REVIEW ARTICLE

CrossMark

Nivolumab for Treating Metastatic or Unresectable Urothelial Cancer: An Evidence Review Group Perspective of a NICE Single Technology Appraisal

Sabine E. Grimm[1] · Nigel Armstrong[2] · Bram L. T. Ramaekers[1] · Xavier Pouwels[1] · Shona Lang[2] · Svenja Petersohn[1] · Rob Riemsma[2] · Gillian Worthy[2] · Lisa Stirk[2] · Janine Ross[2] · Jos Kleijnen[2] · Manuela A. Joore[1]

Abstract

As part of its single technology appraisal (STA) process, the National Institute for Health and Care Excellence (NICE) invited the manufacturer (Bristol-Myers Squibb) of nivolumab (Opdivo®) to submit evidence of its clinical and cost effectiveness for metastatic or unresectable urothelial cancer. Kleijnen Systematic Reviews Ltd, in collaboration with Maastricht University Medical Centre+, was commissioned to act as the independent Evidence Review Group (ERG), which produced a detailed review of the evidence for the clinical and cost effectiveness of the technology, based on the company's submission to NICE. Nivolumab was compared with docetaxel, paclitaxel, best supportive care and retreatment with platinum-based chemotherapy (cisplatin plus gemcitabine, but only for patients whose disease has had an adequate response in first-line treatment). Two ongoing, phase I/II, single-arm studies for nivolumab were identified, but no studies directly compared nivolumab with any specified comparator. Evidence from directly examining the single arms of the trial data indicated little difference between the outcomes measured from the nivolumab and comparator studies. A simulated treatment comparison (STC) analysis was used in an attempt to reduce the bias induced by naïve comparison, but there was no clear evidence that risk of bias was reduced. Multiple limitations in the STC were identified and remained. The effect of an analysis based on different combinations of covariates in the prediction model remains unknown. The ERG's concerns regarding the economic analysis included the use of a non-established response-based survival analysis method, which introduced additional uncertainty. The use of time-dependent hazard ratios produced overfitting and was not represented in the probabilistic sensitivity analysis. The use of a treatment stopping rule to cap treatment cost left treatment effectiveness unaltered. A relevant comparator was excluded from the base-case analysis. The revised ERG deterministic base-case incremental cost-effectiveness ratios based on the company's Appraisal Consultation Document response were £58,791, £78,869 and £62,352 per quality-adjusted life-year gained versus paclitaxel, docetaxel and best supportive care, respectively. Nivolumab was dominated by cisplatin plus gemcitabine in the ERG base case. Substantial uncertainties about the relative treatment effectiveness comparing nivolumab against all comparators remained. NICE did not recommend nivolumab, within its marketing authorisation, as an option for treating locally advanced, unresectable or metastatic urothelial carcinoma in adults who have had platinum-containing therapy, and considered that nivolumab was not suitable for use within the Cancer Drugs Fund.

✉ Sabine E. Grimm
   Sabine.grimm@mumc.nl

1 Department of Clinical Epidemiology and Medical Technology Assessment, School for Public Health and Primary Care (CAPHRI), Maastricht University Medical Centre+, P. Debyelaan 25, PO Box 5800, 6202 AZ Maastricht, The Netherlands

2 Kleijnen Systematic Reviews Ltd, 6 Escrick Business Park, Riccall Road, York YO19 6FD, UK

△ Adis

Key Points for Decision Makers

In the absence of comparative evidence from randomised controlled trials (RCTs) comparing nivolumab and its comparators, a simulated treatment comparison (STC) can potentially mitigate the bias introduced by comparing observational data from single-arm studies. However, given there were no comparative data (unanchored analysis), the results obtained from the STC should be treated with caution because unless all baseline characteristics that might be prognostic variables and effect modifiers are incorporated, it is unclear what the size of any bias might be.

A fractional polynomial model provides great flexibility when used in a network meta-analysis of survival estimates. However, different specifications could lead to vastly different model outcomes, and time-dependent hazard ratios may produce overfitting. Together with the use of a non-established response-based survival analysis method, which was inappropriately justified and that introduced assumptions and additional uncertainty, model outcomes may therefore suffer from substantial methodological and structural uncertainty.

This submission epitomises the increasing uncertainty in NICE technology assessments caused by an evidence base that is immature, often based on single-arm studies, and with a complex statistical analysis that attempts to adjust for bias, but is impossible to validate. In this assessment, this increased uncertainty is not reflected in the analyses and the perceived risk lead to a 'no'. Appropriate risk assessment is needed as this might have improved transparency of, and subsequently aided consistency in, decision making.

NICE issued guidance that did not recommend nivolumab, within its marketing authorisation, as an option for treating locally advanced, unresectable or metastatic urothelial carcinoma in adults who have had platinum-containing therapy. NICE issued an update to this in which it stated that neither data collection from clinical practice nor the ongoing single-arm trials would resolve the identified uncertainty, and that nivolumab for this patient group was not suitable for use within the Cancer Drugs Fund.

1 Introduction

The National Institute for Health and Care Excellence (NICE) is an independent organisation responsible for providing national guidance on promoting good health and preventing and treating ill health in priority areas with significant impact.

