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

Produced by Southampton Health Technology Assessments Centre

Authors Keith Cooper, Senior Research Fellow, SHTAC
Neelam Kalita, Research Fellow, SHTAC
Oluchukwu Onyimadu, Research Fellow, SHTAC
Jill Colquitt, Senior Researcher, Effective Evidence LLP
Emma Loveman, Senior Researcher, Effective Evidence LLP
Geoff Frampton, Senior Research Fellow, SHTAC

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

Date completed 22nd March 2017

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 16/56/09.

Declared competing interests of the authors

None

Acknowledgements

We would like to thank Professor James Catto (Professor in Urological Surgery, University of Sheffield and Honorary Consultant Urological Surgeon, Sheffield Teaching Hospitals) for providing clinical advice to the ERG and for commenting on a draft of the report. We also thank Karen Welch, Information Scientist, SHTAC, for appraising the company's search strategy and running a search update; and Dr Joanna Picot, Senior Research Fellow, SHTAC, for acting as internal editor for the draft report.

Rider on responsibility for report

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

This report should be referenced as follows:

Cooper K, Kalita N, Onyimadu O, Colquitt JL, Loveman E, Frampton GK. Atezolizumab for treating locally advanced or metastatic urothelial carcinoma: A Single Technology Appraisal. Southampton Health Technology Assessments Centre, 2017.

Contributions of authors

Keith Cooper (Senior Research Fellow) critically appraised the health economic systematic review, critically appraised the economic evaluation and drafted the report. Neelam Kalita (Research Fellow) critically appraised the health economic systematic review, critically

appraised the economic evaluation and drafted the report. Jill Colquitt critically appraised the clinical effectiveness systematic review and drafted the report. Emma Loveman critically appraised the clinical effectiveness systematic review and drafted the report. Olu Onyimadu (Research Fellow) critically appraised the health economic systematic review, critically appraised the economic evaluation and drafted the report. Geoff Frampton (Senior Research Fellow) critically appraised the clinical effectiveness systematic review, drafted the report, project managed the review and is the project guarantor.

Word count: 45,758

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

1L	First-line
2L	Second-line
AE	Adverse event
ALT	Alanine aminotransferase
AIC	Akaike information criterion
ASCO	American Society of Clinical Oncology
AST	Aspartate aminotransferase
BIC	Bayesian information criterion
BSC	Best supportive care
CAR	Carboplatin
CD	Could not be determined
CHMP	Committee for Medicinal Products for Human Use
CI	Confidence interval
CrI	Credible interval
CS	Company's submission
CSR	Clinical study report
DIC	Deviance information criterion
DOC	Docetaxel
DOR	Duration of response
EAU	European Association of Urology
ECOG	Eastern Cooperative Oncology Group
eMit	Pharmaceutical electronic market information tool
EORTC	European Organisation for Research and Treatment of Cancer
ERG	Evidence review group
FDA	Food and Drug Administration
GEM	Gemcitabine
GFR	Glomerular filtration rate
HR	Hazard ratio
HRQoL	Health-related quality of life
HTA	Health technology assessment
ICER	Incremental cost-effectiveness ratio
IRF	Independent review facility

K-M	Kaplan Meier
LYG	Life years gained
MAA	marketing authorisation application
MAIC	Matching-adjusted indirect comparison
M-CAVI	Methotrexate, carboplatin, vinblastine
MHRA	Medicines and Healthcare products Regulatory Agency
MRI	Magnetic resonance imaging
mUC	Metastatic urothelial carcinoma
MVAC	Methotrexate, vinblastine, doxorubicin and cisplatin
Nab-PTX	Nanoparticle albumin bound paclitaxel
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NIH	National Institute of Health
NIHR	National Institute of Health Research
NMA	Network meta-analysis
NR	Not reported
ORR	Objective response rate
OS	Overall survival
PBAC	Pharmaceutical Benefits Advisory Committee
PBO	Placebo
PD-L1	Programmed death-ligand 1
PFS	Progression-free survival
PPV	Personalised peptide vaccine
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PS	Performance status
PSS	Personal social services
PTX	Paclitaxel
QALY	Quality-adjusted life year
QLQ	Quality of Life Questionnaire
RCT	Randomised controlled trial
RECIST	Response evaluation criteria in solid tumors
SE	Standard error
SEER	Surveillance, epidemiology and end results program

SmPC	Summary of product characteristics
STC	Simulated treatment comparison
TTD	Time to discontinuation
TTP	Time to progression
UK	United kingdom
VFL	VInflunine

SUMMARY

Scope of the company submission

The company's submission (CS) generally reflects the scope of the appraisal issued by the National Institute for Health and Care Excellence (NICE). This is to appraise the clinical effectiveness and cost-effectiveness of atezolizumab (an intravenous immunotherapy) within its marketing authorisation for treating locally advanced or metastatic urothelial carcinoma in people whose disease has progressed after prior chemotherapy or for whom cisplatin-based chemotherapy is unsuitable. The comparators specified in the scope are:

- Cisplatin-ineligible people (first-line therapy): gemcitabine + carboplatin; or best supportive care.
- People whose disease has progressed after platinum-based therapy: re-treatment with first-line platinum therapy (adequate responders only); docetaxel; paclitaxel; or best supportive care.
- People who are cisplatin-ineligible and whose disease has progressed after platinum-based therapy: re-treatment with first-line platinum therapy (adequate responders only); docetaxel; paclitaxel; or best supportive care.

The company's decision problem differs from the NICE scope in three respects: best supportive care is not considered as a comparator for cisplatin-ineligible people receiving first-line therapy; a distinction is not made between cisplatin-eligible and cisplatin-ineligible people who have progressed after previous platinum-based therapy; and re-treatment with first-line platinum therapy is not considered in the second-line setting. Justifications for these differences are provided, mainly reflecting lack of available evidence.

The current submission is based on immature clinical effectiveness data (single-arm studies only) and lacks data on health-related quality of life. These data are expected to become available when phase III ongoing randomised controlled trials comparing atezolizumab against chemotherapy are completed in November 2017 (second-line setting) and June 2020 (first-line setting). For the present technology appraisal the company has requested that their submission is considered by NICE for the Cancer Drugs Fund.

Summary of submitted clinical effectiveness evidence

The CS includes:

- A systematic review of clinical effectiveness studies for atezolizumab and a systematic search for studies on a wide range of comparators;
- A network meta-analysis comparing atezolizumab to comparators in the NICE scope, based on a simulated treatment comparison.

A systematic search was conducted by the company to identify studies on atezolizumab and any comparator chemotherapy drugs that could be relevant in first-line or second-line treatment settings. The search identified only one study on atezolizumab. This was an ongoing single-arm phase II study (Imvigor 210) which included chemotherapy-naive cisplatin-ineligible patients receiving first-line treatment (cohort 1) and platinum chemotherapy pre-treated patients receiving second-line treatment (cohort 2). The search identified 41 studies of comparators that were deemed eligible for inclusion in a feasibility assessment for network meta-analysis, of which seven comparator studies were finally included in meta-analyses. Assessment of the atezolizumab study followed a systematic review process but the review of comparators was more superficial, with few details of the studies provided.

At the last available data-cut, and based on independent review facility assessment using RECIST v1.1 tumour assessment criteria, first-line patients in Imvigor 210 had a median overall survival of 15.9 months, median progression-free survival 2.7 months, an objective response rate of 22.7%, and the median duration of response had not yet been achieved (median follow-up was 17.2 months and median treatment duration 15 weeks [range 0 to 102 weeks]). Second-line patients had a median overall survival of 7.9 months, progression-free survival of 2.1 months, an objective response rate of 15.8%, and the median duration of response had not yet been achieved (maximum duration of response at the latest data cut was 22.6 months). Median follow-up was 21.1 months and median treatment duration 12 weeks [range 0 to 104 weeks)].

Comparison of the clinical effectiveness of atezolizumab against comparator chemotherapy drugs was limited by a lack of primary evidence, as the relevant comparators were either single-arm studies or single arms within controlled trials. To enable a network to be formed for a network meta-analysis, the company employed a simulated treatment comparison to ‘predict’ a matching atezolizumab arm for each comparator study. The resulting comparisons of atezolizumab against each comparator were then included in a network meta-analysis. The

company selected a fractional polynomial model approach for the network meta-analysis since higher-order fractional polynomial models do not require the assumption of proportional hazards. This approach to network meta-analysis is relatively new but is well-suited to the data format available to the company, which consisted of individual patient data for atezolizumab and aggregate population data for the comparators.

The CS presents network meta-analyses on overall survival and progression-free survival and appropriately acknowledges that these have limitations and their results are uncertain, producing clinically implausible results when used directly without adjustment in the economic model. None of the meta-analysis results are discussed in support of the clinical effectiveness of atezolizumab.

In addition, the ERG has identified a number of methodological issues with how the company has conducted the simulated treatment comparison and network meta-analysis which cast further doubt on the validity of the results of these analyses (see 'Commentary on the robustness of the submitted evidence' below).

Summary of submitted cost effectiveness evidence

The CS includes:

- A review of published economic evaluations of treatments for patients with metastatic or locally advanced urothelial carcinoma,
- An economic evaluation undertaken for the NICE STA process to assess atezolizumab for patients with locally advanced or metastatic urothelial carcinoma. The cost effectiveness of atezolizumab is compared with gemcitabine + carboplatin for patients for whom cisplatin-based chemotherapy is unsuitable as a first-line treatment and compared with docetaxel, paclitaxel and best supportive care for patients whose disease has progressed after prior chemotherapy.

A systematic review was conducted by the company to identify economic evaluations of treatments for patients with metastatic or locally advanced urothelial carcinoma. The review identified seven studies but reported that none of these were relevant to the current submission.

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 its comparators. 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 + paclitaxel (Table 1). The ICERs for second-line atezolizumab compared to

docetaxel, paclitaxel and best supportive care are £131,579, £104,850, £98,208 per QALY gained respectively (Table 2).

Table 1 First-line base case cost effectiveness results

Intervention / comparator	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
Atezolizumab	£77,211	2.69			
Gemcitabine + carboplatin	£18,106	1.35	£59,106	1.34	£44,158

ICER: incremental cost-effectiveness ratio; QALY: quality-adjusted life year

Table 2 Second-line base case cost effectiveness results

Intervention / comparator	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
Atezolizumab	£71,868	1.23			
Docetaxel	£9,439	0.76	£62,430	0.47	£131,579
Paclitaxel	£16,606	0.71	£55,262	0.53	£104,850
Best supportive care	£4,836	0.55	£67,032	0.68	£98,208

ICER: incremental cost-effectiveness ratio; QALY: quality-adjusted life year

In probabilistic sensitivity analyses, the probability of first-line atezolizumab being cost-effective is 10.9% and 53.9% at willingness to pay thresholds of £30,000 and £50,000 per QALY respectively. The probability of second-line atezolizumab being cost-effective is 0% and 0% at willingness to pay thresholds of £30,000 and £50,000 per QALY respectively

The company conducted sensitivity analyses and scenario analyses and concluded that the key drivers to the cost-effectiveness results were the price of atezolizumab and the utility of patients in the progressed disease health state. However, the company did not include sensitivity analyses for overall survival or time to treatment discontinuation

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:

- The company suggests that the proportional hazards assumption is unlikely to hold for comparisons of atezolizumab against standard chemotherapy drugs; however, they based their network meta-analysis for first-line comparisons on a zero-order version of the fractional polynomial model which assumes proportional hazards. The company does not discuss the plausibility of this model.
- 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.

Areas of uncertainty

The company has not provided any 'reality checks' to gauge whether their network meta-analysis analysis results might be reasonable or subject to bias. Uncertainties arising at different steps of the simulated treatment comparison and meta-analysis are not discussed or propagated through to the final results so the cumulative impact of small errors and inconsistencies identified by the ERG is unclear.

The fractional polynomial network meta-analysis approach is a relatively complex method that involves numerous computational steps, and it is important that the analysis approach is reported clearly and as fully as possible for the method to be adequately understood. The company's description of the methods is rather limited and several key aspects of the methodology not reported in the CS were revealed indirectly by the company in responses to clarifications. Due to the limited reporting it is possible that some methodological issues might have gone undetected by the ERG.

There is considerable uncertainty regarding the extent to which the clinical benefits of atezolizumab exceed those of comparator treatments. The uncertainty is due to the immaturity of the evidence base for atezolizumab and because there are no direct randomised controlled trials between atezolizumab and its comparators.

The company has not fully explored uncertainty around the model results through sensitivity and scenario analyses. In particular, they have not included sensitivity analyses varying the treatment effect of atezolizumab or varying parametric survival distribution for overall survival and time to treatment discontinuation.

Summary of additional work undertaken by the ERG

In order to address the issues identified above we undertook a series of sensitivity analyses that varied the treatment effect of atezolizumab, the parametric survival distributions used for overall survival and time to treatment discontinuation, and the utility values used for model health states.

Our base case contained the following elements: changes to utility values and parametric survival distributions used for overall survival and time to treatment discontinuation.

The first-line and second-line results are shown in Table 3 and Table 4. The ERG base case ICER for first-line atezolizunab compared to gemcitabine + carboplatin is £93,948 per QALY gained. The ERG base case ICERs for second-line atezolizumab compared to docetaxel, paclitaxel and best supportive care are £288,247, £180,901 and £166,805 per QALY gained respectively. The ERG cautions that there is considerable uncertainty in the model results.

Table 3 ERG first-line base case analysis results

Intervention / comparator	Costs	Incremental costs	QALYs	Incremental QALYs	ICER (£/QALY)
Atezolizumab	£60,650		1.32		
Gemcitabine + carboplatin	£12,469	£48,181	0.81	0.51	£93,948

ICER: incremental cost-effectiveness ratio; QALY: quality-adjusted life year

Table 4 ERG second-line base case analysis results

Intervention / comparator	Costs	Incremental costs	QALYs	Incremental QALYs	ICER (£/QALY)
Atezolizumab	£66,254		0.84		
Docetaxel	£8,196	£58,059	0.64	0.20	£288,247
Paclitaxel	£13,615	£52,640	0.55	0.29	£180,901
Best supportive care	£4,090	£62,164	0.47	0.37	£166,805

ICER: incremental cost-effectiveness ratio; QALY: quality-adjusted life year

1 INTRODUCTION TO THE ERG REPORT

This report is a critique of the company's submission (CS) to NICE from Roche on the clinical effectiveness and cost effectiveness of atezolizumab for treating locally advanced or metastatic urothelial carcinoma. It identifies the strengths and weaknesses of the CS. A clinical expert was consulted to advise the ERG and to help inform this review.

Clarification on some aspects of the CS was requested from the company by the ERG and NICE on 8th February 2017. A response from the company via NICE was received by the ERG on 27th February 2017 and this can be seen in the NICE committee papers for this appraisal.

2 BACKGROUND

2.1 Summary and critique of the company's description of the underlying health problem

The company has provided an accurate overview of urothelial carcinoma (CS section 3), including a very brief overview of the condition (CS section 3.1), information on the course of disease and prognosis (CS section 3.2), the burden of illness (CS section 3.3), and an explanation of the unmet clinical need (CS section 3.4).

The CS refers both to 'bladder cancer' and 'urothelial carcinoma', although the condition defined in the scope of the current technology appraisal is, strictly, urothelial carcinoma. The majority of bladder cancers (~90%) in the UK are attributable to urothelial carcinoma,[1] and the majority of urothelial carcinomas (90-95%) develop in the bladder.[2] The remaining urothelial carcinomas (10-15%) develop in the renal pelvis and the ureters (referred to as upper tract urothelial carcinomas) and also in the urethra. Although not mentioned in the CS, occurrence of urothelial carcinomas at these different sites is not independent: in 17% of cases of upper tract urothelial carcinoma there will be concurrent bladder cancer present, and 22-47% of the upper tract urothelial carcinomas which develop will recur in the bladder.[2]

Note that the term 'bladder cancer' as used in the scientific literature and clinical guidance documents can have several meanings: it may refer to any cancer of the urinary bladder; or urothelial carcinoma; or both.

Development and classification of urothelial carcinoma

Urothelial carcinoma (also commonly referred to as transitional cell carcinoma) begins in transitional cells (also called urothelial cells), which are flexible cells forming the inner lining (urothelium) of the bladder and upper urinary tract. The CS points out that patients are classified according to the stage of development of the carcinoma, as having either early non-muscle-invasive bladder cancer, muscle-invasive bladder cancer, or metastatic cancer (CS section 3.1).

The CS does not describe the staging or grading of urothelial carcinoma, although this information is readily available from organisations such as Cancer Research UK, Macmillan Cancer Support, and the European Association of Urology.[1-5] The stage of bladder cancer is commonly represented using the Tumour-Node-Metastasis classification (TNM).[5] CS Table 7 shows how non-muscle invasive disease, muscle-invasive disease and metastatic disease relate to the different stages of cancer on the TNM classification.

In non-muscle-invasive bladder cancer the tumour remains confined to the lining of the bladder wall, i.e. it remains within the urothelium (stage Tis or Ta) or has invaded the adjacent connective tissue layer (stage T1) but has not penetrated into the underlying muscle layer. Tumours that have penetrated into the muscle layer (stages T2-T4) are referred to as muscle-invasive bladder cancer. These may spread locally or regionally, or metastasise to distant parts of the body, and are then referred to as metastatic bladder cancer.

The CS does not explicitly define 'locally advanced' urothelial carcinoma. However, it is specified (using the TNM classification) in the inclusion criteria for the company's pivotal atezolizumab study (Imvigor 210) as '*T4b and any N; or any T and N2-3'* (CS Table 25). According to Cancer Research UK, 'locally advanced bladder cancer' refers to cancer that has grown through the bladder wall or has spread only to lymph nodes.[6]

Risk factors for urothelial carcinoma

The CS correctly points out that well-known risk factors for bladder cancer are advanced age, smoking, and exposure to some industrial chemicals. Cancer Research UK lists a wider range of risk factors, including (among others) exposure to ionizing radiation, exposure to chlorinated water, use of certain drugs (e.g. pioglitazone, cyclophosphamide), and a history of bladder infections or inflammation.[7, 8] However, according to the European Association of Urology, there is consensus that the most important modifiable risk factor for urothelial carcinoma is smoking.[9]

Cancer Research UK estimates that 42% of bladder cancer cases in the UK could be preventable due to their link to lifestyle factors.[8]

Incidence rates

The CS reports that bladder cancer is the 10th most common cancer in the UK, although Cancer Research UK state that it is the 7th most common.[8] The latest data cited are from 2014, when there were 10,063 new cases (CS section 3.1). These figures are consistent with the current incidence data available from Cancer Research UK,[8] although it is not clear which type(s) of bladder cancer the data refer to. The incidence of bladder cancer is higher in males (around 7,300 cases in 2014) than in females (around 2,800 cases in 2014), and is more common in White than Asian or Black people, and in people living in deprived areas.[8] From 2012 to 2014, more than half of bladder cancers (55%) were diagnosed in people aged 75 years and over.[8] As mentioned in the CS, the incidence of bladder cancer has decreased by 27% in the UK since the late 1970s and has also decreased in other European countries, and this trend is thought to reflect changing smoking habits and stricter controls on exposure to industrial chemicals.[10]

Course and prognosis

The CS provides an accurate description of the symptoms, course and prognosis of bladder cancer (CS section 3.2). Haematuria (blood in the urine) is the most frequent presenting symptom of bladder cancer, occurring in approximately 80% of cases. Patients may also experience increased frequency and urgency of urination and pain when passing urine. These symptoms mean that bladder cancer is often diagnosed at an early stage, with 75-85% of urothelial carcinomas of the bladder being classed as not invasive at diagnosis (although only 40% of urothelial carcinomas of the upper urinary tract are classed as non-invasive at diagnosis).[2]

The CS points out that non-muscle-invasive bladder cancer is highly treatable but has a high risk of recurrence. The high recurrence rate means that follow-up is a crucial component in effective management.[11] Literature cited by the CS suggests that up to 45% of patients with non-muscle-invasive bladder cancer will eventually progress to muscle-invasive bladder cancer, and that 20-50% of those with muscle-invasive bladder cancer will progress further to metastatic disease.

The CS reports survival rates from the SEER (Surveillance, Epidemiology, and End Results Program) of the US National Cancer Institute (Howlader et al.[12]), which cover the period 1975-2008. According to the SEER data, the 5-year survival rate for localised non-muscle-invasive bladder cancer was 69%, dropping to 34% for those with regional spread, and 6% for metastatic disease. Cancer Research UK provides overall mortality rates[13] and survival rates[8] for bladder cancer, but not specifically for non-muscle-invasive, muscle-invasive, or metastatic disease. Age-specific bladder cancer mortality rates in the UK rise steeply from around age 55-59, with the highest rates being in the 90+ age group.[13] According to Cancer Research UK, males have better survival than females,[14] yet mortality rates are considerably (2.1 times) higher for males,[13] reflecting the higher prevalence of bladder cancer in males.

Section 3.5.3 of the CS presents a table showing bladder cancer 5-year survival rates and the probabilities of recurrence separately for different cancer stages at diagnosis (CS Table 7). The CS credits these data to Howlader et al. 2011,[12] Kaufman et al. 2009,[11] National Collaborating Centre for Cancer 2015 (which reflects NICE guideline NG2[15]), Sharma et al. 2009,[16] de Vos & de Wit 2010,[17] and the American Cancer Society 2015.[18] The data in CS Table 7 appear to be from the SEER program; however, we could not find the source data for CS Table 7 in any of these cited references. The American Cancer Society[18] reported that 5-year survival rates from the SEER program for bladder cancer stages 0, 1, 2 and 3 were about 98%, 88%, 63% and 46% respectively.

2.2 Summary and critique of the company's overview of current service provision

The CS provides a description of the current first line treatment for people with locally advanced or metastatic urothelial bladder cancer (CS section 3.5.1). This is in line with the NICE recommendations.[15] For patients who are otherwise physically fit (performance status 0 or 1) and have adequate renal function, a cisplatin-based chemotherapy such as cisplatin with gemcitabine or accelerated MVAC (methotrexate, vinblastine, doxorubicin and cisplatin) with granulocyte-colony stimulating factor is recommended. For those in whom cisplatin is unsuitable (e.g. if performance status is poor, or they have inadequate renal function), NICE recommends carboplatin with gemcitabine. The company cites evidence from a randomised controlled trial by De Santis et al. (2009)[19] which they say estimates that carboplatin with gemcitabine are used in up to 50% of patients in the first line setting. However, this is a secondary reference which cites

four studies which were published between 2000 and 2006. We note that the 2014 European Society for Medical Oncology practice guidelines[20] concur with this figure, although no source is cited. It is therefore unclear if the estimate of 50% is still valid. The CS concludes in Section 3.5.2 that a significant proportion of patients therefore do not receive the most effective first-line therapy (cisplatin with gemcitabine) and in these patients alternatives are needed.

