an accurate estimate.

As identified in the NICE DSU TSD 18, there are no standard methods for estimating the residual bias. The 'out-of-sample' method described in NICE DSU TSD 18 involves comparing the observed between-study variability in the comparator studies, to the between-study variability in the predicted results for nivolumab. The NICE DSU TSD 18 suggests that if the prediction model includes all of the key effect modifiers and prognostic variables then the nivolumab predictions will be as variable as the observed results in the comparator studies.[23]

For the STCs presented in this submission, we concluded that the 'out-of-sample' method as described in NICE DSU TSD 18 would not provide a good estimate of the residual bias. There were two key reasons for this:

1. The method described NICE DSU TSD 18 involves a comparison of the between-study variability in the observed and predicted data. However, in this case, there was very limited data to estimate the between-study variability. As noted on page 103 in the company submission and in the NICE DSU TSD 18, estimation of the between-study variance is difficult when there is limited data. In order to estimate the between-study variance, we need at least one comparator to be informed by at least two studies (in order to estimate a robust estimate of the between-study variance, we would ideally have several comparators that are each informed by several studies). For both the OS and PFS networks, docetaxel is the only comparator treatment that is informed by more than one study, and even this treatment is only informed by two studies. Thus for these networks, we have only the minimum amount of information that is required for estimating the between-

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

study variance. Similarly, the objective response network also has limited data to estimate the between-study variance (gemcitabine + cisplatin and docetaxel are each only informed by two studies). The estimates of the between-study variability and the subsequent estimates of the residual bias will be very uncertain because of the limited information available to estimate the between-study variabilities. Whilst the lack of data will have an unfavourable impact on the precision, we do not think the lack of data will bias the estimate of the residual bias in any particular direction.

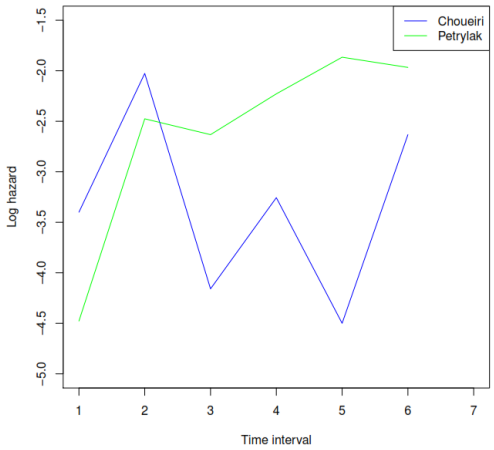
2. In this case, the 'out-of-sample' method is likely to overestimate the amount of residual bias for the survival outcomes. The NMAs for OS and PFS were conducted using a fractional polynomial model. Thus, in order for the residual bias analysis to reflect the NMA, the fractional polynomial model should be used to estimate the between-study variances. The fractional polynomial model synthesises the log hazard rates over time. Following Jansen 2011,[22] the fractional polynomial models used in this submission only allow for between-study variability in the β_0 parameter. This means that the fractional polynomial model assumes a constant between-study variability over time and that the between-study variability is effectively estimated from the variability at each time point. For the predicted nivolumab results, the between-study variability is exactly the same at each time point, since the nivolumab predictions come from a proportional hazards model. However, for the observed comparator results, the observed between-study variability varies between time points since the observed results do not necessarily follow the proportional hazards assumption. For the observed comparator results, the between time-point variability in the log hazard rates, also potentially contributes to the estimate of the between-study variability. This is illustrated in Figure 5 for PFS.

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

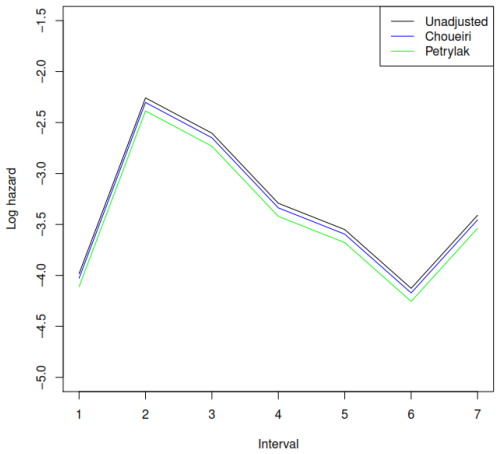
+44 (0)300 323 0140

Figure 5: Observed log hazard rates by time interval for docetaxel and predicted log hazards rates by time interval for nivolumab (PFS)

Observed log hazard rates by time interval (docetaxel)



Predicted log hazards rates by time interval (nivolumab)



The differences between the observed and predicted rates, as described above, imply that the estimates of the between-study variability for the observed results are likely to be higher than the estimates of the between-study variability for the predicted results. Thus the 'out-of-sample' method is likely to overestimate the amount of residual bias for the survival outcomes.

As outlined above, the 'out-of-sample' method described in the NICE DSU TSD 18 will not provide a good estimate of the residual bias. However, as per the request for clarification, we have presented these results below. It is also important to note that the NICE DSU TSD 18 was only released in December last year, and as such, there are no established benchmarks for the ratio produced by the 'out-of-sample' method.

Methods

For each outcome, the 'out-of-sample' method described in the NICE DSU TSD 18 was implemented as follows:

1. A naïve indirect comparison was conducted using a Bayesian random effects model. The analysis was based only on the observed data from the comparator studies. This analysis provides an estimate of the posterior distribution of the between-study variability based on the observed data in the comparator studies (τ^2).
2. A naïve indirect comparison was conducted using a Bayesian random effects model. The analysis was based only on the predicted results for nivolumab in each of the

comparator studies. This analysis provides an estimate of the posterior distribution of the between-study variability based on the predicted results for nivolumab (τ_*^2).

3. As suggested in the NICE DSU TSD 18, the ratio of the between-studies variance (τ_*^2/τ^2) was calculated. We estimated the posterior distribution of the ratio using the samples from the posterior distributions of the between-study variabilities. We summarised the posterior distribution of the ratio using the median and its 95% credible interval.

For OS and PFS, the naïve indirect comparisons were based on second order fractional polynomial models with P1=0, P2=0. This form of the fractional polynomial model was selected, in line with the base case for the cost-effectiveness model.

Results

The results of the residual bias analysis, using the 'out-of-sample' method are provided in Table 11. The ratio of the between-studies variance ranges from 0.02 for overall survival, to 0.43 objective response rate. For all of the outcomes, the 95% credible intervals indicate that the estimate of the ratio is highly uncertain. For the reasons described above, the 'out-of-sample' method is likely to overestimate the amount of residual bias for the survival outcomes (i.e. it is likely to underestimate the ratio).

Table 11: Results of the residual bias analysis ('out-of-sample' method). Estimates and 95% credible intervals

	Between-studies variance		Ratio of the between-studies variance (τ_*^2/τ^2)
	Predicted results for nivolumab (τ_*^2)	Observed results for comparators (τ^2)	
Overall survival (fractional polynomial model, second order, P1=0, P2=0)	0.00516 (0.00001, 0.14968)	0.25381 (0.00055, 3.49406)	0.02058 (0.00002, 19.900)
Progression free survival (fractional polynomial model, second order, P1=0, P2=0)	0.01413 (0.00002, 1.54993)	0.21480 (0.00031, 3.48311)	0.07732 (0.00006, 173.73)
Objective response rate	0.0605 (0.0001, 1.2588)	0.1328 (0.0002, 2.9958)	0.4348 (0.0004, 436.5)

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

A19. Pooled nivolumab data were simulated to match characteristics in the comparator studies. Therefore, it is important that inclusion criteria and population characteristics of comparator studies match the population described in the scope. Please discuss each of the comparator studies and describe whether they reflect the UK population described in the scope.

A detailed overview of the studies included in the indirect treatment comparison is provided in Section D.2.4 "Methods and outcomes of trials included in indirect or mixed treatment comparison". This provides an overview of the inclusion criteria and the population characteristics. A limitation already highlighted is relating to the difference in first-line treatments, whereby the prior chemotherapy treatments varied widely across the trials and it was not always clear what combinations of treatments the patients had received (see Table 20 and Table 21 in Appendix D). One trial specified that the first-line treatment was MVAC, with gemcitabine plus cisplatin given in the second line setting, whereas gemcitabine plus platinum is the standard of care in the first-line setting in the UK.[3] Joly et al. (2009) did not name the type of first-line chemotherapy and Ozawa et al. (2007) did not mention first-line treatment in their inclusion criteria.[5, 21]

A20. Studies such as pazopanib vs. docetaxel, or docetaxel vs. BSC, or docetaxel+ ramucurimab vs. vinflunine could have been used to provide indirect comparisons in the meta-analysis conducted by the company i.e. the so called ITC. Were such studies searched for in the systematic literature review? Is it possible that such studies exist, but not found through the systematic literature review?

As detailed in Appendix D.1.2 of our original submission, docetaxel, vinflunine and BSC were searched for in the SLR and hence studies on these interventions, regardless of their comparator therapies in the study, would have been identified by the SLR. The SLR search did not specifically look for pazopanib or ramucurimab as these were not interventions of interest, and hence studies with treatment arms involving these therapies would only have been captured if another treatment arm in the study was considered an eligible intervention.

As part of the feasibility assessment for the ITC, only treatment arms with eligible interventions were considered for the ITC. On this basis, the following treatment arms for studies identified by the systematic literature review, and their patients and outcome data, were not be considered in the feasibility assessment for the ITC:

Petrylak *et al.* (2016):

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

- docetaxel and icrucumab;
- docetaxel and ramucirumab.

Choueiri *et al.* (2012):

- docetaxel and vandetanib.

Jones *et al.* (2017):

- pazopanib

Sharma *et al.* (2016):

- nivolumab and ipilimumab.

A21. **Priority question:** It is clear across all outcomes (including ORR, PFS and OS) that patients with PD-L1 < 1% expression do less well with nivolumab than those with PD-L1 >=1% expression.

A. Could the company please provide a justification as to why the 'Indirect treatment comparison' for this subgroup only was not performed?

B. Could the company please perform the 'Indirect treatment comparison' for this subgroup only?

The indirect treatment comparison for this subgroup was not performed because neither baseline PD-L1 data nor outcomes split by PD-L1 sub-group were available for the comparator trials. There is some evidence that PD-L1 expression could be a prognostic factor[24, 25] but it has not been reported in the comparator trials and therefore cannot be controlled for.

A22. **Priority question**: The code for the 'indirect or mixed treatment comparison' is shown in Appendix D.2.7. Could the company also provide all of the data necessary for running these models so that the ERG can validate the results?

All analysis was conducted using R and WinBUGS. The R2WinBUGS package was used to call WinBUGS from within R. The R code, WinBUGS code and data for running the models is provided in an accompanying file. Code and data were provided for:

- Fractional polynomial models for OS and PFS,
- Constant HR models for OS and PFS, and
- Binomial models for objective response rate.

For each model, the package includes the following:

- The main R code for running the analyses. To run the analyses, you should open this code. You will need to specify the directory where you have saved the files. This code will then load the appropriate data and call an R function that sets up the initial values and calls the appropriate WinBUGS code. These files are labelled as follows:
    - **Run fractional polynomial survival models.R**: fractional polynomial models for OS and PFS
    - **Run constant HR survival models.R**: Constant HR models for OS and PFS
    - **Run binomial models.R**: Binomial models for objective response rate.
- The R functions for running the analyses.
- The WinBUGS code.
- The .Rdata files containing the required data.

A23. Please provide evidence based on the effectiveness analyses that nivolumab provides an extension of life of at least three months compared to the comparators in order to fulfill the end-of-life criteria.

The economic model estimates a difference in mean life years per patient with nivolumab of 2.78 years (33.36 months) vs 1.19 years (14.28 months) with paclitaxel, 1.40 years (16.80 months) with docetaxel and 1.01 years (12.12 months) with BSC.

Section B: Clarification on cost-effectiveness data

Treatment effectiveness

B1. To derive nivolumab treatment effectiveness, the CheckMate 275 and 032 studies were pooled. This is inconsistent with how utilities, resource use and adverse event rates were derived (from the CheckMate 275 study only). Please justify why the treatment effectiveness data were derived from the pooled CheckMate 275 and 032 studies, but utilities, resource use and adverse event rates were derived from CheckMate 275 only.

Utility

The data from the CheckMate 032 study was not considered at the time of the original analysis. However, this pooled analysis has now been undertaken and shows a small increase in utility values for both pre-progressed and post-progressed patients. In terms of

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

the cost-effectiveness model this increase in utility values has caused a small decrease in the ICERs for nivolumab versus each comparator. This is discussed further in Section B.15.

Adverse events

Adverse event rates were taken from CheckMate 275 only in order to simplify the analysis. The adverse event rates have no meaningful impact on the results of the cost-effectiveness analysis so use of pooled data is not expected to alter the conclusions from the overall analysis.

Resource use

Data from CheckMate 275 and CheckMate 032 was used to estimate treatment duration, one of the key components of resource use with nivolumab. Data from CheckMate 275 only on subsequent radiotherapy and surgery was included in the economic model, though the overview from both trials indicates the rates were relatively similar across the trials (see Table 12).

Table 12: Resource use in CheckMate 275 and CheckMate 032

Resource component	CheckMate 275	CheckMate 032
Subsequent radiotherapy	9.3%	11.5%
Subsequent surgery	3.3%	6.4%

B2. The company states that a response-based modelling approach was adopted in order to reflect the mechanism of action of nivolumab and to reflect that the nivolumab survival curve changes over time as the hazard changes. The company furthermore claims that standard parametric models are unlikely to be flexible enough to characterise this change in the hazard. However, most parametric distributions (except the exponential) can be used to incorporate changing hazards over time. Additionally, standard models (e.g. log-logistic, log-normal and generalised gamma distributions) even include a hazard function that is non-monotonic with respect to time (initially an increasing hazard, followed by a decreasing hazard).[26] Moreover, the NICE technical support document on survival analysis suggests spline-based models as useful, more flexible alternatives.[26] Please provide further justification for the response-based approach, and why landmark analysis was performed, in particular:

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

A. Please provide justification for why a response-based approach was necessary, including whether standard parametric curves (as described in the NICE technical support document on survival analysis) were tested and why they were deemed to not appropriately reflect nivolumab survival.[26]

BMS aims to address previous concerns from appraisals regarding nivolumab, where appropriate. Based on extensive Committee criticism as part of ID971, they were not deemed to be suitable for modelling survival. Standard parametric curves were tested and are provided in Figures 109 to 120 to in Appendix L.

B. Were other methods, such as spline-based models (see also TSD 14), or mixture cure models, considered?[26, 27] If so, why was a landmark analysis preferred? If not, please consider the advantages and disadvantages of these methods compared to the landmark analysis and consider implementation of the most suitable approach.

Other methods, such as those mentioned in the question, were not explicitly investigated. Spline-based models have been used by BMS in other appraisals though they have generally not been accepted by NICE in other nivolumab appraisals (ID811 and ID900). This is in contrast to the Scottish Medicines Consortium who have accepted more flexible survival modelling approaches for nivolumab. Mixture cure models are an area of research for BMS but their applicability to HTA bodies is yet unknown.
As noted, the proposed approach allows for a more flexible shape to the nivolumab survival curve whilst adhering to the Committee's previous preference of using the trial data for a proportion of the survival curves.

C. Please provide justification for the choice of the selected 8-week landmark using clinical expert opinion.

The choice of the 8-week landmark was based on the clinical evidence, whereby the majority of patients had responded by 8 weeks.

D. Please analyse the impact of using different landmarks, by providing scenario analyses results (disaggregated) of alternative landmarks: 12 and 20 weeks.

Due to time constraints, the provision of other analyses has been prioritised by BMS. This request cannot be fulfilled at this time.

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

E. Please provide justification for why no parametric curve was fitted to the Kaplan–Meier estimates prior to the 8-week landmark point. Please also provide the results of an analysis where a parametric curve is fitted to the data before the landmark point.

It was not necessary to fit a parametric curve prior to the landmark point because this part of the data was not used for extrapolation. Fitting a parametric curve to this data would add the unnecessary complexity of selecting which parametric curve best fitted this part of the data. This approach was adopted to adhere to the Committee's previous preference of using the trial data for a proportion of the survival curves.

F. Please provide a scenario analysis where nivolumab patients are not analysed separately by response (i.e. OS and PFS curves fitted to all patients regardless of response status).

An analysis where nivolumab patients are not analysed separately by response can be found by unticking the box labelled "Use response-based approach?" on the tab "PFS & OS" in the economic model.

B3. **Priority question**: For the analysis of responders versus non-responders, proportionality of hazards was discarded, even though no analysis was presented to justify this. Furthermore, the responders' and non-responders' curves were combined using an average weighted by the 8-week responder proportion, thus artificially over-estimating the weight of non-responders in later periods.

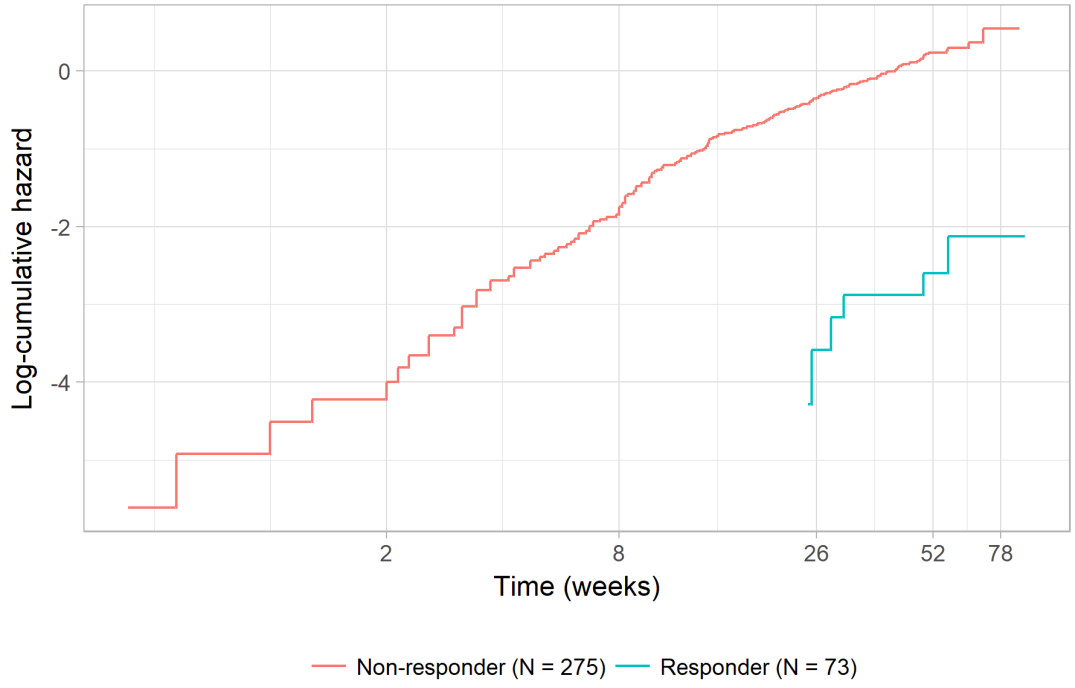
A. Please explore whether proportionality of hazards is violated between responders and non-responders, using log cumulative hazard plots.

The log-cumulative hazard plots split by responder status for OS (Figure 6) and PFS (Figure 7) show that the proportional hazards assumption could be valid for OS, but is unlikely to be valid for PFS. However, proportional hazards were not assumed as this meant there was no requirement to assume the same distribution to be appropriate for both responder and non-responder curves.

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

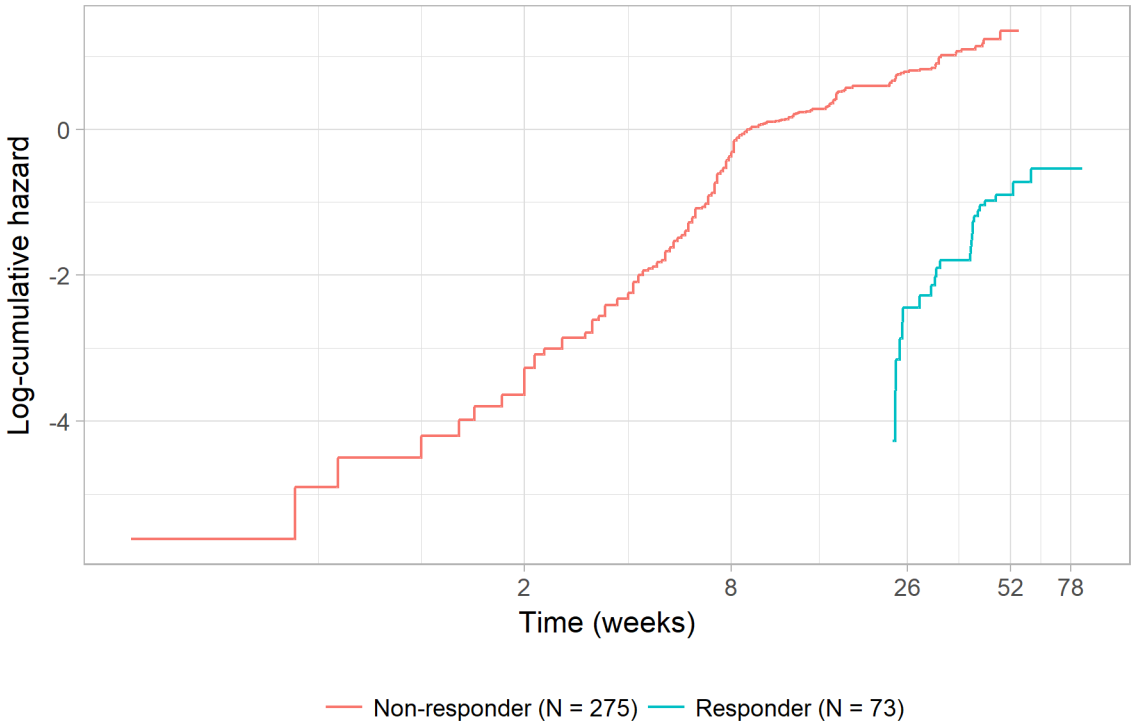
Figure 6: Log-cumulative hazard plot for overall survival for Checkmate 275 and Checkmate 032 – by responder status



Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Figure 7: Log-cumulative hazard plot for progression free survival for Checkmate 275 and Checkmate 032 – by responder status



B. Please provide justification for, and describe the methods used for, combining the responders and non-responders' curves instead of modelling them separately by using additional health states in the model, and provide comment on the impact of this approach.

To generate a combined curve, at each cycle after the landmark point (i.e. 8 weeks in the base case analysis), the PFS and OS estimates for responders and non-responders were multiplied by the proportion of people deemed to be responders or non-responders at the landmark point. For example, using a generalised gamma distribution, the estimates for PFS are 59% for responders and 36% for non-responders at week 12 (cycle 3). Further, at the landmark point the proportion of people classified as responders is 35% and non-responders is 65%. Therefore, a weighted average was generated by multiplying 59% by 35% (responders) and 36% by 65% (non-responders) and summating. This equates to a PFS value of 44%, which is the estimate applied in the combined PFS curve. This calculation was undertaken for both PFS and OS at each post-landmark time point.

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

It was necessary to generate a combined curve, as opposed to have separate responder and non-responder health states in the model, because the hazard ratios for each comparator were generated by comparing each comparator to all nivolumab patients. This was undertaken as it was determined that there was insufficient data to allow separate responder and non-responder nivolumab patient groups to be compared with the comparators using the prediction models discussed previously. As the hazard ratios for the comparators were based on all nivolumab patients the combined curve was generated to facilitate accurate predictions of PFS and OS for all comparators (i.e. it would not be valid to apply these hazard ratios to separate responder and non-responder curves given the approach adopted for the prediction models).

It is expected that the current approach (i.e. generating a combined curve and apply time-varying hazard ratios to estimate PFS and OS for the comparators) will allow the innovative mechanism of action for nivolumab to be appropriately captured whilst allowing a suitable estimation of PFS and OS for the comparators, given the available data.

B4. **Priority question**: The time-varying hazard ratios are calculated by predicting survival of patients from the comparator studies if they would receive nivolumab based on a prediction model estimated on the pooled data from the CheckMate studies (i.e. not divided into the groups of responders vs non-responders). The hazard ratios obtained are then applied to the newly calculated survival curves that combined responders and non-responders. The model parameterisation of the fractional polynomial approach (i.e. which polynomials are chosen) has a large impact on model outcomes.

A. Please discuss the potential bias induced by deriving hazard ratios from one survival curve (fitted to all patients irrespective of response status) and then applying it to a different survival curve (the one that was derived from combining the responders and non-responders curves using a weighted average). Please provide justification for this approach.

The use of the responder-based approach to estimating survival curves was adopted to better characterise the shape of the curves in the unobserved portion, i.e. the tail. By not adopting this approach the long term impact of nivolumab on PFS and OS would likely be underestimated, hence the inclusion of the responder-based approach in the economic model. The reason for the application of hazard ratios to a combined nivolumab curve is discussed further in the answer to the previous question.

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

As with any models, the introduction of assumptions can result in increased uncertainty. In this case, it is possible that, as the model progresses, the 'mix' of responders and non-responders will change (most likely in favour of a greater proportion of responders, due to their increased length of survival). As hazard ratios were derived from a mixed cohort (of responders and non-responders), it was necessary to ensure that the ratio was similarly applied to a mixed group, hence the combined curve. If we suspect that responders will observe a more beneficial hazard ratio than non-responders, then it is possible that the long-term ratio will be underestimating the relative effectiveness of treatment.

B. Please provide hazard ratios derived for responders and non-responders separately.

See answer to question A above.

C. Please provide a scenario analysis incorporating hazard ratios that are estimated independent of time (i.e. fixed over time).

Methods

The prediction models for OS and PFS (as described in Appendix D.2.6.1) were used to predict the HRs with respect to nivolumab in each of the comparator trials. The predicted HRs were then synthesized using a standard NMA model. Further details of this approach are provided below.

For each outcome, and comparator trial:

- A Cox proportional hazards model was used to calculate a naïve HR for nivolumab versus the comparator based on the pooled CheckMate 032 and CheckMate 275 data, and the reconstructed data from the comparator trial (see Appendix D.2.5.1 of our original submission)
- The baseline characteristics of the patients in the comparator trial were simulated using the approach described in Appendix D.2.5.4 of our original submission. For each of the simulated comparator trial patients:
    - The prediction model for the outcome was used to predict their HR relative to an average patient in the CheckMate 032 and CheckMate 275 trials (i.e. a patient with average values of the baseline characteristics).
    - Their HR relative to the comparator treatment was calculated by multiplying their HR relative to an average patient in the CheckMate trials and the naïve HR calculated in step 1.

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

The log HR for nivolumab versus the comparator was then estimated by the mean log HR of all of the simulated patients. The input data for the NMA was the mean log HR and its precision.
For each outcome, standard fixed effect and random effects NMA models for treatment differences (as per NICE DSU TSD 2, Example 7) were then applied to the mean log HRs. The WinBUGS code for these models is included below.

Results – Overall survival

Both fixed and random effects models were evaluated. Table 13 provides a summary of the model fit statistics. The table indicates that the fixed effect model had the lowest DIC. This suggests that there is minimal between-study heterogeneity. It was not possible to evaluate inconsistency because the network does not include any comparisons informed by both direct and indirect evidence. The results of the fixed effect and random effects models are shown below.

Table 13: Model fit statistics for overall survival (standard NMA model with constant HRs)

	$\bar{D}_{res}$	pD	DIC
Fixed effect model	4.0	4.0	8.0
Random effects model	4.5	4.5	9.0

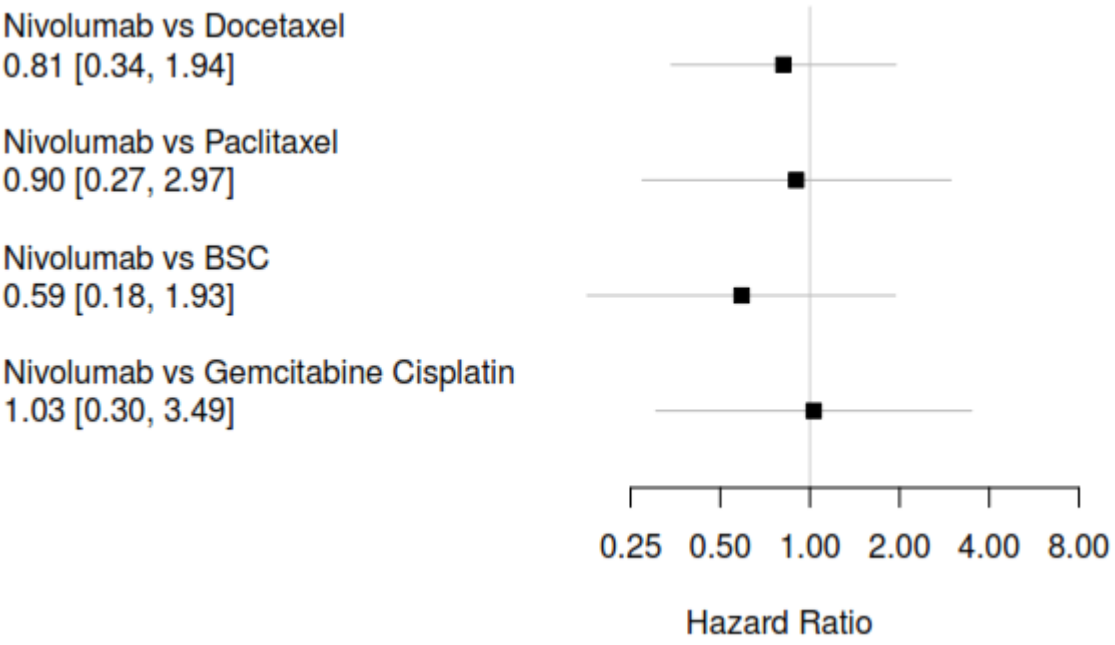
Abbreviations: $\bar{D}_{res}$: residual deviance; DIC: deviance information criterion; pD: number of effective parameters.

For the fixed effect model, Figure 8 illustrates the HRs for nivolumab versus each of the comparators and Table 14 provides the estimates of the HRs and their 95% credible intervals. There is no evidence of a difference between nivolumab and any of the comparators.

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Figure 8: Overall survival: network meta-analysis results (standard fixed effect NMA model with constant HRs): HRs for nivolumab versus each of the comparators



The trials informing the comparison versus cisplatin plus gemcitabine were conducted in gemcitabine-naïve populations, and therefore cannot be considered to provide relevant data for the retreatment with first-line platinum-based chemotherapy. This comparison is included as a scenario analysis only and results should be interpreted with caution.
Abbreviations: BSC: best supportive care.

Table 14: Overall survival: network meta-analysis results (standard fixed effect NMA model with constant HRs): HRs and 95% credible intervals for each pairwise comparison

	Nivolumab	Docetaxel	Paclitaxel	BSC
Docetaxel	0.81 (0.34, 1.94)			
Paclitaxel	0.90 (0.27, 2.97)	1.10 (0.25, 4.84)		
BSC	0.59 (0.18, 1.93)	0.72 (0.17, 3.12)	0.65 (0.12, 3.52)	
Cisplatin plus gemcitabine	1.03 (0.30, 3.49)	1.26 (0.28, 5.64)	1.15 (0.21, 6.28)	1.75 (0.32, 9.62)

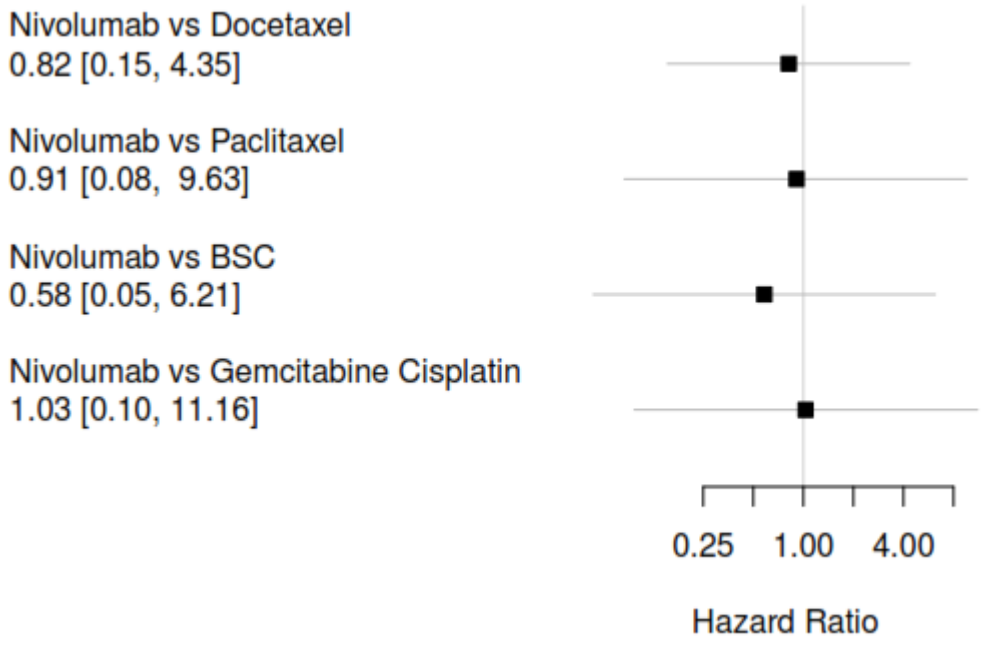
Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

HRs less than 1 favour the column treatment. The trials informing the comparison versus cisplatin plus gemcitabine were conducted in gemcitabine-naïve populations, and therefore cannot be considered to provide relevant data for the retreatment with first-line platinum-based chemotherapy. This comparison is included as a scenario analysis only and results should be interpreted with caution.
Abbreviations: BSC: best supportive care; HR: hazard ratio.

For the random effects model, Figure 9 illustrates the HRs for nivolumab versus each of the comparators and Table 15 provides the estimates of the HRs and their 95% credible intervals. There is no evidence of a difference between nivolumab and any of the comparators.

Figure 9: Overall survival: network meta-analysis results (standard random effects NMA model with constant HRs): HRs for nivolumab versus each of the comparators



The trials informing the comparison versus cisplatin plus gemcitabine were conducted in gemcitabine-naïve populations, and therefore cannot be considered to provide relevant data for the retreatment with first-line platinum-based chemotherapy. This comparison is included as a scenario analysis only and results should be interpreted with caution.
Abbreviations: BSC: best supportive care.

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Table 15: Overall survival: network meta-analysis results (standard random effects NMA model with constant HRs): HRs and 95% credible intervals for each pairwise comparison

	Nivolumab	Docetaxel	Paclitaxel	BSC
Docetaxel	0.82 (0.15, 4.35)			
Paclitaxel	0.91 (0.08, 9.63)	1.11 (0.06, 20.25)		
BSC	0.58 (0.05, 6.21)	0.72 (0.04, 13.18)	0.65 (0.02, 18.60)	
Cisplatin plus gemcitabine	1.03 (0.10, 11.16)	1.26 (0.07, 23.54)	1.14 (0.04, 33.17)	1.77 (0.06, 51.21)

HRs less than 1 favour the column treatment. The trials informing the comparison versus cisplatin plus gemcitabine were conducted in gemcitabine-naïve populations, and therefore cannot be considered to provide relevant data for the retreatment with first-line platinum-based chemotherapy. This comparison is included as a scenario analysis only and results should be interpreted with caution.
Abbreviations: BSC: best supportive care; HR: hazard ratio.

Model assessment

The fixed effect and random effects models were run with three chains, each for a total of 250,000 iterations. The first 150,000 iterations were discarded. Plots of the Brooks-Gelman-Rubin diagnostic indicated satisfactory convergence.

Results – Progression free survival

Both fixed and random effects models were evaluated. Table 16 provides a summary of the model fit statistics. The table indicates that the fixed effect model had the lowest DIC. This suggests that there is minimal between-study heterogeneity. It was not possible to evaluate inconsistency because the network does not include any comparisons informed by both direct and indirect evidence. The results of the fixed effect and random effects models are shown below.

Table 16: Model fit statistics for progression free survival (standard NMA model with constant HRs)

	$\bar{D}_{res}$	pD	DIC
Fixed effect model	2.2	2.0	4.2
Random effects model	2.6	2.6	5.2

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Abbreviations: $\bar{D}_{res}$: residual deviance; DIC: deviance information criterion; pD: number of effective parameters.

For the fixed effect model, Figure 10 illustrates the HRs for nivolumab versus each of the comparators and Table 17 provides the estimates of the HRs and their 95% credible intervals. There is no evidence of a difference between nivolumab and any of the comparators.

Figure 10: Progression free survival: network meta-analysis results (standard fixed effect NMA model with constant HRs): HRs for nivolumab versus each of the comparators

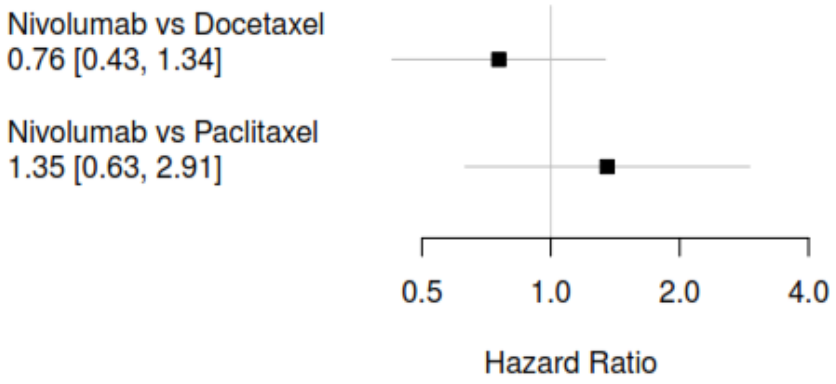


Table 17: Progression free survival: network meta-analysis results (standard fixed effect NMA model with constant HRs): HRs and 95% credible intervals for each pairwise comparison

	Nivolumab	Docetaxel
Docetaxel	0.76 (0.43, 1.34)	
Paclitaxel	1.35 (0.63, 2.91)	1.79 (0.69, 4.67)

HRs less than 1 favour the column treatment.
Abbreviations: HR: hazard ratio.

For the random effects model, Figure 11 illustrates the HRs for nivolumab versus each of the comparators and Table 18 provides the estimates of the HRs and their 95% credible intervals. There is no evidence of a difference between nivolumab and any of the comparators.

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Figure 11: Progression free survival: network meta-analysis results (standard random effects NMA model with constant HRs): HRs for nivolumab versus each of the comparators

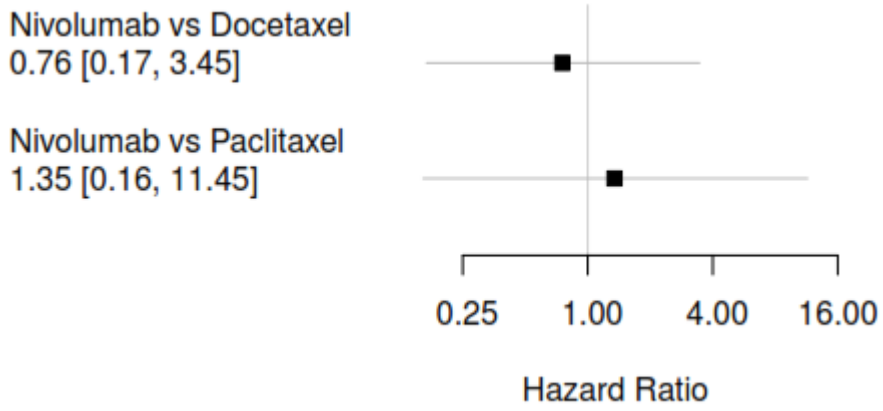


Table 18: Progression free survival: network meta-analysis results (standard random effects NMA model with constant HRs): HRs and 95% credible intervals for each pairwise comparison

	Nivolumab	Docetaxel
Docetaxel	0.76 (0.17, 3.45)	
Paclitaxel	1.35 (0.16, 11.45)	1.78 (0.13, 24.61)

HRs less than 1 favour the column treatment.
Abbreviations: HR: hazard ratio.

Model assessment

The fixed effect and random effects models were run with three chains, each for a total of 250,000 iterations. The first 150,000 iterations were discarded. Plots of the Brooks-Gelman-Rubin diagnostic indicated satisfactory convergence.

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

WinBUGS code

The WinBUGS code for the fixed effect model and random effects model is provided below. Both models were adapted from Example 7 of the NICE DSU TSD 2. This code, and the associated data, is also provided in the package of code.

Standard NMA model for HR data (random effects)

```
model{ # *** PROGRAM STARTS

	for(i in 1:ns) { # LOOP THROUGH 2-ARM STUDIES

		# y[i] is the predicted log HR
		# prec[i] is the precision of the log HR
		y[i] ~ dnorm(delta[i],prec[i]) # normal likelihood for 2-arm trials

		delta[i] ~ dnorm(md[i],tau) # trial-specific treat effects distributions
		md[i] <- d[t[i]] - d[b[i]] # mean of treat effects distributions
		dev[i] <- (y[i]-delta[i])*(y[i]-delta[i])*prec[i] #Deviance contribution

	}

	totresdev <- sum(dev[]) #Total Residual Deviance

	d[1]<-0 # treatment effect is zero for reference treatment
	for (k in 2:nt){ d[k] ~ dnorm(0,.0001) } # vague priors for treatment effects

	sd ~ dunif(0,2) # vague prior for between-trial SD
	tau <- pow(sd,-2) # between-trial precision = (1/between-trial variance)

	for (k in 1:nt) {
		for (c in 1:nt) {
			hr[c,k] <- exp(d[c]-d[k])
		}
	}
} # *** PROGRAM ENDS
```

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Standard NMA model for HR data (fixed effect)

```
model{ # *** PROGRAM STARTS

        for(i in 1:ns) { # LOOP THROUGH 2-ARM STUDIES

                # y[i] is the predicted log HR
                # prec[i] is the precision of the log HR
                y[i] ~ dnorm(md[i],prec[i]) # normal likelihood for 2-arm trials

                md[i] <- d[t[i]] - d[b[i]] # mean of treat effects distributions
                dev[i] <- (y[i]-md[i])*(y[i]-md[i])*prec[i] #Deviance contribution
         }
        totresdev <- sum(dev[]) #Total Residual Deviance

        d[1]<-0 # treatment effect is zero for reference treatment
        for (k in 2:nt){ d[k] ~ dnorm(0,.0001) } # vague priors for treatment effects

        for (k in 1:nt) {
                        for (c in 1:nt) {
                                hr[c,k] <- exp(d[c]-d[k])
                        }
        }
} # *** PROGRAM ENDS
```

Economic model – Results

In total, two different sets of constant hazard ratios (HRs) were incorporated into the economic model (fixed effects and random effects). The results in the economic model when these HRs are adopted, with all other base case parameters remaining unchanged, are summarised in Table 19 and Table 20. These results show an increase in the ICERs compared with the base case. However, it should be noted that these scenarios are likely to substantially overestimate the gains in progression-free and overall survival with each of the comparators, based on the currently understanding of patient outcomes with second-line urothelial cancer. To illustrate the implausibility of the predictions with constant HRs (fixed effect, with random effects showing very similar outputs) the OS curves for paclitaxel, docetaxel and best supportive care are plotted, in which the predicted OS values are shown against the available Kaplan-Meir data. They indicate that OS is substantially overestimated,

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

from around one year onwards and especially versus long term survival estimates of current therapies.

Given the clinical implausibility of the survival estimates for the comparator arm, combined with the fact that a lack of proportional-hazards make the standard NMA approach invalid, this scenario should be considered implausible and inappropriate for decision-making.

Please note, during the updates that were made to the model to address the clarification requests a few minor errors were identified in the model calculations. These have been corrected in the latest version, which has caused very small changes to the overall results (maximum change in the ICER is an increase of approximately £140 for the comparison of nivolumab versus best supportive care). All results reported here use the corrected version of the model. The changes to the model calculations, and latest base case results, are reported in Appendix 1.

Table 19: Results with constant HR, fixed effects

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus Nivolumab (£/QALY)
Nivolumab	[redacted]	2.78	[redacted]				
Paclitaxel	£15,942	2.39	1.66	[redacted]	0.39	[redacted]	£272,284
Docetaxel	£14,033	2.03	1.28	[redacted]	0.75	[redacted]	£76,095
BSC	£9,296	1.14	0.78	[redacted]	1.64	[redacted]	£43,279

Table 20: Results with constant HR, random effects

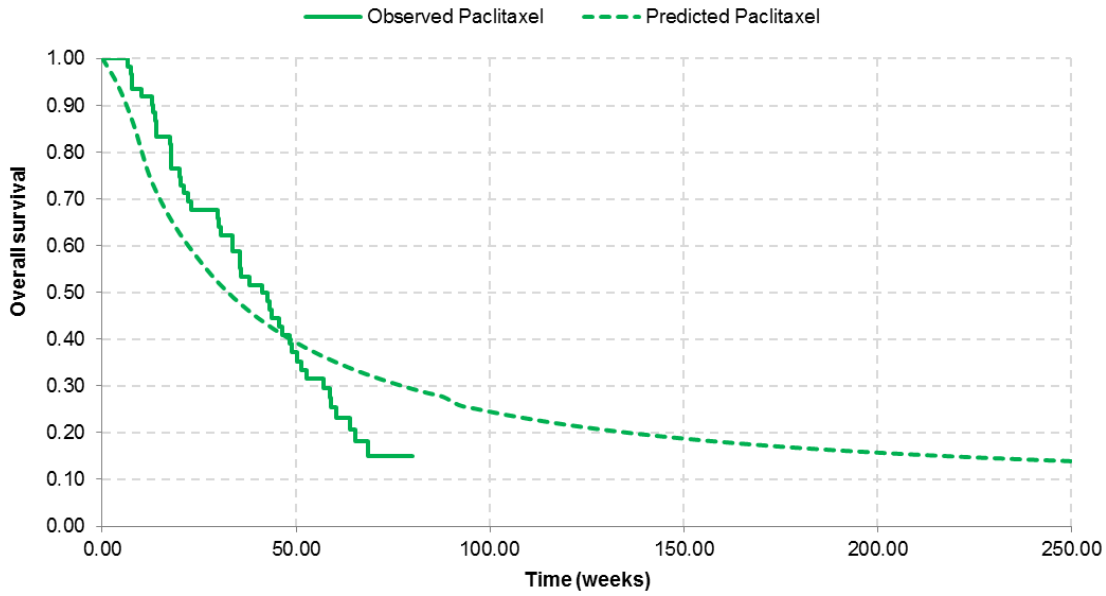
Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus Nivolumab (£/QALY)
Nivolumab	[redacted]	2.78	[redacted]				
Paclitaxel	£16,008	2,42	1.68	[redacted]	0.36	[redacted]	£317,625
Docetaxel	£14,033	2.03	1.28	[redacted]	0.75	[redacted]	£76,095
BSC	£9,245	1.11	0.77	[redacted]	1.67	[redacted]	£42,677

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

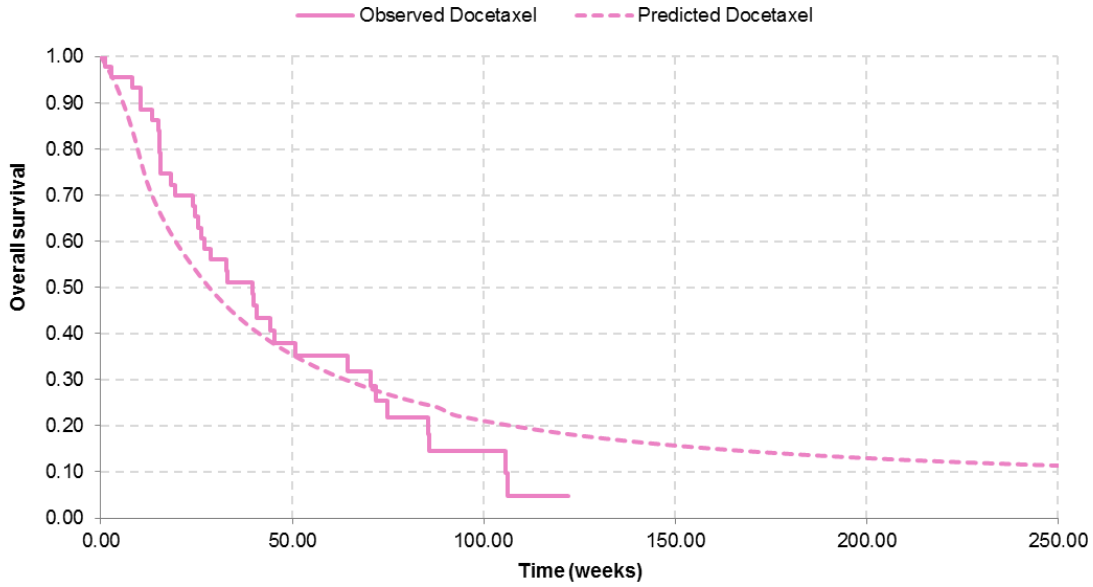
+44 (0)300 323 0140

Figure 12: Overall survival for paclitaxel, docetaxel and BSC with fixed effects constant hazard ratios

Paclitaxel



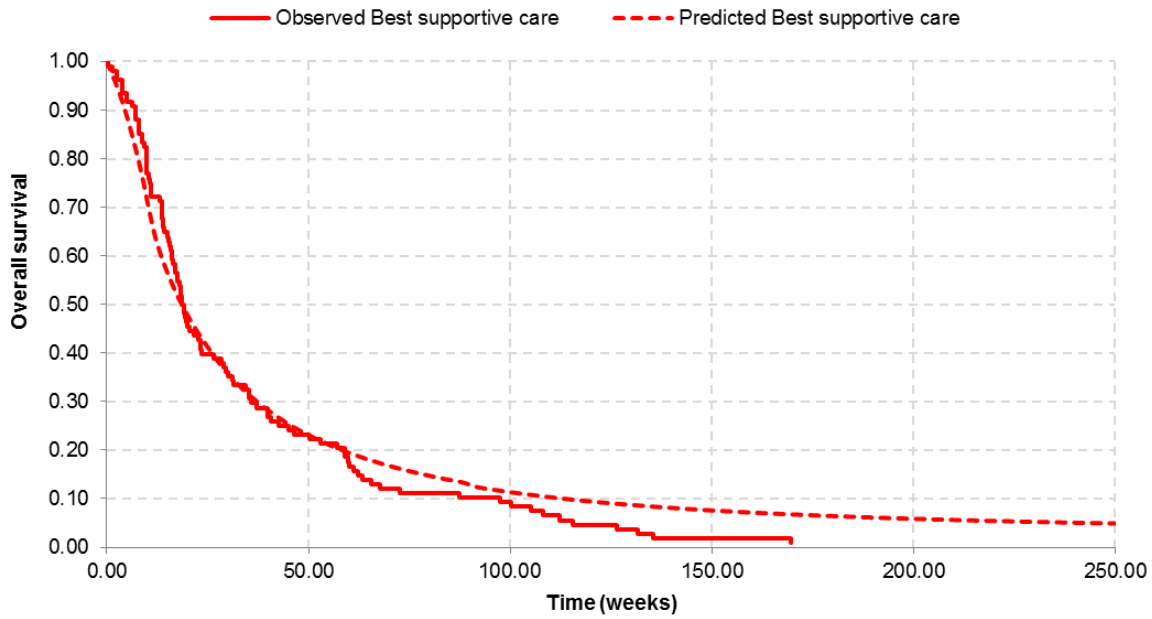
Docetaxel



Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Best supportive care



D. Please provide scenario analysis using further alternative model specifications, with polynomials other than p1=0, p2=0 and p1=1, p2=1. Please also describe how and justify why these particular polynomials are chosen.

At the NMA stage, 5 different fractional polynomial models were fitted, including two first order models and three second order models. These models were selected because they allowed the hazard rates to follow a range of different patterns over time. The two first order models are equivalent to the Weibull and Gompertz models, respectively. The Weibull and Gompertz models allow for hazard functions that are either constant or increase or decrease monotonically. The three second order models allow the hazard function to take a wider variety of forms, including U-shaped and inverted U-shape curves. For each survival outcome, we fitted each of these 5 models as both a fixed effect model and a random effects model. Model fit statistics and clinical plausibility were then used to select the most appropriate models for inclusion in the cost-effectiveness model. We found that the models with p1=0, p2=0 and p1=1, p2=1 were the most appropriate, and thus these were included in the model.

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

The economic model has been updated such that all 10 fractional polynomial models are now included. Given 10 polynomials are included and for both PFS and OS this equates to 100 different results combinations. Therefore, for brevity they are not reported but can be viewed using the economic model.

B5. It is not clear why it is necessary to use the same survival model (generalised gamma distribution) for responders and non-responders.

A. Please provide justification for why it was deemed necessary to use the same survival model for responders and non-responders.

This was a simplifying assumption. Using different distributions for OS and PFS would mean there are 36 different survival curves for each endpoint to choose from.

B. Were clinical experts consulted to support the choice of survival model? If so, please provide the methods for eliciting expert opinion including the number of experts and questions asked as well as the results.

Clinical experts were consulted during an advisory board which included six clinicians and two health economists.[28] Parametric survival curves were presented based on CheckMate 275 data only and included a selection of survival curves (as stated in the advisory board report, included in the reference pack). As a result of feedback from the clinicians and health economists during the advisory board, the survival analysis was re-estimated to include the CheckMate 032 data and adopting the response-based survival analysis, to capture the subpopulation of patients that live for an extended period of time. Further validation of the final overall survival estimates was then done using longer-term data for nivolumab in lung cancer, following clinical expert advice.

C. Please provide an implementation in the model by which it is possible to use differential curves for responders and non-responders. Please also provide a scenario analysis, in which the best fitting curves are chosen separately for responders and non-responders, e.g. using the Weibull for non-responders' OS and PFS and, in two separate scenarios, the exponential and the generalised gamma for responders' OS and PFS.

The model has been updated so that it is now possible to select different distributions for responders and non-responders. Based on the statistical goodness-of-fit two additional scenarios have also been analysed. For the first scenario the Weibull and generalised

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

gamma distributions have been selected for non-responders and responders respectively (both PFS and OS). For the second scenario, again Weibull has been selected for non-responders whereas for responders the exponential has been adopted for OS and generalised gamma for PFS. The results for these scenarios are presented in Table 21 and Table 22.

Table 21: Results with PFS/OS separated by response, Weibull for non-responders and generalised gamma for responders

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus Nivolumab (£/QALY)
Nivolumab	[redacted]	2.71	[redacted]				
Paclitaxel	£14,276	1.12	0.72	[redacted]	1.59	[redacted]	£37,750
Docetaxel	£13,731	1.30	0.87	[redacted]	1.40	[redacted]	£44,388
BSC	£8,938	0.95	0.61	[redacted]	1.76	[redacted]	£38,631

Table 22: Results with PFS/OS separated by response, Weibull for non-responders and generalised gamma (PFS)/Exponential(OS) for responders

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus Nivolumab (£/QALY)
Nivolumab	[redacted]	2.01	[redacted]				
Paclitaxel	£13,983	0.97	0.63	[redacted]	1.04	[redacted]	£54,208
Docetaxel	£12,980	1.15	0.76	[redacted]	0.86	[redacted]	£68,256
BSC	£8,790	0.88	0.56	[redacted]	1.14	[redacted]	£55,946

D. PFS and OS curves were adjusted to account for general population mortality using age-adjusted annual mortality rates. Please discuss the method to implement this and provide justification for this approach. Please also justify why both, PFS and OS, had to be adjusted instead of just OS. Please also discuss whether this has any impact on the plausibility of the OS estimates.

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

The general mortality data were calculated in the 'General mortality data' sheet in the model. Because age-related mortality is non-linear, it is not appropriate to assume a mean age. Therefore, a weighted mortality rate was estimated for each cycle in the model. The spread of ages was calculated based on the mean (65) and standard deviation (9.38) in the trial, and assuming a normal distribution. The weights applied to each age are shown in row 2 (columns F to BN). For each cycle in the model, general population survival was reduced by a proportion equivalent to the weighted four-week mortality rate for the cohort (shown in 'column CB'). As the cycles in the model progressed, the mortality rate changed to reflect the ageing population. 'Column BP' shows the estimated survivorship for the cohort for each cycle in the model. This proportion was used to multiply the extrapolated (not observed) part of the PFS and OS curves. Over a relatively short trial period (<2 years), the impact of an ageing cohort is very unlikely to be picked up in parametric fits to the data. This feature was added to address previous committee concerns that certain parametric distributions would lead to constantly decreasing hazards, which was criticised.

B6. Time to treatment discontinuation (TTD) was estimated irrespective of response status (inconsistent with OS and PFS). However, treatment is discontinued when patients no longer benefit from it. It could therefore be suspected that TTD differs significantly for responders versus non-responders.

A. **Priority question**: Please implement survival models for TTD using the same response-based survival analysis as for PFS and OS (currently landmark analysis with 8-week landmark) in the cost effectiveness model and provide the results of this in a scenario analysis.

Survival models for TTD have been implemented using the 8-week landmark analysis. Model summaries and probabilities have been included with this package. As with the original analysis six different parametric distributions were included for TTD. The AIC and BIC scores for the statistical goodness-of-fit of each distribution are presented in Table 23. They indicate that there is no clear choice of distribution in terms of providing the best fit. However, it should be noted that the generalised gamma, which was adopted for TTD in the base case analysis, provides a fit which is in the middle-range of the six possible distributions when using the response-based approach. Based on the AIC/BIC scores the most suitable distributions are Gompertz and log-logistic. Therefore, the results with these scenarios are presented in Table 24 and Table 25.

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Of note, the resulting ICERs are associated with a considerable proportion of patients continuing to be treated with nivolumab after 4, 5 or more years, which is unlikely to transpire in clinical practice. Scenarios with a proportion of eligible patients continuing treatment after 2 years are likely to be much more reflective of use in clinical practice.

Table 23: AIC and BIC scores for response-based survival curves for TTD, 8 week landmark

Distribution	Responders		Non-responders	
	AIC	BIC	AIC	BIC
Weibull	290.85	295.43	987.98	994.00
Exponential	289.65	291.94	1002.33	1005.34
Gompertz	291.64	296.22	985.21	991.24
Log-logistic	289.60	294.18	987.59	993.61
Lognormal	288.41	292.99	994.45	1000.47
Generalised gamma	290.13	297.00	988.85	997.88

Table 24: Results with response-based TTD curves, Gompertz

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus Nivolumab (£/QALY)
Nivolumab	[redacted]	2.78	[redacted]				
Paclitaxel	£14,430	1.19	0.76	[redacted]	1.60	[redacted]	£57,022
Docetaxel	£13,913	1.40	0.92	[redacted]	1.38	[redacted]	£67,859
BSC	£9,052	1.01	0.64	[redacted]	1.77	[redacted]	£55,626

Table 25: Results with response-based TTD curves, Log-logistic

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus Nivolumab (£/QALY)
Nivolumab	[redacted]	2.78	[redacted]				
Paclitaxel	£14,430	1.19	0.76	[redacted]	1.60	[redacted]	£52,998
Docetaxel	£13,913	1.40	0.92	[redacted]	1.38	[redacted]	£63,111
BSC	£9,052	1.01	0.64	[redacted]	1.77	[redacted]	£52,028

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

A 26 week landmark analysis can also be selected for treatment discontinuation in the model. Please use cell BX28 on the 'Discontinuation' page to switch between the 8 week and 26 week landmark. The AIC/BIC scores for the response-based survival curves for TTD using the 26 week landmark (when using this scenario for OS and PFS) are presented in Table 26.

Table 26: AIC and BIC scores for response-based survival curves for TTD, 26 week landmark

Distribution	Responders		Non-responders	
	AIC	BIC	AIC	BIC
Weibull	153.98	156.10	267.83	269.83
Exponential	149.49	155.83	266.72	272.75
Gompertz	155.83	160.05	263.67	267.68
Log-logistic	154.48	158.70	264.53	268.54
Lognormal	152.76	156.98	264.83	268.85
Generalised gamma	155.38	159.60	265.73	269.74

B. Please provide justification for the survival model choice for TTD, with description of the clinical expert opinion and methods to elicit this.

As stated in section B.3.3.3:

- Discontinuation from treatment is not based solely on patient progression
- The generalised gamma distribution was chosen for the base case. This was to ensure consistency with the choice made for PFS and OS. Furthermore, whilst two distributions produced lower AIC/BIC scores (Gompertz and log-logistic), indicating a better fit, these two distributions also produced very long tails with a percentage of patients on treatment at 5 and even 10 years. This is not in keeping with the expected clinical use of nivolumab and therefore these distributions lack clinical validity.

B7. The ERG noticed several inconsistencies between the Checkmate (032 and 275) trials, the company submission and the cost effectiveness model concerning the number of responders used for the OS landmark analysis. Objective response was achieved in 19 and 52 patients in CheckMate 032 and 275, respectively, totalling 71

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

responders.[29, 30] However, the number of responders at the 8-week landmark for OS estimation, cells DD10 and DE10 of the cost effectiveness model, is 73 patients. In addition, the numbers of responders and non-responders provided in Figure 35 of the company submission for the PFS landmark analysis do not correspond to the number of responders used in the cost effectiveness model. Please clarify which figures are correct for the PFS and OS landmark analyses, amend the cost effectiveness model if necessary and provide the cost effectiveness results using the correct number of responders.

The objective response rate used in the economic model is based on the most recent database lock for both trials. This corresponds to 19 and 54 patients in CheckMate 032 and CheckMate 275, respectively. The number of responders at the 8 week landmark for OS is therefore correct.

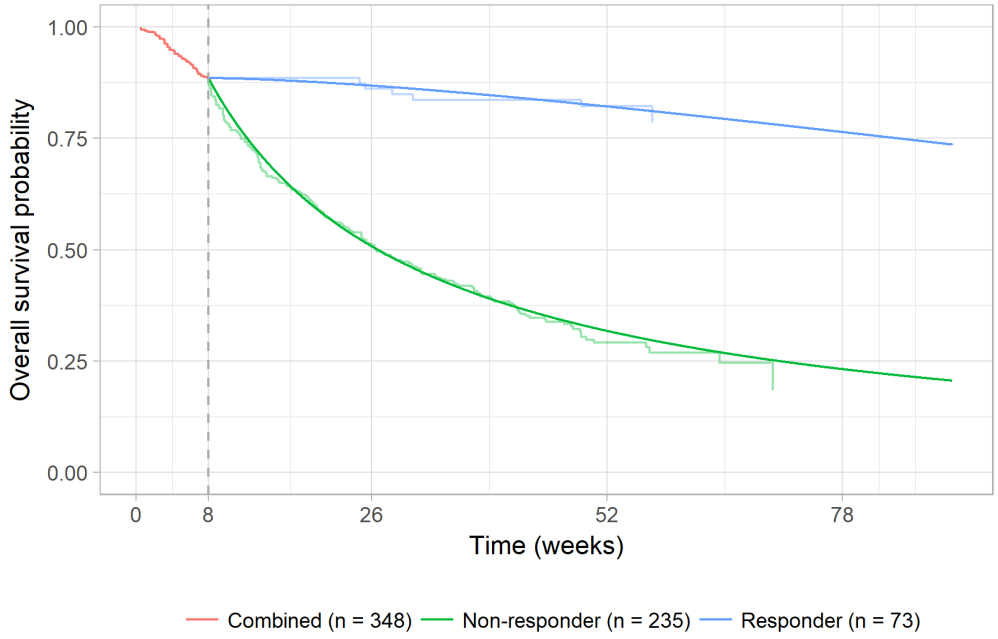
An overview of the clinical data from CheckMate 275, detailing the primary analysis (May 2016) and the updated results (September 2016) is provided in Section B.2.6.2.