Health technologies must be shown to be clinically effective and to represent a cost-effective use of National Health Service (NHS) resources in order for NICE to recommend their use within the NHS in England. The NICE Single Technology Appraisal (STA) process usually covers new single health technologies within a single indication, soon after their UK market authorisation [1]. Within the STA process, the company provides NICE with a written submission, alongside a mathematical model that summarises the company's estimates of the clinical and cost effectiveness of the technology. This submission is reviewed by an external organisation independent of NICE [the Evidence Review Group (ERG)], which produces a report. After consideration of the company's submission (CS), the ERG report, and testimony from experts and other stakeholders, the NICE Appraisal Committee (AC) usually formulates preliminary guidance, the Appraisal Consultation Document (ACD), which indicates the initial decision of the AC regarding the recommendation (or not) of the technology. Stakeholders are then invited to comment on both the submitted evidence and the ACD, after which a further ACD may be produced or a Final Appraisal Determination (FAD) issued, which is open to appeal.

This paper presents a summary of the ERG report [2] for the STA of nivolumab, a human, monoclonal immunoglobulin (Ig) G4 antibody that acts as a PD-1 inhibitor, blocking the interaction of PD-1 with PD-L1 and PD-L2, for treating metastatic or unresectable urothelial cancer (UC); and a summary of the subsequent development of the NICE guidance for the use of this technology in England. Full details of all relevant appraisal documents (including the appraisal scope, CS, ERG report, consultee submissions, ACD, FAD and comments from consultees) can be found on the NICE website [3].

2 The Decision Problem

UC 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 [4]. It accounts for approximately 90% of all bladder cancers [5], and has a considerable impact on urinary, bowel and sexual functions, therefore impacting 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 the late stages of 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 [6].

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) [7, 8]. One previously published technology appraisal in locally advanced unresectable or metastatic UC related to 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 [9]. For patients with locally advanced or metastatic bladder cancer who are physically fit and have adequate renal function, the current standard of care in the first-line setting is platinum-based chemotherapy, namely cisplatin plus gemcitabine (cis + gem). For patients who progress on, or after first-line platinum-based chemotherapy, effective and tolerated treatment options in the second-line setting are limited [4].

The population, according to the final scope issued by NICE, was defined as "adults with metastatic or unresectable UC whose disease has progressed after platinum-based chemotherapy". The definition of the population in the CS [4] was in agreement with the definition in the final scope. The definition of the intervention in the CS was also in line with the definition in the final scope and according to the licensed dose and schedule of nivolumab monotherapy in UC, which, according to the summary of product characteristics (SmPC) by the European Medicines Agency (EMA), is 3 mg/kg administered as an intravenous infusion over 60 min every 2 weeks, and consistent with the existing approved dose and schedule of nivolumab monotherapy in adults, in other indications [10].

The scope issued by NICE lists four comparators: paclitaxel, docetaxel and best supportive care (BSC), as well as retreatment with first-line platinum-based chemotherapy (only for patients whose disease has had an adequate response in first-line treatment). The CS deviated from the scope in that retreatment with first-line platinum-based chemotherapy (cis + gem) was only considered in a scenario analysis.

3 Independent Evidence Review Group (ERG) Review

In accordance with the process for STAs, the ERG and NICE had the opportunity to seek clarification on specific points in the CS [11], in response to which the company provided additional information [12]. The ERG also modified the company's decision analytic model to produce an ERG base case and to assess the impact of alternative parameter values and assumptions on model results. Sections 3.1–3.4 below

summarise this evidence, as well as the ERG's review of that evidence.

3.1 Clinical-Effectiveness Evidence Submitted by the Company

A systematic literature review was carried out to identify studies providing direct and indirect clinical evidence on the use of nivolumab in metastatic or unresectable UC. The company did not identify any randomised controlled trials (RCTs) for nivolumab, but 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 [13, 14].

Data from the individual trials indicated that for Check Mate 275 ($n = 270$), nivolumab led to a confirmed objective response rate (ORR) in 54 (20.0%) patients (95% credible interval [CI] 15.4–25.3) [14], while in CheckMate 032 ($n = 78$), nivolumab led to a confirmed ORR in 19 (24.4%) patients (95% CI 15.3–35.4) [13]. For CheckMate 275, at the then latest database lock of 2 September 2016 ($n = 270$ analysed), nivolumab led to a median overall survival (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) [13, 14]. For CheckMate 275, at the then latest database lock of 2 September 2016 ($n = 270$ analysed), nivolumab led to a median progression-free survival (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) [13, 14].

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%] [14], and, for CheckMate 032 (March 2016 database lock), 76.9% of patients discontinued study treatment (disease progression, 64.1%; nivolumab toxicity, 2.6%) [13]. In the CheckMate 275 trial, 51.1% of patients died (1.1% attributed to nivolumab toxicity), while in the CheckMate 032 trial, 46.2% of patients died (2.6% attributed to nivolumab toxicity) [13, 14]. In the CheckMate 275 trial, 64.4% of patients had a drug-related AE, while in the CheckMate 032 trial, 83.3% of patients had a drug-related AE [13, 14].

The identification of two single-arm studies for nivolumab precluded any conventional mixed treatment comparison (MTC) or indirect meta-analysis via a common comparator. As a consequence, the company decided to perform an unanchored (no common comparator) simulated treatment comparison (STC). Data for the CheckMate trials were pooled for the STC but the pooled results or method were not provided. Single-arm data were provided as an alternative to the STC to allow naïve comparisons to the single-arm

data of nivolumab. More detail on the comparator data can be found in the committee papers [3].

The STC approach used nivolumab individual patient-level data (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 was intended to adjust the outcomes observed in the nivolumab study 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, as well as 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 was 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. The NICE Decision Support Unit (DSU) Technical Support Document (TSD) 18 [15] 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. This was provided by the company upon request.