The CS mentions that most patients will experience disease progression and may require second-line therapy, citing Bellmunt et al. 2013[21] which is a randomised controlled trial (CS section 3.5.1). There is no reference for this statement in the Bellmunt 2013 paper; however, on the basis of evidence presented on the course and prognosis of bladder cancer (CS section 3.2), the ERG agrees that most patients will experience disease progression. The CS correctly states that there is only one treatment (vinflunine) with a licensed indication for second line treatment for urothelial cancer but that it is not recommended by NICE.[22] The CS states there is therefore a wide variety of practice in the choice of second line treatment for these patients citing two sources (the 2014 European Society for Medical Oncology practice guidelines[20] and a UK survey by Lamb et al.[23]) and the view of their clinical experts (CS section 3.5.1). The variety of practice is not discussed in the guideline document; the UK survey shows variability in practice, but the survey was conducted in 2011. The CS concludes (CS section 3.5.2) that no treatment has been shown to improve survival in the second-line setting, and the ERG concurs.

2.3 Summary and critique of the company's definition of the decision problem

Population

The population defined in the company's decision problem is adults with locally advanced or metastatic urothelial carcinoma:

- for whom cisplatin-based chemotherapy is unsuitable
- whose disease has progressed after prior chemotherapy

This corresponds with the final scope issued by NICE and the draft Summary of Product Characteristics (SmPC) for atezolizumab.

The CS refers to first-line (1L) and second-line (2L) treatment, which correspond to two treatment cohorts of the company's key clinical effectiveness study for atezolizumab (Imvigor 210). The populations specified in Imvigor 210 were defined as patients with advanced urothelial cancer:

- who were cisplatin-ineligible (medically ineligible to receive cisplatin chemotherapy), and were either previously untreated or had disease progression at least 12 months after their last dose of treatment with a platinum-containing neoadjuvant or adjuvant chemotherapy regimen (cohort 1, 1L);
- who had disease progression following treatment with a platinum-based chemotherapy regimen (cohort 2, 2L).

The ERG considers that the population described in the decision problem is appropriate for the NHS, although notes that the final wording of the indication may change when the Medicinal Products for Human Use (CHMP) opinion is released.

The ERG notes that atezolizumab has FDA approval for the treatment of patients with locally advanced or metastatic urothelial carcinoma who have disease progression during or following platinum-containing chemotherapy or have disease progression within 12 months of neoadjuvant or adjuvant treatment with platinum-containing chemotherapy.[24]

Intervention

The intervention specified by the NICE scope and the company's decision problem is atezolizumab (Tecentriq), a monoclonal antibody that binds to programmed death ligand 1 (PD-L1). Committee for Medicinal Products for Human Use (CHMP) opinion is anticipated [redacted] and regulatory approval is expected in [redacted]. The recommended dose is 1200 mg administered intravenously every three weeks until loss of clinical benefit or unmanageable toxicity (CS Table 6 and draft SmPC). This is the same dose as used in the Imvigor 210 study, although treatment in the study was continued until disease progression per RECIST (Response Evaluation Criteria for Solid Tumours) v1.1 in first-line patients, or until lack of clinical benefit in second-line patients. The ERG considers that the intervention in the decision problem reflects its anticipated use in the UK and is appropriate for the NHS.

Comparators

The comparators are listed in the final scope issued by NICE according to the patient population.

For first-line patients for whom cisplatin-based chemotherapy is unsuitable, the comparators specified in the NICE scope are gemcitabine + carboplatin, or best supportive care. However,

the company's decision problem includes only gemcitabine + carboplatin. The company states that according to their expert clinical advisor panel, all patients willing and able to receive therapy would receive a first-line treatment option and that those receiving best supportive care would be unable or unwilling to receive any active treatment, including atezolizumab. The company states that these patients would represent a small minority, and also notes that best supportive care has not been assessed in a clinical trial in the first line setting, so that a comparison with atezolizumab would not be possible. However, the company does not provide evidence of the numbers of patients receiving best supportive care as a first-line treatment. The ERG's clinical advisor suggested that as atezolizumab is an immunotherapy, which would have a better safety profile than chemotherapy, then patients unable or unwilling to receive chemotherapy might be able and willing to receive atezolizumab.

For people whose disease has progressed after prior chemotherapy (i.e. second-line), the NICE scope refers specifically to platinum-based prior chemotherapy. The NICE scope separates second-line patients into those who are suitable and unsuitable for cisplatin-based chemotherapy, and for both groups the following comparators are specified:

- Retreatment with first-line platinum-based chemotherapy (only for people whose disease has had an adequate response); for cisplatin-ineligible patients retreatment would be with gemcitabine +carboplatin
- Docetaxel
- Paclitaxel
- Best supportive care

The CS decision problem for the comparators differs from the NICE scope in that the company has removed retreatment with first-line platinum-based chemotherapy as a comparator. The company's justification is that their expert clinical advice was that retreatment with first-line therapy is an option for only a very small proportion of people, is not considered standard of care within England, and '*has not been the subject of a systematic clinical evaluation'*. The ERG notes that the company does not provide any evidence regarding the proportion of people undergoing retreatment with first-line therapy to justify the exclusion. However, the ERG's clinical advisor suggested that the company's approach seems reasonable, given the limited evidence base.

As a result of having removed retreatment as a comparator, the CS decision problem for the comparators differs from the NICE scope as the company does not distinguish between the second-line cisplatin-eligible and cisplatin-ineligible groups. The company's justification for combining the groups is that treatment patterns and response rates for patients receiving second-line treatment with docetaxel, paclitaxel or best supportive care are '*not anticipated to be different based on their eligibility for cisplatin and receiving 2L treatment*' (CS section 1.1). The ERG's clinical advisor suggested that it is difficult to know whether cisplatin-eligible and ineligible patients would fare differently on second-line treatment, given the limited evidence base; and the ERG notes that the studies on relevant second-line comparators did not report whether any of their patients were cisplatin-ineligible (see section 3.1.3 below).

The CS does not define best supportive care, for patients in either the first-line or second-line settings. In response to a clarification request from the ERG and NICE (clarification response A2), the company stated that: '*Patients will receive best supportive care when they are not suitable for active second-line treatment due to clinical considerations of their disease, co-morbidities, or performance status. For these patients, the aim of treatment is to relieve symptoms of their disease, and would include support from oncology and palliative care teams including consultants and specialist nurses, palliative radiotherapy for the relief of symptoms, analgesia, support in the community, and hospice admission*.'

In their clarification of best supportive care (clarification response A2) the company also provided an explanation of their definition of second-line treatment, as follows:
"For clarity, the second line (2L) population includes the following:

- *Patients whose disease has progressed after platinum-based chemotherapy*
- *Patients for whom cisplatin-based chemotherapy is unsuitable and whose disease has progressed after* ***non-****platinum-based therapy"* (ERG bold)

As we note above, this is inconsistent with the NICE scope which specifically refers to patients whose disease has progressed after platinum-based chemotherapy. However, the CS does not refer to non-platinum first-line therapy, so the extent of any deviation from the scope is unclear.

Outcomes

The outcomes listed in the company's decision problem are the same as those specified by the NICE final scope: overall survival, progression-free survival, response rates, adverse effects

and health-related quality of life. However, the CS does not actually report health-related quality of life; therefore, the company's decision problem is misleading. 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.

3 CLINICAL EFFECTIVENESS

3.1 Summary and critique of the company's approach to systematic review

3.1.1 Description of the company's search strategy

The CS states that a wide search was conducted for clinical effectiveness evidence, although the search strategy is not provided (CS section 4.1). Upon request from the ERG and NICE, the company provided a detailed search strategy for each of their information sources (clarification response A4) and these appear to be appropriate and fit for purpose. Overall, the systematic search process is well described, and the information sources and search dates are clearly reported (CS Table 9). The sources included MEDLINE, EMBASE, and the Cochrane Library (searched in June 2016), study registries and conference abstracts (searched in July 2016), and HTA and drug regulatory agencies (searched in August 2016). The CS states that no time limits were applied to the bibliographic searches, except for conference abstracts which were restricted to 2015-2016 (CS section 4.1.3.1). The eligibility criteria (CS Table 10) indicate that reviews (systematic and non-systematic) and meta-analyses were excluded, but the CS does not report whether any were used as a source of references. The CS does, however, report that reference lists of the included primary studies were checked by two reviewers to identify any trials directly comparing atezolizumab versus any comparator (CS section 4.1.5).

The CS states that the goal of the clinical effectiveness search was '*to capture current and upcoming treatments for all relevant markets in the relevant indications for atezolizumab*' (CS section 4.1.1). As such, the search is likely to have been considerably wider than the scope of the current technology appraisal.

The clinical effectiveness search was 5-7 months out of date when the ERG received the CS. We therefore ran a search for the period 2016-2017 on MEDLINE, EMBASE, and the Cochrane Library, covering the condition (bladder or urothelial carcinoma) linked to the following comparators (alone or in combination): paclitaxel, docetaxel, gemcitabine, carboplatin, vinflunine, MVAC (methotrexate + vinblastine + doxorubicin + cisplatin), and best supportive care. We also checked clinicaltrials.gov and the UK Clinical Trials Gateway for potentially relevant studies of atezolizumab or comparators. We identified five systematic reviews or meta-analyses covering possible comparators[25-30] that are not cited in the CS and which appear to

have been published after the company's searches were conducted. We did not find any additional completed or ongoing studies of atezolizumab.

In addition to the update searches, the ERG checked the reference lists of key guidance documents,[5, 9, 15] an evidence review for NICE,[31] recent review articles[32, 33] and a meta-analysis[30] for any potentially relevant studies. We identified 18 studies on comparators (published from 1997 to 2017) which are not cited or listed in the CS but appear, based on their titles and abstracts, to be potentially relevant according to the company's eligibility criteria (CS Table 10). Upon request from the ERG and NICE (clarification question A11), the company confirmed that 16 of these references had been identified and screened for eligibility, and were subsequently excluded, whilst two had not been identified as they had been published after the company's searches were conducted. The potential relevance of these references, and whether they were excluded appropriately, are discussed below in section 3.1.3.

The searches for economic evaluations and utilities (HRQoL) were conducted in September 2016 and resource-use searches were conducted in December 2016. Well-documented and comprehensive search strategies are provided for these searches in CS Appendices 8.7, 8.9 and 8.10. In summary, the ERG considers that the searches and methodology employed by the company to support the systematic reviews of economic evaluations (section 4.1 below), HRQoL (section 4.3.6 below) and resources (section 4.3.7 below) were comprehensive and fit for purpose.

3.1.2 Statement of the inclusion/exclusion criteria used in the study selection

The CS reports eligibility criteria for the population, intervention, comparators, outcomes, and study design in CS Table 10. The company confirmed (clarification question A6) that all of the eligibility criteria were specified a priori.

Eligible population

The population eligibility criteria (but not the scope or decision problem), specifically exclude adjuvant and neoadjuvant stages of the treatment pathway, although the ERG notes that the studies which were ultimately included by the company differed in whether they reported adjuvant and neoadjuvant therapy (section 3.1.3). The eligibility criteria report PD-L1 expression subgroups ("2/3") but no subgroups are specified in the scope and decision problem; the CS

does not explain the rationale for these subgroups, although we understand that efficacy of atezolizumab is likely to vary according to PD-L1 expression status.

Eligible intervention and comparators

The eligibility criteria for the comparators are not fully clear in the CS, since CS Table 10 lists two different sets of eligible comparators, under both the 'Intervention' eligibility criteria domain and the 'Comparators' eligibility criteria domain:

- The 'Intervention' domain in CS Table 10 lists (in addition to atezolizumab) examples of 38 eligible comparators. These include platinum-based, taxane-based and other non-platinum chemotherapies, and monoclonal antibody therapies. The CS also states that '*any other applicable chemotherapies, immunotherapies, antineoplastic agents, antineoplastic protocols, molecular-targeted therapies, cancer vaccines, protein kinase inhibitors, angiogenesis inhibitors, taxanes, taxoids, etc*.' would be eligible.
- The 'Comparators' domain in CS Table 10 specifies '*any pharmacological intervention used*', placebo, and best supportive care.

These two lists of eligible comparators in CS Table 10 are both considerably broader than the comparators specified in the scope and decision problem. However, the CS implies (CS Figure 3; CS section 4.1.4) that the eligibility criteria in CS Table 10 were those used for initial screening of titles and abstracts, and that different, smaller, sets of comparators were subsequently considered eligible:

- CS section 4.1.4 (Search results) states that studies were prioritised in terms of the importance of the comparators, based on clinical guidelines and standards of care in the UK, France, Australia, Canada, and Sweden, with studies on the following comparators being eligible: '*best supportive care, carboplatin + paclitaxel, docetaxel, paclitaxel, nab-paclitaxel, vinflunine, gemcitabine, gemcitabine + paclitaxel, MVAC, carboplatin, cisplatin, oxaliplatin (platinum-based re-challenge if >12 months since last dose), pembrolizumab, nivolumab, and gemcitabine + cisplatin for 2nd line as well as gemcitabine + carboplatin, gemcitabine + paclitaxel and best supportive care for the first-line cisplatin-ineligible population*.' According to CS Figure 3, this prioritisation took place at the full-text screening step. This list of comparators is still broader than the list in the decision problem.
- CS section 4.10.3 (Comparators of interest), which refers to the assessment of studies for the network meta-analysis, lists the eligible comparators as being gemcitabine +

- carboplatin for first-line treatment, and paclitaxel, docetaxel or best supportive care for second-line treatment. The CS does not explain why this list of comparators is different 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 numbers of excluded 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.

Network meta-analysis feasibility assessment

The CS reports that a two-stage process was then used to identify potential bridging studies which might enable indirect linking between relevant comparators in the network meta-analysis (CS section 4.10.5):

In stage 1, 233 studies were assessed and excluded if they did not report one or more of the four outcomes of interest. After this step 74 publications reporting 43 studies remained (i.e. 159 were excluded). There is a discrepancy in that CS section 4.10.5 implies the 159 publications had been excluded due to ineligible outcomes whilst CS Figure 3 and the company's clarification response A8 state that the reason for exclusion was '*interventions not first priority*'.

In stage 2, studies were selected according to their feasibility for inclusion in the network meta-analysis, based on '*building the study networks and their connectivity*', '*assessing the availability of baseline factors associated with the clinical outcomes of interest*', and, for the overall survival and progression-free survival analyses, '*assessing the presence of Kaplan-Meier curves in the corresponding publications*' (CS section 4.10.5). However, the CS does not provide explicit objective criteria for how eligibility decisions were made at stage 2. At this stage 27 publications were excluded, leaving 47 publications for inclusion in the analysis, and these reported on 28 individual studies. The reasons for exclusion are listed in CS Appendix 8.2, but the descriptions are inconsistent and imprecise. The company provided clarification upon request from the ERG and NICE (clarification response A9). The remaining 28 studies which were included after the network meta-analysis feasibility assessment are listed in CS Table 13 (2 studies on first-line therapies) and CS Table 14 (26 studies on second-line therapies).

One of the studies listed as being excluded is the single-arm atezolizumab study Imvigor 210, although the company has included Imvigor 210 in their network meta-analysis. Imvigor 210 has both first-line (1L) and second-line (2L) cohorts and is therefore listed twice in CS Appendix 8.2, meaning that the actual number of excluded studies of comparators was 14 (6 on first-line therapies, 8 on second-line). The ERG has checked and concurs with the company's reasons for excluding these 14 studies, with the exception of a study by Meluch et al. (2001).[34] Appendix 8.2 of the CS states that the Meluch study was excluded due to having no predictors; however, age (median and range), sex, and ECOG performance status were reported (the study contained a mix of first-line and second-line patients so we believe it would not meet the eligibility criteria).

The CS states that for time-to-event analyses (i.e. overall survival and progression-free survival), Kaplan-Meier curves were required, and any studies listed in CS Table 13 and CS Table 14 which were not included in network meta-analysis had been excluded due to unavailability of Kaplan-Meier curves (CS section 4.10.5 and clarification response A10). By comparing these tables in the CS it can be deduced that 21 studies (all on second-line therapies) had been excluded due to 'unavailability of Kaplan-Meier curves'. The ERG checked these 21 studies (listed in Appendix 1) and we found that 13 of them did report Kaplan-Meier curves and, therefore, appear to have been inappropriately excluded from the network meta-analysis. However, these studies were on second-line comparators which do not appear to meet the company's final criteria for inclusion (gemcitabine, MVAC, gemcitabine + paclitaxel, carboplatin + paclitaxel, vinflunine). A possible exception is a study by Ko et al. 2013[35] which was on nab-paclitaxel. This study appears to meet the inclusion criteria, since paclitaxel is a relevant comparator (Appendix 1); however, the ERG's clinical expert advisor suggested that the nab (nanoparticle albumin bound) formulation of paclitaxel is rarely, if ever, used for urothelial carcinoma and as such it would be reasonable to exclude it as a comparator.

A further discrepancy in the screening process is that the company's network meta-analyses of overall survival and progression-free survival included different comparators, despite data being available for both outcomes in several studies. As shown in Table 5 below, the overall survival analysis included docetaxel, paclitaxel, and best supportive care, which is consistent with the NICE scope and the company's decision problem. However, in addition to these comparators the company's progression-free survival analysis included gemcitabine, carboplatin + paclitaxel, and vinflunine which are not NICE scoped comparators. Inconsistently, the company's progression-free survival analysis did not include studies by Vaishampayan et al. 2005[36] on carboplatin + paclitaxel or Vaughn et al. 2009[37] on vinflunine (Appendix 1).

As noted above (section 3.1.1), the ERG identified 18 further publications (each describing a single study) which appeared, on title and abstract, to be potentially eligible for inclusion but which are not cited or referenced anywhere in the CS. The company explained (clarification request A11) that 16 of these publications had been identified and screened, then were subsequently excluded; and two were not published at the time of the company's searches in June 2016. The company's clarification response explains the reasons for exclusion; after consulting the full publications the ERG agrees that these studies would be excluded, although in some cases for different reasons to those stated by the company.

The ERG notes that two studies on second-line paclitaxel are available. One by Ko et al. 2013,[35] was on nab-paclitaxel and (as mentioned above) was excluded from the company's overall survival analysis. The other, by Lee et al. 2012,[38, 39] was on a polymeric micelle formulation of paclitaxel and was included in the overall survival analysis. The CS does not discuss the relevance to current clinical practice of any of the chemotherapy formulations in the studies that they included, and we are unclear whether the polymeric micelle formulation of paclitaxel would have similar effectiveness and tolerability compared to standard paclitaxel chemotherapy.

ERG conclusion on the company's screening process

The eligibility screening process is poorly reported and has been applied inconsistently, with: 16 of the screened studies not being referenced in the CS; 13 studies apparently being excluded for reasons other than those stated in the CS; and inconsistent inclusion/exclusion of studies according to the outcome being analysed.

The bottom line for the overall survival analysis appears to be that no key studies have been missed. It is unclear, however, whether the only included paclitaxel study, which used a polymeric micelle formulation, is representative of standard paclitaxel chemotherapy.

For the progression-free survival analysis, the company included comparators which are not specified in the NICE scope or the company's decision problem. The ERG believes this is not a major concern for the current technology appraisal, since the progression-free survival analysis is not used by the company to support the clinical effectiveness of atezolizumab or to inform the economic analysis.

3.1.3 Identified studies

Following the eligibility screening process reported above (section 3.1.2), the company included one single-arm atezolizumab study (Imvigor 210) and 10 comparator studies in their network meta-analysis (Table 5). No RCTs of atezolizumab were identified. As well as being used in the network meta-analysis, Imvigor 210 is reported separately in the CS as being the primary source of efficacy and safety data for atezolizumab (CS section 4.11).

Some RCTs with one or more relevant comparator arms were identified but the majority of the comparator studies which met the eligibility criteria were single-arm studies. As described above

(section 3.1.2), the progression-free survival analysis included studies on comparators which are not specified in the NICE scope or company's decision problem.

Table 5 Comparator study arms included in network meta-analysis

Outcome	Position in the treatment pathway	
	First-line (cohort 1 in atezolizumab study Imvigor 210 [40])	**Second-line (cohort 2 in atezolizumab study Imvigor 210 [41])**
OS (informs company's economic model)	GEM + CAR (Bamias et al.[42]) GEM + CAR [a] (De santis et al.[19, 43, 44])	BSC [a] (Bellmunt et al.[21, 45]) BSC [a] (Noguchi et al.[46, 47]) DOC (Kim et al.[48, 49]) DOC + PBO [a] (Choueiri et al.[50]) PTX (Lee et al.[38, 39])
PFS (does not inform company's economic model)	GEM + CAR (Bamias et al.[42])	BSC [a] (Bellmunt et al.[21, 45]) BSC [a] (Noguchi et al.[46, 47]) DOC (Kim et al.[48, 49]) DOC + PBO a (Choueiri et al.[50]) GEM (Albers et al.[51]) [b] PTX (Lee et al.[38, 39]) Nab-PTX (Ko et al.[35]) [b] CAR + PTX (Kouno et al.[52]) [b] VFL [a] (Bellmunt et al.[21, 45])

BSC: best supportive care; DOC: docetaxel; CAR: carboplatin; GEM: gemcitabine; Nab: nanoparticle albumin bound; PBO: placebo; OS: overall survival; PFS: progression-free survival; PTX: paclitaxel; STC: simulated treatment comparison

[a] single arm from a randomised controlled trial

[b] reports both OS and PFS curves but included only in the PFS analysis (CS Appendix 8.5)

The CS additionally reports a single-arm phase Ia study of atezolizumab (PCD49089g) which the company has cited as a source of some supporting information on atezolizumab efficacy and safety. We note that since PCD49089g is a phase I study it does not meet the company's eligibility criteria (as listed in CS Table 10) and also the cohort was heavily pre-treated and most patients did not receive the licensed dose of atezolizumab (as indicated by the company in clarification response A41). We have summarised the characteristics and effectiveness results of study PCD4989g in Appendix 2.

Only those studies which met the eligibility criteria for network meta-analysis according to the NICE scope and company's decision problem are summarised here.

3.1.3.1 Atezolizumab study: Imvigor 210

Design characteristics of the Imvigor 210 study are given in CS Table 27, which we have summarised below in Table 6.

Eligibility criteria

The CS provides an extensive list of the inclusion and exclusion criteria for the first-line and second-line cohorts of Imvigor 210 (CS Table 28). Due to the large number of criteria provided these are not reproduced fully here, but key criteria are summarised in Table 7 and Table 8.

Participant flow

Section 4.11.5 of the CS (Participant flow) does not provide the flow of the study participants (i.e. the numbers of participants who were screened, enrolled, treated and analysed) in Imvigor 210. However, diagrams showing the participant flow are provided in the study publications for the first-line cohort[40] and the second-line cohort.[41] Due to copyright restrictions these flow diagrams are not reproduced here.