The correct version of Figure 35 is provided below (Figure 13), whereby the incorrect labelling has been amended.

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Figure 13: Corrected "Figure 35: Week 8 landmark – progression-free survival with generalised gamma"



B8. **Priority question**: Please provide the comparison of nivolumab against cisplatin + gemcitabine in the base-case (see also question A15).

The results of a comparison of nivolumab versus cisplatin + gemcitabine can be found in Appendix O.

B9. The company assumes that the hazard ratio of BSC versus vinflunine can be applied to the paclitaxel PFS curve in order to calculate PFS for the BSC comparator. The company justifies this assumption by stating that the outcomes between vinflunine and paclitaxel/docetaxel are expected to be similar and therefore that this hazard ratio can be applied to the paclitaxel PFS curve in order to obtain PFS estimations for BSC. However, no evidence is provided to support this assumption. Please provide clinical evidence to support this assumption.

BMS are unaware of any clinical data to support this. The similar need for an assumption to estimate PFS was also required for gemcitabine + cisplatin. Again, this is in the absence of any clinical data.

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Model structure

B10. **Priority question**: Please provide an implementation of the model, in which there are separate health states for responders and non-responders (instead of using PFS and OS based on weighted averages).

A. Please add an implementation with differential hazard ratios for OS and PFS for responders and non-responders.

Please see the response to question B4.

B. Please discuss the plausibility of applying differential utility values and resource use for responders and non-responders, and apply these if applicable.

Separate health states have not been created for responders and non-responders; therefore it is not plausible to implement this request.

B11. The company uses a partitioned survival model approach. Could the company provide additional justification for this approach, other than that it has been used in previous appraisals on metastatic cancers and TA272, especially in the light of criticism of partitioned survival models compared with state transition models according to TSD19, which includes that endpoints are treated as independent and that intermediate health states are not reflected?[31]

The model was developed prior to the publication of TSD19, where use of partitioned survival model approach for consistency with previous models in related disease areas was a common justification. Further, as with any model, there are a variety of different ways in which the structure could be approached. The DSU document raises some key points. The independence of endpoints is a necessary assumption in partitioned survival analysis, but this is only likely to affect the representation of the ranges in the probabilistic sensitivity analysis, rather than the base case deterministic results. Whilst other approaches (such as discrete event simulation models) may have advantages in terms of built-in 'memory' and the ability to use prognostic indicators for individual patients, they are only advantageous if they are supported by reliable data. If patient-level prognostic data do not exist, or cannot be agreed upon by experts, then a discrete event simulation will not differ from a partitioned survival model (since probabilities would cease to vary from patient-to-patient and would not vary following different timings of events).

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Adverse events

B12. The company uses adverse event rates for the comparators from sources that are only named briefly in the company submission, without explanation. Please provide an overview of and justification for the chosen sources.

The sources for the adverse event rates are provided in the company submission. For these rates the same sources were used as those adopted in the prediction model, which was used to estimate the hazard ratios for each comparator versus nivolumab. This was to ensure consistency between the sources used for treatment effectiveness and adverse event rates. It should also be noted that these adverse event rates do not have any significant impact on the overall results of the economic model.

Probabilistic sensitivity analysis

B13. The probabilistic sensitivity analysis (PSA) produces errors, does not include all input parameters, uses a small number of simulations, produces results that are different from the deterministic analysis and does not appear to be reproducible. Full incremental results are not provided.

A. **Priority question**: Please provide full incremental analysis with all comparators included in the PSA simultaneously, showing incremental costs and QALYs of nivolumab and all comparators.

The model has been updated to include an incremental analysis in which nivolumab can be compared with docetaxel, paclitaxel and BSC simultaneously. In line with the original submission, cisplatin plus gemcitabine has not been included in the analysis as it is not considered a relevant comparator in the context of second-line UK clinical practice.

B. In addition to the incremental costs and QALYs provided in Table 47 of the company submission, please also provide absolute costs and QALYs resulting from the PSA.

Absolute costs and QALYs have been added as outputs for the PSA.

C. Please include the Kaplan–Meier estimates and hazard ratios in the PSA.

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

The PSA has been updated to include extra parameters, namely the: gender ratio, unit cost of docetaxel, unit cost of paclitaxel and Kaplan-Meier estimates for nivolumab. However, it has been determined that it is not appropriate to include the hazard ratios. As noted in the company submission the inclusion of hazard ratios would generate illogical results due to the time-varying nature of the hazard ratios. For example, if included in the PSA the hazard ratio may start at a value of 2 in cycle 1, decrease to 1 for the next cycle and then increase to 3 for the following cycle. This would generate changes in PFS and OS that are not clinical plausible. Further, as noted in the company submission changes in the survival curves for each comparator are already captured in the PSA through the inclusion of the nivolumab PFS and OS curves in the PSA, which are the reference curves.

D. After re-running the PSA, the PSA results in #NUM errors in both nivolumab and comparator costs and QALYs. Please provide a corrected PSA, which does not produce errors.

The #NUM error has been corrected in the latest version of the model.

E. Please comment on the reasons for which OS and PFS may be lower in the probabilistic compared with the deterministic analysis.

It is expected that the difference is due to the ranges and distributions applied in the probabilistic sampling of PFS and OS, which are generating a lower probabilistic treatment effect compared with the deterministic treatment effect.

F. The ERG did not obtain probabilistic results similar to those reported by the company. Could the company ensure the reproducibility of the PSA and provide a version of the model with identical PSA results as provided in the company submission?

The latest version of the model should generate more reproducible results.

G. Please increase the number of PSA simulations to at least 10,000 PSA simulations (or more if needed to provide reproducible PSA results).

The number of simulations has been increased to 10,000 in the latest version of the model.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Health-related quality of life

B14. The company explains that 204 observations of utilities were missing because of the immaturity of the dataset. The dataset was finalised 2nd September 2016.

A. Please justify the choice to impute future observations (i.e. questionnaires that were due to be administered after the data-cut of 2nd September 2016) and discuss the implications.

The observations beyond the cut-off were treated as missing in the analysis as they can be considered analogous to censored observations, in that an individual patients' EQ-5D profile over the course of the trial is only partially known. As multiple imputation has been shown to improve estimates when implemented for censored data, it was deemed appropriate to impute these future observations in the present analysis.

An implication of this is the assumption that all patients not fully observed to the length of follow-up were assumed to survive for the duration of the trial. As those at the 60 and 72-week follow-ups were associated with much lower health-related quality of life, the imputation procedure would predict low utility for future observations at these periods when it would be reasonably expected that a proportion would have died. Therefore, the mean utility values would be biased downwards for both pre-progression and progressed disease states.

B. Please provide the analyses presented in the company submission using a more recent data-cut.

The last data cut available from Checkmate 275 is from 2nd September. As stated in the submission, the next scheduled database lock is [redacted]

B15. Utility estimates are derived from CheckMate 275 only, disregarding CheckMate 032 and NICE TA 272.

A. Please comment on the reasons for disregarding the utility estimates used in NICE TA 272, discuss how the utility estimates derived from CheckMate 275 differ, and discuss the implications for model outcomes.

As noted by BMS in the decision problem pro forma, utility data was not collected in study 302. Pre-progression utilities were estimated using responses to 1 item from the EORTC

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

QLQ-C30 questionnaire used in study 302, transformed to health-state utilities using a published regression model relating this measure to utility values from a time-trade-off analysis in a sample of US cancer patients and their relatives. Post-progression utilities were taken from a study reporting EQ-5D values in 1270 terminally ill cancer patients with painful bone metastases or poor-prognosis non-small-cell lung cancer.

The utility estimates used in NICE TA 272 were subject to criticism from the ERG and the appraisal committee:

"it (the ERG) also stated that the utility values used did not fit with the preferred NICE reference case, and that there is considerable uncertainty around these estimates because standard methods were not used".

"The Committee noted that neither utility used in the economic model conformed to the NICE reference case and concluded that the lack of appropriate utility data contributed to uncertainty in the model".

Available at https://www.nice.org.uk/guidance/ta272/documents/transitional-cell-carcinoma-of-the-urothelial-tract-vinflunine-appraisal-consultation-document

EQ-5D data was available from patient receiving nivolumab in a patient population which matches the decision problem, providing a strong rationale for using this data to estimate utility.

In terms of the use of these utility values for the comparator therapies rather than the values for TA272 – a summary is provided in Table 64 of Appendix H. This shows the value for PFS used in this appraisal is higher than the value for PFS used for vinflunine and BSC in TA272. This may bias against nivolumab, if the utility value used in TA272 for vinflunine and BSC progression-free survival is correct. However, given the limitations highlighted with the utility values from TA272, BMS were cautious of using them to model the chemotherapy treatments.

Nivolumab has been shown to be associated with an improved quality of life compared to chemotherapy across a number of indications (Harrington *et al.* [2017],[32] Reck *et al.* [2015],[33] Reck *et al.* [2016][34]). However, comparator evidence supporting a quality of life benefit for nivolumab is not yet available in urothelial carcinoma. Therefore, in order to be conservative the same health state utility values have been used across treatments, adjusted only with adverse event disutilities.

B. Please comment on the reasons for disregarding the utility estimates from CheckMate 032, discuss how the utility estimates derived from CheckMate 275 differ, and discuss the implication for model outcomes.

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

The CheckMate 032 utility data has been included in the analysis and presented below.

C. Please provide an analysis using the utility estimates from both CheckMate studies.

When the utility estimates from CheckMate 032 are also included (i.e. pooled CheckMate -275 and -032 data), and analysed using the imputation and mixed model methods described in the company submission, the values of 0.736 and 0.632 are calculated for the pre-progression and post-progression health states respectively. They show there is a small increase in utility, for both states, when compared with the CheckMate -275 data only. A scenario analysis using these values has been undertaken and the results are summarised in Table 27.

Table 27: Results with pooled CheckMate- 275 & 032 utility data

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus Nivolumab (£/QALY)
Nivolumab	[redacted]	2.78	[redacted]				
Paclitaxel	£14,430	1.19	0.79	[redacted]	1.60	[redacted]	£36,567
Docetaxel	£13,913	1.40	0.95	[redacted]	1.38	[redacted]	£43,662
BSC	£9,052	1.01	0.66	[redacted]	1.77	[redacted]	£37,216

B16. Please provide justification for, and discuss the suitability of, the approach used to obtain utility values.

A. Please report the deviation in time for the interpolated observations (i.e. number of cases, mean, median, standard deviation, minimum and maximum).

There were a total of 117 observations encoded as ‘unscheduled follow-ups’; that is, as either ‘Follow-up 1’, ‘Follow-up 2’, ‘Survival follow-up 1’ or ‘Survival follow-up 2’. These ranged from five to 66 weeks with a mean of 32. A preliminary comparison showed that when these fell near a scheduled follow-up time, the latter was missing. We therefore used these unscheduled questionnaires to replace the missing observations if they fell within two weeks either side of the scheduled time. The maximum and minimum deviation from the

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

scheduled time was therefore 2 and -2, with a median deviation of 0 weeks. The mean deviation was -0.047 weeks, with a standard deviation of 1.50 weeks.

B. Please specify why it was deemed necessary to impute data despite the use of a mixed model (which has methods to deal with missing data).

BMS aims to address previous concerns from appraisals regarding nivolumab, where appropriate. Based on feedback from the same ERG and the same Committee as part of ID971, multiple imputation was explored.

There is a minimal difference in the pre-progression values between the non-imputed and imputed values when using the mixed model (0.0055), and a larger difference in the post-progression values (-0.0484).

C. Please discuss the differences in methods and results of observed and imputed values in Table 35 in the company submission.

Table 35 shows, when using the complete case and imputed datasets, the utility values of the pre-progressed patients and the disutility associated with being in the progressed state. These values are taken directly from linear mixed model, in which the constant represents the pre-progressed utility and the fixed effect coefficient for progression represents the utility decrement of progressed disease. For the imputed data analysis, this involves estimating the linear mixed model on each of the imputed datasets and pooling the estimates.

The disutility associated with being in the progressed disease state is greater in the imputed data than the complete case analysis. This is because a greater proportion missing observations are for progressed patients who have discontinued treatment. Increasing the proportion of patients who have discontinued treatment in the progressed group therefore lowers the average utility of that group.

D. Please justify why predictive mean matching was chosen as the imputation method and the limitations associated with this method in the context of a large amount of missing data. Additionally, please provide additional details regarding the imputation methods (e.g. which variables were included) and discuss the plausibility of the imputed data.

Predictive mean matching was adopted due to the nature of the variables being imputed (i.e. the EQ-5D dimensions). Firstly, as they are ordinal, a standard linear regression approach

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

was not adopted. Predictive mean matching was selected over ordinal logistic regression as it is less sensitive to model misspecification, including heteroscedasticity in the error term and non-linear associations with regressors. The robustness of this technique has not been systematically verified under all conditions, although studies have indicated that performance reduces when the proportion of missing data is greater than 30%.[35] However, we do not expect the bias to substantial despite the large proportion missing data in this case due to (i) the large sample size of over 750 complete observations and (ii) the number of 'donor' observations from which the imputed value is selected is set at three rather than five to limit dissimilarities between the donor and the missing observation.

The variables included in the imputation model were a mixture of patient characteristics and observation-specific variables, as well as the values of each EQ-5D dimension at baseline. The following variables are included in the imputation model:

- Age
- Gender
- Ethnicity
- Country
- PD-L1 category
- Visit (i.e. baseline, week 9, etc)
- EQ-5D dimensions at baseline
- EQ-5D dimensions
- Progression status
- Treatment status

The proportion of missing observations for the EQ-5D dimensions ranged from 46.1% (usual activities) to 46.6% (mobility). Baseline EQ-5D scores also had missing values, ranging from 4.6% (pain, anxiety, self-care and usual activities) to 5.3% (mobility). All remaining variables were complete for all observations.

As the predictive mean matching process matches missing values with 'similar' observed values, the distribution of the utility values is in line with what is expected given the nature of

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

the missing data. Since many more missing values are from those who have progressed disease and who have discontinued treatment, which is associated with lower health-related quality of life. As such, a lower proportion have higher utility values.

E. Please justify why a mixed model was used and provide diagnostics of the mixed model.

A mixed model was used because has been used and accepted in previous nivolumab appraisals. Namely, ID971 and TA417. Further, a mixed model was used to account for the random effect of each individual on from the estimate of the fixed effect of being in the progressed disease state. A linear mixed model is a simple approach that deals with this nested structure. Due to the model being run on multiple imputed datasets, standard diagnostic measures such as Akaike and Bayesian Information Criteria, log likelihood or the intraclass correlation coefficient are not provided by the software. Instead, the following output is provided:

	EQ-5D	t	df	Pr(>|t|)	Lower 95% CI	Upper 95% CI
Constant	0.718	43.49	818.85	0	0.686	0.751
	(0.0165					
Progressed	-0.115	-3.94	62.19	0.000	-0.173	-0.0565
	(0.0291)					

F. Please explore adding a variable for a patient being on treatment or not to the mixed model and adding a variable for response status to the mixed model, provide results and discuss the impact on model outcomes.

The effect of including treatment discontinuation as an additional variable is shown in Table 28. The independent effect of progression on health-related quality of life decreases from -0.115 to -0.057, indicating that the decrement for those who have discontinued treatment is greater than those remaining on treatment. This is verified by the negative coefficient on the interaction term between progression and treatment discontinuation. However, neither the treatment discontinuation variable nor this interaction term are statistically significant. The final utility values, separated by both progression status and treatment status, are presented in Table 29.

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

The expected impact of these values on the model outcomes is an increase in the cost-effectiveness of nivolumab versus docetaxel and paclitaxel. This is because, in general, nivolumab patients remain on treatment for longer (particularly compared with paclitaxel as all patients stop treatment after week 24). Therefore, nivolumab patients are likely to have a greater number of QALYs overall due to the utility scores whilst on treatment (i.e. the on treatment utility scores of 0.723 and 0.666 are higher than the values of 0.718 and 0.603 that are applied in the base case, regardless of treatment status). The utility values presented in Table 29 are unlikely to be valid for best supportive care as these patients are not given an active treatment so treatment status is not a relevant covariate.

Response status could not be included as a variable in the mixed model as the relevant data was unavailable at the time of the analysis and in the timeframe required.

Table 28: Results of linear mixed model including treatment discontinuation as a covariate

	Coefficient	t	df	Pr(>|t|)
Constant	0.723 (0.0167)	43.17	716.22	0
Progressed	-0.0570 (0.0265)	-2.153	229.21	0.032
Discontinued	-0.0734 (0.0506)	-1.451	139.89	0.149
Progressed*Discontinued	-0.0192 (0.0518)	-0.371	282.67	0.711

Table 29: Final utility values with linear mixed model including treatment discontinuation as a variable

	Pre-progression	Post-progression
On treatment	0.723	0.666
Off treatment	0.650	0.573

G. The company submission states that 'the generated utility values for post-progression patients was seen to increase and decrease in a manner that would not be expected in clinical practice.' Please explain how it was determined that the time-dependent utilities obtained were not in line with clinical practice.

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Given the data, it was felt that using health state specific utility values would be more appropriate.

Resource use and costs

B17. With regard to the calculation of drug and administration costs:

A. Please comment on the reasons for disregarding resource use (e.g. for calculating drug and administration costs) from CheckMate 032, discuss how the resource use derived from CheckMate 275 differs, and discuss the implications of this.

BMS are unsure of the nature of this question. To estimate time on treatment, a parametric survival curve for time to treatment discontinuation (TTD) was plotted to predict the proportion of patients receiving treatment at each cycle. This curve was based on individual patient data (IPD) from the CheckMate 275 and CheckMate 032 trials and, in line with the approach to PFS and OS, six distributions were plotted.

B. Please provide justification for classing delayed doses as missed doses and discuss the impact of this assumption.

No dose modifications were allowed, but predefined dose delays were permitted for adverse events in the CheckMate 275 and CheckMate 032 trials. Dose delays which exceed the duration of a treatment cycle can be reasonably assumed not to have been given at the same time as the next cycle.

This has been accepted in previous nivolumab appraisals, namely TA417, ID811, ID900 and ID971. The impact of this assumption is that it models the cost of treatment with nivolumab as it was administered in the trial.

C. Please provide justification (other than absence of evidence) for assuming that the dose intensity for docetaxel, paclitaxel, gemcitabine plus cisplatin was equal to that of nivolumab.

The publications available for these comparators were checked for dose intensity information; none was provided. Other options included assuming 100% dose intensity for docetaxel, paclitaxel and gemcitabine plus cisplatin but this would bias in favour of nivolumab and so would be open to criticism by the ERG and appraisal committee.

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

D. Please provide justification that the weight and body surface area from CheckMate 275, used to calculate drug costs, is applicable to patients in the UK setting.

The weight and body surface area values used in the economic model are from the European patients in the CheckMate-275 trial, specifically. Given the similarities in demographics, and in the absence of further UK specific data, the European values were chosen as the most appropriate, available source. A similar value for BSA was used in TA272 (BSA = 1.85 m^2) though that appraisal did not report an average weight value.

E. In the company submission it states 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)". Please provide justification why 3 gemcitabine administrations were assumed per 4 weeks instead of 2.67 administrations per 4 weeks (=2 × 28 / 21).

The administration of cisplatin plus gemcitabine in the model was based on the regimen adopted in the Gondo (2011) study as this was the key source of efficacy data for the analysis. If an alternative regimen was assumed in the model (i.e. 2 doses of gemcitabine per cycle as oppose to 3) then the efficacy data may no longer be applicable as the effectiveness of treatment is likely to be dependent on dosing. Gondo and colleagues report that gemcitabine was given on days 1, 8 and 15 with cisplatin being given on days 1 and 2. Therefore, 3 gemcitabine administrations per cycle were included in the model. The approach in UK clinical practice is noted in the submission (i.e. 2 doses of gemcitabine per cycle) to highlight that in the UK, cisplatin plus gemcitabine is given at a different dosing schedules compared to the Gondo (2011) study. This is in addition to the (already noted) concern that cisplatin plus gemcitabine is given to a different patient population in the UK than in the Gondo (2011) study. That is, one which is not gemcitabine naive. This does not preclude the need to use the efficacy data reported by Gondo (2011), given the paucity of relevant data, and therefore to adopt the treatment regimen reported by Gondo and colleagues.

F. Please provide justification for why administration costs for cisplatin were incorporated in addition to the gemcitabine administration costs for the cisplatin plus gemcitabine regime (as cisplatin and gemcitabine are both administered on day 1).

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

As noted in in the previous answer the dosing schedule for cisplatin plus gemcitabine is based on information presented in Gondo (2011). Gondo and colleagues report that gemcitabine was given on days 1, 8 and 15 with cisplatin being given on days 1 and 2. This equates to 3 gemcitabine and 2 cisplatin administrations per cycle. However, as both cisplatin and gemcitabine were given on day 1 it is assumed only one administrative cost is required here so for this treatment regimen there are a total of 4 administration episodes per cycle in the model.

B18. TA272 (the only other NICE submission in this indication) was identified in the systematic review for cost-effectiveness evidence.

A. Please provide justification for why TA272 was not used to inform costs and resource use.

TA272 assessed the use of vinflunine versus best supportive care after prior platinum therapy. The resource use information included in the appraisal was informed by a combination of sources including: the pivotal clinical trial Study 302, expert clinical opinion elicited (BMS presumes) in 2010 and data from a published survey of 17 breast cancer specialists. Vinflunine is not recommend for use in the UK and is not available via the Cancer Drugs Fund. Therefore, using resource use data from a trial of vinflunine to estimate UK clinical practice, where vinflunine is not used, was not considered appropriate. To use feedback from breast cancer specialists, when experts in the treatment of bladder cancer have been consulted was also considered suboptimal.

Costs are taken from the National Reference Costs for 2007/08, again which would be inappropriate to use in 2017.

Finally, the assumptions regarding the allocation of costs do not fit with the model structure for nivolumab, as outlined below.

Treatment-related monitoring costs, such as oncologist follow-up visits and CT scanning, are not included in TA272's definition of monitoring costs. In the BMS model treatment duration is estimated separately to progression status for nivolumab. Therefore, it makes sense that monitoring and follow-up costs related to treatment are determined by *treatment status*. TA272 assumes these costs (including oncologist care) are determined by progression

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

status, which would be inconsistent with the BMS modelling approach and lead to inconsistencies (i.e. patients still on treatment but not being monitored by their oncologist).

BSC costs in TA272 are based on expert opinion and an estimate of hospice care from a survey of breast cancer specialists. Again, this is inconsistent with the modelling approach for nivolumab, whereby hospice (or other terminal care) is provided to patients at the end of their life, not upon progression.

BMS have estimated resource use associated with comparator therapies and nivolumab via input from six clinical experts (details provided in the Clinical advisory board minutes in the reference pack). Supplementary information was also provided from nivolumab clinical trial evidence (adverse events, usage of radiotherapy and surgery usage) and comparator clinical trial evidence (adverse events) available.

Costs were taken from the most recent NHS reference costs (2015/16).

B. Please provide explanations for discrepancies with TA272 in terms of monitoring costs in the present assessment that range from £272.44 to £555.50 per 4 weeks while in TA 272 this is £3.18 per treatment cycle of 21 days (see company submission of TA272 Table B37).

See answer to question A.

C. Please provide explanations for discrepancies with TA272 in terms of BSC costs of £170.21 per 4 weeks in the present assessment while in TA 272 this is £580 and £1,253 per month for pre progression and post progression respectively.

See answer to question A.

Model validation

B19. The company states that expert opinion has been used to validate OS and PFS predictions of the model for nivolumab and the comparator. However, the company does not provide any information on the number and identification of experts, or the methods used. Could the company please provide the number of experts that were consulted, the methods used, and questions asked to elicit opinion about OS and PFS predictions for nivolumab and the comparators, and an overview of each expert's opinion/statement.

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

The main validation for the evidence synthesis (i.e. the prognostic model, indirect treatment comparison and economic model) was done via an advisory board including six clinicians and two health economists. The minutes of the advisory board are provided in the reference pack, which includes details of the attendees, the topics covered and the expert advice elicited. Additional clinical uncertainties were compiled before submission and directed to two further clinicians – again the minutes of these discussions are provided in the reference pack.[36]

Subgroup analysis

B20. Referring to Question A21, please provide a subgroup analysis using PFS and OS for patients with PD-L1 < 1% and those with PD-L1 >=1% expression along with the corresponding probabilistic results (expected ICER and cost-effectiveness acceptability curves.

As discussed on the clarification questions call with the ERG and NICE technical team, it was not possible to complete this request in the time frame. This is due to the scope of work required to re-run the simulated treatment comparison, network meta-analysis, survival analysis and revise the economic model. The limitations of this analysis, given the lack of PD-L1 subgroup data for all comparators has already been highlighted in question A13.

Section C: Textual clarifications and additional points

C1. The (blue and green) curves presented in Figures 34 and 35 of the company submission and Figures 32 to 41 of Appendix L are all labelled ‘non-responder’. Please correct the labels such that they correspond to their respective subgroups.

Plots updated with the correct labels (no other changes) are presented in Appendix 2. For clarity, figure numbers have been retained as per the original figure numbers in the NICE submission and as referenced in this question, rather than renumbered to correspond to the figure number in this response document.

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

References:

1. Sharma P, Retz M, Siefker-Radtke A, et al. Nivolumab in metastatic urothelial carcinoma after platinum therapy (CheckMate 275): a multicentre, single-arm, phase 2 trial. Lancet Oncology 2017;25:25.
2. Sharma P, Callahan MK, Bono P, et al. Nivolumab monotherapy in recurrent metastatic urothelial carcinoma (CheckMate 032): a multicentre, open-label, two-stage, multi-arm, phase 1/2 trial. Lancet Oncology 2016;17:1590-1598.
3. Gondo T, Ohori M, Hamada R, et al. The efficacy and safety of gemcitabine plus cisplatin regimen for patients with advanced urothelial carcinoma after failure of M-VAC regimen. International Journal of Clinical Oncology 2011;16:345-51.
4. Joly F, Houede N, Noal S, et al. Do patients with advanced urothelial carcinoma benefit from weekly paclitaxel chemotherapy? A GETUG phase II study. Clinical Genitourinary Cancer 2009;7:E28-33.
5. Ozawa A, Tanji N, Ochi T, et al. Gemcitabine and cisplatin for advanced urothelial carcinomas: the Ehime University Hospital experience. International Journal of Clinical Oncology 2007;12:279-83.
6. Kim YS, Lee SI, Park SH, et al. A phase II study of weekly docetaxel as second-Llne chemotherapy in patients with metastatic urothelial carcinoma. Clinical Genitourinary Cancer 2016;14:76-81.
7. McCaffrey JA, Hilton S, Mazumdar M, et al. Phase II trial of docetaxel in patients with advanced or metastatic transitional-cell carcinoma. Journal of Clinical Oncology 1997;15:1853-7.
8. Vaughn DJ, Broome CM, Hussain M, et al. Phase II trial of weekly paclitaxel in patients with previously treated advanced urothelial cancer. Journal of Clinical Oncology 2002;20:937-40.
9. Clopper CJ, Pearson ES. The Use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial. Biometrika 1934;26:404-413.
10. Brookmeyer R, Crowley J. A Confidence Interval for the Median Survival Time. Biometrics 1982;38:29-41.
11. Bristol-Myers Squibb. CheckMate 275 Addendum: Clinical Study Report for Study CA209275 (25th October 2016).
12. European Medicines Agency. European Public Assessment Report nivolumab. Procedure No. EMEA/H/C/003985/II/0019. Available at http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Assessment_Report_-_Variation/human/003985/WC500231654.pdf [Accessed 28 July 2017]. 2017.
13. Bristol-Myers Squibb. CA209-9GH mUC Report UK, 2016.
14. Bellmunt J, Orsola A, Leow JJ, et al. Bladder cancer: ESMO Practice Guidelines for diagnosis, treatment and follow-up. Annals of Oncology 2014;25:iii40-iii48.
15. European Association of Urology. Muscle-invasive and Metastatic Bladder Cancer Guidelines. Available at: http://uroweb.org/guideline/bladder-cancer-muscle-invasive-and-metastatic/#1 [Accessed 3 Feb 2017], 2016.

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

16. National Institute for Health and Care Excellence (NICE). [NG2]: Bladder cancer: diagnosis and management. Available: https://www.nice.org.uk/guidance/NG2 [Accessed 3 Feb 2017].
17. Bellmunt J, Theodore C, Demkov T, et al. Phase III trial of vinflunine plus best supportive care compared with best supportive care alone after a platinum-containing regimen in patients with advanced transitional cell carcinoma of the urothelial tract. Journal of Clinical Oncology 2009;27:4454-4461.
18. Choueiri TK, Ross RW, Jacobus S, et al. Double-blind, randomized trial of docetaxel plus vandetanib versus docetaxel plus placebo in platinum-pretreated metastatic urothelial cancer. Journal of Clinical Oncology 2012;30:507-12.
19. Jones R, Hussain S, Protheroe A, et al. Randomized Phase II Study Investigating Pazopanib Versus Weekly Paclitaxel in Relapsed or Progressive Urothelial Cancer. Journal of Clinical Oncology 2017;0:JCO.2016.70.7828.
20. Petrylak DP, Tagawa ST, Kohli M, et al. Docetaxel As Monotherapy or Combined With Ramucirumab or Icrucumab in Second-Line Treatment for Locally Advanced or Metastatic Urothelial Carcinoma: An Open-Label, Three-Arm, Randomized Controlled Phase II Trial. Journal of Clinical Oncology 2016;34:1500-9.
21. Joly F, Houede N, Noal S, et al. Do patients with advanced urothelial carcinoma benefit from weekly paclitaxel chemotherapy? A GETUG phase II study. Clinical Genitourinary Cancer 2009;7:E28-E33.
22. Jansen JP. Network meta-analysis of survival data with fractional polynomials. BMC medical research methodology 2011;11:61.
23. Phillippo DM, Ades AE, Dias S, et al. NICE DSU Technical Support Document 18: Methods for population-adjusted indirect comparisons in submissions to NICE. Volume 2016. Sheffield: NICE Decision Support Unit, 2016.
24. Huang Y, Zhang S-D, McCrudden C, et al. The prognostic significance of PD-L1 in bladder cancer. Oncology reports 2015;33:3075-3084.
25. Nakanishi J, Wada Y, Matsumoto K, et al. Overexpression of B7-H1 (PD-L1) significantly associates with tumor grade and postoperative prognosis in human urothelial cancers. Cancer Immunology, Immunotherapy 2007;56:1173-1182.
26. Latimer N. NICE DSU Technical Support Document 14: Undertaking survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data. Sheffield: Decision Support Unit, ScHARR, 2017:52.
27. Othus M, Bansal A, Koepl L, et al. Accounting for cured patients in cost-effectiveness analysis. Value Health 2017;20:705-709.
28. Bristol-Myers Squibb. Meeting minutes, Clinical Advisory Board Meeting: 06 March 2017, 0930-1600 [Not in the public domain]. 30 Euston Square, London, UK: Bristol-Myers Squibb, 2017.
29. Bristol-Myers Squibb Pharmaceuticals Ltd. CheckMate 032: Clinical Study Report for Study CA209032 (29th June 2016). 2016.
30. Bristol-Myers Squibb Pharmaceuticals Ltd. CheckMate 275: Clinical Study Report for Study CA209275 (25th July 2016), 2016.

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

31. Woods B, Sideris E, Palmer S, et al. NICE DSU Technical Support Document 19: Partitioned survival analysis for decision modelling in health care: a critical review. Sheffield: Decision Support Unit, ScHARR, 2017:72.
32. Harrington KJ, Ferris RL, Blumenschein G, Jr., et al. Nivolumab versus standard, single-agent therapy of investigator's choice in recurrent or metastatic squamous cell carcinoma of the head and neck (CheckMate 141): health-related quality-of-life results from a randomised, phase 3 trial. The Lancet Oncology 2017;18:1104-1115.
33. Reck M, Coon C, Taylor F, et al. 3011 Evaluation of overall health status in patients with advanced squamous non-small cell lung cancer treated with nivolumab or docetaxel in CheckMate 017. European Journal of Cancer 2015;51:S599-S600.
34. Reck M, Brahmer J, Bennett B, et al. Overall health status (HS) in patients (pts) with advanced (adv) non-squamous (NSQ) NSCLC treated with nivolumab (nivo) or docetaxel (doc) in CheckMate 057. Annals of Oncology 2016;27:mdw594.006-mdw594.006.
35. Kleinke K. Multiple Imputation Under Violated Distributional Assumptions: A Systematic Evaluation of the Assumed Robustness of Predictive Mean Matching. Journal of Educational and Behavioral Statistics 2017:1076998616687084.
36. Bristol-Myers Squibb. Data on File: Expert Clinician Opinion.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Appendix 1

A number of updates have been made to the original economic model following the clarification requests. During these updates the following minor errors were identified and corrected:

- On the 'Comparator' page (now separated into four pages, one for each individual comparator, due to the fully incremental PSA) there was an error in column V. In this column the equation linked to column ER in the 'PFS & OS' sheet when it should have been column ET.
- On the 'Treatment Costs' and 'Comparator Costs' pages (now separated into four pages, one for each individual comparator, due to the fully incremental PSA) the unit cost of CT scans (column H) was applied at every cycle from cycle 280 (row 292) onwards rather than every other cycle.
- On the 'Treatment Costs' page discontinuation costs (column G) of -£3 were incurred in cycles 416 to 420 (rows 429-433). The value should be £0.
- On the 'Comparator Costs' page (now separated into four separate pages for the fully incremental PSA analysis) adverse event costs were incurred from cycle 281 (row 293) onwards starting at £1 and increasing in value by £1 for each additional cycle. Costs should only be incurred in cycle 1.

As noted previously, when updated there were very minor changes to the overall results of the model. The base case results with this latest version of the model are summarised in Table 30.

Table 30: Base case results – with latest model

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER versus Nivolumab (£/QALY)
Nivolumab	[redacted]	2.78	[redacted]				
Paclitaxel	£14,430	1.19	0.76	[redacted]	1.60	[redacted]	£37,643
Docetaxel	£13,913	1.40	0.92	[redacted]	1.38	[redacted]	£44,996
BSC	£9,052	1.01	0.64	[redacted]	1.77	[redacted]	£38,302

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Appendix 2

Corrected figures from original submission main body

Figure 34: Week 8 landmark - overall survival with generalised gamma

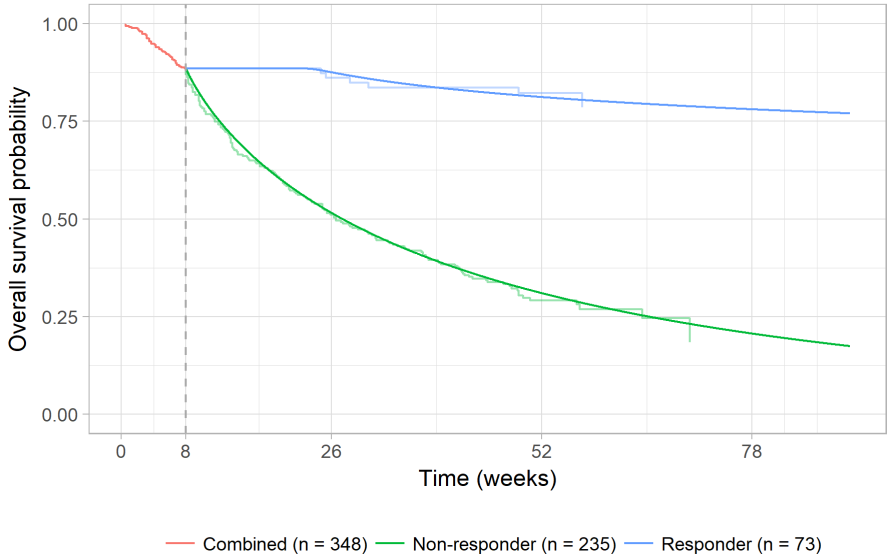
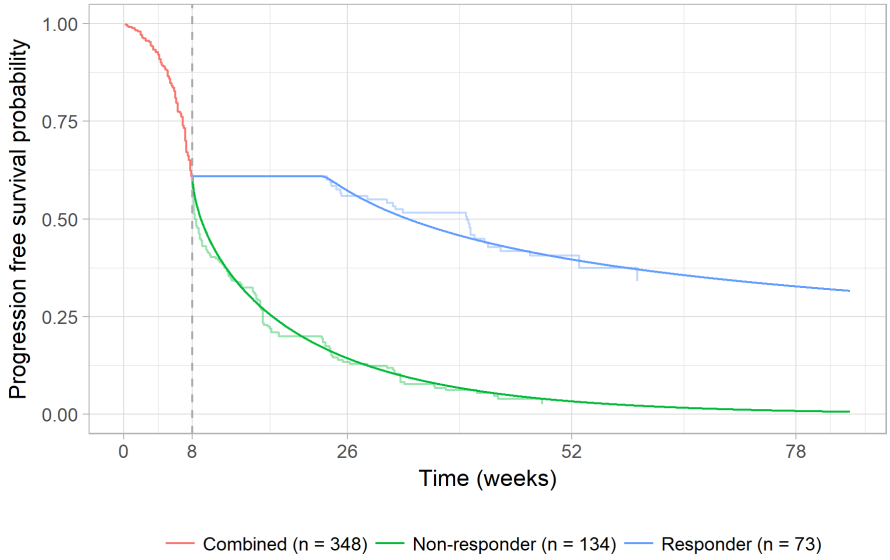


Figure 35: Week 8 landmark – progression-free survival with generalised gamma



Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Corrected figures from original submission Appendix L

Figure 32: Week 8 landmark - overall survival with Weibull survival distribution

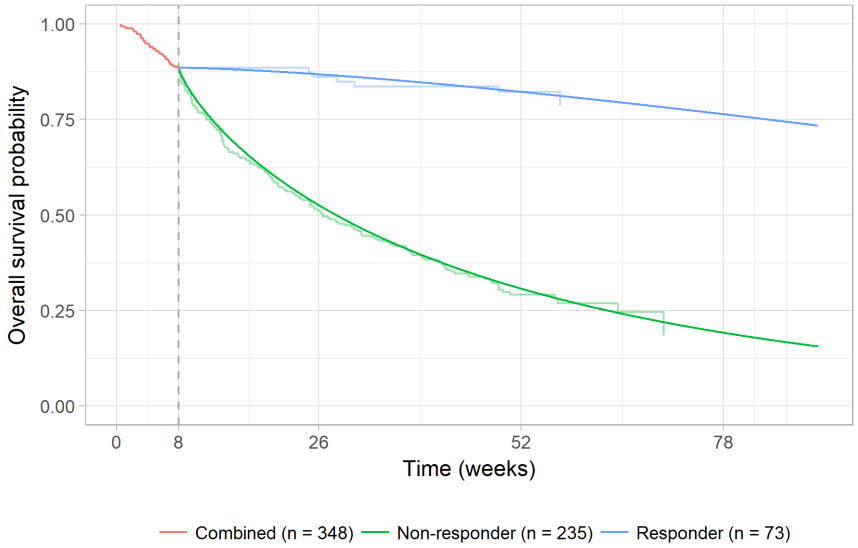
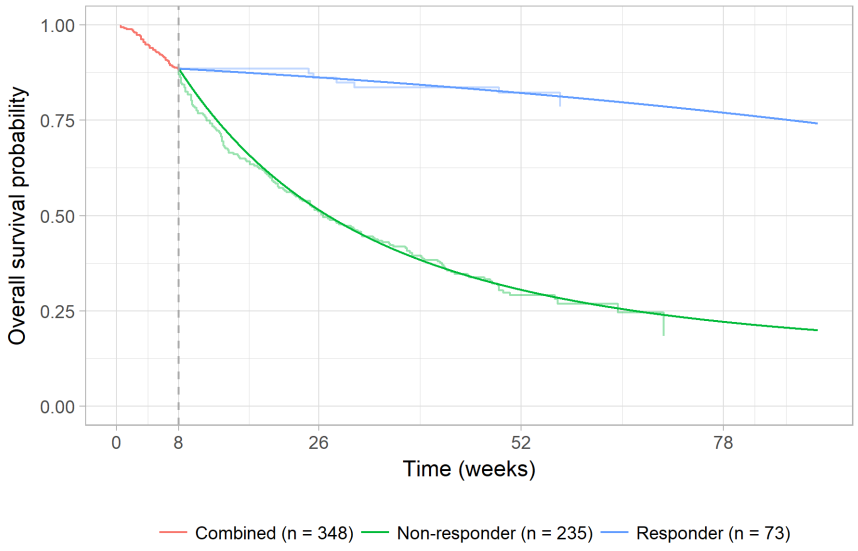


Figure 33: Week 8 landmark - overall survival with Gompertz survival distribution



Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Figure 34: Week 8 landmark - overall survival with lognormal survival distribution

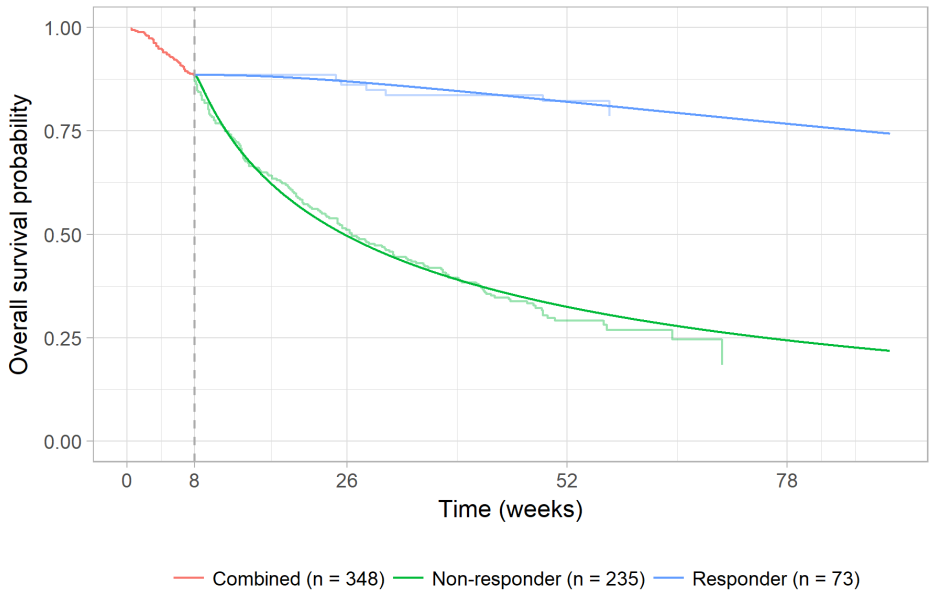
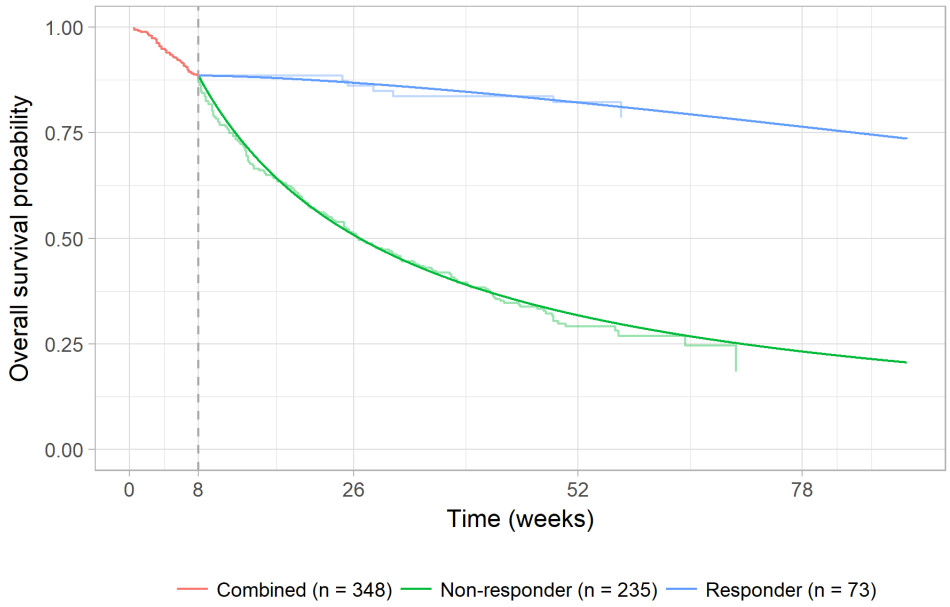


Figure 35: Week 8 landmark - overall survival with log-logistic survival distribution



Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Figure 36: Week 8 landmark - overall survival with exponential survival distribution

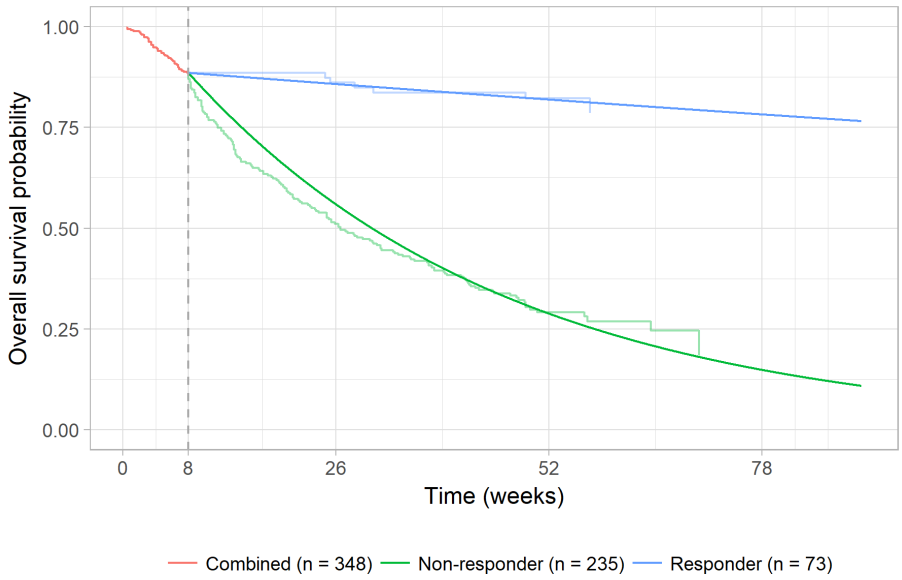
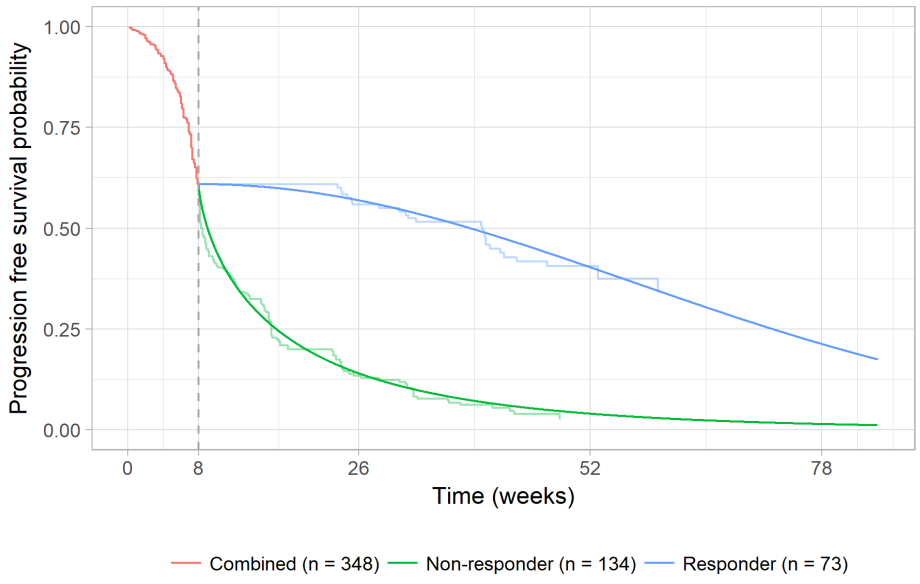


Figure 37: Week 8 landmark – progression-free survival with Weibull survival distribution



Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Figure 38: Week 8 landmark – progression-free survival with Gompertz survival distribution

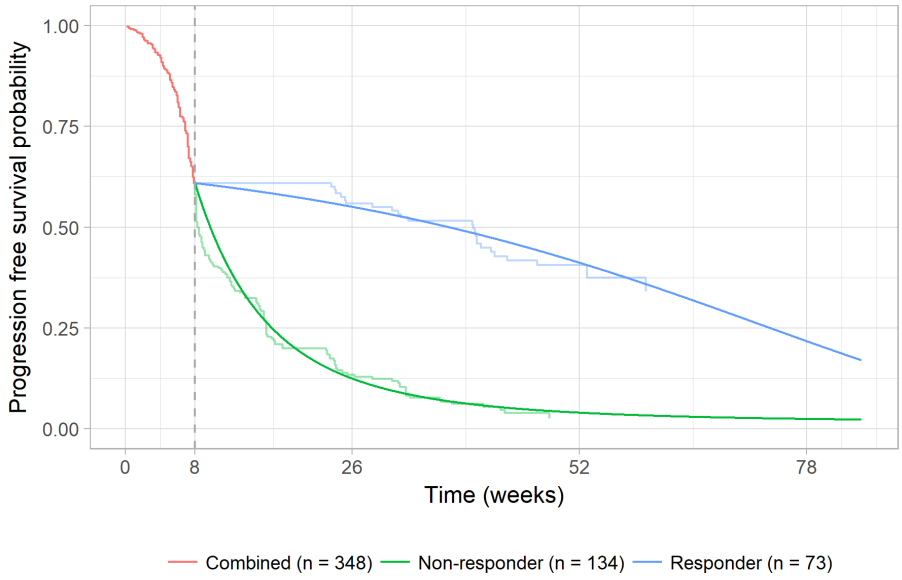
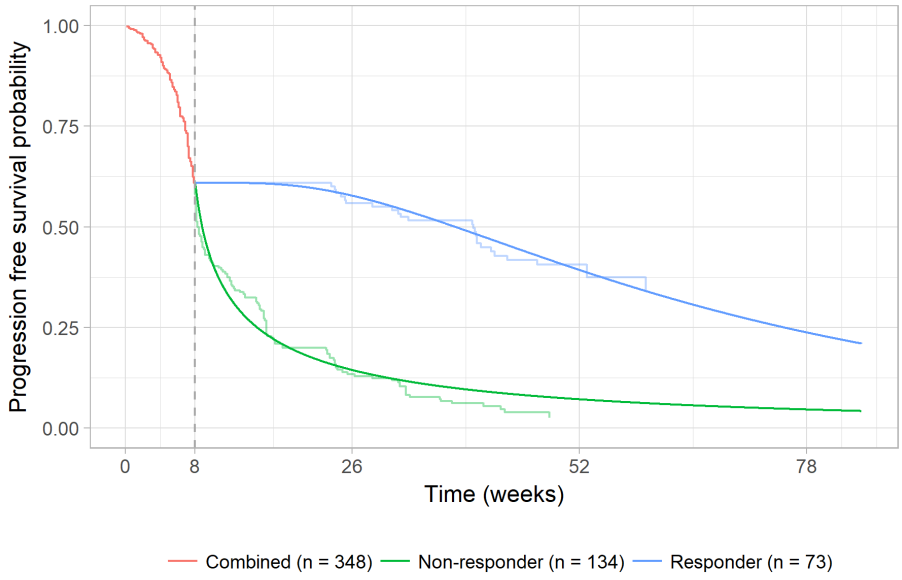


Figure 39: Week 8 landmark – progression-free survival with lognormal survival distribution



Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)300 323 0140

Figure 40: Week 8 landmark – progression-free survival with log-logistic survival distribution

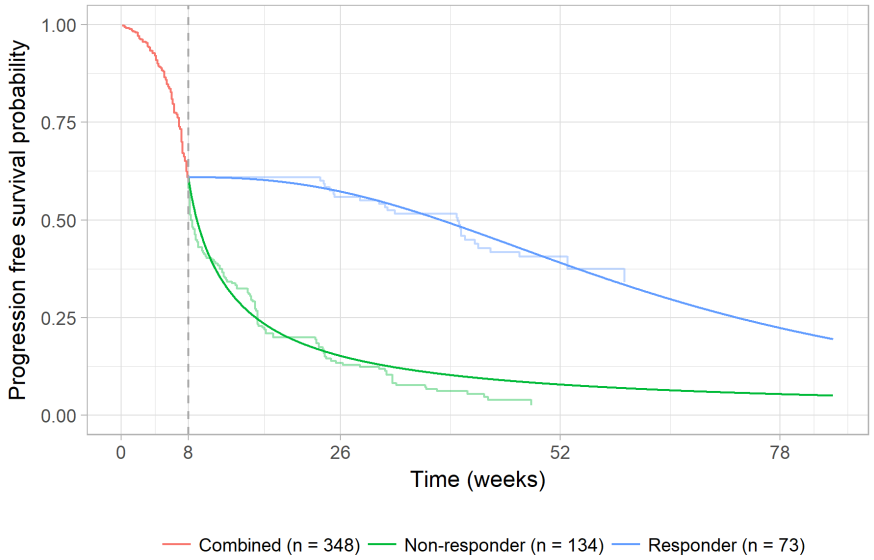
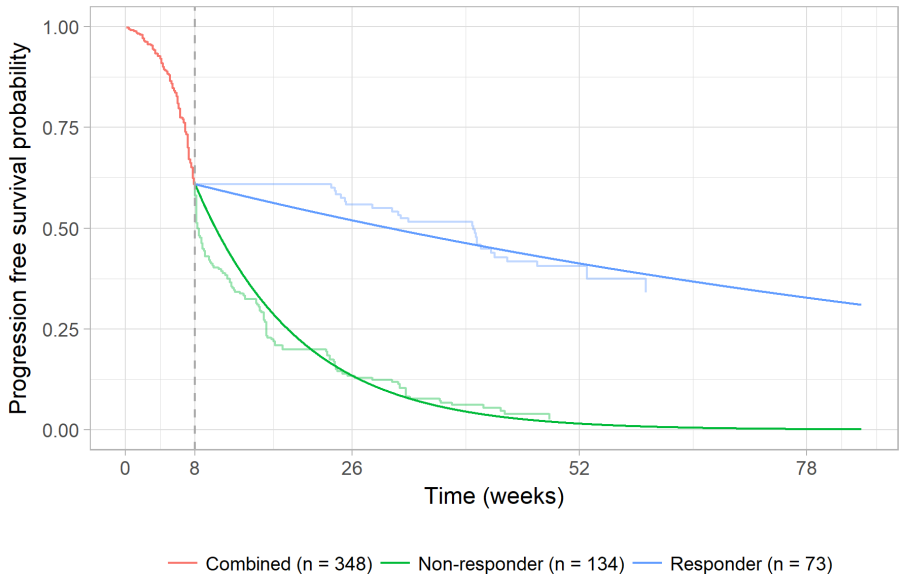


Figure 41: Week 8 landmark – progression-free survival with exponential survival distribution



NICE National Institute for Health and Care Excellence

Patient organisation submission

Single Technology Appraisal (STA)

Nivolumab for treating metastatic or unresectable urothelial cancer after platinum-based chemotherapy [ID995]

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on conditions and their treatment that is not typically available from other sources.

To help you give your views, please use this questionnaire with our guide for patient submissions.

You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 10 pages.

About you

NICE National Institute for Health and Care Excellence

1.Your name	[redacted]
2. Name of organisation	Action Bladder Cancer UK
3. Job title or position	[redacted]
4a. Brief description of the organisation (including who funds it). How many members does it have?	UK Bladder Cancer charity. We have three main strands to our work: • Improving outcomes for bladder cancer patients • Improving research into bladder cancer • Improving patient support We are working to improve outcomes for bladder cancer patients by: • Raising awareness of the signs and symptoms among the public so they seek advice sooner • Improving awareness and investigation techniques among health professionals to improve early diagnosis • Improving the treatment and management of bladder cancer to increase patient survival rates in line with that achieved for other common cancers We are working to improve research into bladder cancer by: • Identifying the key research priorities • Encouraging, contributing to and funding research • Improving research data and statistics We are working to improve patient support through:

- Our high quality information materials and resources library
- Actively increasing the number of bladder cancer patient support groups across the UK
- Providing advice and support to both new and existing groups and helping to bring groups together
- Helping to give bladder cancer patients a voice

Funded by donations, fundraising events and by corporate donations. Our corporate donors are bound by our corporate statement as follows:

CORPORATE STATEMENT Action Bladder Cancer UK is a charity working to support those with bladder cancer and to improve outcomes for patients. We are committed to working in ethical collaboration with commercial and corporate partners in the interest of people affected by bladder cancer. We will accept funding from appropriate corporate and industry supporters. Neither our work, our campaigning nor our information materials will be influenced by accepting any corporate donations or sponsorship. We feel it is important to work with companies that manufacture drugs, treatments or devices which will treat or support bladder cancer patients. We will work in a transparent partnership with appropriate pharmaceutical companies and the medical device industry where these relationships will help promote and improve the interests of bladder cancer patients and fit within the objectives of our charity. We would not accept support from any pharmaceutical or medical industry company for work that we consider to that lie outside the agreed objectives of our charity. We are happy to accept funding, or support in kind, from appropriate corporate supporters outside the health or pharmaceutical sectors. Each corporate collaboration will be assessed and agreed on an individual basis by the charity executive. We are grateful for the support shown by our existing corporate supporters which help us in our work.

	ABC UK has 8 Trustees including a healthy mix of clinicians, urology consultants, cancer nurse specialist, GP with interest in bladder cancer, researchers and patients. We have one employee and outsourced secretariat.
4b. Do you have any direct or indirect links with, or funding from, the tobacco industry?	No
5. How did you gather information about the experiences of patients and carers to include in your submission?	**All our Trustees and staff work closely with patients, both directly and via our network of support groups. In addition, three of our trustees and many of our volunteers and fundraisers are patients or carers. It is absolutely fundamental to our work that we have a deep and current understanding of our patients, their hopes and fears and their treatment options, current and future.**
Living with the condition	
6. What is it like to live with the condition? What do carers experience when caring for someone with the condition?	Awareness is so poor that initial diagnosis is invariably a shock and bc remains a difficult disease to talk about due to general lack of awareness. The fact that recurrence is so high makes it a difficult condition to live with, despite treatment for NMIBC being relatively straightforward and effective. The particular condition for this consultation is the advanced case of metastasised bc where platinum chemotherapy has already been given and where survival rates are known to be exceptionally poor. Therefore the specific condition is very difficult, verging on hopeless, for both patient and carer. This new drug represents an innovative treatment and potential lifeline for patients.

Current treatment of the condition in the NHS	
7. What do patients or carers think of current treatments and care available on the NHS?	Treatment of this specific condition is by platinum based chemotherapy and/or palliative care. These are readily available but response rates and quality of life are very poor. Many patients with metastatic bladder cancer are not suitable for cisplatin and so there is an urgent need for alternatives. The NHS does not have an effective treatment option for this condition.
8. Is there an unmet need for patients with this condition?	Yes. Patients with metastatic bladder cancer have an average life expectancy of only a few months. Many are unable to tolerate the preferred cisplatin chemotherapy. Side effects of cisplatin are severe, even when combined with other drugs, leading to a poor quality of life. About 5,000 patients die each year from this condition, and this has not improved in over 30 years. So there is a huge unmet need and bc patients generally feel overlooked.
Advantages of the technology	
9. What do patients or carers think are the advantages of the technology?	In its simplest form the treatment represents hope to many for whom other treatment options have been exhausted. Therefore the main benefits include: - complete response - prolonging life - improved quality of life for patient, carer, family, friends

	Trials have shown that the treatment does prolong life and for about 20% of patients the effects are enduring. Side effects for the majority of patients are minor and tolerable. The treatment is relatively easy to administer. If the treatment is licenced and similar outcomes to those observed in trials are experienced here, there may be scope to use the treatment at other stages of the disease or as a primary treatment.
Disadvantages of the technology	
10. What do patients or carers think are the disadvantages of the technology?	None. The treatment is widely regarded as an innovative, breakthrough treatment and ABC UK is not aware of any disadvantages perceived by patients or carers. Care would need to be taken to manage patient and carer expectations – the treatment will not save everyone!
Patient population	
11. Are there any groups of patients who might benefit more or less from the technology than others? If so, please describe them and explain why.	The mechanism for this treatment is not known precisely, although there is some improvement in patients that express higher levels of PD1/PDL1. It may be possible to develop biomarkers that could more accurately predict which patients would respond best (or even which may not respond or experience a serious adverse event), leading to a precision medicine. Currently about 5,000 patients die each year in the UK from metastatic bc. All of these could potentially benefit and approximately 20% could be expected to show an enduring and high quality of life benefit.

	By stimulating the body's own immune system, the treatment has also shown great benefit in the group of patients who are not suitable for cisplatin, leading to a first line application for the treatment. It is our hope at ABC UK that the treatment may prove effective earlier in the treatment pathway, for instance instead of BCG to treat HR NMIBC (High Risk Non Muscle Invasive Bladder Cancer). This could avoid the need for cystectomies in a significant minority of patients.
Equality	
12. Are there any potential equality issues that should be taken into account when considering this condition and the technology?	**None known**
Other issues	
13. Are there any other issues that you would like the committee to consider?	Bladder Cancer has had relatively little research and new treatment development in recent decades. Despite it being the 4th most prevalent cancer in men and 7th overall, and very expensive for the NHS to treat, mortality rates of c50% have shown NO improvement in the past 30 years. The mechanism of this new drug is different from anything available to treat BC today, hence the treatment is highly innovative. ABC UK supports the licencing and use of the treatment within the NHS. Ideally more research could be commissioned to optimise the treatment regimen and to better understand the mechanism of treatment,

	ultimately leading to biomarkers to identify patients for whom the treatment would be most effective/ineffective. It would also be useful for patients to contribute to the 'Life and Bladder Cancer' PROMS (Patient Reported Outcome Measures Study), being run by Leeds/Sheffield.