An evidence synthesis model was used to synthesize the results of the STC, i.e. adjusted hazards (for OS and PFS) and odds (for ORR) from the nivolumab trials with those from the comparator trials. The output of this model was hazard ratios (HRs) for OS and PFS, and odds ratios for ORR. 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 4-week period. This model is known as fractional polynomial (FP) [16], and, through variation in a set of up to two key parameters, permits a wider variation in the form of the survival curves. The choice of FP model was reported to have been determined by best statistical fit.

The systematic review identified nine trials for inclusion in the STC. In addition to the two nivolumab studies, two comparator studies of paclitaxel, two of docetaxel, one of BSC, and two of cis + gem were identified [13, 14, 17–23]. Because not all studies reported all outcomes, only five studies were used for OS—one per comparator for all comparators except docetaxel, for which there were two [17, 18, 20, 21, 23]. The comparator studies were a mix of RCTs or single-arm studies. For PFS, only three studies were used—two for docetaxel and one for paclitaxel [18, 20, 23]. For ORR, 6 of 7 studies were synthesized, with only one paclitaxel study not being included [17, 20–24]. 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 found that for OS, nivolumab is superior to all comparators but only at certain time points; the CIs 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% CI 1.17–5.52, 68–72 weeks); docetaxel at time points between 20 and 72 weeks (HR 2.01, 95% CI 1.14–3.37, 68–72 weeks); and BSC at time points between 20 and 72 weeks (HR 1.86, 95% CI 1.17–2.85, 68–72 weeks). Nivolumab was superior to cis + gem above 20 weeks but never reached statistical significance.

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

The STC analysis of ORR using a fixed-effect model found that nivolumab was significantly better than BSC (OR 106.70, 95% CI 6.72–49,820) or docetaxel (OR 3.12, 95% CI 1.06–9.49), although the uncertainty was large. No significant differences were found for nivolumab compared with paclitaxel or gemcitabine plus cisplatin. In the random-effects model, nivolumab was only statistically superior to BSC (OR 108.1, 95% CI 4.17–52,240).

3.1.1 Critique of Clinical-Effectiveness Evidence and Interpretation

Searches were carried out in accordance with the NICE guide to the methods of technology appraisal, using a good range of databases. Additional searches of conference proceedings were reported, along with trial registers and 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 with its comparators. Failing this, a network meta-analysis (NMA) of RCTs using a set of common comparators would be the preferred approach. However, the submission relied on two single-arm studies of nivolumab, which were entered into an STC together with

the single arms of comparator studies. Single-arm studies are considered to be of low quality in estimating the relative effect of one treatment versus another. 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:

- There was no STC analysis for AEs or HRQoL; therefore, the value of any potential extension to life could not be judged in relation to any changes to the patients' quality of life.
- The analysis relied on two small single-arm nivolumab studies—one included 78 patients and the other included 270 patients; therefore, any statistical analyses had increased uncertainty due to the small sample size.
- The numbers of patients were small for all comparator studies (33–117) and not all studies provided data for all outcomes.
- There were no common comparators, therefore an unanchored STC had to be performed.
- The company pooled the two nivolumab trials despite each one using different methods of outcome assessment—CheckMate 275 used independent review (BIRC) assessment, and CheckMate 032 used investigator assessment. The results of this pooling (and its variability) were not reported.
- Ideally, the results of the STC would be based on 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.
- The major assumption for unanchored STC is that all effect modifiers or prognostic variables are accounted for. Not all key characteristics (possible effect modifiers or prognostic variables) for the STC were reported for all comparator trials, therefore imputations were required for these characteristics that were based on correlations to the baseline characteristics in the nivolumab trials.
- 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 was not clear that the most parsimonious model was 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 FP model 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, in a series of letters critiquing an unanchored comparison by Di Lorenzo et al. 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'" [15].

Additionally, although the STC adjustment made nivolumab appear more effective, it was unclear that the comparator trials were in patients with a poorer prognosis, irrespective of treatment, than those in the nivolumab trials. Indeed, unadjusted, there was little difference in survival, at least at the median, between nivolumab (8.74 and 9.72 months in CheckMate 275 and 032, respectively) and docetaxel or paclitaxel (9.2 or 8 months, respectively). The median survival for cis + gem was higher than the value reported for nivolumab at 10.5 months.

The ERG found that the FP model for synthesizing data for OS and PFS was 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, while 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. The company did provide the results for PH models in response to the clarification request, but the method used had questionable validity and was not the method recommended in the paper on which the FP approach was based [16]. The ERG was able to reproduce the base-case FP model 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, i.e. as if estimated without the STC for these base-case FP models. However, the uncertainty in these unadjusted HRs was not estimable without the original nivolumab IPD. Finally, the ERG found that the HRs estimated using a PH model according to Jansen [16] were different to those provided by the company by an amount that did not seem explicable by uncertainty.

In conclusion, it was difficult to be sure what the effectiveness of nivolumab was in comparison to the comparators in the scope. Evidence from directly examining the single arms of the trial data indicated little difference between the

outcomes measured from the nivolumab and comparator studies. Such a naïve comparison carries a high risk of bias. STC analysis was used to try to reduce this bias, but there was 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, failed to validate the results (if at all applicable, given the lack of data). 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.

