CONFIDENTIAL UNTIL PUBLISHED

Evidence Review Group Report commissioned by the NIHR HTA Programme on behalf of NICE

Atezolizumab for treating locally advanced or metastatic urothelial carcinoma

ERRATUM to the Evidence Review Group Final Report

This document contains an erratum to the Evidence Review Group (ERG) report following the factual accuracy check by Roche

Produced by Southampton Health Technology Assessments Centre

Correspondence to Dr Geoff Frampton
Southampton Health Technology Assessments Centre
University of Southampton
First Floor, Epsilon House
Enterprise Road, Southampton Science park
Southampton SO16 7NS
www.southampton.ac.uk/shtac

Date completed 12th April 2017 (updated with company erratum 19th April 2017)

Changes made to the ERG report are indicated in *blue italicised text* for the following pages: 14, 16*, 27, 31, 51, 52, 55, 61*, 79, 87*, 91, 96, 97, 100*, 104, 109, 113, 115 (* indicates changes in response to a company erratum)

The company constructed two partitioned survival models in Microsoft Excel with identical model structure. The models compared first-line atezolizumab with gemcitabine + carboplatin; and second-line atezolizumab with docetaxel, paclitaxel and best supportive care. The models have a lifetime time horizon of 20 years, with discounting of 3.5% per annum for costs and health benefits, a weekly cycle length and a half-cycle correction. The perspective of the analysis is for the NHS and Personal Social Services. The models have three health states: 'progression-free survival', 'progressed disease' and 'death'.

The models use clinical trial data for atezolizumab from IMvigor 210, a single-arm phase II study. Clinical trial data for the comparators are derived from studies found through a systematic search of the clinical literature. The model uses parametric survival modelling to fit survival curves to the observed data for progression-free survival and overall survival for atezolizumab. *The company assumes that progression-free survival for atezolizumab is equivalent to that of the first-line comparators, and to the second-line comparators docetaxel and paclitaxel. Proportional hazards were assumed for comparisons against second-line best supportive care.* For the comparators' overall survival, the overall survival curves for atezolizumab are adjusted using the results of the company's fractional polynomial model. The model derives the proportion of patients in the progressed disease state as the difference between the progression-free survival and overall survival curves. The generalised gamma distribution was used for progression-free survival and overall survival for first-line and second-line comparisons.

Utility estimates were taken from the Australian Pharmaceutical Benefits Advisory Committee (PBAC) cost-utility analysis for vinflunine, in which quality of life values from the EORTC QLQ Q30 questionnaire for patients with advanced urothelial carcinoma who had received vinflunine were mapped to EQ-5D values. Atezolizumab is administered intravenously every three weeks and the recommended dose is 1200mg at a proposed list price of £3807.69 per dose. The cost of comparator treatments are taken from the pharmaceutical electronic market information tool (eMit) and their doses are as recommended by their Summaries of Product Characteristics. Health state costs are based on those used in the NICE technology appraisal for vinflunine (TA272).

The results of the economic model are presented as incremental cost effectiveness ratios (ICERs), measured as the incremental cost per quality-adjusted life-year (QALY). For the base case the incremental cost per QALY gained is £44,158 for first-line atezolizumab compared to gemcitabine + *carboplatin* (Table 1). The ICERs for second-line atezolizumab compared to

Commentary on the robustness of submitted evidence

Strengths

The company has conducted thorough searches and, despite some inconsistencies in application and reporting of the eligibility screening process appears to have identified all of the key studies on atezolizumab and the scoped comparators.

The model structure is representative of the clinical pathway for patients with advanced or metastatic urothelial carcinoma. The company conducted a systematic review to identify cost-effectiveness, HRQoL and cost studies and values from this review were utilised in the model. The models are intuitive and user-friendly.

Weaknesses and areas of uncertainty

Weaknesses

The ERG has the following concerns regarding the simulated treatment comparison:

- It is based on a very small set of covariates.
- Some aspects of the analysis are unclear, including how the company accounted for missing covariate values.
- The cumulative impact of small errors and inconsistencies in the data is unclear.

The ERG has the following concerns regarding the network meta-analysis:

- Hazard ratios for overall survival were not used to inform clinical effectiveness of atezolizumab and were considered to be clinically implausible when applied in the economic analysis without adjustment.
- Hazard ratios for progression-free survival were considered to be clinically implausible and were not used to inform the clinical effectiveness or cost-effectiveness evaluation of atezolizumab.

and health-related quality of life. *However, the CS systematic review of health-related quality of life did not identify any relevant data for this outcome.* The outcomes are appropriate and clinically meaningful to patients, and the ERG considers that all important outcomes, other than quality of life, have been included in the decision problem.