Key messages

14. In up to 5 bullet points, please summarise the key messages of your submission:

- ABC UK supports the licencing and use of the treatment within the NHS
- The treatment is highly innovative
- The treatment gives hope to many for whom other treatment options have been exhausted
- Further research/trials to optimise the treatment and develop biomarkers would be highly desirable
- Consideration should be given for research/trails for use of the treatment earlier in the disease progress and/or as a primary treatment

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

NICE National Institute for Health and Care Excellence

Patient organisation submission

Single Technology Appraisal (STA)

Nivolumab for treating metastatic or unresectable urothelial cancer after platinum-based chemotherapy [ID995]

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on conditions and their treatment that is not typically available from other sources.

To help you give your views, please use this questionnaire with our guide for patient submissions.

You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 10 pages.

About you

1.Your name	[redacted]
2. Name of organisation	FIGHT BLADDER CANCER
3. Job title or position	[redacted]
4a. Brief description of the organisation (including who funds it). How many members does it have?	Patient advocacy group and charity for bladder cancer. UK based. 5000+ supporters, majority are patients or carers 90%+ of annual income is from individual donations and fundraising
4b. Do you have any direct or indirect links with, or funding from, the tobacco industry?	None
5. How did you gather information about the experiences of patients and carers to include in your submission?	We have a confidential online forum of 3000+ patients and carers, offer telephone and email support, a 1 to 1 peer support service and conduct regular surveys.

Living with the condition	
6. What is it like to live with the condition? What do carers experience when caring for someone with the condition?	Bladder cancer (metastatic or unresectable urothelial cancer) has a very poor prognosis and mostly results in a continuous round of treatments between diagnosis and death. The current treatments can often have quite serious side effects that significantly reduce the quality of life for the final months. For carers, it is a period of ultimate worry and exhaustion as you care for your loved one as the patient and their medical team fight to preserve life for as long as possible.
Current treatment of the condition in the NHS	
7. What do patients or carers think of current treatments and care available on the NHS?	Current treatments are very limited and have great limitations in extending life for these patients. The knowledge that there has been no new treatments in this cancer for over 35 years adds to the feeling that a diagnosis with this cancer isn't cared about as much as other cancers.
8. Is there an unmet need for patients with this condition?	There is a significant unmet need for patients with this condition. If a platinum based treatment is ineffective there is currently no second line treatment available.
Advantages of the technology	
9. What do patients or carers think are the advantages of the technology?	There is a great amount of interest amongst the bladder cancer community about the new immunotherapy oncology treatments and we feel that, at last, there might be a light at the end of the tunnel. Ideally we will see an improvement in prognosis but it is also important to see if there are improvements in QoL with or without improvements in prognosis.

Disadvantages of the technology	
10. What do patients or carers think are the disadvantages of the technology?	Too early to know this.
Patient population	
11. Are there any groups of patients who might benefit more or less from the technology than others? If so, please describe them and explain why.	Too early to know this but it appears that certain patients benefit more than others. Why and how we predict this will be a key challenge.
Equality	
12. Are there any potential equality issues that should be taken into account when considering this condition and the technology?	None known

Other issues	
13. Are there any other issues that you would like the committee to consider?	None

Key messages
14. In up to 5 bullet points, please summarise the key messages of your submission: • Current prognosis for these patients is very poor • Current first line treatments have significant side effects • Quality of Life is an essential part of this evaluation • There is a clear unmet need • For a cancer with so few advances in decades, this gives hope to many.

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

NICE National Institute for Health and Care Excellence

Clinical expert statement

Nivolumab for treating metastatic or unresectable urothelial cancer after platinum-based chemotherapy [ID995]

Thank you for agreeing to give us your views on this technology and its possible use in the NHS.

You can provide a unique perspective on the technology in the context of current clinical practice that is not typically available from the published literature.

To help you give your views, please use this questionnaire. You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this expert statement

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 13 pages.

About you	
1. Your name	**Dr Simon Crabb**
2. Name of organisation	**University of Southampton**

3. Job title or position	**Associate Professor in Medical Oncology**
4. Are you (please tick all that apply):	☐ an employee or representative of a healthcare professional organisation that represents clinicians? ☒ a specialist in the treatment of people with this condition? ☒ a specialist in the clinical evidence base for this condition or technology? ☐ other (please specify):
5. Do you wish to agree with your nominating organisation's submission? (We would encourage you to complete this form even if you agree with your nominating organisation's submission)	☐ yes, I agree with it ☐ no, I disagree with it ☐ I agree with some of it, but disagree with some of it ☐ other (they didn't submit one, I don't know if they submitted one etc.)
6. If you wrote the organisation submission and/ or do not have anything to add, tick here. (If you tick this box, the rest of this form will be deleted after submission.)	☐ yes

The aim of treatment for this condition	
7. What is the main aim of treatment? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability.)	To extend survival (in a group of patients with incurable disease). To relieve symptoms and improve or maintain quality of life. To extend the time until subsequent cancer progression.
8. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by x cm, or a reduction in disease activity by a certain amount.)	Stabilisation or reduction of cancer volume (normally based on cross sectional imaging) for a period of time of 3 months or longer. Relief of disease related symptoms.
9. In your view, is there an unmet need for patients and healthcare professionals in this condition?	Yes. Outcomes for this patient group are currently dire and we desperately need improved treatment options. Without treatment, patients with progressive disease after prior platinum based chemotherapy will deteriorate and die within a median of 4 to 5 months (Bellmunt et al, J Clin Oncol, 2009;27:4454-61).

	Progression free survival in clinical trials of various systemic therapies over the last 20 years has been in the order of 2-4 months. Until the last 12 months there had never been an intervention that had improved overall survival outcomes for this group of patients. Immunotherapy is the first approach to do so and the first new treatment approach in decades.
What is the expected place of the technology in current practice?	
10. How is the condition currently treated in the NHS?	Patients with metastatic or unresectable urothelial cancer are treated with platinum based chemotherapy in the first line setting if they are fit enough to do so. In approximately half of patients who receive chemotherapy this would be cisplatin based (combined with gemcitabine or the MVAC regimen combined with methothrexate, doxorubicin and vinblastine). In the remainder, who are unable to receive cisplatin, an alternative of carboplatin and gemcitabine is normally used. In the second line setting that this appraisal address, most clinicians have used chemotherapy in those fit to receive it. The commonest regimen in the UK is single agent paclitaxel. Other regimens that have been used in the UK include carboplatin/paclitaxel, gemcitabine/paclitaxel, vinflunine and MVAC. Many centres have made use of expanded access programs in the last year to deliver PD1/PD-L1 inhibitors to this patient group instead of chemotherapy and most clinicians would now view it as the appropriate standard of care.

• Are any clinical guidelines used in the treatment of the condition, and if so, which?	NICE guidelines: Bladder cancer: diagnosis and management (https://www.nice.org.uk/guidance/ng2) The Association of Cancer Physicians: Multi-disciplinary Team (MDT) Guidance for Managing Bladder Cancer (http://www.bug.uk.com/publications.php) European Association of Urology: Oncology Guidelines - Muscle-invasive and Metastatic Bladder Cancer (http://uroweb.org/guideline/bladder-cancer-muscle-invasive-and-metastatic/#7_8)
• Is the pathway of care well defined? Does it vary or are there differences of opinion between professionals across the NHS? (Please state if your experience is from outside England.)	Generally there is agreement and consistency on the treatment pathway for this group of patients with limited variation in the UK or internationally.
• What impact would the technology have on the current pathway of care?	There would be an almost universal switch from chemotherapy (paclitaxel in most cases) to nivolumab as the standard second line treatment option if it were available for use.
11. Will the technology be used (or is it already used) in the same way as current care in NHS clinical practice?	As described above, nivolumab would be used as the standard second line treatment option in place of chemotherapy. We don't currently have routine access to immunotherapy in this treatment setting. However immunotherapy has been used by many centres through expanded access programs or clinical trials.

• How does healthcare resource use differ between the technology and current care?	Patients who are suitable to receive either chemotherapy or nivolumab are not significantly different and so the impact on the number of patients treated should be minimal. Nivolumab is given to the point of disease progression whereas paclitaxel is a usually a fixed duration course of treatment. There might be a modest increase in the duration of treatment therefore but this would probably impact mostly on a minority who had prolonged benefit from nivolumab.
• In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.)	Second line treatment of metastatic or unresectable urothelial cancer. This will be delivered within specialist oncology clinics for urothelial cancer in secondary care. Patients will normally already be known to these services having (by definition) already received first line chemotherapy delivered by the same clinical team.
• What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	Very minimal as these patients would otherwise be receiving IV chemotherapy delivered by the same teams who will are already using nivolumab and other immunotherapies for other cancers. As a result the infrastructure to administer nivolumab is established and active. Although there is a different pattern of toxicity to chemotherapy it is unlikely to make a significant difference in the demands on the oncology teams and institutions that would administer this treatment.
12. Do you expect the technology to provide clinically meaningful benefits compared with current care?	

• Do you expect the technology to increase length of life more than current care?	Yes in the sense that this has been demonstrated for pembrolizumab and I expect the same benefit for nivolumab. However we have yet to see a randomised trial to test this formally in this setting for nivolumab.
• Do you expect the technology to increase health-related quality of life more than current care?	Yes based on the duration of response and QOL data in the CheckMate 275 trial.
13. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	Not that we have strong evidence for and I would not restrict it to a particular subset of patients. There has been interest, with this and other similar agents, regarding the use of immunohistochemical (e.g., PD-L1 expression) or more complex predictive biomarkers. Although these may enrich for patient subsets with a higher response rate or duration of response these have not been shown to be predictive for treatment benefit separately from their potential prognostic impact. The appropriate approach, based on current data, is therefore to treat all comers, allowing for common sense assessment of patient fitness and co-morbidities. The CheckMate 275 eligibility criteria are a reasonable summary of suitability for treatment in terms of diagnosis, prior treatment, fitness and blood result parameters.
The use of the technology	
14. Will the technology be easier or more difficult to use for patients or healthcare	Roughly the same compared to chemotherapy. They are both intravenous drugs that require specialist administration in a dedicated oncology clinic for urothelial cancers. They (paclitaxel and nivolumab) are

professionals than current care? Are there any practical implications for its use (for example, any concomitant treatments needed, additional clinical requirements, factors affecting patient acceptability or ease of use or additional tests or monitoring needed.)	both fairly well tolerated but with side effects that can occasionally be serious and very occasionally require admission. As described elsewhere, the infrastructure and personnel are established.
15. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	Generally we will start and stop at the point of objective disease progression based on cross sectional imaging. A small minority of patients will discontinue for overt clinical progression or for toxicity/intolerance. These considerations are all also the case for chemotherapy that is currently used.
16. Do you consider that the use of the technology will result in any substantial health-related benefits that are unlikely to be included in the	No.

quality-adjusted life year (QALY) calculation?	
17. Do you consider the technology to be innovative in its potential to make a significant and substantial impact on health-related benefits and how might it improve the way that current need is met?	
• Is the technology a 'step-change' in the management of the condition?	Yes. Immunotherapy is the first new approach to treatment for this disease in decades and the only alternative to chemotherapy.
• Does the use of the technology address any particular unmet need of the patient population?	It addresses the fact that without treatment, patients die on average in 4.5 months and chemotherapy has modest efficacy in this setting. A notable proportion of patients derive prolonged disease control with maintenance of quality of life in response to nivolumab.
18. How do any side effects or adverse effects of the technology affect the	This is usually well tolerated and on average a little better than with chemotherapy. However a minority will experience significant treatment related toxicity (18% grade 3 or 4 adverse event rate in the Checkmate 275 trial of which diarrhoea and fatigue were most common). Toxicity can commonly be addressed with

management of the condition and the patient's quality of life?	dose delay and use of corticosteroids. Treatment discontinuation is required occasionally but most patients will tolerate treatment to allow it to continue until the point of disease progression.
Sources of evidence	
19. Do the clinical trials on the technology reflect current UK clinical practice?	Yes. The Checkmate 275 trial, and other trials of similar agents, have reflected UK practice in that they defined a group of 'second line' patients, following prior use of platinum based chemotherapy, that we would previously have given chemotherapy to. In this sense UK clinicians would feel able to directly extrapolate these results to their current practice and would switch directly to nivolumab from chemotherapy if it were available for use.
• If not, how could the results be extrapolated to the UK setting?	
• What, in your view, are the most important outcomes, and were they measured in the trials?	Overall survival Response duration Quality of life Each of these were assessed in the CheckMate 275 trial.
• If surrogate outcome measures were used, do they adequately predict	N/A

long-term clinical outcomes?	
• Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	No. Our use of these agents in trials, expanded access programs and in other cancers mirrors what has been reported for nivolumab and similar agents in the trials for urothelial cancer.
20. Are you aware of any relevant evidence that might not be found by a systematic review of the trial evidence?	No.
21. How do data on real-world experience compare with the trial data?	We have not seen much 'real world' data for immunotherapy in bladder cancer yet. What there has been, and my own experience, has been fully consistent with trial data.
Equality	
22a. Are there any potential equality issues that should be taken into account when considering this treatment?	None specifically.

22b. Consider whether these issues are different from issues with current care and why.	Not different to current care.

Key messages

25. In up to 5 bullet points, please summarise the key messages of your statement.

- Outcomes in this clinical setting are dire.
- Immunotherapy with agents such as nivolumab provide an entirely novel approach to treatment of urothelial cancer.
- These drugs are efficacious, generally well tolerated, and provide a proportion of patients with long term durable responses.
- UK clinicians now consider immunotherapy to be the appropriate standard of care for second line treatment of urothelial cancer and would fully support introduction of nivolumab.
- The trial data to support the use of nivolumab in urothelial cancer is fully consistent with UK practice, and with our experience of using these agents in clinical trials, expanded access programs and in other cancers.

Thank you for your time.

Please log in to your NICE Docs account to upload your completed statement, declaration of interest form and consent form.

NICE National Institute for Health and Care Excellence

Clinical expert statement

Nivolumab for treating metastatic or unresectable urothelial cancer after platinum-based chemotherapy [ID995]

Thank you for agreeing to give us your views on this technology and its possible use in the NHS.

You can provide a unique perspective on the technology in the context of current clinical practice that is not typically available from the published literature.

To help you give your views, please use this questionnaire. You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this expert statement

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 13 pages.

About you	
1. Your name	**Dr Yvonne L Rimmer**
2. Name of organisation	**Cambridge University Hospital** **On behalf of British Uro-Oncology Group**

3. Job title or position	**Consultant Clinical Oncologist**
4. Are you (please tick all that apply):	☐ an employee or representative of a healthcare professional organisation that represents clinicians? **x** a specialist in the treatment of people with this condition? ☐ a specialist in the clinical evidence base for this condition or technology? ☐ other (please specify):
5. Do you wish to agree with your nominating organisation's submission? (We would encourage you to complete this form even if you agree with your nominating organisation's submission)	☐ yes, I agree with it ☐ no, I disagree with it ☐ I agree with some of it, but disagree with some of it **x** other (they didn't submit one, I don't know if they submitted one etc.)
6. If you wrote the organisation submission and/ or do not have anything to add, tick here. (If you tick this box, the rest of this form will be deleted after submission.)	☐ yes

The aim of treatment for this condition	
7. What is the main aim of treatment? (For example, to stop progression, to improve mobility, to cure the condition, or prevent progression or disability.)	To palliate symptoms, improve quality of life and delay time to further progression of disease and improve survival.
8. What do you consider a clinically significant treatment response? (For example, a reduction in tumour size by x cm, or a reduction in disease activity by a certain amount.)	Improvement of symptoms with acceptable toxicity and so quality of life from the technology.
9. In your view, is there an unmet need for patients and healthcare professionals in this condition?	Yes, following first line treatment (NICE guidelines recommend Platinum based chemotherapy) there is no consensus on subsequent therapies.
What is the expected place of the technology in current practice?	

10. How is the condition currently treated in the NHS?	
• Are any clinical guidelines used in the treatment of the condition, and if so, which?	NICE guidance on bladder cancer (NG2 2015) recommend sequencing chemotherapy
• Is the pathway of care well defined? Does it vary or are there differences of opinion between professionals across the NHS? (Please state if your experience is from outside England.)	In the first line metastatic setting the pathway would be standard with little variation from NICE recommendations. If a patient is fit (performance status 0-1) with good renal function (EGFR> 60mls/min) then cisplatin based combination chemotherapy is used. If renal function is poorer or other co-morbidities then consideration is given for the use of carboplatin combination chemotherapy. Approximately 40-50% of patients respond, although response durations are short and disease progresses in most with a median survival of approximately 15 months. In the second line setting there is less consensus amongst professionals and more variation in NHS practice. NICE guidelines suggest options of either re-challenge with platinum or alternative chemotherapy, however response rate in the second line setting is only around 10% and hence discussions with patients regarding their balance of quality of life from toxicities of treatment with chance of response can be challenging and many patients decline.
• What impact would the technology have on the current pathway of care?	The immunotherapies would be used instead of second line chemotherapy.
11. Will the technology be used (or is it already used) in	

the same way as current care in NHS clinical practice?	
• How does healthcare resource use differ between the technology and current care?	The trials with immunotherapy have shown benefit in patients with good performance status 0-1, and it would be in this group of patients that I would consider second line chemotherapy hence if immunotherapy was available then patient numbers may be similar. The infusions are given 2 weekly. Current second line schedules vary between 1-3 weekly.
• In what clinical setting should the technology be used? (For example, primary or secondary care, specialist clinics.)	In the same setting as systemic chemotherapy within specialist clinics in secondary care.
• What investment is needed to introduce the technology? (For example, for facilities, equipment, or training.)	Facilities and equipment would all be in place in larger centres that already have experience of using immunotherapies in other tumour types (renal, melanoma and lung). Education and training of clinicians and nurses who are not familiar with these drugs would be required as the side effects and advice offered is different than systemic chemotherapy.
12. Do you expect the technology to provide clinically meaningful benefits compared with current care?	
• Do you expect the technology to increase	Yes, studies have suggested a favourable median overall survival using Nivolumab (in Single arm Phase 2 trial) compared to a meta-analysis of single agent second line chemotherapy studies.

length of life more than current care?	
• Do you expect the technology to increase health-related quality of life more than current care?	Yes, with reported objective response rates double or more that of standard single agent chemotherapy to improve symptoms from metastatic bladder cancer including pain and haematuria I would expect to increase health-related quality of life more than current care.
13. Are there any groups of people for whom the technology would be more or less effective (or appropriate) than the general population?	PD-L1 expression was assessed retrospectively in studies, whilst there are objective responses across all PDL1 subgroups compared to standard treatment, it is higher where PDL1- expression of 5% or greater occurred.
The use of the technology	
14. Will the technology be easier or more difficult to use for patients or healthcare professionals than current care? Are there any practical implications for its use (for example, any concomitant	I do not think immunotherapy is more difficult to use for professionals or patients when compared with systemic chemotherapy- but as noted earlier side effect profiles do differ.

treatments needed, additional clinical requirements, factors affecting patient acceptability or ease of use or additional tests or monitoring needed.)	
15. Will any rules (informal or formal) be used to start or stop treatment with the technology? Do these include any additional testing?	Rules for stopping as with systemic chemotherapy in the metastatic patient are defined firstly by patient tolerability and side effects. If the treatment is being tolerated the benefit clinically assessed regarding their symptoms and used in combination with radiological imaging.
16. Do you consider that the use of the technology will result in any substantial health-related benefits that are unlikely to be included in the quality-adjusted life year (QALY) calculation?	
17. Do you consider the technology to be innovative in	Yes, metastatic bladder cancer has unfortunately had very few breakthroughs or changes of treatment in the past 20 years. The metastatic bladder cancer population in the majority can be more elderly with a

its potential to make a significant and substantial impact on health-related benefits and how might it improve the way that current need is met?	number of co-morbidities, however for fitter patients who are suitable for second line treatment there are very few options and immunotherapy could provide this.
• Is the technology a 'step-change' in the management of the condition?	It could be if initially used in this setting and then considered earlier as well within the radical bladder cancer treatment options to try and improve survival.
• Does the use of the technology address any particular unmet need of the patient population?	Current assessment in metastatic disease in second line setting- useful to consider the in first line setting or in combination with chemotherapy in the radical treatment setting.
18. How do any side effects or adverse effects of the technology affect the management of the condition and the patient's quality of life?	The side effect profile has been found to be acceptable, in study 5% of patients discontinued treatment due to toxicity.
Sources of evidence	

19. Do the clinical trials on the technology reflect current UK clinical practice?	There is no use of Nivolumab outside of trials in metastatic bladder cancer.
• If not, how could the results be extrapolated to the UK setting?	The trial is a single arm phase 2 study and hence does not directly compare with second line chemotherapy and so is limited by comparisons being with historical data.
• What, in your view, are the most important outcomes, and were they measured in the trials?	Response rates for patients, and survival were measured. Quality of life from patient data has not been reported at this time. The length of follow up at the time of publication was only 6 months and hence long – term outcomes are awaited.
• If surrogate outcome measures were used, do they adequately predict long-term clinical outcomes?	
• Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently?	Not that I am aware of
20. Are you aware of any relevant evidence that might	

not be found by a systematic review of the trial evidence?	
21. How do data on real-world experience compare with the trial data?	Real world experience within our own department on the use of immunotherapies with other tumour groups supports their use but in patients with good performance status – they would not be recommended in patients who were felt to be 'unfit for chemotherapy'.
Equality	
22a. Are there any potential equality issues that should be taken into account when considering this treatment?	No
22b. Consider whether these issues are different from issues with current care and why.	
Key messages	

25. In up to 5 bullet points, please summarise the key messages of your statement.

- Favourable patient outcomes
- Manageable toxicity
- For use in good performance status patients
- Significant unmet clinical need
-

Thank you for your time.

Please log in to your NICE Docs account to upload your completed statement, declaration of interest form and consent form.

in collaboration with:

ERASMUS UNIVERSITEIT ROTTERDAM
INSTITUTE OF HEALTH POLICY & MANAGEMENT

Maastricht University

Nivolumab for treating metastatic or unresectable urothelial cancer

Produced by	Kleijnen Systematic Reviews Ltd. in collaboration with Erasmus University Rotterdam (EUR) and Maastricht University
Authors	Nigel Armstrong, Health Economist, Kleijnen Systematic Reviews Ltd, UK Sabine Grimm, Health Economist, Maastricht UMC Bram Ramaekers, Health Economist, Maastricht UMC Xavier Pouwels, Health Economist, Maastricht UMC Shona Lang, Systematic Reviewer, KSR Ltd Debra Fayter, Systematic Reviewer, KSR Ltd Svenja Petersohn, Health Economist, Maastricht UMC Rob Riemsma, Reviews Manager, KSR Ltd Gill Worthy, Statistician, KSR Ltd Lisa Stirk, Information Specialist, KSR Ltd Janine Ross, Information Specialist, KSR Ltd Manuela Joore, Health Economist, 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	30/08/2017

Source of funding: This report was commissioned by the NIHR HTA Programme as project number STA 16/108/11.

Declared competing interests of the authors

None.

Acknowledgements

None.

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

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

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, Grimm S, Ramaekers BLT, Pouwels X, Lang S, Fayter D, Petersohn S, Riemsma R, Worthy G, Stirk S, Ross J, Joore MA, Kleijnen J. Nivolumab for treating metastatic or unresectable urothelial cancer: a Single Technology Assessment. York: Kleijnen Systematic Reviews Ltd, 2017.

Contributions of authors

Nigel Armstrong acted as project lead and systematic reviewer on this assessment, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Sabine Grimm acted as health economic project lead, critiqued the company's economic evaluation and contributed to the writing of the report. Bram Ramaekers, Xavier Pouwels and Svenja Petersohn acted as health economists on this assessment, critiqued the company's economic evaluation and contributed to the writing of the report. Debra Fayter, Shona Lang and Rob Riemsma acted as systematic reviewers, 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. Lisa Stirk and Janine Ross critiqued the search methods in the submission and contributed to the writing of the report. Manuela Joore acted as health economist 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

Ab	Antibody
AE	Adverse Events
AIC	Akaike information criterion
ALT	Alanine transaminase
BI	Budget impact
BIC	Bayesian information criterion
BIRC	Blinded independent review committee
BNF	British National Formulary
BOR	best overall response
BSA	body surface area
BSC	Best supportive care
CDF	Cancer Drugs Fund
CD28	Cluster of differentiation 28
CE	Cost Effectiveness
CEA	Cost-effectiveness Analysis
CEAC	Cost effectiveness Acceptability Curve
CHMP	Committee for Medicinal Products for Human Use
CI	Confidence Interval
CIS	Carcinoma in situ
Cis	Cisplatin
CR	Complete response
CRD	Centre for Reviews and Dissemination
CrI	Credible interval
CS	Company's submission
CSR	Clinical study report
CHMP	Committee for Medicinal Products for Human Use
CT	Computer tomography
CTCAE	Common Terminology Criteria for Adverse Events (NCI)
CTLA-4	Cytotoxic T-lymphocyte-associated protein 4
"D" ¯ "res"	Residual deviance
DIC ¯	Deviance information criteria
DOR	Duration of response
DSU	Decision Support Unit
EAU	European Association of Urology
ECOG	Eastern Cooperative Oncology Group
EMA	European Medicines Agency
eMIT	Electronic market information tool
EORTC	European Organisation for Research and Treatment of Cancer
EORTC QLQ-C30	European Organisation for Research and Treatment of Cancer quality of life questionnaire
EPAR	European public assessment report
EQ-5D	European Quality of Life-5 Dimensions
EQ-5D-3L	European Quality of Life-5 Dimensions, three-level scale
ERG	Evidence Review Group
ESMO	European Society for Medical Oncology
EUR	Erasmus University Rotterdam
FDA	Food and Drug Administration
FP	Fractional polynomial
G-CSF	Granulocyte colony stimulating factor
GCP	Good Clinical Practice
Gem	Gemcitabine
GFR	Glomerular filtration rate
GP	General practitioner

HR	Hazard ratio
HRQoL	Health-related Quality of Life
HTA	Health Technology Assessment
IC	Indirect Comparison
ICD	International Classification of Diseases
ICER	Incremental Cost Effectiveness Ratio
IFNγR	Interferon gamma receptor
IPD	Individual patient data
IQR	Interquartile range
ITC	Indirect treatment comparison
ITT	Intention to Treat
IV	Intravenous
KM	Kaplan–Meier
KSR	Kleijnen Systematic Reviews
LPFT	Last patient first treatment
LYG	Life years gained
LYS	Life Year Saved
LYs	Life years
MedDRA	Medical Dictionary for Regulatory Activities
MeSH	Medical Subject Headings
MHC	Major histocompatibility complex
MHRA	Medicines and Healthcare Products Regulatory Agency
MICE	Multiple imputation by chained equations
MRI	Magnetic resonance imaging
MVAC	Methotrexate, vinblastine, doxorubicin and cisplatin
NA	Not applicable
NCI	National Cancer Institute
NF-κB	Nuclear transcription factor-κB
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
NMA	Network meta-analysis
NMB	Net monetary benefit
NR	Not Reached/Not Reported
NSCLC	Non-small cell lung cancer
OR	Odds Ratio
ORR	Objective response rate
OS	Overall survival
PAS	Patient access scheme
PD	Progressive disease
pD	Number of effective parameters
PD-1	Programmed death 1
PD-L1	Programmed death ligand 1
PD-L2	Programmed death ligand 2
PFS	Progression-free survival
PH	Proportional hazards
PP	Post-progression
PR	Partial response
PRESS	Peer Review of Electronic Search Strategies
PROs	Patient-reported outcomes
PS	Performance status
PSA	Probabilistic Sensitivity Analyses
PSS	Personal Social Services
PSSRU HCHS	Personal and Social Services Research Unit Hospital and Community Health Services

PI3K	Phosphoinositide 3-kinase
Q2W	Every 2 weeks
Q3W	Every 3 weeks
QALY(s)	Quality-adjusted Life Year(s)
QoL	Quality of life
RCT	Randomised Controlled Trial
RECIST	Response Evaluation Criteria In Solid Tumours
SAE	Serious Adverse Events
SD	Stable disease/Standard deviation
SE	Standard error
Shp-2	Src homology 2 domain-containing protein tyrosine phosphatase 2
SLR	Systematic literature review
SmPC	Summary of product characteristics
STA	Single Technology Appraisal
STC	Simulated treatment comparison
TEAEs	Treatment-emergent adverse events
TESAEs	Treatment-emergent serious adverse events
TNM	Tumour-node-metastasis
TTD	Time to treatment discontinuation
TTR	Time to response
TURBT	Transurethral resection of the bladder tumour
UC	Urothelial carcinoma
UICC	Union for International Cancer Control
UK	United Kingdom
UMC	University Medical Centre
USA	United States of America
VAS	Visual analogue scale
VEGF	Vascular endothelial growth factor
VGPR	Very good partial response
WHO	World Health Organisation

Table of Contents

Table of Tables

Table of Figures

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. 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=275) 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 (BIRC) 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 had 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 data for BSC was not identified.

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 median 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 outcomes for which this method was applied were OS, PFS and ORR. Key characteristics were identified using literature searches and using discussions with clinical advisors. Eleven characteristics were initially identified, but no more than four characteristics were used per outcome. It was reported that stepwise model selection suggested that the best Cox Proportional hazards (PH) model for OS is based on Eastern

Cooperative Oncology Group (ECOG) performance status (PS), haemoglobin level, visceral metastases and liver metastases. For PFS the same approach showed the best model is based on ECOG PS, age, visceral metastases and liver metastases. Stepwise model selection suggested that the best logistic regression model for objective response is based on age and visceral metastases. The basis of selection was reported to be parsimony as indicated by the Akaike information criterion (AIC). No models other than the final and presumably most parsimonious models (no more than four covariates) were presented despite the consideration of 11 possible covariates. Since an unanchored STC relies on the major assumption that all effect modifiers and prognostic factors are accounted for, the NICE Decision Support Unit (DSU) Technical Support Document (TSD) 18 recommends caution in the application of the method. It also recommends a so-called 'out-of-sample' method for estimating the residual bias of any STC, due to effect modifiers or prognostic variables that are not accounted for in the prediction models. The company provided such an analysis in their response to the request for clarification.

Finally an evidence synthesis model was used to synthesise the results of the STC i.e. adjusted hazard ratios (HRs) (for OS and PFS) and odds ratios (for ORR) across all trials. For OS and PFS this enabled the adoption of an evidence synthesis model that did not require a PH assumption i.e. a fixed HR of nivolumab versus each comparator, but instead allowed the HR to vary over time, one HR per four-week period. This model, based on a paper by Jansen, 2011, is known as fractional polynomial (FP) and through variation in a set of up to two key parameters (P1 and P2) permits a wider variation in the form of the survival curves. Choice of FP model was reported to have been determined by best statistical fit, although the results of only two other sets of parameter values out of many possible were presented in Appendix D. The company also presented the results of analyses based on a PH model for OS and PFS i.e. fixed HRs in response to the request for clarification. The company were also requested in the clarification letter to present the results by Programmed death ligand 1 (PD-L1) subgroup, but they declined citing lack of baseline data in the comparator studies.

The systematic review identified 12 trials for inclusion in the STC; three were excluded as the dose and/or treatment regimens did not correlate with current UK clinical practice. In addition to the two nivolumab studies, two comparator studies were identified of paclitaxel, two of docetaxel, one of BSC, and two of cisplatin plus gemcitabine. Because not all studies reported all outcomes, only five were used for OS, one per comparator for all comparators except docetaxel for which there were two. The comparator studies were a mix of randomised controlled trials or single arm studies. For PFS only three were used, two for docetaxel and one for paclitaxel. For ORR six of seven studies were synthesised, only one paclitaxel study not being included. There was much variability in patient populations between the included studies of the STC.

The analysis based on the STC and using a fixed effect FP model with P1=0 and P2=0 found that for OS nivolumab is superior to all comparators but only at certain time points; the credible intervals for the HRs were quite wide and indicated the results were not always statistically significant. For OS nivolumab was statistically superior to: paclitaxel at time points between 44 and 72 weeks (HR 2.63, 95% CrI 1.17 to 5.52, 68 -72 weeks); docetaxel at time points between 20 and 72 weeks (HR 2.01, 95% CrI 1.14 to 3.37, 68 -72 weeks); BSC at time points between 20-72 weeks (HR 1.86, 95% CrI 1.17 to 2.85, 68 -72 weeks). Nivolumab was superior to cisplatin plus gemcitabine above 20 weeks but never reached statistical significance.

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 275. 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.
6. Ideally the results of the STC would be based on independent review (BIRC) assessment methods. Given that the BIRC method was only available for CheckMate 275 at a minimum it

would have been useful to perform the STC using only the CheckMate 275 data. This was suggested to the company but was not performed.

7. The major assumption for unanchored STC is that all effect modifiers or prognostic variables are accounted for. Not all of the key characteristics (possible effect modifiers or prognostic variables) for the STC were reported for all comparator trials, therefore imputations were required for these characteristics which were based on correlations to the baseline characteristics in the nivolumab trials.
8. The method used for the prediction models lacked transparency; the results at each stage of the stepwise selection process were not provided. In particular, it is not clear that the most parsimonious model is the best model. It would have been useful to see an STC that was based on prediction models with more covariates including all 11 considered. The only external test of validity of the STC i.e. the 'out-of-sample' method seemed to either show insufficient reduction in bias or be inapplicable given the use of the fractional polynomial model that was used for survival analysis. As stated on page 56 of TSD 18: '*The size of this systematic error can certainly be reduced, and probably substantially, by appropriate use of…STC. Much of the literature on unanchored … STC acknowledges the possibility of residual bias due to unobserved prognostic variables and effect modifiers; however, it is not made clear that the accuracy of the resulting estimates is entirely unknown, because there is no analysis of the potential magnitude of residual bias, and hence no idea of the degree of error in the unanchored estimates. It is, of course, most unlikely 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]

Analysis of the single arm studies alone indicates that there is little difference in survival at least at the median between nivolumab at 8.74 and 9.72 respectively and either docetaxel and paclicaxel, at 9.2 or 8 months respectively. The value for gemcitabine plus paclitaxel was higher at 10.5 months.

The ERG found that the FP model for synthesising HRs for OS and PFS is supportable partly because of its flexibility in permitting a wide variety of functional forms from fixed HRs (PH assumption) to time varying HRs with different shaped survival curves. However, whilst the company stated that they chose the base-case models on the basis of best fit, the results of only two of many parameter sets were presented in Appendix D. The company did provide the results for PH models in response to the clarification request, but the method used has questionable validity and was not the one recommended in the paper on which the FP approach was based. The ERG was able to reproduce the base-case PF model (fixed effect, P1=0, P2=0) results for OS and PFS at least close enough that any difference could be explained by uncertainty. The ERG was also able to produce results that were based on unadjusted values of hazards for nivolumab by applying the fixed HR, one for each comparator trial reported in Appendix D i.e. as if estimated without the STC for these base case PF models. This confirmed that the model used for the adjustment had been a PH model as described by the company. However, the uncertainty in these unadjusted HRs was not estimable without the original nivolumab IPD. Finally, the ERG did find that the HRs estimated using a PH model according to Jansen, 2011 were different to those provided by the company by an amount that did not seem explicable by uncertainty.

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 lacked success in reducing the bias (if it is applicable at all given the lack of data and FP 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.

1.4 Summary of cost effectiveness submitted evidence by the company

Systematic literature review

The company performed a SLR with the objective to identify evidence to support the development of a cost effectiveness model for nivolumab as a treatment for locally unresectable or metastatic urothelial cancer (UC). Although economic evaluations were identified with populations that matched the population described in the final scope of this appraisal, these did not consider the cost effectiveness of nivolumab.

Model structure and main modelling decisions

The company developed a de novo economic model using a cohort-based partitioned survival model. The model consists of three mutually exclusive health states: progression-free (PF) and post-progression (PP) disease states and death. Patients enter the model in the PF state and are treated with nivolumab or one of its comparators. Patients remain in the PF state until disease progression or death. The proportion of patients in each health state is determined by overall survival (OS) and progression-free survival (PFS) curves.

The model includes patients with metastatic or unresectable UC who have progressed following first-line platinum-based chemotherapy. Patient characteristics included in the model were age, gender, weight and body surface area (BSA) based on the CheckMate 275 study.

Nivolumab is implemented as per its EMA Summary of Product Characteristics (SmPC) posology and method of administration for second-line UC (i.e. 3mg/kg Q2W).

The company considered the following comparators in their base-case:

- Paclitaxel: 80mg/m^2 Q3W of a four week cycle
- Docetaxel: 75mg/m^2 Q3W
- Best supportive care (BSC)

The company also presented a scenario analysis, in which cisplatin plus gemcitabine was added as a comparator. The company justified this deviation from the scope (i.e. not including cisplatin plus gemcitabine in its base-case) by stating that there was limited evidence for retreatment with first-line platinum-based chemotherapy regimens for patients with locally advanced unresectable or metastatic UC.

The model adopts the perspective of the NHS and Personal and Social Services (PSS) in England and Wales. The cycle length is four weeks to account for the length of treatment cycles. 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.

Treatment and relative effectiveness

Treatment effectiveness estimates were derived from the CheckMate 275 and CheckMate 032 studies. The time-to-event data of both studies were combined for the survival analyses, but the pooling method was not stated. Parametric time-to-event models were used to estimate overall survival (OS), progression-free survival (PFS) and time-to-treatment discontinuation (TTD) in the company's cost effectiveness model. A response-based approach was adopted to estimate OS and PFS, but not for TTD in the company's base-case. In response to clarification questions, the company also enabled a response-based analysis for TTD for scenario analysis. The response-based analysis was used because, according to the company, standard survival modelling approaches would not appropriately characterise the novel mechanism of action of nivolumab and standard parametric time-to-event models were not deemed flexible enough to characterise the change in hazard over time resulting from having (long-term) responders, and non-responders (no supporting evidence provided). In its response-based analysis, the company used a landmark analysis to prevent the occurrence of immortal-time bias. In this landmark analysis, OS and PFS of both groups (responders and non-responders) were estimated together until a specified landmark point (eight weeks in the company's base-case, 26 weeks explored in scenario analysis) based on the Kaplan-Meier estimates, after which different survival curves were fitted for each group and adjusted for background mortality. The parametric time-to-event models used to estimate OS and PFS after the landmark were selected based on statistical fit (Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC)) and visual inspection. Out of exponential, Weibull, Gompertz, lognormal, log-logistic and generalised gamma, the generalised gamma was chosen to estimate OS and PFS of both responders and non-responders. OS and PFS estimates obtained from the parametric time-to-event models estimated for responders and non-responders separately were combined by using a weighted average, with the weighting based on the proportion of responders in patients being progression-free and alive at the eight-week landmark point. This weighting was held constant throughout the model time horizon. The adjustment for background mortality was based on UK life tables and incorporated using a distribution around the mean UK age (instead of the mean age of the cohort).

The relative effectiveness of nivolumab versus the comparators was modelled through time-varying HRs obtained mainly via the STC. 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. The company explained that the predicted OS and PFS of the comparators were mostly lower than the observed OS and PFS, especially for docetaxel, because of the differences in patient characteristics between the comparator trials and the CheckMate studies. Data not available from the STC relied on the following assumptions: PFS for BSC was derived assuming that the HR for BSC versus paclitaxel was equivalent to that of BSC versus vinflunine for second-line UC patients, and then applying this HR to the paclitaxel PFS curve. This HR was held constant during the time horizon of the cost effectiveness model, due to the absence of alternative data. PFS estimates for cisplatin plus gemcitabine were derived by assuming equivalence of cisplatin plus gemcitabine PFS with that of paclitaxel. No evidence was provided to support this assumption.

Time-to-treatment discontinuation

TTD was estimated through a parametric time-to-event model that was selected based on statistical fit (AIC and BIC) with the pooled CheckMate studies, as well as other, unspecified, considerations. In the CS, TTD was estimated independent of response status but response-based TTD analysis was enabled in response to clarification questions. Even though the Gompertz and the log-logistic distributions showed a better fit, the generalised gamma distribution was selected to estimate TTD in the base-case

analysis, with the company claiming that this was to ensure consistency with OS and PFS and that these two distributions produced long tails with some patients still on treatment after five and 10 years. TTD of the comparators was based on their respective PFS curves because it was assumed that comparator treatment would continue until disease progression. For paclitaxel, only six cycles of treatment were assumed (24 weeks). BSC was assumed to be administered until death.

Adverse events

The company stated that grade 3-4 adverse events were incorporated in the model if their incidence was ≥5%. The impact of adverse events on quality of life and resource use and costs were incorporated in the first cycle of the model.

Health-related quality of life

None of the studies identified by the SLR were consistent with the NICE reference case and therefore EQ-5D-3L data valued with UK preference weights were taken from the CheckMate 275 trial. These utility estimates were stratified according to progression-free and post-progression health states. Utility estimates were derived using a mixed-effects model to reflect within subject variance, after interpolating for measurement times deviating from the measurement schedule and adjusted for missing data using multiple imputation. This resulted in health state utilities of 0.718 and 0.604 pre-progression and post-progression respectively.

The company applied disutilities to several AEs based on studies reporting utilities in patients with non-small cell lung cancer, pancreatic cancer and leukaemia. Disutilities were not treatment-specific and were applied as one-off events at the beginning of treatment, based on the proportion of patients experiencing the adverse event and the duration of the adverse event.

Resource use and costs

Resource use and unit costs data to inform the economic model were based on a number of sources, including CheckMate 275, national databases, published sources (both sources identified and not identified in the SLR), clinical advice and assumptions. British National Formulary (BNF) was used to obtain unit prices for nivolumab (40mg and 100mg), which were adjusted by a Patient Access Scheme (PAS), [redacted] Unit prices for docetaxel, paclitaxel and gemcitabine plus cisplatin were taken from the electronic market information tool (EMIT). The dose/number of vials required per administration were estimated based on dosage scheme and dose intensity (reflecting missed doses), using estimations of patient average weight and body surface area (both based on the CheckMate 275 study) and calculating dose intensity based on data from CheckMate 275 and CheckMate 032 assuming that all delayed doses represent missed doses. Dose intensity for all comparators was assumed equal to that of nivolumab. Administration costs were added to each dose. Monitoring cost (while on treatment) estimates were based on resources estimated using expert opinion and unit prices derived from NHS reference costs. Best supportive care costs were incurred until death after treatment discontinuation. Although not described in the CS, treatment dependent AE costs were incorporated as one-off event costs for patients on treatment during the first cycle of the model based on their occurrence.

Cost effectiveness results

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 docetaxel, paclitaxel, cisplatin plus gemcitabine and BSC 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.

Furthermore, the ERG wishes to express strong concerns about the appropriateness of response-based analysis, implemented through landmark analysis. The need for response-based analysis was inadequately justified, with the company failing to demonstrate how standard parametric survival analysis methods failed to describe the mechanism of action of nivolumab in urothelial carcinoma. In contrast to what the company stated, most standard parametric time-to-event models do include changing hazards over time and some allow for non-monotonic changing hazard functions over time. No mathematical reasoning was provided and based on visual inspection of the conventional, not

response-based, conventional, survival analysis alone, it is the ERG's view that the need for response-based analysis could not be established. The ERG considers that a standard approach should be shown to be inappropriate in the particular decision problem at hand before discarding it.

If, however, the need for alternative methods to conventional survival analysis could be justified, it is the ERG's view that the methods recommended in NICE DSU TSD 14 should be considered before adopting a landmark analysis. However, the company stated that these alternatives, such as spline-based or mixture cure models, were not considered. In summary, the company (a) did not provide sufficient evidence to demonstrate that conventional parametric time-to-event models failed to describe nivolumab survival, (b) did not provide evidence to support that the committee's criticisms on previous nivolumab appraisals applied to the current appraisal, and (c) did not provide evidence to demonstrate that the landmark analysis provided more valid results than standard survival modelling analyses or alternative methods recommended in TSD 14 (for example, no expert opinion was used to validate the resulting survival curves).

The use of response-based landmark analysis introduced further assumptions and additional uncertainty into the cost effectiveness analysis. These assumptions include (a) the choice of the eight-week landmark, with alternative choices causing unpredictable changes in cost effectiveness (the company only provided one alternative landmark and declined to provide others upon request); (b) the use of Kaplan-Meier estimates for the period up to the landmark instead of fitting a parametric curve until then may result in overfitting; (c) fitting parametric models to the responder and non-responder groups also results in larger uncertainty about these fitted curves: the sample size used is significantly smaller because of the splitting up of the study population into two groups and because only the available data after the landmark is used; (d) responder and non-responder groups were then combined for the indirect comparison casting further doubt over whether the response-based analysis has any benefits, especially given that hazard ratios are derived from the overall population and are then applied in a combined responder and non-responder population. The combination of curves was implemented using a weighted average, with the weight being the proportion of responders at the landmark, which was held constant. This inflated the proportion of non-responders in later periods because the proportion of responders is expected to increase over time compared to the proportion of non-responders; (e) response-based and conventional approaches result in vast differences in the predicted life years for nivolumab, with a predicted mean of 2.80 life years in the response-based analysis and 1.84 life years in the conventional, not response-based, approach (deterministic estimates). No explanation for this deviation was provided, and none of the response-based model predictions were validated using expert opinion. The use of response-based, and landmark, analysis had by far the biggest impact on the ICERs, with ICERs being significantly decreased in all comparisons when using the response-based approach.

The ERG's concerns about the selection of parametric time-to-event models include the rejection of the proportional hazard assumption between responders and non-responders without sufficient justification, and the simultaneous selection of parametric time-to-event models for responders and non-responders, which stands in contrast to the company's statement that there was *'no requirement to assume the same distribution to be appropriate for both responder and non-responder curves'*. This led to selection of the generalised gamma distribution, despite it not making the best statistical fit for non-responders (the Weibull makes a better fit). The company provided an updated model allowing the selection of differential distributions for responders and non-responders. Of further concern is that, despite NICE DSU TSD 14 recommendations, the choice of parametric time-to-event models for the response-based approach was not supported by expert opinion. Furthermore, the company was inconsistent in not using response-based analysis for estimating TTD. For TTD, the company chose the generalised gamma distribution despite it not having the best statistical fit and justified their choice by stating that the better

fitting Gompertz and log-logistic distributions would result in implausible numbers of patients still on treatment at five years. The choice of differential parametric time-to-event curves for responder and non-responder OS, PFS and TTD was shown to significantly increase the ICERs in ERG scenario analyses.

The cost effectiveness analysis model suffers from significant uncertainty and bias induced by comparing single-arm studies through the STC. It is the ERG's opinion that the discrepancy in populations in which relative effectiveness estimates were derived (adjusted CheckMate 275 and CheckMate 032 population) and applied (i.e. the combined but separately estimated responder and non-responder survival curves) induced bias that could not be quantified and that the company declined to comment on, despite the ERG's request. The ERG would have preferred to apply separate HRs to responders and non-responders. However, the company did not provide these, stating that small numbers in responder and non-responder groups did not allow separate estimation of relative effectiveness.

The company did not sufficiently justify the need for time-dependent HRs to model the relative effectiveness of nivolumab versus the comparators, providing log-cumulative hazard plots that showed the separate CheckMate studies, while the HRs were derived based on the pooled CheckMate studies dataset. The ERG considers that therefore proportionality of hazards could not be ruled out. Time-independent HRs were provided by the company in response to clarification questions but these could not be replicated by the ERG. The use of the time-independent HRs produced by the ERG increased all cost effectiveness estimates in ERG scenario analysis. The ERG notes that using time-independent HRs has the advantage of preventing over-parameterisation which might occur when estimating time-dependent HRs with the relatively limited amount of data submitted by the company.

Assumptions that were not supported by clinical evidence were made around the relative effectiveness of nivolumab versus cisplatin plus gemcitabine and BSC in terms of PFS to make up for lack of data to inform these. Alternative assumptions in ERG scenario analysis only had a small effect on the ICERs in these comparisons.

The parameterisation of the fractional polynomial model that informs the NMA was found to have a large impact on cost effectiveness outcomes. In a PSA only varying the parameter values of the FP model between those parameter values that were provided as possible parameter combinations by the company resulted in substantial differences in incremental costs and QALYs for all comparators (for instance, incremental QALYs of nivolumab vs docetaxel had a credible interval of [redacted] to [redacted]).

Adverse events

Only the CheckMate 275 trial was used to inform the adverse event rates in the cost effectiveness model while the clinical effectiveness of nivolumab was estimated based on both CheckMate studies. The selection of sources for adverse events associated with comparators was not appropriately justified. The inclusion of both neutropenia and leukopenia was questionable, given that neutropenia 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 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.

Overall the systematic review process was well documented and appeared to be performed well.

The ERG considers the adopted perspective, time horizon and discounting used in the model to be appropriate for this appraisal. Incorporation of costs, resource use, and HRQoL data was appropriate, with a few minor errors and questionable judgements. The model structure followed that of past NICE technology appraisals in metastatic cancers. The company explored a range of different parametric time-to-event models to model survival data.

1.6.2 Weaknesses and areas of uncertainty

All nivolumab trial data were based on March, May and September 2016 database locks. More up-to-date data was requested but was not provided.

The ERG was concerned that limiting the MEDLINE and Embase clinical effectiveness searches to English language only publications may have introduced potential language bias.

No randomised controlled trials (RCTs) were identified for nivolumab.

There were no studies that directly compared nivolumab with any specified comparator. Furthermore, there were no studies that could provide a common comparator to support indirect comparison or MTC.

There are serious concerns regarding the representativeness of the nivolumab trial patients to the UK population. Firstly, only six patients from one trial were from the UK. Secondly, as few as 18.8% of patients in the UK might have and ECOG performance status of 0, as opposed to over 50% in the two nivolumab trials. Thirdly, there is a mismatch in terms of prior therapies, as many as over 75% of patients in the UK would have previously taken a gemcitabine platinum-based combination compared to fewer than 40% in the trials. Finally, there is a question of the applicability to those with locally advanced unresectable as opposed to metastatic disease given the very small proportion of such patients in the trials.

Risk of bias was not assessed appropriately for the single arm studies (which include those for nivolumab). Single arm studies are by definition low down in the hierarchy of study design and therefore the quality of these studies is low to start with and risk of bias tools have not been widely developed for this study design. With this is mind risk of bias was judged to be high for all data used in the STC given that only single arms were used.

No STC analysis for AEs or HRQoL was performed.

The STC analysis is compromised by many limitations (listed earlier) which impairs the ability to critique the presence of residual bias. Given that TSD 18 states 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"* the ERG does not think the STC methods are sufficiently reported nor validated to sustain the companies claims.

The company did not provide the comparison of nivolumab with cisplatin plus gemcitabine in their base-case model, despite it being in the scope.

With regards to a response-based modelling approach, the use of unconventional, response-based, landmark survival analysis, without sufficient justification for its need necessitated further assumptions and thereby substantially increased uncertainty. Assumptions introduced include the choice of the eight-week landmark, with alternative choices causing unpredictable changes in cost 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 £87,709, £68,519 and £69,515 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 1.1: 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

	Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Nivolumab ICER (£/QALY)
	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	£12,493	0.74	[redacted]	[redacted]	£87,709
	Paclitaxel	£13,866	0.63	[redacted]	[redacted]	£68,519
	Cis+gem	£29,384	1.24	[redacted]	[redacted]	Dominated
	BSC	£8,696	0.56	[redacted]	[redacted]	£69,515
Alternative parametric TTE models (lognormal for OS, log-logistic for PFS) (A.1)[b]	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£13,173	1.01	[redacted]	[redacted]	£45,721
	Paclitaxel	£14,654	0.89	[redacted]	[redacted]	£39,286
	Cis+gem	£29,736	1.58	[redacted]	[redacted]	£72,732
	BSC	£9,235	0.72	[redacted]	[redacted]	£38,147
Naïve comparison data instead of STC results (A.3)[b]	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£13,005	0.77	[redacted]	[redacted]	£92,335
	Paclitaxel	£13,914	0.60	[redacted]	[redacted]	£64,914
	Cis+gem	£30,910	1.56	[redacted]	[redacted]	Dominated
	BSC	£8,630	0.52	[redacted]	[redacted]	£65,593
Response-based analysis (B.1)[c]	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£12,783	0.84	[redacted]	[redacted]	£53,273
	Paclitaxel	£14,163	0.73	[redacted]	[redacted]	£44,877
	Cis+gem	£30,310	1.39	[redacted]	[redacted]	£103,186
	BSC	£8,811	0.59	[redacted]	[redacted]	£44,183
Response-based analysis using alternative TTE models for OS, PFS and TTD (B.3)[c]	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£12,452	0.77	[redacted]	[redacted]	£77,597
	Paclitaxel	£13,948	0.67	[redacted]	[redacted]	£67,608
	Cis+gem	£29,880	1.25	[redacted]	[redacted]	£143,923
	BSC	£8,662	0.55	[redacted]	[redacted]	£64,282

	Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Nivolumab ICER (£/QALY)
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						

2. BACKGROUND

In this section the ERG provides a review of the evidence submitted by Bristol-Myers Squibb in support of nivolumab, trade name Opdivo® for the treatment of metastatic or unresectable UC after platinum-based chemotherapy. We outline and critique the company's description of the underlying health problem and the overview of current service provision. The information is taken mainly from Chapter B.1.3 of the company submission (CS) with sections referenced as appropriate.

2.1 Critique of company's description of underlying health problem.

The underlying health problem of this appraisal is metastatic or unresectable UC in adult patients who have received platinum-based chemotherapy.

The company described the origin of UC from the urothelium or epithelial lining of the urinary tract which extends from the renal pelvis to the ureter, bladder and proximal urethra. Urothelial cancer can also be known as transitional cell carcinoma. As described in Table 3 on staging, the bladder is the main organ that is affected. Indeed, the CS states that UC *'accounts for approximately 90% of all bladder cancer'.*[2]

Common presenting symptoms of UC include painless haematuria (blood in the urine), dysuria, frequency, urgency, feeling of incomplete voiding, and straining. In addition, urinary, bowel and sexual functions are affected and therefore impacts on overall health-related quality of life (HRQoL), daily life and sleeping patterns.

The CS states that '*Locally advanced and metastatic disease refers to tumours that have grown through the bladder wall and/or have spread to lymph nodes or other distant sites.'*[2]

The CS outlines the impact of advanced or metastatic UC on patients. This includes symptoms of disease such as limited mobility, abdominal, bone or pelvic pain, anorexia, wasting and pallor.

The CS states that *'UC is the 10th most common cancer in the UK, and is 3–4 times more commonly found in males than females.*[36] *In 2014, there were 9,021 patients newly diagnosed with UC in England and Wales, of which 7,307 (73%) were in males and 2,756 (27%) were in females. The disease is also more common in older adults, with more than half (54%) of UC cases in the UK each year diagnosed in patients aged 75 and over.*

The majority of patients with UC are diagnosed in stages I and II (62%), with approximately 20% diagnosed at the advanced, metastatic stage.[36,2]

In section B.1.3.4, the CS states that *'Based on available data from Cancer Research UK and expert clinician feedback, the number of patients in England and Wales eligible for treatment with nivolumab, as per the licensed indication for locally advanced unresectable or metastatic UC whose disease has progressed following platinum-containing chemotherapy, is estimated to be 894 patients.'*[2]

ERG comment:

The ERG checked the references cited by the company to support the statements made above and considered the company to have provided an appropriate description of the underlying health problem. In addition the ERG would like to indicate death and survival statistics. Around 10% will survive their

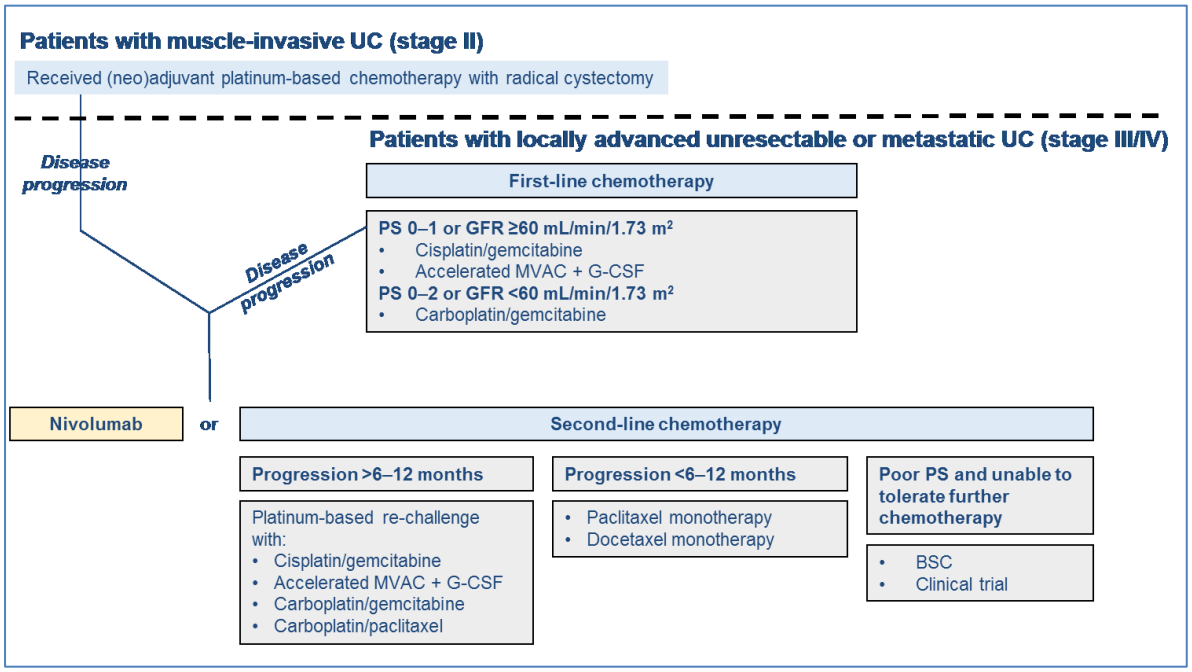
cancer for five years or more after diagnosis with T4 bladder cancer.[3] In 2014 there were 5,369 deaths from bladder cancer in the UK (3% of total cancer deaths).

The ERG notes that the projected numbers (894) eligible for nivolumab treatment were based on clinical expert opinion and could not be verified by the ERG, although the calculations for this figure (Table 56 of the CS) appear to be appropriate.

2.2 Critique of company's overview of current service provision

Figure 2.1 shows the CS current treatment pathway for persons with locally advanced or metastatic bladder cancer as well as the proposed position of nivolumab, based on NICE and EAU/ESMO guidelines and expert clinician feedback.[2]

Figure 2.1: Adapted treatment pathway to show potential position of nivolumab in the treatment of locally advanced or metastatic bladder cancer



Source: Figure 7 of CS

BSC = best supportive care; G-CSF = Granulocyte-colony stimulating factor; GFR = glomerular filtration rate; MVAC = methotrexate, vinblastine, doxorubicin and cisplatin; PS = performance status; UC = urothelial carcinoma

The company quote the NICE guidance for persons with locally advanced or metastatic bladder cancer. They state that '*For patients with locally advanced unresectable or metastatic UC whose condition has progressed after first-line therapy and who are physically fit [ECOG PS 0 or 1] with adequate renal function [GFR 60 ml/min/1.73 m^2], NICE recommends retreatment with cisplatin in combination with gemcitabine, or accelerated (high-dose) MVAC in combination with G-CSF. Patients for whom cisplatin-based chemotherapy is unsuitable (i.e. GFR <60 ml/min/1.73 m^2) may be treated with carboplatin plus paclitaxel in this setting.*'[2]

More specifically NICE guidance (NG2) states: 'Consider second-line chemotherapy with carboplatin in combination with paclitaxel or gemcitabine in combination with paclitaxel'.[4]

The company quote additional input from clinical experts.[5] Feedback from expert clinicians who were in UK clinical practice indicated that *'the vast majority of patients with locally advanced unresectable or metastatic UC following prior platinum-based chemotherapy would be treated with paclitaxel*

monotherapy, with docetaxel monotherapy also used in some centres. Of those patients considered fit enough to be offered second-line treatment with paclitaxel monotherapy, approximately one third to one half of these patients would typically refuse further chemotherapy treatment, and this figure may be even higher in some smaller centres. These patients would therefore currently opt for best supportive care (BSC), which may include painkillers, steroids and blood transfusions. Some patients would also be unsuitable for chemotherapy altogether, and would therefore be offered BSC instead of taxane-based chemotherapy.'[5]

In addition with reference to patients deemed physically fit, the expert clinicians added *'they would only consider retreatment with platinum-based chemotherapy for patients they considered fit enough and who had been progression-free for at least 9–12 months (or 6 months in some centres) following prior platinum-based chemotherapy; as such, this would very much be the minority of patients, representing only 5–10% of cases in the second-line setting.'*[5]

With reference to patients recommended for second line treatment of gemcitabine plus paclitaxel, the expert clinicians added that '*this regimen is used rarely in few centres across the UK and only for patients who have progressed quickly following first-line platinum chemotherapy and are very symptomatic'*[5]

The company suggest two potential positions for nivolumab in the treatment of for locally advanced or metastatic UC after failure of prior platinum-containing chemotherapy:[2]

1. In first-line locally advanced unresectable or metastatic disease, following disease progression after prior platinum-containing therapy received as (neo)adjuvant therapy with radical cystectomy in the muscle-invasive disease stage
2. In second-line unresectable or metastatic disease, following disease progression after prior platinum-containing therapy received in the locally advanced unresectable or metastatic disease stage.

ERG comment:
The company's description of the treatment pathway and options was based on existing NICE guidance (NICE guideline NG2; Bladder cancer: diagnosis and management) which is appropriate and relevant to the decision problem.[4] In particular the second-line treatment options for the management of locally advanced or metastatic bladder cancer were most relevant for the position of nivolumab in the treatment pathway. The company provided an adapted pathway based on inputs from clinical experts, this appears to be sensible, assuming the expert opinions are correct (this data could not be verified by the ERG as it is not in the public domain).

The ERG draws the attention of the committee to the potential placement of nivolumab at second-line for patients with locally advanced unresectable or metastatic UC, which is in accordance with the scope. However, the placement following progression subsequent to muscle-invasive disease (stage II) is not within scope.

The ERG notes the following ongoing appraisals relevant to the decision problem, as mentioned in the scope:[6]

Atezolizumab for treating metastatic urothelial bladder cancer after platinum-based chemotherapy [ID939] Publication expected September 2017.

Pembrolizumab for previously treated advanced or metastatic urothelial cancer (ID 1019) Publication expected October 2017.

3. 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.
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.	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 will be considered from an NHS and Personal Social Services perspective.	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 is not in accordance with the population defined in the scope. 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]

In the CheckMate 275 trial, nivolumab (BMS-936558) was administered intravenously over 60 minutes at 3 mg/kg every two weeks until progression or unacceptable toxicity. This is in line with the decision problem.[10]

In the CheckMate 032 trial, patients were given nivolumab (3 mg/kg administered by intravenous infusion every two weeks) as monotherapy or in combination with ipilimumab. For the purposes of the CS only the nivolumab monotherapy patients were included, however they could switch to ipilimumab. Eighteen (23%) of 78 patients (receiving nivolumab monotherapy) switched to combination treatment with ipilimumab upon disease progression.[9] Therefore the ERG considers that the intervention in CheckMate 032 is not in line with the intervention described in the final scope.