3.2 Cost-Effectiveness Evidence Submitted by the Company

A de novo cohort partitioned survival model (cycle length, 4 weeks; time horizon, lifetime) was developed in Microsoft Excel (Microsoft Corporation, Redmond, WA, USA). The model consisted of three mutually exclusive health states: progression-free (PF) and post-progression (PP) disease states, and death (Fig. 1). Patients entered the model in the PF state and were treated with nivolumab or one of its comparators. Patients remained in the PF state until disease progression or death. The proportion of patients in each health state was determined by OS and PFS curves. The model included patients with metastatic or unresectable UC who have progressed following first-line platinum-based chemotherapy. Nivolumab was implemented as per its EMA SmPC posology and method of administration for second-line UC (i.e. 3 mg/kg every 2 weeks), and offered a confidential Patient Access Scheme (PAS). In the base case, the following comparators were considered: paclitaxel (80 mg/m^2 every 3 weeks of a 4-week cycle), docetaxel (75 mg/m^2 every 3 weeks) and BSC. A scenario analysis additionally considered cis + gem (cisplatin 35 mg/m^2 on days 1 and 2, plus gemcitabine 1000 mg/m^2 on days 1, 8, and 15) as a comparator. The model adopted the perspective of the NHS and Personal and Social Services (PSS) in England and Wales. All costs and utilities were discounted at a rate of 3.5% per year.

Treatment-effectiveness estimates were derived from the CheckMate 275 and 032 studies combined. Parametric time-to-event models were used to estimate OS, PFS and time to treatment discontinuation (TTD). A response-based approach was adopted to estimate OS and PFS for responders and non-responders separately, based on the proportion of responders observed at a fixed point in time in the CheckMate studies. This response-based approach was subsequently also enabled for TTD. The use of a response-based analysis was justified by standard survival modelling approaches not appropriately characterising the novel mechanism of action of nivolumab and changing hazards over time resulting from a mix of long-term responders and

non-responders. The response-based approach was implemented using landmark analysis to prevent the occurrence of immortal-time bias. Kaplan–Meier estimates for OS and PFS of responder and non-responder groups together were used until the specified landmark point (8 weeks), 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 (AIC and the Bayesian Information Criterion [BIC]) and visual inspection. The generalised gamma distribution was chosen to estimate OS and PFS of both responders and non-responders, and OS and PFS estimates for both groups were then combined by using a weighted average, with the weighting based on the proportion of responders in patients being progression-free and alive at the 8-week landmark point.

The relative effectiveness of nivolumab versus the comparators was modelled through time-varying HRs obtained mainly from the STC. The STC was performed based on the pooled CheckMate 032 and 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. To estimate PFS for BSC in the absence of data from the STC, it was assumed that the HR for BSC versus paclitaxel was equivalent to that of BSC versus vinflunine (constant HR over time) for second-line UC patients. This constant HR was then applied to the paclitaxel PFS curve. PFS estimates for cis + gem were derived by assuming equivalence of cis + gem PFS with that of paclitaxel.

TTD of nivolumab was estimated using the generalised gamma distribution, to ensure consistency with OS and PFS, and to avoid the long tails of the better fitting Gompertz and log-logistic distributions, which would result in some patients still receiving treatment after 5 and 10 years. TTD of the comparators was based on their respective PFS curves, assuming 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.

Resource use and unit costs data to inform the economic model were based on a number of sources, including the main studies, national databases and public sources, and were, if necessary, inflated to 2015/2016 costs. Prices for nivolumab with its PAS and all comparators were based on the British National Formulary (BNF) and the electronic market information tool (EMIT) [25, 26]. Further costs in the economic model included monitoring costs and BSC costs, as well as one-off event costs for AEs, subsequent treatment costs and terminal care costs.

None of the utility studies identified by the review were consistent with the NICE reference case, therefore EQ-5D-3L

Fig. 1 Model structure of partitioned survival analysis for nivolumab in metastatic or unresectable urothelial cancer
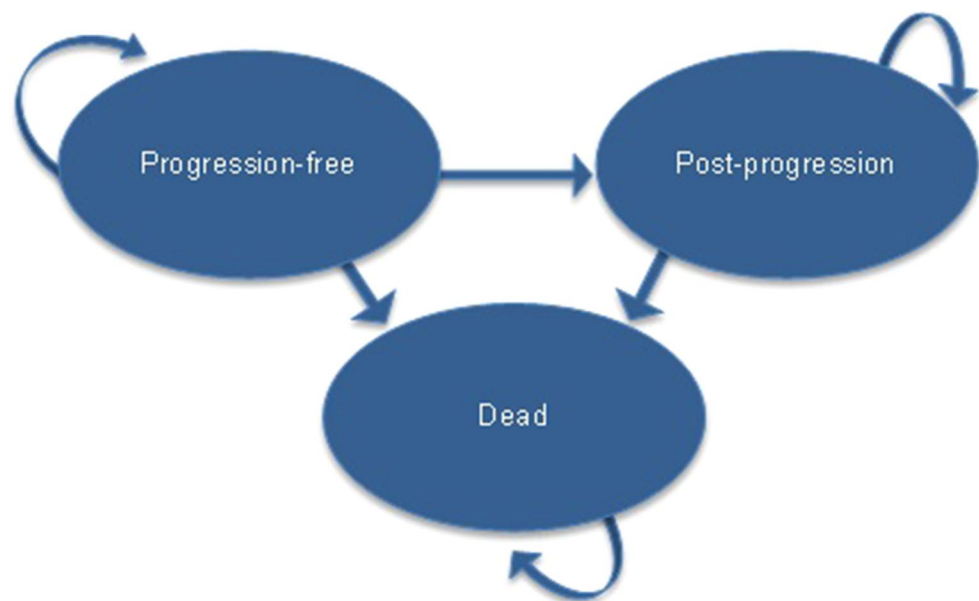
data valued with UK preference weights were taken from the CheckMate 275 trial [27]. These utility estimates were stratified according to PF and PP health states. Utility estimates were derived using a mixed-effects model to reflect within-subject variance, after interpolating for measurement times deviating 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- and post-progression, respectively. Disutilities were applied 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 AE and the duration of the AE.