Economic analysis

The economic analysis described in the decision problem conforms with the NICE reference case and is appropriate for the NHS. The company conducted a cost-utility analysis with a 20-year time horizon, which is considered sufficiently long to reflect any differences in costs or outcomes. Costs are considered from the NHS and Personal Social Services perspective

Other relevant factors

The NICE scope does not specify any subgroups that should be considered, and in line with this none are considered in the company's cost-effectiveness analysis, although clinical effectiveness evidence is presented according to PD-L1 expression subgroups.

No issues related to equity or equality have been identified by the NICE scope, the company decision problem, or the ERG.

- to the “priority” comparators specified in CS section 4.1.4, and no reasons are given in CS Figure 3 as to why studies were excluded at these screening steps.

Eligible outcomes

The CS lists 12 eligible outcomes (CS Table 10), and these are reflective of the NICE scope and the company’s decision problem. However, the CS states that only four of these outcomes were considered for the network meta-analysis: overall survival, 12-month survival, progression-free survival and objective response rate (CS section 4.10.5). No reason is given in the CS for focusing on these outcomes, although the ERG agrees that overall survival and progression-free survival are important outcomes for the evaluation of urothelial cancer treatments.

Eligible study designs

Randomised controlled trials (RCTs), non-randomised trials, and single-arm studies were eligible, and this seems appropriate. Phase I studies were excluded.

Summary of the screening process

CS section 4.1.3.2 (Review strategy) briefly describes the eligibility screening process, and provides a PRISMA flow chart (CS Figure 3). In CS Figure 3 the *reason for exclusion of* publications is incomplete (373 of 631 recorded only). The company clarified that the remaining 258 records were excluded because no outcomes of interest were reported (clarification response A7).

The CS does not state how many reviewers conducted the eligibility screening process but the company confirmed (clarification question A6) that titles/abstracts and full texts were assessed by two reviewers. The CS does not report whether any types of bias may have arisen during the eligibility screening.

According to the CS, the literature was initially screened on titles and abstracts using the eligibility criteria listed in CS Table 10. The remaining publications and internet search results were then assessed based on the full-text versions, yielding a data set of n=233 publications for inclusion in a ‘qualitative synthesis’ to ascertain feasibility of a network meta-analysis.

3.1.5 Description and critique of the company's outcome selection

The NICE scoped outcomes of overall survival, progression-free survival, response rates and adverse effects of treatment were measured in IMvigor*210* and PCD4989g. The NICE scoped outcome of HRQoL was not reported in any of the primary studies making up the evidence base, although this is not made clear in the company's decision problem.

Efficacy results are presented in the CS for various data-cuts (which we have summarised in section 3.3). In the Imvigor 210 study, objective response rate was the primary outcome. This was assessed by an independent review facility (IRF) using the RECIST (Response Evaluation Criteria In Solid Tumours) v1.1 criteria which is a standard approach for determining tumour size.[55] In cohort 2 investigator-assessed modified RECIST immune response criteria were also used which quantify only the viable portions of the tumour (references are provided[56, 57]). The CS states that the modified criteria are not yet used in standard practice (CS section 4.13.2). In clarification response A35 the company stated that the rationale for using the modified RECIST criteria was to account for the possibility of 'pseudoprogression' (i.e. where tumour size reflects immune cell infiltration rather than active cancer), and the potential for delayed anti-tumour activity.

The ERG has focused on reporting outcomes for the most recent data-cut and, where reported, we present results obtained using both RECIST methods. We have focused on the assessments by the independent review facility because these should be at lower risk of bias than investigator assessments. However, the CS does not report whether the independent review facility was blinded to any aspects of the Imvigor 210 study design, and does not explain whether the independent review facility was related to an independent data monitoring committee which is described in CS section 4.11.6. The CS states that there was a high concordance rate between independent review facility and investigator assessments (94%; CS section 4.11.10.3), but does not report results from both assessment approaches for the latest data-cut (20-month follow-up).

Secondary outcomes were the duration of response and progression-free survival assessed using RECIST v1.1 criteria by the independent review facility and investigator; overall survival; and 1-year survival; and these are appropriate endpoints.

Safety outcomes reported in the CS include treatment-emergent adverse events (no definition is provided in the CS or the clinical study report), serious adverse events, and adverse events of special interest. Those of special interest were immune-mediated adverse events and renal function events which are anticipated effects of using a monoclonal antibody therapy. Another possible adverse event of special interest could be infusion related reactions. Rates of these are presented for both cohorts of the Imvigor 210 study, although the CS does not list them as specific events of special interest. Overall, the safety outcomes reported are those that the ERG would expect to be provided for a monoclonal antibody anticancer therapy.