3.3 Comparators

The NICE scope indicates four possible comparators: retreatment with first-line platinum-based chemotherapy (only for people whose disease has had an adequate response), paclitaxel, docetaxel and best supportive care. The company submission presents evidence for three comparators only: paclitaxel, docetaxel and best supportive care.

Both included nivolumab trials were single arm studies and therefore no direct or indirect comparators could be included. The company submission used a simulated treatment comparison (STC) to provide comparisons of nivolumab to paclitaxel, docetaxel and best supportive care; cisplatin plus gemcitabine were included only as part of a scenario analysis.[1] Cisplatin plus gemcitabine were only included in a scenario analysis because the company submission stated they had limited generalisability to the decision problem, the specific reasons given were:

1) patients '*had received MVAC in first-line treatment and are therefore not considered to be directly comparable to those receiving cisplatin plus gemcitabine retreatment in current UK clinical practice, as they are gemcitabine naïve'* (section B.2.9.1 CS) Gondo *et al.* (2011).[13]

2) inclusion of '*chemotherapy-naïve patients in addition to patients who had previously undergone first-line treatment*' (section B.2.9.1 CS) Ozawa *et al.* (2007).[14]

3) '*the two trials did not use the standard dosing regimen typically used for cisplatin plus gemcitabine in the UK'* (section B.2.9.1 CS)[2]

According to NICE guidelines [NG2] gemcitabine and cisplatin or MVAC and G-CSF can be given as both first line and second line treatments, for locally advanced and metastatic bladder cancer.[4] Also, whilst it is true that for one trial patients who were chemotherapy naïve were included,[14] this was not the trial that informed OS.[13] Therefore the ERG would not consider cisplatin and gemcitabine to be unsuitable for inclusion in the STC, especially given the limitations of the nivolumab and other comparator trials.

3.4 Outcomes

The company states that it assessed all the outcomes of the decision problem (overall survival, progression-free survival, response rates, adverse effects of treatment, health-related quality of life). However there were no direct or indirect comparators and the company submission used a STC to provide evidence of effectiveness to the comparators listed above. For the STC only three outcomes were considered; overall survival, PFS and ORR (section B.2.9 CS).[2]

There was no comparative data for adverse events or for quality of life. Note that adverse events and quality of life were reported for the two trials, but since these were single arm trials these results were

not informative. Adverse event data were provided in the response to clarification.[7] However, unlike for effectiveness, no evidence synthesis was performed for either of these two outcomes.

3.5 Other relevant factors

As stated by the company: 'A PAS [patient access scheme] is already in place with the Department of Health for inclusion in this technology appraisal, representing a simple discount of [redacted] on the list price of nivolumab' (CS, page 18).[2]

According to the company this STA fulfils the end-of-life criteria because:

- No studies identified in the SLR reported in Appendix D of the CS provided evidence of OS estimates for this patient population that approached 24 months.
- The economic analysis predicted mean life years (LYs) per patient with nivolumab of 2.78 years (33.36 months). In comparison, predicted mean LYs per patient with comparator therapies were 1.19 years (14.28 months) with paclitaxel, 1.40 years (16.80 months) with docetaxel and 1.01 years (12.12 months) with BSC. Nivolumab was therefore predicted to offer an extension to life of considerably greater than three months versus each of these comparators. Furthermore, in the context of the average survival of patients receiving paclitaxel, docetaxel or BSC, the survival gains offered by nivolumab represent a significant extension to life.

ERG comment: It appears that life expectancy is less than 24 months. However, given the absence of comparative trial data it is impossible to be confident of the extension to life resulting from treatment with nivolumab versus any of the comparators. The company bases the claim of extension to life on the economic model, which is informed by the STC, which attempts to estimate the treatment effect of nivolumab versus the comparators. However, as indicated in Section 4.3 and 4.4, the STC methods used to make the adjustment to reduce bias are not completely transparent, are accompanied by several limitations and are likely to result in residual bias (as argued in the methods guide followed by the company, NICE DSU TSD 18).[1] It is clear is that there is little difference in survival at least at the median between nivolumab (CheckMate 275 and CheckMate 032 trials)[10, 11] at 8.74 and 9.72 respectively and either docetaxel and paclicaxel, at 9.2 or 8 months respectively.[15, 16] The value for gemcitabine plus paclitaxel was even higher at 10.5 months.[13] It is true that the differences in these values are subject to potential bias given that the trial data represents observational data, but it is also true that the evidence provided by the STC to reduce this bias is far from clear.

4. CLINICAL EFFECTIVENESS

4.1 Critique of the methods of review(s)

The company conducted a systematic review to identify relevant direct and indirect clinical evidence on the use of nivolumab in metastatic or unresectable UC. This section critiques the methods of the review including searching, inclusion criteria, data extraction, quality assessment and evidence synthesis.

4.1.1 Searches

The following paragraphs contain summaries and critiques of all searches related to clinical effectiveness presented in the company submission. The Canadian Agency for Drugs and Technologies in Health (CADTH) evidence based checklist for the Peer Review of Electronic Search Strategies (PRESS), was used to inform this critique. [17] The submission was checked against the Single Technology Appraisal (STA) specification for company/sponsor submission of evidence. [18] The ERG has presented only the major limitations of each search strategy in the report.

The company submission stated that systematic review searches were undertaken in March 2017. Search strategies were reported in Appendix D of the CS for the following databases: Embase, MEDLINE, MEDLINE In-Process, MEDLINE Epub Ahead of Print and the Cochrane Library CENTRAL, DARE, NHS EED and HTA databases. In response to clarification the company confirmed that PubMed was not searched for this review and therefore should not have been listed in Appendix D.1.1.

Additional searches of the following conference proceedings were reported for the last four years: American Society of Clinical Oncology (ASCO), Genitourinary Cancers Symposium (GUCASYM), American Urological Association (AUA), European Association of Urology (EAU), European Society of Medical Oncology (ESMO).

The CS reported that bibliographies of eligible studies were searched for further relevant studies, and the reference lists of any systematic reviews and HTAs were scanned for further studies. ClinicalTrials.gov and WHO International Clinical Trials Registry Platform (WHO ICTRP) were also searched for ongoing clinical trials.

ERG comment:

- The database searches were clearly documented and reproducible, using a wide range of resources to identify published and unpublished literature. Database hosts and dates of searches were all reported. The database searches used combinations of indexing terms appropriate to the resource searched, free text and a number of synonyms for the condition. Study design filters were not applied.
- The search strategies contained some redundancy in their structure, but this will not have affected recall of studies.
- A typographical error in the Cochrane Library database searches noted by the ERG was amended, and searches were re-run by the company in response to clarification. No new relevant records were found.
- The ERG was concerned that limiting the MEDLINE and Embase clinical effectiveness searches to English language only studies may have introduced potential language bias. Current best practice states that 'Whenever possible review authors should attempt to identify and assess for eligibility all possibly relevant reports of trials irrespective of language of publication'.[19]

- All conference searches were conducted via Embase. The ERG has some concerns that relevant abstracts may have been omitted by searching using a biomedical database rather than directly searching conference proceedings, however this is unlikely to have affected the recall of relevant studies.

4.1.2 Inclusion criteria

A systematic literature review was conducted to identify clinical evidence on the efficacy and safety of nivolumab for the treatment of unresectable or metastatic UC. The full text documents were then assessed against the eligibility criteria by two independent reviewers, with disagreements adjudicated by a third reviewer.

The eligibility criteria used in the search strategy for clinical effectiveness are presented in Table 4.1.

Table 4.1: Eligibility criteria used in search strategy for clinical effectiveness

	Inclusion criteria	Exclusion criteria

Population	• Male and female adults aged 18 and over
	• Any ethnicity
	• Trials assessing patients with Stage III or Stage IV advanced, metastatic or unresectable urothelial carcinoma
	• Eligible patients must have progression or recurrence:
	  o After treatment with at least 1 platinum-containing chemotherapy regimen for metastatic urothelial cancer or surgically unresectable locally advanced urothelial cancer, OR
	  o Within 12 months of peri-operative (neo-adjuvant or adjuvant) treatment with platinum agent in the setting of cystectomy for localised muscle-invasive urothelial cancer
	• Trials with mixed populations of patients receiving first and second line treatment will only be eligible if results are reported separately for second line treatment or if more than 50% of the population are receiving second line treatment
		• Paediatric population
		• Patients with Stage I or II urothelial carcinoma
		• Patients undergoing first-line treatment
		• Trials without a defined population
		• Trials with an unclear population

Interventions	• Retreatment with platinum-based chemotherapy (e.g. cisplatin plus gemcitabine, accelerated MVAC (methotrexate, vinblastine, adriamycin/doxorubicin and cisplatin), carboplatin plus gemcitabine or carboplatin plus paclitaxel) • Gemcitabine plus paclitaxel • Docetaxel monotherapy • Paclitaxel monotherapy • Gemcitabine monotherapy	

	Inclusion criteria	Exclusion criteria
	• Vinblastine monotherapy • Vinflunine monotherapy • Best supportive care	
Comparators	• Placebo • Any intervention of interest • Any other treatment that may facilitate an indirect comparison • Best supportive care	

Outcomes	• Overall survival (OS)
	• Progression-free survival (PFS) or time to tumour progression (TTP)
	• Objective response rate (ORR)
	• Complete response (CR)
	• Partial response (PR)
	• Duration of response (DoR)
	• Treatment-related adverse event (AEs):
	  ○ Rates of overall Grade 3 or 4 AEs
	  ○ Rates of specific Grade 3 or 4 AEs including:
	  1. Neutropenia
	  2. Anaemia
	  3. Thrombocytopenia
	  4. Febrile Neutropenia
	  5. Asthenia (Fatigue)
	  6. Nausea
	  7. Vomiting
	  8. Diarrhoea
	  9. Pruritus
	  10. Pneumonia
	  11. Lung infiltration
	  12. Alanine aminotransferase (ALT) increase
	  13. Hepatitis
	• Discontinuation/withdrawals due to AE
	• Health-related quality of life (HRQoL)
		No outcomes of interest

Study design	• Randomised controlled trials • Non-randomised prospective controlled clinical trials or single-arm trials • Systematic reviews – will be eligible for reference checking only • Conference abstracts only to provide supplementary information	• Retrospective trials • Case reports • Case series of fewer than 5 people • Editorials, letters or news articles • Conference abstracts – as the primary trial reference
Language restrictions	English language only	Non-English

	Inclusion criteria	Exclusion criteria
Publication year	NR	NR
Source: CS, Table 7, pages 39-40		

ERG comment:

- The population of the systematic review is in line with the scope.
- The interventions and comparators for the inclusion criteria are appropriate for identifying treatments to facilitate a network analysis of nivolumab versus the comparators of the scope. A separate review for nivolumab only does not appear to have been performed. It is noticeable that nivolumab is not included as an intervention; the ERG assumes this is an oversight by the company given that nivolumab studies are included.
- All the outcomes outlined in the decision problem were included; however the company has limited the inclusion of adverse events to those that are grade 3 or 4. This will preclude assessment of 'all adverse events'.
- Randomised controlled trials, non-randomised controlled trials and single arm trials were all included in the review.

4.1.3 Critique of data extraction

According to Appendix D.1.4 of the CS data extraction was '*carried out by two independent reviewers with disagreements adjudicated by a third reviewer*'.[20]

ERG comment: The ERG believes that overall the data extraction was carried out appropriately.

4.1.4 Quality assessment

According to Appendix D.1.5 of the CS quality assessment was '*carried out by two independent reviewers with disagreements resolved through discussion with a third reviewer*'.[20]

Quality assessment was performed for prospective cohort trials using the CRD Cohort Trial Checklist (reference 21 of the CS) and for randomised controlled trials using the guidance of the Centre for Review and Dissemination (reference 22 of the CS).

There were 12 trials included in the STC. Two single arm studies were identified for nivolumab; both were open label and single arm studies. The remainder trials were a mix of randomised controlled trials or single arm studies.

For the quality assessment of the randomised controlled trial the following domains were assessed: randomisation, allocation concealment, comparability of groups, blinding, drop out, selective reporting of outcomes and use of intention to treat analysis and appropriate methods for dealing with missing data (summarised in Table 14, D.1.5. of the CS)

Cohort studies are classed as a comparison of outcomes between a group of participants who have received an intervention and a group who have not. This is clearly not appropriate for a single arm study. For the quality assessment of cohort studies the following domains were assessed: comparability of groups, were the groups assessed at similar time points of disease progression, was the intervention reliably ascertained, comparable confounding variables, adequate adjustment of confounding variables, was a dose response relationship between intervention and outcome demonstrated, blinding, adequate follow-up, proportion of the cohort followed up, comparable drop-out rates. (Summarised in Table 13, D.1.5. of the CS). From this list it is clear that most questions are concerned with the comparability between groups, thereby illustrating that this risk of bias tool is not appropriate for the single arm studies

identified within the CS. Single arm studies are by definition low down in the hierarchy of study design and therefore the quality for these studies is low to start with and risk of bias tools have not been widely developed for this study design.

ERG comment: Study quality appeared to be appropriately assessed for randomised trials but not for the single arm studies (which include those for nivolumab). However, risk of bias has to be deemed to be high for all data used in the STC given that only single arms were used.

4.1.5 Evidence synthesis

According to the company, '*Data from CheckMate 275 and CheckMate 032 were pooled in the context of the STC presented in Section B.2.9 and Appendix D'* (CS, section B.2.8, page 59).[20] However, no methods are presented for the pooling of results, and results themselves have not been reported either. We asked the company to provide details of the statistical method(s) used for pooling the data from Checkmate 275 and CheckMate 032 and to explain which data were used (BIRC or investigator-assessed)**.** We also asked the company to conduct pooled analyses using data from each method separately.[21]

In the response to the clarification letter, the company did not state how the two nivolumab trials were pooled. They did clarify that the BIRC method was chosen for CheckMate 275 and only the investigator-assessed results were available for CheckMate 032.[7] They also stated the following on page 26 of the response: *'As agreed with the ERG on the preliminary teleconference to discuss the clarification questions, analyses using each method separately have not been provided.'* However, no such agreement was made. The ERG continues to believe that results derived from performing the STC twice using a) only BIRC or b) only investigator-led methods would provide valuable insight into the variability of the data. Given that the BIRC method was only available for CheckMate 275 this would imply using only the CheckMate 275 data for STC. This was suggested to the company during the teleconference but the analysis was not provided.

4.2 Critique of trials of the technology of interest, their analysis and interpretation (and any standard meta-analyses of these)

The company conducted a systematic literature review to identify relevant clinical evidence on the efficacy and safety of nivolumab for the treatment of unresectable or metastatic urothelial carcinoma. Two trials investigating nivolumab were found: CheckMate 275[8, 10] and CheckMate 032[9, 11].

An overview of CheckMate 275 and CheckMate 032 is provided in Table 4.2.

Table 4.2: Clinical effectiveness evidence for nivolumab

Study	CheckMate 275 (NCT02387996)	CheckMate 032 (NCT01928394)
Publications (primary reference in bold)	**Sharma *et al.* (2017)[8]** Clinical study report[10]	**Sharma *et al.* (2016)[9]** Clinical study report[11]
Study design	Multicentre, open-label, single-arm phase II study	Multicentre, open-label, two-stage, multi-arm, phase I/II[a]
Population	Patients with locally advanced unresectable or metastatic UC who had progressed or recurred after at least one previous line of platinum-containing chemotherapy (N=270)	Patients with locally advanced unresectable or metastatic UC who had progressed or recurred after treatment with at least one platinum-

				containing chemotherapy regimen (N=78)		
Intervention(s)	Nivolumab (IV 3 mg/kg Q2W)			Nivolumab (IV 3 mg/kg Q2W)		
Comparator(s)	N/A (single-arm)			N/A[a]		
Indicate if trial supports application for marketing authorisation	Yes	Indicate if trial used in the economic model	Yes	Yes	**Indicate if trial used in the economic model**	Yes
Reported outcomes specified in the decision problem	ORR OS PFS HRQoL via the European Organisation for Research and Treatment of Cancer (EORTC) general cancer module (QLQ-C30) and the EuroQoL-5 dimensions-3 levels (EQ-5D-3L) questionnaires Adverse events (AEs)			ORR OS PFS EQ-5D-3L AEs		
All other reported outcomes	Duration of response and additional safety outcomes			Duration of response and additional safety outcomes		

Source: CS, Table 4, pages 27-28

[a]CheckMate 032 investigated nivolumab or nivolumab combined with ipilimumab in patients with UC, triple-negative breast cancer, gastric cancer, pancreatic adenocarcinoma, small cell lung cancer, and ovarian cancer. Here, presentation of CheckMate 032 refers only to the nivolumab monotherapy UC cohort (n=86) of relevance to this submission.

BIRC = blinded independent review committee; CSR = clinical study report; EORTC QLQ-C30 = European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30; EQ-5D-3L = 3-level EuroQoL 5-Dimensions; HRQoL: health-related quality of life; IV= intravenous; N/A = not applicable; ORR = objective response rate; OS = overall survival; PD-L1 = programmed death ligand 1; PFS = progression-free survival; Q2W = every two weeks; UC = urothelial carcinoma.

4.2.1 Study design and methodology of the nivolumab studies

CheckMate 275

CheckMate 275 is an ongoing, phase II single-arm clinical trial investigating the efficacy and safety of nivolumab in patients with locally advanced unresectable or metastatic UC who had failed at least one previous line of therapy.[8]

Patients with histologically confirmed metastatic or surgically unresectable UC with disease progression or recurrence after at least one platinum-based chemotherapy were enrolled and assigned to a cohort according to tumour PD-L1 expression status (PD-L1 ≥5%, PD-L1 < 5%, or indeterminate). Enrolment in the trial continued until approximately 70 patients with confirmed PD-L1 expression of ≥5% were treated. Enrolment continued further in Japan until approximately 25 Japanese patients were treated, or until November 2015, whichever occurred sooner.

Enrolled patients were treated with IV nivolumab 3mg/kg Q2W until documented disease progression (based on RECIST v1.1 criteria) and clinical deterioration, unacceptable toxicity, or other protocol-defined reasons. Treatment beyond initial investigator-assessed RECIST v1.1-defined progression was

permitted if the patient had an investigator-assessed clinical benefit, did not have rapid disease progression, and was tolerating the study drug.

The primary endpoint of CheckMate 275 was objective response rate (ORR) based on Blinded Independent Review Committee (BIRC) assessment using RECIST v1.1 in the all-treated population, in patients with PD-L1 expression ≥1%, and in patients with PD-L1 expression ≥5%. Objective response was defined as the proportion of patients with best overall response of confirmed complete response (CR) or partial response (PR) assessed by the BIRC. Time to response and duration of response were estimated in patients with a confirmed CR or PR. Responses were confirmed at the second scan at least four weeks after criteria for objective response were met.

The trial consisted of three phases: screening, treatment, and follow-up. Treated patients were evaluated for response according to the RECIST v1.1 guidelines beginning eight weeks (±1 week) after the first dose of nivolumab and then every eight weeks (±1 week) thereafter up to 48 weeks, then every 12 weeks (±1 week) until disease progression (investigator-assessed RECIST v1.1-defined progression) or treatment discontinuation, whichever occurred later. Patients were followed for OS every three months until death, lost to follow-up, or withdrawal of study consent.

CheckMate 032

CheckMate 032 is an ongoing phase I/II multi-arm trial investigating the efficacy and safety of nivolumab monotherapy or nivolumab in combination with ipilimumab in patients with one of the following tumour types: UC, triple-negative breast cancer, gastric cancer, pancreatic adenocarcinoma, small cell lung cancer, and ovarian cancer.[9] The company used a subgroup of patients enrolled in this study in their analyses: 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). Therefore, reference to CheckMate 032 in the CS refers only to this subgroup of UC patients.[9]

A total of 86 patients were enrolled in the nivolumab monotherapy treatment arm of CheckMate 032, of whom 78 patients received at least one dose of nivolumab. All 78 patients who received at least one dose of nivolumab were included in the safety and efficacy analyses. The subgroup of UC patients included in the company analyses (N=78) does include 18 patients who crossed-over to nivolumab in combination with ipilimumab.

Eligible patients with histologically or cytologically confirmed carcinoma of the renal pelvis, ureter, bladder, or urethra and disease progression after at least one previous platinum-based chemotherapy treatment were treated with IV nivolumab 3 mg/kg Q2W until documented disease progression (based on RECIST v1.1 criteria), unacceptable toxicity, or other protocol-defined reasons.

The primary endpoint of CheckMate 032 was the proportion of patients with a confirmed investigator-assessed objective response, defined as the number of patients with a best overall response of a CR or PR as per the RECIST v1.1 criteria divided by the number of treated patients. Patients were evaluated for response at baseline, six weeks after the first dose of nivolumab, continuing every six weeks for the first 24 weeks, and then every 12 weeks until disease progression or treatment discontinuation, whichever occurred later. Patients receiving nivolumab monotherapy could switch to nivolumab plus ipilimumab (nivolumab 1 mg/kg and ipilimumab 3 mg/kg or nivolumab 3 mg/kg and ipilimumab 1 mg/kg intravenously, every three weeks for four cycles) following disease progression if they met prespecified criteria.

For a CR or PR to be judged to be a best overall response, the assessment needed to be confirmed by a second scan no less than four weeks after the criteria for response was first met. Patients who did not meet response-evaluable criteria (i.e. at least one target lesion at baseline and at least one on-study assessment) were judged to be not assessable. Treatment beyond initial investigator-assessed RECIST v1.1-defined progression was permitted if the patient had an investigator-assessed clinical benefit and was tolerating the study drug.

A summary of the methodology and trial design of CheckMate 275 and CheckMate 032 is presented in Table 4.3. Further details of the methodology of CheckMate 275 and CheckMate 032, including the full eligibility criteria can be found in Appendix M of the CS.

ERG comment: The main problem with the design of the nivolumab trials is the absence of a comparator arm. No analysis can estimate the influence of bias in any outcome in these single arm trials in comparison to the outcomes of other comparator trials.

Table 4.3: Summary of CheckMate 275 and CheckMate 032 study methodology

Trial name	**CheckMate 275 (n=270)**	**CheckMate 032 (n=78)**
Location	International: 63 sites across 11 countries in North America (USA), Europe, Australia and Asia	International: 16 sites in 5 countries: Finland, Germany, Spain, UK and USA
Trial design	Multicentre, open-label, single-arm phase II study	Multicentre, open-label, multi-arm, phase I/II study[b]
Eligibility criteria for participants	Key inclusion criteria Males and females ≥18 years of age with an ECOG PS 0 or 1 Histologically or cytologically confirmed metastatic or surgically unresectable transitional cell carcinoma of the urothelium involving the bladder, urethra, ureter, or renal pelvis Measurable disease by CT or MRI per RECIST v1.1 criteria Progression or recurrence after treatment With at least 1 platinum-containing chemotherapy regimen for metastatic or surgically unresectable locally advanced urothelial cancer, or Within 12 months of peri-operative (neo-adjuvant or adjuvant) treatment with platinum agent in the setting of cystectomy for localised muscle-invasive urothelial cancer Patients that had received more than 2 prior lines of chemotherapy must not have had liver metastases Availability of tumour samples for PD-L1 expression analysis[a] Previous palliative radiotherapy must have been completed at least 2 weeks before administration of the study drug	Key inclusion criteria Males and females ≥18 years of age with an ECOG PS 0 or 1 Measurable disease by CT or MRI per RECIST v1.1 criteria Locally advanced or metastatic urothelial cell carcinoma Progression or recurrence After at least 1 previous platinum-containing chemotherapy treatment for metastatic or locally advanced unresectable urothelial cancer, or Recurrence within 1 year of completing previous platinum-based neoadjuvant or adjuvant treatment After previously refusing standard treatment with chemotherapy for the treatment of metastatic (stage IV) or locally advanced disease Key exclusion criteria Active brain metastases or leptomeningeal metastases Any serious or uncontrolled medical disorder History of or active, known or suspected autoimmune disease (vitiligo, type 1 diabetes mellitus, residual hypothyroidism caused by auto immune thyroiditis, and disorders not expected

Trial name	**CheckMate 275 (n=270)**	**CheckMate 032 (n=78)**
	Key exclusion criteria Active brain or leptomeningeal metastases Active, known or suspected autoimmune disease Previous malignancy active within the previous 3 years (except locally curable cancers that appeared to have been cured or carcinoma in situ) Any serious or uncontrolled medical disorder Autoimmune disease (vitiligo, type 1 diabetes mellitus, residual hypothyroidism due to an autoimmune condition only requiring hormone replacement, psoriasis not requiring systemic treatment, or conditions not expected to recur in the absence of an external trigger were permitted) Systemic treatment with either corticosteroids (>10 mg daily prednisone equivalents) or other immunosuppressive medications within 14 days of first study drug administration Prior treatment with an anti-PD-1, anti-PD-L1, anti-PD-L2, anti-CTLA-4 antibody, anti-CD137, or any other antibody or drug specifically targeting T-cell co-stimulation or immune checkpoint pathways Treatment with any chemotherapy, radiation therapy, biologics for cancer, or investigational therapy within 28 days of first study drug administration All toxicities attributed to previous anticancer therapy other than neuropathy, alopecia, and fatigue must have resolved to grade 1 or baseline before administration of study drug. A full list of inclusion and exclusion criteria is presented in Appendix M.	to recur in the absence of an external trigger were permitted) Need for immunosuppressive doses of systemic corticosteroids (>10 mg daily prednisone equivalents) for at least 2 weeks before study drug administration Prior treatment with experimental anti-tumour vaccines or any modulator of T-cell function or checkpoint pathway A full list of inclusion and exclusion criteria is presented in Appendix M.
Settings and locations where the data were collected	The study was conducted in a secondary care (hospital) setting at 63 sites across 11 countries worldwide The study was conducted in accordance with Good Clinical Practice guidelines by qualified investigators using a single protocol to promote consistency across sites	The study was conducted in a secondary care (hospital) setting at 16 sites across 5 countries worldwide The study was conducted in accordance with Good Clinical Practice guidelines by qualified investigators using a single protocol to promote consistency across sites

Trial name	**CheckMate 275 (n=270)**	**CheckMate 032 (n=78)**
Method of study drug administration	Nivolumab 3mg/kg Q2W via IV infusion over 60 minutes Treatment was continued until progression, unacceptable toxicity, or withdrawal of consent Patients were permitted to continue treatment beyond investigator-assessed RECIST v1.1-defined progression if they were experiencing a clinical benefit, as determined by the investigator, and were tolerating the study drug No dose modifications were allowed, but predefined dose delays were permitted for adverse events	Nivolumab 3mg/kg Q2W via IV infusion over 60 minutes Treatment was continued until progression, unacceptable toxicity, or withdrawal of consent. Patients were permitted to continue treatment beyond investigator-assessed RECIST v1.1-defined progression if they were experiencing a clinical benefit, as determined by the investigator, and were tolerating the study drug Patients could switch to nivolumab plus ipilimumab (nivolumab 1 mg/kg and ipilimumab 3 mg/kg or nivolumab 3 mg/kg and ipilimumab 1 mg/kg intravenously, every 3 weeks for four cycles) after progression if they met pre-specified criteria.
Permitted and disallowed concomitant medication	The following medications were prohibited during the study: Immunosuppressive agents (except to treat a drug-related adverse events) or systemic corticosteroids (>10 mg daily prednisone equivalent) within 14 days of study drug administration[b] Any antibody or drug specifically targeting T-cell co-stimulation or checkpoint pathways, or chemotherapy, radiation therapy, biologics for cancer, or investigational therapy within 28 days of first study drug administration	The following medications were prohibited during the study: Immunosuppressive agents (except to treat a drug-related adverse event) Systemic corticosteroids >10 mg daily prednisone equivalent[b] Any concurrent antineoplastic therapy (i.e. surgery, chemotherapy, hormonal therapy, immunotherapy, radiation therapy except for palliative radiation therapy described above or standard or investigational agents for treatment of cancer) Supportive care for disease-related symptoms was permitted to be offered to all patients on the trial. Palliative (limited-field) radiation therapy and palliative surgical resection were permitted if the certain protocol-defined criteria were met.
Primary endpoint	The primary endpoint of CheckMate 275 was BIRC-assessed ORR (as per RECIST v1.1) in the all-treated population, in patients with PD-L1 expression ≥1%, and in patients with PD-L1 expression ≥5% ORR was defined as the number of patients with a best overall response (BOR) of confirmed complete response (CR) or partial response (PR) divided by the number of all-treated patients, PD-L1 ≥1% patients or PD-L1 ≥5% subjects, respectively	The primary endpoint of CheckMate 032 was confirmed investigator-assessed ORR ORR was defined as the number of patients with a BOR of CR or PR as per RECIST v1.1 divided by the number of treated patients

Trial name	**CheckMate 275 (n=270)**	**CheckMate 032 (n=78)**
Secondary and exploratory endpoints	**Secondary endpoints:** BIRC-assessed PFS OS Investigator-assessed ORR (in the all-treated population, patients with PD-L1 expression ≥1%, and patients with PD-L1 expression ≥5%) **Exploratory endpoints:** Investigator-assessed PFS Safety HRQoL via the EORTC QLQ-C30 questionnaire General health status via the EQ-5D-3L questionnaire Pharmacokinetics and exploration of exposure-response relationships* Immunogenicity* Pharmacodynamic activity in the peripheral blood and tumour tissue as measured by flow cytometry, immunohistochemistry, soluble factor analysis, and gene expression (microarray technology, quantitative RT-PCR)* Association between biomarkers in the peripheral blood and tumour tissue with safety and efficacy* **Outcomes not considered relevant to present in this submission*	**Secondary endpoints:** Investigator-assessed PFS OS DOR Safety **Exploratory endpoints:** Assessed by PD-L1 expression (≥1% and <1%): ORR OS PFS HRQoL via the EQ-5D and EQ-VAS questionnaires
Timing of assessments	Tumour assessments were scheduled at 8 weeks from the date of first dose (±1 week), then every 8 weeks (±1 week) thereafter up to 48 weeks, then every 12 weeks (±1 week) until documented disease progression or treatment discontinuation (whichever occurred last). Assessments were performed using CT or MRI and included the pelvis, chest, abdomen and all known sites of disease Survival assessment was scheduled every 3 months until death, lost to follow-up or withdrawal of study consent AEs were assessed during treatment visits and were included in the safety analyses if they occurred within 30 days from the day of the last dose received	Treated subjects were evaluated for response by the investigator according to the RECIST v1.1 at baseline and then every 6 weeks (±1 week) from first dose for the first 24 weeks, then every 12 weeks (±1 week) until disease progression or treatment was discontinued (whichever occurred later) Assessments were performed using CT or MRI and included the pelvis, chest, abdomen and all known sites of disease AEs were assessed during treatment visits. Safety was defined as the incidence of treatment-related adverse events leading to drug discontinuation within the first 12 weeks of treatment in patients who had at least one dose of study drug HRQoL was assessed before study drug administration through Week 13, then

Trial name	**CheckMate 275 (n=270)**	**CheckMate 032 (n=78)**
	HRQoL and general health status were assessed before each dose at Week 1, then every 8 weeks up to 48 weeks, then every 12 weeks until disease progression or treatment discontinuation (whichever occurred later) Two follow-up visits and subsequent survival follow-up visits were also scheduled for AEs and HRQoL measures[c]	at the same time of subsequent tumour assessments, during Follow-Up Visit 1 and 2 and survival visits Two follow-up visits and subsequent survival follow-up visits were also scheduled (AEs and HRQoL)[c]
Pre-planned subgroups	A pre-planned analysis of the primary and secondary endpoints in patients with PD-L1 expression <1% and ≥1% was conducted Further subgroup analyses were conducted to assess the impact of pre-specified baseline characteristics, site of original tumour origin (bladder, renal pelvis/ureter), number of Bellmunt risk factors, and prior cancer therapy regimens (number of prior regimens in a metastatic setting, time from completion of most recent prior regimen to study treatment) on confirmed ORR per BIRC, PFS and OS	As part of the exploratory endpoints, ORR, OS and PFS were analysed in subgroups defined by PD-L1 expression (<1% and ≥1%). In addition, ad-hoc subgroup analyses were conducted to assess the impact several key baseline factors such as ECOG-PS, metastases, or haemoglobin on investigator-assessed ORR
Duration of study and follow-up	The first patient was treated on the 9th March 2015 and the trial is currently ongoing. The last patient last visit date for the primary database lock of the 30th May 2016, data from which are presented in this submission, was the 15th April 2016. The median follow-up for OS was 11.5 months. A further database lock took place on 2nd September 2016 and data from this are also presented in this submission.	The first patient was treated on the 5th June 2014 and the trial is currently ongoing. The last patient last visit date for the primary database lock of 24th March 2016 was the 11th February 2016, data from which are presented in this submission. The median follow-up for OS was 9.69 months.

Source: CS, Table 5, pages 30-35

[a]Patients were required to have an evaluable tumour tissue sample for PD-L1 expression testing at screening, but were not excluded based on PD-L1 status. [b]Several advanced or metastatic solid tumour types were studied in CheckMate 032, but only the urothelial carcinoma arm treated with nivolumab monotherapy is presented in this submission. [c]Patients were followed for at least 100 days after the last dose of study drug. Follow-up Visit 1 was scheduled for 35 days from the last dose ±7 days or coincided with the date of discontinuation (± 7 days) if date of discontinuation was >35 days after last dose. Follow-up Visit 2 was scheduled for 80 days (±7 days) from follow-up Visit 1. Survival follow-up visits were scheduled for every 3 months (± 7 days) from Follow-up Visit 2.

AEs = adverse events; BIRC = blinded independent review committee; BOR = best overall response; CR = complete response; CT = computer tomography; CTLA-4 = cytotoxic T-lymphocyte-associated protein 4; ECOG = Eastern Cooperative Oncology Group; EORTC QLQ-C30 = European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30; EQ-5D-3L = 3-level EuroQoL 5-Dimensions; GCP = Good Clinical Practice; HRQoL = health-related quality of life; IV = intravenous; MRI = magnetic

Trial name	**CheckMate 275 (n=270)**	**CheckMate 032 (n=78)**
resonance imaging; ORR = objective response rate; OS = overall survival; PD-1 = programmed death 1; PD-L1 = programmed death ligand 1; PD-L2 = programmed death ligand 2; PFS = progression-free survival; PR = partial response; PROs = patient-reported outcomes; PS = performance status; RECIST = response evaluation criteria in solid tumours.		

4.2.2 Baseline characteristics of the nivolumab studies

Baseline demographics, disease characteristics and a summary of prior therapies of the patients included in CheckMate 275 and CheckMate 032 are presented in Table 4.4.

In CheckMate 275, median age was 66 years, the majority of patients were white and male, and over 70% were current or former smokers. The vast majority of patients (96.7%) had metastatic disease. Overall 71.5% of patients had received at least one prior regimen in the metastatic disease setting, and 29.3% had received two or more prior regimens for metastatic disease. Prior systemic cancer therapy was less common in the neoadjuvant and adjuvant settings, with 22.2% receiving at least one neoadjuvant regimen and 30.7% of patients receiving prior regimen(s) in the adjuvant setting.

The median age of the patient population in CheckMate 032 was 66 years; the majority were white (92.3%) and male (69.2%). The vast majority (91%) of patients had metastatic (stage IV) disease, and 75.6% of patients had at least two disease sites.

The company provided the following additional information based on feedback from clinical experts: *'Expert clinician feedback was that the patient populations of CheckMate 275 and CheckMate 032 were very similar, and could be considered generally representative of the patient population expected to receive nivolumab in UK clinical practice. Across both trials, expert clinician feedback was that the proportion of patients with PS 0 was perhaps slightly over-representative of the number of patients likely to have PS 0 in this setting, and that the median age of the patients in both trials may be slightly lower than the age of the average UC patient treated in the second-line setting in UK clinical practice. However, a recent chart review conducted in UK clinical practice of patients with locally advanced unresectable or metastatic UC initiating second-line therapy found that the mean patient age was in fact very similar, albeit slightly lower (mean of 62.8 years), than in both CheckMate trials.'*[5]

In response to the clarification request, the company stated that there were no UK sites in CheckMate 275 and in CheckMate 032, there were 6 patients (7.7%) treated in the study in the UK.[7]

ERG comment:
There are serious questions regarding the representativeness of the nivolumab trial patients to the UK population. Firstly, almost no patients in the UK were included and none in the largest trial (CheckMate 275).[10] Secondly, in response to the clarification request, the company confirmed that as few as 18.8% of patients in the UK might have and ECOG PS of 0, as opposed to over 50% in the two nivolumab trials.[7] Thirdly, there is a mismatch in terms of prior therapies, as confirmed in Table 8 of the response to clarification, which shows that, in a chart review, as many as over 75% of patients in the UK would have previously taken a gemcitabine platinum-based combination compared to fewer than 40% in the trials.[7] Finally, there is a question of the applicability to those with locally advanced unresectable as opposed to metastatic disease given the very small proportion of such patients in the trials. The company stated in the response to clarification that type of disease in these terms was not prognostic given no mention of this at their advisory board. However, lack of comment at the advisory board does not mean that clinical experts do not believe this to be the case.

Table 4.4: Baseline characteristics of patients in the all-treated population of CheckMate 275 and CheckMate 032

	CheckMate 275	CheckMate 032
Characteristic	**Total (n=270)**	**Total (n=78)**
Demographics		
Age, median years (range)	66 (38–90)	66 (31–85)
Age categorisation, n (%)		
<65	122 (45.2)	37 (47.4)
≥65 and <75	110 (40.7)	31 (39.7)
≥75 and <85	35 (13.0)	N/A
≥75	N/A	10 (12.8)
>85	3 (1.1)	N/A
Male, n %	211 (78.1)	54 (69.2)
Race, n %		
White	231 (85.6)	72 (92.3)
Asian	30 (11.1)	1 (1.3)
Black	2 (0.7)	4 (5.1)
Other	3 (1.1)	1 (1.3)
Not reported	4 (1.5)	N/A
Region, n (%)		
US	106 (39.3)	59 (75.6)
Japan	23 (8.5)	0 (0.0)
Rest of world	141 (52.2)	19 (24.4)
Tobacco use, n (%)		
Current/former smoker	194 (71.9)	48 (61.5)
Never smoked	67 (24.8)	29 (37.2)
Unknown	9 (3.3)	1 (1.3)
Disease characteristics		
ECOG PS, n (%)		
0	145 (53.7)	42 (53.8)
1	124 (45.9)	36 (46.2)
3	1 (0.3)	0
Bellmunt risk factors, n (%)		
0	98 (36.3)	27 (34.6)
1	111 (41.1)	39 (50.0)
2	46 (17.0)	8 (10.3)
3	15 (5.6)	4 (5.1)
Site of primary tumour, n (%)		
Urinary bladder	197 (73.0)	NR
Renal pelvis	46 (17.0)	NR

	CheckMate 275	CheckMate 032
Characteristic	**Total (n=270)**	**Total (n=78)**
Ureter	19 (7.0)	NR
Urethra	8 (3.0)	NR
Disease setting, n (%)		
Metastatic	261 (96.7)	71 (91.0)
Locally unresectable/non-metastatic	9 (3.3)	7 (9.0)
Baseline metastases, n (%)		
Any visceral involvement	227 (84.1)	61 (78.2)
Liver	75 (27.8)	20 (25.6)
Lymph node only	43 (15.9)	13 (16.7)
PD-L1 expression, n (%)		
Assessable	N/A	67 (85.9)
<1%	N/A	42 (53.8)
≥1%	124 (45.9)	25 (31.8)
<5%	N/A	53 (67.9)
≥5%	83 (30.7)	14 (17.9)
Number of sites with ≥1 lesion, n (%)		
1	85 (31.5)	19 (24.4)
2	94 (34.8)	30 (38.5)
3	51 (18.9)	24 (30.8)
4	29 (10.7)	3 (3.8)
≥5	11 (4.1)	2 (2.6)
Prior therapy		
Prior systemic therapy regimen setting, n (%)		
Metastatic	193 (71.5)	N/A
Adjuvant	83 (30.7)	33 (42.3)
Neo-adjuvant	60 (22.2)	14 (17.9)
Previous therapies in metastatic setting, n (%)		
0	77 (28.5)	N/A
1	114 (42.2)	26 (33.3)
2	57 (21.2)	N/A
2-3	N/A	42 (53.8)
>3	N/A	10 (12.8)
≥3	22 (8.1)	N/A
Prior surgery related to cancer, n (%)	250 (92.6)	71 (91.0)
Prior radiotherapy, n (%)	85 (31.5)	25 (32.1)

Source: CS, Table 5, pages 35-37
ECOG PS = Eastern Cooperative Oncology Group performance status; N/A = not applicable; NR = not reported; PD-L1 = programmed death ligand 1.

4.2.3 Statistical analyses in the nivolumab studies

The statistical analyses used for the primary and secondary endpoints alongside sample size calculations and methods for handling missing data are summarised in Table 4.5.

ERG comment: The ERG believes that the statistical methods used within the nivolumab studies were appropriate. The ERG notes that the primary design of CheckMate 275 was to evaluate ORR based on assessments of nivolumab monotherapy in patients with tumour expressing PD-L1 (membranous staining in ≥ 5% and ≥ 1% tumour cells) and overall patients. CheckMate 32 was primarily designed to evaluate the ORR of nivolumab monotherapy in patients with advanced or metastatic UC. Neither study design was appropriate for comparative analysis.

Table 4.5: Statistical methods for the primary analysis of CheckMate 275 and CheckMate 032

Trial name	CheckMate 275	CheckMate 032
Hypothesis objective	Treatment with nivolumab monotherapy would lead to clinical benefit in patients with metastatic or surgically unresectable UC who have progressed post platinum treatment as demonstrated by a clinically meaningful ORR	Treatment with nivolumab monotherapy will have clinical activity in subjects with advanced or metastatic tumours
Statistical analysis	ORRs (both BIRC- and investigator-assessed) were summarised by a binomial response rate and their corresponding two-sided 95% exact CIs using the Clopper-Pearson method.**[**CS REF 45**]** BOR was summarised by response category Median values of DOR were calculated along with two-sided 95% CI using Brookmeyer and Crowley method.**[**CS REF 46**]** TTR was summarised using descriptive summary statistics for the responders Time-to-event distributions were estimated using Kaplan-Meier techniques. This was done for PFS, OS and DOR (note that time to response was analysed using summary statistics such as mean, SD, median, min, max). Median survival time along with 95% CIs were constructed based on a log-log transformed CI for the survivor function S(t)**[**CS REF 46+47**]** Rates at fixed time points were derived from the Kaplan-Meier estimate and corresponding confidence interval were derived based on Greenwood formula**[**CS REF 48**]** for variance derivation and on log-log transformation applied on the survivor function S(t)**[**CS REF 49**]**	ORR was summarised by a binomial response rate and corresponding two-sided 95% exact CI using the Clopper-Pearson method. Time-to-event distributions (DOR, PFS and OS) were estimated using Kaplan-Meier techniques When appropriate, the median along with 95% CI was provided using Brookmeyer and Crowley methodology (using the log-log transformation for construction of CIs). Rates at fixed time points (e.g. OS at 12 months) were derived from the Kaplan Meier estimate along with their corresponding log-log transformed 95% CIs.
Sample size, power calculation	The primary objective was to estimate ORR as per BIRC assessment for: • All treated patients • Patients with PD-L1 expression ≥1% • Patients with PD-L1 expression ≥5%	The primary objective was to estimate investigator-assessed ORR An ORR of 10% or less was considered not of clinical value, and an ORR of 25% or greater was considered of strong clinical interest A sample size of 60–100 treated subjects would provide 90% to 97% power to reject the null hypothesis of 10% response rate if

Trial name	CheckMate 275	CheckMate 032
	For all treated patients, a sample size of 242 would provide 90% power to reject the null hypothesis that ORR was 10% at a two-sided 5% type I error if the true ORR in this population was 16.9%. Assuming ORR is 30%, 70 treated patients with PD-L1 expression ≥5% would provide 99.1% power at 5% type 1 error to reject the null hypothesis of a two-sided test that the true ORR was 10%, based on historical control data for single-agent chemotherapy,**[**CS REF 34, 35, 50**]** a threshold below which was considered not clinically meaningful in this population, and 90% power at 5% type I error to reject the null hypothesis of a two-sided test that the true ORR was 14.7%. Under the assumption of 32% prevalence rate of PD-L1 ≥5% among all PD-L1 evaluable patients, approximately up to 220 PD-L1 evaluable patients would be treated. Assuming an additional 10% of treated patients with PD-L1 indeterminate status, the total sample size was expected to be approximately 242. Under the assumption of 50% prevalence rate of PD-L1 ≥1% among all PD-L1 evaluable patients, approximately up to 110 patients with PD-L1 expression ≥1% would be treated. This would provide 90% power to reject the null hypothesis that ORR was 10% at a two-sided 5% type 1 error if the true ORR in this population was 20.6%.	the true response rate was 25% with a two-sided Type I error rate of 5%
Data management, patient withdrawals	The final analysis of the primary endpoint ORR (based on BIRC assessments) was to be performed six months after approximately 70 patients with PD-L1 expression of ≥5% had been treated (i.e. six months after last patient first treatment)	All 78 patients who received at least one dose of nivolumab were included in the safety and efficacy analyses

Source: CS, Table 5, pages 38-40

BOR = best overall response; CI = confidence interval; ORR = overall response rate; PD-L1 = programmed death ligand 1; TTR = time to response.

4.2.4 Quality assessment of the nivolumab studies

The company considered the CheckMate 275 and CheckMate 032 studies to be of satisfactory quality based on the CRD cohort study checklist.[22]

ERG comment: The ERG considers both studies as low-level evidence in the hierarchy of clinical study designs, and not suitable for comparisons with other interventions.

4.2.5 Results of the nivolumab studies

CheckMate 275

The primary endpoint in Checkmate 275 was ORR (based on BIRC assessments) and the primary database lock was 30 May 2016. The company responded to the clarification request by stating that the next database locks for CheckMate 275 and CheckMate 032 in [redacted] and [redacted], respectively.[7]

Treatment with nivolumab led to a confirmed objective response per blinded independent review committee (BIRC) in a total of 52 (19.6%) patients (95% CI: 15.0 to 24.9) and 6 (2.3%) patients achieved a complete response (CR) (see Table 4.6). Patients in the PD-L1≥1% cohort achieved an objective response rate (ORR) of 23.8% (95% CI: 16.5 to 32.3) and patients with <1% PD-L1 expression had a confirmed ORR of 16.1% (15.8% at the second database lock).

As reported in Sharma et al. (2017),[8] 177 high-quality gene expression profiles have been generated from patients' tumour tissues. Higher values of the 25-gene interferon-γ signature were associated with a greater proportion of responders to nivolumab and higher PD-L1 expression. Patients with high interferon-γ signature were more likely to respond to nivolumab than were those with low interferon-γ signature (p=0. 0003).

Time to response (TTR) and duration of response (DOR) were estimated in patients with a confirmed partial response (PR) or complete response (CR). Median TTR as per BIRC was 1.87 months (interquartile range (IQR): 1.81 to 1.97 months) and the majority of responders achieved their response at the time of first tumour assessment (Week 8).

At the time of the clinical database lock (30 May 2016), median DOR as per BIRC had not been reached in the efficacy-treated population and across the <1% and ≥1% PD-L1 subgroups. The majority of responders (76.9%) were still continuing to respond and [redacted] of patients had a DOR of at least three months.

Table 4.6: Primary efficacy results of CheckMate 275

Tumour response	Efficacy-treated population (n=265)	PD-L1 <1% (n=143)	PD-L1 ≥1% (n=122)
ORR, n (%) **95% CI**	52 (19.6) 95% CI: 15.0–24.9	23 (16.1) 95% CI: 10.5–23.1	29 (23.8) 95% CI: 16.5–32.3
BOR			
CR	6 (2.3)	1 (0.7)	5 (4.1)
PR	46 (17.4)	22 (15.4)	24 (19.7)
SD	60 (22.6)	25 (17.5)	35 (28.7)
PD	104 (39.2)	67 (46.9)	37 (30.3)
Unable to determine[a]	49 (18.5)	28 (19.6)	21 (17.2)
Median TTR (n=52), months; IQR	1.87	1.94	1.87

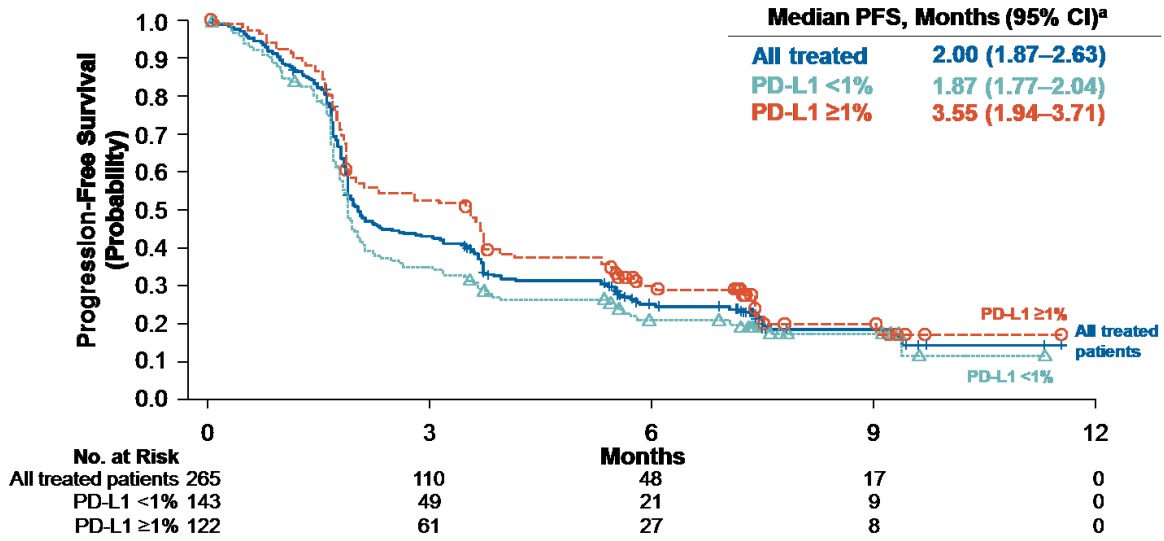
Tumour response	Efficacy-treated population (n=265)	PD-L1 <1% (n=143)	PD-L1 ≥1% (n=122)
	IQR: 1.81–1.97	IQR: 1.81–2.10	IQR: 1.81–1.97
Median DOR (n=52), months; 95% CI	NR 95% CI: 7.43–NR	NR 95% CI: 7.43–NR	NR 95% CI: 7.52–NR

Source: CS, Table 12, page 43-44
[a]BOR was reported as unable to determine in 49 patients (18.5%); main reasons were because the patient had died or started subsequent therapy before the first scan visit at Week 8.
BOR = best overall response; CI = confidence intervals; CR = complete response; DOR = duration of response; IQR = interquartile range; ORR = objective response rate; PD = progressive disease; PD-L1 = programmed death ligand 1; PR = partial response; SD = stable disease; TTR = time to response NR = not reached.

At the time of the primary clinical database lock (30 May 2016), 201 patients (75.8%) had experienced a PFS event. Median PFS in the efficacy-treated population was 2.00 months (95% CI: 1.87 to 2.63), and the PFS rates at three and six months were 43.1% (95% CI: 37.0 to 49.1) and 25.2% (95% CI: 20.0 to 30.8), respectively. Median PFS for patients in the PD-L1 ≥1% cohort was longer than in the all-treated population at 3.55 months (95% CI: 1.94 to 3.71), and in the PD-L1 <1%, median PFS was 1.87 months (95% CI: 1.77 to 2.04) (see Figure 4.1).

Results for investigator-assessed ORR were investigated as a secondary outcome and the results were consistent with BIRC-assessed ORR. A total of [redacted] patients ([redacted]) achieved an objective response of which [redacted] patients ([redacted]) achieved a CR.

Figure 4.1: Kaplan-Meier plot for progression-free survival in CheckMate 275

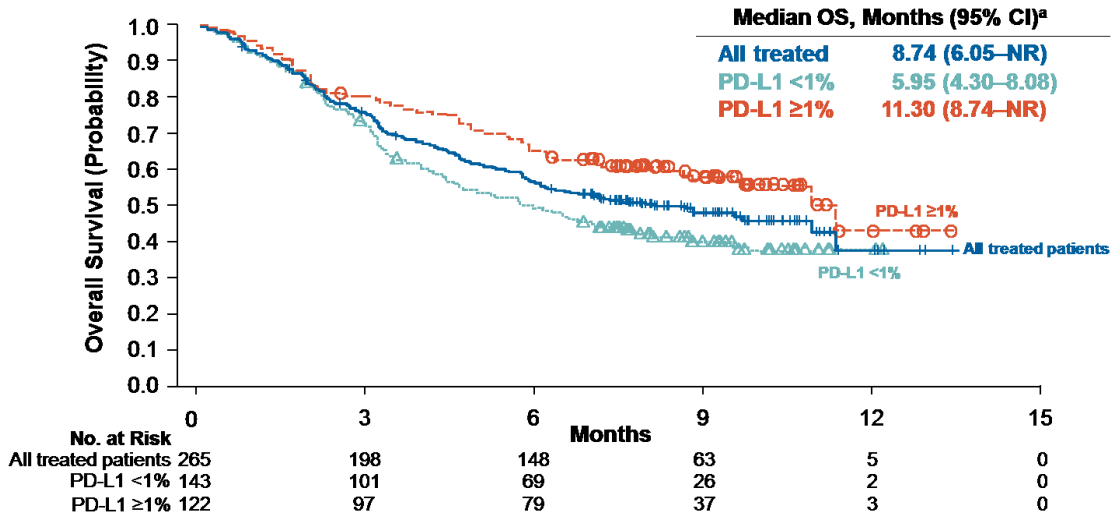


Source: CS, Figure 11, page 46
CI = confidence interval; PD-L1 = programmed death ligand 1; PFS = progression-free survival.

Median follow-up for overall survival (OS) (time between first dose and last known date alive or death) was 7.00 months (IQR: 2.96 to 8.77 months). At the primary analysis database lock (30 May 2016), 138 patients (51.1%) had died. Median OS in the efficacy-treated population was 8.74 months (95% CI: 6.05 to N/A); three-month and six-month OS rates were 75.8% (95% CI: 70.2 to 80.5) and 57.0% (95% CI: 50.7 to 62.7).

The Kaplan-Meier plot for OS is presented in Figure 4.2. Median OS for patients in the PD-L1 ≥1% cohort was longer than in the all-treated population at 11.30 months (95% CI: 8.74 to NR), and in the PD-L1 <1%, median OS was 5.95 months (95% CI: 4.30 to 8.08).

Figure 4.2: Kaplan-Meier plot for overall survival in CheckMate 275



Source: CS, Appendix E, Figure 26, page 146
CI = confidence interval; OS = overall survival; PD-L1 = programmed death ligand 1.

Results from the second database lock of CheckMate 275 (2 September 2016) were consistent with those from the primary analysis database lock in terms of ORR, PFS and OS. In total, 54 patients (20.0%) had achieved an ORR (95% CI: 15.4 to 25.3), and two more patients had achieved a CR. Median DOR was 10.35 months (95% CI: 7.52 to NR). A further six patients had died, taking the total to 154 (57%). A comparison of the main results between database locks and trials is shown in Table 11 of the CS and reproduced in Table 4.7. There also continued to be a statistically signification difference in median OS between PD-L1 <1% and PD-L1 >= 1% (5.95 months (95% CI: 4.37 to 8.08), and in the PD-L1 <1%, median OS was 11.63 months (95% CI: 9.10 to NA).

Table 4.7: Overview of clinical effectiveness results from CheckMate 275 and CheckMate 032

Outcome	CheckMate 275		CheckMate 032
	Initial database lock: 30 May 2016 n=265[c]	Latest database lock: 2 Sep 2016 n=270[c]	n=78
ORR, n (%), [95% CI]	52 (19.6), [15.0–24.9]	54 (20.0), [15.4–25.3][b]	19 (24.4) [15.3–35.4]
TTR, median (IQR), months	1.87 (1.81–1.97)[a]	1.94 (1.84–2.50)[b]	1.48 (1.25–4.14)
DOR, median (95% CI), months	NR (7.43–NR)[a]	10.35 (7.52–NR)[b]	NR (9.92–NR)
PFS, median (95% CI), months	2.00 (1.87–2.63)[a]	2.00 (1.87–2.63)[b]	2.78 (1.45–5.85)
OS, median (95% CI), months	8.74 (6.05–NR)[a]	8.57 (6.05–11.27)[b]	9.72 (7.26–16.16)
Source: CS, Table 11, page 43			

[a]Minimum follow-up of 6 months from the date of first dose. [b]Minimum follow-up of 8.3 months. [C]Follow-up for the latest database lock was sufficient to include 5 patients from Japan who were not included in efficacy analyses in the initial database lock.
CI = confidence intervals; DOR = duration of response; NR = not reached.ORR = objective response rate; OS = overall survival; PFS =progression free survival; TTR = time to response

Patient-reported outcomes data for the measurement of HRQoL was assessed via the EORTC QLQ-C30 questionnaire and the EQ-5D-3L questionnaire in CheckMate 275. Due to the limited study follow-up, interpretations of EORTC QLQ-30 results are limited to the first 41 weeks of follow-up for the all-treated population. Overall, patient HRQoL continued to increase or was maintained throughout the trial from baseline to Week 41.

The mean baseline EQ-5D VAS score was 60.2, and mean scores were higher at Week 9 on treatment (67.5). By Week 41, the average EQ-5D VAS was more than 80 points. However, by this time data was based on only n=24 patients.

A total of [redacted] of patients received ≥90% of the planned nivolumab dose intensity, and the median number of doses received was [redacted] (range: [redacted]). The median duration of therapy was [redacted] months. At the time of the 30 May 2016 database lock, 75.6% of patients had discontinued treatment with nivolumab. The most common reasons for discontinuation were disease progression (53.3%), AEs unrelated to nivolumab (12.6%), and nivolumab toxicity (5.2%).

A summary of the safety results from CheckMate 275 and CheckMate 032 is presented in Table 4.8. The majority of treated patients experienced at least one AE regardless of causality, during treatment with nivolumab or within 30 days of the last nivolumab dose. As of their respective clinical database locks, a total of 138 (51.5%) patients and 36 (46.2%) patients in the CheckMate 275 and CheckMate 032 trials had died, respectively. The proportion of deaths due to study drug toxicity was 1.1% and 3%, respectively. All-cause AEs leading to treatment discontinuation were reported in 20.7% and 7.7% of patients in CheckMate 275 and CheckMate 032, respectively.

Table 4.8: Summary of safety analysis in CheckMate 275 and CheckMate 032

Adverse event, n (%)	**CheckMate 275 (n=270)[a]**		**CheckMate 032 (n=78)[b]**	
Deaths	138 (51.1)		36 (46.2)	
Deaths due to study drug toxicity	3 (1.1)[c]		2 (2.6)[d]	
	Any grade	Grade 3-4	Any grade	Grade 3-4
All causality AEs	267 (98.9)	137 (50.7)	78 (100)	43 (55.1)
Drug-related AEs	174 (64.4)	48 (17.8)	65 (83.3)	18 (23.1)
All-causality serious AEs	147 (54.4)	99 (36.7)	36 (46.2)	23 (29.5)
Drug-related serious AEs	[redacted]	[redacted]	8 (10.3)	[redacted]
All-causality AEs leading to treatment discontinuation	56 (20.7)	42 (15.6)	6 (7.7)	4 (5.1)
Drug-related AEs leading to treatment discontinuation	13 (4.8)	8 (3.0)	2 (2.6)	2 (2.6)

Source: CS, Table 23, page 72-73
[a] AEs were coded using the MedDRA version 19.0 and were graded for severity according to the NCI CTCAE version 4.0. [b] AEs were coded using the MedDRA version 18.1 and were graded for severity according to the NCI CTCAE version 4.0. C Three deaths (Grade 5 pneumonitis, Grade 5 acute respiratory failure, and Grade

5 cardiovascular failure) were judged as study drug-related. [d] Two deaths (Grade 4 pneumonitis and Grade 4 thrombocytopenia) were assessed as study drug-related.
AEs = adverse events; MedDRA = Medical Dictionary for Regulatory Activities; NCI CTCAE = National Cancer Institute Common Terminology Criteria for Adverse Events; SAEs = serious adverse events.