Of 167 participants screened for eligibility for the first-line (cisplatin-ineligible) cohort, 44 were ineligible and were excluded before enrolment. Ineligibility reasons were clearly reported and appear appropriate for 24 of these people, but were reported only as 'all other reasons' for the remaining 20.[40] A total of 123 participants were enrolled in the first-line cohort, but four participants were excluded after enrolment, with reasons reported. One of these exclusions was due to disease progression before cycle 1, although this does not appear to be one of the pre-specified exclusion criteria (as listed in CS Table 28). The remaining 119 participants received at least one dose of atezolizumab. Of these, 102 subsequently discontinued treatment. Reasons for discontinuation were disease progression (n=77), patient withdrawal (n=12), adverse events (n=11) and unspecified other reasons (n=2). The number of participants remaining on-treatment at the July 2016 data-cut (median follow-up 17.2 months; clarification response A34) was n=17.[40]

Table 6 Key design characteristics of the IMvigor 210 study

Location	Patients were recruited from 70 centres in North America and Europe, including 3 sites in the UK.
Design	Single-arm open-label phase II study
Eligibility criteria for participants	Patients with locally advanced or metastatic urothelial carcinoma were enrolled regardless of their PD-L1 expression, or number of prior therapies (from first-line cisplatin-ineligible patients to heavily-treated patients with exposure to multiple prior regimens). Patients were enrolled into one of two cohorts: **Cohort 1:** chemotherapy-naïve patients who are cisplatin-ineligible (N=119) **Cohort 2:** patients who have progressed during or after at least one platinum chemotherapy regimen (N=310)
PD-L1 subgroups	Baseline PD-L1 expression in tumour specimens was centrally evaluated using the VENTANA PD-L1 (SP142) immunohistochemistry assay (Ventana Medical Systems, Mountain View, California, US). PD-L1 expression on IC was evaluated based on three scoring levels: • IC2/3, ≥5% PD-L1 expression in immune cells • IC1, ≥1% and <5% PD-L1 expression in immune cells • IC0, <1% PD-L1 expression in immune cells
Trial drugs, permitted and disallowed concomitant medication	Single-agent atezolizumab 1200 mg administered by intravenous infusion on day 1 of each 21-day cycle until disease progression according to RECIST v1.1 criteria (Cohort 1 only) or until lack of clinical benefit (Cohort 2)
Patient monitoring	Patients had tumour assessments at baseline, every 9 weeks for 12 months, and every 12 weeks thereafter. Patients who discontinued treatment continued follow-up assessments for survival and subsequent anti-cancer therapy every ≈3 months until death, loss to follow-up, withdrawal of consent, or study termination, whichever occurred first.
Primacy outcomes	Co-primary endpoint: [a] • Independent review facility-assessed ORR (confirmed) according to RECIST v1.1 criteria (central independent review; Cohort 1 & 2), and; • Investigator-assessed ORR (according to modified RECIST criteria; immune-related response criteria [Cohort 2 only]).
Secondary outcomes	DOR and PFS assessed by the independent review facility and investigator according to RECIST v1.1 criteria, OS, and 1-year OS. DOR and PFS according to modified RECIST criteria will be additional secondary endpoints. The efficacy endpoints as assessed by modified RECIST criteria are applicable only to Cohort 2.

DOR: duration of response; ORR: objective response rate; OS: overall survival; PD-L1: programmed death ligand 1; PFS: progression-free survival; RECIST: response evaluation criteria in solid tumours
[a] Reference for the primary outcome was a 10% historical control rate (see section 3.1.6)

Table 7 Key inclusion criteria for the Imvigor 210 atezolizumab study

First-line (1L) cohort	Second-line (2L) cohort
● ECOG performance status 0, 1 or 2. ● No prior chemotherapy for inoperable locally advanced or metastatic or recurrent urothelial carcinoma. ● For patients who received prior adjuvant/neoadjuvant chemotherapy or chemo-radiation for urothelial carcinoma, > 12 months treatment free between the last treatment administration and the date of recurrence was required in order for participants to be considered treatment naive in the metastatic setting. ● Ineligible ('unfit') for cisplatin, as per specified criteria in CS Table 28.	● ECOG performance status of 0 or 1. ● Disease progression during or following treatment with at least one platinum containing regimen (e.g., GEM, MVAC, GEM + CAR) for inoperable locally advanced or metastatic urothelial carcinoma or disease recurrence. A regimen is defined as patients receiving ≥2 cycles of a platinum containing regimen. Patients who received one cycle of a platinum-containing regimen but discontinued due to Grade 4 hematologic toxicity or Grade 3 or 4 non-hematologic toxicity may also be eligible ● Patients who received prior adjuvant/neoadjuvant chemotherapy and progressed within 12 months of treatment with a platinum-containing adjuvant/ neo-adjuvant regimen will be considered as 2L patients. Patients with progression after chemo-radiotherapy must demonstrate progression outside the prior radiotherapy port.
● Historically or cytologically documented advanced or metastatic urothelial carcinoma of the bladder, renal pelvis, ureters or urethra; locally advanced bladder cancer must be inoperable. ● Availability of viable tumour specimens as defined in CS Table 28. ● Life expectancy ≥12 weeks. ● Measurable disease as defined by RECIST v.1.1. ● Adequate haematologic and end-organ function (not defined).	

CAR: carboplatin; ECOG: Eastern Cooperative Oncology Group; GEM: gemcitabine; MVAC: methotrexate, vinblastine, doxorubicin & cisplatin combination; RECIST: response evaluation criteria in solid tumours

Of 486 participants screened for eligibility for the 2L (platinum-treated) cohort, 170 were excluded before enrolment. Ineligibility reasons are clearly reported for 134 of these and appear appropriate, but are reported as 'other reason' with no further detail for the remaining 36.[41] A total of 316 participants were enrolled in the 2L cohort, of which 311 received atezolizumab treatment. The five who did not receive atezolizumab were stated not to have met the eligibility criteria, but no reasons are given). One participant was excluded after receiving atezolizumab, due to being not evaluable because of incorrect cohort assignment, although results are reported in the CS for all 311 patients. Of the remaining participants, 248 subsequently discontinued treatment, due to disease progression (n=211), adverse events (n=13), patient withdrawal (n=9) and unspecified other reasons (n=15). The number of participants remaining on-treatment at the September 2015 data-cut (median follow-up 11.7 months; clarification response A34) was n=62.[41]

According to the publications,[40, 41] one first-line cohort participant was re-assigned to the second-line cohort and two second-line participants were re-assigned to the first-line cohort between the May 2015 and September 2015 data cuts.

Table 8 Key exclusion criteria for the Imvigor 210 atezolizumab study

1L and 2L cohorts
● Any approved anti-cancer therapy, including chemotherapy, or hormonal therapy within 3 weeks prior to initiation of study treatment (exceptions: palliative radiotherapy for bone metastases or soft tissue lesions should be completed > 7 days prior to baseline imaging; hormone-replacement therapy or oral contraceptives). ● Active or untreated CNS metastases as determined by CT or MRI evaluation during screening and prior radiographic assessments (patients with treated asymptomatic CNS metastases are eligible, provided they meet all of the criteria specified in CS Table 28). ● Uncontrolled tumour-related pain. ● Comorbidities as specified in CS Table 28, including: leptomeningeal disease; uncontrolled pleural effusion; pericardial effusion; ascites requiring recurrent drainage procedures (≥1 per month); active tuberculosis; active hepatitis B or C; positive test for HIV; severe infections within 4 weeks of starting atezolizumab, or infection signs and symptoms within 2 weeks; history of autoimmune disease; history of specified respiratory diseases including idiopathic pulmonary fibrosis, pneumonia, pneumonitis. ● Allergic hypersensitivity to specified antibodies or biopharmaceuticals. ● Uncontrolled hypercalcaemia, or symptomatic hypercalcaemia requiring specified therapies. ● Low serum albumin as defined in CS Table 28. ● Any other evidence of or suspicion of diseases or metabolic dysfunction that would contraindicate use of an investigational drug, affect the interpretation of the results, or render the patient at high risk from treatment complications. ● Medication-related exclusion criteria as specified in CS Table 28 for stated time periods prior to the initiation of atezolizumab treatment (or anticipated need for): CD137 agonists or immune checkpoint blockade therapies; systemic immune-stimulatory agents (e.g. interferons); systemic corticosteroids or other systemic immunosuppressive medications; antibiotics. ● Prior allogeneic stem cell or solid organ transplant. ● Significant cardiovascular disease as specified in CS Table 28. ● Major surgical procedure other than for diagnosis within 4 weeks of starting atezolizumab, or anticipated need for such procedure during the study. ● Receipt of live attenuated vaccine within 4 weeks of starting atezolizumab, or anticipated need for vaccine during the study.

Baseline characteristics

Baseline characteristics of the participants in Imvigor 210 are reported in CS Table 29 and reproduced here in Table 9, including additional information reported in the publications.[40, 41]

Imvigor 210 was a multinational study conducted in the USA, Canada, France, Germany, Italy, Spain, The Netherlands, and the UK. Five of the participants in the first-line cohort (4.2%) and 17 of the participants in the second-line cohort (5.5%) were in the UK. The CS gives a brief overview of the participants' characteristics (CS section 4.11.5) and concludes that the demographic profiles of each of the first-line and second-line cohort populations are consistent with those observed in the general urothelial carcinoma population in clinical practice, and consistent with patient populations in other recent clinical trials. The ERG's clinical expert advisor agreed that the two populations in Imvigor 210 are generalisable to those with advanced or metastatic bladder cancer in England. Median age was 73 years for participants in the first-line cisplatin-ineligible cohort and 66 years in the second-line platinum-treated cohort, with the youngest patients in each cohort being aged 51 years and 32 years respectively. In both cohorts the majority of the participants were male, and in the second-line cohort the majority were of white ethnicity, although ethnicity is not reported for the first-line cohort.

The CS points out that in the first-line cohort the most common reason for patients being cisplatin ineligible was impaired renal function (69.7% of participants had GFR <60 mL/min), and that the baseline characteristics are representative of patients with poor prognostic factors, including ECOG performance status =2 (20.2%), visceral metastasis (65.5%), liver metastasis (21.0%) and creatinine clearance < 60 mL/min (70.6%).

As shown in Table 9, 15.1% of the patients in cohort 1 (cisplatin-ineligible) had received prior cisplatin therapy. The CS states that this is likely to be due to treatment with cisplatin in the neoadjuvant setting, and following progression patients are subsequently deemed cisplatin ineligible at the time of selecting first-line treatments in the metastatic setting.

In the 2L cohort the majority of participants had visceral metastases (78.4%), with approximately one third having liver metastases (31.0%) and two thirds having ECOG performance status of 1. The CS points out that approximately 40% of participants in the 2L cohort had received ≥2 regimens in the metastatic setting, indicative of a heavily pre-treated population. A prior cisplatin-based regimen had been received by 73% of participants, whilst 26% had received carboplatin alone. The CS states, and the ERG's clinical expert advisor agreed, that this is broadly representative of UK clinical practice in metastatic urothelial carcinoma.

We note that a relatively high proportion of the participants in Imvigor 210 had upper tract urothelial carcinoma, i.e. the primary tumour site was the renal pelvis or ureters: 27.7% in the first-line cohort and 22.2% in the 2L cohort. This is higher than the expected 'real world' proportion of upper tract urothelial carcinomas which is usually given as being around 5-10%.[2] Upper tract carcinomas are more likely to be invasive at diagnosis and have a worse prognosis than those which arise in the bladder.[2]

Table 9 Baseline characteristics of participants in the Imvigor 210 study

	Cohort 1 (1L) Cisplatin-ineligible n=119	**Cohort 2 (2L) platinum-treated n=310**
Age, years, median (range)	73.0 (51–92)	66.0 (32–91)
Age ≥ 80 years, n (%)	25 (21.0)	24 [a] (7.7)
Sex, male, n (%)	96 (80.7)	241 (77.7)
Primary tumour site, n (%) [b]		
Bladder or urethra	85 (71.4)	239 [a] (76.8)
Renal pelvis or ureter	33 (27.7)	69 [a] (22.2)
Metastatic disease, n (%)	110 (92.4)	291 [a] (93.9)
Visceral sites [c]	78 (65.5)	243 (78.4)
Liver only	25 (21.0)	96 (31.0)
Lymph node only	31 (26.1)	43 (13.9)
Prior therapy, n (%)		
Radiotherapy	12 (10.1)	99 [a] (31.9)
Perioperative chemotherapy [d]	22 (18) [e]	56 [a] (18.0)
Cisplatin-based	18 [a] (15.1)	227 (72.9)
Carboplatin-based	1 [a] (0.8)	80 (26.1)
Number of prior regimens (metastatic setting)	n=0, 98.3% n=1, 1.7%	n=0, 18.1% n=1, 39.0% n=2, 21.3% n≥3, 21.6%
Prior cystectomy or nephroureterectomy	80 (67.2) [f]	228 [a] (73.5)
Haemoglobulin ≤ 10 g/dl	19 (16.0) [f]	69 [a] (22.3)
PD-L1 expression immunohistochemistry subgroups (%)		
IC0 (PD-L1 expression <1%)	32.8	33.2%
IC1 (PD-L1 expression ≥1 but <5%)	40.3	34.5

IC2/3 (PD-L1 expression ≥5%)	26.9	32.2
IC1/2/3 (PD-L1 expression ≥1%)	67.2	66.8

Table 9 continued	**Cohort 1 (1L) Cisplatin-ineligible n=119**	**Cohort 2 (2L) platinum-treated n=310**
ECOG PS 0	45 [a] (37.8)	117 (37.7)
ECOG PS 1	50 [a] (42.0)	193 (62.3)
ECOG PS 2	24 (20.2)	1 [a] (0.3)
Renal impairment, GFR <60 and >30 mL/min	83 (69.7)	108 or 109 [a] (35)
Hearing loss, 25 dB[f]	17 (14.3)	Not reported
Peripheral neuropathy, ≥Grade 2	7 (5.9)	Not reported
Renal impairment and ECOG PS 2	8 (6.7)	Not reported

ECOG: Eastern Cooperative Oncology Group; GFR: glomerular filtration rate; PS: performance status
[a] number not reported in CS or publication; estimated from percentage by ERG
[b] excluding 1 participant with primary tumour site prostatic urethra
[c] liver, lung, bone, any non-lymph node or soft tissue metastasis
[d] adjuvant or neoadjuvant treatment with first disease progression beyond 12 months (except for 1 participant who received targeted therapy)
[e] as reported in the publication[40] (CS reports percentage = 20.2)
[f] provided by the company in clarification response A44

3.1.3.2 Comparator studies

The CS does not provide the baseline characteristics of the comparator studies, except in relation to whether the studies reported four prognostic variables (proportion with age > 65 years, proportion male, proportion with liver metastases, and proportion with ECOG performance status ≥1) (CS Table 17). In response to a clarification request by the ERG and NICE (clarification response A25), the company provided tables summarising the characteristics of the comparator studies. However, the tables focus mainly on methodological aspects of the studies and they report very little information on the participants' characteristics. Whilst this partly reflects a paucity of information reported by the primary studies, there is more information available in the study publications that could have been provided. The ERG has consulted the study publications and we have summarised the available information on the participants' characteristics for the first-line studies in Table 10 and for the second-line studies in Table 11.

Table 10 Baseline characteristics of participants in the first-line comparator studies

Baseline characteristic	Bamias et al.[42]	De Santis et al.[19, 43, 44]
Study design	Single arm	RCT
Regimen (number of participants) Data are reported for **bold** arms	**GEM + CAR (n=34)**	**GEM + CAR (n=119)** M-CAVI (n=119)
Age, years, median (range)	75.5 (57–84)	70 (36–87)
Age, proportion >65 years	-	-
Sex, male, n (%)	28 (82)	90 (75.6)
Primary tumour site, n (%): bladder	30 (88)	90 (75.6)
renal pelvis	3 (9)	12 (10.1)
ureter	1 (3)	12 (10.1)
urethra	0 (0)	3 (2.5)
other (unspecified)	0 (0)	2 (1.7)
ECOG PS 0, n (%)	11 (32)	20 (16.8) [a]
ECOG PS 1, n (%)		46 (38.7) [a]
ECOG PS 2, n (%)	-	53 (44.5) [a]
ECOG PS ≥2, n (%)	23 (68)	-
With comorbidities, n (%)	22 (65) [b]	59 (49.6) [c]
Haemoglobin <10 mg/dl, n (%)	5 (15)	-
Any metastases, n (%)	-	-
Visceral metastases, n (%)	15 (44)	55 (46.2)
Liver metastases, n (%)	-	20 (16.8)
Median follow up, months	8	54
Adjuvant or neoadjuvant therapy, n (%)	6 (17.6)	0 (0)

- (dash) indicates data not reported; CAR: carboplatin; GEM: gemcitabine; ECOG: Eastern Cooperative Oncology Group; MCAVI: methotrexate + carboplatin + vinblastine; PS: performance status; WHO: World Health Organisation

[a] reported as WHO PS score (which is the same as ECOG PS score)

[b] described as comorbidities precluding cisplatin therapy

[c] described as associated chronic disease

As can be seen in Table 10, it is difficult to determine whether the two studies of first-line gemcitabine + carboplatin were homogeneous because the studies used different criteria for describing the participants' characteristics, or did not report some key characteristics. Both studies enrolled predominantly men (75.6% to 82%); most primary tumours (75.6% to 88%) were in the bladder; and just under half the participants in each study (44% to 46.2%) had

visceral metastases. Participants in the Bamias et al. study[42] were older than those in the De Santis et al. study[44] (median age 75.5 versus 70 years) and a higher proportion had comorbidities (65% versus 49.6%), although the proportion of patients with ECOG performance status 0 or 1 was lower in the Bamias et al. study (32% versus 55.5%). However, the definitions of comorbidities were not identical in the studies. Only Bamias et al. permitted prior adjuvant or neoadjuvant therapy (received by 17.6% of patients). We note that the study by Bamias et al. had a relatively small sample size (n=34) compared to that of De Santis et al. (n=119).

The summary of participant characteristics for the five studies of second-line comparators (Table 11) shows that it is difficult to compare these studies in detail, as different criteria were used to describe the participant populations, and some studies did not report key information. For age, the studies either reported the proportion of participants aged ≥65 years (three studies; range 45.8 to 49 years), or the median age (three studies; range 57 to 65 years) (one study reported both measures). Four studies reported that the participants were predominantly male (68.1% to 80%) whilst Bellmunt et al.[21] did not report this. Only the studies by Kim et al.[49] and Lee et al. [39] reported the primary sites of the carcinoma, which was most frequently in the bladder (58% to 70%), though in both these studies nearly a quarter (23% to 24%) of the tumours originated in the ureters. The studies all included patients with ECOG performance score 0 or 1, apart from Lee et al.[39] which included 14% of patents with performance score 2. Four studies (except Lee et al.[39]) reported the proportion with haemoglobin concentration <10 mg/dl, and this ranged from 8.5% to 22%. Four studies (except Noguchi et al.[47]) reported the proportion with visceral metastases, which ranged from 61% to 74%, whilst three studies (excluding Bellmunt et al.[21] and Noguchi et al.[47]) reported liver metastases, which ranged from 30% to 37.5%. The median follow-up in the second-line comparator studies varied considerably, from 3.2 to 45 months. Sample sizes were relatively small in three studies by Kim et al., Lee et al. and Noguchi et al. (31 to 41 participants) but larger in the studies by Choueiri et al. (n=75) and Bellmunt et al. (n=117).

The bottom half of Table 11 contains sparse information because several characteristics of the study populations (e.g. the composition, duration and frequency of previous chemotherapy and radiotherapy treatment regimens, and patients' responses to these), which might have a bearing on patients' tolerance of or response to subsequent therapy, were not reported in the primary studies. Although the NICE scope mentions cisplatin-ineligible patients in the second-line setting, the CS and the company's clarification response A25 do not state whether any of the

patients in the included second-line studies were cisplatin-ineligible. Upon checking the publications we found that none of the five included studies provided this information, and only two of the studies reported the proportion of patients who had received prior cisplatin.

By comparing Tables 9, 10 and 11 it can be seen that the first-line atezolizumab cohort had a greater percentage of patients with ECOG PS = 0-1 compared to the first-line gemcitabine + carboplatin studies (79.8% versus 32.0 % in Bamias and 55.5% in De Santis) and a greater percentage with visceral metastases (65.5% versus 44.0% in Bamias and 46.2 in De Santis). The second-line atezolizumab cohort had a greater percentage of patients with visceral metastases than the four comparator studies where this outcome was reported (78.4% versus a range of 61% to 74% in the comparators) but the percentage with liver metastases was within the range of the four comparator studies which reported this (31% versus a range of 15% to 37.5%).

Of the two second-line studies that included best supportive care arms, Bellmunt et al.[21, 45] did not provide a definition of best supportive care, whilst Noguchi et al.[46, 47] stated '*BSC was including palliative radiotherapy, antibiotics, analgesics, corticosteroids, and transfusion*.'

Overall, due to the paucity and inconsistency of the available information on participants' baseline characteristics, it is difficult to be certain whether the second-line studies were adequately homogeneous to be eligible for the company's network meta-analysis; or whether any individual studies had particularly better or worse prognostic characteristics that might suggest a need for further exploration in sensitivity analyses.

Table 11 Baseline characteristics of participants in the second-line comparator studies

	Bellmunt et al.[21, 45]	**Choueiri et al.[50]**	**Kim et al.[48, 49]**	**Lee et al.[38, 39]**	**Noguchi et al.[46, 47]**
Study design	RCT	RCT	Single arm	Single arm	RCT
Regimen (number of participants) *Data in this table are for the arms shown in* ***bold***	VFL + BSC (n=253) **BSC (n=117)**	DOC + vandetanib (n=74) **DOC + PBO (n=75)**	**DOC (n=31)**	**PTX (n=37)**	PPV + BSC (n=39) **BSC (N=41)**
Age, years, median (range)	-	-	64 (40-79)	57 (44-78)	65 (46-81)
Age, proportion ≥65 years	57 (49)	33 (45.8)	15 (48)	-	-
Sex, male, n (%)	-	49 (68.1)	24 (77)	29 (78)	33 (80)
Primary site, n (%): bladder	-	-	18 (58)	26 (70)	-
renal pelvis	-	-	6 (19)	2 (5)	-
Ureter	-	-	7 (23)	9 (24)	-
Urethra	-	-	0 (0)	0 (0)	-
other (unspecified)	-	-	0 (0)	0 (0)	-
ECOG PS 0, n (%)	45 (38)	37 (47.2) [a]	0 (0)	14 (38)	33 (80)
ECOG PS 1, n (%)	72 (62)	38 (52.8)	31 (100)	18 (48)	8 (20)
ECOG PS 2, n (%)	0 (0)	0 (0)	0 (0)	5 (14)	0 (0)
Haemoglobin <10 mg/dl, n (%)	14 (12)	6 (8.5)	7 (23)	-	9 (22)
Any metastases, n (%)	-	-	-	-	41 (100)
Visceral metastases, n (%)	87 (74)	46 (63.9)	19 (61)	23 (62)	-
Liver metastases, n (%)	-	27 (37.5)	10 (32)	11 (30)	Liver/bone 6 (15)
Median follow up, months	45	7.1	37.6	16.6	3.2
1L setting, n (%): metastatic	-	-	29 (94)	30 (81)	15 (37) [b]
perioperative (neo/adjuvant)	-	-	2 (6)	17 (46) [c]	14 (34) [b]
1L response: complete or Partial	-	-	17 (55)	11 (30)	-

stable disease	-	-	9 (29)	5 (14)	-
progressive disease	-	-	3 (10)	6 (16)	-
not evaluable	-	-	2 (6)	15 (41) [d]	-
Cisplatin-ineligible	-	-	-	-	-
Prior cisplatin, %	72.6 [e]	-	94 [f]	- [g]	-
Prior taxane	-	Only PTX (11.1%)	-	None permitted	-
Prior palliative chemotherapy	-	-	-	31 (84)	-
Prior radiotherapy, %	-	20.8	-	-	12
Platinum-free interval	-	-	< 3 months: 36%	< 3 months: 43% ≥6 months: 27%	-
Treatment sequence information	-	Prior therapies: >1: 28 (38.9) >2: 10 (13.9)	Setting: 2L: 26 (84) 3L: 5 (16)	-	Prior therapies: [h] 1: 10 (35) ≥2: 5 (17)

- (dash) indicates data not reported; BSC: best supportive care; DOC: docetaxel; ECOG: Eastern Cooperative Oncology Group; PBO: placebo; PPV: personalised peptide vaccination; PS: performance score; PTX: paclitaxel; VFL: vinflunine

[a] deduced by ERG from eligibility criteria

[b] publication reports prior chemotherapy setting for only 29 of the 41 patients; % values calculated by ERG

[c] publication reports 2 values; ERG believes this one is correct

[d] not evaluable or no evidence of disease

[e] 72.6% received cisplatin and no other platinum; 19.7% carboplatin; 7.7% other (unspecified) platinum combination

[f] of which 87% had received GEM + CIS

[g] 11.1% had received salvage MVAC (methotrexate + vinblastine + doxorubicin + cisplatin) after prior failure of GEM + CIS

[h] prior chemotherapy for advanced bladder cancer

Ongoing trials

The CS reports that there are two ongoing phase III studies, one planned phase III study and one ongoing phase Ib/2 study which have relevance to the current appraisal (CS section 4.14).