In summary, the ERG considers that the selected outcomes are appropriate to the NICE scope, with the exception that no data on HRQoL were available.

3.1.6 Description and critique of the company's approach to trial statistics

The CS states that effectiveness analyses in IMVigor 210 were performed on the intention-to-treat (ITT) population. This is not defined in the CS but the company explained (clarification response A37) that it refers to enrolled patients who received any amount of study drug. The company also stated in the clarification response that an exception to this involves objective response rate analyses, which were performed on the objective response-evaluable population, defined as ITT patients who have measureable disease per RECIST v1.1 criteria at baseline. *According to footnotes for CS Tables 31 and 34, the response evaluable population was 99/119 patients in cohort 1 and all patients in cohort 2.*

The CS reports using a hierarchical fixed-sequence testing procedure to compare the primary endpoint, objective response rate, between atezolizumab and a historical response rate of 10%. Hypothesis testing was carried out on three pre-defined populations (based on decreasing proportion of PD-L1 expression) sequentially on the basis of independent review-assessed objective response rate according to RECIST v1.1 followed by investigator assessed objective response rate according to modified RECIST criteria. If no statistical significance was detected at a particular level in the hierarchy, no further hypothesis testing was done. The ERG agrees that this is an appropriate statistical approach and is consistent with statistical recommendations of the EMEA.[58]

the resulting 'predicted controlled trials' can then be incorporated into a network meta-analysis, with atezolizumab as the common link.

The company's approach to the STC prediction model is described briefly in CS section 4.10.8. The first step in the STC analysis approach is to identify the covariates (i.e. the prognostic factors and effect modifiers for survival) that will be used in the prediction model. We note that the assumption of an unanchored STC is that all effect modifiers and prognostic factors are accounted for, which is considered 'largely impossible' to meet, leading to an unknown amount of bias in the unanchored estimate.[60] It is important therefore that as many of the key covariates as possible can be identified and included in the analysis to reduce the bias.

STC prediction covariates

The CS specifies four covariates which they used in their prediction model: the proportions of patients who: were aged > 65 years; were male; had liver metastases; and had ECOG performance status ≥1 (equivalent to Karnofsky performance status ≤90%[61]) (CS Table 17). *Limited justification is given in the CS for these covariates being prognostic factors* or effect modifiers. The CS states that due to the limited amount of data available in metastatic urothelial cancer, studies were included when ≥1 out of the four predictors were reported, although included studies for comparators of interest all reported a minimum of three of the four factors (CS section 4.10.4).

The CS states (section 4.10.13) that where trials did not report baseline values for the covariates of interest, the missing values were imputed by generating*, at every bootstrap iteration,* random values from a uniform distribution, with boundaries defined by the range of reported values across the studies included in the analysis. As the company acknowledges in the CS (and also in clarification response A31) this approach has limitations. The ERG believes that this approach *may not have captured the full range of clinically plausible values and* a *more extensive* multiple imputation approach would have been more appropriate. Multiple imputation aims to allow for the uncertainty about the missing data by creating several different plausible imputed data sets and appropriately combining results obtained from each of them.[62] *The company does not report or discuss the distributions of the imputed covariates.*

In response to a clarification request from the ERG and NICE, the company explained that the age cut-off of ≥65 years was selected as this was considered a clinically important age cut-off, but they did not give any empirical evidence for this (clarification response A17). The company also provided a description of a targeted literature search, not reported in the CS, which they

explored in sensitivity analysis (CS section 4.10.11.12); however, no sensitivity analysis is reported (this, together with sensitivity analysis of the heterogeneity priors was subsequently provided by the company in clarification response A30).

Model selection for first-line comparisons

For first-line treatment comparisons of overall survival the company selected the *first-order Gompertz fractional polynomial model.* The CS states that second-order fractional polynomial models were not considered due to the limited evidence base. Given the fit of the zero-order model it might be assumed that hazards were proportional in the comparison of atezolizumab to gemcitabine + carboplatin, although this is not stated in the CS. Visual inspection of overall survival curves (CS Figures 8 and 9) suggests that hazards may not have been proportional (in one study the curves cross) but the CS does not comment on this. The network meta-analysis section of the CS does not provide any information about time-dependency of the hazard ratio. However, in reporting the economic analysis (CS section 5.3.6) the CS states that the hazard ratio increased linearly over time and required capping to avoid clinically implausible values (see section 4.3.5).