Treatment-related adverse events occurred in 174 (64%) of 270 patients. The most common treatment-related adverse event of any grade was fatigue, which was noted in 45 patients (17%). Grade 3 or 4 treatment-related adverse events occurred in 48 patients (18%) – most commonly grade 3 fatigue and diarrhoea, each of which occurred in five patients (Table 4.9). Thirteen patients (5%) discontinued treatment because of nivolumab toxicity, including four (1%) from pneumonitis, two (1%) from pemphigoid, and one each (<1%) from dyspnoea, interstitial lung disease, maculopapular rash, pruritic rash, abdominal pain, diarrhoea, and circulatory collapse. The most common treatment-related select (immuno mediated) adverse events (any grade) were skin (47 [17%]) and endocrine (39 [14%]). Most select adverse events resolved and were manageable with immune-modulating drugs (mostly systemic corticosteroids; data not shown). Some drug-related endocrinopathies were not deemed to be resolved because of ongoing hormone replacement therapy.[8]

Of the 270 patients in the safety population, 138 deaths (51%) were reported, of which 121 (88%) were due to disease progression. Of the 53 patients who died within 30 days of their last nivolumab dose, 39 (74%) died of disease progression. Of the 14 deaths not related to disease progression, 11 were attributed to other reasons and three were attributed by investigators to treatment, all of which occurred in patients with metastatic disease. One patient died of pneumonitis, one of acute respiratory failure, and one of cardiovascular failure.[8]

Table 4.9: Drug-related adverse events in ≥5% patients in CheckMate 275 and CheckMate 032

Adverse event	CheckMate 275 (n=270)		CheckMate 032 (n=78)	
	Any grade	Grade 3-4	Any grade	Grade 3-4
Total patients with an event	174 (64.4)	48 (17.8)[a]	65 (83.3)	18 (23.1)[b]
General disorders and administration site conditions	80 (29.6)	10 (3.7)	29 (37.2)	2 (2.6)
Fatigue	45 (16.7)	5 (1.9)	28 (35.9)	2 (2.6)
Asthenia	16 (5.9)	4 (1.5)	N/A	N/A
Pyrexia	15 (5.6)	0 (0.0)	N/A	N/A
Gastrointestinal disorders	54 (20.0)	7 (2.6)	24 (30.8)	2 (2.6)
Diarrhoea	24 (8.9)	5 (1.9)	7 (9.0)	0 (0.0)
Nausea	19 (7.0)	1 (0.4)	10 (12.8)	1 (1.3)
Skin and subcutaneous tissue disorders	54 (20.0)	6 (2.2)	34 (43.6)	3 (3.8)
Pruritus	25 (9.3)	0 (0.0)	23 (29.5)	0 (0.0)
Rash	16 (5.9)	3 (1.1)	5 (6.4)	0 (0.0)
Rash maculo-papular	N/A	N/A	14 (7.9)	2 (2.6)
Dry skin	N/A	N/A	5 (6.4)	0 (0.0)
Investigations	N/A	N/A	26 (33.3)	8 (10.3)
Lipase increased	N/A	N/A	11 (14.1)	4 (5.1)
Amylase increased	N/A	N/A	7 (9.0)	3 (3.8)
Lymphocyte count decreased	N/A	N/A	5 (6.4)	2 (2.6)

Adverse event	CheckMate 275 (n=270)		CheckMate 032 (n=78)	
	Any grade	Grade 3-4	Any grade	Grade 3-4
Blood creatinine increased	N/A	N/A	4 (5.1)	0 (0.0)
Endocrine disorders	31 (11.5)	1 (0.4)	6 (7.7)	0 (0.0)
Hypothyroidism	21 (7.8)	0	4 (5.1)	0 (0.0)
Musculoskeletal and connective tissue disorders	N/A	N/A	13 (16.7)	1 (1.3)
Arthralgia	N/A	N/A	9 (11.5)	0 (0.0)
Metabolism and nutrition	27 (10.0)	3 (1.1)	10 (12.8)	2 (2.6)
Decreased appetite	22 (8.1)	0	5 (6.4)	0 (0.0)
Hyperglycaemia	N/A	N/A	5 (6.4)	1 (1.3)
Blood and lymphatic system disorders	N/A	N/A	11 (14.1)	1 (1.3)
Anaemia	N/A	N/A	8 (10.3)	0 (0.0)
Respiratory, thoracic and mediastinal disorders	N/A	N/A	11 (14.1)	1 (1.3)[b]
Dyspnoea	N/A	N/A	6 (7.7)	2 (2.6)
Nervous system disorders	N/A	N/A	7 (9.0)	0 (0.0)

Source: CS, Table 23, page 74-75

[a]Grade 5 events reported in 3 (1.1%) patients (1 death due to pneumonitis, 1 death due to acute respiratory failure, 1 death due to cardiovascular failure). [b] 1 (1.3%) Grade 5 drug-related AE (pneumonitis).

AEs = adverse events; N/A = not applicable.

Select AEs were defined as AEs of special clinical interest that are potentially associated with the use of nivolumab, and were identified based on the following principles:

- AEs that may differ in type, frequency, or severity from AEs caused by non-immunotherapies
- AEs that may require immunosuppression (e.g. Corticosteroids) as part of their management
- AEs whose early recognition and management may mitigate severe toxicity
- AEs for which multiple event terms may be used to describe a single type of AE, thereby necessitating the pooling of terms for full characterisation.

Considering the AEs already observed across other studies of nivolumab therapy, the AEs considered as select AEs were endocrinopathies, diarrhoea/colitis, hepatitis, pneumonitis, interstitial nephritis, rash and hypersensitivity/infusion reactions.

Most select AEs were considered drug-related by the investigator, with the exception of hepatic and renal events, where a lower proportion of select AEs were deemed to be drug-related. The most frequently reported any-grade drug-related select AE categories were skin (17.4%) and endocrine (14.4%) – see Table 4.10.

Table 4.10: Drug-related select adverse events in CheckMate 275 and CheckMate 032

Select adverse event, n (%)	CheckMate 275		CheckMate 032	
	Any grade	Grade 3–4	Any grade	Grade 3–4
Total patients with an event, by category				
Skin	47 (17.4)	4 (1.5)	33 (42.3)	2 (2.6)
Endocrine	39 (14.4)	1 (0.4)	6 (7.7)	0 (0.0)
Gastrointestinal	25 (9.3)	6 (2.2)	8 (10.3)	1 (1.3)
Hepatic	10 (3.7)	5 (1.9)	4 (5.1)	1 (1.3)
Pulmonary	11 (4.1)	3 (1.1)	2 (2.6)	0 (0.0)
Renal	3 (1.1)	1 (0.4)	7 (9.0)	1 (1.3)
Hypersensitivity/infusion reactions	3 (1.1)	1 (0.4)	2 (2.6)	0 (0.0)
Drug-related 'select' AEs, by category				
Skin				
Pruritis	25 (9.3)	0 (0.0)	23 (29.5)	0 (0.0)
Rash	16 (5.9)	3 (1.1)	5 (6.4)	0 (0.0)
Rash maculo-papular	4 (1.5)	1 (0.4)	14 (17.9)	2 (2.6)
Erythema	2 (0.7)	0 (0.0)	N/A	N/A
Pruritis generalised	2 (0.7)	1 (0.4)	N/A	N/A
Rash macular	2 (0.7)	0 (0.0)	N/A	N/A
Rash pruritic	2 (0.7)	1 (0.4)	N/A	N/A
Rash erythematous	N/A	N/A	2 (2.6)	0 (0.0)
Rash papular	N/A	N/A	1 (1.3)	0 (0.0)
Palmar-plantar erythrodysaesthesia syndrome	N/A	N/A	1 (1.3)	0 (0.0)
Blister	1 (0.4)	0 (0.0)	N/A	N/A
Dermatitis	1 (0.4)	0 (0.0)	N/A	N/A
Eczema	1 (0.4)	0 (0.0)	N/A	N/A
Rash generalised	1 (0.4)	0 (0.0)	N/A	N/A
Skin exfoliation	1 (0.4)	0 (0.0)	N/A	N/A
Skin irritation	N/A	N/A	1 (1.3)	0 (0.0)
Urticaria	1 (0.4)	0 (0.0)	N/A	N/A
Endocrine				
Thyroid disorder	35 (13.0)	0 (0.0)	6 (7.7)	0 (0.0)
Hypothyroidism	21 (7.8)	0 (0.0	4 (5.1)	0 (0.0)
Hyperthyroidism	11 (4.1)	0 (0.0	3 (3.8)	0 (0.0)
Blood thyroid stimulating hormone increased	10 (3.7)	0 (0.0	1 (1.3)	0 (0.0)
Blood thyroid stimulating hormone decreased	5 (1.9)	0 (0.0	N/A	N/A
Thyroiditis	2 (0.7)	0 (0.0	N/A	N/A
Thyroxine increased	2 (0.7)	0 (0.0	N/A	N/A

Select adverse event, n (%)	CheckMate 275		CheckMate 032	
	Any grade	Grade 3–4	Any grade	Grade 3–4
Autoimmune thyroiditis	1 (0.4)	0 (0.0	N/A	N/A
Thyroxine decreased	1 (0.4)	0 (0.0	N/A	N/A
Thyroxine free increased	1 (0.4)	0 (0.0	N/A	N/A
Adrenal disorder	2 (0.7)	0 (0.0)	N/A	N/A
Adrenal insufficiency	2 (0.7)	0 (0.0)	N/A	N/A
Pituitary disorder	2 (0.7)	1 (0.4)	N/A	N/A
Hypophysitis	2 (0.7)	1 (0.4)	N/A	N/A
Diabetes	1 (0.4)	0 (0.0)	N/A	N/A
Type I diabetes mellitus	1 (0.4)	0 (0.0)	N/A	N/A
Gastrointestinal				
Diarrhoea	24 (8.9)	5 (1.9)	7 (9.0)	0 (0.0)
Colitis	2 (0.7)	1 (0.4)	1 (1.3)	1 (1.3)
Hepatic				
Alanine aminotransferase increased	8 (3.0)	2 (0.7)	3 (3.8)	0 (0.0)
Aspartate aminotransferase increased	6 (2.2)	3 (1.1)	1 (1.3)	1 (1.3)
Blood alkaline phosphatase increased	3 (1.1)	2 (0.7)	1 (1.3)	0 (0.0)
Blood bilirubin increased	2 (0.7)	1 (0.4)	1 (1.3)	0 (0.0)
Liver function test increased	2 (0.7)	1 (0.4)	N/A	N/A
Transaminases increased	2 (0.7)	0 (0.0)	N/A	N/A
Hyperbilirubinaemia	1 (0.4)	0 (0.0)	N/A	N/A
Pulmonary				
Pneumonitis	10 (3.7)	2 (0.7)	2 (2.6)	0 (0.0)
Interstitial lung disease	1 (0.4)	1 (0.4)	N/A	N/A
Renal				
Acute kidney injury	1 (0.4)	0 (0.0)	1 (1.3)	1 (1.3)
Blood creatinine increased	1 (0.4)	1 (0.4)	4 (5.1)	0 (0.0)
Renal failure	1 (0.4)	0 (0.0)	N/A	N/A
Blood urea increased	N/A	N/A	3 (3.8)	0 (0.0)
Hypersensitivity/infusion reactions				
Infusion related reaction	2 (0.7)	1 (0.4)	1 (1.3)	0 (0.0)
Hypersensitivity	1 (0.4)	0 (0.0)	1 (1.3)	0 (0.0)

Source: CS, Table 26, page 76-78
Includes events reported between first dose and 30 days after last dose of study therapy.
AEs = adverse events; N/A = not applicable.

CheckMate 032

An overview of the primary efficacy results (primary database lock: 24 March 2016) from the UC cohort of CheckMate 032 is presented in Table 4.11. A confirmed investigator-assessed objective response was achieved in 19 (24.4%) patients (95% CI: 15.3 to 35.4) of 78 treated patients, with five patients (6%) achieving a CR and 14 patients (18%) achieving a PR.

Patients in the PD-L1≥1% cohort achieved an objective response rate (ORR) of 24.0% and patients with <1% PD-L1 expression had a confirmed ORR of 26.2%.

Table 4.11: Overview of clinical effectiveness results from CheckMate 032

Tumour response	Nivolumab (n=78)	PD-L1 <1% (n=42)	PD-L1 ≥1% (n=25)
ORR, n (%)	19 (24.4) [95% CI 15.3–35.4]	11 (26.2)	6 (24.0)
BOR, n (%)			
CR	5 (6.4)	1 (2.4)	4 (16.0)
PR	14 (17.9)	10 (23.8)	2 (8.0)
SD	22 (28.2)	11 (26.2)	8 (32.0)
PD	30 (38.5)	18 (42.9)	8 (32.0)
Unable to determine	7 (9.0)	2 (4.8)	3 (12.0)
Median TTR, months (IQR)	1.48 (1.25–4.14)	[redacted]	[redacted]
Median DOR, months (95% CI)	NR (9.92–NR)	[redacted]	[redacted]

Source: CS, Table 15, page 51; and CS, Appendix E, Table 56, page 148
BOR = best overall response; CI = confidence intervals; CR = complete response; DOR = duration of response; IQR = interquartile range; ORR = objective response rate; PD = progressive disease; PD-L1 = programmed death ligand 1; PR = partial response; SD = stable disease; TTR = time to response NR = not reached.

The Kaplan-Meier plots for PFS and OS in CheckMate 032 are presented in Figure 4.3 and Figure 4.4.

Median PFS was 2.78 months (95% CI 1.45 to 5.85) and 60 (77%) of 78 patients had disease progression or died by data cut-off. Of 18 (23.1%) censored patients, [redacted] had their PFS time censored on either the date of last on-study tumour assessment or date of last assessment prior to subsequent anti-cancer therapy. The most common reason for censoring among these patients was [redacted]. PFS rates (95% CI) were [redacted] at three months, [redacted] at six months and 20.8% (12.3 to 30.9) at 12 months.

Median PFS for patients in the PD-L1 ≥1% cohort was longer than in the all-treated population at 5.45 months (95% CI: 1.41–11.17), and in the PD-L1 <1%, median PFS was 2.76 months (95% CI: 1.41–6.51)

Figure 4.3: Kaplan-Meier plot for progression-free survival in subgroups of CheckMate 032

FIGURE REDACTED

Source: CS, Appendix E, Figure 27, page 148
CI = confidence interval; PD-L1 = programmed death ligand 1.

Median OS was 9.7 months (95% CI 7.3 to 16.2) and 46 (59%) of 78 patients had died at the time of data cut-off. OS rates (95% CI) were [redacted] at three months, [redacted] at six months, and 45.6% (34.2 to 56.3) at 12 months. Median follow-up for OS (time between dose date and last known date alive or death) for all nivolumab monotherapy treated UC patients was 9.69 months (range: 0.7 to 20.7 months).

Median OS for patients in the PD-L1 ≥1% cohort was longer than in the all-treated population at 16.16 months (95% CI: 7.59 to N.A.), and in the PD-L1 <1%, median OS was 9.89 months (95% CI: 7.03 to N.A.)

Figure 4.4: Overall survival in subgroups of CheckMate 032

FIGURE REDACTED

Source: CS, Appendix E, Figure 28, page 149
CI = confidence interval; PD-L1 = programmed death ligand 1.

Patient-reported outcomes data for the measurement of HRQoL was assessed via the EQ-5D-3L questionnaire in CheckMate 032. A total of 73 (93.5%) UC patients treated completed the EQ-5D VAS questionnaire at baseline and the mean baseline EQ-5D VAS score was 72.4 (SD 24.5). Overall, the mean EQ-5D VAS score increased over time. By Week 19, clinically meaningful improvements (>7-point change from baseline) were reported and the average EQ-5D VAS score was >80 points. The EQ-5D VAS continued to improve through Week 61. After week 61, the sample size was too small to interpret (<10).

In CheckMate 032, the majority ([redacted]) of patients received ≥90% of the planned nivolumab dose intensity; the median number of nivolumab doses received was 8.5 with [redacted] receiving >4 doses. The median duration of therapy was [redacted] months (95% CI: [redacted]). At the time of the 24 March 2016 database lock, 76.9% of patients in the UC cohort of CheckMate 032 had discontinued study treatment; the most common reason was disease progression (64.1%). Two (2.6%) patients discontinued due to study drug toxicity.

ERG comment: The outcomes for nivolumab in CheckMate 275 are generally worse than in the CheckMate 032 trial; given the low sample sizes of the studies this could be explained by sampling error. There appeared to be little change between the May and September database locks, although median OS did come down slightly. The company were asked to provide the most recent data in addition to those submitted in the CS, given that the survival data is from an analysis that is over a year old.[21] The company did not provide further data.[7] There was a statistically significant difference in OS between the PD-L1 < 1% and PD-L1 >= 1% subgroups. The company were requested to perform the

indirect treatment comparison (STC) for these subgroups, but they declined citing unavailability of PD-L1 status in the comparator trials as a reason.[7] The ERG would argue that, whilst PD-L1 status might be prognostic, it would be unlikely to affect the effectiveness of the comparator treatments given their different mode of action to nivolumab. Therefore the ERG considers PD-L1 status is unimportant for the comparator. Moreover, lack of information on other baseline characteristics did not preclude their inclusion in the prediction model for the STC (see Section 4.4.1 below) since such missing data was imputed by the company.

4.2.6 Meta-analyses of the nivolumab studies

According to the company, *'Data from CheckMate 275 and CheckMate 032 were pooled to perform the STC presented in Section B.2.9 and Appendix D'* (CS, section B.2.8, page 59).[2, 20] However, no methods or results are presented for the pooling of data.

ERG comment: The ERG asked the company to provide details of the statistical method(s) used for pooling the data from Checkmate 275 and CheckMate 032 and to explain which data were used (BIRC or investigator-assessed).[21] The ERG also asked the company to conduct pooled analyses using data from each method separately.

In the response to the clarification letter, the company did not state how the two nivolumab trials were pooled.[7] They did clarify that the BIRC method was chosen for CheckMate 275, but only the investigator-assessed results were available for CheckMate 032. They also stated the following on page 26 of the response: '*As agreed with the ERG on the preliminary teleconference to discuss the clarification questions, analyses using each method separately have not been provided.*'[7] However, no such agreement was made and the ERG continues to believe that the results of the STC using only BIRC or only investigator-led methods would provide valuable insight into the variability of those results. Given that the BIRC method was only available for CheckMate 275 this would imply a minimum of performing the STC using only the CheckMate 275 data. This additional analysis was suggested to the company during the teleconference (to which the company refer in the response to clarification) but was not performed.

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

The systematic literature review (SLR) identified no RCTs directly comparing the efficacy and safety of nivolumab in the patient population of interest versus any of the comparators relevant to this submission or placebo.

Three studies were excluded because the dose and/or treatment regimens did not correlate with current UK clinical practice.[5] The trials not considered for further assessment were Kim et al. (2016),[23] McCaffrey et al. (1997)[24] and Vaughn et al. (2002).[25]

Nine trials, including the two nivolumab trials, were considered eligible for STC.[8, 9, 13-16, 26-28] (See Table 4.12). Note that the single arm study design of the nivolumab studies prevented standard indirect comparison or mixed treatment comparisons since there was an incomplete network. To allow any comparison of nivolumab effectiveness to any eligible comparator the company performed an unanchored (no common comparator) STC. An unanchored STC relies on the major assumption that absolute outcomes can be predicted from a set of covariates; therefore it assumes that all effect modifiers and prognostic factors are accounted for.[1] In addition to the two nivolumab studies,[8, 9] a further seven studies were found to be used in the STC. The seven studies looked at paclitaxel,[15, 28] docetaxel,[16, 27] BSC,[26] and cisplatin plus gemcitabine.[13, 14]

Table 4.12: Summary of trials included in simulated treatment comparisons

Trial ID	Study design	Interventions (n patients assigned)	Treatment included in STC
Bellmunt 2009	RCT	Vinflunine + BSC (253) vs. BSC (117)	BSC
Choueiri 2012	RCT	Docetaxel + vandetanib (74) vs. Docetaxel +placebo (75)	Docetaxel
Gondo 2011	Single arm	Gemcitabine + cisplatin (33)	Gemcitabine + cisplatin
Joly 2009	Single arm	Paclitaxel (45)	Paclitaxel
Jones 2017	RCT	Pazopanib (66) vs. Paclitaxel (65)	Paclitaxel
Ozawa 2007	Single arm	gemcitabine + cisplatin (55)	Gemcitabine + cisplatin
Petrylak 2016	RCT	Docetaxel (49) vs. Docetaxel + ramucirumab (49) vs. Docetaxel + icrucumab (50)	Docetaxel
Sharma 2016	Single arm	Nivolumab (78)	Nivolumab
Sharma 2017	Single arm	Nivolumab (270)	Nivolumab

In the two trials identified for cisplatin plus gemcitabine (Gondo et al. (2011)[13] and Ozawa et al. (2007)[14]), all patients in Gondo et al. (2011)[13] had received MVAC in first-line treatment and are, according to the company, therefore not comparable to those receiving cisplatin plus gemcitabine retreatment in current UK clinical practice, as they are gemcitabine naïve.[5] The Ozawa et al. (2007)[14] trial included chemotherapy-naïve patients in addition to patients who had previously undergone first-line treatment. Although outcome data are reported separately for these two populations, patient baseline characteristic data are reported for the two populations combined. Therefore, it is not possible to determine baseline characteristics for patients who had only received first-line treatment, precluding an adjusted (STC) comparison with patients in other studies included in this analysis. Additionally, the two trials did not use the standard dosing regimen typically used for cisplatin plus gemcitabine in the UK. Furthermore, the study by Gondo et al. (2011)[13] provided no PFS data, and the study by Ozawa et al. (2007)[14] provided neither OS not PFS data. As the only identified evidence for cisplatin plus gemcitabine, these trials were taken forwards for the ITC, but the company used the comparison between nivolumab and cisplatin plus gemcitabine for the purposes of a scenario analysis only.

Only one of the studies was conducted exclusively in the UK,[15] one study included some patients from the UK (CheckMate 032: six out of 78),[7] one study was conducted in multiple countries, but it was unclear whether this included the UK[26] and the remaining six studies did not include UK patients.[8, 13, 14, 16, 27, 28].

All trials reported some inclusion criteria. All trials except Ozawa et al. (2007)[14] reported inclusion criteria relating to previous treatment. Six trials required patients to have shown evidence of recurrence or progression following first-line platinum therapy.[8, 9, 15, 26-28] One trial specified that the first-line treatment was MVAC.[13] Joly et al. (2009) did not name the type of first-line chemotherapy.[28] Ozawa et al. (2007) did not mention first-line treatment in their inclusion criteria.[14]

Although some of these studies are RCTs, the company used single arms only from each study. Therefore, all the advantages of comparability between groups in a RCT have been lost. The company

tried to use a STC to adjust for some of the differences between the included studies. As stated by the company, the network for nivolumab and its comparators is disconnected. Hence the indirect comparison was conducted using STC methodology. Ideally, for each outcome, the STC should adjust for all the effect modifiers and prognostic variables. However, this is rarely possible, as some effect modifiers and prognostic variables may not be reported by all of the trials or may not be known (for example, as yet undiscovered genetic markers). The company followed the recommendations in the NICE DSU TSD 18.[1] However, we reiterate an unanchored STC '…effectively assumes that absolute outcomes can be predicted from the covariates; that is, it assumes that all effect modifiers and prognostic factors are accounted for This assumption is very strong, and largely considered impossible to meet. Failure of this assumption leads to an unknown amount of bias in the unanchored estimate.'[1].

The nine studies included in the STC are described in Table 4.13. Details of prior chemotherapy received are reported in Table 4.14 (Patient characteristics). As can be seen in Table 4.14, patient populations in the studies differed in terms of Eastern Oncology Cooperative Group (ECOG) performance status at baseline, tumour location, presence and location of metastases, previous adjuvant treatment, prior chemotherapy treatments, prior radiotherapy, prior surgery and prior response to chemotherapy. In addition patient populations differ in BMI, ethnicity, smoking status, time since diagnosis, PDL-1 expression, haemoglobin level, platelet level, neutrophil level, CD8 count, and lactate dehydrogenase level. Baseline variables are available for some of the trials, but in many case cases no data are available.

The statistical analysis data for studies included in the STC are reported in Tables 22 and 23 of the CS (CS, Appendix D, pages 91-93).[2]

ERG comment:
There was much variability in patient populations between the included studies and so it is unlikely that they can be considered as comparable. The company did adjust for differences in performing the STC (see Section 4.4.). However, many characteristics were not reported for the comparator studies, thus leading to the likelihood of persistent imbalance in both prognostic factors and effect modifiers.[1] The majority of data for nivolumab or the eligible comparators did not come from UK patients.

Table 4.13: Single arms of studies included in the simulated treatment comparison

Trial ID	Population	Intervention	Reported outcomes specified in the decision problem	All other reported outcomes
Sharma et al. (2017)[8] CheckMate 275*	Histologically or cytologically confirmed metastatic or surgically unresectable transitional cell carcinoma of the urothelium involving the bladder, urethra, ureter, or renal pelvis, age ≥18 years, and ECOG PS of 0 or 1. Progression or recurrence after treatment either: o With at least 1 platinum-containing chemotherapy regimen for metastatic or surgically unresectable locally advanced urothelial cancer, or o Within 12 months of peri-operative (neo-adjuvant or adjuvant) treatment with platinum agent in the setting of cystectomy for localised muscle-invasive urothelial cancer.	Nivolumab 3 mg/kg Q2W	OS, PFS, ORR	BIRC-assessed PFS, OS and investigator-assessed ORR, PFS, safety, HRQoL (EORTC QLQ-C30 and EQ-5D-3L)
Sharma et al. (2016)[9] CheckMate 032*	Locally advanced or metastatic urothelial cell carcinoma, age ≥18 years, and ECOG PS of 0 or 1. Progression or recurrence either: o After at least 1 previous platinum-containing chemotherapy treatment for metastatic or locally advanced unresectable urothelial cancer, or o Recurrence within 1 year of completing previous platinum-based neoadjuvant or adjuvant treatment o After previously refusing standard treatment with chemotherapy for the treatment of metastatic (stage IV) or locally advanced disease	Nivolumab 3 mg/kg Q2W	OS, PFS, ORR	Investigator-assessed PFS, OS, DOR, safety, HRQoL (EQ-5D)

Trial ID	Population	Intervention	Reported outcomes specified in the decision problem	All other reported outcomes
Bellmunt et al. (2009)[26]	Patients with histologically confirmed locally advanced or metastatic TCC of urothelial tract, documented progression after first-line platinum-containing chemotherapy, age ≥18 years, and ECOG PS of 0 or 1.	BSC (including palliative radiotherapy, antibiotics, analgesics, corticosteroids and/or transfusions); 3-week cycle;	OS, ORR	Disease control rate, clinical benefit, QoL
Choueiri et al. (2012)[27]	Eligible patients required histologically or cytologically confirmed locally advanced or metastatic UC, progression of disease documented by the investigator after platinum-containing chemotherapy, age ≥18 years, and ECOG PS of 0 or 1.	Docetaxel (75mg/m^2 D1) + Placebo (100mg daily); 21-day cycle;	PFS, ORR	Safety and disease control rate
Jones et al. (2017)[15]	Histologically confirmed TCC of the bladder, renal pelvis, ureter or urethra which was locally advanced or metastatic; Progressive disease during or after one prior platinum-based chemotherapy regimen for advanced disease	Paclitaxel (80mg/m^2 IV over 1 hour, D1, D8, D15); 28 day course;	OS, PFS, Grade 3 and Grade overall AEs	PR, SD, QoL, toxicity
Petrylak et al. (2016)[16]	Patient had histologically or cytologically confirmed TCC of the bladder, urethra, ureter, or renal pelvis, locally advanced or metastatic and unresectable TCC of the bladder, urethra, ureter, or renal pelvis and had received treatment with a platinum-containing regimen.	Docetaxel (75 mg/m^2 IV; D1); 3-week cycle,	OS, PFS, ORR	DoR, safety, pharmacokinetics, pharmacodynamics and immunogenicity
Gondo et al. (2011)[13]	Patients with histologically confirmed advanced and metastatic UC. All patients had evidence of disease progression, relapse or no response after MVAC chemotherapy as first-line treatment.	Gemcitabine (1,000 mg/m2; D1, D8, D15); Cisplatin (35 mg/m2; D1, D2); 28 day-cycle;	OS, ORR	Toxicity
Joly et al. (2009)[28]	Patients had urothelial carcinoma of the bladder, or urothelial tract, with a progressive measurable disease after previous line of chemotherapy for	Paclitaxel (80mg/m2 IV over 1 hour, D1, D8, D15); 28 day course;	ORR	CR, PR, SD

Trial ID	Population	Intervention	Reported outcomes specified in the decision problem	All other reported outcomes
	advanced disease (neoadjuvant, adjuvant, or metastatic therapy), life expectancy ≥3 months, WHO performance status of 0-2			
Ozawa et al. (2007)[14]	Patients had histological or cytological proof of UC, at least one bi-dimensionally measurable lesion according to WHO criteria, and a WHO performance status <2	Gemcitabine (1000mg/m2 D1, D8, D15) Cisplatin (70mg/m2 D2); Every 28 days	ORR	Toxicity

Source: CS, Appendix B, Tables 16 and 17, pages 67-68 and *response to clarification letter.
AE = adverse event; BSC = best supportive care; CR = complete response; D = day; DoR = duration of response; ECOG = Eastern Cooperative Oncology Group; IV = intravenous; MVAC = methotrexate, vinblastine; adriamycin (doxorubicin) and cisplatin; NA = not applicable; NICE = The National Institute for Health and Care Excellence; ORR = objective response rate; OS = overall survival; PFS = progression-free survival; PR = partial response; PS = performance status; QoL = quality of life; SD = stable disease; TCC = transition cell carcinoma; TTR = time to response; UC = urothelial carcinoma; WHO = World Health Organization.

Table 4.14: Patients' characteristics in studies included in the simulated treatment comparison

Trial, treatment arm and population size	Age (years)	Males n (%)	ECOG status n (%)	Location of urothelial cancer n (%)	Presence and location of metastasis n (%)	Prior neo-adjuvant or adjuvant treatment	Prior type of chemotherapy n (%)	Prior radio-therapy n (%)	Prior surgery n (%)	Response to prior chemo-therapy n (%)
Sharma et al. (2017)[8] CheckMate 275*	Median 66 (38-90)	211 (78.1)	0: 145 (53.7) 1: 124 (45.9) 3: 1 (0.3)	Urinary bladder: 197 (73.0) Renal pelvis: 46 (17.0) Ureter: 19 (7.0) Urethra: 8 (3.0)	Visceral: 227 (84.1) Liver: 75 (27.8) Lymph node only: 43 (15.9)	Adjuvant: 83 (30.7) Neo-adjuvant: 60 (22.2)	Cisplatin and gemcitabine: 87 (32.2) Carboplatin and gemcitabine: 54 (20.0) MVAC: 16 (5.9) Vinflunine 20 (7.4) Paclitaxel 18 (6.7) *Therapies used in ≥5% patients in metastatic setting listed*	85 (31.5)	250 (92.6)	CR: 23 (8.6) PR: 44 (16.4) SD: 51 (19.0) PD: 88 (32.7) N/A, UtD, NR: 63 (23.3)[a] *Percentage based on prior platinum containing regiment associated with recurrence/regression (n=72)*
Sharma et al. (2016)[9] CheckMate 032*	Median 66 (31-85)	54 (69.2)	0: 42 (53.8) 1: 36 (46.2)	NR	Visceral: 61 (78.2) Liver: 20 (25.6) Lymph node only: 13 (16.7)	Adjuvant: 33 (42.3) Neo-adjuvant: 14 (17.9)	Cisplatin and gemcitabine: 23 (29.5) Carboplatin and gemcitabine: 15 (19.2) MVAC: 7 (9.0)	25 (32.1)	71 (91.0)	CR: 2 (2.8) PR: 15 (20.8) SD: (19 (26.4)

Trial, treatment arm and population size	Age (years)	Males n (%)	ECOG status n (%)	Location of urothelial cancer n (%)	Presence and location of metastasis n (%)	Prior neo-adjuvant or adjuvant treatment	Prior type of chemotherapy n (%)	Prior radio-therapy n (%)	Prior surgery n (%)	Response to prior chemo-therapy n (%)
							Carboplatin and paclitaxel: 5 (6.4) Vinflunine: 4 (5.1) *Therapies used in ≥5% patients in metastatic setting listed*			PD: 24 (33.3) N/A, UtD: 12 (16.7)[a] *Percentage based on prior platinum containing regiment associated with recurrence/regression (n=72)*
Bellmunt et al. (2009)[26] BSC n=117	65+: n=57 (48.7%)	NR	Grade 0: 45 (38.5); Grade 1: 72 (61.5); Grade 2: 0; Grade 3: 0	NR	Visceral involvement: 87 (74.4)	NR	Cisplatin and no other platinum: 85 (7.26) Carboplatin and no other platinum: 12(19.7) Other platinum combination: 9(7.7)	NR (22)	NR	NR
Choueiri et al. (2012)[27] Docetaxel and placebo n=72	≥65: n=33 (45.8%)	49 (68.1)	Grade 0: NR; Grade 1: 38 (52.8); Grade 2: NR;	NR	Visceral: 46 (63.9); Liver: 27 (37.5)	NR	Previous treatment with platinum-based chemotherapy was a requirement of the eligibility criteria.	15 (21)	Cystectomy: 36 (50)	NR

Trial, treatment arm and population size	Age (years)	Males n (%)	ECOG status n (%)	Location of urothelial cancer n (%)	Presence and location of metastasis n (%)	Prior neo-adjuvant or adjuvant treatment	Prior type of chemotherapy n (%)	Prior radio-therapy n (%)	Prior surgery n (%)	Response to prior chemo-therapy n (%)
			Grade 3: NR				Prior paclitaxel: 8 (11.1).			
Jones et al. (2017)[15] Paclitaxel n=65	Median 70 (IQR: 63-77)	49[a] (75)	Grade 0: (39); Grade 1: (52); Grade 2: (9); Grade 3: (0)	Bladder primary: NR (66)	Nodal: NR (45); Liver: NR (29) Visceral (non-lymph node): 49 (75.4)[b]	NR	Platinum based: 65 (100)	NR	NR	NR
Petrylak et al. (2016)[16] Docetaxel n=45	Median 69 (IQR: 29-84)	35 (78)	Grade 0: 17 (38); Grade 1: 26 (58); Grade 2: 1 (2.2); Grade 3: 0; Missing: 1 (2.2)	NR	Visceral: 29 (64); Liver: 12 (NR)	NR	Platinum-based therapy (cisplatin or carboplatin): 45 (100); Gemcitabine: 42 (93); Cisplatin: 31 (69); Carboplatin: 20 (44); Doxorubicin: 4 (9); Methotrexate: 4 (9); Vinblastine: 4 (9); Investigational drug: 1 (2); Paclitaxel: 4 (9); Capecitabine: 0; Fluorouracil: 1 (2); Ifosfamide: 1 (2); Mitomycin: 0; Pemetrexed: 1 (2).	5 (11)	40 (89)	44 (98)

Trial, treatment arm and population size	Age (years)	Males n (%)	ECOG status n (%)	Location of urothelial cancer n (%)	Presence and location of metastasis n (%)	Prior neo-adjuvant or adjuvant treatment	Prior type of chemotherapy n (%)	Prior radio-therapy n (%)	Prior surgery n (%)	Response to prior chemo-therapy n (%)
Gondo et al. (2011)[13] Gemcitabine and Cisplatin n=33	Median 66 (40-82)	26 (78.8)	Inclusion criteria: ECOG PS <1 n: NR	Bladder alone: 19 (57.6); Ureter: 7 (21.2); Renal pelvis: 7 (21.2).	Bone: 5 (15.2); Bone only: 1 (3) Lymph nodes only: 10 (30.3); Lymph nodes and lung: 5 (15.2); Lymph nodes and local recurrence: 4 (12.1); Lymph nodes and liver: 2 (6.1); Lymph nodes and bone: 1 (3.0); Evaluable lymph nodes: 24 (72.7) Lung only: 3 (9.1); Evaluable lung: 11 (33.3); Lung and local recurrence: 2 (6.1) Liver: 5 (15.2); Liver and peritoneum: 1 (3.0); Visceral lesions: 23; Other: 10 (30.3).	Adjuvant: 14 (42)	MVAC. Number of courses: 1: 2 (6.1); 2: 10 (30.3); 3: 10 (30.3); 4: 14 (12.1); ≥5: 7 (21.2).	NR	32 (97)	NR

Trial, treatment arm and population size	Age (years)	Males n (%)	ECOG status n (%)	Location of urothelial cancer n (%)	Presence and location of metastasis n (%)	Prior neo-adjuvant or adjuvant treatment	Prior type of chemotherapy n (%)	Prior radio-therapy n (%)	Prior surgery n (%)	Response to prior chemo-therapy n (%)
Joly et al. (2009)[28] Paclitaxel n=45	Mean 64 (47-79)	36 (80a)	NR	Bladder alone: 38 (84); Non-bladder cancer reported as other: 7 (16a)	Bone: 14 (33); Visceral: 26 (58); Nodes: 23 (55); Pulmonary: 22 (52); Liver: 16 (38); Other: 11	Adjuvant: 32 (71)	Gemcitabine and Cisplatin: 40(89) MVAC: 5(11) Paclitaxel with cisplatin: 1; Paclitaxel with cisplatin and gemcitabine: 1 first-line adjuvant: 32 (71) first-line for metastasis: 13 (29).	16 (36)	Total: 39 (87); Radical surgery: 28 (NR); Transurethral resection of the bladder: 7 (NR)	NR (62)
Ozawa et al. (2007)[14] Gemcitabine n=55	Median 71 (32-84)	44 (80)	NR	Bladder alone: 28 (50.9); Ureter: 16 (29.1); Renal pelvis: 11 (20)	Lymph nodes: 23; Lymph node and lung: 6; Lymph node and liver: 3; Lymph node and bone: 4; Lymph node, lung and liver: 1; Lymph node, lung, liver and bone: 1; Lung: 5; Lung and liver: 1; Lung and bone: 1;	NR	20/47 patients with metastatic disease received prior chemo MVAC: 14 (25a); MEC: 5 (9a); Low dose cisplatin: 1 (2a)	NR	NR	NR

Trial, treatment arm and population size	Age (years)	Males n (%)	ECOG status n (%)	Location of urothelial cancer n (%)	Presence and location of metastasis n (%)	Prior neo-adjuvant or adjuvant treatment	Prior type of chemotherapy n (%)	Prior radio-therapy n (%)	Prior surgery n (%)	Response to prior chemo-therapy n (%)
					Lung, liver and bone: 2					
Source: CS, Appendix D, Tables 20 and 21, pages 84-87 and response to clarification letter* [a]Reviewer-calculated value, [b]Data provided by study author on request. ECOG: Eastern Cooperative Oncology Group; IQR: interquartile range; MEC: methotrexate, epirubicin and cisplatin; MVAC: methotrexate, vinblastine, adriamycin (doxorubicin) and cisplatin; NR: not reported; PD: progressive disease.										

4.4 Critique of the indirect comparison and/or multiple treatment comparison

4.4.1 Methodology of the simulated treatment comparison

The company used a population-adjusted method (STC) to conduct comparisons between nivolumab and eligible comparators with respect to OS, PFS and ORR outcomes.[1]

The STC was informed by individual patient data (IPD) from the two nivolumab studies[8, 9] and published data from the other seven studies of comparator treatments.[13-16, 26-28]

The methods followed the recommendations of the NICE DSU TSD 18: Methods for Population-Adjusted Indirect Comparisons in Submissions to NICE.[1]

For each outcome, the key steps of the STC approach were:

1. Use the nivolumab IPD to develop a model that predicts how patients respond to treatment based on key baseline patient characteristics.
2. For each comparator trial in the network, use the baseline characteristics from the comparator trial to predict how patients in the comparator trial might have responded to nivolumab. Compare the real data from the comparator, to the predicted data for nivolumab.
3. Use a meta-analysis to synthesise the results across all of the comparator trials.

Details of each of the steps are shown in Appendix D of the CS.[20]

For Step 1, prognostic factors and effect modifiers were identified via a targeted literature search and via discussion with clinicians at the advisory board meeting.[5] The Prediction models were estimated on the pooled CheckMate 275 and CheckMate 032 data. It was reported that stepwise model selection suggested that the best Cox Proportional hazards (PH) model for OS is based on ECOG PS, haemoglobin level, visceral metastases and liver metastases. Note that this model includes all three of the key prognostic factors identified by Bellmunt et al. (2010)[26] (ECOG PS, haemoglobin level and liver metastases). For PFS the same approach showed the best model is based on ECOG PS, age, visceral metastases and liver metastases. Stepwise model selection suggested that the best logistic regression model for objective response is based on age and visceral metastases. The basis of selection was reported to be parsimony as indicated by the Akaike information criterion (AIC). No models other than the final and presumably most parsimonious models (no more than four covariates) were presented despite the consideration of eleven possible covariates.

For Step 2, because not all of these baseline characteristics were reported for all comparator trials, for each comparator trial, any baseline characteristics that were in the final prediction models, but not reported for the comparator trial, were then predicted using the correlations between baseline characteristics in the nivolumab trials.

This method essentially adjusts the outcomes estimated from the nivolumab trials to attempt to simulate how they might be observed in each of the comparator trials. Therefore, there is one adjusted value (for nivolumab) for each outcome, e.g. ORR, for each comparator trial. This means that there can be more than one adjusted value for nivolumab per comparator. For example, as shown in Table 4.17, ORR is estimated for docetaxel from two trials, Choueiri et al. (2012)[27] and Petrylak et al. (2016).[16] Therefore, there will be two adjusted values of ORR for nivolumab to compare to these trials and to estimate the treatment effect in terms of a relative risk. For OS and PFS adjusted hazards are predicted with one for

each of a set of four-weekly time intervals. As with ORR, there are two trials for docetaxel and so this means two sets of adjusted hazards, each one of which goes into the meta-analysis model in Step 3.

For Step 3, OS and PFS were evaluated using a fractional polynomial approach, which permits the estimation of hazard ratios (HRs) that vary over time. ORR was evaluated using an evidence synthesis model for binomial outcomes.[29] For all outcomes, both fixed effect and random effects models were applied. For the survival outcomes, different types of fractional polynomial model (according to variation in two parameters that determine the shape of the survival curves) were also explored. The deviance information criterion (DIC) was used to evaluate model fit and guide the best choice of model. For the survival outcomes, clinical plausibility of the extrapolated HRs was also considered based on expert clinical feedback elicited via an advisory board and further clinician interviews.

In addition the company stated that they conducted naïve indirect comparisons alongside STCs as recommended by the DSU.[1] Although not explicitly stated, one can presume that this means that the hazards for nivolumab were not adjusted using Steps 1 and 2 above.

In order to investigate how well the STC method performed the company also compared the docetaxel versus docetaxel plus vandetanib results from Choueiri et al. (2012) [27] to the results of an STC using data from this trial.

For STCs, the NICE DSU TSD 18 recommends estimating the residual bias.[1] This is the bias due to effect modifiers or prognostic variables that are not accounted for in the prediction models because they are not available for either the nivolumab and/or the comparator studies. The NICE DSU TSD 18 emphasises that there are no standard methods for estimating the residual bias and that this is a key area for further research. The NICE DSU TSD 18 suggests two general options for evaluating residuals bias: 'in-sample' methods, which use the same data that was used to develop the prediction model, and 'out-of-sample' methods which incorporate additional data.

ERG comment: As stated above the DSU report mentions that an unanchored STC *'effectively assumes that absolute outcomes can be predicted from the covariates; that is, it assumes that all effect modifiers and prognostic factors are accounted for. This assumption is very strong, and largely considered impossible to meet. Failure of this assumption leads to an unknown amount of bias in the unanchored estimate.'*[1] The ERG believes the STC was limited by the following issues:

1. The method used for the prediction models lacked transparency; the results at each stage of the stepwise selection process were not provided. In particular, it is not clear that the most parsimonious model is the best model. It would have been useful to see an STC that was based on prediction models with more covariates including all 11 considered.
2. There was a lack of information from the comparator studies on possible effect modifiers or prognostic variables, which led to the company imputing the missing values in Step 2.
3. 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.
4. In an ideal scenario, the results of the STC using only BIRC or only investigator-led methods would have provided valuable insights into the variability of the results. Given that the BIRC method was only available for CheckMate 275 at a minimum it would have been useful to perform the STC using only the CheckMate 275 data. This was suggested to the company during the teleconference but was not performed.

An attempt was made by the company to validate the STC. It is the understanding of the ERG that the data from the trial by Choueiri et al. (2012)[27] was used to compare docetaxel or docetaxel plus vandetanib to nivolumab using unadjusted meta-analysis and using STC. However, this comparison is bound to produce almost identical results because in both the STC and the non-STC meta-analysis the data to inform the comparison was the same i.e. from this trial. The only difference between the STC and the direct method is that in the STC data on other trials was entered, but none of this data informs the comparison between docetaxel and docetaxel plus vandetanib. Therefore, this is essentially a spurious test of validity.

The company performed an 'in-sample' method to evaluate the residual bias. However, this method is likely to underestimate the residual bias.[1] Hence, the use of an 'out-of-sample' method is strongly recommended in NICE DSU TSD 18. This relies on the idea that, if the STC has accounted for all prognostic variables then the variance of the predictions (in this case based on the model estimated from the nivolumab data and combined with the comparator baseline characteristics) should be the same as that observed in the trial data. Unfortunately, the company concluded that the 'out-of-sample' method described in NICE DSU TSD 18 would not provide an accurate estimate of the residual bias. In the clarification letter, the company was asked to perform this analysis and in response, the company stated that in this appraisal the data was too limited to estimate the between-study variability.[7] They also argued that the fractional polynomial model constrains the between-studies variance.[29] However, they did perform the analysis and it did show much lower variance in the STC model predictions. Whether this is due to the lack of data or a limitation of the fractional polynomial model it does illustrate the point made in TSD 18 that: *'...the accuracy of the resulting estimates is entirely unknown, because there is no analysis of the potential magnitude of residual bias, and hence no idea of the degree of error in the unanchored estimates. It is, of course, most unlikely that systematic error has been eliminated.'*[1]

4.4.2 Results of the simulated treatment comparison

All studies reported data for at least one outcome. Outcome data was considered eligible for the STC analysis if a Kaplan-Meier curve was provided in addition to numerical data.

OS was reported by seven studies, including five for the four comparators with two for docetaxel.[8, 9, 13, 15, 26, 27] All of the studies except Bellmunt et al. (2009) reported a definition of OS.[26] Median survival was reported in all of the studies except Gondo et al. (2011), which reported a mean OS of 10.5 months.[13] Median OS ranged from 4.6 months in response to BSC[26] to 9.7 months in response to nivolumab.[9]

As well as in the CheckMate 032 and CheckMate 275 trials, PFS was reported by three comparators studies, for docetaxel and paclitaxel.[8, 9, 15, 27] Jones et al. (2017)[15] did not report a definition for PFS. The median PFS ranged from 1.58 months[27] in response to docetaxel and placebo to 4.1 months in response to paclitaxel.[15]

Eight studies reported ORR, including six for the four comparators.[8, 9, 13, 14, 16, 26-28] Only one study of paclitaxel by Jones et al. (2017) did not.[15] Four comparator studies did not report a definition of ORR.[8, 9, 13, 26] The ORR ranged from 0% in response to BSC[26] to 40% in response to gemcitabine and cisplatin.

The individual results of the comparator trials included in the STC are given in tables 4.15 to 4.17. The pooled results for nivolumab were not reported and were not provided in the response to the 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 4.15: 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	265	From first dose and last known date alive or death	8.74 (95%CI 6.05 to NR)
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 4.16: 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	265	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 4.17: 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	265	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.	52 (19.6) (95% CI 15.0 to 24.9)
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

For each comparator trial, and each outcome, the response to nivolumab was estimated by applying the final prediction model to the baseline characteristics in the trial in order to produce adjusted values of the outcome. Tables (see Tables 4.18 and 4.19) of hazard ratios simulated as the adjusted hazard of nivolumab in each of the trials compared to the unadjusted hazard of nivolumab in the Checkmate trials were provided by the company.[20]

Table 4.18: Overall survival. Simulated hazard ratios for response to nivolumab in each of the comparator trials versus response to nivolumab in CheckMate 275 and CheckMate 032

• Trial	Mean HR[a]	Mean log HR	SD log HR
Bellmunt *et al.* (2009)	1.04	0.043	0.608
Choueiri *et al*. (2012)	0.99	-0.010	0.635
Gondo *et al.* (2011)	0.85	-0.162	0.624
Jones *et al.* (2017)	1.04	0.043	0.609
Petrylak *et al.* (2016),	0.98	-0.025	0.618

Source: Table 35 of CS Appendix D
[a] Mean HR, back-transformed from the log scale.
Abbreviations: HR: hazard ratio; SD: standard deviation

Table 4.19: Progression-free survival. Simulated hazard ratios for response to nivolumab in each of the comparator trials versus response to nivolumab in CheckMate 275 and CheckMate 032

• Trial	Mean HR[a]	Mean log HR	SD log HR
Choueiri *et al.* (2012)[30]	0.96	-0.045	0.421
Jones *et al.* (2017)[31]	0.95	-0.056	0.391
Petrylak *et al.* (2016)[32]	0.88	-0.128	0.405

Source: Table 35 of CS Appendix D
[a] Mean HR, back-transformed from the log scale.
Abbreviations: HR: hazard ratio; SD: standard deviation

In terms of OS, these data suggested that patients in Choueri et al. (2012) (docetaxel and placebo), Petrylak et al. (2016) (docetaxel) and Gondo et al. (2011) (Gemcitabine and cisplatin) would have had on average a better response to nivolumab than patients in the nivolumab trials.[13, 16, 27] However patients in Bellmunt et al. (2009) (BSC) and Jones et al. (2017) (paclitaxel) would have had on average a poorer response.[15, 26]

In all three studies evaluating PFS (Choueri et al. (2012), Jones et al. (2017) (paclitaxel) and Petrylak et al. (2016)) patients would have had a better response to nivolumab than patients in the nivolumab trials[15, 16, 27].

The simulation suggested that patients in each of the six comparator trials evaluating objective response would have had a better response to nivolumab than in the nivolumab trials.

For OS, the company stated that the second order (P1=0, P2=0) fixed effect model was used in the base case in the cost effectiveness model analysis because it provided the most clinically plausible extrapolations out of the three best fitting models. Therefore we present in Table 4.20 the results of this model as in the main company submission. It should be noted that HRs greater than 1 favour nivolumab.

Table 4.20: Overall survival: STC results (second order (P1=0, P2=0) fixed effect model): HRs and 95% credible intervals for each of the comparators versus nivolumab for selected time intervals

Comparison	Time Interval (weeks)	HR (95% CrI)
Paclitaxel versus nivolumab	0-4	0.13 (0.02–0.64)
	8-12	0.69 (0.36–1.26)
	20-24	1.43 (0.86–2.31)
	44-48	2.27 (1.41–3.56)
	68-72	2.63 (1.17–5.52)
	92-96	2.75 (0.82–8.52)
Docetaxel versus nivolumab	0-4	0.31 (0.09–0.84)
	8-12	1.15 (0.75–1.72)
	20-24	1.81 (1.25–2.62)
	44-48	2.11 (1.46–3.00)
	68-72	2.01 (1.14–3.37)
	92-96	1.83 (0.8–3.87)
BSC versus nivolumab	0-4	0.81 (0.33–1.79)
	8-12	2.05 (1.36–3.08)
	20-24	2.51 (1.69–3.72)
	44-48	2.27 (1.57–3.25)
	68-72	1.86 (1.17–2.85)
	92-96	1.51 (0.82–2.66)
Cisplatin plus gemcitabine versus nivolumab (scenario analysis only)	0-4	0.06 (0.00–0.70)
	8-12	0.61 (0.21–1.37)
	20-24	1.33 (0.66–2.49)
	44-48	1.75 (0.96–2.99)
	68-72	1.61 (0.68–3.31)
	92-96	1.36 (0.37–4.05)
Source: Table 18 of CS BSC = best supportive care; CrI = credible interval; HR = hazard ratio		

For PFS, the second order (P1=0, P2=0) fixed effect model was taken forward for the base case analysis in the cost effectiveness model because it had clinical plausibility and the lowest DIC. No PFS data were available for cisplatin plus gemcitabine or BSC. Therefore we present the results of this model as in the main company submission. It should be noted that HRs greater than 1 favour nivolumab (Table 4.21).

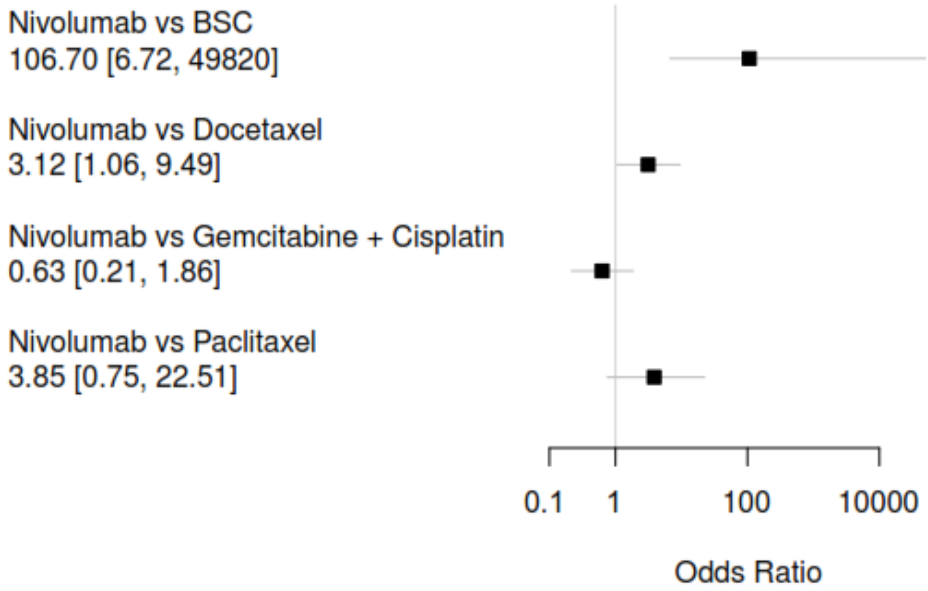
Table 4.21: Progression-free survival: STC results (fixed effect second order (P1=0, P2=0) model): HRs and 95% credible intervals for each of the comparators versus nivolumab for selected time intervals

Comparison	Time Interval (weeks)	HR (95% CrI)
Paclitaxel versus nivolumab	0-4	0.07 (0.01, 0.36)
	8-12	0.53 (0.30, 0.90)
	20-24	1.63 (1.04, 2.52)
	44-48	4.36 (1.84, 9.08)
	68-72	7.26 (1.40, 28.85)
	92-96	10.21 (0.91, 76.04)
Docetaxel versus nivolumab	0-4	1.24 (0.61, 2.42)
	8-12	1.72 (1.18, 2.49)
	20-24	1.36 (0.78, 2.20)
	44-48	0.75 (0.16, 3.19)
	68-72	0.45 (0.04, 4.82)
	92-96	0.29 (0.01, 6.93)

Source: Table 20 of the CS:
CrI = credible interval; HR = hazard ratio

For ORR the fixed effect model was used in the base case analysis so network meta-analysis results for this model are presented here (Figure 4.5, Table 4.22). However the random effects model results are also presented (Figure 4.6, Table 4.23).

Figure 4.5: Objective response rate: STC results (fixed effect model): Odds ratios for nivolumab versus each of the comparators

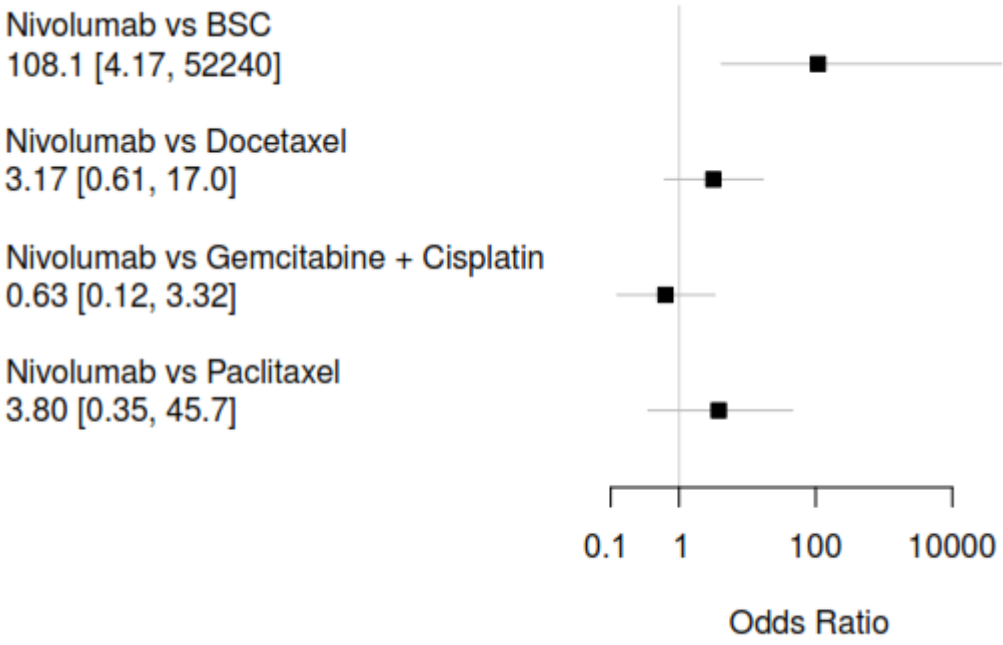


Source: Figure 30 of the CS

Table 4.22: Objective response rate: STC results (fixed effect model): Odds ratios and 95% credible intervals for each pairwise comparison

	Nivolumab	BSC	Docetaxel	Cisplatin plus gemcitabine
BSC	106.7 (6.72, 49820)			
Docetaxel	3.12 (1.06, 9.49)	0.03 (0.00, 0.59)		
Paclitaxel	3.85 (0.75, 22.5)	0.03 (0.00, 1.00)	1.23 (0.17, 9.74)	6.15 (0.87, 48.4)
Cisplatin plus gemcitabine	0.63 (0.21, 1.86)	0.01 (0.00, 0.12)	0.20 (0.04, 0.93)	
Source: Table 22 of the CS BSC = best supportive care				

Figure 4.6: Objective response rate: STC results (random effects model): Odds ratios for nivolumab versus each of the comparators



Source: Figure 19 of the CS appendices

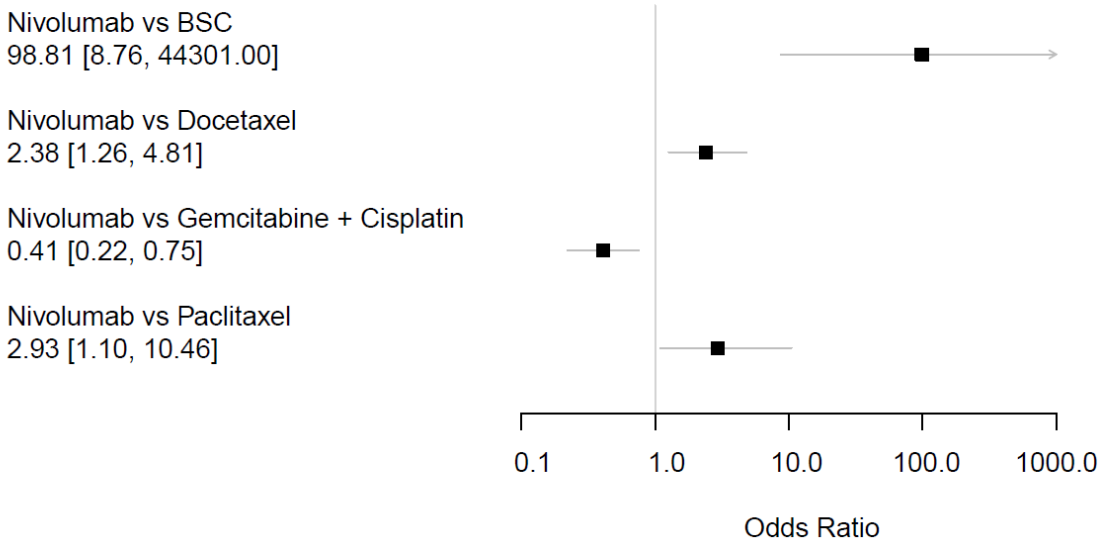
Table 4.23: Objective response rate: STC results (random effects model): Odds ratios and 95% credible intervals for each pairwise comparison

	Nivolumab	BSC	Docetaxel	Gemcitabine + cisplatin
BSC	108.1 (4.17, 52240)			
Docetaxel	3.17 (0.61, 17.0)	0.03 (0.00, 1.16)		
Gemcitabine + cisplatin	0.63 (0.12, 3.32)	0.01 (0.00, 0.23)	0.20 (0.02, 2.04)	
Paclitaxel	3.80 (0.35, 45.7)	0.03 (0.00, 2.17)	1.20 (0.07, 23.3)	6.02 (0.32, 118.1)

Source: Table 45 of the CS appendices
BSC = best supportive care

Finally, the results of a naïve indirect comparison conducted by the company in a sensitivity analysis for the outcome of objective response are presented below (both fixed effect and random effects models). Results for OS and PFS were not reported for the naïve indirect comparison: only model fit statistics were presented in the CS.[2] The results for ORR are presented in Tables 4.24 and 4.25 and Figures 4.7 and 4.8.

Figure 4.7: Naïve indirect comparison forest plot with the estimated odds ratio and its 95% credible interval, for the fixed effect model of objective response of nivolumab versus comparator treatments



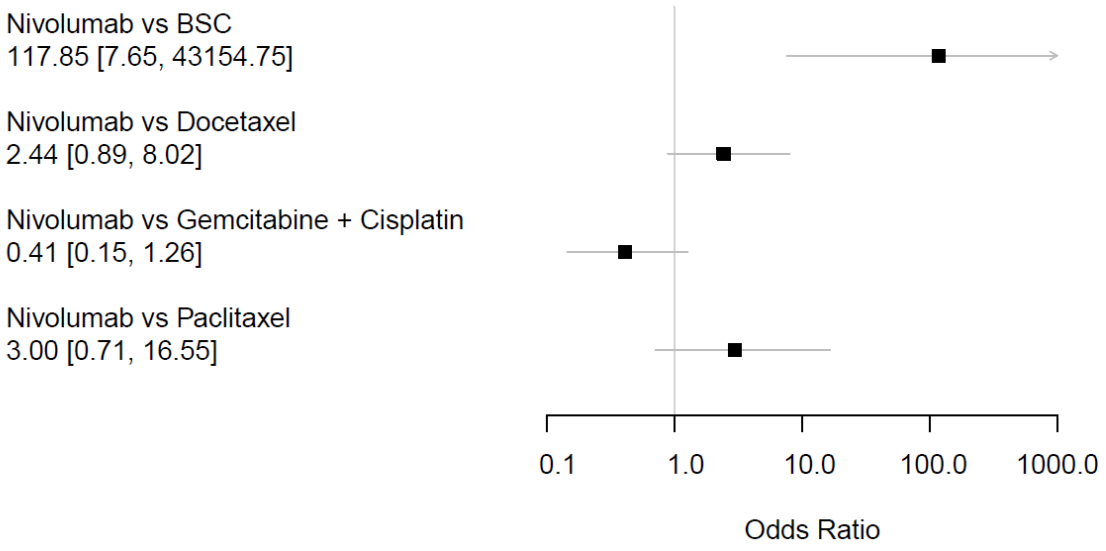
Source: Figure 20 of the CS appendices

Table 4.24: Naïve indirect comparison estimated odds ratio and 95% credible interval of the fixed effect model for the pairwise comparison of objective response between treatments

	Nivolumab	BSC	Docetaxel	Gemcitabine + Cisplatin
BSC	98.8 (8.76, 44301.00)			
Docetaxel	2.38 (1.26, 4.81)	0.02 (0.00, 0.31)		
Gemcitabine + Cisplatin	0.41 (0.22, 0.75)	0.00 (0.00, 0.05)	0.17 (0.07, 0.38)	
Paclitaxel	2.93 (1.10, 10.5)	0.03 (0.00, 0.47)	1.24 (0.39, 4.87)	7.26 (2.43, 27.8)

Source: Table 46 of the CS appendices
BSC = best supportive care

Figure 4.8: Naïve indirect comparison forest plot with the estimated odds ratio and its 95% credible interval, for the random effects model of objective response of nivolumab versus comparator treatments



Source: Figure 21 of the CS appendices

Table 4.25: Naïve indirect comparison estimated odds ratio and 95% credible interval of the random effects model for the pairwise comparison of objective response between treatments

	Nivolumab	BSC	Docetaxel	Cisplatin plus gemcitabine
BSC	117.8 (7.65, 43154.75)			
Docetaxel	2.44 (0.89, 8.02)	0.02 (0.00, 0.37)		
Paclitaxel	3.00 (0.71, 16.55)	0.03 (0.00, 0.57)	1.21 (0.26, 6.77)	7.38 (1.56, 43.0)
Cisplatin plus gemcitabine (scenario analysis only)	0.41 (0.15, 1.26)	0.00 (0.00, 0.05)	0.17 (0.05, 0.55)	

Source: Table 47 of the CS appendices
BSC = best supportive care

4.4.3 Adverse events

No formal comparison was made of AEs between the comparators. However, three studies reported overall adverse events.[8, 9, 15] Jones et al. (2017)[15] reported 27% of patients had Grade 3 or higher adverse events based on the Common Terminology Criteria for Adverse Events (CTCAE) v4.02 criteria. CheckMate 032[9] and CheckMate 275[8] both used the National Cancer Institute - Common Toxicity Criteria (NCI-CTC) v4.0 and reported the number of overall adverse events separately for Grade 3 and Grade 4. In CheckMate 032[9] and in CheckMate 275[8] the number of Grade 3 adverse event was 17 (22%) and 44 (16%) respectively. No Grade 4 adverse events were reported by CheckMate 032.[9] CheckMate 275 reported 4 (1%) Grade 4 adverse events.[8] In response to the request for clarification the company provided some more details of AEs in the comparator trials, as shown in Table 4.26.[7] A comparison can be made between these results and those reported for the CheckMate 032 and CheckMate 275 trials shown in Table 4.9. However, the AEs incorporated in the CEA and thus probably of most importance were summarised in the CS in the cost effectiveness section and reproduced in Table 5.7 below.[2] This shows that 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 again cisplatin plus gemcitabine. The rate of asthaenia was also lower than all comparators except cisplatin plus gemcitabine.