- IMvigor 211 (phase III, ongoing) is comparing atezolizumab against investigator's choice of chemotherapy (vinflunine, docetaxel or paclitaxel) in metastatic urothelial carcinoma in a second and third line setting. Completion is expected in November 2017.
- IMvigor 130 (phase III, planned) will evaluate the safety and efficacy of atezolizumab ± gemcitabine/carboplatin compared against gemcitabine/carboplatin in cisplatin-ineligible patients with metastatic urothelial carcinoma in a first-line setting. Completion is expected in June 2020.
- WO29635 (phase Ib/2, ongoing) is a study in non-muscle-invasive bladder cancer of the safety, pharmacokinetics, immunogenicity, patient reported outcomes, and preliminary anti-tumour activity of atezolizumab administered as a single agent and in combination with Bacille Calmette-Guérin vaccine (BCG) in patients with BCG-unresponsive non-muscle-invasive bladder cancer, and in combination with BCG in patients with BCG relapsing, and very high risk, BCG-naive non-muscle-invasive bladder cancer.
- IMvigor 010 (WO29636) (phase III, ongoing) is a study in muscle-invasive bladder cancer on patients selected according to their PD-L1 status which is comparing atezolizumab as an adjuvant therapy against observation alone.

The ERG's update searches (section 3.1.1) did not identify any further ongoing studies.

3.1.4 Description and critique of the approach to validity assessment

Section 4.11.6 of the CS is titled 'Quality assessment of non-randomised evidence' but does not report quality assessment. Section 4.11.7 of the CS is titled 'Methods for assessing risk of bias' but states only that the risk of bias was not assessed for Imvigor 210 as it was a single-arm study. However, an earlier section describing the methods of the network meta-analysis (CS section 4.10.6) describes a method for quality assessment, referring to CS Appendix 8.3 where a quality assessment for the Imvigor 210 study is reported. This was undertaken separately for the two publications for each cohort in the study and also for the clinical study report, and was adapted from a National Institutes for Health (NIH) tool for case series studies.[53] There is no discussion of why this was adapted, or the appropriateness of using a case series study quality

tool when the NIH has tools available for all types of observational studies. One question from the tool that was not applied by the company was ‘were the cases consecutive?’.

The CS provides several summary tables of quality assessment in CS Appendix 8.3. Several of the tables are un-numbered and no explanation of the different tables is provided. The CS conducted quality assessments of the studies used in their network-meta analysis, using three approaches:

- The adapted NIH questions for case series studies, applied to the single-arm studies;
- NICE risk of bias questions for RCTs (also applied to the single-arm studies if an individual question was appropriate);
- Cochrane risk of bias questions for RCTs (also applied to the single-arm studies if an individual question was appropriate).

Given that the CS states that no RCT evidence was identified (CS section 4.11.9) it is unclear why so much emphasis was placed on tools for assessing RCTs, and why both the NICE and Cochrane tools for RCTs were considered necessary. The CS presents the results of the three approaches for assessing study quality separately, which makes it difficult to identify the ‘bottom’ line key issues about study quality. An overall summary of the quality assessments is provided in CS Figure 4 but this arbitrarily classifies studies as having ‘high’, ‘moderate to high’. ‘moderate’, ‘low to moderate’ or ‘low’ quality. These categories are not explained and do not indicate whether there are threats to validity (i.e. risks of systematic errors or lack of generalisability). The ERG and NICE requested explanation of the quality assessment process (clarification question A26) but the company’s response does not define their decision criteria for the different study quality classes.

The company has not used their quality assessment to inform other aspects of the submission, and there is no discussion provided as to whether study quality would affect the eligibility of studies for inclusion in the network meta-analysis. According to the company’s summary in CS Figure 4 there was ‘moderate to high’ heterogeneity within studies which were subsequently included in their meta-analysis. The company stated (clarification response A24) that it was necessary to include studies of heterogeneous populations due to the lack of alternative data.

The ERG has assessed the relevant arm of each included study using the NIH tool because the studies are used as single-arm studies in the company’s analysis and the questions regarding

risk of bias for RCTs are therefore redundant. The ERG considers that three questions from an NIH appraisal tool for before-and-after studies[54] are also relevant (enrolment of all eligible participants, blinding of outcome assessors, and sample size) as these address potential threats to validity or reliability. These additional three questions have therefore been assessed by the ERG for each study.

The ERG's detailed quality assessment of the atezolizumab studies, first-line comparator studies and second-line comparator studies is tabulated in Appendix 3 (Table 54 to Table 56). For Imvigor 210 the ERG has checked the criteria for the study as a whole, not the individual publications as assessed by the company. Quality assessment of the phase I study PCD4989g is not reported in the CS and the company explained that this was because the study only provides descriptive supporting information (clarification response A27).

As shown in Appendix 3, the ERG generally agrees with the company's assessment (where reported) of these studies, although our assessment differs on the question of whether all subjects were comparable. This is because the NIH tool gives no guidance on what this question is assessing; for this question we have assessed studies on how comparable the populations are to the NICE scope.

For the additional ERG questions, it is unclear whether sample sizes in the Bamias et al. and Noguchi et al. studies would be adequate to provide confidence in the findings since they were determined on response rates rather than survival outcomes. If based on sample sizes, confidence in the findings would be highest for the studies by De Santis et al., Bellmunt et al. and Chouieri et al., which had 75-119 participants, than the remaining studies which had only 31-41 participants. None of the studies reported using blinded outcome assessors. Three studies (De Santis 2012, Choueiri 2012, Noguchi 2014) were assessed as enrolling all eligible participants that met the pre-specified entry criteria into the study. The remaining studies did not present enough information to assess this question.

In summary, the ERG believes that the main validity issue for the included studies is the lack of direct head-to-head randomised studies comparing atezolizumab to relevant comparators and uncertainty as to how similar the baseline characteristics of the studies are, given the limited available information.

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 IMvigor201 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. The ERG notes that the CS does not present the numbers for the response-evaluable population in cohorts 1 and 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 source and justification of the selected 10% historical control response rate is not specified in the CS or the associated publications. In response to a clarification question from the ERG and NICE (clarification A36), the company provided a justification for using this response rate as a reference and noted that a recent study of nivolumab in metastatic urothelial carcinoma also used a 10% historical control rate to assess effectiveness.[59] The ERG's clinical expert advisor agreed that the company's justification for using a historical control response rate of 10% is reasonable.

The CS reports the statistical power of Imvigor 210 in section 4.11.4, although this appears to relate to the study as a whole, rather than to the individual cohorts on which the analyses are based.

According to the study protocol, [redacted]

A number of data-cuts were conducted for both cohorts in IMVigor 210, which included interim analyses, primary analyses, updated analyses and follow-up analyses (CS Table 26). The CS clearly states which data-cuts are presented throughout the results section.

For cohort 1, primary analyses were undertaken when the last patient enrolled had a minimum of 6 months follow-up (median follow-up 8.5 months, range 0.2 to 14.3 months) and follow-up analyses were undertaken at 15 months (the company stated in clarification response A34 that median follow-up was 17.2 months, range 0.2 to 23.5 months). Response rates at the primary analysis and 15-month follow-up are presented in the CS; overall survival, progression-free survival and adverse events are presented at the 15-month follow-up only. Interim analyses of cohort 1 were also undertaken but results are not presented in the CS.

For cohort 2, primary analyses were undertaken when the last patient enrolled had a minimum of 6 months follow-up (the company stated in clarification response A34 that median follow-up

was 7.1 months, range 0.23 to 10.61 months) and follow-up analyses were undertaken at 20 months (median 21.1 months, range 0.2 (censored) to 24.5 months). Response rates at the primary analysis and 20-month follow-up are presented in the CS; overall survival, progression-free survival and adverse events are presented at the 15-month follow-up only. 'Updated analyses' of cohort 2 (median follow-up 11.7 months) were also undertaken (CS section 4.11.1) but results are not presented in the CS.

In summary, the company's approach to trial statistics in Imvigor 210 appears appropriate.

Limited details on study PCD4986g are provided in the CS; the company provided the study protocol in response to clarification request A42. Similar methods to IMVigor 210 were used for calculation of overall survival, progression free survival and duration of response.

3.1.7 Description and critique of the company's approach to the evidence synthesis

3.1.7.1 Simulated treatment comparison

In the absence of direct head-to-head comparisons of atezolizumab with the scoped comparators, the company conducted a simulated treatment comparison (STC), also referred to as a 'prediction model' by the CS. An STC can be used to carry out 'unanchored' indirect comparisons, where there is a disconnected treatment network or single-arm studies, and allows adjustment for differences across trials.[60] It is a form of outcome regression and is appropriate for the current evidence base, i.e. where individual patient data are available in one (atezolizumab) population and only aggregate data are available for the comparator populations. The company briefly justifies why they chose to use STC rather than unadjusted (naive) comparisons or a matching-adjusted indirect comparison (MAIC) (CS section 4.10.7). The ERG agrees that STC is the most suitable approach for the available data structure, although, as noted below, the method is strongly dependent on assumptions.[60]

In an STC, a statistical model describing the outcomes in terms of the covariates is fitted to the individual patient data for a treatment of interest (in this case, the intervention, atezolizumab), and used to predict the outcomes that would have been observed in the aggregate target population.[60] This effectively creates an atezolizumab arm within each comparator study, and

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). No justification is given in the CS for any of 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 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 a 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]

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

had conducted to identify relevant prognostic factors (clarification response A16). This search identified liver involvement, ECOG performance status and haemoglobin concentration (<10g/dL) as being relevant prognostic factors based on the literature, and age and sex were thought to be relevant prognostic factors according to the opinion of one Roche internal clinical expert.

The company stated in clarification response A16 that haemoglobin concentration <10g/dL was identified as a prognostic factor but excluded from analysis since trials typically excluded all patients with low baseline haemoglobin. The ERG notes that four out of the five second-line studies that were included by the company for their network meta-analysis did report the proportion of patients with haemoglobin <10g/dL, which ranged from 8.5% to 23% (Table 11).

The ERG's clinical advisor agreed that performance status and age are important prognostic factors; however, they are correlated and the impact of this is unclear. The advisor also suggested that re-treatment interval could be considered as a prognostic factor, if reported.

The company stated in clarification response A18 that IMVigor10 included patients at second and later lines of treatment. Therefore a cut-off of two or more prior chemotherapies was used to assess the impact of having a larger or lower proportion of patients being third-line or more, in contrast to only second-line. This prognostic factor was not selected in the base-case model as it did not improve predictive performance.

The ERG noticed some discrepancies in the proportions of patients with each covariate that are reported in CS Table 17 and we queried these with the company (clarification question A22):

- Data for the proportion with liver metastases in a randomised controlled trial conducted by Bellmunt et al.[21, 45] are available in an abstract but excluded from the analysis. The company stated that this was because the abstract did not meet the inclusion criteria, and they suggested that this omission would not affect the overall results. Liver metastasis data were also omitted for a study conducted by Lee et al.;[38, 39] the company stated this was due to a typographical error but would not affect the overall results.
- The company did not use ECOG performance scores which are reported in a publication for the best supportive care and vinflunine arms of the Bellmunt et al. RCT. The company stated, without providing a rationale, that this was because they instead calculated the weighted mean of covariates across both treatment arms – to adjust for

the study rather than each arm separately, as the prediction model aimed at imputing a hypothetical missing atezolizumab arm for the study as a whole. The ERG notes that this calculation increased the proportion of patients in the best supportive care arm with poorer prognosis (ECOG performance status ≥1) from 0.62 to 0.69, i.e. a slight worsening of the population's prognostic characteristics.

- The ERG queried why the age data for the Bellmunt et al. RCT differ in CS Table 17 from those reported in the publication. The company explained that they imputed the proportion of patients aged >65 years for those studies where only the mean or median were reported (clarification response A21), but exceptionally, for the Bellmunt et al. RCT, the imputed data for age >65 were used as the data available in the paper were unfortunately overlooked.

The CS does not discuss whether the extent of systematic error due to imbalance in unaccounted for covariates is acceptable and no estimates are presented for the degree of likely bias. The CS does, however, note caveats around the estimates and that the outcomes of the network meta-analysis are uncertain, producing 'clinically implausible' results when applied in their economic model without correction.

There is imbalance between the study populations in the four prognostic factors listed in CS Table 17 and this is noted in the CS (CS section 4.10.6). The resulting potential bias reduction that STC would provide compared with an unadjusted (naive) comparison is not reported.

STC prediction models

The CS focuses on the STC for overall survival and progression-free survival outcomes. The company also analysed other endpoints that were not of relevance to the economic model, to more broadly assess the comparative effectiveness of atezolizumab versus other interventions (clarification response A12 and A20). However, the analyses of objective response rate and 12-month survival rate were not used to inform the company's assessment of clinical effectiveness, they did not provide parameters for the economic analysis, and no results for these binary outcomes are provided in the CS. Therefore, only overall survival and progression-free survival analyses are described here.

Cox regression models based on the selected covariates were used to simulate an atrezolizumab arm for each comparator study. The models were fitted to bootstrap samples

from the individual patient data from the atezolizumab Imvigor 210 study. The company tested the fit of nine competing models which included different combinations of the covariates and their interaction terms (CS Table 18). Model selection was based on the best predictive performance as judged using the concordance index (indicating the probability that a patient with longer survival time will have a lower risk score). Model parameters and concordance indices for the overall survival outcome are given in CS Table 20 for first-line treatment comparisons and in CS Table 22 for second-line comparisons. For both the first-line and second-line treatment comparisons the company chose the four-covariate model (age, sex, liver metastasis, performance status) without interactions, as adding interactions did not improve fit. For second-line comparisons the company tested including the number of prior chemotherapies (proportion of patients receiving ≥2 prior chemotherapies) as a fifth covariate but this did not improve fit. The ERG notes that, both for first-line and second-line comparisons, the model fit based on the concordance index did not differ tangibly between the four-covariate model and a model which included only liver metastases and performance status, although this is not discussed in the CS.

The Cox models generated predicted log-hazards over time with their associated standard errors and these were used as predicted atezolizumab data points in the network meta-analysis.

Summary of the ERG's appraisal of the STC

The NICE Decision Support Unit provides recommendations on the methods of simulated treatment comparison analysis.[60] A table showing the ERG's appraisal of the company's approach compared against these recommendations is provided in Appendix 4.

In summary, the ERG has the following concerns about the company's approach to the STC:

- Relatively few covariates were used in the prediction model;
- The selection of the covariates in the prediction model is not well justified and is subject to a number of uncertainties.
- The company used a single data calculation method for imputing missing data; multiple imputation would have been preferable to clarify uncertainty around the plausibility of imputed data values.
- The cumulative impact of small errors and inconsistencies in the data is unclear.

3.1.7.2 Network meta-analysis

Analysis of survival data depends on the assumption of proportional hazards being satisfied, and violations of the assumption can lead to severely biased estimates of expected survival.[63] The validity of the proportional hazards assumption has not been ascertained for comparison of atezolizumab against traditional chemotherapy in urothelial carcinoma. Based on appraisals of immunotherapies in melanoma and non-small cell lung cancer, the company reasons, appropriately, that the proportional hazards assumption is unlikely to hold for comparisons involving atezolizumab (CS section 4.10.9).

Fractional polynomial models

Given the possible violation of the proportional hazards assumption, the company developed fractional polynomial models for their network meta-analysis. Whereas traditional survival analysis represents the treatment by a single parameter, i.e. the hazard ratio, the fractional polynomial approach models the hazard over time and represents the treatment effect with multiple parameters. As such, fractional polynomial models can be an appropriate way to model survival data where the proportional hazards assumption is violated, and are suitable for comparisons where both individual patient data and aggregate patient data are available.[63]

The company conducted their network meta-analysis using a Bayesian framework. The time-to-event data for the comparators were obtained by digitising Kaplan-Meier curves for overall survival and progression-free survival reported in the included studies. The survival proportions for each monthly time interval were extracted and used to calculate the number of patients at risk at the start of each interval and the incident number of deaths. The CS briefly reports that the event probability for each time interval was obtained from a binomial likelihood distribution based on the underlying hazard function modelled by the fractional polynomial analysis. The predicted log-hazard for each comparison with atezolizumab at multiple time points was fitted with a normal distribution. The approach described by the company is broadly consistent with an approach for using fractional polynomial models in network meta-analysis as outlined by Jansen (2011),[63] except that fewer details of the analysis are provided in the CS. The CS does not provide any data on the number of events and patients at risk that they obtained from the included studies, but this information was provided by the company in clarification response A32.

Three orders of fractional polynomial models were considered: zero-order, which corresponds to an exponential model and assumes proportional hazards; first-order, which corresponds to the Weibull model (where exponent P1=0) or the Gompertz model (where exponent P1=1); and second-order with exponents P1 and P2 (giving possible combinations of P1=P2=0, P1=0, P2=1, and P1=P2=1). According to the CS and the company's response to clarification question A13, the zero-order fractional polynomial model was included to allow assessment of the proportional hazards assumption ('*e.g. through the deviance information criterion (DIC)), which was possible as the model was fitted to the same data as the more complex models'*.

Study heterogeneity assumptions

The company states that there was a 'limited evidence base' with which to estimate the between-study standard deviation and they therefore used informative prior distributions for the fixed-effects model parameters, taken from Turner et al.,[64] to account for between-study heterogeneity (CS Table 19). The ERG agrees that this is an appropriate approach. However, whilst the CS states that three priors (informative, weakly informative and vague) were compared in sensitivity analyses, no sensitivity analysis results are reported. These were subsequently provided by the company, for second-line comparisons only (clarification response A30).

The CS states that fixed-effects models were first fit, with random-effects models subsequently fit if the data allowed. It is important to consider the plausibility of model assumptions rather than basing decisions solely on model fit,[65] but the choice of fixed or random effect models was justified only on model fit (in clarification response A28 the company stated that random-effects models were included to allow for between-study heterogeneity; however, fixed-effects were subsequently chosen based on model fit). The process of assessing model fit is not clearly explained in the CS, which mentions that, in addition to the deviance information criterion, 'additional criteria' were used, but these are not specified (CS Section 4.10.10).

According to the CS, a fixed-effects model was used for the first-line treatment comparisons. The ERG requested an explanation from the company via NICE as to why a random-effects model was not used (clarification response A29). The company provided DIC values for comparisons of the fixed-effects and random-effects models for each of the three between-study heterogeneity priors and explained that the choice of fixed-effects model was based on the DIC. For the second-line treatment comparisons, the CS states that a random-effects model was

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 zero-order fractional polynomial model, as this had the lowest DIC among three fixed-effects models that were compared (CS Table 21), indicating that the more complex first-order fractional polynomial models did not perform better. 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).

The CS states that the clinically implausible values of hazard ratios are likely to reflect the sparse nature of the evidence base and results of the network meta-analysis are therefore subject to uncertainty. Hazard ratios for overall survival were employed in the economic analysis (subject to capping). However, the CS states that hazard ratios from network meta-analysis of progression-free survival could not be used in the economic analysis due to being clinically implausible (CS section 4.10.11) and results of these analyses are provided separately in CS Appendix 8.5. Given that the analyses of progression-free survival were not used by the company to support either the clinical effectiveness or cost-effectiveness of atezolizumab, these are not considered in detail in the current report.

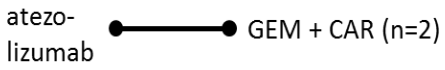
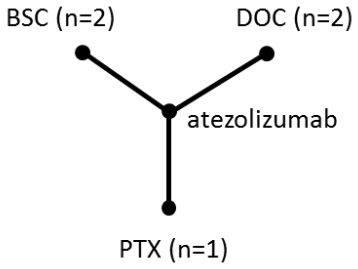
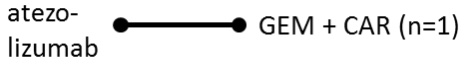
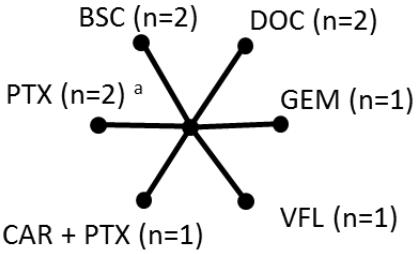
Network structure

The CS does not present network diagrams; however, the networks are simple (summarised in Table 12). As all comparisons are against atezolizumab, there are no indirect comparisons involved. Results of the network meta-analysis for each comparison would therefore be identical to those obtained by performing separate pairwise comparisons under the same statistical model (confirmed by the company for the fractional polynomial model in clarification responses A14 and A15). As noted above (section 3.1.2) the company has been inconsistent in applying their eligibility criteria such that they have included more comparators for their analysis of progression-free survival than for their analysis of overall survival.

Output of the network meta-analysis

The fractional polynomial analysis generates results which reflect the time course of the log-hazard function and as such can be expressed as log-hazard function curves and their parameters (intercept and slope). An explanation of the relationship between the log hazard function and hazard ratio is given below in section 4.3.5.2. When reporting the results of the network meta-analysis (see section 3.3.6), the company does not provide any guidance on the clinical interpretation of these parameters or any discussion of any of the clinical effectiveness results from the network meta-analysis.