In the company's deterministic base-case analysis, nivolumab was associated with larger quality-adjusted life-years (QALYs) and life-year (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 cis + gem, respectively. The ICERs resulting from the probabilistic sensitivity analysis (PSA) were substantially larger than the deterministic ICERs, driven by a reduction in PFS and OS in the PSA (compared with the deterministic analysis), but this was not explored further.

3.2.1 Critique of Cost-Effectiveness Evidence and Interpretation

The choice of partitioned survival analysis for this decision problem was in line with other appraisals in metastatic cancer, but it should be noted that the recent NICE DSU TSD 19 [28] advocates for alternative model structures that can more accurately reflect interdependent survival functions and use transition probabilities for each possible transition between health states when extrapolating beyond the trial data. The use of response-based analysis lacked sufficient justification, but, if considered appropriate, its implementation was considered by the ERG to be flawed as it should have been incorporated in the model via separate responder and non-responder health states. The ERG considered 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 032 studies, as well as the final scope issued by NICE for this appraisal. The decision to not provide the comparison of nivolumab with cis + gem in the base case was justified, citing expert opinion wherein the population in the only available cis + gem study differed from the UK population in that the study population did not receive cis + gem as a comparator. The ERG considered this argument 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 naïvety to gemcitabine might therefore be irrelevant.

The ERG considered the lack of clarity on how pooled estimates were obtained from the CheckMate 032 and 275 studies as one of the main issues; however, this issue was not clarified. Furthermore, the ERG was concerned about the appropriateness of the 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, as recommended in NICE DSU TSD 14 [29], were

considered inadequate to describe the mechanism of action of nivolumab in UC. 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, survival analysis alone, it was the ERG's view that the need for response-based analysis could not be established. The ERG considered that a standard approach should be shown inappropriate in the particular decision problem at hand before discarding it. However, if the need for alternative methods to conventional survival analysis could be justified, it was the ERG's view that the methods recommended in NICE DSU TSD 14 [29] (including spline models) should be considered before adopting a response-based landmark analysis. The ERG therefore considered that there was insufficient evidence to demonstrate that conventional parametric time-to-event models failed to describe nivolumab survival and to demonstrate that the landmark analysis provided superior results to standard survival modelling analyses or alternative methods recommended in TSD 14.

The use of response-based landmark analysis introduced further assumptions and additional uncertainty into the cost-effectiveness analysis.

- The need for selecting a cut point, here the 8-week landmark, with alternative choices causing unpredictable changes in cost effectiveness (only one alternative landmark was provided and further analyses declined), which was viewed as problematic [30].
- The use of Kaplan–Meier estimates for the period up to the landmark, instead of fitting a parametric curve until that time, which may result in overfitting.
- Fitting parametric models to the responder and non-responder groups separately resulted in larger uncertainty about these fitted curves due to significantly smaller sample sizes in the two groups.
- Inconsistency in using the response-based analysis for OS and PFS, but not for estimating TTD.
- Selection of a parametric model that did not make the best statistical fit (the Weibull model made a better fit in the non-responder group) to maintain the same parametric models for both responders and non-responders. 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.
- Combining responder and non-responder groups for the indirect comparison, which cast further doubt over whether the response-based analysis had any benefits, especially given that HRs were derived from the overall population and then applied in a combined responder and non-responder population.

- Response-based and conventional approaches resulted in vastly different estimates for predicted LYs when treated with nivolumab, with a predicted mean of 2.80 LYs in the response-based analysis and 1.84 LYs 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 both 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 cost-effectiveness analysis model suffered from significant uncertainty. Uncertainty likely driving decision uncertainty the most was located in the (relative) treatment-effectiveness estimation. It stemmed partly from the availability of only single-arm studies that were compared via the STC. This uncertainty was not reflected in the PSA. The parameterisation of the FP model that informed the NMA was found to have a large impact on cost-effectiveness outcomes (scenario analyses around the parameters of the FP model alone resulted in, for instance, an incremental QALY range for nivolumab versus docetaxel of 0.18–0.82). Furthermore, the need for time-dependent HRs to model the relative effectiveness of nivolumab versus the comparators was not appropriately justified. The ERG considered that proportionality of hazards could not be ruled out. Time-independent HRs were provided by the company, but these could not be replicated by the ERG. The use of the time-independent HRs produced by the ERG in scenario analysis increased all cost-effectiveness estimates. The ERG noted that using time-independent HRs had the advantage of preventing overparameterisation, which might occur when estimating time-dependent HRs with the relatively limited amount of data submitted by the company.

The number of iterations (1000) used in the PSA was shown to not yield stable results; even 10,000 simulations did not achieve stability. The marked differences between the deterministic and probabilistic results in the company's base case were largely resolved by removing the response-based analysis. Important parameters regarding relative effectiveness were excluded from the PSA, but inappropriate parameters, such as patient characteristics (age, weight) and comparator treatment costs were included. The exclusion of HRs from the PSA was justified by potential counterintuitive results when sampling the time-dependent HRs in each period independently. However, this issue could have been circumvented, for example, by using the same random numbers for all time points in each PSA run. Because relative effectiveness estimates were expected to be the largest contributor to decision uncertainty, the PSA was deemed to be insufficient.