Table 4.26: Comparator adverse events

Study	Treatment	Safety population n	Neutropenia n (%)	Febrile Neutropenia n	Anaemia n (%)	Thrombocytopenia n (%)	Asthenia n (%)	Nausea n (%)	Vomiting n (%)	Diarrhoea n (%)	Pruritus n (%)	Pneumonia n (%)	Lung infiltration n (%)	ALT increase n (%)	Hepatitis n (%)	Abdominal pain with	Fever n (%)	Leukopenia n (%)	Constipation n (%)
Bellmunt et al. (2009)[26]	Vinflunine and BSC	248 at base line	123 (50)	15 (6)	47 (19.1)	14 (5.7)	48 (19.3)	6 (2.4)	7 (2.8)	NR	NR	NR	NR	NR	NR	NR	NR	NR	40 (16.1)
	BSC	117 at base line	1 (0.9)	0 (0)	9 (8.1)	1 (0.9)	21 (17.9)	1 (0.9)	0 (0)	NR	NR	NR	NR	NR	NR	NR	NR	NR	1 (0.9)
Choueiri *et al*. (2012)[27]	Docetaxel and Vandetanib	142	10 (14)	NR	1 (1)	NR	4 (6)	NR	NR	0 (0)	NR	NR	NR	NR	NR	NR	NR	NR	NR
Jones *et al*. (2017)[15]	Paclitaxel	129	Grade 3>: (6)	NR	NR	Grade 3≥ 0 (0)	Grade 3≥: NR (5)	Grade 3≥: 0 (0)	NR	Grade 3≥: NR (2)	NR	NR	NR	Grade 3>: NR (2)	NR	NR	NR	NR	NR
Petrylak *et al*. (2016)[16]	Docetaxel	140	Grade 3>: 16 (36)	Grade 3≥: 6 (13)	Grade 3≥: 3 (6.7)	Grade 3≥: 0	Grade 3≥: 6 (13)	Grade 3≥: 0 (0)	Grade 3≥: 0 (0)	Grade 3>: 1 (2.2)	NR	Grade 3>: 4 (8.9)	NR	NR	NR	NR	NR	Grade 3>: 6 (13)	NR
Gondo *et al*. (2011)[13]	Gemcitabine and cisplatin	33	Grade 3:	NR	Grade 3:	Grade 3: 5	Grade 3: 0	Grade 3: 0	Grade 3: 0	NR	NR	NR	NR	NR	NR	NR	Grade 3: 0	Grade 3:	NR

Study	Treatment	Safety population n	Neutropenia n (%)	Febrile Neutropenia n	Anaemia n (%)	Thrombocytopenia n (%)	Asthenia n (%)	Nausea n (%)	Vomiting n (%)	Diarrhoea n (%)	Pruritus n (%)	Pneumonia n (%)	Lung infiltration n (%)	ALT increase n (%)	Hepatitis n (%)	Abdominal pain with	Fever n (%)	Leukopenia n (%)	Constipation n (%)
			19 (57.6); Grade 4: 3 (9.1)		12 (36.4); Grade 4: 2 (6.1)	(15.2); Grade 4: 6 (18.2)	(0); Grade 4: 0 (0)	(0); Grade 4: 0 (0)	(0); Grade 4: 0 (0)								(0); Grade 4 0 (0)	14 (42.4); Grade 4: 1 (3).	
Joly *et al*. (2009)[28]	Paclitaxel	44	Grade 3: 1 (2); Grade 4: 2 (4)	NR	Grade 3: 3 (7); Grade 4: 2 (4).	NR	Grade 3: 6 (14); Grade 4: 0 (0)	Grade 3: 1; Grade 4: 0.	Grade 3: 1 (2); Grade 4: 0 (0).	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Ozawa *et al*. (2007)[14]	Gemcitabine and cisplatin	55	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR

Source: response to clarification

BSC: Best supportive care; NR: not reported

ERG comment:
In terms of ORR the main analysis using the fixed effect model presented finds that nivolumab is significantly better than BSC and docetaxel. No significant differences were found for nivolumab paclitaxel and gemcitabine. In the random effects model nivolumab is only statistically significantly superior to BSC. In the naïve indirect comparison nivolumab is superior to all three comparators in the fixed effect model but only to BSC in the random effects model. The results of the STC show that for OS and PFS nivolumab is superior to all comparators at most time points. However, the credible intervals for the HRs are quite wide, crossing 1 in many cases. The results of the naïve indirect comparison i.e. with the fractional polynomial model, but without the STC, were not reported. Results for other functional forms of the fractional polynomial model were presented in Appendix D, but of many functional forms, the results of only two more were presented.[20] The company was also asked to provide the results assuming proportional hazards i.e. one HR (fixed with respect to time) per comparator.[21] In response, the company provide the results of both random and fixed effects models. The method described appeared to be ad hoc. They first estimated so-called 'naïve' HRs using a proportional hazards model, but not using adjusted data i.e. apparently using the CheckMate trial data. They then adjusted these HRs to produce those intended to be as a result of the STC by the following method:

1) For each patient they calculated an adjusted HR by multiplying this 'naïve' HR by a factor calculated as the ratio of the hazard predicted by the prediction model (given the patient's characteristics) and the hazard of a patient with characteristics at the average CheckMate values

2) They then took the average of the log of this adjusted HR to get the mean adjusted log HR for each trial i.e. five values, which was then entered in the meta-analysis model.

No formal comparison was made of AEs and perhaps the most important AE data was reported in the cost effectiveness section of the CS.[2] However, it appears that the rates for nivolumab were either lower or comparable to those for the comparators.

4.5 Additional work on clinical effectiveness undertaken by the ERG

The company did not show the unadjusted hazards (estimated directly from the CheckMate 032 and CheckMate 275 trials), but they did state that they used a proportional hazard model, which suggests that hazards at all time points would be increased by the same amount, as indicated by the HRs in Tables 4.18 and 4.19. In order to check the reproducibility of the STC the data and code for running the models was requested by the ERG.[21] In response the company supplied this as an R script, with a Winbugs script embedded. However, the ERG could not run this without it generating errors and so requested it purely as a Winbugs script i.e. with the data incorporated in Winbugs format. The ERG has been able to run the meta-analyses and reproduced results only different by an amount that could be attributed to random error. The ERG can also verify that the data for OS and PFS includes the adjusted log hazards for nivolumab i.e. as a result of the STC. Because the company failed to show the unadjusted values i.e. those estimated directly from the CheckMate 032 and CheckMate 275 trials the ERG sought a method of estimating these. For OS, it was found that there were 110 values of the log hazard in five sets of 22 (corresponding to 22 four–weekly time intervals), one set for each of the five comparator trials shown in Table 4.18 (Table 35 in the CS). It was shown that by re-adjusting each of the five sets of the log hazards by the mean log HRs in Table 4.18, a single set of 22 hazards could be obtained. This verified the proportional hazard assumption since only one log HR per set was required to obtain the same original set of hazards. This single set, by definition, must be those without adjustment by the STC and which can thus be considered as having been estimated directly from the CheckMate 032 and CheckMate 275 trials.

The ERG was also able to perform the same analysis for PFS as for OS described above. In this case there were 36 nivolumab log hazards in three sets corresponding to the three PFS studies, as shown in Table 4.19 and used for only two comparators, paclitaxel and docetaxel.

The ERG was also able to check the last stage i.e. the evidence synthesis by which the fixed HRs were estimated, which revealed that this was essentially pointless in that the HRs that acted as inputs ended up being identical to the outputs, except for that versus docetaxel. This is because there was only one input per comparator, except for docetaxel for which there were two i.e. from two trials, Choueiri et al. (2012)[27] and Petrylak et al. (2016).[16] The ERG was also not convinced that the method prior to this final stage i.e. adjusting the naïve HRs was valid. Instead, for OS, the ERG performed the method advocated by Jansen, which sets the time dependent parameters in the fractional polynomial model to zero, thus allowing only a difference in the time-independent hazard.[29] This should then allow the estimation of fixed HRs. Following this method produced HRs that were quite dissimilar to those reported in the response to clarification.

4.6 Conclusions of the clinical effectiveness section

Ideally, in order to determine the relative benefits of nivolumab and its comparators there would be a series of randomised controlled trials comparing nivolumab and its comparators. Failing this, a network meta-analysis of RCTs using a set of common comparators would be the preferred approach. This would be the clearest way of determining if there was a gain in PFS or OS. However the submission relies on two single arm studies of nivolumab, one of which is small, which are then entered into a STC together with the single arms from some RCTs. Comparisons based on single arms from RCTs and studies are by their nature far less reliable than those made using the difference between arms from RCTs; in effect a comparison of observational data. 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.[1] As TSD 18 makes clear, unless all baseline characteristics that might be prognostic variables and effect modifiers are incorporated in any model to adjust for bias, it is unclear what the size of any bias might be. The ERG found the following limitations in the STC analysis:

1. Although the company stated that they had tested the fit of prediction models with various sets of baseline characteristics, it is not entirely clear how this was done: the final model had far fewer covariates than originally considered and no models with more covariates were presented or incorporated in the STC as part of a sensitivity analysis.
2. Many baseline characteristics were not available across all comparator trials and had to be imputed
3. The only external test of validity of the STC i.e. the 'out-of-sample' method seemed to either show insufficient reduction in bias or be inapplicable given the use of the fractional polynomial model that was used for survival analysis.
4. To compound the uncertainty, the numbers of actual patients are small for all comparisons and not all studies provided data for all outcomes.
5. The survival data are not fully mature in the nivolumab trials. The latest database lock provided updated OS data with a median follow-up time of 11.5 months, and at this point, only 57% of patients had died. The ERG did ask for an analysis based on more recent data, but none was provided.[7]
6. Not all study outcomes are based on independent review. An analysis based only on BIRC derived data from the nivolumab trials was also requested.[21] However, in the response to the clarification letter, the company declined to do this.[7] They also stated the following on page 26 of the response: 'As agreed with the ERG on the preliminary teleconference to discuss the

clarification questions, analyses using each method separately have not been provided.' However, no such agreement was made. Given that the BIRC method was only available for CheckMate 275 the best analysis would use only the CheckMate 275 data. This was suggested to the company during the teleconference to which the company refer in the response to clarification.

7. The company also stated that a naïve indirect comparison was performed, which the ERG understands to be without the STC, but still using the fractional polynomial model for OS and PFS. Given the ERG's opinion that the fractional polynomial model was probably appropriate, and there is doubt as to the validity of the STC, the ERG considers that the results of this naïve indirect comparison should be presented. The ERG did attempt this, but only by back-calculation and with no estimate of uncertainty.
8. The ERG would accept that the polynomial fraction model appears to be a valid and highly flexible approach to estimating HRs. However, the results of very few functional forms were presented, leaving some doubt as to the most appropriate. Also, one legitimate form is to assume proportional hazards i.e. a fixed HR with respect to time. The company did attempt this, but the methods are questionable and the method, which uses the same model as that with time-dependent HRs was not employed. Its employment by the ERG, at least for OS, seemed to produce quite different results.

Although the pooled nivolumab trial data that was used for the STC was not presented in the CS, one can compare at least crudely (without any adjustment for baseline characteristics) the outcomes of the nivolumab trials (in Tables 4.15 to 4.17) with those of the comparator trials. In particular, OS and PFS do appear to be superior for nivolumab than for BSC. However, there appears to be almost complete overlap in the 95% CIs for PFS and OS between CheckMate 275 and the docetaxel trial.[10, 16] Of course, this is without any adjustment, but even the STC, which includes the CheckMate 032 trial, which is more favourable to nivolumab, shows considerable uncertainty.[11] It is also the belief of the ERG that the comparison with gemcitabine plus cisplatin is legitimate despite the differences to the scope identified by the company in the treatment history of the patients in these trials.[13, 14] The main reason for this is that it appears to the ERG that these differences affect comparability in the same way as in all of the other comparator trials and which the company has attempted to adjust for using the STC.

It should also be highlighted that no evidence synthesis of AEs or HRQoL was performed, although the rates for nivolumab did appear to be similar or lower than for the comparators.

In conclusion, it is difficult to be sure what the effectiveness of nivolumab is in comparison to the comparators in the scope. There is evidence from directly examining the single arms of the trial data that there is little difference between the outcomes measured from the nivolumab and comparator studies. Of course, naïve comparison of single arms clearly carries a high risk of bias. However, there is also no clear evidence that this risk of bias would be reduced by the STC analysis. Multiple limitations in the STC were identified and a judgment of the influence of the adjustment due to the STC cannot be evaluated because the company did not present an unadjusted (naïve) analysis. The ERG was able to estimate the unadjusted hazards, but not with estimates of uncertainty. The effect of an analysis based on a different prediction model remains unknown. As stated on page 56 of TSD 18, and used by the company for the basis of the STC: '*The size of this systematic error can certainly be reduced, and probably substantially, by appropriate use of MAIC or STC. Much of the literature on unanchored MAIC and STC acknowledges the possibility of residual bias due to unobserved prognostic variables and effect modifiers; however, it is not made clear that the accuracy of the resulting estimates is entirely unknown, because there is no analysis of the potential magnitude of residual bias, and hence no idea of the degree of error in the unanchored estimates. It is, of course, most unlikely 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]

5. COST EFFECTIVENESS

5.1 ERG comment on company's review of cost effectiveness evidence

This section pertains mainly to the review of cost effectiveness analysis studies. However, the search section (5.1.1) also contains summaries and critiques of other searches related to cost effectiveness presented in the CS. Therefore, the following section includes searches for identifying economic evaluations; studies reporting utility values and; studies reporting cost/resource use data.

5.1.1 Objective and searches performed for cost effectiveness section

The following paragraphs contain summaries and critiques of all searches related to cost effectiveness evidence presented in the CS.

Objective of cost effectiveness analysis search and review

The company performed an SLR with the objective to identify evidence to support the development of a cost effectiveness model for nivolumab as a treatment for locally unresectable or metastatic UC. With a single review, the company aimed to identify relevant UC studies in terms of published:

1. economic evaluations;
2. studies reporting utility values and;
3. studies reporting cost/resource use data.

The CS reported that searches were carried out in December 2016. Searches were not limited by date or by language. A single review was performed to identify relevant studies in UC that included published economic evaluations, studies reporting cost/resource use data, and studies reporting utility values

Searches were carried out on the following databases: Embase, MEDLINE, MEDLINE In-Process, MEDLINE Epub Ahead of Print, HTA and NHS EED via the Cochrane Library and EconLit. Searches were carried out in line with the NICE 2013 guide to the methods of technology appraisal sections 5.2.2 and 5.2.4.[33]

Supplementary searches of the following conference proceedings for 2014-2016 were reported: American Society of Clinical Oncology (ASCO), European Association of Urology (EAU), European Multidisciplinary Meeting on Urological Cancers (EMUC), European Society for Medical Oncology (ESMO), International Society for Pharmacoeconomics and Outcomes Research (ISPOR - Europe and International). The CS also reported searches of the following resources: NICE, SMC and NCPE websites, Cost-Effectiveness Analysis (CEA) Registry, University of Sheffield Health Utilities Database (ScHARRHUD) and EQ-5D Publications Database.

Bibliographies of identified systematic reviews, meta-analyses and HTA submissions were searched for relevant articles.

ERG comment:

- The searches in Appendix G were clearly structured, documented and reproducible, using a wide range of resources to identify published and unpublished literature. Database hosts and dates of searches were all reported. Most database searches used combinations of indexing terms appropriate to the resource searched, free text and a number of synonyms for the condition. Language limits were not applied.
- The EconLit strategy was limited, however due to the database content this is unlikely to have resulted in missed relevant studies.

- Study design filters were applied to the Embase and MEDLINE searches, and although these do not appear to be published validated filters, they contain a wide range of search terms and are therefore unlikely to have missed any relevant studies.
- Search strategies were missing from the CS for NHS EED, the HTA database and EconLit, and for the conference and website searches, however these were supplied in full by the company following a request for clarification.

5.1.2 Inclusion/exclusion criteria used in the study selection

Full details regarding the inclusion/exclusion criteria are provided in Appendix G of the CS (Table 60). In summary the following criteria were used:

- **Patient:** Patients with advanced, metastatic or unresectable UC (mixed populations were excluded unless results were presented separately for those with advanced, metastatic or unresectable)
- **Intervention and comparator:** any intervention or comparator except non-pharmacological interventions, which were excluded
- **Outcomes:** 1) LYs, quality adjusted life years (QALYs) or costs (UK perspective); 2) original health state utility data or; 3) original costs or resource use data relevant to the UK NHS or social work in Scotland or the Health Service Executive in Ireland
- **Study design:** original research or SLR
- **Other:** English language only

5.1.3 Included/excluded studies in the cost effectiveness review

In total 676 references were identified in the SLR. Duplicates (n=100) were excluded, resulting in 576 references for the title and abstract screening. During this process 539 references were excluded (22 due to reference not being in English/not in human participants). After full-text screening of the remaining 37 references, another 31 references were excluded (see Appendix G of the CS (Table 61) for the reason for exclusion per study). After including three references identified by hand search, nine references (seven unique studies) were included, including three economic evaluations.[34-36] See Appendix G of the CS (Figure 29) for the PRISMA diagram. The included studies are summarised in Appendices G.2.1, G.2.2, H and I of the CS.

5.1.4 Conclusions of the cost effectiveness review

Although economic evaluations were identified with populations that matched the population described in the final scope of this appraisal, these did not consider the cost effectiveness of nivolumab and therefore a de novo health economic analysis was conducted for the purposes of this appraisal.

In the vast majority of the studies that report original health-state utility data, no EQ-5D health state descriptions were used, and the studies did not report full details of the elicitation and valuation methods. Therefore, none of the included utility studies were deemed consistent with the NICE reference case for use in the health economic model. To inform the utility values for the economic model, the company used EQ-5D-3L data collected from the CheckMate 275 trial. Additionally, the disutilities for Grade 3 and 4 AEs were derived from the literature (CS Table 35). However, it was unclear how these studies were identified (as these studies were not retrieved from the SLR).

One of the identified resource use and cost studies was used to retrieve the AE costs for leukopoenia (CS Table 41). Although other literature sources were used (e.g. for terminal care costs and costs for other AEs), it was unclear how these studies were identified (as these studies were not retrieved from the SLR).

ERG comment: Since the identified cost effectiveness studies were not performed using the intervention of interest, the ERG agrees that conducting a de novo health economic analysis was necessary. Relevant health-state utility, as well as resource use and cost studies were identified by the company. It was however unclear why the company used literature sources not identified in the SLR to inform the model and not for instance TA272 (the only other NICE submission in this indication), which was identified in the SLR. Additionally, it was unclear how these alternative literature sources were identified.

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

Table 5.1: Summary of the company's economic evaluation (with signposts to CS)

	Approach	Source / Justification	Signpost (location in CS)
Model	A cohort-based partitioned survival model was implemented in Excel	To capture the progressive nature of UC disease and to provide consistency with previous NICE submissions relating to metastatic cancers.	Section B.3.2.2
States and events	Health states: - Progression-free state - Progressed disease state - Death	To be in line with previous NICE submissions relating to metastatic cancers, including the only previous submission in this indication (TA272, 2013)[34]	Section B.3.2.2
Comparators	- Paclitaxel - Docetaxel - Best supportive care - Cisplatin + gemcitabine (only in scenario analysis)	Paclitaxel, docetaxel and BSC were included to be consistent with the scope. The scope also specified cisplatin + gemcitabine as a comparator but this was only included in scenario analysis because of limited evidence on cisplatin + gemcitabine for retreatment with first-line platinum-based chemotherapy.	Section B.3.2.3
Population	Patients with metastatic or unresectable UC who have progressed following first-line platinum-based chemotherapy.	This is consistent with the population of the CheckMate 275 and 032 trials, as well as the final scope issued by NICE.	Section B.3.2.1
Treatment effectiveness	Treatment effectiveness was estimated in terms of gains in OS and PFS that nivolumab could provide over the comparators. Estimates were informed by the CheckMate 275 and 032 single-arm studies, using response-based survival analysis	A response-based modelling approach to estimate OS and PFS was adopted in order to reflect the mechanism of action of nivolumab and that the nivolumab survival curve changes over time as the hazard changes. According	Sections B.3, B.3.3.1 and B.3.3.2

	Approach	Source / Justification	Signpost (location in CS)
	implemented using landmark analysis, where responders and non-responders were modelled separately from the chosen 8-week landmark. A simulated treatment comparison informed time-varying hazard ratios for nivolumab versus each comparator.	to the company, standard parametric models were deemed unlikely to be flexible enough to characterise this change in the hazard. To overcome immortal time bias, landmark analysis was used. It was necessary to generate time-varying hazard ratios as the proportional hazard assumption did not hold for these comparators given the unique mechanism of action of nivolumab.	
Adverse events	Resource use, costs and utility decrements were considered for Grade 3 and 4 AEs.	To represent those AEs that are more likely to have an effect on quality of life.	Sections B.3.4.4, B.3.4.5 and B.3.5.1
Health related QoL	The HRQL data used in the cost effectiveness analysis for the progression-free and the progressed disease state were derived from EQ-5D-3L data collected in CheckMate 275 and analyses using a mixed model. Disutilities for AEs were also included; these were derived from the literature.	None of the studies identified through the SLR were deemed to be consistent with the NICE reference case.	Section B.3.4
Resource utilisation and costs	Resource use and costs in the model consisted of drug acquisition costs and drug dosing, drug administration and monitoring, costs associated with best supportive care, treatment discontinuation, terminal care and AEs. These were based on information from CheckMate 275, the BNF, EMIT, published sources identified in the SLR and expert clinician feedback.	CheckMate and published sources were used when they provided estimates of resource use and costs. In the absence of such estimates, assumptions were made and validated through discussions with clinicians.	Section B.3.5
Discount rates	Discount rate of 3.5% for utilities and costs	As per NICE reference case	Table 42
Sub groups	None	As per NICE scope	Section B.3.9

	Approach	Source / Justification	Signpost (location in CS)
Sensitivity analysis	Both DSA and PSA were performed as well as scenario analyses	The PSA excluded key parameters.	Sections B.3.8

Source: CS
Abbreviations: AE, adverse events; BNF, British National Formulary; BSC, best supportive care; CS, company submission; DSA, deterministic sensitivity analysis; EMIT, electronic market information tool; HRQL, health-related quality of life; OS, overall survival; PFS, progression-free survival; PSA, probabilistic sensitivity analysis; SLR, systematic literature review; UC, urothelial cancer

5.2.1 NICE reference case checklist

Table 5.2: NICE reference case checklist

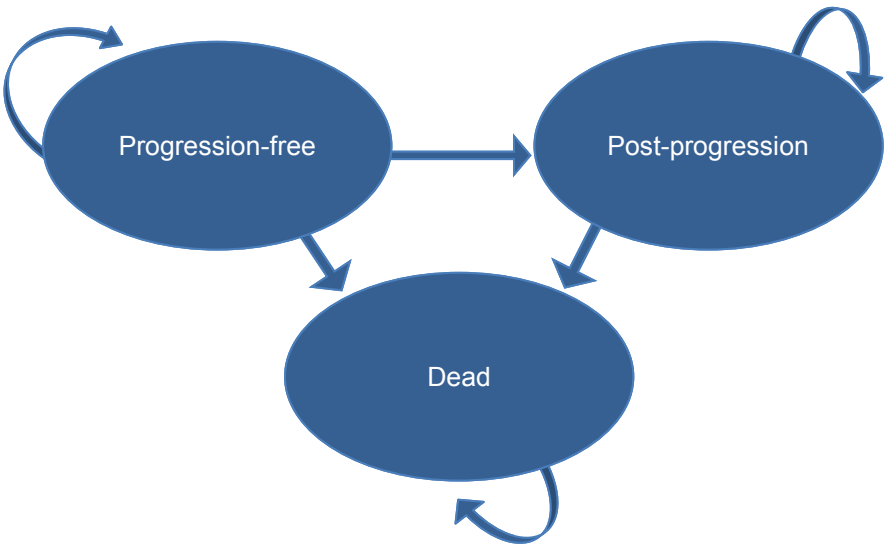
Elements of the economic evaluation	Reference Case	Included in submission	Comment on whether *de novo* evaluation meets requirements of NICE reference case
Population	As per NICE scope	Yes	
Comparator(s)	Therapies routinely used in the National Health Service (NHS), including technologies regarded as current best practice	Partly	Comparator cisplatin + gemcitabine was identified in NICE scope but only included in scenario analysis.
Type of economic evaluation	Cost effectiveness analysis	Yes	
Perspective on costs	NHS and Personal Social Services (PSS)	Yes	
Perspective on outcomes	All health effects on individuals	Yes	
Time horizon	Sufficient to capture differences in costs and outcomes	Yes	
Synthesis of evidence in outcomes	Systematic review	Yes	
Measure of health effects	Quality adjusted life years (QALYs)	Yes	
Source of data for measurement HRQoL	Described using a standardised and validated instrument	Yes	
Source of preference data for valuation of changes in HRQoL	Time-trade off or standard gamble	Yes	
Discount rate	An annual rate of 3.5% on both costs and health effects	Yes	

Elements of the economic evaluation	Reference Case	Included in submission	Comment on whether *de novo* evaluation meets requirements of NICE reference case
Equity weighting	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	Yes	
Sensitivity analysis	Probabilistic modelling	No	The PSA does not incorporate all relevant parameters (the HRs, a key parameter in the model, are not reflected in the PSA).
Source: CS Abbreviations: HR, hazard ratio; NHS, National Health Service; PSA, probabilistic sensitivity analysis; PSS, Personal Social Services; QALY, quality-adjusted life year			

5.2.2 Model structure

The company developed a de novo economic model using a cohort-based partitioned survival model. The model consists of three mutually exclusive health states: progression-free (PF) and post-progression (PP) disease states and death. Patients enter the model in the PF state and are treated with nivolumab or one of its comparators. Patients remain in the PF state until disease progression or death. The proportion of patients in each health state changes over time and is determined by the OS and PFS curves, which are treatment dependent. Patients cannot move from the PP state back to the PF state. This model structure was chosen to capture the progressive nature of UC disease and to be consistent with previous submissions to NICE relating to metastatic cancers, including the previous submission in this indication (TA272, 2013)[34]. The model structure is depicted in Figure 5.1.

Figure 5.1: Partitioned survival model structure



ERG comment: The ERG's comments include (1) a critique of the choice of partitioned survival analysis for this decision problem and (2) the use of response-based analysis without reflecting responder and non-responder states in the model structure.

(1) The recent TSD 19 critiques partitioned survival analysis modelling in cancer appraisals.[37] It is stated that it is the most commonly used decision modelling approach in advanced or metastatic cancer. Limitations of the method include that (1) survival functions are modelled independently even though there are dependencies such as that progression is a prognostic factor for mortality, (2) transition probabilities are not estimated for each possible transition between health states. These limitations are especially evident in the extrapolation beyond trial data (before that, dependencies are reflected in the data) and can lead to inappropriate extrapolation [37]. This can, for example, be caused by mortality hazards being extrapolated independently of progression, whilst the mix of progressed and non-progressed patients changes over time (at a certain time all patients will have progressed), or by inappropriate reflection of the treatment effect mechanism in the estimated long-term hazards. Alternatives include other types of transition models, as well as a hybrid modelling approach, by which patients were first allocated to a treatment response category using a decision tree, and second a partitioned survival analysis approach was used. The company, in response to clarification questions, stated that other model structures were not explored.[7] Based on TSD 19, the ERG considers that alternative model structures should and will be considered more frequently in the future, but the company's approach is consistent with past technology appraisals.

(2) The company used a response-based approach to modelling overall survival (OS) and progression-free survival (PFS), but does not reflect the resulting responder and non-responder groups in their model structure. The combination of these groups introduces a superfluous assumption, which is that the proportions of responders and non-responders remain the same throughout the model time horizon. This assumption is unrealistic given that responders are likely to survive longer compared to non-responders, resulting in an increase in the proportion of responders over time. Had the company kept these two groups separate by allowing for differential responder and non-responder health states, the change in responder and non-responder proportions over time would have been reflected automatically. The company argued in their response to clarification questions that it was not possible to keep these two groups separate because the STC required a larger sample size to estimate HRs for responders and non-responders separately.[7] The ERG wishes to highlight that it is not necessary to estimate separate HRs for the two groups and that this was explained in detail at the preliminary teleconference to discuss the clarification questions. The same HR could have been applied to both groups, as is done in the model currently.

5.2.3 Population

The model includes patients with metastatic or unresectable UC who have progressed following first-line platinum-based chemotherapy. Patient characteristics included in the model were age, gender, weight and body surface area (BSA). These were based on the CheckMate 275 study[10].

ERG comment: This patient group is consistent with the population of the CheckMate 275 and CheckMate 032 trials, as well as the final scope issued by NICE for this appraisal. Age and gender estimates are relevant for the calculation of background mortality and are further discussed in Section 5.2.6. Weight and BSA influence the calculation of dose and there is a discussion about this in Section 5.2.9.

5.2.4 Interventions and comparators

Nivolumab is implemented as per its EMA Summary of Product Characteristics (SmPC) posology and method of administration for second-line UC (i.e. 3mg/kg Q2W).

The company considered the following comparators in their base-case:

- Paclitaxel: 80mg/m^2 Q3W of a four-week cycle
- Docetaxel: 75mg/m2 Q3W
- Best supportive care (BSC)

The company also presented a scenario analysis, in which cisplatin + gemcitabine was added as a comparator. The company justified this deviation from the scope (i.e. not including cisplatin + gemcitabine in its base-case) by stating that there was limited evidence for retreatment with first-line platinum-based chemotherapy regimens for patients with locally advanced unresectable or metastatic UC. The SLR had not identified any relevant trials for this comparator. The only available data stemmed from a trial in which cisplatin + gemcitabine was used in re-challenge[13], assuming a gemcitabine-naïve patient population. The company argued that this study was non-generalisable to the UK, where it is standard clinical practice that patients would receive cisplatin plus gemcitabine as first-line treatment, and where different dosing schedules from the ones in the study are used.

ERG comment: The ERG requested that the company provide the comparison of nivolumab with cisplatin plus gemcitabine in the base-case, but the company did not provide this analysis within the base-case analysis. The company justified this in their response to clarification question A15[7] citing expert opinion stating that the population in the Gondo (2011) study[13] differed from the UK population in that the study population received MVAC in first line instead of cisplatin plus gemcitabine. The ERG challenges the position of the company in that patients in the Gondo (2011) study[13] would have had exposure to platinum-based therapy (part of MVAC is cisplatin) 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. Indeed, if that was a valid argument, the other comparisons could not be performed either because no RCTs were available. The company could have adjusted the available data based on expert opinion. It is the ERG's view that the company did not present valid arguments to exclude cisplatin plus gemcitabine as a comparator and the ERG will therefore include this comparison in its base-case based on the data from Gondo (2011)[13].

5.2.5 Perspective, time horizon and discounting

The model adopts the perspective of the NHS and Personal and Social Services (PSS) in England and Wales. The cycle length is four weeks to account for the length of treatment cycles. 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.

ERG comment: The ERG considers the adopted perspective, time horizon and discounting to be appropriate for this appraisal.

5.2.6 Treatment effectiveness and extrapolation

Parametric time-to-event models were used to estimate OS, PFS and TTD in the company's cost effectiveness model. A response-based approach was adopted to estimate OS and PFS, but not for TTD in the company's base-case.

5.2.6.1 OS and PFS of nivolumab

The parametric time-to-event models representing OS and PFS of nivolumab were informed by the CheckMate 032 and CheckMate 275 trials, which are both single arm trials.[10, 11] The time-to-event data of both trials were combined (pooling method not stated) to perform the survival analyses described in the following sections.

Response-based and landmark analyses

The company implemented a response-based analysis to estimate OS and PFS of the nivolumab arm because it claimed that standard survival modelling approaches would not appropriately characterise the novel mechanism of action of nivolumab, i.e. responders may have long and durable response to treatment leading to extended survival. Therefore the company suggested that standard parametric time-to-event models were not deemed flexible enough to characterise the change in hazard over time resulting from having (long-term) responders, and non-responders (no supporting evidence provided).[2]

The company used a landmark analysis to prevent the occurrence of the immortal-time bias. In this landmark analysis, OS and PFS of both groups (responders and non-responders) were estimated together until a specified landmark point after which different survival curves were fitted separately for each group. For the base-case analysis, the company chose an eight-week landmark point, which corresponds to the median time to response in both CheckMate 032 and CheckMate 275 trials (1.87 and 1.48 months in CheckMate 275 and CheckMate 032, respectively). Before this eight-week landmark point, the Kaplan-Meier estimates for the whole group were used to estimate OS and PFS. After the landmark point, parametric time-to-event models were fitted to the responders' and non-responders' survival data for the remainder of the time horizon, and adjusted for background mortality.

A sensitivity analysis explored the impact of using a 26-week landmark point, with the justification that, at that time point, '*all patients had responded while leaving a sufficiently long observational period for further extrapolation.'*[2]

ERG comment: The main concerns of the ERG were (1) the method used for pooling both CheckMate 032 and CheckMate 275 trials, (2) the use of response-based analysis, (3) the use of landmark analysis to model PFS and OS of nivolumab, and (4) the use of KM estimates up to the chosen landmark.

(1) The CS reported that data from both CheckMate 032 and CheckMate 275 studies were pooled without stating which method was used to pool the data. Upon request from the ERG, the company explained that OS and PFS data from both studies were combined without adjustments because there was no evidence of differences between the studies based on a Wald test. Hence, the pooled CheckMate studies dataset contained 348 patients (78 patients from CheckMate 032 and 270 patients from CheckMate 275).[7] Concerns with pooling from CheckMate 032 and CheckMate 275 studies were outlined in Section 4.2.6.

(2) The company justified the use of a response-based approach stating that standard parametric time-to-event models were not flexible enough to characterise the change in hazard over time due to possible sustained and long-term response to treatment. However, the ERG noted that most standard parametric time-to-event models include changing hazards over time; some standard parametric time-to-event models allow for non-monotonic changing hazard functions over time (i.e. log-logistic, log-normal and generalised gamma distributions). The company did not provide any mathematical reasoning to support their argument that a different response cannot be accurately described by standard parametric survival models. The ERG considers that based on visual inspection of the not response-based, conventional survival analysis alone, the case for response-based analysis might not be supported, as the parametric time-to-event model fitted to OS made a good fit and the model for PFS could be regarded as providing a reasonable fit (see Figures 5.2 and 5.3).

The company's second argument in favour of the landmark analysis was that it was implemented to address concerns from previous appraisals of nivolumab in which standard parametric time-to-event models were not deemed suitable to model survival with nivolumab treatment.[7] The company argued in response to clarification questions that landmark analysis *'allows for a more flexible shape to the*

nivolumab survival curve whilst adhering to the Committee's previous preference of using the trial data for a proportion of the survival curves.'[7] The ERG considers that a standard approach should be shown to be inappropriate in the particular decision problem at hand before discarding it and the company failed to do so, as described in the previous paragraph.

The ERG requested that the company justify whether alternative methods (e.g. spline models, mixture cure models) were considered instead of the landmark analysis because spline models are suggested in the NICE DSU TSD 14 as a flexible alternative to standard parametric time-to-event models (while the landmark approach is not mentioned).[38] The company responded that spline models were generally not accepted in previous appraisals of nivolumab and that the acceptability of mixture cure models for HTA bodies is yet unknown. The ERG considers that this is not a valid argument given that spline models and mixture cure models are recommended in the TSD.

In conclusion, the company (a) did not provide sufficient evidence to demonstrate that conventional parametric time-to-event models failed to describe nivolumab survival, (b) did not provide evidence to support that the committee's criticisms on previous nivolumab appraisals applied to the current appraisal, and (c) did not provide evidence to demonstrate that the landmark analysis provided more valid results than standard survival modelling analyses or alternative methods recommended in NICE DSU TSD 14 (for example, no expert opinion was used to validate the resulting curves).

(3) The ERG's third concern is the choice of the eight-week landmark. The choice of the eight-week landmark was based on the collected evidence while it is advised to determine the landmark point a priori to the analysis in order *'to safeguard the analysis against the danger of a data-driven decision'*.[39] Therefore, the ERG asked the company to investigate the influence of a 12- or 20-week landmark point on the results but these analyses were not provided by the company due to time constraints. As demonstrated in a previous nivolumab appraisal, the choice of the landmark point may not have a linear relationship with the ICER.[40] Hence, the influence of this assumption, i.e. the arbitrarily post-hoc selected landmark point, on the results is highly unpredictable.

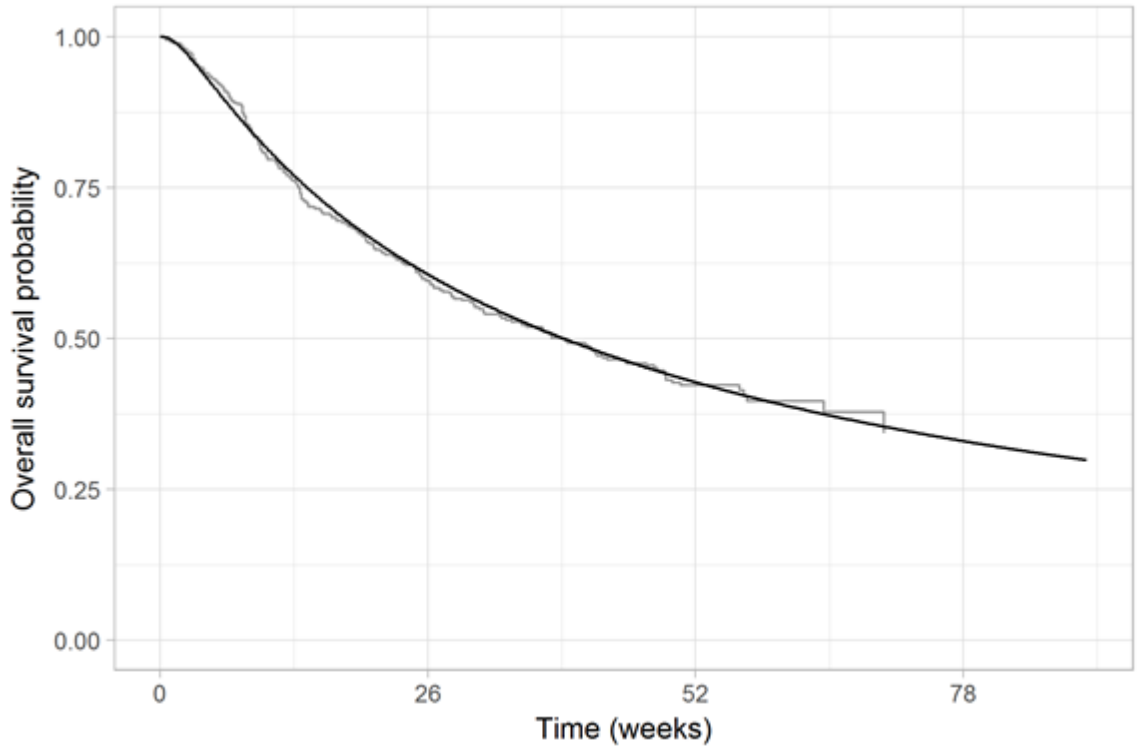
(4) The ERG asked the company to justify why the Kaplan-Meier estimates were used until the landmark point instead of a parametric time-to-event model, and to provide the results of an analysis using a parametric time-to-event model until the landmark point. The company did not provide the results of such analysis and responded that using the Kaplan-Meier estimates until the landmark point reflected the *'Committee's previous preference of using the trial data for a proportion of the survival curves.'*, not clearly referring to a specific technology appraisal.[7] According to the company, using a parametric time-to-event model would also add unnecessary complexity to the model. The ERG does not consider these arguments to be valid: a previous precedent does not relieve the company from demonstrating appropriateness of their method, and fitting a distribution to the data up to the landmark does not present more complexity than making Kaplan-Meier estimates probabilistic. The ERG therefore prefers the use of a parametric time-to-event model to estimate survival until the landmark point to avoid the problem of overfitting when using Kaplan-Meier estimates. The possibility for this analysis was, however, not included in the company's model.

In conclusion, the company deviated from the NICE TSD recommendations by using a response-based analysis. However, the company did not demonstrate (1) that conventional modelling approaches of survival failed to correctly characterise the OS and PFS of nivolumab, and (2) that the response-based approach resulted in estimates that could be considered more realistic than the standard approach. The uncertainty about whether this approach more accurately reflects prognosis for patients treated with nivolumab was exacerbated by additional assumptions required for response-based analysis, such as, most crucially, the choice of the landmark point, which has an unpredictable effect on results. Fitting

parametric models to the responder and non-responder groups also results in larger uncertainty about these fitted curves: the sample size used is significantly smaller, a) because of the splitting up of the study population into two groups and b) because only the available data after the landmark is used. The fact that responder and non-responder groups had to be combined for the indirect comparison casts further doubt over whether the response-based analysis has any benefits (hazard ratios are derived from the overall population and are then applied in a combined responder and non-responder population, as described below in the section on relative treatment effectiveness). It should also be noted that response-based and conventional approaches result in vast differences in the predicted life-years for nivolumab, with a predicted mean of 2.80 life years in the response-based analysis and 1.84 life years in the conventional, not response-based, approach (deterministic estimates). No explanation for this deviation was provided, and these estimates were not validated using expert opinion.

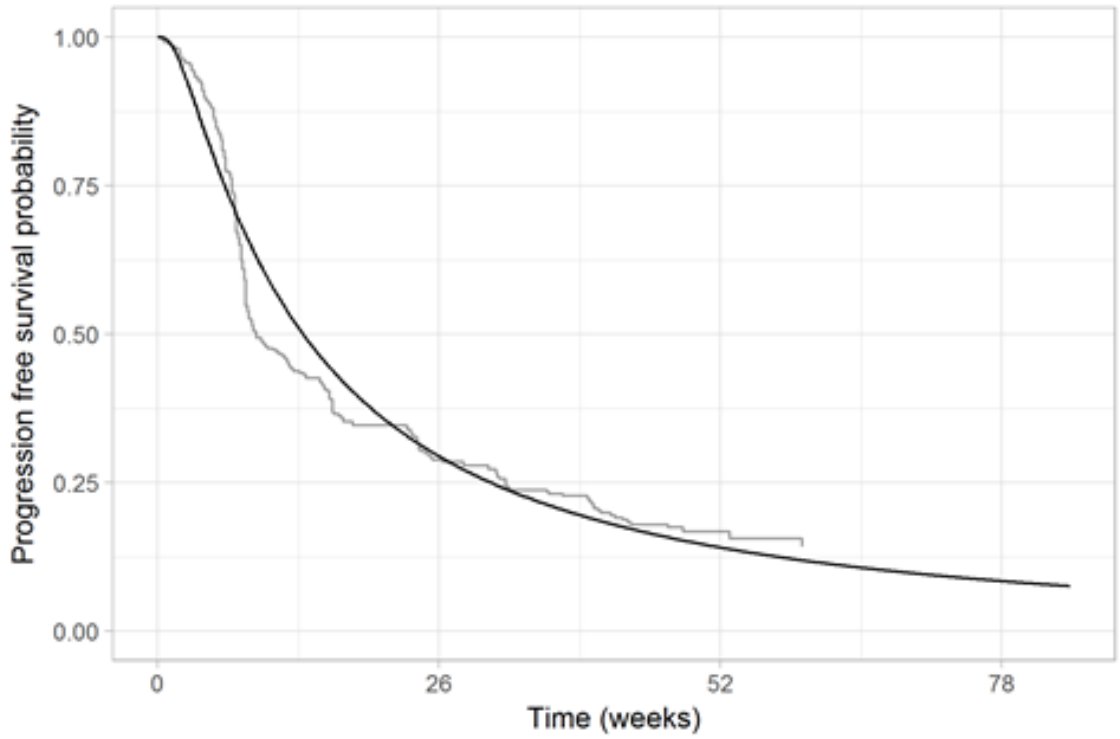
For the aforementioned reasons, and in line with the TSD recommendations, the ERG used the conventional approach of fitting parametric time-to-event models to the overall population in its base-case analysis. Based on statistical fit and visual inspection, the ERG considers the distributions preferred by the company (i.e. the generalised gamma for both OS and PFS) to be the most plausible in its base-case analysis (Figures 5.2 and 5.3). Alternative distributions are explored in scenario analyses. The ERG also explored the use of a response-based analysis in scenario analyses.

Figure 5.2: Standard parametric time-to-event model for overall survival (generalised gamma distribution)



Source: Appendix L, figure 114

Figure 5.3: Standard parametric time-to-event model for progression-free survival (generalised gamma distribution)



Source: Appendix L, figure 120

Time-to-event models selection for OS and PFS estimations of nivolumab

Parametric time-to-event models were fitted separately to the OS and PFS data of the responder and non-responder groups (without investigating the proportional hazard assumption through log-cumulative hazard plots). The company stated that the following six parametric distributions were fitted to the OS and PFS data as recommended by the NICE Decision Support Unit Technical Support Document 14[38]:

- Exponential
- Weibull
- Gompertz
- Lognormal
- Log-logistic
- Generalised gamma

The parametric time-to-event models used to estimate OS and PFS were selected based on statistical fit (Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC)) and visual inspection. Table 5.3 provides an overview of the statistical fit of the different distributions for OS and PFS in the responder and non-responder groups.

The company considered the model selection for OS and PFS (in both responders and non-responders groups) simultaneously and selected the generalised gamma distribution to represent OS and PFS of both responder and non-responder groups. The generalised gamma distribution was selected because 1) it was the best fitting distribution based on 3 out of 8 criteria (see numbers printed in bold in Table 5.3), and 2) the Weibull distribution (which was the best fitting distribution based on 4 out of 8 criteria) provided a poor fit to the responders' OS and PFS (unclear how this was determined). Hence, the

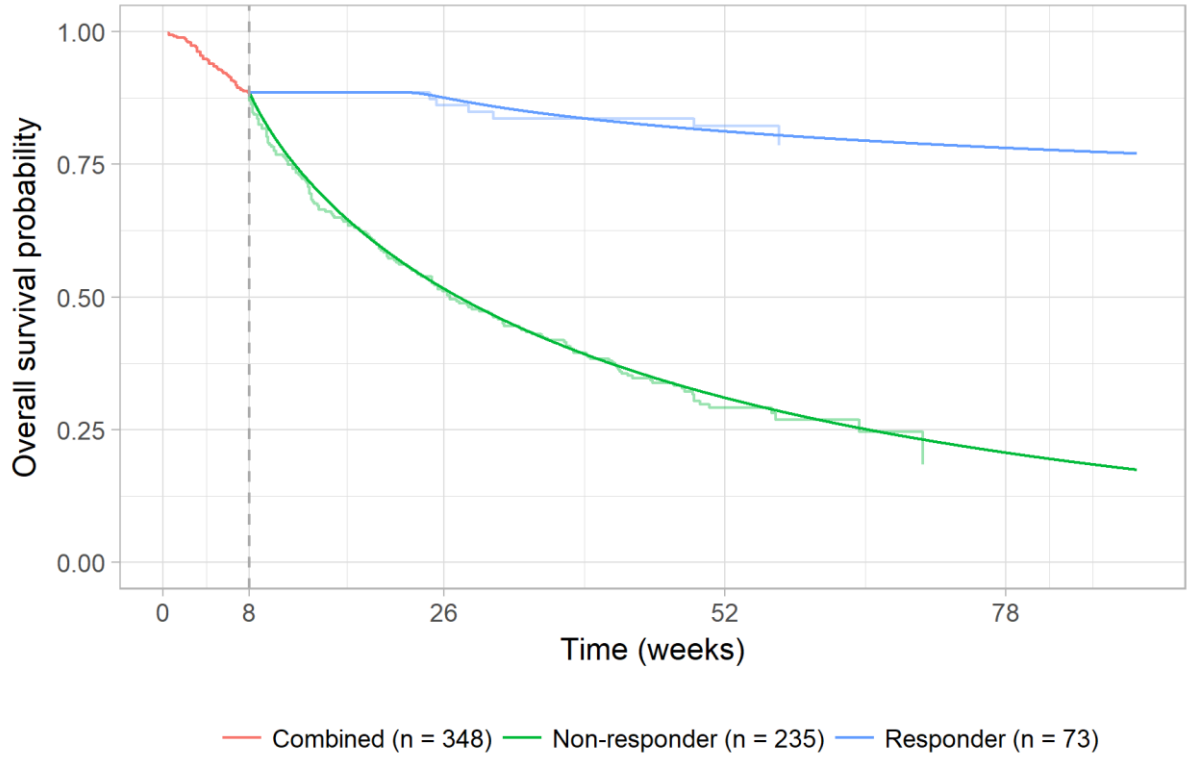
company concluded that the generalised gamma provided the best fit overall. Experts were not consulted to support the selection of the parametric time-to-event models applied to the responder and non-responder groups. Figures 5.4 and 5.5 present the landmark analyses for OS and PFS based on responders' status.

Table 5.3: Statistical fit measures of the distributions representing OS and PFS in the responder and non-responder groups at the eight-week landmark

Distribution	OS				PFS			
	Responders		Non-responders		Responders		Non-responders	
	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC
Exponential	90.1	**92.4**	1402.7	1406.1	276.9	279.2	787.8	790.6
Weibull	91.1	95.7	**1393.3**	**1400.2**	266.9	271.5	**763.4**	**769.2**
Gompertz	91.9	96.4	1395.4	1402.3	273.1	277.7	780.7	786.4
Lognormal	90.4	95.0	1397.4	1404.3	262.4	267.0	773.1	778.8
Log-logistic	91.0	95.6	1394.4	1401.3	264.6	269.2	776.7	782.4
Generalised gamma	**87.9**	94.8	1394.5	1404.8	**256.6**	**263.5**	765.0	773.5

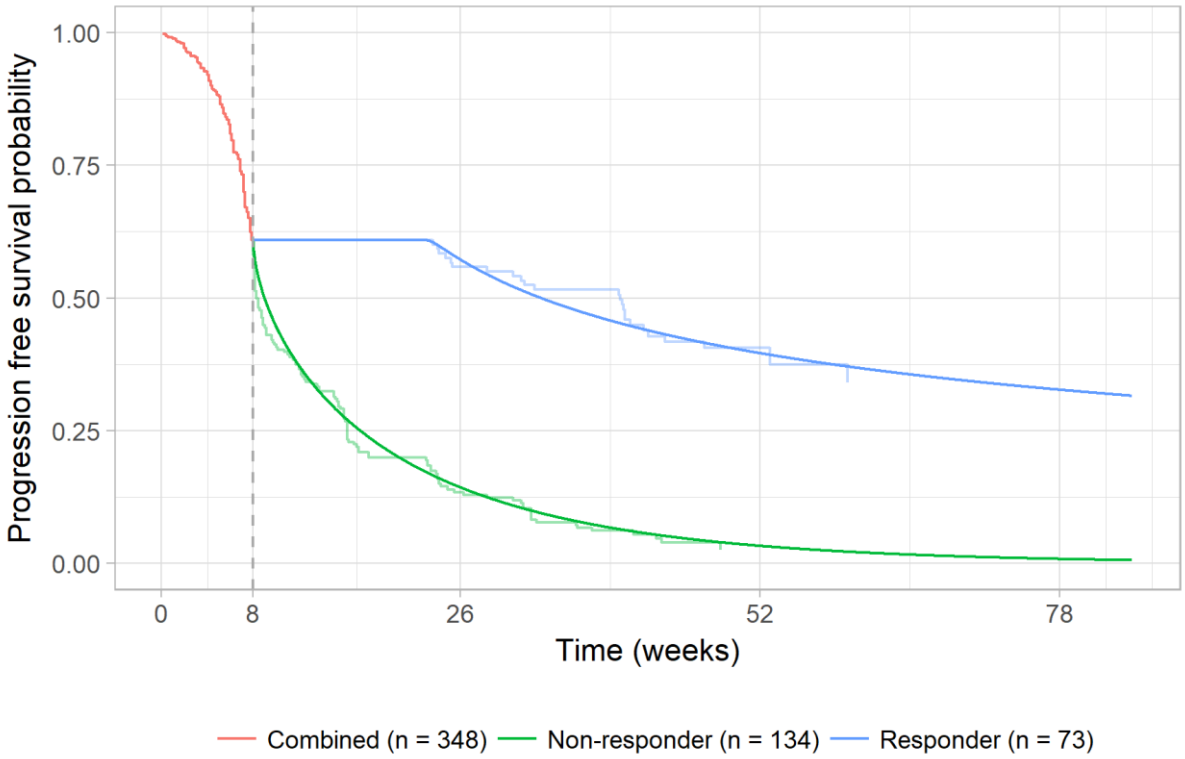
Source: Adapted from Table 29 of the CS[2]
Abbreviations: AIC, Akaike Information Criterion; BIC, Bayesian Information Criterion; OS, overall survival; PFS, progression-free survival.
Bold printed values represent the distributions with the lowest AIC or BIC (i.e. the 'best fitting' time-to-event models)

Figure 5.4: Week 8 landmark – overall survival with generalised gamma[a]



Source: Response to clarification letter, Figure 34[7]
[a] The ERG requested corrected figures because the number of responder was incorrect in the original CS

Figure 5.5: Week 8 landmark – progression-free survival with generalised gamma[a]



Source: Response to clarification letter, Figure 35[7]

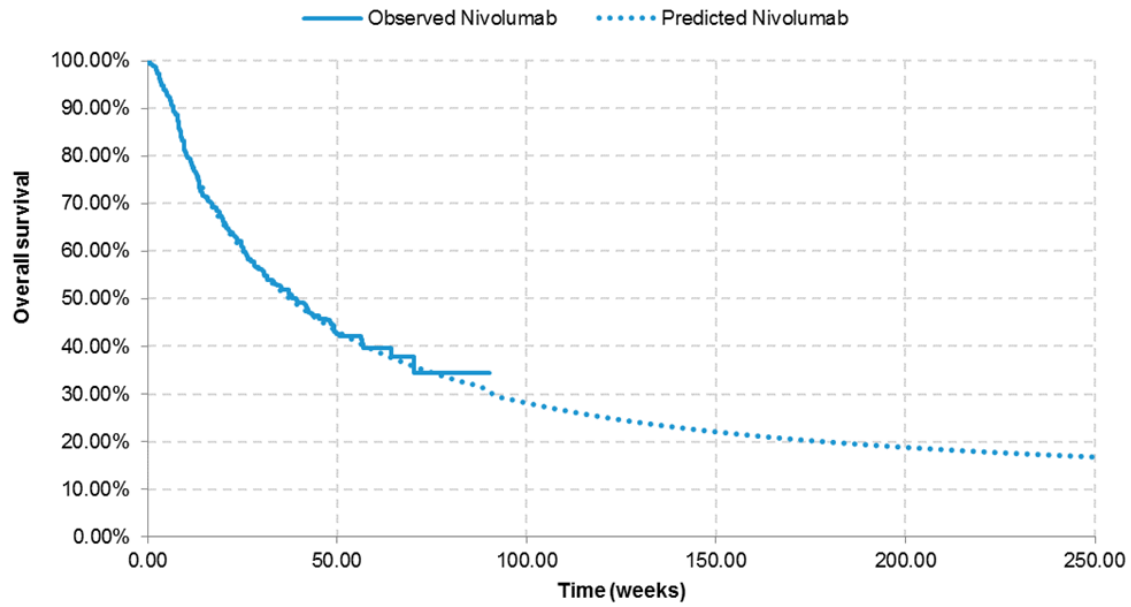
[a] The ERG requested corrected figures because the number of responder was incorrect in the original CS

In order to implement the parametric time-to-event models in the cost effectiveness model, OS and PFS estimates obtained from the parametric time-to-event models estimated for responders and non-responders separately were combined by using a weighted average). This weighting was based on the proportion of responders in patients being progression-free and alive at the eight-week landmark point (based on both CheckMate 032 and CheckMate 275 trials[10, 11], and was assumed to stay constant for the remainder of the time horizon.

Figures 5.6 and 5.7 present the survival curves as used in the base-case analysis for OS and PFS, respectively, compared to the observed OS and PFS obtained with nivolumab. These curves are the result of the weighted average of the responders' and non-responders' OS and PFS estimates, and are compared to the OS and PFS Kaplan-Meier estimates of the pooled CheckMate studies dataset.

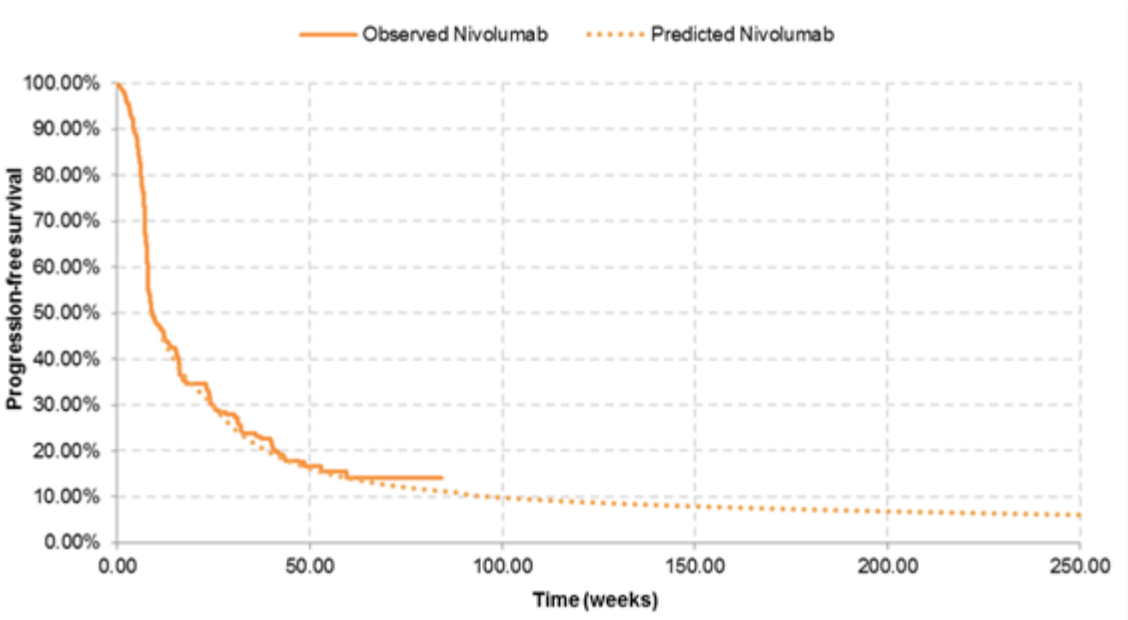
Figure 5.6: Kaplan-Meier estimate of OS with nivolumab, based on the pooled CheckMate 032 and CheckMate 275 trials dataset ('Observed Nivolumab') compared to the predicted values

based on the landmark and response-based analysis (generalised gamma distribution) ('Predicted Nivolumab')



Source: Figure 37 of the CS[2]

Figure 5.7: Kaplan-Meier estimate of PFS with nivolumab, based on the pooled CheckMate 032 and CheckMate 275 trials dataset ('Observed Nivolumab') compared to the predicted values based on the landmark and response-based analysis (generalised gamma distribution) ('Predicted Nivolumab')



Source: Figure 36 of the CS[2]

ERG comment: The main issues concerning the selection of the parametric time-to-event models are (1) the rejection of the proportional hazard assumption between responders and non-responders, (2) the simultaneous selection of the parametric time-to-event models, (3) the lack of expert consultation, and

(4) the combination of the responders' and non-responders' curves at a weight which stays constant over time.

(1) The company assumed in its base-case analysis that the proportional hazard assumption did not hold between responders and non-responders, but did not provide log-cumulative hazard plots to support this assumption. Upon the ERG's request, the company provided the log-cumulative hazard plots and concluded that the proportional hazard assumption could potentially be valid for OS but not for PFS. However, the company did not assume proportional hazards *'as this meant there was no requirement to assume the same distribution to be appropriate for both responder and non-responder curves'*.[7] The ERG does not agree with this argument since the proportional hazard assumption seemed to hold for OS, and could potentially also hold for PFS, based on the examination of the log-cumulative hazard plots. No additional evidence was provided to discard the proportional hazard assumption based on clinical implausibility of the assumption. The influence of assuming proportionality of hazards and using a hazard ratio on one of the curves on the results was not investigated by the company.

(2) In the base-case model, the company selected the same distributions (generalised gamma) for responder and non-responder groups without justifying why. This contradicts the company's argument that there was *'no requirement to assume the same distribution to be appropriate for both responder and non-responder curves'*.[7] This decreased the flexibility allowed by the different parametric time-to-event models. In response to the clarification questions, an updated model was provided by the company, which allowed the selection of different parametric time-to-event models for responders and non-responders.

(3) The NICE DSU TSD 14 recommends to consult clinical experts to support the choice of the parametric time-to-event models besides using statistical fit and visual inspection.[40] According to the CS and response to clarification questions,[7] clinical experts were only consulted during an advisory board. The survival curves presented during this advisory board were fitted to the CheckMate 275 trial only and did not include response-based analysis.[5] The final parametric time-to-event models were therefore not validated using expert opinion.