Table 12 Summary of simulated treatment comparisons in the network meta-analysis

Outcome	First-line treatment (cohort 1 in atezolizumab study Imvigor 210)	Second-line treatment (cohort 2 in atezolizumab study Imvigor 210)
Overall survival (informs company's economic model)		
Progression-free survival (does not inform company's economic model)		

BSC: best supportive care; CAR: carboplatin; GEM: gemcitabine; DOC: docetaxel; PTX: paclitaxel; VFL: vinflunine

[a] nanoparticle albumin bound paclitaxel in one study

Summary of the ERG's appraisal of the network meta-analysis

The ERG has assessed the company's network meta-analysis using a critical appraisal checklist which we have based on published reporting guidelines (Jansen et al.,[65] inter alia). Our appraisal is provided in Appendix 5.

In summary, the ERG has the following concerns regarding the company's approach to the network meta-analysis:

- The simulated treatment comparison which informs the network meta-analysis has several limitations (as noted above we identified concerns around the selection of covariates and handling of missing data; see also Appendix 4);
- It is unclear whether the included studies were adequately homogeneous to permit valid meta-analysis; some aspects of prior therapies received by patients were not reported in the primary studies, and best supportive care was not adequately defined;

- A lack of sensitivity analyses to test the robustness of both the simulated treatment comparison and the network meta-analysis methods means that specific uncertainties are not propagated through to aid interpretation of the final clinical effectiveness results;
- The meta-analysis produced clinically implausible hazard ratios (which, as explained in section 4.3.5, resulted in the need for capping of the hazard ratios in the economic analysis);
- The meta-analysis is not used to provide any evidence for the clinical effectiveness of atezolizumab.

3.2 Overall summary statement of the company's approach

A summary of the ERG's appraisal of the company's approach to the evidence synthesis is given in Table 13.

The company conducted extensive searches which appear to have identified all relevant studies. The eligibility screening process is described in stages, making it somewhat difficult to follow, but the inclusion/exclusion criteria are deducible. The eligibility criteria have not been consistently applied, although this does not appear to have resulted in any major inclusion/exclusion errors. The screening process is described only briefly in the CS, but in clarification response A9 the company stated that screening was conducted by two reviewers.

Whilst the overall systematic review process appears reasonable, there are several issues with the meta-analysis methods applied by the company (simulated treatment comparison and network meta-analysis) which mean that the results of the analyses are uncertain. These are explained in detail in section 3.1.7 and summarised in section 3.4.

The submitted evidence is consistent with the decision problem

Table 13 Quality assessment (CRD criteria) of the CS review

CRD Quality Item: score Yes/ No/ Uncertain with comments	
1. Are any inclusion/exclusion criteria reported relating to the primary studies which address the review question?	**Yes.** But these were not applied consistently, with some studies being excluded for reasons other than those stated.
2. Is there evidence of a substantial effort to search for all relevant research? i.e. all studies identified	**Yes.** The search was broad and comprehensive and a detailed search strategy was provided in a clarification request. The ERG identified 18 studies that appeared to be eligible but were not cited or referenced in the CS. The company clarified that 16 of these had been identified, screened and excluded and two were published later than the company's searches. Overall, no relevant studies appear to have been missed.
3. Is the validity of included studies adequately assessed?	**Partly.** The company used a NIH checklist for single-arm studies which does not cover some potential biases. Decisions on study quality are summarised narratively and difficult to interpret in relation to whether there are threats to internal or external validity. The quality assessment does not appear to inform any decisions about study eligibility.
4. Is sufficient detail of the individual studies presented?	**Partly.** Yes for the atezolizumab study, but no details of the comparator studies are provided in the CS. Some details of

	study drug dosing, study design, eligibility criteria and age, sex and ethnicity (but not other baseline characteristics) were provided by the company in clarification response A25.
5. Are the primary studies summarised appropriately?	**Partly.** Yes for the atezolizumab study, but no details of the comparator studies are provided in the CS. The company conducted a simulated treatment comparison but the CS does not summarise the characteristics, or specify the sample size, of the simulated study arms; limitations of the available data and the need for assumptions mean that the results may not be reliable.

3.3 Results

Results from the two cohorts of the Imvigor 210 study and the PCD4989 study are summarised in CS Sections 4.11.10 and 4.11.11. The main source of evidence is from the Imvigor 210 study, where efficacy results are presented in the CS for various data-cuts (CS Table 26). The ERG has reproduced the most recent data cut for each cohort. For cohort 1 (first-line treatment), this was at 15 months (4th July 2016 data cut with a median follow-up of 17.2 months [range 0.2 to 23.5 months]; company's clarification response A34) and with 14% of participants remaining on treatment. For cohort 2 (second-line treatment), this was at 20 months follow-up (4th July 2016 data cut with a median follow-up 21.1 months [range 0.2 to 24.5 months]; company's clarification response A34). The ERG has focused on results from the independent review facility assessment of outcomes; investigator-assessed outcomes are only reported where independent review facility assessments are unavailable. Where available, all data presented in the CS have been checked with the publications and the clinical study report.

Patients in PCD4989g received second-line treatment with atezolizumab, but not with the licensed dose, and therefore the results from this study should be interpreted with caution. A summary of study PCD4989g and its clinical effectiveness results is provided in Appendix 2.

3.3.1 Effectiveness of first-line atezolizumab

Results are reported for cohort 1 (first-line therapy) in CS section 4.11.10.2.

Survival

Overall survival and progression-free survival were secondary outcomes in the Imvigor 210 study. The median overall survival in cohort 1, assessed by independent review facility using RECIST v1.1 was 15.9 months, and 57.2% of patients had 12-month survival (Table 14). The Kaplan-Meier overall survival curve for first-line atezolizumab treatment (cohort 1) in Imvigor 210 (CS Figure 19) is shown in Figure 1.

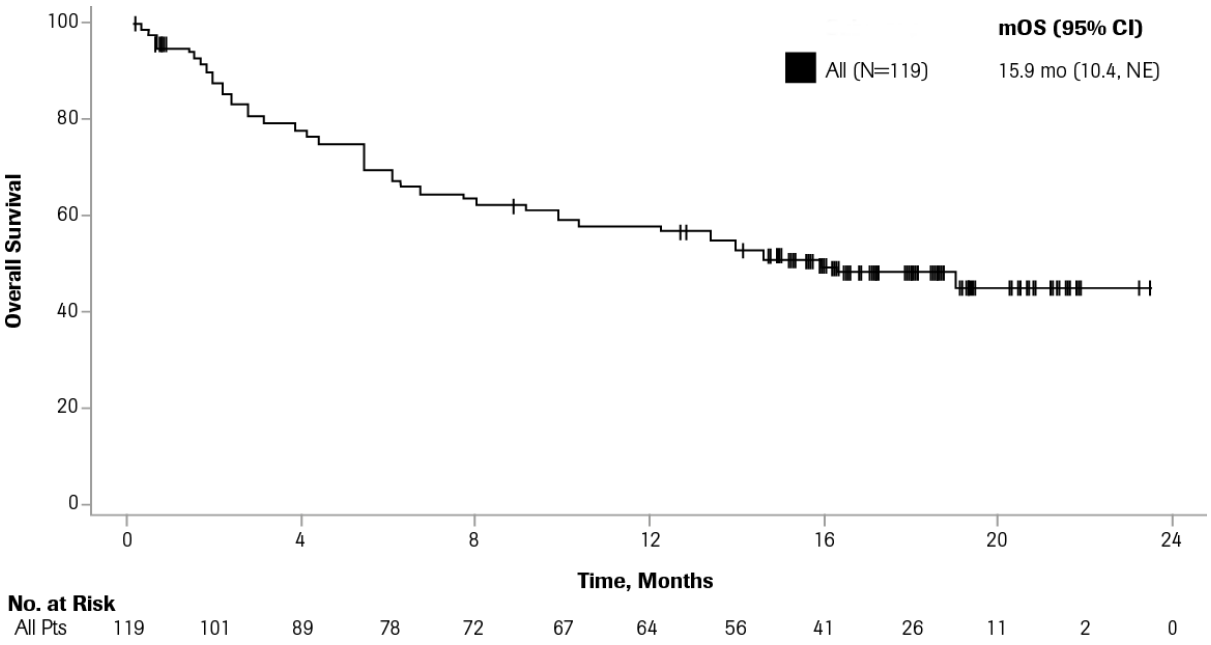


Figure 1 Kaplan-Meier overall survival curve for first-line atezolizumab (Imvigor 210 cohort 1)

Progression free survival at the 15 month analysis was 2.7 (95% CI 2.1, 4.2) months (Table 14). The CS does not report a Kaplan-Meier curve for first-line progression-free survival.

Table 14 Survival outcomes for cohort 1 of Imvigor 210

Outcome (95% CI) **(RECIST v1.1; IRF assessed)**	**Imvigor 210 cohort 1** **All patients, N = 119**
Overall survival, median, months	15.9 (10.4, NE)
12 months survival, %	57.2% (48.2%, 66.3%)
Progression-free survival, median, months	2.7 (2.1, 4.2)

CI: confidence interval; IRF: independent review facility; NE: not estimable

Response rates

Objective response rate as assessed by the independent review facility using RECIST v1.1 was the primary endpoint in cohort 1 of the Imvigor 210 study. Results are reported in CS Section 4.11.10.2. At the 15-month follow-up analysis, 22.7% achieved an objective response, and a complete response was seen in 9.2% of patients (Table 15). The CS states in Section 4.11.10.2 that the 15-month follow-up analysis confirms the findings from the primary analysis (data cut at a median follow-up of 8.5 months) that the objective response rates exceed the 10% historical control.

Median duration of response had not been reached in cohort 1 of Imvigor 210. CS Section 4.11.10.2 and CS Figure 18 show that the majority of responses were longer than one year, with many still ongoing at the 15-month data cut. The median treatment duration was 15 weeks (range 0 to 102 weeks).

Table 15 Response outcomes for cohort 1 of Imvigor 210

Outcome (95% CI) **(RECIST v1.1; IRF assessed)**	**Imvigor 210 cohort 1** **All patients, N = 119** [a]
ORR, %	22.7 (15.52, 31.27)
Complete response, %	9.2
Median time to onset of first response, months	2.1 (range 1.8 – 10.5)

CI: confidence interval; IRF: independent review facility; ORR, objective response rate
[a] Includes 20 patients with missing/unevaluable responses. All treated patients had measurable disease at baseline per investigator-assessed RECIST v1.1.

3.3.2 Effectiveness of second-line atezolizumab

Results are presented for cohort 2 of the Imvigor 210 study in CS Section 4.11.10.3.

Survival

Overall survival and progression-free survival were secondary outcomes in cohort 2 of the Imvigor 210 study (CS Section 4.11.10.3). At the 20-month follow-up assessment in Imvigor 210 cohort 2, the overall survival was 7.9 months as assessed by the independent review facility using RECIST v1.1 (Table 16). The Kaplan-Meier overall survival curve for second-line atezolizumab treatment in Imvigor 210 (CS Figure 21) is shown in Figure 2.
Twelve month survival was 36.9% and median progression free survival 2.1 months. The CS does not report a Kaplan-Meier curve for second-line progression-free survival.

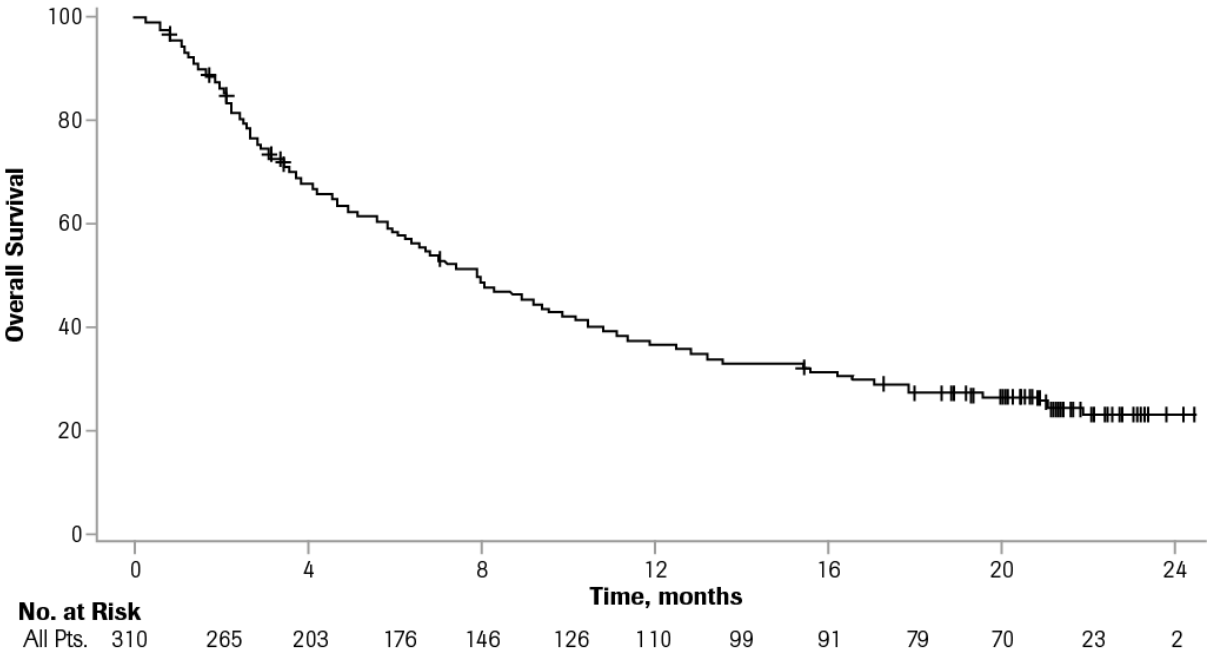


Figure 2 Kaplan-Meier overall survival curve for second-line atezolizumab (Imvigor 210 cohort 2)

Table 16 Survival outcomes for cohort 2 of Imvigor 210

Outcome (95% CI) **(RECIST v1.1; IRF assessed)**	**Imvigor 210 cohort 2: All patients, N = 310**
Overall survival, median, months	7.9 (6.7–9.3)
12 months survival, %	36.9% (31.4–42.3)
Progression-free survival, median, months	2.1 (2.1–2.1) [a]

CI: confidence interval; IRF: independent review facility; NE: not estimable
[a] ERG unclear why confidence interval as reported in the CS has zero range

Response rates

Objective response rate as assessed by the independent review facility using RECIST v1.1 was a co-primary endpoint in cohort 2 of the Imvigor 210 study, alongside objective response rate assessed by the investigator using modified RECIST criteria. Results are reported in CS Section 4.11.10.3. At the 20-month follow-up analysis, objective response rate was 15.8 months by independent review facility assessment and a complete response was seen in 6.1% of patients (Table 17). The median treatment duration was 12 weeks (range 0 to 104 weeks).

Median duration of response had not been reached in cohort 2 of Imvigor 210. The maximum duration of response at the latest follow-up analysis (which had median follow-up 21.1 months) was 22.6 months. The median time to response was 2.1 months (95% CI 2.0, 2.2). At the time of the 12-month analysis 65.3% of participants were ongoing with a response (not reported for the 20-month analysis).

Table 17 Response outcomes for cohort 2 of Imvigor 210

Outcome (95% CI) **(RECIST v1.1; IRF assessed)** [a]	**Imvigor 210 cohort 2: All patients, N = 310**
ORR, per RECIST, %	15.8 (11.9–20.4)
ORR per immune-modified RECIST, %	19.7 (15.4–24.6)
Complete response, %	6.1% (3.7–9.4)
Duration of response, maximum months	22.6

CI: confidence interval; IRF: independent review facility; ORR, objective response rate
[a] CS Table 36 implies both RECIST and modified RECIST assessments were done by the IRF; the company's response to clarification request A35 states, however, that the standard RECIST criteria were applied by the IRF whereas the modified RECIST criteria were investigator-assessed.
.

3.3.3 HRQoL results

The Imvigor 210 study and the PCD4989g study did not include HRQoL outcomes.

3.3.4 Sub-group analysis results

Response rate outcomes for atezolizumab in the Imvigor 210 study are reported according to PD-L1 expression subgroups for the first-line cohort (CS section 4.11.10.2) and second-line cohort (CS section 4.11.10.3). These subgroups are not discussed here as they are not reported for survival outcomes.

The CS states (narratively only) that results for subgroups defined by demographic and baseline characteristics showed positive results on objective response rates (CS sections 4.11.10.2 and 4.11.10.3). In response to a clarification request by the ERG and NICE, the company noted that because of different data cuts the subgroup results are inconsistent with those reported in the CS (clarification response A39). The ERG agrees that the subgroup results for cohort 1 broadly agree with the narrative summary in the CS, but that there is more uncertainty in the subgroup data than in the whole-population analyses. For cohort 2 the results data provided by the company in their clarification response are not structured by baseline characteristics and the ERG has not been able to compare these with the narrative summary in the CS.

The NICE scope and company's decision problem do not specify any subgroups.

3.3.5 Effectiveness of comparators

3.3.5.1 First-line comparators

The CS does not provide effectiveness results for the two studies of first-line comparator treatments which were included in the company's network meta-analysis. The ERG has summarised these from the study publications in Table 18 (for the company's meta-analysis results see section 3.3.6 below).

Overall survival

Median overall survival on first-line gemcitabine + carboplatin was 9.3 months in the De Santis et al. study and 9.8 months in the Bamias et al. study (Table 18). The Kaplan-Meier curves for

overall survival on first-line gemcitabine + carboplatin in these studies (CS Figures 8 and 9) are included below in Figure 3 and Figure 4.

Progression-free survival

Median progression-free survival on first-line gemcitabine + carboplatin was 4.4 months in the Bamias et al. study and 5.8 months in the De Santis et al. study. A Kaplan-Meier curve for progression-free survival on first-line gemcitabine + carboplatin in the study by Bamias et al. is reported in CS Appendix 8.5 (not reproduced here).

Response rates

Objective response rates on first-line gemcitabine + carboplatin ranged from 24% to 41.2%, but the rate of complete responses was only 3% to 3.4% (Table 18)

Table 18 Survival and response rates in first-line comparator studies

Outcome (95% CI)	De Santis 2012[44] Gemcitabine + carboplatin	Bamias 2007[42] Gemcitabine + carboplatin
Overall survival, median, months	9.3 (CI not reported)	9.8 (4.7, 14.9)
Progression free survival, median, months	5.8 (CI not reported)	4.4 (1.03, 7.75)
Overall response, % [a]	41.2 (CI not reported)	24 (11 to 41)
Complete response, %	3.4 (CI not reported)	3 (0 to 15)
Partial response, %	37.8 (CI not reported)	21 (9 to 38)

[a] referred to as objective response (Bamias) or overall response (De Santis)

3.3.5.2 Second-line comparators

The CS does not provide effectiveness results for the five studies of second-line comparator treatments which were included in the company's network meta-analysis. The ERG has summarised these from the study publications in Table 19 (for the company's meta-analysis results see section 3.3.6 below).

Overall survival

Median overall survival on second-line best supportive care was reported in two studies (Bellmunt et al. and Noguchi et al.) and ranged from 4.1 months to 4.6 months (Table 19). The

Kaplan-Meier curves for overall survival on best supportive care in these studies (CS Figures 10 and 11) are included below in Figure 5 and Figure 6.

Median overall survival on second-line docetaxel was reported in two studies (Chouieri et al. and Kim et al.) and ranged from 7.03 months to 8.3 months (Table 19). Note that the docetaxel arm in the Chouieri et al. study was a combination of docetaxel + placebo. The CS does not comment on the nature of the placebo or whether incorporation of a placebo in the arm would affect interpretation. The Kaplan-Meier curves for overall survival on second-line docetaxel in these studies (CS Figures 12 and 14) are included below in Figure 7 and Figure 8.

Median overall survival on second-line paclitaxel, reported in one study (Lee et al.), was 6.5 months (Table 19). The corresponding Kaplan-Meier curve (CS Figure 13) is included below in Figure 9.

Progression-free survival

Median progression free survival on second-line best supportive care was 1.8 months in the Noguchi et al. study (not reported for the Bellmunt et al. study) (Table 19). The Kaplan-Meier curves for progression-free survival on best supportive care in these studies are reported in CS Appendix 8.5 (not reproduced here).

Median progression-free survival on second-line docetaxel in the studies by Chouieri et al. and Kim et al. ranged from 1.4 months to 1.58 months (Table 19). The Kaplan-Meier curves for progression-free survival on second-line docetaxel in these studies are reported in CS Appendix 8.5 (not reproduced here).

Median progression-free survival on second-line paclitaxel, reported in one study (Lee et al.) was 2.7 months (Table 19). The corresponding Kaplan-Meier curve is reported in CS Appendix 8.5 (not reproduced here).

Response rates

No responses were achieved in best supportive care study arms. The overall response rate on second-line docetaxel ranged from 6% to 7%, but the rates of complete responses were not

reported. Overall response rate was higher in the paclitaxel study, at 21%, but only 3% of the patients receiving paclitaxel were complete responders (Table 19).

Table 19 Survival and response rates in second-line comparator studies

Outcome (95% CI)	Noguchi 2016[47] BSC	Bellmunt 2013[21] BSC	Choueiri 2012[50] Docetaxel	Kim 2016[49] Docetaxel	Lee 2012[39] Paclitaxel
Overall survival median, months	4.1 (2.8, 6.9)	4.6 (4.1, 6.6)	7.03 (NR)	8.3 (5.9, 10.6)	6.5 (5.0, 8.0)
Progression free survival, median, months	1.8 (1.3, 2.3)	NR	1.58 (NR)	1.4 (1.3, 1.6)	2.7 (0.9, 4.6)
Overall response, % [a]	0	0	7 (NR)	6 (1 to 21)	21 (7 to 34)
Complete response, %	0	0	NR [a]	NR	3 (NR)
Partial response, %	0	0	NR	NR	18 (NR)

BSC: Best supportive care; NR: not reported
[a] referred to as objective response (Kim, Lee) or overall response (Bellmunt, Chouieri)

3.3.6 Network meta-analysis results

As explained above (section 3.1.7), the ERG is concerned that the company's approach to network meta-analysis enables violation of the proportional hazards assumption. The results of the analysis may therefore be incorrect and should be considered uncertain. However, we have reproduced the company's results here for consideration.

The company presents the results of the network meta-analysis as hazard ratios and also visually in a series of Figures, in which the survival curves for the simulated atezolizumab arm, the observed atezolizumab arm in Imvigor 210, and the comparator arm can be compared for each treatment comparison.

Hazard ratios

For first-line comparison of atezolizumab against gemcitabine + carboplatin the CS reports a hazard ratio for overall survival of 0.6 (credible interval 0.47 to 0.82), i.e. in favour of atezolizumab (CS section 4.10.11.1). However, the time point to which this hazard ratio refers is not stated.