Model selection for second-line comparisons

For second-line treatment comparisons of overall survival the company selected the Gompertz (i.e. first-order) fractional polynomial model, as this had the lowest DIC among three fixed-effects models that were compared (CS Table 23). Second-order models were considered, and had lower DIC values indicating better fit, but the CS states these exhibited large posterior correlations (>0.9) indicative of over-fitting and so were not used. Posterior correlations were also relatively large (>0.8) for the selected Gompertz model but the CS does not discuss this. Hazard ratio time curves are presented for comparisons of atezoluzumab against best supportive care, paclitaxel and docetaxel (CS Figures 15-17) with the corresponding parameter estimates (CS Table 24), and these indicate that the hazard ratio for the atezolizumab-docetaxel comparison decreased with time. In reporting the economic analysis (CS section 5.3.6) the CS states that the hazard ratios for second-line comparisons increased linearly over time and required capping to avoid clinically implausible values (see section 4.3.5).

Overall survival on second-line paclitaxel, compared with second-line atezolizumab (one study), is shown in Figure 9.

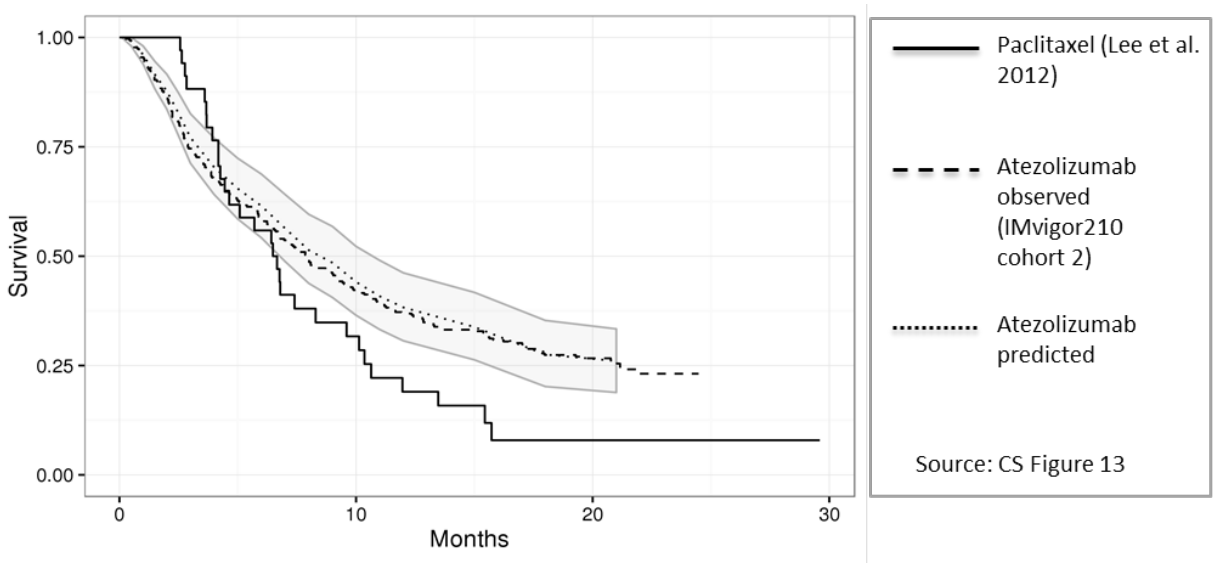

Figure 9 Overall survival curves for second-line paclitaxel (Lee et al. 2012) and atezolizumab

3.3.7 Adverse events

The CS presents safety endpoints from the two cohorts of the Imvigor 210 study and the PCD4989g study (minimal data) in CS section 4.12.3. We have summarised adverse event information from the PCD4989g study here, although the company stated that patients in PCD4989g received less than the licensed atezolizumab dose (see Appendix 2). No pooled adverse event data from the three sources of evidence are presented in the CS.

The rate of any adverse event was around 96-98% in the Imvigor 210 study (Table 20). Rates were generally similar across the two cohorts, where reported. *The most frequent treatment-related adverse events (affecting >10% of patients) were: Cohort 1: fatigue (30%), diarrhoea (12%) and pruritis (11%) (CS Table 43); Cohort 2: fatigue (30.6%), nausea (26.5%), pyrexia (22.3%), vomiting (19.4%), arthralgia (17.7%), pruritis (11.9%), rash (11.6%), decreased appetite (11.3%) and chills (10.6%) (CS Table 46)* Serious adverse events were experienced in 38% of patients in cohort 1 and 47% in cohort 2. The most commonly reported serious adverse events, reported in at least 2.5% of participants, were acute kidney injury, small intestinal obstruction, renal failure, sepsis and diarrhoea in cohort 1 (proportions are not reported in the CS). In cohort 2 the most commonly reported serious adverse events, reported in at least 3 participants, [redacted] (data from

The ERG has the following concerns regarding the network meta-analysis:

- Hazard ratios for overall survival were not used to inform clinical effectiveness of atezolizumab and were considered to be clinically implausible when applied in the economic analysis without adjustment.
- Hazard ratios for progression-free survival were considered to be clinically implausible and were not used to inform the clinical effectiveness or cost-effectiveness evaluation of atezolizumab.