(4) The parametric time-to-event models were fitted separately to responders and non-responders and were weighted based on the proportions of responders and non-responders at the landmark point. This inflated the proportion of non-responders in later periods because the proportion of responders is expected to increase over time compared to the proportion of non-responders. This assumption is likely to be conservative but it is not clear, and, as described in Section 5.2.6.1, using different landmark points may have an unpredictable influence on the results.

In conclusion, most issues identified in the selection of parametric time-to-event models are avoided by using conventional analysis, as opposed to response-based analysis. These issues include the pooling of responder and non-responder groups, making assumptions about proportional hazards between the two groups and the potential for using differential curves for responders and non-responders. Therefore, the ERG used the conventional approach in its base-case analysis using the company's base-case and alternative parametric time-to-event models. As mentioned before, the influence of using a response-based analysis will be explored in the ERG's scenario analyses, using the company's base-case and alternative parametric time-to-event models.

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.

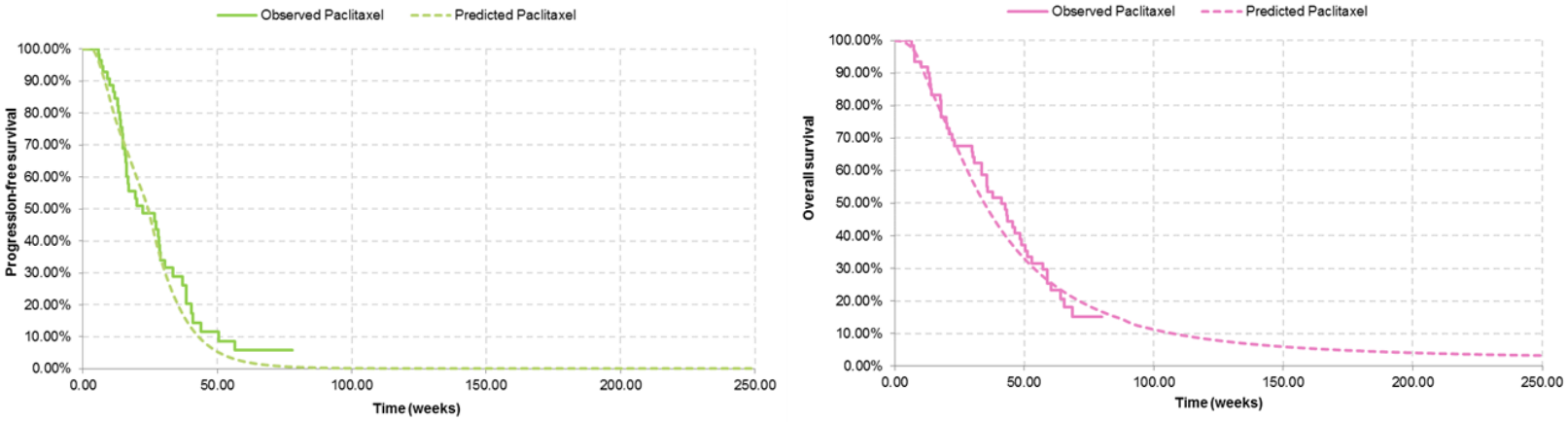
(2) Not in line with conventional methods of incorporating background mortality in parametric survival models, the company used a distribution of age instead of a fixed mean age, to reflect patient heterogeneity. This resulted in slightly higher background mortality compared to standard background mortality 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. 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. PFS was not directly adjusted using the general population mortality data but a minimum function was implemented to ensure that PFS did not become higher than OS.

5.2.6.2 Relative effectiveness of nivolumab

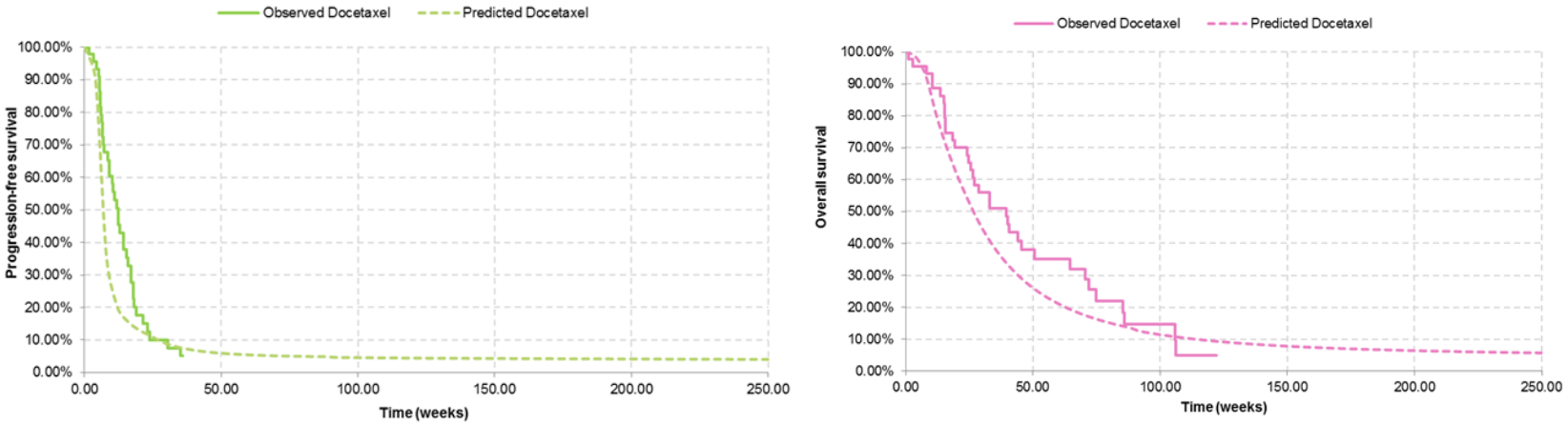
The relative effectiveness of nivolumab versus the comparators was modelled through time-varying hazard ratios (HRs) because the '*proportional hazard assumption did not hold for these comparisons given the unique mechanism of action for nivolumab*'.[2] No evidence was provided to support the violation 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 docetaxel, because of the differences in patient characteristics between the comparator trials and the CheckMate 032 and CheckMate 275 studies.

Figure 5.8: Progression-free survival and overall survival with paclitaxel – observed and predicted values with the generalised gamma distribution



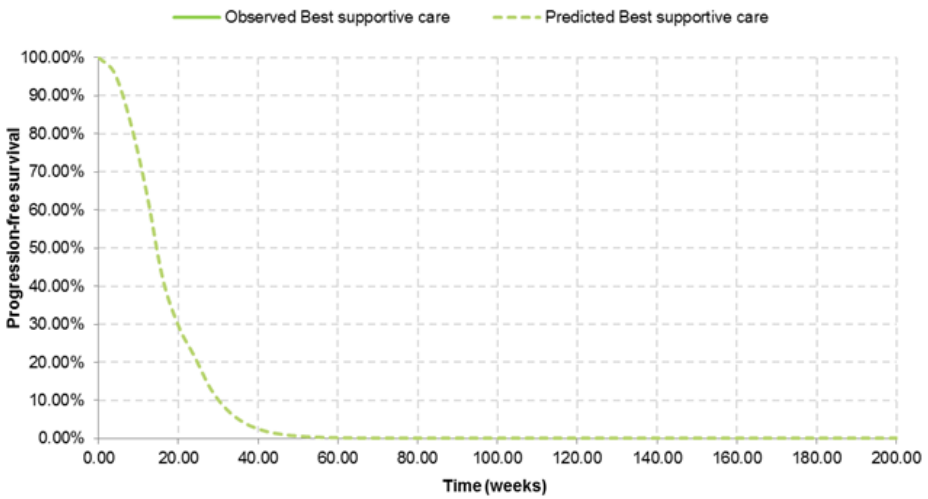
Source: Figure 40 of the CS

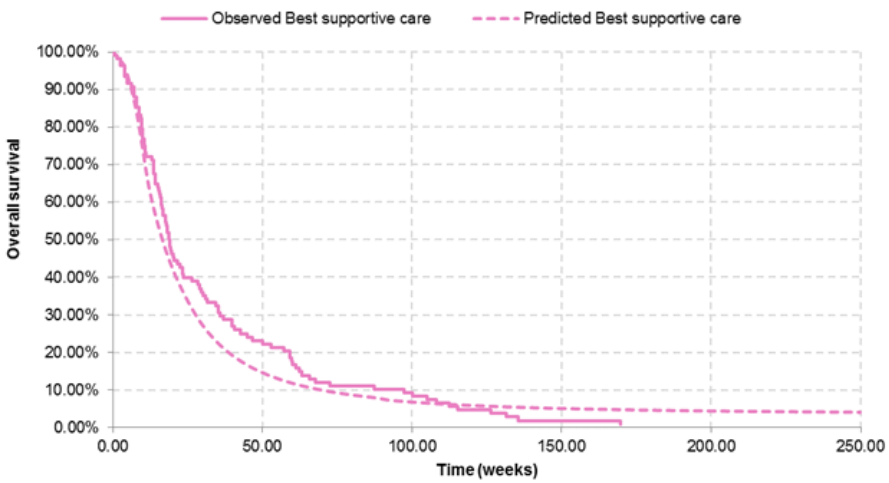
Figure 5.9: Progression-free survival and overall survival with docetaxel – observed and predicted values with the generalised gamma distribution



Source: Figure 40 of the CS

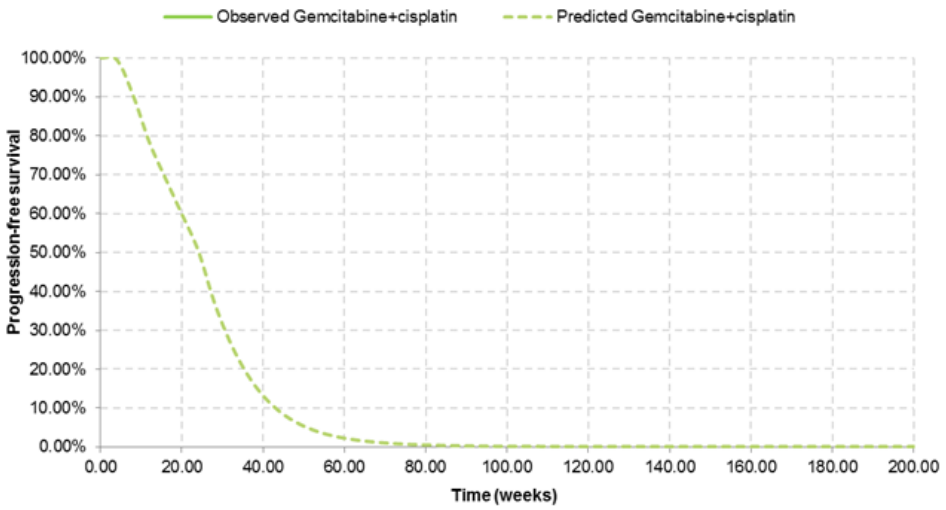
Figure 5.10-free survival and overall survival with best supportive care – observed and predicted values with the generalised gamma distribution[a]

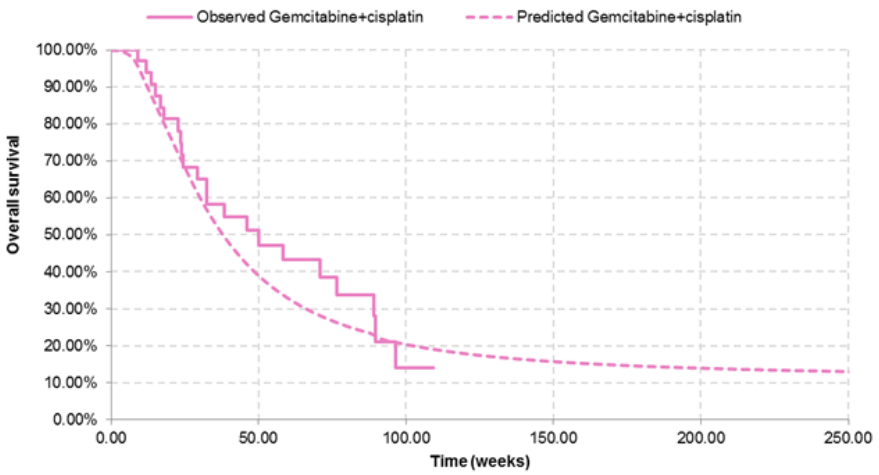




[a] No observed progression-free survival data were identified for best supportive care
Source: Figure 42 of the CS

Figure 5.11: Progression-free survival and overall survival with cisplatin plus gemcitabine – observed and predicted values with the generalised gamma distribution[a]





[a] Cisplatin plus gemcitabine treatment was analysed as a scenario analysis. No observed progression-free survival data were identified for cisplatin plus gemcitabine
Source: Figure 43 of the CS

Best supportive care (BSC) was not included in the STC for PFS due to a lack of relevant PFS data identified in the clinical SLR (Section 4.3). Therefore, the company assumed that the HR for BSC versus paclitaxel (1.47) was equivalent to that of BSC versus vinflunine for second-line UC patients (Bellmunt et al. (2009)[26]). The company assumed that this HR could be applied to the paclitaxel PFS curve to estimate the PFS of BSC, due to the similarities in terms of outcomes between vinflunine and paclitaxel/docetaxel. This HR was held constant during the time horizon of the cost effectiveness model, due to the absence of alternative data. No evidence was provided to support these assumptions.

Cisplatin plus gemcitabine was not included in the STC for PFS due to a lack of relevant PFS data identified in the clinical SLR (Section 4.4). The HR of paclitaxel versus nivolumab was applied to estimate the PFS of cisplatin plus gemcitabine because the company expected that paclitaxel and cisplatin plus gemcitabine would provide similar PFS results since they are all chemotherapy agents. No evidence was provided to support this assumption.[20]

ERG comment: The ERG's concerns include (1) the uncertainty and bias induced by comparing single-arm studies, (2) the discrepancy in populations in which relative effectiveness estimates are derived and applied, (3) the need for and effect of applying time-dependent HRs instead of time-independent HRs, (4) the estimation of HRs for PFS of BSC and cisplatin plus gemcitabine, and (5) the large impact of the parameter values used for the fractional polynomial NMA model.

(1) As described in Section 4.6, the STC and NMA performed by the company to obtain time-dependent HRs were associated with considerable uncertainty and the introduced bias associated with the STC was not quantified. For these reasons, the cost effectiveness analysis performed by the company suffers from significant uncertainty and potential bias. As stated in NICE DSU TSD 18 for STC's incorporating one-arm studies only, the accuracy of the resulting estimates is entirely unknown and without any evidence that the STC reduces the systematic error, the results *'are not worthy of consideration'*.[1]

(2) An additional concern is that the time-dependent HRs were obtained based on a comparison using the pooled CheckMate 032 and CheckMate 275 trials dataset, which did not take response status into account. Instead, the HRs for all patients (regardless of response status) were applied to the combined parametric time-to-event models, which accounted for response status. More specifically, the same time-dependent HRs were applied to the combined survival curves based on the weighted average of the responders and non-responders time-to-event models. Hence, there is a discrepancy between the a priori population on which the relative effectiveness is based on the a posteriori population in which the HRs are applied. The potential bias introduced by this methodology was not investigated by the company, despite a request in the clarification questions.[7] The ERG notes that applying HRs to the combined survival curves may underestimate the relative effectiveness in the responders group, but overestimate the relative effectiveness in the non-responders group. The ERG would have preferred to apply separate HRs to responders and non-responders, however, these were not provided by the company. This concern is redundant when using the conventional, not response-based, approach. The ERG further noticed that the code supplied to estimate the time-dependent HRs only estimated them up to a time horizon of 256 weeks, ending much before the end of the model time horizon. It is not clear where the time-dependent HRs implemented after 260 weeks were sourced from.

(3) The company applied time-dependent HRs to model the relative effectiveness of nivolumab versus the comparators because it assumed that the proportional hazard assumption did not hold. The company did not consult log-cumulative hazard plots to support this assumption, as recommended by the NICE DSU TSD 14[38]. Upon the ERG's request, the company provided the log-cumulative hazard plots of nivolumab versus the comparators. Based on these plots, the company confirmed that the proportional hazard assumption did not hold. The ERG considers that the proportionality of hazards could not be

ruled out based on the company's analyses because both CheckMate 032 and CheckMate 275 trials were presented separately in these plots, while the HRs were derived based on the pooled CheckMate 032 and CheckMate 275 trials dataset. Therefore, these plots did not allow investigation as to whether the proportional hazard assumption held for the analysis performed by the company. Because the company did not provide sufficient evidence to support the violation of the proportional hazard assumption and to support the need for time-dependent HRs, the ERG requested scenario analyses using time-independent HRs (i.e. fixed for the entire time horizon) to estimate the relative effectiveness of nivolumab versus the comparators. The company provided a network meta-analysis using fixed and random effects to estimate time-independent HRs in its response to the clarification letter.[7] These time-independent HRs were still in favour of nivolumab, except for cisplatin plus gemcitabine, which became more effective than nivolumab. The use of these time-independent HRs increased all cost effectiveness estimates (Section 5.2.10). The company did not consider these scenario analyses to be appropriate for decision making because a) the survival estimates for the comparator arm were considered to be implausible overestimations and b) the proportional hazard assumption was violated. The ERG considers that these claims were not strongly supported by the evidence submitted by the company. In response to a), the company only presented a single parametric time-to-event model to illustrate the overestimation of survival in the comparator arms but different parametric time-to-event models could lead to different results and a better fit with the data. In addition, the ERG notes that using time-independent HRs has the advantage of preventing over-parameterisation which might occur when estimating time-dependent HRs with the relatively little amount of data submitted by the company. In response to b), as stated above, the violation of the proportional hazard assumption was not demonstrated sufficiently by the company.

(4) Finally, the HRs used to estimate PFS of BSC and cisplatin plus gemcitabine were not obtained through the STC but were based on assumptions, which were not supported by clinical evidence (i.e. same HRs for BSC vs paclitaxel as for BSC vs vinflunine and same HRs for cisplatin plus gemcitabine versus nivolumab as for paclitaxel versus nivolumab).[2, 7] The assumption that PFS when treated with cisplatin plus gemcitabine is the same as when treated with paclitaxel is likely non-conservative. The ERG performed scenario analyses to investigate the influence of alternative time-dependent HRs for BSC and cisplatin plus gemcitabine PFS on the cost effectiveness results. In these scenario analyses, the time-dependent HRs obtained for OS of BSC and cisplatin plus gemcitabine were used. These time-dependent HRs were selected because they were based on evidence concerning the drug of interest instead of being based on assumptions lacking supporting evidence. However, the ERG is aware that the relative effectiveness of a treatment compared to another may change across different outcomes.

(5) The use of the fractional polynomial model introduces some uncertainty into the cost effectiveness analysis. The company showed the effects of a set of alternative p1 and p2 values on the ICERs, showing that their base-case ICERs increased significantly. In response to clarification questions the company enabled in the model 10 different p1 an p2 values, resulting in 100 possible combinations. It is the ERG's concern that these different combinations could have an unpredictable effect on model outcomes and the ERG therefore explored the range of ICERs that could be obtained through a 'mini-PSA', in which 10,000 draws from different combinations of these parameter values are used in the model. Whilst implementing this, the ERG noted that certain combinations of parameter values result in extreme hazard ratios and survival estimates above 100%, showing that not all of these are plausible candidates. The ERG adjusted survival estimates to prevent this problem from occurring in their mini-PSA.

5.2.6.3 Time to treatment discontinuation

Treatment with nivolumab should continue *'as long as clinical benefit is observed or treatment is no longer tolerated by the patient.'*[2] Time-to-treatment discontinuation (TTD) was estimated through a parametric time-to-event model. The same (six) distributions as for OS and PFS were fitted to the pooled CheckMate 032 and CheckMate 275 studies' dataset and statistical fit of the different curves was assessed through the AIC and BIC (Table 5.4). In the CS, TTD was estimated independent of response status.

The generalised gamma distribution was selected to estimate TTD in the base-case analysis, with the company claiming that this was done to ensure consistency with the curves selected to represent OS and PFS. The Gompertz and log-logistic distributions showed better statistical fit than the generalised gamma distribution but the company argued that these two distributions produced long tails with patients still being on treatment after 5 and 10 years, which lacked clinical validity (Table 5.5). The impact of using alternative distributions to estimate TTD was explored in sensitivity analyses.

Table 5.4: TTD estimation based on different parametric time-to-event models

Time	TTD estimation		
	Generalised gamma	Gompertz	Log-logistic
1 year	17.6%	21.4%	22.1%
2 year	8.3%	16.7%	12.7%
3 year	5.1%	15.9%	8.9%
4 year	3.2%	15.8%	6.9%
5 year	2.1%	15.8%	6.0%
10 year	0.2%	15.8%	2.8%

Source: company's cost effectiveness model
[a] Parametric time-to-event model used in the company's base-case analysis

TTD of the comparators was based on their respective PFS curves because it was assumed that comparator treatment would continue until disease progression or unacceptable toxicity. Treatment with paclitaxel was assumed to stop after 6 (model) cycles (if treatment was not discontinued yet), i.e. 24 weeks. This represented the clinical use of paclitaxel in the UK[15] and was confirmed by clinical experts.[5] The company assumed that BSC was administered until death.

Table 5.5: Statistical fit measures of the distributions representing time to treatment discontinuation

Endpoint	Distribution	AIC	BIC
Time to treatment discontinuation	Exponential	2381.86	2385.71
	Weibull	2329.96	2337.67
	Gompertz	**2318.29**	**2325.99**
	Lognormal	2341.69	2349.40
	Log-logistic	2322.93	2330.63
	Generalised gamma	2328.48	2340.04

Source: Table 30 of the CS[2]
Abbreviations: AIC, Akaike Information Criterion; BIC, Bayesian Information Criterion
Bold printed values represent the distributions with the lowest AIC or BIC (i.e. the 'best fitting' time-to-event models)

ERG comment: The ERG's concerns include (1) inconsistency in estimating TTD compared with estimating OS and PFS (the use of a conventional, not response-based, approach to estimate TTD), and (2) the choice of parametric distributions for TTD.

(1) Unlike OS and PFS, the parametric time-to-event models estimating TTD were not estimated based on a landmark and response-based analysis but on the pooled CheckMate 032 and CheckMate 275 trials dataset. This was inconsistent with the analysis of OS and PFS and no justification was provided. The ERG requested from the company to implement a response-based, landmark, analysis for TTD, assuming that treatment duration may be influenced by response status, especially given that treatment with nivolumab should continue *'as long as clinical benefit is observed or treatment is no longer tolerated by the patient.'*[2] The company provided an updated cost effectiveness model in which TTD can be estimated in the same way as OS and PFS, i.e. using a response-based analysis. However, the ERG noticed that the company calculated the proportion of responders and non-responders based on the sum of patients in the OS and PFS health states, thereby double-counting patients. The ERG considered it more appropriate to use all responders alive for the calculation of proportion of responders.

(2) The company justified the use of the generalised gamma distribution by the lack of clinical plausibility of the alternative parametric time-to-event models (e.g. Gompertz and log-logistic distributions). This argument was not supported by clinical expert opinion, and the ERG considers there to be uncertainty about the likely treatment duration. Within the response-based analysis provided in response to the clarification questions,[7] the company explored the influence of using Gompertz or log-logistic distributions for both responders and non-responders. Both scenario analyses increased the ICERs (Section 5.2.10). However, the company considered that the proportion of patients who were still receiving treatment after five years or more was not representative of clinical practice in both scenarios (Table 5.6).

In conclusion, the ERG adopted a conventional, non-response based approach in the base-case, using the generalised gamma distribution for estimating TTD, in line with the CS. The ERG furthermore explored the influence of using a response-based and landmark analysis for OS, PFS and TTD in a scenario analysis. In this scenario analysis, the generalised gamma was used to estimate TTD of the responders and non-responders, and in a second analysis, the Gompertz and log-logistic distributions were used for responders and non-responders, respectively.

Table 5.6: TTD estimation based on different parametric time-to-event models (landmark and response-based analysis)

Time	TTD estimation				
	Generalised gamma[a]	Generalised gamma[b]	Gompertz	Log-logistic	Best fitting parametric time-to-event models[c]
1 year	17.6%	19.6%	20.1%	20.7%	20.0%
2 year	8.3%	11.4%	13.2%	11.8%	13.2%
3 year	5.1%	8.4%	10.1%	8.0%	10.6%
4 year	3.2%	7.0%	8.4%	5.9%	9.1%
5 year	2.1%	6.1%	7.3%	4.6%	8.2%
10 year	0.2%	4.1%	2.2%	2.1%	3.1%

Source: updated cost effectiveness model submitted with the response to the clarification letter
[a] Used in the company base-case
[b] Estimation based on the landmark and response-based analysis

[c] Based on the landmark and response-based analysis, the log-normal and the Gompertz distributions were the best fitting parametric time-to-event models for the responders and non-responders, respectively.

5.2.7 Adverse events

Table 5.7 presents the adverse events that were included in the cost effectiveness model. Grade 3-4 adverse events were incorporated in the model if their incidence was ≥5%. The impact of adverse events on quality of life and costs were incorporated in the first cycle of the model (see sections 5.2.8 and 5.2.9 for more details).

Table 5.7: Adverse event rates incorporated in the cost effectiveness model

Adverse event	Nivolumab	Docetaxel	Paclitaxel	BSC	Cisplatin plus gemcitabine[a]
Neutropenia	1.00%	14.00%	6.00%	0.90%	66.67%
Anaemia	1.48%	1.00%	0.00%	8.10%	42.42%
Thrombocytopenia	NR	NR	0.00%	0.90%	33.33%
Asthenia	1.48%	6.00%	5.00%	17.90%	0.00%
Nausea/vomiting	0.37%	NR	0.00%	0.90%	0.00%
Diarrhoea	1.85%	0.00%	2.00%	NR	NR
ALT increase	0.74%	0.00%	2.00%	NR	NR
Leukopenia	0.00%	0.00%	0.00%	NR	45.45%
Source	Checkmate 275[10]	Choueiri *et al.* (2012)[27]	Jones *et al.* (2017)[15]	Bellmunt *et al.* (2009)[26];Bellmunt *et al.* (2013)[41]	Gondo *et al.* (2011)[13]

Source: adapted Table 31 of the CS[2]

[a] The trials informing the comparison versus cisplatin plus gemcitabine were conducted in gemcitabine-naïve populations, and therefore cannot be considered to provide relevant data for the retreatment with first-line platinum-based chemotherapy. This comparison has been briefly included in Appendix O as a scenario analysis only and results should be interpreted with caution.

Abbreviations: ALT, alanine transaminase; BSC, best supportive care; NR, not reported

ERG comment: The ERG's concerns relate to (1) the selection of sources for AEs associated with nivolumab, (2) selection of sources for AEs associated with the comparators, (3) the inclusion of both neutropenia and leukopenia, and (4) an inconsistency between the inclusion criteria for AEs and the actually included AEs.

(1) For the nivolumab arm, the CheckMate 275 trial was the only source informing the adverse event rates in the cost effectiveness model while the clinical effectiveness of nivolumab was estimated based on both CheckMate 032 and CheckMate 275 studies. The company justified this choice in its response to the clarification letter by stating that it simplified the analysis and that adverse events did not have a meaningful impact on the results[7]. Hence, the use of both trials instead of CheckMate 275 only for the estimation of adverse event rates would not affect the conclusions of the analysis. The company did not provide evidence to support this argument.

(2) Another issue with the adverse events are that the company did not justify the selection of the source used to estimate AE rates of the comparator. In the response to the clarification letter, the company

explained that these sources were selected to ensure consistency by using the same sources as for the relative effectiveness estimation of nivolumab versus the comparators. The company did not argue why these sources were the most appropriate.

(3) Both neutropenia and leukopenia were incorporated in the cost effectiveness model. The ERG was unsure whether this was appropriate, given that neutropenia is a subtype of leukopenia. However, this is not likely to have a significant impact on model outcomes.

(4) Finally, AEs were included in the cost effectiveness model when their incidence was ≥5%. However, nausea/vomiting, diarrhoea, and ALT increase have an incidence <5% for all treatments included in the cost effectiveness model. Hence it is inconsistent to include these AEs in the cost effectiveness model. The ERG removed these adverse events from its analyses.

5.2.8 Health-related quality of life

Within the economic SLR, six records of four unique studies were identified that included HRQoL in locally advanced or metastatic UC.[34, 36, 42-45] None of these studies were consistent with the NICE reference case and therefore data to inform utilities of the economic evaluation were taken from the CheckMate 275 trial where the EQ-5D-3L was used and valued with UK preference weights.

5.2.8.1 EQ-5D-3L data from CheckMate 275 trial

In absence of alternative data that was consistent with the NICE reference case, the utilities derived from the CheckMate 275 study were deemed most appropriate for this appraisal. Utility estimates derived from the CheckMate 275 study were stratified according to progression-free and post progression health states. Data were available at baseline for 261/270 (96%) patients. During follow-up, the completion-rate declined but remained above 70% at 49 weeks (Table 5.8).

Table 5.8: EQ-5D-3L questionnaire completion rates over time (total enrolled population)

	EQ-5D-3La	
Assessment	**n/N**	**%**
Week 1 (baseline)	261/270	96.7
Week 9	144/167	86.2
Week 17	97/116	83.6
Week 25	75/91	82.4
Week 33	54/70	77.1
Week 41	24/32	75.0
Week 49	6/7	85.7

[a] Completion rates = patients who completed the PRO with ≥1 score at the assessment time point/expected population (total population minus patients who have died or dropped out)
Abbreviations: EQ-5D-3L: EuroQoL 5-Dimensions 3-Levels; PRO: patient reported outcomes.
Source: Table 32 of the CS[2]

In total 794/1,465 (54%) observations were missing. After interpolation of observations made for measurement times deviating from the measurement schedule, 788/1,465 (54%) of observations were available. The remaining missing observations were partly (204/1,465 = 14%) due to the immaturity of the dataset, i.e. patients had not reached all follow-up measurements yet. The company acknowledged that discontinued treatment, progressive status and female gender seemed to be predictors of missing observations, and thus data might not have been missing completely at random. All missing observations were imputed using multiple imputation by chained equations and predictive mean matching, where the number of imputations was set to 40.

The company used a mixed-effects model to reflect within subject variance. This resulted in health state utilities of 0.718 and 0.604 pre-progression and post-progression respectively (Table 5.9). It is noteworthy that imputed pre-progression utilities were similar to observed utilities, but imputed post-progression utilities were lower than the observed utilities. The company furthermore explored the effect of time on progression effect. The pattern seen in post-progression utilities however was deemed different from what was seen in clinical practice, and the company therefore used one set of time-independent utilities.

Table 5.9: Summary of utility values for cost effectiveness analysis

State	Utility/disutility value: mean (standard error)	95% CI	Source
Pre-progression	Imputed value: 0.718 (0.016) Observed value: 0.713 (0.017)	Imputed value: 0.686 to 0.75 Observed value: 0.679 to 0.747	Imputed from Checkmate 275
Change in utility – pre-progression to post-progression	Imputed value: -0.115 (0.0291) Observed value: -0.061 (0.0167)	Imputed value: -0.143 to -0.087 Observed value: -0.123 to -0.055	Imputed from Checkmate 275
Post-progression	Imputed value 0.603 (N/A) Observed value: 0.623 (N/A)	N/A	Checkmate 275

Abbreviations: ALT: alanine transaminase; CI: confidence interval; N/A: not applicable; NR: not reported.
Source: Table 35 of the CS [2]

Adverse event disutilities

The company applied disutilities to several AEs (see Table 5.10); these were based on studies reporting utilities in patients with non-small cell lung cancer, pancreatic cancer and leukaemia. Disutilities were not treatment-specific and were applied as one-off events at the beginning of treatment, based on the proportion of patients experiencing the adverse event and the duration of the adverse event.

Table 5.10: Disutilities used in comparison to previous nivolumab appraisal ID971

Adverse event	Disutility ID995	Source	Disutility ID971	Source
Neutropenia	-0.18	Attard et al. (2014)[46]	-0.09	Nafees (2008)[47]
Anaemia	-0.09	Beusterien et al. (2010)[48]	-0.07	Nafees (2008)[47]
Thrombocytopenia	-0.18	Attard et al. (2014)[46]		
Asthenia/Fatigue	-0.12	Attard et al. (2014)[46]	-0.07	Nafees (2008)[47]
Nausea/vomiting	-0.05	Nafees et al. (2008)[47]	-0.05	Nafees (2008)[47]
Diarrhoea	-0.29	Attard et al. (2014)[46]		
ALT increase	-0.05	NICE TA347 (2015)[49]		
Leukopenia	-0.09	Frederix et al. (2013)[50]		

Sources: Table 35 of the CS [2], previous nivolumab appraisal ID971 [51]

ERG comment: The ERG identified several inconsistencies and choices lacking justification in the handling of utility values. The main issues include (1) inconsistencies in reported observations, (2) the use of utilities derived only from CheckMate 275, (3) the imputation of immature data, (4) the use of multiple imputation instead of the mixed model to adjust for missing data, (5) lack of justification for not using time-dependent utilities, and (6) disutilities for adverse events were inconsistent with those used for a previous nivolumab appraisal.

(1) The ERG noted a small inconsistency in the reported number of observations. As they were reported in the response to the clarification letter, the number of interpolated observations (117), imputed observations (677) and valid observations (661) do not add up to the total of observations (1465), but deviated by 10 observations.[7]

(2) The exclusion of utilities of the CheckMate 032 trial, which was in accordance with the reference case, is inconsistent with the pooling of other outcomes from CheckMate 275 and CheckMate 032 trials. In response to clarification question B16.C, the company reported utilities pooled from both CheckMate 032 and CheckMate 275 trials.[7] In this analysis pre- and post-progression utilities were higher compared to the utilities used by the company, and this resulted in a decrease in the ICERs for all nivolumab comparisons.[7]

(3) The ERG considers the company's decision to impute immature data as unjustified, and is concerned that it works with the unlikely assumption that none of the immature observations will be censored due to death of patients. The ERG wants to stress that the appropriateness of imputation as a substitution for follow-up is highly questionable. The impact of this on utility values is unclear, especially given that the company did not explore the assumptions made and the uncertainty surrounding the immaturely imputed utilities.

(4) The ERG considers the approach to adjust for missing data not sufficiently justified. The company could have used the mixed model, employed to calculate health state utilities, to adjust for missing observations, but instead used multiple imputation. In response to clarification question B16.B, the company presented utilities using only a mixed model.[7] These closely resembled the utilities produced using multiple imputation and led to only a small difference in ICERs. The ERG was satisfied that the use of multiple imputation to adjust for missing data did not have a large impact on model outcomes.

(5) Unfortunately, the company did not respond to the ERG request for an explanation how it was determined that time-dependent utilities were '… *seen to increase and decrease in a manner that would not be expected in clinical practice*' and were not used in the economic evaluation[2] (clarification question B16.G[7]). However, the company additionally added a variable of on- and off-treatment into the mixed model. The utilities presented were thus for four health states: pre- and post-progression, before and after treatment discontinuation, respectively. In this scenario, the disutility of treatment discontinuation was larger than the disutility of progression (Table 5.11), which was in line with the expectation of the ERG. This analysis raises the question whether on- and off-treatment are better predictors of utility values than pre- or post-progression. However, for consistency with other TAs and because progression is commonly accepted to be a predictor for health state values, the ERG maintained the company's pre- and post-progression utility values.

Table 5.11: Final utility values with linear mixed model including treatment discontinuation as a variable

	Pre-progression	Post-progression
On treatment	0.723	0.666
Off treatment	0.650	0.573
Source: Table 29 of the company's response to request for clarification from the ERG [7]		

(6) AE disutilities used were inconsistent with those used in ID971.[51] The disutilities used in the CS stemmed from multinational trials on various cancers, were not evaluated in UK UC patients and are larger than in ID971.[51] Given the prevalence of AEs, it can be expected that the disutilities used favour the cost effectiveness of nivolumab. This is explored in the ERG's sensitivity analysis. It is of note that leukopenia was not associated with a utility decrement or cost. The company did not apply a cost to leukopenia because of the overlap of leukopenia with neutropenia and because no cost was applied in ID971[2]. For consistency, a utility decrement for leukopenia should therefore also not be applied. However, this inconsistency is not influential.

In conclusion, the ERG adopted the pooled utility estimates in its base-case and explored alternative AE disutilities in an exploratory analysis.

5.2.9 Resources and costs

Resource use and unit costs data to inform the economic model were based on a number of sources, including:

- CheckMate 275;
- national databases;
- published sources (both sources identified and not identified in the SLR described in Section 5.1 of this report) and;
- clinical advice.

Additionally, assumptions were necessary in the absence of evidence. These assumptions were validated through discussions with clinicians.

Drug, administration and monitoring costs

The British National Formulary (BNF) was used to obtain unit prices for nivolumab (40mg and 100mg). A PAS, [redacted], was incorporated in the model. The unit prices for docetaxel, paclitaxel and gemcitabine plus cisplatin were taken from the electronic market information tool (EMIT).

The dose/number of vials required per administration were estimated based on the dosage scheme and the dose intensity (reflecting missed doses). For this calculation an average weight of 77.3 kg (SD 16.34) and Body surface area (BSA) of 1.90 m^2 (SD 0.205) were assumed (both based on the CheckMate 275 trial). Using a normal distribution the proportions of patients in different weight and BSA categories were calculated (see CS Tables 36 and 37[2]). Additionally, the calculation of the dose intensity (93.4%) was based on data from the CheckMate 275 and CheckMate 032 trials and based on the assumption that all delayed doses represent missed doses. In absence of evidence, the company assumed that the dose intensity for docetaxel, paclitaxel, gemcitabine plus cisplatin was equal to that of nivolumab.

The average drug costs per patient per four weeks were calculated by combining the drug unit prices, the vials required per administration, the dose intensity and the number of administrations per four weeks.

In addition to the drug costs, administration costs of £198.94 per dose were incorporated (derived from NHS reference costs 2015-16). These costs were incorporated independent of the dose intensity as it was assumed that for missed doses, the chair time would still have been reserved for the patient. The total drug and administration costs per 4 weeks ranged between £304 for docetaxel and [redacted] for nivolumab (see Table 5.12).

Monitoring costs (while on treatment) included in the model (Table 5.13) consisted of regular follow-up visits with an oncologist, CT scans and various blood tests (full blood count, hepatic function test, renal function test, thyroid function test, pituitary function test). The resource use was based on expert opinion (i.e. advisory board feedback) while the unit prices were based on NHS reference costs 2015-16. The total monitoring costs per four weeks ranged between £272 for docetaxel and £556 for gemcitabine plus cisplatin (see Table 5.13).

Table 5.12: Drug and administration costs for nivolumab (with PAS) and comparators

	Per vial		Per dose			Per 4 weeks			
	Vial size (mg)	Costs per vial	Dosage scheme	Dose intensity	Average dose[a]	Number of administrations	Drug costs	Administration costs	Total costs
Nivolumab	40	[redacted]	3 mg/kg	93.4%	260.27	2.00	[redacted]	£397.88	[redacted]
	100	[redacted]							
Docetaxel	80	£12.47	75 mg/m2	93.4%	185.02	1.33	£38.45	£265.25	£303.71
Paclitaxel	100	£8.50	80 mg/m2	93.4%	200.17	3.00	£51.04	£596.82	£647.86
Gemcitabine	1000	£178.56	1000 mg/m2	93.4%	2312.79	3.00	£1,238.92	£596.82	£2,057.66[b]
Cisplatin	50	£6.99	70 mg/m2	93.4%	164.36	1.00	£22.98	£198.94	

[a]This includes wastage (as no vial sharing is assumed) and dose intensity (reflecting missed doses)

[b]Total costs of cisplatin + gemcitabine

Table 5.13: Monitoring costs

	Oncologist follow-up visit per 4 weeks		CT scans per 4 weeks		Various blood tests[a] per 4 weeks		Total per 4 weeks
	Frequency	Costs	Frequency	Costs	Frequency	Costs	Costs
Nivolumab	2.00	£326.00	0.50	£57.50	10.00	£10.00	£393.50
Docetaxel	1.33	£217.33	0.44	£51.11	4.00	£4.00	£272.44
Paclitaxel	3.00	£489.00	0.44	£51.11	9.00	£9.00	£549.11
Gemcitabine plus cisplatin	3.00	£489.00	0.50	£57.50	9.00	£9.00	£555.50

[a]Full blood count, hepatic function test, renal function test, thyroid function test, pituitary function test (all costing £1)

Best supportive care costs

For the BSC comparator, BSC costs were administered until death. For the remaining comparators, BSC costs were incorporated after treatment discontinuation (i.e. discontinuation of nivolumab, docetaxel, paclitaxel or cisplatin + gemcitabine) until death.

BSC costs included GP home visits, community nurse specialist visits and blood transfusions as well as drug costs for prednisolone, morphine, gabapentin and alendronic acid. The total BSC costs per 4 weeks amounted to £170.21 (see CS Table 39[2]).

Adverse event costs

Although not described in the CS, treatment dependent AE costs were incorporated as one-off event costs for patients on treatment during the first cycle of the model based on the occurrence (See Table 5.7) and costs (CS Table 41[2]) of AE. The sum of these costs is provided per treatment in Table 5.14.

Table 5.14: Total AE event costs

	Total AE event costs	Total AE event costs (alternative costs per AE event)[a]	Difference
Nivolumab	£147.24	£115.33	-£31.91
Docetaxel	£773.55	£284.67	-£488.88
Paclitaxel	£408.62	£205.91	-£202.71
Cisplatin + gemcitabine	£5,389.57	£2,477.92	-£2,911.65
BSC	£819.63	£847.52	£27.89

Source: economic model submitted by the company

[a]Alternative costs per AE were retrieved from ID971[51] (nivolumab for treating recurrent or metastatic squamous-cell carcinoma of the head and neck after platinum-based chemotherapy). Moreover, the costs for leukopenia were set to £0 given 1) the overlap with neutropenia; 2) AE occurrence was missing for all comparators except one and; 3) given that this is consistent with ID971 as in this assessment no costs for leukopenia were considered. Finally, the fatigue AE costs from ID971 were assumed to be applicable for asthenia.

Subsequent treatment costs

Following discontinuation of nivolumab, docetaxel, paclitaxel or cisplatin plus gemcitabine, a proportion of patients received subsequent radiotherapy and/or surgery (9.3% and 3.3% respectively based on CheckMate 275). The unit prices were based on NHS reference costs 2015-16 and amounted to £128.22 and £3,201.68 for radiotherapy and surgery respectively. The costs were incorporated as one-off event costs after treatment discontinuation.

Terminal care costs

Terminal care costs were incorporated in the model as event costs of £6,152.64 related to the transition to death. These costs were an average of the acute care and community costs for cancer patients in their last eight weeks of life.[52]

ERG comment: The ERG identified several technical errors, inconsistencies and assumptions that lacked justification. These included a technical error (1) in calculating the dose intensity; inconsistencies, namely (2) using the average weight and BSA from CheckMate 275 (not using CheckMate 032), (3) using the subsequent treatment proportions from CheckMate 275 (not using CheckMate 032), (4) not using cost and resource use data from TA272 (identified in the SLR), and (5) using different AE unit costs compared with ID971[51] (nivolumab for treating recurrent or metastatic squamous-cell carcinoma of the head and neck after platinum-based chemotherapy); as well as three

assumptions, namely (6) assuming an administration scheme that is inconsistent with UK clinical practice for cisplatin + gemcitabine, (7) Assuming that all delayed doses are missed doses for calculating nivolumab dose intensity, and (8) assuming that the dose intensity for the comparators is equal to that of nivolumab.

(1) The identified technical errors entailed the incorporation of dose intensity in the economic model. Drug dose intensity was incorporated in the calculation of the total dose required per weight category, which was subsequently used to calculate the number of vials per weight category. This is incorrect as the dose intensity is related to the number of missed doses and not to the number of vials per weight category. Hence the dose intensity should be applied after calculating the number of vials per weight category. This is corrected in the ERG base-case.

(2) The company assumed a weight and BSA of 77.3 kg and 1.90 m^2 respectively to calculate the dose/number of vials per administration. This was based on CheckMate 275 only. This is inconsistent given that the company combined data from the CheckMate 275 and CheckMate 032 in the majority of their analyses, presumably assuming that the combined population is most relevant for the decision problem being considered. Although the ERG requested clarification on this inconsistency (clarification question B17.A[7]), no further details were provided. Moreover, given that the mean weight was 83.51 kg in CheckMate 032 (mean BSA was not provided in the CSR[11]), this inconsistency resulted in an underestimation of the nivolumab drug costs. Hence, an average weight of 80.405 kg based on both CheckMate 275 and CheckMate 032 was used in the ERG analyses. Given that mean BSA from CheckMate 032 was not provided, the mean BSA of 1.90 m^2 from CheckMate 275 was retained. Moreover, this seems appropriate given that in TA272[34] a similar BSA (of 1.85 m^2) was used (as stated by the company in response to clarification question B17.D[7]).

(3) Similar to the previous inconsistency, the proportions of patients receiving subsequent radiotherapy and/or surgery (9.3% and 3.3% respectively), following discontinuation of nivolumab, docetaxel, paclitaxel or cisplatin + gemcitabine, was retrieved from CheckMate 275 only. These proportions were 11.5% and 6.4% in CheckMate 032. For consistency, average proportions based on both CheckMate 275 and CheckMate 032 were used in the ERG analyses (10.40% and 4.85% for patients receiving subsequent radiotherapy and/or surgery respectively).

(4) The company identified TA272 (the only other NICE submission in this indication) in its SLR. This source was nevertheless not used to inform costs and resource use.[34] The company stated (response to clarification question B18.A[7]) that NHS reference costs for 2007/2008 (from TA272) would be inappropriate to use in 2017. This argument is inconsistent with other costs used by the company (e.g. the leukopenia cost estimate was derived from a paper published in 2004[35]). However, considering the response to clarification question B18[7], it seems reasonable not to use the monitoring and BSC costs from TA272. In response to this clarification question the company states that treatment-related monitoring costs in TA272 did not include oncologist visits and CT scans and were dependent on progression status (instead of treatment status as preferred by the company). Regarding BSC costs, the company stated in TA272 these costs included hospice costs while the company prefers to incorporate these costs as part of the terminal care costs.[34]

(5) The AE unit costs are reported in CS Table 41[2]. These AE unit costs however differ from previous nivolumab assessments (e.g. ID971[51]) and no justification is provided for the sources used to obtain the AE unit costs. This is of particular concern for the AE unit costs for neutropenia and nausea and vomiting as these were based on NHS reference costs for paediatrics. To illustrate the impact of the inconsistency with ID971[51] (nivolumab for recurrent or metastatic head and neck cancer), the ERG calculated alternative AE costs based on ID971[51] (Table 5.14; see footnote for calculation details).

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) × 36.6% (the proportion of dose delays that exceeded 14 days; averaged for CheckMate 275 and CheckMate 032) = 2.4% (i.e. dose intensity of 97.6%).

(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 docetaxel, paclitaxel and BSC respectively (Table 5.17).

Table 5.15: Summary of quality-adjusted life year gains by health state

	Nivolumab		Docetaxel			Paclitaxel			Cis+ gem			BSC		
	QALYs	LYG	QALYs	LYG	Incremental QALYs vs. Nivolumab	QALYs	LYG	Incremental QALYs vs. Nivolumab	QALYs	LYG	Incremental QALYs vs. Nivolumab	QALYs	LYG	Incremental QALYs vs. Nivolumab
Health state														
Pre-progression	[redacted]	1.06	[redacted]	0.75	[redacted]	[redacted]	0.47	[redacted]	[redacted]	0.47	[redacted]	[redacted]	0.32	[redacted]
Post-progression	[redacted]	1.72	[redacted]	0.65	[redacted]	[redacted]	0.71	[redacted]	[redacted]	1.99	[redacted]	[redacted]	0.70	[redacted]
Adverse events	[redacted]		[redacted]		[redacted]	[redacted]		[redacted]	[redacted]		[redacted]	[redacted]		[redacted]
Total	[redacted]	2.78	[redacted]	1.40	[redacted]	[redacted]	1.19	[redacted]	[redacted]	2.47	[redacted]	[redacted]	1.01	[redacted]

Abbreviations: QALY: quality-adjusted life year; LYG: life years gained; Cis+gem: cisplatin plus gemcitabine; BSC: best supportive care.
Source: Table 67 of the CS Appendix J [20]

Table 5.16: Summary of costs by health state

	Nivolumab	Docetaxel		Paclitaxel		Cis+ gem		BSC	
	Costs	Costs	Incremental costs vs. Nivolumab	Costs	Incremental costs vs. Nivolumab	Costs	Incremental costs vs. Nivolumab	Costs	Incremental costs vs. Nivolumab
Treatment	[redacted]	£3,113	[redacted]	£3,515	[redacted]	£12,381	[redacted]	£2,310	[redacted]
Monitoring	[redacted]	£2,716	[redacted]	£2,734	[redacted]	£3,455	[redacted]	£0	[redacted]
Post-progression	[redacted]	£1,521	[redacted]	£1,864	[redacted]	£4,492	[redacted]	£0	[redacted]
Adverse events	[redacted]	£739	[redacted]	£411	[redacted]	£5,378	[redacted]	£806	[redacted]
Terminal care	[redacted]	£5,857	[redacted]	£5,902	[redacted]	£5,630	[redacted]	£5,940	[redacted]
Total	[redacted]	£13,945	[redacted]	£14,426	[redacted]	£31,337	[redacted]	£9,056	[redacted]

Abbreviations: Cis+gem: cisplatin plus gemcitabine; BSC: best supportive care
Source: Table 68 of the CS Appendix J [20]

Table 5.17: Base-case results – with PAS

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER of nivolumab vs each comparator (£/QALY)
Nivolumab	[redacted]	2.78	[redacted]				
Paclitaxel	£14,426	1.19	0.76	[redacted]	1.60	[redacted]	£37,647
Docetaxel	£13,945	1.40	0.92	[redacted]	1.38	[redacted]	£44,960
BSC	£9,056	1.01	0.64	[redacted]	1.77	[redacted]	£38,164
Cis+gem	£31,337	2.47	1.49	[redacted]	0.31	[redacted]	£71,608

Abbreviations: BSC: best supportive care; Cis+gem: cisplatin plus gemcitabine; ICER: incremental cost-effectiveness ratio; LYG: life years gained; QALYs: quality-adjusted life years.
Source: Table 44 of the CS [2]

ERG comment: The ERG comments relate to (1) the exclusion of cisplatin plus gemcitabine from the base-case, and (2) the driving factor of incremental QALYs being the extended post-progression survival.

(1) Cost effectiveness results were not presented for nivolumab compared with cisplatin plus gemcitabine within the company's base-case. This is not in line with the scope. The ERG requested this analysis in the clarification letter but the company continued to exclude this analysis from the base-case, arguing in their response to question B13.A[7], that '*... it is not considered a relevant comparator in the context of second-line UK clinical practice*'[7]. The ERG disagrees with this statement, especially given that this comparator was named in the scope. More detail on this is presented in Section 4.

(2) In a previous nivolumab appraisal ID971,[51] it has been discussed that incremental QALYs were mainly driven by extended survival post-progression and after treatment discontinuation. Such a pronounced effect of nivolumab after progression or treatment discontinuation had not been seen in clinical practice, [51] thus the extrapolation in the model has been criticised in previous committee appraisals. The ERG wishes to flag up that in the company's base-case the issue of the QALY gain coming almost entirely from the post-progression health state was less pronounced but still accounted for over 50% of incremental gains for all comparators in the company's base-case.

5.2.11 Sensitivity and scenario analyses

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 (further information in Table 46 of the CS[2]).

Results of the PSA using 1,000 iterations are shown in Table 5.18. Incremental costs increased and incremental QALYs decreased compared to the deterministic results, resulting in ICERs of £46,209 and £44,698 per QALY gained for nivolumab versus paclitaxel and BSC, and an ICER of £54,220 per QALY gained for nivolumab versus docetaxel. 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). As PFS and OS are greater in nivolumab than in the comparators, the effect on nivolumab was more pronounced than on the comparators. Probability of cost effectiveness at a threshold of £50,000 per QALY gained was 72.1% versus paclitaxel, 49.0% versus docetaxel, 76.3% versus BSC and 6.9% versus gemcitabine plus cisplatin.

Table 5.18: Probabilistic CS results

Technologies	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)	Probability of cost effectiveness[a]
Paclitaxel	[redacted]	[redacted]	£46,209	72.10%
Docetaxel	[redacted]	[redacted]	£54,220	49.00%
BSC	[redacted]	[redacted]	£44,698	76.30%
Cis+gem	[redacted]	[redacted]	£103,568	6.9%

[a]The probability of nivolumab being cost-effective versus the stated comparator at a cost-effectiveness threshold of £50,000/QALY.
Abbreviations: Cis+gem: cisplatin plus gemcitabine; BSC: best supportive care, ICER: incremental cost-effectiveness ratio; QALYs: quality-adjusted life years.
Sources: Table 47 of the CS [2], Table 79 of the CS Appendix O [20]

The company stated that individual one-way DSAs were conducted including all parameters other than survival curves. The parameters were varied within their respective 95% CI or, if not applicable, within a ± 50% range of the deterministic base-case value. The DSA results including the PAS were presented using tornado diagrams with the 10 key model drivers (CS Figures 46-48 [2]). Ranked by importance, the following parameters were identified as most influential on the cost effectiveness of nivolumab versus paclitaxel, docetaxel and BSC:

1. Mean age (65; 47-84)
2. Cost per 100mg Nivolumab (£1,097; £548.50-£1,645.50)
3. Mean weight (77,3; 45-100)
4. Nivolumab dose intensity (93%; 47%-100%)

The company performed six deterministic scenario analyses, which are presented in Table 5.19. In summary, the 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 were major drivers of model results, mostly resulting in less favourable cost effectiveness estimates for nivolumab versus its comparators (see Table 5.19).

Table 5.19: Deterministic scenario analyses

Scenario			**ICER vs. Paclitaxel**	**ICER vs. Docetaxel**	**ICER vs. BSC**
Base case		**Gen. gamma**	**£37,647**	**£44,960**	**£38,164**
1 Survival curves	Landmark week 8	Weibull	£101,994	£114,823	£91,372
		Gompertz	£49,010	£59,858	£50,201
		Lognormal	£52,900	£72,044	£53,634
		Log-logistic	£58,279	£78,063	£59,695
		Exponential	£57,998	£70,582	£59,564
	Landmark week 26	Gen. Gamma	£34,541	£40,246	£34,774
		Weibull	£50,060	£62,866	£51,378
		Gompertz	£35,655	£41,933	£35,269
		Lognormal	£38,834	£48,610	£38,192
		Log-logistic	£42,475	£54,235	£43,097
		Exponential	£60,279	£76,786	£61,389
2 Fractional polynomial model[a]		p1=1, p2=1	£56,073	£59,504	£43,554

Scenario		ICER vs. Paclitaxel	ICER vs. Docetaxel	ICER vs. BSC
3 Exponential piecewise model	Piecewise exponential at 8 weeks	£53,616	£65,450	£55,597
	Piecewise exponential at 26 weeks	£55,681	£71,147	£57,293
4 Vial sharing		£35,651	£42,630	£36,333
5 Stopping rule[b]		£31,561	£37,781	£32,743
6 Alternative TTD parametric curves	Weibull	£33,562	£40,141	£34,525
	Gompertz	£183,467	£216,984	£168,053
	Lognormal	£61,810	£73,465	£59,688
	Log-logistic	£61,994	£73,683	£59,851
	Exponential	£28,331	£33,971	£29,866

[a] Second-best fitted fractional polynomial model
[b] Stopping rule applied where are the end of 2 years treatment, 75% of patients still receiving treatment will discontinue treatment
Sources: Tables 48 – 54 [2]

ERG comment: The ERG identified several inconsistencies and limitations regarding the DSA and PSA presented by the company. These relate to (1) the exclusion of parameters from the DSA, (2) the exclusion of parameters from the PSA, (3) the number of iterations used in the PSA, along with (4) the unexplained differences between deterministic and probabilistic results, and (5) the absence of cisplatin plus gemcitabine from the fully incremental PSA.

(1) In the DSA, the contribution of survival curves were not explored and even though stated by the company, HRs were not varied either. The ERG concludes the DSA does not accurately reflect uncertainty of the cost effectiveness of nivolumab versus the comparators.

(2) The PSA excluded HRs and Kaplan-Meier estimates used to estimate nivolumab survival before the landmark, and erroneously included patient characteristics. In response to the clarification questions, the company included Kaplan-Meier curves in the PSA, but stated that it did not include hazard ratios because *'inclusion of hazard ratios would generate illogical results due to the time-varying nature of the hazard ratios […]'* resulting in *'changes in PFS and OS that are not clinical plausible'*[7]. This was not further elaborated on and methods to correct for this were not explored. The ERG agrees that varying the HR in each time period could result in counterintuitive results but the ERG also thinks that this could have been corrected for, for example, by using a fixed set of random numbers. The company furthermore stated that the comparators' OS was accounted for via the OS estimates of nivolumab. However, it is the relative effectiveness that has the greatest effect on the model and on uncertainty and the ERG therefore does not consider this to be a valid argument and concludes that the PSA does not fulfil the NICE reference case and does not reflect a significant part of the uncertainty. The ERG therefore chose not to present the CEACs.

(3) The PSA presented by the company used 1,000 iterations, a number criticised as too small by the ERG. In response to the clarification letter, the company increased the number of iterations to 10,000, which is considered to be more appropriate. However, the ERG tested the use of 20,000 in its base-case and still noted discrepancies in incremental costs and QALYs between two runs (not in excess of £100 in costs and third decimal place utility values), thus indicating that a large number of PSA iterations is required to achieve stable results.

(4) Unfortunately, the company did not provide further information in response to the ERG clarification question on why nivolumab OS and PFS in the PSA might be lower compared to the deterministic analysis. The discrepancy between probabilistic and deterministic results persisted in the ERG's response-based analysis, with probabilistic results being the more conservative. However, the ERG noticed that using conventional, not response-based, survival analysis resulted in probabilistic model outcomes that reflected much more closely the deterministic results. The large discrepancy between probabilistic and deterministic results is likely the result of a combination of the increased uncertainty associated with the response-based approach (which in turn is caused by fitting parametric models to smaller sample sizes based on responder and non-responder groups and only using data after the landmark), the skew of the used distributions and the quantitative difference in survival between the response-based and conventional approaches (response-based approach yields an average of 2.45 and 2.8 probabilistic and deterministic nivolumab life years respectively and the conventional approach an average of 1.82 and 1.84 probabilistic and deterministic nivolumab life years respectively).

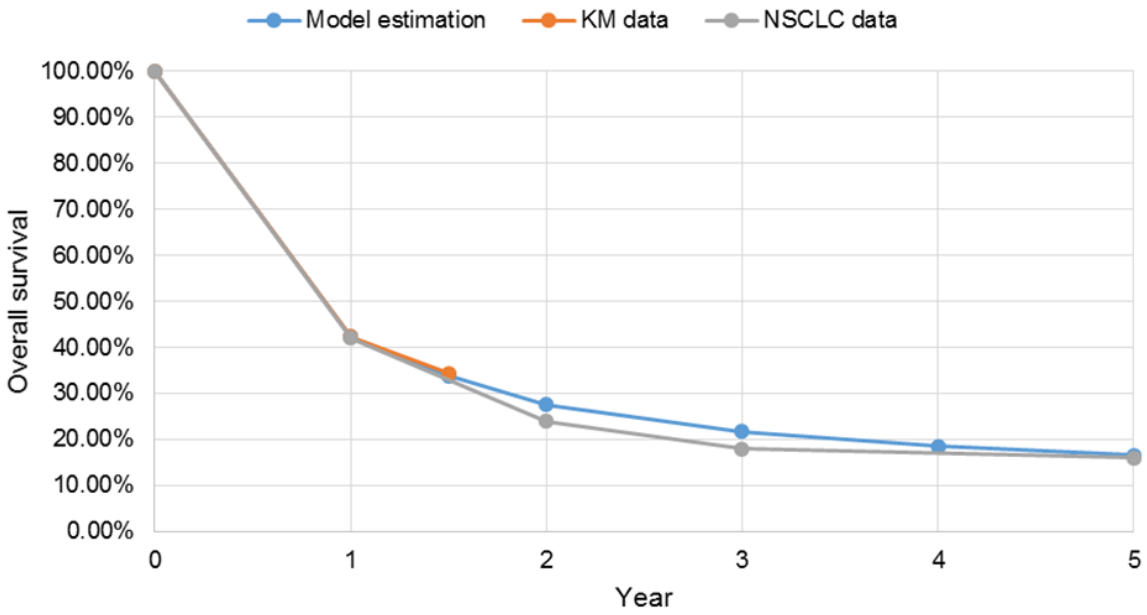
(5) In response to clarification question B13.A[7], the company provided a model that allowed for a simultaneous comparison of nivolumab to docetaxel, paclitaxel and BSC in fully incremental analysis. Despite the ERG's request to include the comparator cisplatin plus gemcitabine in the base-case, cisplatin plus gemcitabine remained excluded from the incremental PSA.

In conclusion, the ERG extended the incremental PSA to contain 10,000 iterations and to include cisplatin plus gemcitabine as a comparator.

5.2.12 Model validation and face validity check

The company undertook efforts to validate their cost effectiveness estimates for both nivolumab and comparators. The predictions of the model regarding OS and PFS were compared against expert feedback and other long-term nivolumab data in NSCLC and other solid tumours, using five-years follow up data from the CheckMate 003 study.[53] Clinical experts stated that lung cancer would be the most similar to bladder cancer, in relation to the strong link to smoking, the choice of treatment used in clinical practice, and the poor outcomes associated with both diseases without treatment. A comparison between the prediction of the generalised gamma and the CheckMate 003 data is shown in Figure 5.12.

Figure 5.12: Validation of model predictions of OS with nivolumab



Source: CS Figure 49

Validation of comparator estimates also involved comparison against expert opinion and the KM estimates derived from available clinical data (see Table 5.20). Two clinical experts stated that they would not expect more than 5% of patients to be alive at two years, when treated with the comparators. This feedback was deemed to be most closely aligned with outcomes for paclitaxel, informed by the UK PLUTO trial (see Table 5.20).[15] The company states that, because of this expert opinion, it might be that overall survival may be slightly over-estimated in the model.