For second-line comparisons of atezolizumab against best supportive care, docetaxel and paclitaxel, the CS provides charts showing plots of the posterior median log hazard ratio against time in CS Figures 15-17. These are for a fixed-effects analysis. The curves (not reproduced here) appear to suggest that the log hazard ratio is time-invariant for best supportive care and paclitaxel but time-dependent for docetaxel. However, the CS does not provide any interpretation of these curves. Wide credible intervals indicate that there is considerable uncertainty in the predicted log hazard ratios, especially after month 15.

Note that interpretation of time-dependent hazard ratios can lead to implausible values of the hazard ratio at given time points; for this reason the company capped the overall survival hazard ratio estimates obtained from the network meta-analysis to enable their inclusion in the economic model (see section 4.3.5.2).

The company reports the parameters of the log-hazard function (slope, intercept, and their correlation) for second-line comparisons (CS Table 24) but not for first-line comparisons. The CS does not explain how they should be interpreted in order to draw conclusions on the effectiveness of atezolizumab. Parameters of the log-hazard function are not reproduced here but those which inform the economic analysis are discussed in section 4.3.

Survival curves

Curves for overall survival are provided in CS Figures 8 to 14. Curves for progression-free survival are provided in Figures 1 to 10 in CS Appendix 8.5. The Figures for overall survival are reproduced here for consideration. Any visual comparison of the observed atezolizumab curves against the corresponding comparator curves would effectively be a naive (unadjusted) comparison since differences in the studies' characteristics are not taken into account.

Overall survival: first-line

Overall survival on first-line gemcitabine + carboplatin, compared with first-line atezolizumab (two studies), is shown in Figure 3 and Figure 4.

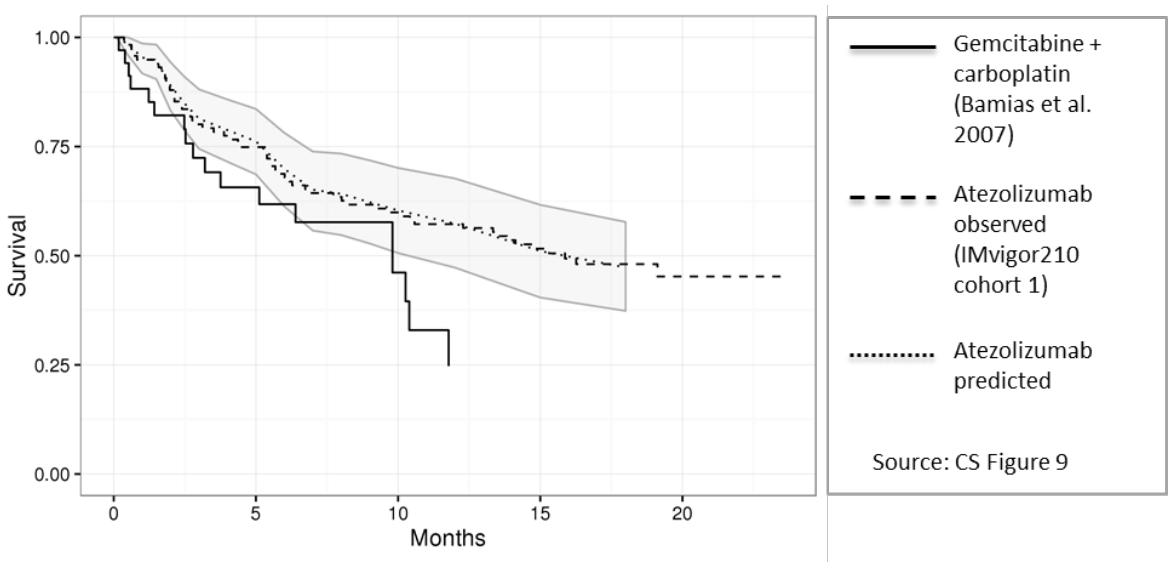


Figure 3 Overall survival curves for first-line gemcitabine + carboplatin (Bamias et al. 2007) and atezolizumab

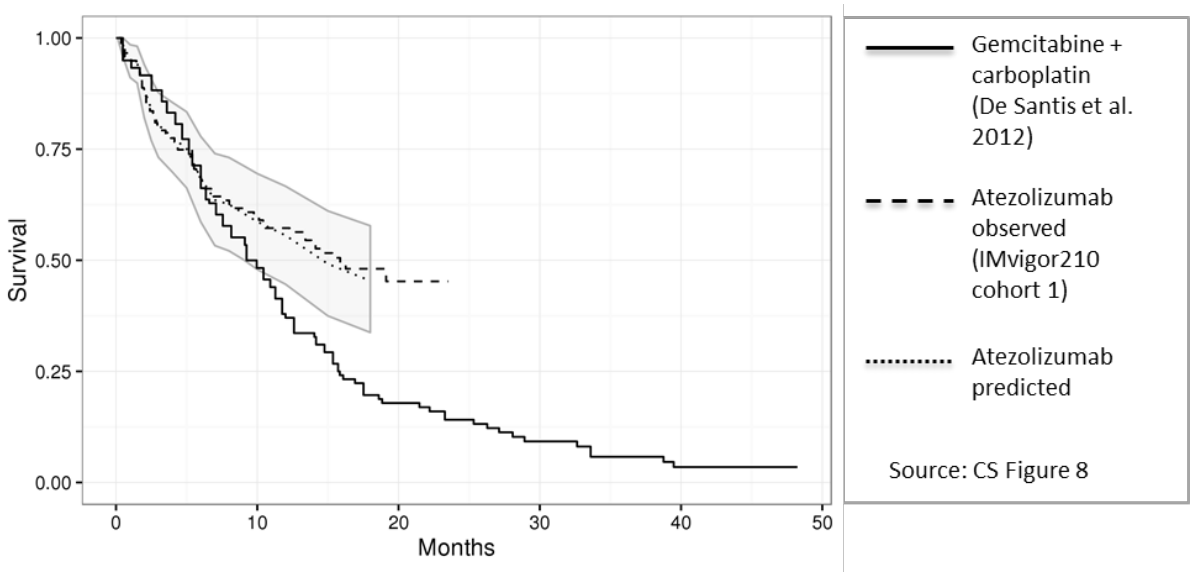


Figure 4 Overall survival curves for first-line gemcitabine + carboplatin (De Santis et al. 2012) and atezolizumab

Overall survival: second-line

Overall survival on second-line best supportive care, compared with second-line atezolizumab (two studies), is shown in Figure 5 and Figure 6.

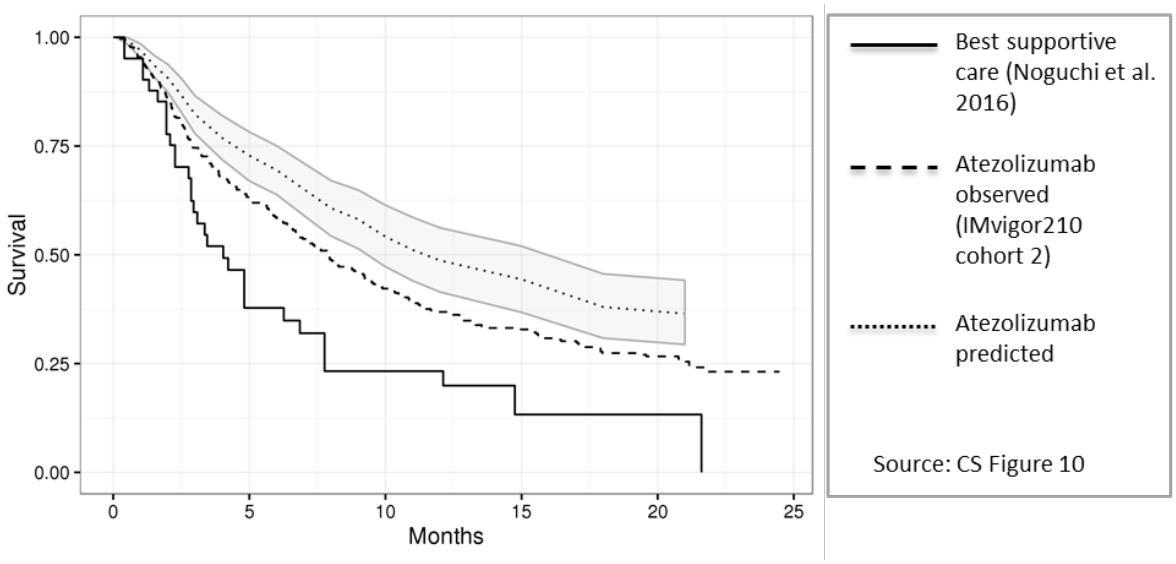


Figure 5 Overall survival curves for second-line best supportive care (Noguchi et al. 2016) and atezolizumab

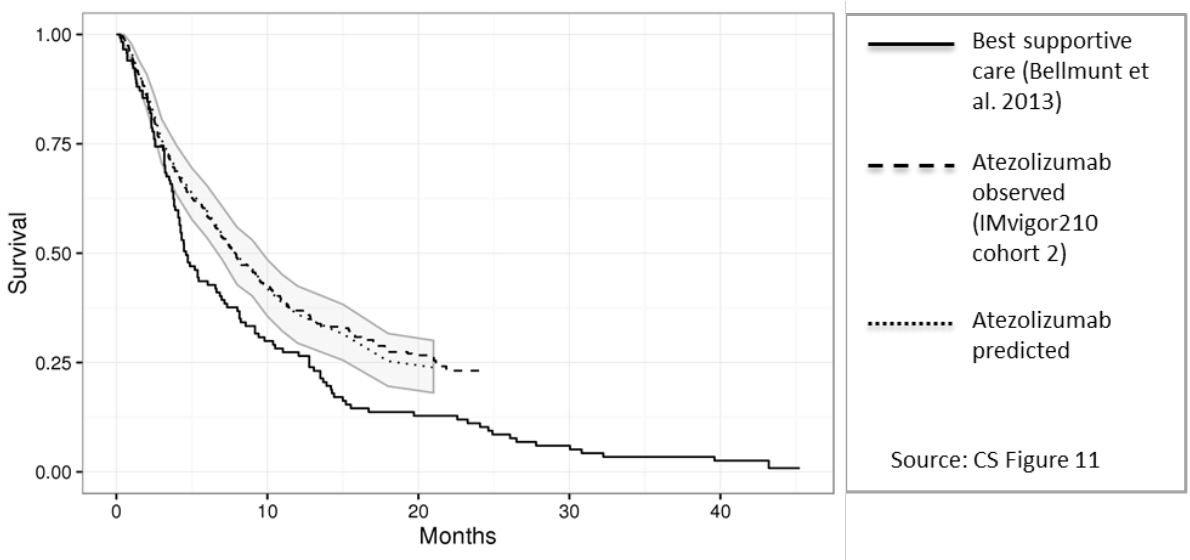


Figure 6 Overall survival curves for second-line best supportive care (Bellmunt et al. 2013) and atezolizumab

Overall survival on second-line docetaxel, compared with second-line atezolizumab (two studies), is shown in Figure 7 and Figure 8.

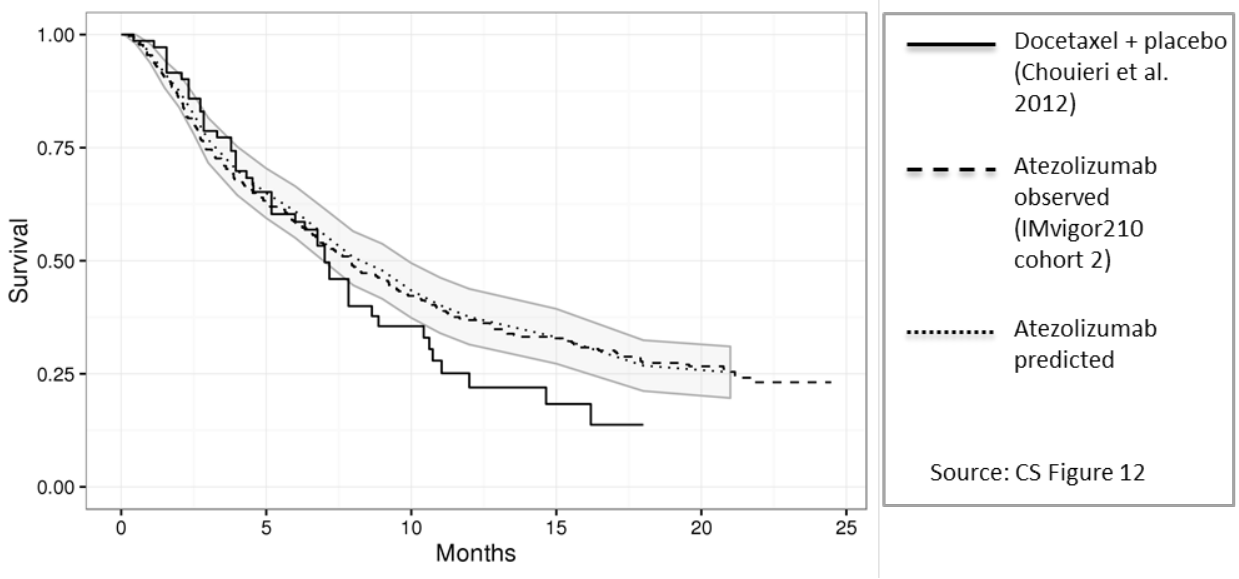


Figure 7 Overall survival curves for second-line docetaxel (Choueiri et al. 2012) and atezolizumab

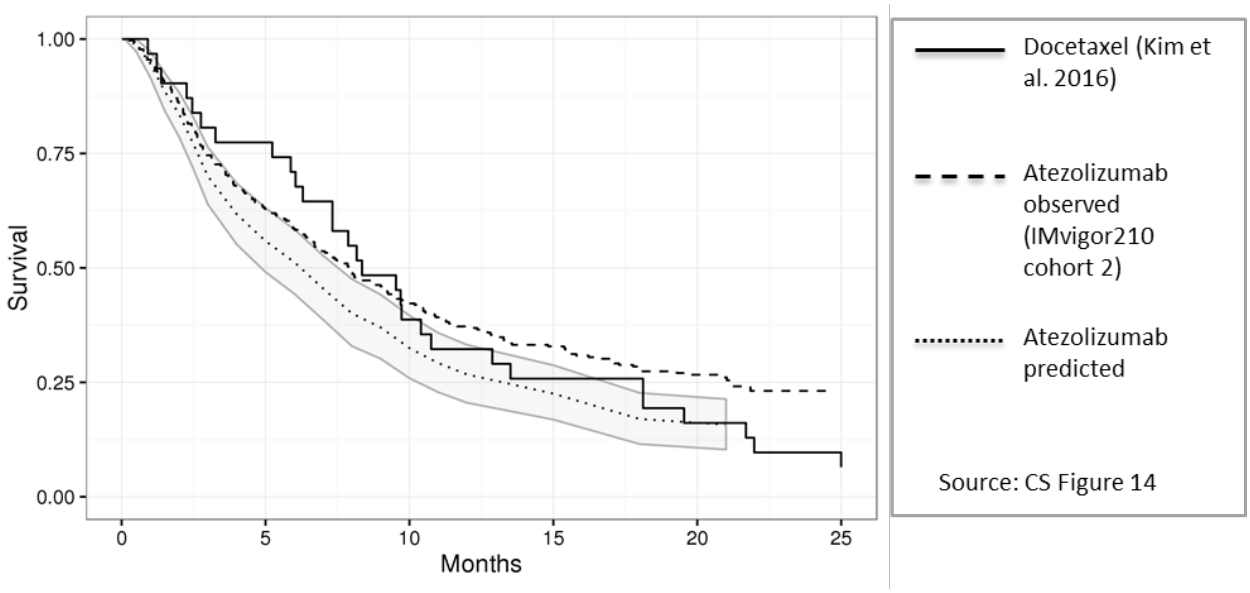


Figure 8 Overall survival curves for second-line docetaxel (Kim et al. 2016) and atezolizumab

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

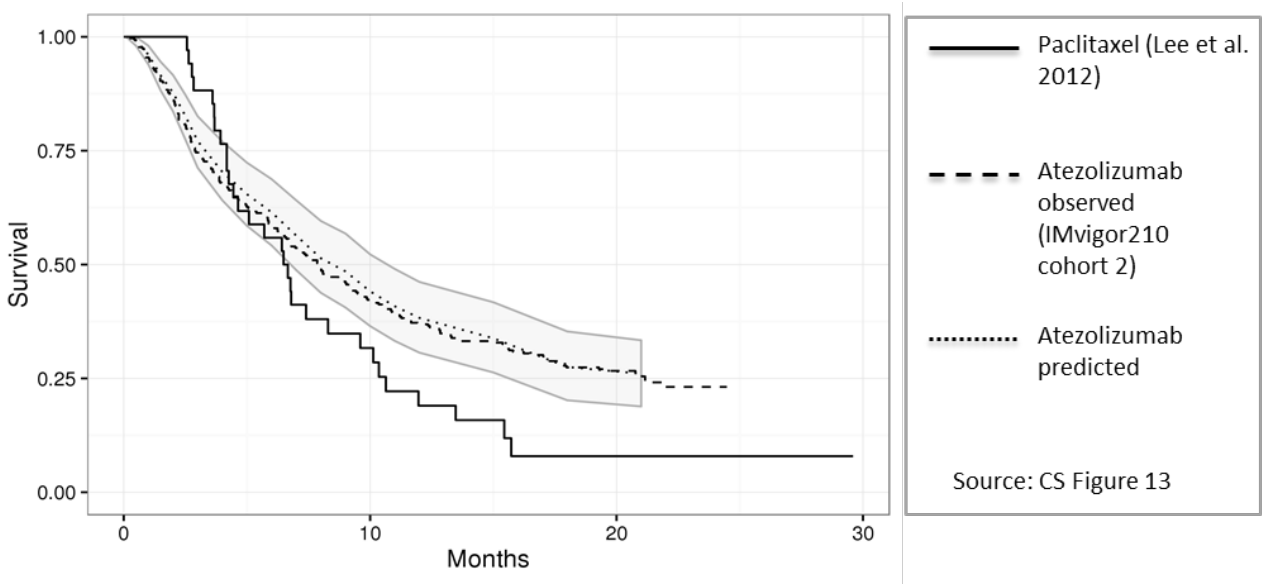


Figure 9 Overall survival curves for second-line paclitaxel (Lee et al. 2012) and atezolizumab

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 side effects, affecting at least 20% of the patients, were fatigue (tiredness), decreased appetite, nausea (feeling sick), and dyspnoea (shortness of breath).[66] 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, were [redacted] (data from

updated clinical study report). The CS states that these were related to underlying disease in most instances.

Grade 3-4 events were experienced in 45%-60% of participants.

Treatment-related adverse events

The CS does not specify how treatment-related was defined. The Imvigor 210 study publication by Balar et al. 2017[40] states only that this was '*deemed to be related to treatment by the investigator*'. No other information about the definition of 'treatment-related' is given in the clinical study reports.

The rate of treatment-related adverse events across the three cohorts was 66-71%.Treatment related serious adverse events were experienced in 10.1% and 12.3% of participants in the two cohorts of Imvigor 210 respectively and in 5.3% of participants in study PCD4989g (Table 20). Rates for individual treatment related serious adverse events were generally low. Most frequently reported across the three cohorts were diarrhoea (2.5% in cohort 1, 0.3% in cohort 2, 0 in PCD4989g), renal failure (1.7% in cohort 1, 0 in cohort 2 and PCD4989g) and pyrexia (0.8% in cohort 1, 0.6% in cohort 2 and 2.1% in PCD4989g) (data provided in clarification response A43). Of the Grade 3-4 adverse events, 16–18% were deemed to be treatment related.

The most commonly reported treatment-related grade 3-4 adverse events in the Imvigor 210 study were fatigue, diarrhoea, anaemia, increases in alanine aminotransferase, aspartate aminotransferase and bilirubin, and renal failure (Table 21; data for cohort 1 are from clarification response A43). In study PCD4989g 9.5% of participants experienced treatment-related grade 3-4 events (data are from clarification request A43). Across the three cohorts the rates of these individual events were low, around 2%.

Table 20 Overview of adverse events

Event, %	**Imvigor 210 Cohort 1, 15-month cutoff, n=119**	**Imvigor 210 Cohort 2, 20-month follow-up, n=310**	**PCD4989g study, n=95[a]**
Adverse event, any	95.8	97.7	97.9
Treatment-related adverse event, any	66.4	71.0	66.3
Serious adverse event, any	37.8	46.5	-
Treatment-related serious adverse event, any	10.1	12.3	5.3 [b]
Grade 3-4 event	45.4	60.0	50.5
Treatment-related grade 3-4 event	16.0	18.1	9.5
Grade 5 event (death related to adverse event)	3.4	1.0	1.1
Treatment-related grade 5 event (death related to adverse event)	0.8	0	-
Adverse event of special interest	31.1	30.0	36.8 [b]
Grade 3-4 adverse event of special interest	7.6	6.5	-
Adverse event leading to dose interruption	34.5	32.3	-
Adverse event leading to study drug withdrawal	7.6	3.9	4.2 [b]

[a] as confirmed in clarification A41, not all participants received the licensed dose, results are supportive data only.
[b] from clarification response A43

Adverse events of special interest

Adverse events of special interest (Table 22) were mostly immune-mediated adverse events and renal function events which are anticipated effects of using a monoclonal antibody therapy. They were experienced by 30-31% of patients in Imvigor 210 and 36.8% in study PCD4989g.

The most frequent adverse events of special interest in cohort 1 were: rash (10.1%), hypothyroidism (7.6%), increased alanine aminotransferase (7.6%) increased aspartate aminotransferase (6.7%), increased bilirubin (3.4%), colitis (2.5%), dermatitis (2.5%) and peripheral neuropathy (2.5%) (data provided by the company in clarification response A43). Twenty-five percent of participants received steroids for an adverse event of special interest in this cohort (Table 22). The CS states that no major decline in median estimated glomerular filtration rate was observed in cohort 1.

Table 21 Treatment related grade 3-4 adverse events

Event, %	Imvigor 210 Cohort 1, 15-month cutoff, n=119	Imvigor 210 Cohort 2, 20-month follow-up, n=310	PCD4989g n=95 [a]
Overall	16.0	18.1	9.5
Fatigue	3.4	1.6	Not reported
Diarrhoea	1.7	0.3	Not reported
Pruritus	0.8	0.3	Not reported
Decreased appetite	0.8	0.6	Not reported
Hypothyroidism	0	0.3	Not reported
Anaemia	0.8	0.6	1.1
Chills	0	0	Not reported
Nausea	0	1.9	Not reported
Pyrexia	0	0.6	Not reported
Rash	0.8	0.3	Not reported
Vomiting	0	1.3	Not reported
Rash, maculopapular	0	0	1.1
ALT increase	3.4	1.9	1.1
Arthralgia	0	1.0	Not reported
AST increase	2.5	1.6	1.1
Blood bilirubin increase	1.7	0.6	Not reported
Blood alkaline phosphatase increase	0.8	1.6	1.1
Dyspnoea	0	0	Not reported
Infusion-related reaction	0	0	Not reported
Lymphocyte count decrease	0	1.0	1.1
Renal failure	1.7	0.6	Not reported
Asthenia	0	0	2.1
Neutropenia	0	0	1.1

ALT, alanine aminotransferase; AST, Aspartate aminotransferase
[a]as confirmed in clarification response A41, not all participants received the licensed dose, results are supportive data only.