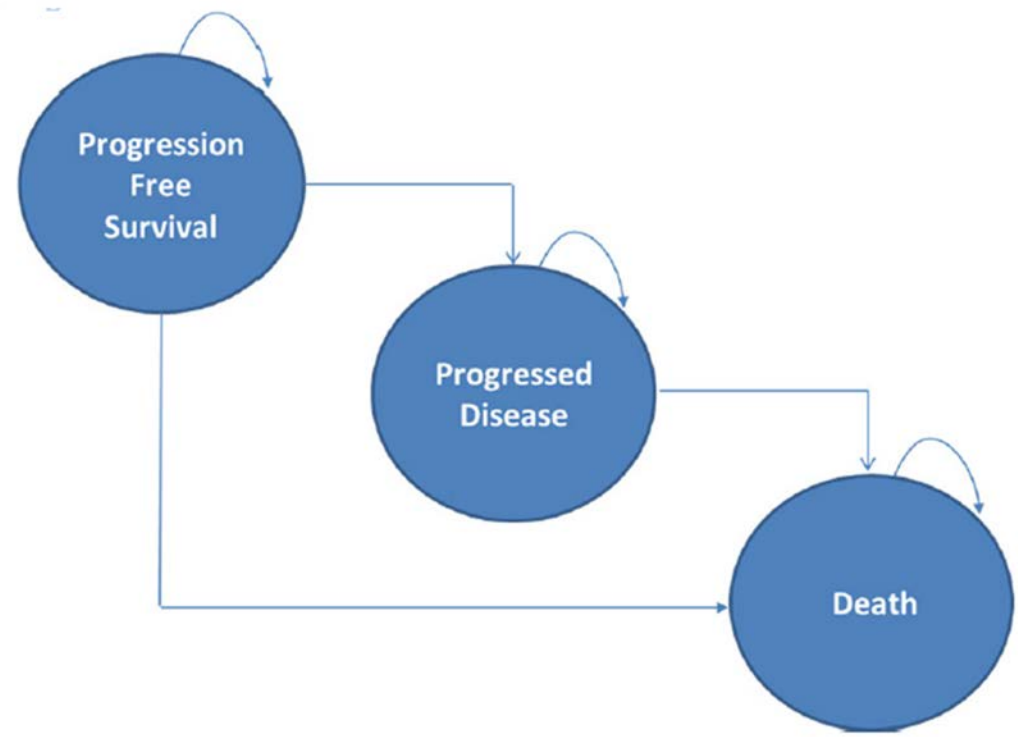

Figure 10 State model schematic (CS Figure 22)

Patients are treated with atezolizumab until *loss of clinical benefit or unmanageable toxicity*. Patients treated with the comparator treatment are treated for a specified number of treatment cycles, according to the marketing authorisation. On the basis of expert clinical advice, the company assumed that there are no subsequent lines of anti-cancer therapy for any treatment arm in either population following progression. The CS states that for second-line treatment this assumption was confirmed by the IMvigor 210 study where only 14.7% of patients receive subsequent treatment with gemcitabine with the majority only receiving palliative radiotherapy. For cisplatin-ineligible patients, the CS states that these might be expected to receive subsequent therapy, for example the NICE guidelines recommend either carboplatin + paclitaxel or gemcitabine + paclitaxel, but that incorporating these treatments is unlikely to have a significant effect on the incremental cost or effectiveness of second-line therapy. The ERG's clinical expert advisor agreed that it is reasonable to assume that most patients on second-line treatment would not receive subsequent anti-cancer therapy following disease progression.

The ERG considers the model structure to be an appropriate representation of the biological processes of advanced or metastatic urothelial cancer and appropriately represents the

ratios but argues that these approaches are not appropriate techniques to obtain progression-free survival for the comparator drugs. So, they applied an assumption that progression-free survival of gemcitabine + carboplatin is equivalent to that of atezolizumab. The CS does not justify this assumption but it mirrors an assumption that the company made for second-line comparisons (explained below) that progression-free survival curves for atezolizumab and the comparators are equivalent.