3.3 Addenda to the Original Company Submission (CS)

In response to the ACD [31], the company submitted an updated model that included updated OS and PFS estimates from CheckMate 032, but, due to time constraints, not from CheckMate 275, even though these were also available. Furthermore, the company implemented a 2-year treatment stopping rule. These updates reduced the originally submitted company ICERs.

After the FAD was issued in January 2018, the company submitted a proposal for recommendation for use in the Cancer Drugs Fund (CDF) [32], including an NHS price reduction and new analyses. These analyses included the ERG's model settings, but with a 2-year treatment stopping rule in place, justifying this by any CDF agreement including such a stopping rule. In scenarios, treatment waning effects by setting OS and PFS HRs to 1 at certain time points, and OS and PFS updates from both CheckMate 275 and 032, were explored.

3.3.1 Critique of Addenda to the Original CS

The analyses submitted by the company in response to the ACD were deemed by the ERG to suffer from potential selection bias as updated OS and PFS estimates from the CheckMate 032 study only, and not from the CheckMate 275 study, were used in the economic model. Furthermore, the modelling of a 2-year treatment stopping rule would reduce the treatment costs, while maintaining the effectiveness of continued treatment artificially. Although it might be biologically plausible for treatment effects to continue after stopping treatment, the exact continued effect was uncertain.

The analyses in the CDF proposal after the FAD could partly be replicated by the ERG but were considered by the ERG to introduce new uncertainty. The use of the data update from both the CheckMate 275 and 032 studies significantly underestimated the comparator OS predictions when compared with the comparator observed OS and PFS data, leaving profound doubts over whether these analyses were implemented correctly. These were therefore not considered further by the ERG. While the ERG acknowledged that the 2-year stopping rule would more appropriately reflect the cost of nivolumab during 2 years of CDF reimbursement, concerns about the overestimation of treatment effectiveness remained. The implementation of treatment waning effects were considered by the ERG to be flawed as these would alter the comparators' survival curves and not the nivolumab survival curves.

3.4 Additional Work Undertaken by the ERG

The ERG defined a new base case that included multiple adjustments to the original base case presented in the CS based on identified errors, violations and alternative judgements [33]. This included correcting two errors—one in the calculation of the background mortality rates, and one in the calculation of dose intensity. In terms of fixing violations, the ERG included cis + gem as a comparator in the base case, excluded AEs with an incidence of < 5% from the analysis, used utility and weight estimates pooled from both CheckMate 275 and 032 instead of using only CheckMate 275, and removed inappropriate parameters from the PSA. As matters of judgement, the ERG used a conventional, not response-based, survival analysis approach and changed the assumption that all delayed doses were missed doses, to doses delayed by 7 days or more were missed doses. The ERG base-case analysis resulted in ICERs versus paclitaxel, docetaxel and BSC that were substantially larger than those produced by the company, and nivolumab was dominated by cis + gem. The revised ERG deterministic base-case ICERs based on the company's ACD response (i.e. including the CheckMate 032 data update and as reported in the first FAD) were £58,791, £78,869 and £62,352 per QALY gained versus paclitaxel, docetaxel and BSC, respectively. Nivolumab remained dominated by cis + gem in the ERG base case.

Furthermore, the ERG performed multiple exploratory analyses, based on both the ERG base case and the ERG base-case assumptions, except for using the company's preferred response-based approach. These analyses included the use of alternative parametric time-to-event models, use of a naïve treatment comparison instead of the STC, use of time-independent HRs for OS and PFS derived by the ERG, and the use of an alternative landmark if the response-based approach was used. All these changes increased the ICERs substantially, with the notable exception of alternative time-to-event models when the conventional survival analysis approach was used (alternative models in the response-based approach also substantially increased the ICERs). Further scenarios aimed to determine the uncertainty introduced by alternative specifications for the FP model used in the NMA, and resulted in substantial variation in absolute costs and QALYs.

3.5 End-of-Life Criteria

The company argued that end-of-life criteria were fulfilled, stating that "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 that "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" [3]. The ERG noted that the evidence from the economic model was weak and that there was a lack of robust evidence supporting the end-of-life claim.

3.6 Conclusion of the ERG Report

In conclusion, given that the revised ERG base-case ICERs were estimated to be above £50,000 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 versus its comparators remained substantial. The uncertainty and the resulting risk to the decision maker and the health system associated with making a recommendation was not appropriately quantified.

4 Key Methodological Issues

When no studies provide a common comparator to support any indirect or mixed treatment comparison, there is increased uncertainty and risk of bias. In this case, the company attempted to reduce this risk of bias by performing an STC; however, the risk could not be eliminated and it remained unclear whether bias could be reduced at all, given that it was unknown whether all prognostic variables and possible effect modifiers were accounted for. The opacity in the selection of these, the combination with an FP model for the NMA of survival, and the lack of additional scenarios hampered the assessment of this particular STC.

Survival analysis comes with its particular challenges when there are groups of responders and non-responders and a potential for cure, including changing hazard patterns over time. Conventional parametric survival analysis may not be appropriate in these instances. However, these conventional methods should not be discarded without presenting the appropriate evidence that conventional parametric curves make a poor fit to the data, for instance using cumulative log hazard plots to illustrate the poor fit. Furthermore, established methods as described in the NICE DSU TSD 14 [29] should be explored before resorting to other, less-established methods such as the response-based landmark analysis used in this appraisal. This analysis introduced many new assumptions, its implementation was questionable, and it generated additional uncertainty about cost-effectiveness estimates.