Table 22 Adverse events of special interest, immune-mediated requiring systemic corticosteroids.

	Imvigor 210 Cohort 1, 15-month cutoff, n=119		Imvigor 210 Cohort 2, 20-month follow-up, n=310	
Event, % (rounded)	Any Grade	Grade 3-4	Any Grade	Grade 3-4
Overall	12[a]	7[b]	13[c]	3
Rash	3	1	[redacted]	[redacted]
ALT increase	2	2	[redacted]	[redacted]
Blood bilirubin increase	2	2	[redacted]	[redacted]
Rhabdomyolysis[d]	2	1	-	-
AST increase	1	1	[redacted]	[redacted]
[redacted]	-	-	[redacted]	[redacted]
Autoimmune colitis	1	1	-	-
Colitis	1	1	[redacted]	[redacted]
Diarrhoea[d]	1	1	-	-
Liver disorder[d]	1	1	-	-
Rheumatoid arthritis	1	1	-	-
Arthralgia[d]	1	0	-	-
Arthritis[d]	1	0	-	-
Hypothyroidism	1	0	-	-
Muscle spasms[d]	1	0	-	-
Rash, maculopapular	1	0	[redacted]	[redacted]
Tenosynovitis[d]	1	0	-	-

a [redacted] b [redacted] c [redacted]
d [redacted]

ALT, alanine aminotransferase; AST, Aspartate aminotransferase
Data for cohort 2 not reported in the CS and have been taken from the updated CSR.

The most frequent adverse events of special interest in cohort 2 were: rash (11.6%), increased alanine aminotransferase (5.2%), increased aspartate aminotransferase (5.2%), hypothyroidism (3.2%), maculo-papular rash (3.2%), peripheral neuropathy (3.2%), pneumonitis (2.6%), and increased bilirubin (2.6%) (data provided by the company in clarification response A43). The CS states that in 63 patients treated with atezolizumab for ≥1 year in cohort 2, 13% experienced an immune-mediated adverse event of any grade, and 3% experienced a Grade 3–4 immune-

mediated adverse event. In these patients, rash, acute kidney injury and influenza-like illness were the most common immune-mediated adverse events of any grade (n=2 each).

In study PCD2989g adverse events of special interest were provided by the company in response to clarification question A43. The most commonly reported events were similar to those seen in the two cohorts of Imvigor 210: rash (12.6%), aspartate aminotransferase increase (10.5%), peripheral neuropathy (8.4%), and alanine aminotransferase increase (7.4%).

Across both cohorts of Imvigor 210 12-13% of patients experienced immune-mediated adverse events of special interest requiring systemic corticosteroids (Table 22).

Rates of infusion-related reactions are reported (CS Tables 43 and 46), although they were not classed by the company as adverse events of special interest. Rates of infusion-related reactions were relatively low, affecting 3% of patients in cohort 1 and 0.6% in cohort 2 (none were Grade 3-4).

Adverse events leading to atezolizumab dose interruption or withdrawal

In cohort 1, 34.5% of patients had an adverse event leading to dose interruption and 7.6% had an adverse event leading to treatment withdrawal. In cohort 2, 32.3% of patients had an adverse event leading to dose interruption and 3.9% had an adverse event leading to treatment withdrawal (CS section 4.13.1). In study PCD4989g 4.2% of patients had an adverse event leading to treatment withdrawal (reported by the company in clarification response A43).

Specific adverse events leading to atezolizumab withdrawal were specified by the company in clarification response A43. Across both cohorts of Imvigor 210 and also in study PCD4989g the reasons for withdrawal were diverse and mostly affected only one patient each. These included, among others: cohort 1: cardiac arrest, myocardial infarction, sepsis, diarrhoea, rheumatoid arthritis, respiratory failure; cohort 2: sepsis, pulmonary sepsis, colitis, fatigue, cerebral haemorrhage, pneumonitis, pruritis; study PCD4989g: increased bilirubin, sepsis, intracranial mass. The rates of withdrawals were not specified in relation to the time on treatment.

Deaths

In cohort 1, there were 59 deaths: 52 were due to progressive disease, five due to grade 5 adverse events (four within 30 days of the last atezolizumab dose; one more than 30 days after the last dose), and two were due to unspecified causes (not progression or adverse event) (CS section 4.12.3.1).

In cohort 2, there were 226 deaths: 211 were due to progression, three due to grade 5 adverse events (CS section 4.12.3.1) and [redacted] (Supplemental Results Report IMVigor 210, pages 737-738).

Four of the participant deaths in cohort 1 and three in cohort 2 were due to Grade 5 adverse events, of which one (unspecified, in cohort 1) was treatment-related.

Summary of adverse events

Overall, atezolizumab appears to be reasonably well-tolerated given the advanced age of the population, and the adverse events data do not raise any safety concerns beyond those expected for an anti-cancer immunotherapy. The majority of deaths in the Imvigor 210 study were due to progressive disease, with only one death (in cohort 1) attributed as being treatment-related. Around one third of patients in each cohort experienced dose interruptions as a result of adverse events, whilst 7.6% of cohort 1 patients and 3.9% of cohort 2 patients had adverse events leading to withdrawal of atezolizumab.

3.4 Summary of the clinical effectiveness evidence

The published evidence base for effectiveness of first-line and second-line atezolizumab is based on a single phase II single-arm study, Imvigor 210. Limited additional supporting information is provided by the company from a phase I study which included patients receiving second-line atezolizumab. However, the patients received on average slightly less than the licensed dose. The primary outcome in Imvigor 210 was the objective response rate, whilst overall survival and progression-free survival were secondary outcomes. At the latest available data-cut, and based on independent review facility assessment using RECIST v1.1 criteria, first-line patients had a median overall survival of 15.9 months, median progression-free survival 2.7 months, an objective response rate of 22.7%, and the median duration of response had not yet

been achieved. Second-line patients had a median overall survival of 7.9 months, progression-free survival of 2.1 months, an objective response rate of 15.8%, and the median duration of response had not yet been achieved.

Overall, atezolizumab appears to be reasonably well-tolerated given the advanced age of the population, and the adverse events data do not raise any safety concerns beyond those expected for an anti-cancer immunotherapy.

Comparison of the clinical effectiveness of atezolizumab against comparator chemotherapy drugs was limited by a lack of evidence to allow the formation of a network, as the relevant comparators were either single-arm studies or single arms within controlled trials. To enable a network to be formed for meta-analysis, the company employed a simulated treatment comparison to 'predict' a matching atezolizumab arm for each comparator study. The resulting comparisons of atezolizumab against each comparator were then included in a network meta-analysis. The company determined that the proportional hazards assumption would be unlikely to hold for comparisons between atezolizumab and standard chemotherapy drugs. They selected a fractional polynomial model analysis approach for the network meta-analysis since higher-order fractional polynomial models are not dependent on the assumption of proportional hazards (but see below).

The company acknowledge that the results of the fractional polynomial network meta-analysis are unreliable and should be interpreted with caution.

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

- A fundamental assumption of a simulated treatment comparison is that, ideally, all covariates (i.e. prognostic factors or effect modifiers for survival) have been included in the analysis. The company has included only three or four binary covariates (from age, sex, liver metastasis, performance status). This may limit how well-matched the simulated atezolizumab arms are to the comparator arms.
- Some aspects of the analysis are unclear, including the imputation approaches used to account 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:

- The company states that the reason for using fractional polynomial models was to allow analysis of comparisons which violate the proportional hazards assumption. However, after assessing model fit, the company selected the zero-order fractional polynomial model which assumes proportional hazards. The company does not discuss the plausibility of this model.
- 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.

4 COST EFFECTIVENESS

4.1 Overview of the company's economic evaluation

The company's submission to NICE includes:

- a review of published economic evaluations of chemotherapy treatment regimens for patients with advanced or metastatic urinary bladder cancer.
- a report of an economic evaluation undertaken for the NICE STA process. The cost effectiveness of atezolizumab is compared with gemcitabine + carboplatin for patients with locally advanced or metastatic urothelial cancer for whom cisplatin-based chemotherapy is unsuitable and compared with docetaxel, paclitaxel and best supportive care for patients whose disease has progressed after prior chemotherapy.

4.2 Summary and critique of the company's review of published economic evaluations

A systematic search of the literature was conducted by the manufacturer to identify economic evaluations of chemotherapy treatment regimens for patients with advanced/metastatic urinary bladder cancer who have progressed after at least one prior chemotherapy (see section 3.1 of this report for our critique of the company's search strategy).

The inclusion and exclusion criteria for the systematic review are listed in CS Appendix 8.7. The inclusion criteria state that economic evaluations of chemotherapy treatment regimes in patients with advanced or metastatic urinary bladder cancer who have progressed after at least one prior chemotherapy regimen (or who are intolerant of cisplatin-based chemotherapy) would be included. No exclusion criteria are reported.

Forty-one studies were identified from screening 844 titles and abstracts, with a further three studies identified through hand-searching. Of these, 37 studies were excluded, mainly as the studies were reviews or editorials or were in the wrong patient population. Seven studies were included for full review (the CS does not report the references for these studies identified; they were provided by the company in clarification response B9). The company reported that none of these studies were relevant to the current submission and the CS does not provide any further details for these studies. The ERG is unclear why the company considered these studies not relevant to the current submission and we note that the company used two of these studies to inform their analyses of resource use[22] and HRQoL.[67]

4.3 Critical appraisal of the company's submitted economic evaluation

4.3.1 NICE reference case

The ERG's critical appraisal of the submitted economic evaluation based on the NICE reference case requirements is summarised in Table 23.

Table 23 NICE reference case requirements

NICE reference case requirements	Included in submission	Comment
Decision problem: As per the scope developed by NICE	Not completely	CS Table 2, CS section 1.1.1. The economic evaluation in the CS has combined two of the populations in the NICE scope to create one population whose disease has progressed (2L).
Comparator: As listed in the scope developed by NICE	Not completely	The CS does not include best supportive care for the 1L cohort and does not include retreatment with 1st line platinum-based chemotherapy for 2L treatment (CS Table 2, section 1.1.1).
Perspective on costs: NHS and PSS	Yes	
Evidence on resource use and costs: Costs should relate to NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS	Yes	
Perspective on outcomes: All direct health effects, whether for patients or, when relevant, carers	Yes	
Type of economic evaluation: Cost utility analysis with fully incremental analysis	Yes	
Synthesis of evidence on outcomes: Based on a systematic review	Yes	CS section 4.1.
Time horizon: Long enough to reflect all important differences in costs or outcomes between the technologies being compared	Yes	Time horizon of 20 years (CS Table 49).
Measuring and valuing health effects: Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of health related quality of life.	Yes	Health effects measured in QALYs. Utilities are mapped from EORTC QLQ C30 results to EQ-5D (CS section 5.4.6)
Source of data for measurement of health related quality of life: Reported directly by patients and/or carers.	Yes	
Source of preference data: Representative sample of the UK population	Yes	
Equity considerations: An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit.	Yes	
Discount rate: 3.5% per annum for costs and health effects	Yes	CS Table 49.

1L: first-line; 2L: second-line

In general, the company's analysis conforms to NICE's reference case requirements, but the analysis differs from the NICE scope with regard to the populations and comparators.

4.3.2 Model Structure

The company constructed two cost-utility models for first-line and second-line treatment with atezolizumab. The model structure was identical for the two models. The models have a lifetime time horizon of 20 years, discounting of 3.5% per annum for costs and health benefits, a weekly cycle length and apply a half-cycle correction. The perspective of the analysis is the NHS and PSS. The CS states that the time horizon was sufficiently long to capture all meaningful differences between the treatments compared and that the perspective and discounting rate were as specified by the NICE reference case.[68] The ERG considers the perspective of the model and the choice of time horizon, cycle length and discounting rate are appropriate.

The models were constructed in Microsoft Excel and each consists of a partitioned survival model with three health states: 'progression-free survival', 'progressed disease and death. A schematic of the model (CS Figure 22) is shown in Figure 10 below. The CS states that this model was chosen as the structure and health states are in line with the clinical pathway and the model structure is consistent with the approaches used in earlier NICE appraisals for treatments with advanced or metastatic carcinoma, including the previous appraisal for urothelial cancer.[22]

The model uses parametric survival modelling to fit survival curves to the observed data for progression-free survival and overall survival (see more details in section 4.3.5). The model derives the proportion of patients in the progressed disease state as the difference between the progression-free survival and overall survival curves.

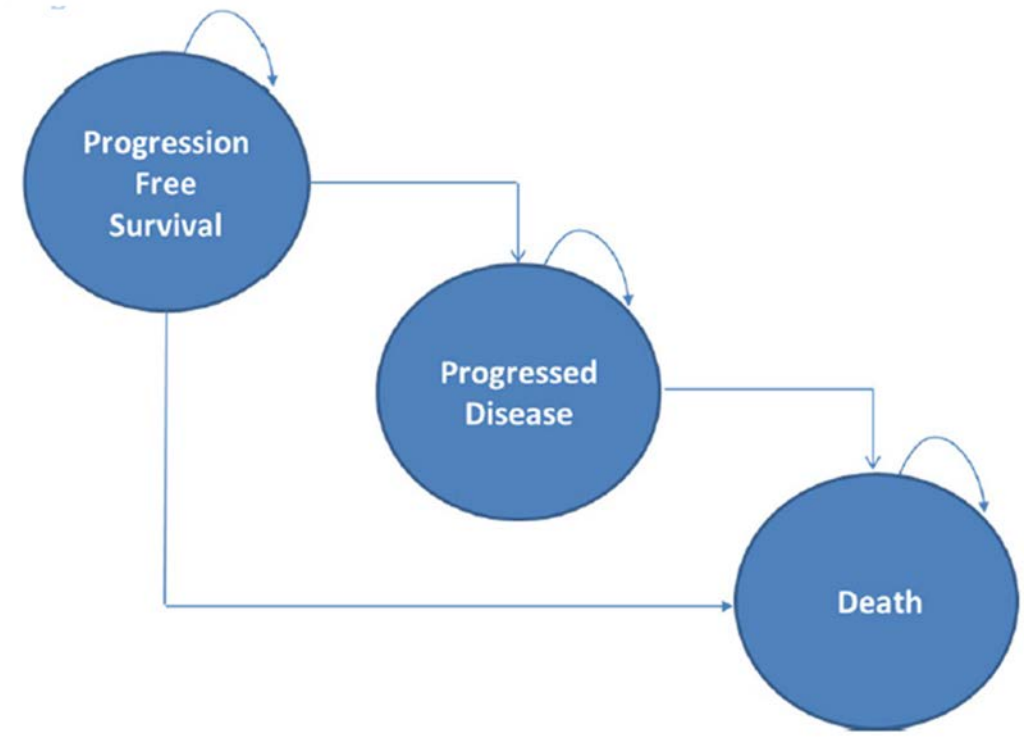


Figure 10 State model schematic (CS Figure 22)

Patients are treated with atezolizumab until disease progression unless they discontinue due to adverse events. 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

treatment pathway. The CS presents the model structure with sufficient justification for the methodological and structural choices (CS Section 5.2). In general, the modelling approach appears appropriate.

4.3.3 Population

The company performed economic analyses for the treatment of two groups of adult patients with locally advanced or metastatic urothelial cancer: i) patients who are unsuitable for cisplatin-based chemotherapy; and ii) patients whose disease has progressed after prior chemotherapy. These patient groups are in accordance with the final scope issued by NICE. The company is anticipating marketing authorisation for these populations being granted [redacted].

The company primarily uses the open-label phase II study, IMvigor 210, as a source of clinical effectiveness parameters for atezolizumab in the economic model. As we describe in more detail above (section 3.1.3), this study includes two patient cohorts: i) cohort 1: patients who are cisplatin-ineligible and received atezolizumab as a first-line treatment option and ii) cohort 2: patients who received atezolizumab as second-line treatment, after progression on chemotherapy. The company aligned their modelled populations in the two cohorts with those in the Imvigor 210 study. The baseline characteristics of these two cohorts are presented in Table 9.

The mean ages of the first-line and second-line cohorts used in the economic models are 71.8 years and 65.6 years respectively, and are consistent with the baseline characteristics of the Imvigor 210 patients (Table 9).

Although Imvigor 210 is an international study (conducted in the USA, Canada, France, Germany, Italy, The Netherlands, Spain, and the UK), only 22 patients were from the UK (first-line: five out of 119; second-line: 17 out of 310). Following expert clinical advice, we do not have any concerns about the generalisability of patients in IMvigor10 to UK NHS patients.

The CS acknowledges that the use of data from the single-arm phase II study has limitations and states that these constraints will be overcome when ongoing phase III studies IMvigor130 (for cohort 1) and IMvigor211 (for cohort 2) are completed in 2020 and 2017 respectively.

For the economic analyses of second-line treatment, the company has merged cisplatin-ineligible and cisplatin-eligible patients into a single group who receive the same comparators i.e. docetaxel, paclitaxel and best supportive care. The CS states that the treatment patterns and response rates for patients in the second-line treatment cohort are unlikely to be different based on patients' eligibility for cisplatin, although no evidence is provided in support of this.

Sub group analysis

The scope does not specify any subgroups for the appraisal and the company has not conducted any subgroup analyses.

4.3.4 Interventions and comparators

The interventions and comparators used in the first-line and second-line patient cohorts within the economic models are summarised in Table 24.

Table 24 List of intervention and comparators used in the company's economic analyses

Patient cohort	Intervention	Comparators
First-line	Atezolizumab	Gemcitabine + carboplatin
Second-line	Atezolizumab	Docetaxel Paclitaxel Best supportive care

In summary, the comparators used in the economic models broadly align with the NICE scope for this appraisal, except for slight deviations, as discussed in section 2.3.

4.3.5 Treatment effectiveness and extrapolation

The clinical outcomes included in the CS model were progression-free survival, overall survival and time to treatment discontinuation (TTD). The company's approach for obtaining clinical effectiveness estimates for comparisons of atezolizumab against first-line and second-line chemotherapy treatments for use in the economic analysis is explained and critiqued in section 3.1 above. In summary, the company did not find any direct head-to-head comparisons of atezolizumab against chemotherapy and only single-arm studies of relevance were identified (as described above, section 3.1.3). To enable comparisons between atezolizumab and chemotherapy drugs the company conducted a simulated treatment comparison in which each individual comparator arm was compared against a 'predicted' atezolizumab arm. The predicted arm was based on a Cox regression prediction model informed by baseline covariates in the

comparator arm. These simulated atezolizumab-chemotherapy comparisons were then included by the company in network meta-analysis (using a fractional polynomial model) to produce survival hazard ratios for atezolizumab versus each comparator. We have summarised and critiqued the included comparator studies above in section 3.1.3 (their baseline characteristics are given in Table 10 and Table 11). We have also provided a summary and critique of the simulated treatment comparison and network meta-analysis methods in section 3.1.7; and the results of the meta-analysis in section 3.3.6.

For the economic analyses, Imvigor 210 was used as the primary data source for the atezolizumab arm in both the first-line and second-line patient cohorts. Estimates of the clinical effectiveness of atezolizumab versus first-line and second-line comparators were provided by hazard ratios from the company's network meta-analysis. The company used five methods to estimate treatment effects in their economic models (see Table 25):

i. Extrapolation from Imvigor 210
ii. Assumption based on the KEYNOTE-045 study: progression-free survival of gemcitabine + carboplatin, and that of docetaxel and paclitaxel are equal to progression-free survival of atezolizumab
iii. Mix cure rate model
iv. Proportional hazards model
v. Fractional polynomial with capped hazard ratio

These methods are described further in the following subsections.

Table 25 Methods to estimate treatment effects

Outcome	Intervention	Comparators		
First-line	**Atezolizumab**	**Gemcitabine + carboplatin**		
PFS	Extrapolation from IMvigor 210	Assumption: PFS of gemcitabine +carboplatin = PFS of atezolizumab		
OS	Mix cure rate model (uses data from IMvigor 210 and Life tables)	Results from fractional polynomial NMA with capped HR		
Second-line	**Atezolizumab**	**Best supportive care**	**Docetaxel**	**Paclitaxel**
PFS	Extrapolation from IMvigor 210	Use of proportional hazards model (by using the HR obtained from the fractional polynomial NMA)	Assumption: PFS of gemcitabine +carboplatin = PFS of atezolizumab	Assumption: PFS of gemcitabine +carboplatin = PFS of atezolizumab
OS	Mix cure rate model (uses data from IMvigor 210 and Life tables)	Results from fractional polynomial NMA with capped HR	Results from fractional polynomial NMA with capped HR	Results from fractional polynomial NMA with capped HR

HR: Hazard Ratio; NMA: network meta-analysis; PFS: progression-free survival

4.3.5.1 Progression-free survival

Atezolizumab (first-line and second-line)

In the company's base case, parametric distributions (exponential, Weibull, log-logistic, log-normal, generalised gamma and Gompertz distributions) were fitted to the observed Kaplan-Meier data from the Imvigor 210 study to extrapolate progression-free survival curves for both first-line and second-line treatments. The company assessed the goodness of fit of these distributions by using the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), visual inspection and clinical plausibility, following which the generalised gamma distribution was chosen for the base case in both the patient cohorts. The company also used the log-normal and log-logistic distributions in scenario analyses, but these did not have any significant impact on the base case ICERs (CS Table 93).

First-line comparator: gemcitabine + carboplatin

The CS states that the results of the fractional polynomial network meta-analysis (discussed earlier in section 3.3.6), when applied to the economic model, provided clinically implausible results. The company explored using the proportional hazards model and capping of hazard

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 an Australian 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 (CrI 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 does not present any rationale for using the KEYNOTE-045 study to inform the progression-free survival parameter for the comparator arms. 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 Australian 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.

4.3.5.2 Overall Survival

Atezolizumab (first-line and second-line)

The company uses a mix-cure rate model to extrapolate overall survival for the atezolizumab arms in both the patient cohorts. The mix -cure rate model estimates decline in mortality risk associated with cancer by accounting for cancer-related mortality risk and background mortality risk. Two populations - those with a low risk of cancer-related death and those with a high risk of cancer-related death are combined to produce an average survival curve for the whole population. The survival equations for these patient groups use 'cure fraction' as a factor determining trial population survival. The CS uses the dataset from the observed survival times in the IMvigor 210 study and background mortality risks from life tables. The CS assumes the cure fraction for atezolizumab is 0% in the base case (which implies 0% of patients will be at a lower risk of death due to the condition). Long-term survival data were extrapolated by fitting a generalised gamma distribution in the base case analyses. The company measured the goodness of fit using AIC and BIC statistics which justify the selection of this distribution (CS Table 53 and Table 54). Different 'cure fraction' rates ranging from 1% to 3% were assessed in scenario analyses. These alternative cure fraction rates do not have a significant impact on the base case ICERs.