Second-line comparators: docetaxel, paclitaxel and best supportive care

For second-line comparisons, the progression-free survival of docetaxel and paclitaxel were assumed to be equivalent to that of atezolizumab. This assumption is based on *a phase III* clinical study KEYNOTE-045[69] which included two patient cohorts: i) those who were treatment naive and ineligible for cisplatin-based chemotherapy; and ii) those who had previously received platinum-based chemotherapy. Although these patient populations align with those in this appraisal, KEYNOTE-045 compared pembrolizumab to investigator's choice of a 'blended comparison' of docetaxel, paclitaxel or vinflunine for which the data indicated a '*non-significant HR of 0.98 for PFS*' for pembrolizumab compared to the blended comparator (CS section 5.3.4). As the hazard ratio was not statistically significant and almost equivalent to 1.0, the company assumed that the progression-free survival curves for the comparators are equivalent to that of atezolizumab.

For best supportive care, the company assumed a proportional hazards model with a hazard ratio of 1.12 (Crl 0.91 to 1.37) based on the fixed-effect zero fractional polynomial model used in the economic analysis.

For validation, the company compared the progression-free survival model results against the observed clinical data from IMvigor 210 (CS Table 75). The CS states that the economic model overestimates median progression-free survival compared to the observed data.

ERG comments on the methods for modelling progression-free survival

The ERG views the standard method adopted to extrapolate progression-free survival data for both the first-line and second-line atezolizumab arms in the IMvigor 210 trial, by fitting parametric distributions, to be appropriate. In both patient cohorts, the gamma distribution is used for data extrapolation which appears to provide a good fit to the progression-free survival data, based upon AIC and BIC values and visual inspection of the survival curves.

The economic models provide an option which enabled the ERG to run the analyses not assuming that atezolizumab is equivalent to its comparators. For this scenario, in first-line treatment comparisons, the model uses parametric curves fitted to the gemcitabine + carboplatin progression-free survival data whereas for the second line treatment comparisons, the relative effects of the comparator arms i.e. docetaxel, paclitaxel and best supportive care are derived from the fractional polynomial models. In both the cases, the impacts on base case ICERs are minimal (see Table 26).

Table 26 Comparison of the CS base case results with the ERG's assumption on progression-free survival

Comparator	**ICER (£/QALY)**	
First-line	**CS Base case**	**ERG scenario: PFS of atezolizumab ≠ PFS of GEM + CAR**
Gemcitabine + carboplatin	£44,158	£43,841
Second-line	**CS Base case**	**ERG scenario: The relative effects of the comparators are obtained from FP models**
Docetaxel	£131,579	£132,250
Paclitaxel	£104,850	£99,996
Best supportive care	£98,208	£98,273

CAR: carboplatin; GEM: gemcitabine; FP: fractional polynomial; PFS: progression-free survival
ICER: incremental cost-effectiveness ratio; QALY: quality-adjusted life year

The CS *justifies the use of the KEYNOTE-045 study to inform the progression-free survival parameter for the comparator arms, based on expert clinical advice.* It is unclear if this study was identified from a systematic search. Further, IMvigor 210 and KEYNOTE-045 consist of different interventions i.e. atezolizumab and pembrolizumab, respectively. To assume that progression-free survival curves of the comparators in the current appraisal are similar to that of atezolizumab based on *this study* implicitly indicates that progression-free survival of atezolizumab is similar to that of pembrolizumab. Whilst we acknowledge that atezolizumab and pembrolizumab belong to the same broad class of drugs, the CS does not provide any evidence that they will have similar effectiveness, and we note that they have different specific modes of action (atezolizumab is a PD-L1 inhibitor whilst pembrolizumab is a PD1 inhibitor). According to the ERG's clinical expert, there is insufficient information available on whether atezolizumab and pembrolizumab differ in effectiveness, but it would be reasonable to assume that they are similar.

used for the intervention arm in both first-line and second-line comparisons. The CS also does not present any sensitivity analyses varying the treatment effect of atezolizumab compared to the comparator arms. Further, the CS does not report any sensitivity analyses varying the contrast estimates used within the fractional polynomial models. To address these issues, we conducted a range of sensitivity analyses, details of which are described below in section 4.4.