Employing a treatment stopping rule and thereby assuming a shortened period of time in which treatment costs are incurred while maintaining continued treatment effectiveness causes bias in model outcomes when the evidence is from studies that did not have such a stopping rule in place. Any scenarios modelling these have the caveat of substantial uncertainty. If these are modelled, some assumptions need to be made with regard to continued treatment effect after treatment is stopped. If a treatment waning effect is implemented, this should be applied to the treatment in question, rather than by setting the HR to 1 after a defined period. If the latter approach is adopted, the comparator survival curves may be altered if the HR is applied to the treatment, causing significant bias in model outcomes, such as flawed estimates of predicted LYs. Instead, it could be assumed that the survival curve of the treatment in question reflects that of a selected comparator (e.g. BSC) at the chosen time point. To present the substantial uncertainty associated with such practice, different time points and shapes of survival curves should be explored.

5 National Institute for Health and Care Excellence Guidance

On 12 January 2018, NICE issued guidance that did not recommend nivolumab, within its marketing authorisation, as an option for treating locally advanced unresectable or metastatic urothelial carcinoma in adults who have had platinum-containing therapy. On 25 May 2018, NICE issued an update to this, in which it reiterated this decision and expanded on the CDF proposal, stating that neither data collection from clinical practice nor the ongoing trials would resolve the identified uncertainty, and that nivolumab was not suitable for use within the CDF for patients with unresectable or metastatic UC after platinum-containing therapy.

5.1 Consideration of Clinical- and Cost-Effectiveness Issues

This section summarises the key issues considered by the AC. The full list can be found in the FAD dated 25 May 2018 [34].

5.1.1 Comparators

NICE concluded that docetaxel, paclitaxel and BSC were clinically relevant comparators, but that retreatment with chemotherapy (i.e. cis + gem) was not because it was predominantly used before these other treatment options had become available.

5.1.2 Considerations of Clinical Effectiveness

The AC noted that the CheckMate studies provided efficacy estimates for nivolumab, but that no RCT evidence was available. The AC was concerned that, without a trial directly comparing nivolumab with other treatments, it was difficult to reliably assess the relative treatment benefit of nivolumab. Furthermore, the AC considered trial data to be immature and based on small numbers of patients, and was therefore associated with considerable uncertainty in the results.

The AC had concerns about the robustness of the unanchored STC, and considered that the results of this analysis should be treated with caution. When not all important prognostic factors are accounted for in an STC, bias is introduced. The AC considered it unlikely that all of the important prognostic factors had been accounted for, therefore affecting the robustness of the results. No external validity tests, such as the out-of-sample method, and no sensitivity analyses to test the effect of using alternative prognostic factors in the predictive model, were performed.

The AC furthermore concluded that the relative effectiveness estimates inferred from the NMA are counter to clinical expectations, with the relative effectiveness of nivolumab decreasing with time, and are associated with uncertainty, and that these limitations needed to be accounted for in its decision making. The AC considered as major issues the fact that the optimal parameterisation of the FP was unknown, with the ICERs being sensitive to alternative ways of parameterising the FP model, and that the network of evidence was sparse.

5.1.3 Considerations of Cost Effectiveness

The AC had a preference for using the standard parametric time-to-event survival analysis approach over the response-based approach, and also noted that it had not seen any firm evidence to show that the response-based model was an adequate method for modelling long-term outcomes. While the AC considered that the response-based approach could be explored for modelling survival, it also considered that it introduced unnecessary complexity into the modelling of survival. The committee concluded that more evidence would be needed to support its appropriateness in preference to established modelling methods. Further to this, the AC considered that the standard survival analysis approach resulted in more plausible model estimates of OS at 5 years than the response-based approach, with the response-based approach potentially overestimating OS of responders. The 5-year OS estimates based on the standard survival analysis were also more in line with clinical expert opinion.

The AC concluded that implementing a treatment stopping rule while assuming lifetime treatment benefit was inappropriate because it would assume that costs stopped, while treatment benefit of nivolumab was unchanged. The AC also noted that the company's scenarios in which the continued treatment effect ended after 3 or 5 years produced counterintuitive results, which were related to the company's implementation of the treatment waning effect.

The AC recognised that all ICERs produced from the analyses needed to be treated with caution, and concluded that the most plausible ICERs, based on the ERG's revised base case, were £58,791 and £78,869 per QALY gained versus paclitaxel and docetaxel, respectively. The AC expected probabilistic ICERs to be higher still.

6 Conclusions

This article describes the STA considering nivolumab for treating metastatic or unresectable UC for adults whose disease has progressed after platinum-based chemotherapy. Given the sparse evidence base and the multiple methodological concerns, the AC concluded that the ICERs of nivolumab versus its comparators were above £50,000 per QALY gained, and did not recommend nivolumab, within its marketing authorisation, as an option for treating locally advanced unresectable or metastatic urothelial carcinoma in adults who have had platinum-containing therapy.

This submission is an example of the increasing practice of submitting single-arm studies for marketing authorisation and reimbursement applications observed in cancer drugs over the past years [35, 36]. This practice increasingly necessitates the use of methods such as the present STC, which results in considerable uncertainty and poses the risk of introducing bias that cannot be assessed. The tension between obtaining early access to innovative treatments, and the challenges in assessing the value of these treatments, should be recognised and formalised using risk-assessment methods [37] that allow for better management of this risk.