Table 5.20: Comparison of overall survival extrapolation in model against observed data

Data source	Survival curve	Proportion alive, %					
		1 year	1.5 years	2 years	3 years	4 years	5 years
Nivolumab							
Model estimates for OS	Gen. Gamma (Base case)	42.34%	33.82%	27.54%	21.66%	18.51%	16.55%
CheckMate 275	Kaplan-Meier data	XXX	XXX	-	-	-	-
CheckMate 003 (NSCLC)	-	42%	-	24%	18%	-	16%
Docetaxel							
Model estimates for OS	Gen. Gamma (Base case)	25.01%	15.67%	11.05%	7.67%	6.36%	5.69%
Choueiri *et al.* (2012)[30]	Kaplan-Meier data	24.33%	13.03%	-	-	-	-

Sideris *et al.* (2016)[54]	Kaplan-Meier data (Bytescout)	19%	8%	6%	-	-	-
Paclitaxel							
Model estimates for OS	Gen. Gamma (Base case)	31.41%	17.40%	10.56%	5.66%	3.94%	3.15%
Jones *et al.* (2017)[31]	Kaplan-Meier data	31.58%	15.08%				
Sideris *et al.* (2016)[54]	Kaplan-Meier data (Bytescout)	19%	8%	6%	-	-	-
BSC							
Model estimates for OS	Gen. Gamma (Base case)	14.00%	8.96%	6.64%	5.03%	4.42%	4.09%
Bellmunt *et al.* (2013)[55]	Kaplan-Meier data	21.30%	10.65%	7.41%	1.39%	-	-
Source: CS table 55 Abbreviations: BSC: best supportive care; NSCLC: non-small cell lung cancer; OS: overall survival.							

ERG comment: The ERG's concerns include (1) the lack of internal and cross validity efforts as well as sparse use of expert opinion, (2) external validation efforts that are based on a lung cancer study, (3) the use of only CheckMate 275 for validating model predictions, as well as (4) transparency issues with the model.

(1) The company focused on external validation only. There is no description of face validity checks or cross validity checks (for instance, model outcomes could have been compared with those from TA 272[34]). It is also noteworthy that clinical experts were only consulted prior to model development at an advisory board. Clinical experts therefore did not provide feedback on the distributions used for estimating OS and PFS in the company's base-case response-based approach.

(2) The CS cites clinical experts as stating that bladder cancer is most similar to lung cancer. However, the ERG considers it questionable whether lung cancer really is similar enough to bladder cancer to enable data from the CheckMate 003 trial to be used for external validation of model predictions in bladder cancer. The cited study also was not identified through a SLR. This is of even more concern given that there are significant molecular differences in the two diseases.[5] The comparison does show a slight over-estimation of longer-term OS using the company's base-case model predictions when compared with longer-term OS data from the NSCLC study.[53]

(3) In the comparison of model predictions for OS in nivolumab patients, the company only provides data of CheckMate 275, and not the pooled estimates from CheckMate 275 and 032. This discrepancy impairs the credibility of this validation effort.

(4) The ERG wishes to highlight a few transparency issues with the submitted model file. Hidden columns on several sheets, the practice of not naming cells, the practice of disabling headings for columns and rows and the missing macro for generating the CEAC caused the ERG unnecessary difficulties in validating and amending the model.

5.3 Exploratory and sensitivity analyses undertaken by the ERG

Table 5.20 summarises all main issues highlighted by the ERG in Section 5.2, indicates the expected direction of bias introduced by these issues and whether these are examined in any analyses/incorporated in the ERG base-case.

Table 5.21: Main ERG critique of company's submitted economic evaluation

Issue	Bias introduced[a]	ERG analyses	Addressed in company analysis?
Model structure (section 5.2.2)			
• Combination of responder and non-responder groups instead of creating separate health states	+/-	NA	Requested, not addressed
Interventions and comparators (section 5.2.4)		ERG base-case (FV)	Requested, not addressed
• Exclusion of cisplatin plus gemcitabine	+		
Treatment effectiveness and extrapolation (section 5.2.6)			
• Pooling of CheckMate studies	+/-	NA	Requested, not addressed
• Response-based analysis and the use of landmark analysis, with the main issues including:	+	ERG base-case (FV), and exploratory analysis: Scenario analysis	Partly addressed, listed for each :
o the choice of landmark of 8 weeks	+/-		Requested, partly addressed
o the use of KM estimates up to the landmark	+/-		Not addressed
o rejection of proportional hazards between responders and non-responders	+/-	NA	Not addressed
o simultaneous choice of parametric time-to-event models for OS, PFS and TTD	+/-	NA	Company enabled differential selection
o a posteriori combination of responder and non-responder groups	+/-	Scenario analysis	Requested, not addressed
o application of HRs (which are derived from all patients) on the a posteriori group	+/-	NA	Not addressed
• Background mortality: error in use of UK life tables and converting rate to probability	+/-	NA	NA
• Effectiveness data derived from single-arm studies using a simulated treatment comparison results in large uncertainty and potential bias that was not quantified	+/-	ERG base-case (FE)	Not addressed
• Use of time-varying HRs	+	Exploratory analysis	Company provided time-fixed HRs, but ERG's own estimates differed
• Estimation of HRs for PFS of BSC and cisplatin plus gemcitabine based on assumptions around similarity in comparative effectiveness to vinflunine and paclitaxel	+/-	Exploratory analysis	Not addressed
• Inconsistency in that TTD analysis was not response-based, when OS and PFS were	+/-	Exploratory analysis	Company provided response-based TTD
• Proportion of responders for TTD analysis based on a sum of PFS and OS patients	-	ERG base-case (FV)	NA
Adverse events (sections 5.2.7-5.2.9)			
• Use of only CheckMate 275	+/-	NA	Requested, not addressed

Issue	Bias introduced[a]	ERG analyses	Addressed in company analysis?
• Choice of source for AE rates used for comparators not justified	+/-	NA	Not addressed
• Inclusion of AEs with incidence of <5% not in line with inclusion criteria	+/-	ERG base-case (FV)	Not addressed
Health-related quality of life (section 5.2.8)			
• Utilities only derived from CheckMate 275	-	ERG base-case (FV)	Company provided pooled utilities
• AE disutilities inconsistent with those used in ID971	+	Exploratory analysis	Not addressed
Resources and costs (section 5.2.9)			
• Technical error incorporating dose intensity	-	ERG base-case (FE)	NA
• Inconsistency in estimating weight and subsequent treatment proportions, based on CheckMate 275 only	+	ERG base-case (FV)	Not addressed
• AE unit costs inconsistent with ID971	+	Exploratory analysis	Not addressed
• Cisplatin plus gemcitabine administration scheme not reflective of UK practice	+	Exploratory analysis	Not addressed
• Assumption that all delayed doses were missed doses	+	ERG base-case (MJ)	Not addressed
• Assumption that dose intensity for the comparators is equal to that of nivolumab	+	NA	NA
Cost-effectiveness analyses (sections 5.2.10 and 5.2.11)			
• Relative effectiveness not considered in the PSA	+/-	NA	Requested, not addressed
• Patient characteristics included in PSA	+/-	ERG base-case (FV)	Not addressed
• OS and PFS under-estimated in PSA compared to deterministic analysis	+/-	NA	Not addressed
Validation (section 5.2.12)			
• Insufficient validation of the model	+/-	NA	Not addressed

Abbreviations: NA, not applicable; FE, fixing error; FV, fixing violations; MJ, matters of judgement

[a]Likely conservative assumptions (of the intervention versus all comparators) are indicated by '-'; while '+/-' indicates that the bias introduced by the issue is unclear to the ERG and '+' in indicates that the ERG believes this issue likely induces bias in favour of the intervention versus at least one comparator.

Based on all considerations from Section 5.2 (summarised in Table 5.21), the ERG defined a new base-case. This base-case included multiple adjustments to the original base-case presented in the previous sections. These adjustments made by the ERG form the ERG base-case and were subdivided into three categories (derived from Kaltenthaler 2016):[56]

- Fixing errors (correcting the model where the company's submitted model was unequivocally wrong)
- Fixing violations (correcting the model where the ERG considered that the NICE reference case, scope or best practice had not been adhered to)
- Matters of judgement (amending the model were the ERG considers that reasonable alternative assumptions are preferred)

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 £87,709, £68,519 and £69,515 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 £83,397, £65,411 and £67,175 per QALY gained, with cisplatin plus gemcitabine dominating nivolumab.

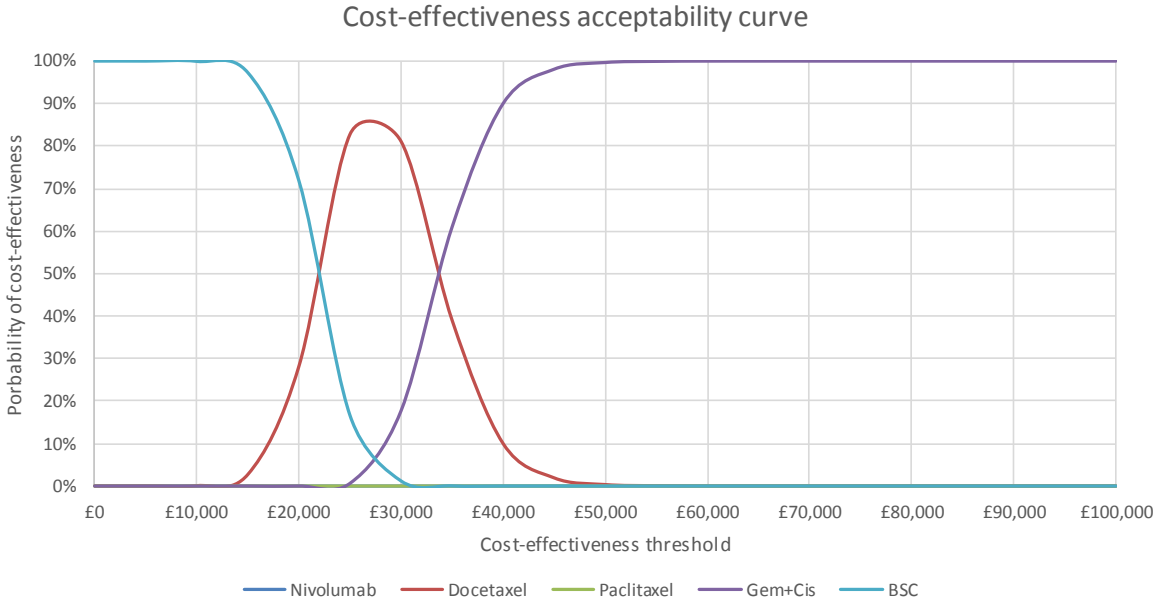
Table 5.22: ERG base-case (probabilistic)

	Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Nivolumab ICER (£/QALY)
ERG base-case	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£12,493	0.74	[redacted]	[redacted]	£87,709
	Paclitaxel	£13,866	0.63	[redacted]	[redacted]	£68,519
	Cis + gem	£29,384	1.24	[redacted]	[redacted]	Nivolumab is dominated
	BSC	£8,696	0.56	[redacted]	[redacted]	£69,515

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 5.13: 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.

5.3.2 Additional exploratory analyses performed based on the ERG base-case

Additional sensitivity analyses were performed to examine the potential impact of the following 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 (ERG base-case apart from 9.). Results are presented in Tables 6.2 in Section 6.

a) Exploratory analyses using the ERG base-case:

1. Alternative parametric time-to-event models: use of the lognormal distribution for OS (best-fitting according to AIC/BIC) and log-logistic for PFS (best fitting according to BIC, second-best according to AIC).
2. Use of alternative specifications for the fractional polynomial model, by employing a 'mini-PSA' across the different p1 and p2 values provided by the company in response to clarification questions. Results are presented as credible intervals about incremental costs and QALYs and the resulting range of ICERs in Table 6.3 in Section 6.
3. Use of naïve comparison performed by the ERG, instead of the STC, to derive HRs for OS and PFS. The ERG noticed that the code supplied to estimate the time-dependent HRs only estimated them up to a time horizon of 256 weeks, ending much before the end of the model time horizon. It is not clear where the time-dependent HRs implemented after 260 weeks were sourced from. The ERG used the company's time-dependent HRs after 260 weeks, which should not be influential and work in favour of nivolumab.
4. Use of time-independent HRs for OS and PFS derived by the ERG instead of time-dependent HRs.
5. Use of HRs for OS as proxy for HR for PFS for the comparisons with BSC and cisplatin plus gemcitabine.
6. Use of adverse event disutilities and resource use from technology appraisal ID971.
7. Use of the UK dosage schedule for cisplatin plus gemcitabine.
8. An extreme scenario of assuming no treatment effect of nivolumab vs comparators.

b) Exploratory analyses on the ERG base-case using response-based analysis for OS, PFS and TTD:

1. Maintaining the company's base-case choice of parametric time-to-event models, i.e. the generalised gamma for responders' and non-responders' OS, PFS and TTD.
2. Use of parametric time-to-event models with the best fit for OS and PFS (based on AIC/BIC) for responder OS and PFS (generalised gamma), non-responder OS and PFS (Weibull), but maintaining responder and non-responder TTD as the generalised gamma.
3. Use of parametric time-to-event models with the best fit (based on AIC/BIC) for responder OS and PFS (generalised gamma), non-responder OS and PFS (Weibull), responder TTD (lognormal) and non-responder TTD (Gompertz).
4. Use of 26-week landmark instead of 8-week landmark

5.3.3 Subgroup analyses performed based on the ERG base-case

No subgroup analyses were performed.

5.4 Conclusions of the cost effectiveness section

The majority of 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 Sections 5.2.2 and 5.2.4.[33]

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 £87,709, £68,519 and £69,515 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 £83,397, £65,411 and £67,175 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 (£92,335, £64,914, dominated, £65,593 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 (£71,639, £95,775, £76,576, £55,577 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 ICERs more favourable to nivolumab (£45,721, £39,286, £72,732, £38,147 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, PFS and TTD, however, resulted in a marked increase in ICERs compared with the response-based company's base-case (£77,597, £67,608, £143,923, £64,282 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 (£75,094, £71,255, £87,022, £61,647 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 6.1: 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
	Nivolumab	[redacted]	[redacted]			

	Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Nivolumab ICER (£/QALY)
Utilities from pooled CheckMate studies (6)[b]	Docetaxel	£12,803	0.84	[redacted]	[redacted]	£49,613
	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	£12,803	0.82	[redacted]	[redacted]	£52,858
	Paclitaxel	£14,198	0.71	[redacted]	[redacted]	£44,330
	Cis+gem	£30,315	1.35	[redacted]	[redacted]	£97,665
	BSC	£8,835	0.58	[redacted]	[redacted]	£43,958
ERG base-case (combining adjustments 1-10)	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£12,493	0.74	[redacted]	[redacted]	£87,709
	Paclitaxel	£13,866	0.63	[redacted]	[redacted]	£68,519

	Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Nivolumab ICER (£/QALY)
	Cis+gem	£29,384	1.24	[redacted]	[redacted]	Nivolumab is dominated
	BSC	£8,696	0.56	[redacted]	[redacted]	£69,515

Note: [a] results have been reproduced by the ERG, based on the economic model submitted by the company in their clarification response; [b] this scenario is conditional on the fixing errors adjustment (adjustments 1 and 2)
ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year

Table 6.2: 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	£12,493	0.74	[redacted]	[redacted]	£87,709
	Paclitaxel	£13,866	0.63	[redacted]	[redacted]	£68,519
	Cis+gem	£29,384	1.24	[redacted]	[redacted]	Dominated
	BSC	£8,696	0.56	[redacted]	[redacted]	£69,515
Alternative parametric TTE models (lognormal for OS, log-logistic for PFS) (A.1)	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£13,173	1.01	[redacted]	[redacted]	£45,721
	Paclitaxel	£14,654	0.89	[redacted]	[redacted]	£39,286
	Cis+gem	£29,736	1.58	[redacted]	[redacted]	£72,732
	BSC	£9,235	0.72	[redacted]	[redacted]	£38,147
Naïve comparison data instead of STC results (A.3)	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£13,005	0.77	[redacted]	[redacted]	£92,335
	Paclitaxel	£13,914	0.60	[redacted]	[redacted]	£64,914
	Cis+gem	£30,910	1.56	[redacted]	[redacted]	Dominated

	Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Nivolumab ICER (£/QALY)
	BSC	£8,630	0.52	[redacted]	[redacted]	£65,593
Time-independent HRs (A.4)	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£10,213	0.60	[redacted]	[redacted]	£71,639
	Paclitaxel	£13,081	0.78	[redacted]	[redacted]	£95,775
	Cis+gem	£26,584	0.86	[redacted]	[redacted]	£76,576
	BSC	£8,173	0.40	[redacted]	[redacted]	£55,577
Alternative assumptions for PFS HRs for BSC and cis+gem (A.5)	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£12,507	0.74	[redacted]	[redacted]	£87,863
	Paclitaxel	£13,858	0.63	[redacted]	[redacted]	£68,679
	Cis+gem	£34,999	1.26	[redacted]	[redacted]	Dominated
	BSC	£8,698	0.55	[redacted]	[redacted]	£68,369
AE disutilities and resource use from TA ID971 (A.6)	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£12,068	0.74	[redacted]	[redacted]	£89,222
	Paclitaxel	£13,695	0.63	[redacted]	[redacted]	£69,051
	Cis+gem	£26,508	1.26	[redacted]	[redacted]	Dominated
	BSC	£8,750	0.56	[redacted]	[redacted]	£69,622
UK dosage schedule for cis+gem (A.7)	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£12,476	0.74	[redacted]	[redacted]	£87,722
	Paclitaxel	£13,852	0.63	[redacted]	[redacted]	£68,621
	Cis+gem	£31,195	1.24	[redacted]	[redacted]	Dominated
	BSC	£8,678	0.56	[redacted]	[redacted]	£69,560
No treatment effect of nivolumab vs comparators (A.8)	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£13,726	1.19	[redacted]	[redacted]	£5,740,183
	Paclitaxel	£14,270	1.19	[redacted]	[redacted]	£11,382,482
	Cis+gem	£32,028	1.15	[redacted]	[redacted]	£415,600
	BSC	£10,635	1.16	[redacted]	[redacted]	£1,168,837
Response-based analysis	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£12,783	0.84	[redacted]	[redacted]	£53,273

	Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Nivolumab ICER (£/QALY)
using ERG base-case (B.1)	Paclitaxel	£14,163	0.73	[redacted]	[redacted]	£44,877
	Cis+gem	£30,310	1.39	[redacted]	[redacted]	£103,186
	BSC	£8,811	0.59	[redacted]	[redacted]	£44,183
Response-based analysis using alternative TTE models for OS, PFS, but not TTD (B.2)	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£12,475	0.77	[redacted]	[redacted]	£78,795
	Paclitaxel	£13,983	0.68	[redacted]	[redacted]	£68,594
	Cis+gem	£29,893	1.25	[redacted]	[redacted]	£146,721
	BSC	£8,678	0.55	[redacted]	[redacted]	£65,249
Response-based analysis using alternative TTE models for OS, PFS and TTD (B.3)	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£12,452	0.77	[redacted]	[redacted]	£77,597
	Paclitaxel	£13,948	0.67	[redacted]	[redacted]	£67,608
	Cis+gem	£29,880	1.25	[redacted]	[redacted]	£143,923
	BSC	£8,662	0.55	[redacted]	[redacted]	£64,282
Response-based analysis using 26-week landmark (B.4)	Nivolumab	[redacted]	[redacted]			
	Docetaxel	£10,849	0.51	[redacted]	[redacted]	£75,094
	Paclitaxel	£13,689	0.52	[redacted]	[redacted]	£71,255
	Cis+gem	£28,678	0.79	[redacted]	[redacted]	£87,022
	BSC	£8,035	0.35	[redacted]	[redacted]	£61,647

Note: [a] results have been reproduced by the ERG, based on the economic model submitted by the company in their clarification response
ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year

Table 6.3. Impact of using different parameter values in the fractional polynomial model for NMA

Technologies	Incremental costs (CI) of nivolumab vs comparators		Incremental QALYs (CI) of nivolumab vs comparators		ICER of nivolumab vs comparators	
	Lower	Upper	Lower	Upper	Range based on CIs for incremental costs and QALYs	
Docetaxel	[redacted]	[redacted]	[redacted]	[redacted]	£178,199	£52,441
Paclitaxel	[redacted]	[redacted]	[redacted]	[redacted]	£160,141	£47,615
Cis + gem	[redacted]	[redacted]	[redacted]	[redacted]	Dominated	£35,146
BSC	[redacted]	[redacted]	[redacted]	[redacted]	£96,636	£43,847

7. END OF LIFE

The company discusses the end-of life criteria in section B.2.13.2 of the CS, arguing that nivolumab fulfils the end-of-life criteria in this appraisal.[2]

This argument is partly based on lack of evidence to argue that it does not – *'no study provided evidence of OS estimates for this patient population that approached the 24 months that represents the threshold for NICE's end of life criteria',* and partly on very weak evidence from the economic model based on a comparison of single arm studies – *'The economic analysis predicted mean life years per patient with nivolumab of 2.78 years (33.36 months). In comparison, predicted mean life years per patient with comparator therapies were 1.19 years (14.28 months) with paclitaxel, 1.40 years (16.80 months) with docetaxel and 1.01 years (12.12 months) with BSC'.*

We agree that there is no evidence to argue that nivolumab does not fulfil the end-of-life criteria in this appraisal. But, at the same time, there is no robust evidence to argue that it does.

8. OVERALL CONCLUSIONS

8.1 Statement of principal findings

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. The systematic review was performed to a good standard.

The identification of two single arm studies for nivolumab, CheckMate 275 and CheckMate 032, 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). In terms of ORR the main analysis using the fixed effect model presented finds that nivolumab is significantly better than BSC and docetaxel. No significant differences were found for nivolumab paclitaxel and gemcitabine. In the random effects model nivolumab is only statistically significantly superior to BSC. In the naïve indirect comparison nivolumab is superior to all three comparators in the fixed effect model but only to BSC in the random effects model. The results of the analysis using fixed effect fractional polynomial model (allowing variation of HRs over time) based on the STC show that for OS and PFS nivolumab is superior to all comparators at most time points. However, the credible intervals for the HRs are quite wide, crossing 1 in many cases. The results of the naïve indirect comparison i.e. with the fractional polynomial model, but without the STC, were not reported. The results assuming a proportional hazards model i.e. fixed HRs were reported in the response to the clarification request, although were derived by a method that lacked validity and were quite different to those obtained by the ERG using a method advocated in the paper on which the company analysis was based. Very few of the many functional forms of the fractional polynomial model were explored.

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.[1] The ERG found several serious limitations in the STC analysis. In particular, the major assumption for unanchored STC is that all effect modifiers or prognostic variables are accounted for. Not all of the key characteristics (possible effect modifiers or prognostic variables) for the STC were reported for all comparator trials, therefore imputations were required for these characteristics which were based on correlations to the baseline characteristics in the nivolumab trials. Also, the method used for the prediction models lacked transparency; the results at each stage of the stepwise selection process were not provided. In particular, it is not clear that the most parsimonious model is the best model. The ERG was able to produce the results based on a naïve comparison (without the STC), which verified the adoption of the PH model used in the STC i.e. all HRs of nivolumab versus each comparator were multiplied by a single factor, the HR of the adjusted (by the STC) vs. unadjusted hazard for nivolumab. However, it would have been useful to see an STC that was based on prediction models with more covariates including all eleven considered. The only external test of validity of the STC i.e. the 'out-of-sample' method seemed to either show insufficient reduction in bias or be inapplicable given the use of the fractional polynomial model that was used for survival analysis. As stated on page 56 of TSD 18: '*The size of this systematic error can certainly be reduced, and probably substantially, by appropriate use of…STC. Much of the literature on unanchored … STC acknowledges the possibility of residual bias due to unobserved prognostic variables and effect modifiers; however, it is not made clear that the accuracy of the resulting estimates is entirely unknown, because there is no analysis of the potential magnitude of residual bias, and hence no idea of the degree of error in the unanchored estimates. It is, of course, most unlikely 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 £87,709, £68,519 and £69,515 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.

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.

8.2 Strengths and limitations of the assessment

The searches for eligible studies in the CS were well documented and easily reproducible. Searches were carried out on a wide range of databases and other resources. Supplementary searches of conference proceedings and organisational websites, and the checking of references lists were undertaken by the company in order to identify additional studies not retrieved by the main searches. However, the search for English language studies only in the MEDLINE and Embase searches in the clinical effectiveness section was felt to be a limitation. The systematic review was well conducted, but no randomised controlled trials (RCTs) were identified for nivolumab and there were no studies that directly compared nivolumab with any specified comparator. Furthermore, there were no studies that could provide a common comparator to support indirect comparison or MTC. The STC analysis is compromised by many limitations (listed earlier) which impairs the ability to critique the presence of residual bias. Given that the TSD 18 states 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…"* the ERG does not think the STC methods are sufficiently reported nor validated to sustain the companies claims.[1]

The economic model had a structure similar to past NICE technology appraisals in metastatic cancer but deviated from conventional survival modelling in that it used a response-based approach. This was inconsistently implemented, insufficiently justified and alternative approaches were not explored. The uncertainty and bias potentially introduced by this approach could not be completely explored. The lack of validation of model predictions raised concerns about the validity of CS model results. Lastly, the exclusion of cisplatin plus gemcitabine from the company's base-case stands in contrast to the scope and lacked appropriate justification.

8.3 Suggested research priorities

The ERG recommends the conduct of an RCT of nivolumab versus at least one of the comparators or perhaps an investigator choice design, which might be lacking in terms of power, depending on time believed to be reasonable to recruit, but would provide at least some unbiased evidence of effectiveness.

9. REFERENCES

[1] Phillippo D, Ades T, Dias S, Palmer S, Abrams KR, Welton N. *NICE DSU Technical Support Document 18: methods for population-adjusted indirect comparisons in submissions to NICE*. Sheffield: NICE Decision Support Unit, 2016. 81p.

[2] Bristol-Myers Squibb Pharmaceuticals Ltd. *Nivolumab for treating metastatic or unresectable urothelial cancer after platinum-based chemotherapy [ID995]. Document B: Company evidence submission. Submission to National Institute of Health and Clinical Excellence. Single technology appraisal (STA)*: Bristol-Myers Squibb Pharmaceuticals Ltd, 2017. 143p.

[3] Cancer Research UK. Cancer statistics for the UK [Internet]. 2017 [accessed 20.7.17]. Available from: http://www.cancerresearchuk.org/health-professional/cancer-statistics

[4] National Institute for Health and Care Excellence. *Bladder cancer: diagnosis and management* London: NICE, 2015. 57p. Available from: http://nice.org.uk/guidance/ng2

[5] Bristol-Myers Squibb. *Meeting minutes, Clinical Advisory Board Meeting: 06 March 2017, 0930-1600 [Not in the public domain]*. 30 Euston Square, London, UK: Bristol-Myers Squibb, 2017

[6] National Institute for Health and Care Excellence. *Nivolumab for treating metastatic or unresectable urothelial cancer. Final scope [Internet]*. London: NICE, 2017 [accessed 3.7.17]. 4p. Available from: https://www.nice.org.uk/guidance/gid-ta10163/documents/final-scope

[7] Bristol-Myers Squibb Pharmaceutical Ltd. *Nivolumab for treating metastatic or unresectable urothelial cancer after platinum-based chemotherapy [ID995] – Response to request for clarification from the ERG*: Bristol-Myers Squibb Pharmaceutical Ltd, 2017. 97p.

[8] Sharma P, Retz M, Siefker-Radtke A, Baron A, Necchi A, Bedke J, et al. Nivolumab in metastatic urothelial carcinoma after platinum therapy (CheckMate 275): a multicentre, single-arm, phase 2 trial. *Lancet Oncol* 2017;18(3):312-322.

[9] Sharma P, Callahan MK, Bono P, Kim J, Spiliopoulou P, Calvo E, et al. Nivolumab monotherapy in recurrent metastatic urothelial carcinoma (CheckMate 032): a multicentre, open-label, two-stage, multi-arm, phase 1/2 trial. *Lancet Oncol* 2016;17(11):1590-1598.

[10] Bristol-Myers Squibb Pharmaceuticals Ltd. *CheckMate 275: Clinical Study Report for Study CA209275 (25th July 2016)*, 2016

[11] Bristol-Myers Squibb Pharmaceuticals Ltd. *CheckMate 032: Clinical Study Report for Study CA209032 (29th June 2016)*. 2016

[12] European Medicines Agency. *Summary of opinion (initial authorisation). Maviret: glecaprevir/pibrentasvir [Internet]*. London: European Medicines Agency, 2017 [accessed 15.8.17]. 1p. Available from:

http://www.ema.europa.eu/docs/en_GB/document_library/Summary_of_opinion_-_Initial_authorisation/human/004430/WC500229790.pdf

[13] Gondo T, Ohori M, Hamada R, Tanaka A, Satake N, Takeuchi H, et al. The efficacy and safety of gemcitabine plus cisplatin regimen for patients with advanced urothelial carcinoma after failure of M-VAC regimen. *Int J Clin Oncol* 2011;16(4):345-51.

[14] Ozawa A, Tanji N, Ochi T, Yanagihara Y, Kikugawa T, Yamaguchi A, et al. Gemcitabine and cisplatin for advanced urothelial carcinomas: the Ehime University Hospital experience. *Int J Clin Oncol* 2007;12(4):279-83.

[15] Jones RJ, Hussain SA, Protheroe AS, Birtle A, Chakraborti P, Huddart RA, et al. Randomized phase II study investigating pazopanib versus weekly paclitaxel in relapsed or progressive urothelial cancer. *J Clin Oncol* 2017;35(16):1770-1777.

[16] Petrylak DP, Tagawa ST, Kohli M, Eisen A, Canil C, Sridhar SS, et al. Docetaxel as monotherapy or combined with ramucirumab or icrucumab in second-line treatment for locally advanced or metastatic urothelial carcinoma: an open-label, three-arm, randomized controlled phase II trial. *J Clin Oncol* 2016;34(13):1500-9.

[17] Canadian Agency for Drugs and Technologies in Health. *CADTH peer review checklist for search strategies [Internet]*. Ottawa: CADTH, 2013 [accessed 17.7.13]. 3p. Available from: http://www.cadth.ca/en/resources/finding-evidence-is

[18] National Institute for Health and Care Excellence. *Single technology appraisal: user guide for company evidence submission template [Internet]*. London: NICE, 2015 [accessed 3.7.17]. 52p. Available from: https://www.nice.org.uk/guidance/pmg24/resources/single-technology-appraisal-user-guide-for-company-evidence-submission-template-pdf-72286715419333

[19] Lefebvre C, Manheimer E, Glanville J. 6.4.9 Language, date and document format restrictions (Chapter 6: Searching for studies) [Internet]. In: Higgins JPT, Green S, editors. Version 5.1.0 [updated March 2011]: The Cochrane Collaboration, 2011 [accessed 7.8.17]. Available from: http://handbook.cochrane.org/

[20] Bristol-Myers Squibb Pharmaceuticals Ltd. *Nivolumab for treating metastatic or unresectable urothelial cancer after platinum-based chemotherapy [ID995]. Appendices. Submission to National Institute of Health and Clinical Excellence. Single technology appraisal (STA)*: Bristol-Myers Squibb Pharmaceuticals Ltd, 2017. 257p.

[21] National Institute for Health and Care Excellence. *Nivolumab for treating metastatic or unresectable urothelial cancer after platinum-based chemotherapy [ID995]. Clarification letter*. London: NICE, 2017

[22] Centre for Reviews and Dissemination. *Quality Assessment: Cohort Studies (Box 5.9). CRD Report 4: Undertaking systematic reviews of research on effectiveness*. York: University of York, 2001

[23] Kim YS, Lee SI, Park SH, Park S, Hwang IG, Lee SC, et al. A phase II study of weekly docetaxel as second-line chemotherapy in patients with metastatic urothelial carcinoma. *Clin Genitourin Cancer* 2016;14(1):76-81.

[24] McCaffrey JA, Hilton S, Mazumdar M, Sadan S, Kelly WK, Scher HI, et al. Phase II trial of docetaxel in patients with advanced or metastatic transitional-cell carcinoma. *J Clin Oncol* 1997;15(5):1853-7.

[25] Vaughn DJ, Broome CM, Hussain M, Gutheil JC, Markowitz AB. Phase II trial of weekly paclitaxel in patients with previously treated advanced urothelial cancer. *J Clin Oncol* 2002;20(4):937-40.

[26] Bellmunt J, Theodore C, Demkov T, Komyakov B, Sengelov L, Daugaard G, et al. Phase III trial of vinflunine plus best supportive care compared with best supportive care alone after a platinum-containing regimen in patients with advanced transitional cell carcinoma of the urothelial tract. *J Clin Oncol* 2009;27(27):4454-61.

[27] Choueiri TK, Ross RW, Jacobus S, Vaishampayan U, Yu EY, Quinn DI, et al. Double-blind, randomized trial of docetaxel plus vandetanib versus docetaxel plus placebo in platinum-pretreated metastatic urothelial cancer. *J Clin Oncol* 2012;30(5):507-12.

[28] Joly F, Houede N, Noal S, Chevreau C, Priou F, Chinet-Charrot P, et al. Do patients with advanced urothelial carcinoma benefit from weekly paclitaxel chemotherapy? A GETUG phase II study. *Clin Genitourin Cancer* 2009;7(2):E28-33.

[29] Jansen JP. Network meta-analysis of survival data with fractional polynomials. *BMC Med Res Methodol* 2011;11:61.

[30] Choueiri TK, Ross RW, Jacobus S, Vaishampayan U, Yu EY, Quinn DI, et al. Double-blind, randomized trial of docetaxel plus vandetanib versus docetaxel plus placebo in platinum-pretreated metastatic urothelial cancer. *Journal of Clinical Oncology* 2012;30(5):507-12.

[31] Jones R, Hussain S, Protheroe A, Birtle A, Chakraborti P, Huddart R, et al. Randomized Phase II Study Investigating Pazopanib Versus Weekly Paclitaxel in Relapsed or Progressive Urothelial Cancer. *J Clin Oncol* 2017;0(0):JCO.2016.70.7828.

[32] Petrylak DP, Tagawa ST, Kohli M, Eisen A, Canil C, Sridhar SS, et al. Docetaxel as monotherapy or combined with ramucirumab or icrucumab in second-line treatment for locally advanced or metastatic urothelial carcinoma: an open-label, three-arm, randomized controlled phase II trial. *J Clin Oncol* 2016;34(13):1500-9.

[33] National Institute for Health and Care Excellence. *Guide to the methods of technology appraisal 2013 [Internet]*. London: NICE, 2013 [accessed 23.8.17]. 93p. Available from: http://publications.nice.org.uk/pmg9

[34] National Institute for Health and Care Excellence. *Vinflunine for the treatment of advanced or metastatic transitional cell carcinoma of the urothelial tract [TA272]*. London: National Institute for Health and Care Excellence, 2013 [accessed 3 Feb 2017] Available from: https://www.nice.org.uk/guidance/ta272

[35] Robinson P, Maase H, Bhalla S, Kielhorn A, Aristides M, Brown A, et al. Cost-utility analysis of the GC versus MVAC regimens for the treatment of locally advanced or metastatic bladder cancer. *Expert Rev Pharmacoecon Outcomes Res* 2004;4(1):27-38.

[36] Scottish Medicines Consortium. *Vinflunine (as ditartrate), 25mg/mL, concentrate for solution for infusion (Javlor(R)) SMC No. (686/11)*, 2015 [accessed 27.02.17] Available from: https://www.scottishmedicines.org.uk/files/advice/vinflunine_Javlor_Resubmission_FINAL_June_2015_for_website.pdf

[37] Woods B, Sideris E, Palmer S, Latimer N, Soares M. *NICE DSU Technical Support Document 19: Partitioned survival analysis for decision modelling in health care: a critical review*. Sheffield: Decision Support Unit, ScHARR, 2017. 72p. Available from: http://scharr.dept.shef.ac.uk/nicedsu/wp-content/uploads/sites/7/2017/06/Partitioned-Survival-Analysis-final-report.pdf

[38] Latimer N. *NICE DSU Technical Support Document 14: Undertaking survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data*. Sheffield: Decision Support Unit, ScHARR, 2017. 52p. Available from: http://scharr.dept.shef.ac.uk/nicedsu/wp-content/uploads/sites/7/2016/03/NICE-DSU-TSD-Survival-analysis.updated-March-2013.v2.pdf

[39] Dafni U. Landmark analysis at the 25-year landmark point. *Circ Cardiovasc Qual Outcomes* 2011;4(3):363-71.

[40] 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 - committee papers [Internet]*, 2017 [accessed 9.8.17] Available from: https://www.nice.org.uk/guidance/gid-ta10080/documents/committee-papers

[41] Bellmunt J, Fougeray R, Rosenberg JE, von der Maase H, Schutz FA, Salhi Y, et al. Long-term survival results of a randomized phase III trial of vinflunine plus best supportive care versus best supportive care alone in advanced urothelial carcinoma patients after failure of platinum-based chemotherapy. *Ann Oncol* 2013;24(6):1466-72.

[42] Berdeaux G, Roskell N, Hemstock M. Influence of the method of analysis on estimates of QALY treatment difference: Phase III trial of vinflunine versus best supportive care in patients with TCCU. *Value Health* 2015;18(7):A467.

[43] von der Maase H, Hansen SW, Roberts JT, et al. Gemcitabine and cisplatin versus methotrexate, vinblastine, doxorubicin, and cisplatin in advanced or metastatic bladder cancer: results of a large, randomized, multinational, multicenter, phase III study. *J Clin Oncol* 2000;18:3068-77.

[44] Davey P, Rajan N, De Souza P, et al. Examining preferences, utility values and cost-effectiveness for gemcitabine plus cisplatin (GEM/cis) for the treatment of bladder cancer - a discrete choice conjoint analysis conducted in Australia. *Eur J Cancer* 2001;37:S163.

[45] Pickard AS, Jiang R, Lin HW, et al. Using patient-reported outcomes to compare relative burden of cancer: EQ-5D and Functional Assessment of Cancer Therapy-General in eleven types of cancer. *Clin Ther* 2016;38:769-777.

[46] Attard CL, Brown S, Alloul K, Moore MJ. Cost-effectiveness of folfirinox for first-line treatment of metastatic pancreatic cancer. *Curr Oncol* 2014;21(1):e41-51.

[47] Nafees B, Stafford M, Gavriel S, Bhalla S, Watkins J. Health state utilities for non small cell lung cancer. *Health Qual Life Outcomes* 2008;6:84.

[48] Beusterien KM, Davies J, Leach M, Meiklejohn D, Grinspan JL, O'Toole A, et al. Population preference values for treatment outcomes in chronic lymphocytic leukaemia: a cross-sectional utility study. *Health Qual Life Outcomes* 2010;8:50.

[49] National Institute for Health and Care Excellence. *Nintedanib for pre Nintedanib for previously treated locally viously treated locally advanced, metastatic, or locally recurrent advanced, metastatic, or locally recurrent non-small-cell lung cancer [TA347]*. London: National Institute for Health and Care Excellence, 2015 [accessed 23.8.17] Available from: https://www.nice.org.uk/guidance/ta347

[50] Frederix GW, Quadri N, Hovels AM, van de Wetering FT, Tamminga H, Schellens JH, et al. Utility and work productivity data for economic evaluation of breast cancer therapies in the Netherlands and Sweden. *Clin Ther* 2013;35(4):e1-7.

[51] 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] [Internet]. 2017 [accessed 17.7.17]. Available from: https://www.nice.org.uk/guidance/indevelopment/gid-ta10080/documents

[52] Addicott R, Dewar S. *Improving choice at end of life: a descriptive analysis of the impact and costs of the Marie Curie Delivering Choice programme in Lincolnshire*. London: Kings Fund, 2008. 45p. Available from: https://www.kingsfund.org.uk/publications/improving-choice-end-life

[53] Brahmer J, Horn L, Jackman D, Spigel D, Antonia S, Hellmann M, et al. Abstract CT077: Five-year follow-up from the CA209-003 study of nivolumab in previously treated advanced non-small cell lung cancer (NSCLC): Clinical characteristics of long-term survivors. *Cancer Res* 2017;77(13 Supplement):CT077-CT077.

[54] Sideris S, Aoun F, Zanaty M, Martinez NC, Latifyan S, Awada A, et al. Efficacy of weekly paclitaxel treatment as a single agent chemotherapy following first-line cisplatin treatment in urothelial bladder cancer. *Mol Clin Oncol* 2016;4(6):1063-1067.

[55] Bellmunt J, Fougeray R, Rosenberg JE, von der Maase H, Schutz FA, Salhi Y, et al. Long-term survival results of a randomized phase III trial of vinflunine plus best supportive care versus best supportive care alone in advanced urothelial carcinoma patients after failure of platinum-based chemotherapy. *Annals of Oncology* 2013;24(6):1466-1472.

[56] Kaltenthaler E, Carroll C, Hill-McManus D, Scope A, Holmes M, Rice S, et al. The use of exploratory analyses within the National Institute for Health and Care Excellence single technology appraisal process: an evaluation and qualitative analysis. *Health Technol Assess* 2016;20(26):1-48.

Appendix 1: Details of ERG analyses (for validation purposes)

Altered cells are printed in *italics*.

Fixing errors

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

Addition of tab "National life table"; General mortality data!CA3:CB73

2. Error in calculating dose intensity

Drug costs!E14:E20; Drug costs!E32:E36; Drug costs!E43:E47; Drug costs!E54:E58; Drug costs!E65:E69; Costs & Resource Use!E32:E35

Fixing violations

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

PSA Simulation!H11;PSA Simulation!R11; PSA Simulation Y14:AC10013

Of note: the ERG added the total LY for each comparator (in each PSA draw) in columns J to N of the PSA Simulation-sheet.

4. Calculation of responder and non-responder proportions for response-based TTD analysis based on OS and PFS, thereby double-counting patients.

Discontinuation!CM23:CN24

5. Adverse events with an incidence <5% were included in the model, despite the company stating that these should be excluded.

Adverse Events!I13; Adverse Events!I17; Adverse Events!J13; Adverse Events!J15; Adverse Events!K13; Adverse Events!K15

6. Use of utilities from CheckMate 275 only.

LIVE!E32:E33

7. Use of BSA and weight from CheckMate 275 only.

Set-Up!E28

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.

PSA Distributions!J13:J16; PSA Distributions!J19:J22

Matters of judgement

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

PFS & OS!BS11; PFS & OS!BP18:BU30; PFS & OS!BY21:CV470; PFS & OS!DL21:DM470; Discontinuation!AH27:AH447; Discontinuation!BD27:BD447

10. The assumption that all delayed doses are missed doses.

Costs & Resource Use!I24:I28

National Institute for Health and Care Excellence

Centre for Health Technology Evaluation

Pro-forma Response

ERG report

Nivolumab for treating metastatic or unresectable urothelial cancer after platinum-based chemotherapy [ID995]

You are asked to check the ERG report from Kleijnen Systematic Reviews Ltd to ensure there are no factual inaccuracies contained within it.

If you do identify any factual inaccuracies you must inform NICE by **5pm on Monday 11 September 2017** using the below proforma comments table. All factual errors will be highlighted in a report and presented to the Appraisal Committee and will subsequently be published on the NICE website with the Evaluation report.

The proforma document should act as a method of detailing any inaccuracies found and how and why they should be corrected.

Issue 1

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
At multiple points throughout the report, the ERG state that "the company did not provide the comparison of nivolumab with cisplatin plus gemcitabine in the base-case, despite it being in the scope". This statement is inaccurate.	It is suggested that these statements should be removed.	The NICE final scope does not explicitly state cisplatin plus gemcitabine as a comparator. The comparators are listed as • Retreatment with 1st line platinum-based chemotherapy (only for people whose disease has had an adequate response) • Docetaxel • Paclitaxel • Best supportive care This statement is therefore inaccurate and should be removed.	This is not a factual inaccuracy. The scope states: '• Retreatment with 1st line platinum-based chemotherapy (only for people whose disease has had an adequate response)' Although cisplatin plus gemcitabine is not explicitly mentioned it is one example of a platinum-based chemotherapy.

Issue 2

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 31 of the report states that "the placement following progression subsequent to muscle-invasive disease (stage II) is not within scope." This statement is inaccurate.	It is suggested that this statement should be removed.	The scope states that the population of the appraisal is "adults with metastatic or unresectable urothelial cancer whose disease has progressed after platinum-based chemotherapy". As the scope does not state at which stage of disease patients must have received prior platinum-based chemotherapy, it is	This is not a factual inaccuracy. It is the most obvious interpretation of the scope given that it states: 'has progressed' and not 'had progressed' suggesting that the progression follows achieving the status of metastatic or

		inaccurate to state that progression subsequent to muscle-invasive disease (stage II) is not within the scope. Patients who have received platinum-based chemotherapy at the muscle-invasive disease stage and have subsequently progressed to locally advanced or metastatic disease would therefore be consistent with the scope.	unresectable.

Issue 3

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 36 of the ERG report states, regarding the inclusion criteria of CheckMate 032, that 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'. This statement is not fully correct and the third criterion here should be "after previously refusing standard treatment with chemotherapy for the treatment of metastatic (stage IV) or locally advanced disease" The report then goes on to state "Therefore it appears that not all patients are required to have had	These statements should be supplemented with the correct figures from the CheckMate 032 CSR, within which it can clearly be found that 75/78 patients (96.2%) had received prior platinum-containing chemotherapy. Only 3 patients had not received prior platinum-containing chemotherapy in the CheckMate 032 trial and therefore the trial is in accordance with the population defined in the scope.	Only 3 patients had not received prior platinum-containing chemotherapy in the CheckMate 032 trial. The figure of 60.2% quoted within the ERG report relates to the number of patients who had received prior platinum-containing chemotherapy in the adjuvant and/or neo-adjuvant therapy settings. Table 6 in the manufacturer submission contains a typographical error; where N/A is reported for the percentage of patients who had received prior platinum-based chemotherapy in the metastatic disease stage, this value should instead be 66/78 (84.6%).	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."

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 is not in accordance with the population defined in the scope.		Nevertheless, it can be clearly found in the CSR for CheckMate 032 that only 3 patients had not received prior platinum-containing chemotherapy. As such, this trial population is in alignment with the population defined in the NICE scope and stating otherwise is considered inaccurate.	

Issue 4

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
At multiple points throughout the report, it is stated that in the CheckMate 032 trial, 23% of patients switched to ipilimumab.	These statements are incorrect and should instead read that 23% of patients switched to receive a combination on nivolumab plus ipilimumab as part of the CheckMate 032 trial protocol.	Accurate reporting of trial information.	Text changed.

Issue 5

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
All of the available data for the current comparators provide evidence to demonstrate that OS for patients with locally advanced unresectable or metastatic urothelial carcinoma is	The conclusions of the ERG should be amended to reflect the data presented to support nivolumab meeting the end of life criteria.	The conclusions of the ERG regarding the end of life criteria are not in alignment with the data presented as part of the submission.	Not a factual inaccuracy.

substantially lower than 24 months. The extension of life with nivolumab of more than 3 months versus the relevant comparators was also demonstrated within the economic model. It is therefore inaccurate to state that there is no robust evidence to argue that nivolumab meets the end of life criteria as part of this appraisal.			

Issue 6

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
On page 13, the report states "Data from the individual trials indicated that for CheckMate 275 (n=275) nivolumab led to…" The number of patients in CheckMate 275 is also reported as being 275 on page 16.	The n number reported here is incorrect and should be n=270.	Accurate reporting of trial information.	Text changed.

Issue 7

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
On page 13, the report states "In CheckMate 032 (n=78) nivolumab led to a confirmed ORR (BIRC) in 19 (24.4%) patients (95% CI: 15.3–35.4)."	ORR in CheckMate 032 was not assessed via BIRC. Instead this statement should read: "In CheckMate 032 (n=78) nivolumab led to a confirmed ORR (investigator assessed) in 19 (24.4%) patients (95% CI: 15.3–35.4)."	Accurate reporting of trial information.	Text changed.

Issue 8

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
On page 14, the report states that ORR data for BSC was not identified. This is incorrect.	ORR data for BSC was identified in one trial (n=85) where 0 patients (0%) achieved an objective response.	As reported in Table 26 of the appendices of the manufacturer submission, ORR data for BSC was identified in the trial by Bellmunt et al. 2009, and this data was used within the STC for ORR. This statement is therefore inaccurate.	Text changed.

Issue 9

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
On page 14, the report states median PFS data for docetaxel and then states that PFS data from other comparators were not available. This is incorrect, as PFS data were available for paclitaxel from the trial by Jones et al. 2017.	Paclitaxel (one trial, n=65) had a median PFS of 4.1 months (80% CI [3 to 5.6]).	As reported in Table 25 of the appendices of the manufacturer submission, PFS data for paclitaxel were available from the trial by Jones et al. 2017, and these data were used within the STC for PFS. This statement is therefore inaccurate.	Text changed.

Issue 10

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
The statement on page 14, relating to the trial of gemcitabine and cisplatin reports “gemcitabine and cisplatin (one trial, n=65), had	This statement should read “gemcitabine and cisplatin (one trial, n=33), had a mean OS of 10.5 months.” The number of patients in the trial (Gondo *et al*.	Accurate reporting of trial information.	Text changed.

a median OS of 10.5 months." This is also misreported on page 38.	2011) was n=33, and the OS data reported was mean OS not median OS.		

Issue 11

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
The ERG highlight the following as transparency issues in the submitted model: "hidden columns on several sheets, the practice of not naming cells, the practice of disabling headings for columns and rows and the missing macro for generating the CEAC" (page 138)	This statement should be removed.	This statement is inaccurate as it implies that the submitted model is designed in an opaque way. This is not true – each of the points relate to formatting and are based on the preferences of the model developers. It is also very simple to make the adjustments the ERG required and therefore this cannot reasonably be referred to as a transparency issue.	This is not a factual inaccuracy. For example, not naming cells makes it more difficult to validate the model, and cannot be easily changed.

Issue 12

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
The ERG state that the company's methods for adjusting for background mortality is inappropriate without adequate justification as to why this is the case (page 114).	This statement should be removed.	It is incorrect to state that the 'population mortality adjustment' should be made to both responders and non-responders before combining the curves. This is because the general population mortality is an average, and it is perfectly plausible that some sub-populations (for instance, those	The company is correct in pointing this out. The ERG considered this approach as inappropriate in a case where a maximum function was used for adjusting OS and PFS for background mortality. However, since the company multiplied background mortality

		most likely to respond to immunotherapy, would have survival greater than the population average). The adjustment should, and was, applied after the combining of the curves to ensure that the average mortality of the whole cohort could not be lower than that of the general population.	with OS and PFS, the approach was appropriate. The following statements have been deleted: "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."

Issue 13

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Data reported in table 4.15, 4.16 and 4.17 for the CheckMate 275 trial is incorrect as the data included in the STC was from the most recent data (n=270) not the primary datacut (n=265).	To update the population assess, outcome definition and results for OS, PFS and ORR to the most recent data cut.	Accurate reporting of information used in simulation treatment comparison.	Text in tables changed.

Issue 14

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
The ERG claim that the company "did not provide sufficient evidence to support the violation of the proportional hazard assumption and to support the need for time-dependent HRs" (p119).	This statement should be removed.	The ERG state that the proportionality of hazards could not be ruled out based on the company's analyses because both CheckMate 032 and CheckMate 275 trials were presented separately in these plots, while the HRs were derived based on the pooled CheckMate 032 and CheckMate 275 trials dataset ". However, as the log hazard plots from both the CheckMate -032 and -275 trials show that the proportional hazards assumption does not hold this should provide amble evidence that it does not hold for the pooled data. The company cannot see how it is feasible that pooling this data would lead to the proportional hazard then holding. Further, in the clarification response the company noted that the observation of the violation of the proportion hazard assumption was supported by clinicians (via the clinical advisory board). No mention of this was made by the ERG, which is misleading. Overall, the statement made by the ERG that not enough evidence was provided to prove the proportional hazards	This is not a factual inaccuracy. Not all implications of the use of time dependent hazard ratios were discussed. Log cumulative hazard plots should be provided for the relevant data used in the analysis.

		assumption is violated is inaccurate.	

Issue 15

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
On page 132 paragraph 3 the report states "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) × *36.6% (the proportion of dose delays that exceeded 14 days*; averaged for CheckMate 275 and CheckMate 032) = 2.4% (i.e. dose intensity of 97.6%)." The figure of 36.6% is incorrect for the proportion of dose delays exceeding 14 days. This figure is the proportion of doses delayed <8 days, assuming a non-weighted average across both trials.	It is not clear what the correct figure should be, given the use of both delays exceeding 7 days and delays exceeding 14 days as criteria in the sentence. Additional reference to this issue indicates that >7 days is the ERGs proposed criteria for inclusion as a missed dose. If so the calculation is also not correct as that includes only the doses delayed <8 days as missed, which is in contradiction to the ERG proposal.	The data presented is not as described and the calculation does not reflect the description provided.	The company is correct. The statement has been changed to: "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%)." Furthermore, the ERG performed the analyses when this error was corrected (ERG base-case and exploratory analyses) and replaced all instances in the text and tables.

Issue 16

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 14, paragraph 4 "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." Also see page 21, paragraph 6 "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." The report states that details of pooling method were requested but not provided. This information was provided in response to question A12 of the ERG's clarification request. The pooled results referred to here were not requested in the clarification letter. [For clarification question A12	Suggest this sentence is deleted.	The statement in the report is factually inaccurate, and contradicts paragraph 5 of page 106 which states that information regarding the pooling was provided as requested.	This is not a factual inaccuracy. It is true that a response was given, as acknowledged in Section 5.2.6.1. However, the evidence synthesis used estimates of the hazards for nivolumab that were adjusted by using the STC, as discovered by the ERG on examining the Winbugs code. The ERG were also able to back-calculate to discover the unadjusted hazards, as described in Section 4.5 of the ERG report. Nowhere in the CS or the response to clarification were these unadjusted hazards, which must have been the result of some pooling, reported and therefore it remains unclear precisely how the two CheckMate trials were pooled in order to estimate them.

asked; On page 58 (company submission, section B.2.8) it is mentioned that data from the CheckMate studies were pooled. Please provide details of the statistical method(s) used for pooling the data from Checkmate 275 and 032 and please explain which data were used (BIRC or investigator-assessed). Please conduct all analyses using data from each method separately.			

Issue 17

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 42, paragraph 4, sentence 1 Report states that the company did not state how the two nivolumab trials were pooled. This information was provided in response to question A12 of the ERG's clarification request.	Suggest this sentence is deleted	The statement in the report is factually inaccurate, and contradicts paragraph 5 of page 106 which states that information regarding the pooling was provided as requested.	See response to Issue 16

Issue 18

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 68, paragraph 4, sentence 1 Report states that the company did not state how the two nivolumab trials were pooled. This information was provided in response to question A12 of the ERG's clarification request.	Suggest this sentence is deleted	The statement in the report is factually inaccurate, and contradicts paragraph 5 of page 106 which states that information regarding the pooling was provided as requested.	See response to Issue 16.

Issue 19

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 81, paragraph 7, sentence 1 Report states that the company did not provide the pooled results in response to the clarification letter. The pooled results referred to here were not requested.	Suggest this sentence is deleted.	The statement in the report is factually inaccurate as those pooled results were not requested in the clarification letter.	See response to Issue 16.

Issue 20

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
On page 95, point 2 of section 4.6:	Suggest removing statement on p153.	These statements are inconsistent with each other. If the use of	Not a factual inaccuracy. It is not inconsistent to identify both

"Many baseline characteristics were not available across all comparator trials and had to be imputed" is listed as a limitation of the analysis. On page 153, paragraph 3: "it would have been useful to see an STC that was based on prediction models with more covariates including all eleven considered"		imputation is a limitation of the analysis, it is not clear how using a model with all eleven covariates (and therefore more dependent on imputation) would help.	lack of information on all covariates and also lack of inclusion of such covariates as limitations. Inclusion of any set will incur the first problem and so the only way to avoid this is not to include any other than those that are fully informed, which is liable to lead to no regression analysis at all or misspecification. There is clearly therefore a trade-off, the effect of which can be tested by examining the effect of inclusion of different sets of covariates.

Issue 21

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
On page 14, the reporting of serious drug-related AEs from CheckMate 275 should be highlighted as commercial in confidence.	In the CheckMate 275 trial 64.4% of patients had a drug related AE ([redacted] serious drug related AE).	It is not anticipated that certain outcomes of the overall safety analysis in CheckMate 275 will be published. These unpublished data are commercially important to Bristol-Myers Squibb.	Changed.

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
On page 14, the reporting of	In the CheckMate 275 trial 64.4% of patients	It is not anticipated that certain	Changed.

serious drug-related AEs from CheckMate 275 should be highlighted as commercial in confidence.	had a drug related AE ([redacted] serious drug related AE).	outcomes of the overall safety analysis in CheckMate 275 will be published. These unpublished data are commercially important to Bristol-Myers Squibb.	

Issue 23

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
On page 21, the report reads "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 docetaxel, paclitaxel, cisplatin plus gemcitabine and BSC respectively." A similar error is repeated on page 132, paragraph 6, excluding the cisplatin plus gemcitabine comparison. This is incorrect.	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.	The wrong figures are quoted for the respective comparators.	Text changed.

Issue 24

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
On page 23 of the report the ERG	"However, the company did not provide these, stating that there was insufficient data to allow	The BMS response states that "it was determined that there was	This is not a factual inaccuracy. Insufficient data were indeed

states that: "However, the company did not provide these, stating that small numbers in responder and non-responder groups did not allow separate estimation of relative effectiveness." This is incorrect and does not reflect the response provided by BMS. Additionally this is misquoted on page 104, paragraph 2, sentence 5. "The company argued in their response to clarification questions that it was not possible to keep these two groups separate because the STC required a larger sample size to estimate HRs for responders and non-responders separately"	separate estimation of relative effectiveness for responder and non-responder groups." "The company argued in their response to clarification questions that it was not possible to keep these two groups separate because the STC required response-based comparator survival curves which were not available to estimate HRs for responders and non-responders separately"	insufficient data to allow separate responder and non-responder nivolumab patient groups to be compared with the comparators using the prediction models discussed previously."	cited as the reason for not estimating separate hazard ratios and the ERG's statements reflect this.

Issue 25

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
On page 24 the report the ERG states that: "external validation efforts that are based on a lung cancer study only and therefore questionable in terms of their relevance"	This statement should be removed.	Detailed validation of the nivolumab and comparator arms was provided on pages 125-127 of the submission. In addition to the validation of the nivolumab predicted outcomes	For clarity, the statement on page 24: "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 that are based

This view is repeated on page 146, paragraph 1. This is incorrect and does not reflect the company submission.		(using long-term nivolumab data from CheckMate-003), further validation not mentioned as part of this statement included - validation against published registry real world data on the use of paclitaxel (submission reference 112) - validation of outcomes for the comparator arm with UK experts (submission reference 67) which was subsequent and in addition to advisory board - validation of outcomes for the comparator arm against the clinical trials informing them	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." has been changed to: "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."

Issue 26

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
On the following pages the report states that the final parametric time-to-event models were not validated using expert opinion: Page 113 paragraph 3, sentence 4	These statements are should be removed.	Clinical expert validation elicited prior to the final analysis does not preclude its relevance or validation purposes given that the opinion elicited (i.e. regarding the shape and scale of the expected curve) remains the same (i.e. it's content	This is not a factual inaccuracy. The final parametric time-to-event models were not validated using expert opinion.

Page 139, paragraph 2, sentence 4 Page 155, paragraph 4, sentence 5 This is incorrect and does not fairly reflect the facts.		does not change) even after additional analysis has been conducted, and this output of this analysis can still be compared to the feedback gained.	

Issues 27

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
On page 131, paragraph 3 “an average weight of 80.405 kg based on both CheckMate 275 and CheckMate 032 was used in the ERG analyses” This description is misleading as does not present full information for reader interpretation.	“an unadjusted average of 80.405 kg based on both CheckMate 275 and CheckMate 032 (i.e. weighted equally and not based on trial size) was used in the ERG analyses”	Given the importance of this change, it should accurately reflect the analysis undertaken.	This is not a factual inaccuracy.

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 Error! No text of specified style in document.**.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 Error! No text of specified style in document.**.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 Error! No text of specified style in document.**.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 Error! No text of specified style in document.**.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 Error! No text of specified style in document.**.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

	Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	Nivolumab ICER (£/QALY)
	BSC	£8,838	0.60	[redacted]	[redacted]	£44,600
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

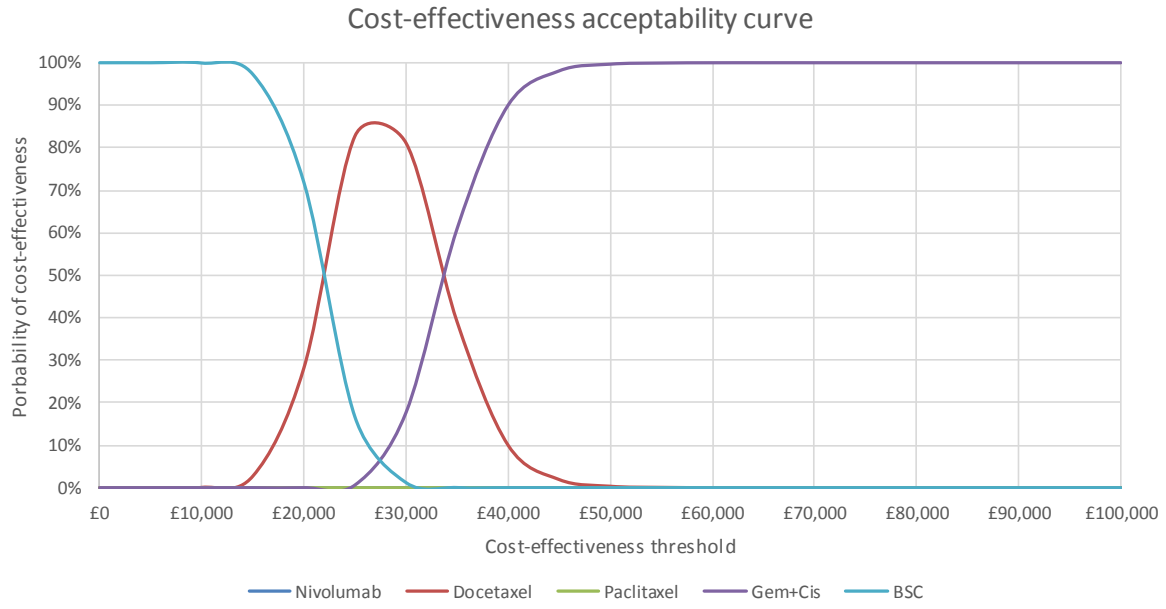
gemcitabine dominated nivolumab. The individual effects of each change on costs, QALYs and ICERs are presented in Section 6, Table 6.1. For comparison, the deterministic ERG base-case ICERs were £82,028, £64,298 and £66,161 per QALY gained, with cisplatin plus gemcitabine dominating nivolumab.

Table Error! No text of specified style in document.**.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 Error! No text of specified style in document.**.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 Error! No text of specified style in document..**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.