Table 28 Hazard ratios used in the company's economic analyses

First-line	OS HR until 8 months	OS HR after 8 months
Atezolizumab vs gemcitabine + carboplatin	0.62 (Crl: 0.47, 0.82) *Results from the first-order FP model are used to estimate the HR at different time points until the time points correspond with the median follow up* (i.e. at 8 months) at which point the HR is capped.	0.54 The economic model uses the value of 1.84 (i.e. HR of gemcitabine + carboplatin vs atezolizumab). This value is used based on the assumption of proportional hazards.
Second-line	**OS HR until 21.16 months**	**OS HR at and after 21.16 months**
Docetaxel vs atezolizumab	Results from the first-order FP model are used to estimate the HR until the time points correspond with the median follow up (i.e. at 21.16 months) at which point the HR is capped.	2.12 (this value is based on the assumption of proportional hazards)
Paclitaxel vs atezolizumab	Same as above	1.49 (this value is based on the assumption of proportional hazards)
BSC vs atezolizumab	Same as above	1.66 (this value is based on the assumption of proportional hazards)

HR: Hazard Ratio; FP: fractional polynomial; OS: overall survival

The company's choice of parametric curves for overall survival is based upon the fit with survival data for atezolizumab, assessed using AIC and BIC values and visual inspection of the parametric curves. The ERG notes that other parametric curves may also provide a good fit with the observed trial data and that the model also provides the option to use the Kaplan-Meier data with a parametric distribution for the tail of the curve. We also note that the AIC and BIC values only provide information on the fit to the observed data and do not inform the choice of the extrapolation beyond the trial, which should be based upon clinical plausibility.

As shown in Figure 12, Figure 13, and Figure 14, the modelled overall survival curves for the second-line comparator arms are comparable with the survival curves reported by the studies of interest. To assess the most plausible distribution for extrapolating overall survival data, we compared different model fits for the atezolizumab arm and the best supportive care arm. The goodness of fit was measured through visual inspection. We chose best supportive care for this comparison due to the available evidence being based on a larger sample size and a longer follow up period (see Table 11) for this comparator among the three comparator arms (docetaxel, paclitaxel and best supportive care) used in the economic analyses. Based on our observation, we view that Kaplan-Meier data and a Weibull curve would provide the most appropriate fit for extrapolating long term survival data. Further details of this analysis and alternative plausible survival distributions are presented in section 4.4.

The ERG notes that the company is inconsistent in the time points used to cap the hazard ratio across the two patient cohorts. As previously mentioned, the first-line hazard ratio is capped at 8 months whereas for the second-line comparisons, the cut-off is 21.16 months. For both first-line and second-line hazard ratios the assumption of proportional hazards is applied after the capping time point. The ERG conducted exploratory analyses for both first-line and second-line comparisons in which we varied the time points at which the assumption of proportional hazards starts (see section 4.4). Secondly, the ERG has concerns about the company's approach to cap the hazard ratio. The CS states this was done to arrive at clinically plausible results. However, this raises questions about whether the results from the fractional polynomial models used in the network meta-analysis are appropriate to inform the economic analyses if it is necessary to cap them in order to provide plausible results. We have performed exploratory analyses to see the effect on overall results of varying the slope of the contrast estimates. This was done to avoid needing to cap the hazard ratios. Further details of the analyses are presented in section 4.4 below.

4.3.5.3 Time to treatment discontinuation

In the CS, TTD for first- and second-line atezolizumab is captured in the model through patients transitioning in the model. Data for TTD for atezolizumab was taken directly from the IMvigor 210 study for the trial period. Beyond this time-frame, the company extrapolated discontinuation data by adopting the standard technique of fitting parametric distributions to the TTD Kaplan-Meier curves. Goodness of fit to the data was assessed using AIC and BIC and graphical

In the company's model, utilities are imputed in a way that is slightly inconsistent with the CS text: as stated in the CS, for atezolizumab, the 'on-treatment' utility in the model is 0.75 and the 'off-treatment' utility is 0.71; however, the base-case utilities for comparators are both set at 0.75. We carried out a scenario analysis where both utilities for atezolizumab are set at 0.75, in line with the assumption that atezolizumab is better tolerated than the comparators (see section 4.4 for details). In the same analysis we set the 'on-treatment' utility of *comparators* to 0.71 and set the 'off-treatment' utility to 0.75 to reflect the disutilities commonly observed during treatment with chemotherapy.

4.3.7 Resource use and costs

The company conducted a systematic literature search for resource use among patients aged 18 years and above with advanced urothelial carcinoma, and their search strategy appears appropriate (section 3.1.1). The inclusion criteria specified that the outcomes of interest were direct costs, total cost, resource cost and cost drivers. The search was not restricted to studies conducted in the UK. The review identified 15 studies that met the broad search criteria of the CS. Twelve studies were further screened out and the rationale for their exclusion is stated in CS Appendix 8.11 (we note this is wrongly mentioned as Appendix 8.10 in the CS). The ERG agrees with company's rationale for excluding these studies. The three studies finally included were selected based on their relevance to the UK population. They are Seal et al. 2015;[73] Huillard et al. 2016;[74] and NICE 2013.[22]

Seal et al. 2015 estimated total all-cause costs attributable to medical services, inpatient visits and emergency department visits spanning a 6-month period pre- and post-metastatic cancer diagnosis. The setting of Seal et al. is in the US. Huillard et al. was a retrospective study that captured the proportion of patients admitted to an intensive care unit, and the utilisation of supportive care, among adults suffering from bladder cancer in their last month of life. The setting for Huillard et al. is France. The ERG notes that, although the CS states that these studies contain data of interest (See Table 64 of the CS and CS Section 5.5.1), they have not been incorporated into the model.