Compliance with Ethical Standards

Conflicts of Interest Sabine E. Grimm, Nigel Armstrong, Bram L.T. Ramaekers, Xavier Pouwels, Shona Lang, Svenja Petersohn, Rob Riemsma, Gillian Worthy, Lisa Stirk, Janine Ross, Jos Kleijnen and Manuela A Joore have no conflicts of interest to declare.

References

1. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal 2013. London: NICE; 2013. http://publications.nice.org.uk/pmg9. Accessed 23 Aug 17.
2. Armstrong N, Grimm S, Ramaekers BLT, Pouwels X, Lang S, Fayter D, et al. Nivolumab for treating metastatic or unresectable urothelial cancer. York: Kleijnen Systematic Reviews Ltd; 2017.
3. National Institute for Health and Care Excellence. Nivolumab for treating metastatic or unresectable urothelial cancer after platinum-based chemotherapy [ID995]. https://www.nice.org.uk/guidance/indevelopment/gid-ta10163. Accessed 15 Jun 2018.
4. 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. https://www.nice.org.uk/guidance/ta530/documents/committee-papers. Accessed 10 Jul 2018.
5. Pasin E, Josephson DY, Mitra AP, Cote RJ, Stein JP. Superficial bladder cancer: an update on etiology, molecular development, classification, and natural history. Rev Urol. 2008;10(1):31–43.
6. Sharma S, Ksheersagar P, Sharma P. Diagnosis and treatment of bladder cancer. Am Fam Phys. 2009;80(7):717–23.
7. National Institute for Health and Care Excellence. Bladder cancer: diagnosis and management London: NICE; 2015. http://nice.org.uk/guidance/ng2.
8. European Association of Urology. Muscle-invasive and metastatic bladder cancer—guidelines. 2016. http://uroweb.org/guideline/bladder-cancer-muscle-invasive-and-metastatic/#1. Accessed 3 Feb 2017.
9. 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. https://www.nice.org.uk/guidance/ta272. Accessed 3 Feb 2017.
10. Bristol-Myers Squibb Pharmaceuticals Ltd. Nivolumab (OPDIVO) 10 mg/mL concentrate for solution for infusion. Summary of product characteristics, 2017. https://www.medicines.org.uk/emc/medicine/30476. Accessed 20 Jun 2017.
11. 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.
12. Bristol-Myers Squibb Pharmaceuticals 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. https://www.nice.org.uk/guidance/ta530/documents/committee-papers. Accessed 10 Jul 2018.
13. 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):1590–8.
14. 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):312–22.
15. 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.
16. Jansen JP. Network meta-analysis of survival data with fractional polynomials. BMC Med Res Methodol. 2011;11:61.
17. 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.
18. 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):1770–7.
19. 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.
20. 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.
21. 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.
22. 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.
23. 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.
24. 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.
25. Joint Formulary Committee. British National Formulary. London: BMJ Group and Pharmaceutical Press; 2017.
26. Department of Health. Drugs and pharmaceutical electronic market information (eMit). London: Department of Health; 2016.
27. Andrea N, Marc-Oliver G, Margitta R, Jose Angel Arranz A, Sergio B, Jens B, 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 Check-Mate 275. J Clin Oncol. 2017;35(15 Suppl):4526.

28. 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. http://scharr.dept.shef.ac.uk/nicedsu/wp-content/uploads/sites/7/2017/06/Partitioned-Survival-Analysis-final-report.pdf.

29. 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. http://scharr.dept.shef.ac.uk/nicedsu/wp-content/uploads/sites/7/2016/03/NICE-DSU-TSD-Survival-analysis.updated-March-2013.v2.pdf.

30. Latimer NR. Response to "Survival analysis and extrapolation modeling of time-to-event clinical trial data for economic evaluation: an alternative approach" by Bagust and Beale. Med Decis Making. 2014;34(3):279–82.

31. Bristol-Myers Squibb Pharmaceuticals Ltd. Nivolumab for treating adults with locally advanced unresectable or metastatic urothelial carcinoma after failure of platinum-based chemotherapy [ID995]: company response to Appraisal Consultation Document (ACD). Middlesex: Bristol-Myers Squibb; 2017.

32. Bristol-Myers Squibb Pharmaceuticals Ltd. BMS Proposal for Recommendation for use in the Cancer Drugs Fund for ID995: Nivolumab for treating metastatic or unresectable urothelial cancer after platinum-based chemotherapy: Bristol-Myers Squibb,

2018. 8p. https://www.nice.org.uk/guidance/ta530/documents/committee-papers-3. Accessed 10 July 2018.

33. 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.

34. National Institute for Health and Care Excellence. Nivolumab for treating metastatic or unresectable urothelial cancer after platinum-based chemotherapy. Final appraisal determination. London: NICE; 2018. https://www.nice.org.uk/guidance/gid-ta10163/documents/final-appraisal-determination-document-2. Accessed 15 Jun 18.

35. Davis C, Naci H, Gurpinar E, Poplavska E, Pinto A, Aggarwal A. Availability of evidence of benefits on overall survival and quality of life of cancer drugs approved by European Medicines Agency: retrospective cohort study of drug approvals 2009-13. BMJ. 2017;359:j4530.

36. Dickson R, Boland A, Duarte R, Kotas E, Woolacott N, Hodgson R, et al. EMA and NICE appraisal processes for cancer drugs: current status and uncertainties. Appl Health Econ Health Policy. 2018;16(4):429–32.

37. Grimm SE, Strong S, Brennan A, Wailoo AJ. The HTA risk analysis chart: visualising the need for and potential value of managed entry agreements in health technology assessment. Pharmacoeconomics. 2017;35(12):1287–96.