Resource use consists of the drug dose and its costs, administration costs per 21 day treatment cycle, adverse event management costs and weekly supportive care costs (health state costs). The CS makes the case that none of the studies identified in the company's search directly quantified costs and healthcare resource use for the population of interest from a UK NHS

referenced in the CS states a 2014 Department of Health cost of £354.72. The ERG notes that these errors have a negligible impact on the results of cost-effectiveness analysis. We also observed that references for certain adverse events (alanine aminotransferase increase, aspartate aminotransferase increase, blood bilirubin increase, diarrhoea, electrolyte abnormalities, hypophosphataemia and infection) are not included in the CS references. The ERG and NICE raised this issue with the company and the company provided the reference for these adverse events (clarification response *B4*).

The company's systematic review did not identify any relevant resource use data associated with health states in metastatic urothelial carcinoma. The CS states that resource use was elucidated through expert clinical advice, and deemed appropriate by the ERG and NICE appraisal committee on vinflunine.[22] The CS uses these same assumptions (summarised in Table B39 of the manufacturer submission for TA272, January 2013) in CS Table 69. We note that the health home visit cost is referenced as Curtis 2016 but that publication does not report this cost. The ERG and NICE queried this with the company and in response the company described the error as typographical (clarification response B1). The company stated that the correct reference for the health home visit cost is the manufacturer's submission for vinflunine. Health state costs are slightly higher in the CS and the company explained further in their clarification that they have been inflated to 2015/16 costs.

Resource utilisation for health states is estimated on a per cycle basis in the CS, calculated from separately stated unit costs and frequency of use per month. In the CS, the pre-progression state costs amounted to £111.85, while the post-progression costs amounted to £146.79. despite the paucity of data, the company's approach is consistent with the reference case. The CS reports one-way sensitivity analysis for monthly atezolizumab off-treatment supportive care costs, and comparator off-treatment supportive care costs, varying between a lower value of half the base case and an upper value increased by 50% of the base case value. The ERG notes that the values used in these sensitivity analyses are arbitrary but in the absence of relevant data they are reasonable to capture the high uncertainty surrounding the cost inputs.

Table 32 Comparison of best supportive care results for the current submission and a previous submission on vinflunine

Comparator	Costs, £	Life years	QALYs
BSC (from vinflunine appraisal)	£8642	0.63	0.234
BSC (from atezolizumab appraisal)	*£4836*	0.75	0.55

BSC: best supportive care; QALY: quality-adjusted life year

4.3.9 Cost effectiveness Results

Results from the economic model (section 5.7 of the CS) are presented as the incremental cost per QALY gained for first-line atezolizumab compared with gemcitabine + carboplatin and for second-line comparisons with docetaxel, paclitaxel and best supportive care.

For the first-line base case an incremental cost per QALY gained of £44,158 per QALY is reported (see Table 33) for atezolizumab compared to gemcitabine + paclitaxel. For the second-line base case, the ICERs for atezolizumab compared to docetaxel, paclitaxel and best supportive care are £131,579, £104,850, £98,208 per QALY gained respectively.

Table 33 First-line base case cost effectiveness results

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£) incremental (QALYs)
Atezolizumab	£77,211	3.74	2.69				
Gemcitabine + carboplatin	£18,106	1.84	1.35	£59,106	1.91	1.34	£44,158

ICER: incremental cost-effectiveness ratio; LYG: life years gained: QALYs: quality-adjusted life years

Table 34 Second-line base case cost effectiveness results

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£) incremental (QALYs)[a]
Atezolizumab	£71,868	1.69	1.23				
Docetaxel	£9,439	1.04	0.76	£62,430	0.65	0.47	£131,579
Paclitaxel	£16,606	0.96	0.71	£55,262	0.73	0.53	£104,850
BSC	£4,836	0.75	0.55	£67,032	0.94	0.68	£98,208

ICER: incremental cost-effectiveness ratio; LYG: life years gained: QALYs: quality-adjusted life years
[a] Pairwise comparison with atezolizumab.