First-line comparator: gemcitabine + carboplatin

To obtain overall survival curves for the comparator arm, the company uses the results from the fractional polynomial model (presented above in section 3.3.6). The CS states that using the data from the network meta-analysis results in the hazard ratio increasing linearly over time, which would inadvertently lead to clinically implausible results as the relative efficacy of atezolizumab continues to increase. As a result, the company capped the hazard ratio at the time point corresponding to the median follow-up duration of the study which, as reported by Bamias et al.[42] for the first-line cohort was at 8 months. Beyond this time point, the company assumed proportional hazards.

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

The company used the same approach and assumption as for the first-line comparison to model overall survival for the second-line comparators. Hazard ratios were capped at the time points

corresponding to the median follow-up of the atezolizumab study which for the latest data-cut (see section 3.3) was at 21.16 months.

Table 27 and Table 28 show the parameter estimates, which the CS refers to as 'contrast estimates', from the fractional polynomial models used in the company's network meta-analysis, and the hazard ratios used in the company's economic analyses (for an overview of the fractional polynomial models see section 3.1.7 above). The derivation of hazard ratios from the contrast estimates is explained as follows:

$$\text{Log HR} = \text{Intercept} + (\text{slope}^* \text{ time points})$$

i.e., $$\text{HR} = e^{\text{Intercept} + (\text{slope}^* \text{ time points})}$$

where, time-points refer to time points in the Markov cycles.

Table 27 Contrast estimates from fractional polynomial models

Treatment	Intercept (median)	Intercept (lower bound)	Intercept (upper bound)	Slope (median)	Slope (lower bound)	Slope (upper bound)	Correlation between intercept and slope
First-line (from clarification response A33)							
Gemcitabine +carboplatin	0.21	-0.242	0.647	0.051	-0.009	0.112	-0.749
Second-line (from CS Table 24)							
BSC	0.547	0.238	0.848	-0.002	-0.038	0.034	-0.736
Paclitaxel	0.333	-0.280	0.901	0.003	-0.073	0.070	-0.738
Docetaxel	-0.168	-0.581	0.234	0.044	-0.008	0.092	-0.787

BSC: best supportive care

ERG comments on the methods for modelling overall survival

The company's approach to modelling survival in patients in the atezolizumab arm using a mix cure rate model appears to be reasonable. The ERG notes that the overall survival model results for atezolizumab compare well with the observed IMvigor 210 trial data (CS Table 76), based upon visual inspection.

Whilst the company has reported validation checks for the modelled overall survival results (by comparing the model results with results from clinical experts as shown in CS Table 77), the CS

does not report any sensitivity or scenario analyses with alternative parametric distributions. This is a major concern as the model results are very sensitive to the parametric distribution 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) The value is obtained from the zero order FP model which is then used to estimate the HR at different time points until the follow up duration for the comparator study (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.

The same parametric distribution is used for atezolizumab and its comparators but the company does not comment on how well the parametric distribution fits with the comparator trial data. The ERG compared the modelled overall survival results with the first-line survival results reported in the study by De Santis et al.[44] We extracted the Kaplan-Meier curve for gemcitabine + carboplatin from De Santis et al. (using Engauge digitiser). We then compared the curve with the modelled overall survival curves obtained using the company's base case results using the estimates from the fractional polynomial model and with the assumption of proportional hazards. As shown in Figure 11, the exponential distribution provides a better fit to the overall survival data in De Santis et al.[44] compared to the cure generalised gamma (i.e., the mix cure-rate model extrapolated using a generalised gamma distribution) used in the base case of the CS. As the follow-up duration for gemcitabine + carboplatin is significantly longer than for atezolizumab, it appears reasonable to base the parametric curve on the best fit for the gemcitabine + carboplatin arm, rather than the atezolizumab arm. Based on this observation, we consider that it would be appropriate to use the Kaplan-Meier data with an exponential tail to extrapolate first-line overall survival. This is explored in section 4.4 below.

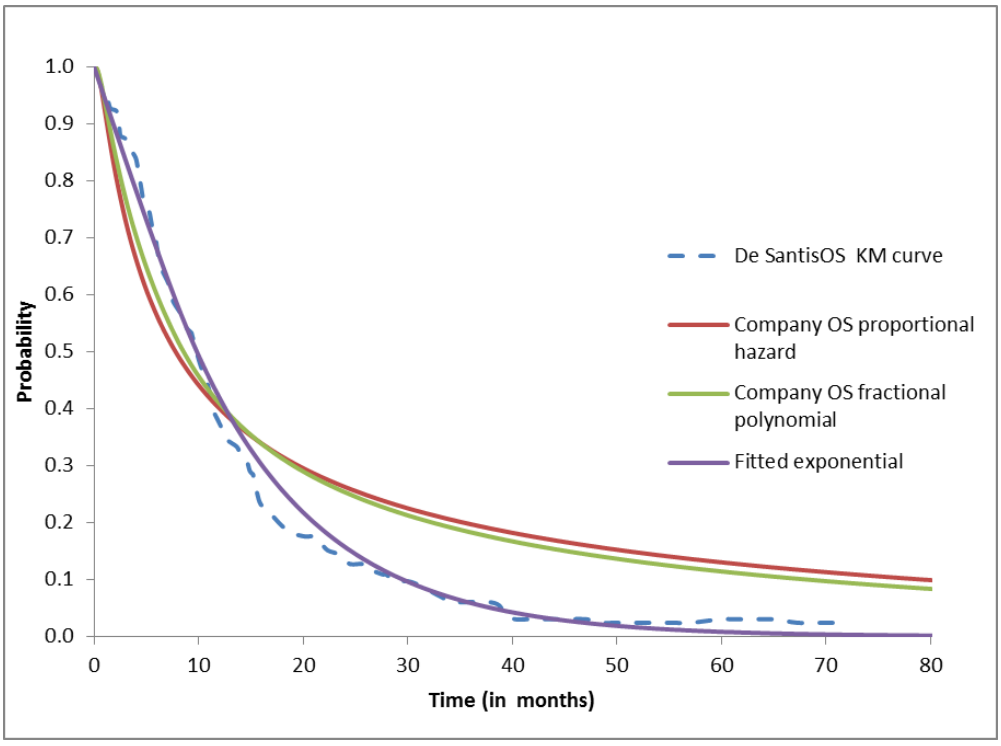


Figure 11 Comparison of the overall survival curves from De Santis et al. and the company's model

Similarly, for the second-line treatment comparisons, we compared the modelled overall survival for each of the comparator arms with the survival data presented by Bellmunt et al.[45] for best supportive care; Kim et al.[49] for docetaxel; and Lee et al[39] for paclitaxel. Of the five second-line comparator studies (i.e. those listed in Table 11), two reported survival data on best supportive care. We chose the study by Bellmunt et al.[45] to compare the modelled overall survival curve for best supportive care due to it having a larger sample size and longer follow up compared to the study by Noguchi et al.[46, 47] For docetaxel, the study by Kim et al.[49] was chosen over the study by Choueiri et al.[50] due to having a longer follow up duration. For paclitaxel, we used the only study that reported a survival curve for paclitaxel, by Lee et al.[39] A similar technique was used to extract Kaplan-Meier data for survival from these studies, as adopted for the first-line comparisons.

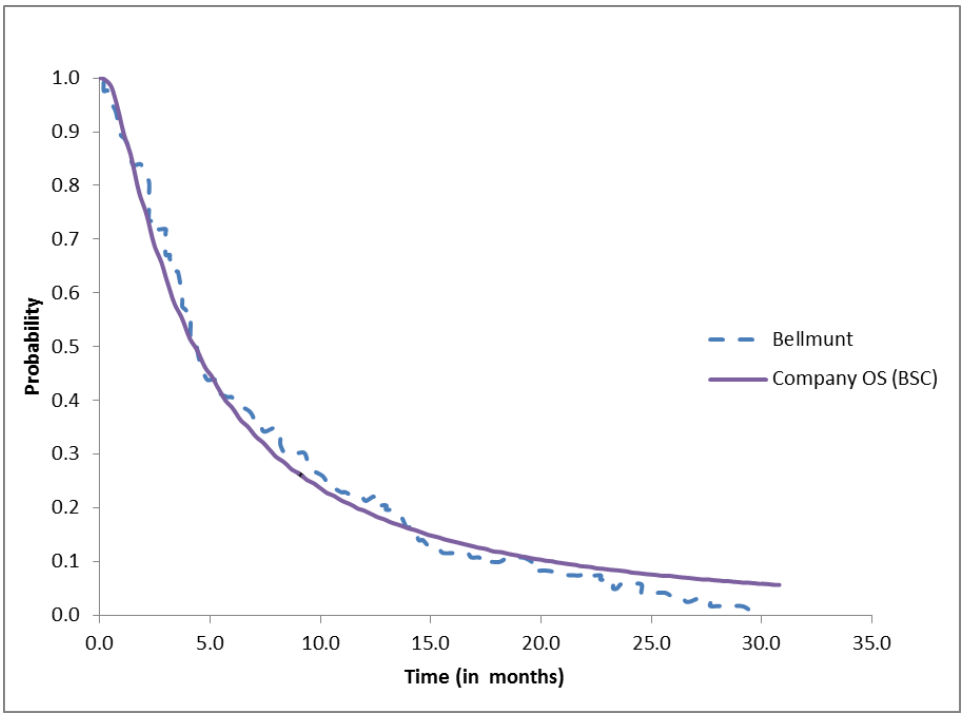


Figure 12 Comparison of the overall survival curve for best supportive care from Bellmunt et al. with the company's modelled curve

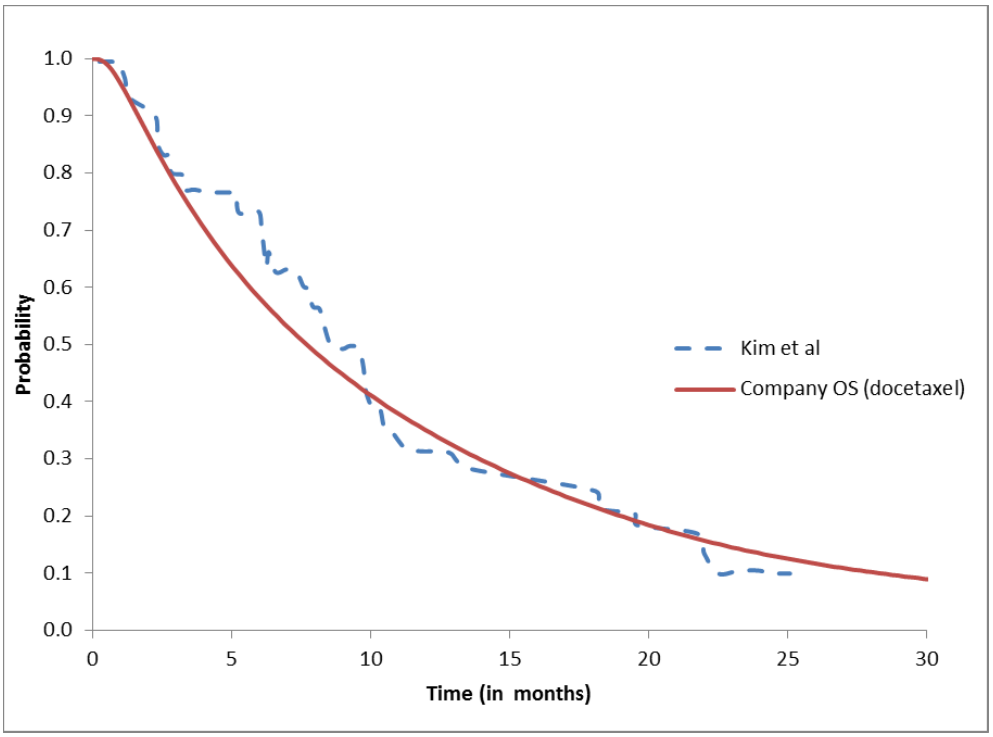


Figure 13 Comparison of the overall survival curve for docetaxel from Kim et al. with the company's modelled curve

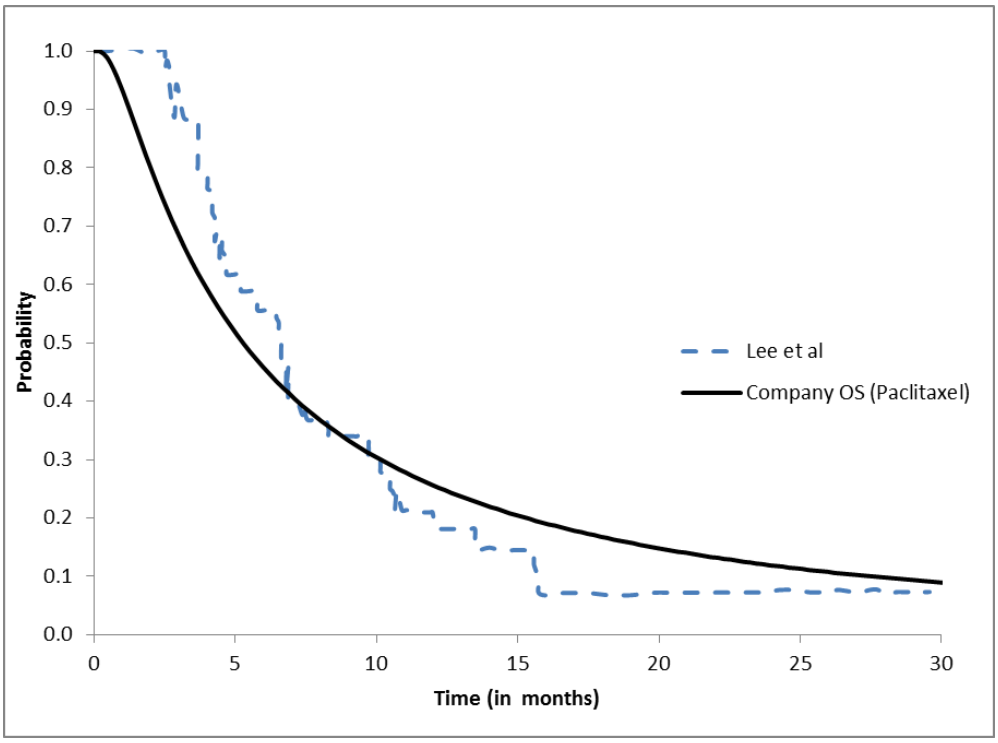


Figure 14 Comparison of the overall survival curve for paclitaxel from Lee et al. with the company's modelled curve

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 primarily 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

assessment. The CS states that for the first-line and second-line comparator arms, progression-free survival is used as a proxy for time on treatment. To assess uncertainty associated with TTD, the company has conducted scenario analyses in which progression-free survival is used as a proxy for time on treatment for the atezolizumab arm. The ICERs indicate that the results are sensitive to the way treatment duration is modelled. These are shown in detail in the ERG's additional analyses (section 4.4).

ERG comments on the methods for modelling time to discontinuation

On balance, for the base case, we agree with the company's approach to extrapolate TTD data for the first-line and second-line atezolizumab arms. However, they have used a generalised gamma distribution in both the patient cohorts, although the findings from the AIC and BIC statistics indicate that a Weibull function for first-line and a log-logistic function for second-line provide the best fit. Visual inspection of fitting different distributions shows that both Weibull and log-logistic for first-line treatment and other curves as stated in CS Table 67 for second-line treatment provide plausible fit to model TTD in the two patient cohorts. We ran the economic models with the alternative plausible distributions in both the patient cohorts, as discussed in the ERG's exploratory analyses in section 4.4.

In estimating TTD for the comparator arms in both the patient cohorts, the company contradicts their statement that '*PFS is not a good surrogate for treatment duration as it is likely to underestimate the true treatment duration expected in clinical practice, and as such, treatment cost*' (CS section 5.5.5, end of 1st paragraph within Atezolizumab section). The ERG notes that patients treated with first-line gemcitabine + carboplatin receive up to a maximum of six cycles of treatment and therefore TTD is not modelled according to progression-free survival. For second-line treatment, TTD associated with docetaxel and paclitaxel is modelled according to progression-free survival. However, as the costs associated with these drugs are minimal, the assumption (using progression-free survival as a proxy for TTD) does not have any significant impact on the overall model results. TTD does not apply to best supportive care as there is no associated treatment cost.

Whilst the company has conducted scenario analyses associated with the atezolizumab arm, no such analyses have been conducted for the comparator arms. This appears to be appropriate, based on the reasons outlined above. In summary, we view that the company's approach to modelling TTD within the current appraisal is reasonable.

4.3.5.4 Adverse events

The company does not model the impact of adverse events on HRQoL. The CS states that there are limited data on adverse events which is coupled with a lack of comparative data for HRQoL in metastatic urothelial carcinoma. These aspects make it challenging to incorporate the effects of adverse events on HRQoL in the economic analyses. The CS notes that EQ-5D which will be collected as part of an ongoing phase III trial (due to complete after the conclusion of the current technology appraisal) should provide more evidence on the impact of adverse events on HRQoL. However, costs associated with adverse events are incorporated in the economic models, details of which are explained below in section 4.3.7. The ERG supports these justifications with respect to adverse events.

4.3.6 HRQoL

The CS reports that HRQoL data specific to the decision problem will be available from ongoing phase III studies. Pending the completion of these studies, and for the purpose of this submission, the company conducted a systematic literature review to identify HRQoL studies for patients with advanced or metastatic urinary bladder cancer who have progressed after at least one prior chemotherapy regimen or who are intolerant of cisplatin-based chemotherapy. The electronic databases searched included Medline In-process, Embase and the Cochrane library and the search strategy (reported in CS Appendix 8.9) appears to have been appropriate according to our appraisal (section 3.1.1). The inclusion criteria specified utilities derived directly from trials, through generic preference-based instruments or through mapping studies. Studies that reported utilities in patients undergoing surgery or receiving chemotherapy were also included. After removing duplicates, the CS identified 127 references as being potentially relevant (CS Figure 30). However, after reviewing these references in detail, the company concluded that they were not relevant to the decision problem and excluded all of them. Following the exclusion of these studies, the company expanded their search criteria to include any publication reporting HRQoL data for patients diagnosed with urothelial/bladder cancer regardless of the line of treatment or the disease severity (CS Table 59). Once more, the CS reports that none of the studies identified were consistent with the reference case and therefore all the studies were excluded. The ERG agrees with the exclusions.

Having identified no relevant studies, the CS reports that relevant HTA submissions and cost-utility analyses identified during the company's review of economic evaluation publications were

re-visited (CS section 5.1; CS Table 60). Generally, the identified studies acknowledged a lack of appropriate utilities for the populations of interest and most of these studies employed mapping or preference-based elicitation from proxy populations.

Based on advice from the company's experts that utility values in the NICE guidance on vinflunine for treatment of transitional cell carcinoma of the urothelial tract[22] were too low, the company used utility values cited in the Australian Pharmaceutical benefits Advisory Committee (PBAC) cost-utility analysis for vinflunine[67] to carry out base-case cost-effectiveness analysis. The ERG's expert clinical advisor agreed that the vinflunine utility values from the NICE appraisal were too low.

The CS does not provide a complete list of the excluded 127 studies in the main text or the appendix. However, the ERG identified one additional study which included measures of HRQoL in patients with advanced urothelial carcinoma and which could potentially be used to estimate or inform utility scores: Soga et al. 2007.[70] The study by Soga et al. was is in a Japanese setting, where paclitaxel + carboplatin therapy was administered as second-line treatment to patients who had become resistant to platinum based chemotherapy. The study reports the EORTC QLQ-30 values at two time points – pre-treatment and post-treatment. We mapped these values to the EQ-5D and estimated single utility scores. The utility scores we estimated (0.707 and 0.673 for pre-treatment and post treatment respectively) indicate that patients have lower HRQoL on treatment than when not on treatment.

In the company submission, two health states account for changes in HRQoL in competing cohorts within the model. They are the 'on-treatment' or progression-free survival state and the 'off-treatment' or progressive state. The health state utility values used in the model are shown in Table 29 (CS Table 62). While utility scores are attached to these health states, the quality of life impact of adverse events is not accounted for in the model. Based on the opinion of the ERG's clinical advisor, the company's decision to ascribe a higher utility value to the 'on-treatment' state is counterintuitive, as patients are expected to have a lowered HRQoL during treatment due to the unpleasant effects of chemotherapy. The CS uses similar utility scores for both the intervention and the comparators. According to the ERG's clinical expert advisor, atezolizumab is likely to be more tolerable than the comparator chemotherapies due to its mechanism of action in the body. Therefore the assumption of similar utilities could possibly bias cost-effectiveness analysis in favour of the comparators.

Table 29 Summary of utility values for cost-effectiveness analysis (CS Table 62)

State	Utility value: mean (standard error)	Reference in submission (section and page number)	Justification
On-treatment	0.75 (0.150)	5.4.6	Derived from mUC patients in vinflunine Australian PBAC assesssment
Off-treatment	0.71 (0.142)	5.4.6	Derived from mUC patients in vinflunine Australian PBAC assesssment

mUC: metastatic urothelial carcinoma; PBAC: Pharmaceutical Benefits Advisory Committee

The PBAC cost-utility analysis for vinflunine[67] cited Rowen et al. 2011[71] as the source of the algorithms used in estimating utility values. Rowen et al. 2011 derived a preference based measure (EORTC-8D), which was applied to EORTC QLQ-C30 scores from a vinflunine trial to derive the utility scores for progression-free survival: vinflunine + best supportive care, 0.75; best supportive care, 0.78; and progressive disease, 0.71. The PBAC analysis also mentions a second paper by Mckenzie et al. 2009[72] which uses a mapping approach to derive preference-based utility scores from EORTC QLQ-C30. The values derived from Mckenzie et al. 2009 are lower and experts (as stated in the PBAC analysis) were said to be of the opinion that values derived from the Rowen et al. algorithm are likely to be more robust.

The CS reports sensitivity analyses that varied the utility scores. For both atezolizumab and the comparators, a lower value from the vinfluine NICE appraisal and an upper value of 1 were explored (CS Table 92). For the 'off-treatment' utility, the CS simply assumes a lower value of 0.5 and an upper value of 1. The CS sensitivity analyses (CS Figures 46 and 47, and CS Table 93) show that utility is one of the main drivers of cost-effectiveness. ERG analysis also confirms this. Therefore the ERG considers that, given the high uncertainty surrounding the base-case utility inputs in the CS model, HRQoL data derived directly from trials with atezolizumab and the comparators would lead to more robust conclusions.

The utility values used in the CS are not adjusted for age and disutilities arising from adverse events are not factored into the model. The CS states that, due to limited data, it was not feasible to model the effects of adverse reactions on HRQoL.

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 atezolizumab 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

perspective. The CS states that, following consultation with experts, the key sources for costs and resource inputs were a NICE appraisal on vinflunine[22] and NICE appraisals on non-small cell lung cancer.[75, 76]

While the CS model has a dose fixed at 1200mg on day one of each 21 day cycle for atezolizumab in line with ongoing IMVigor phase III trials, the dosing for comparators is in mg/m^2. The CS assumes that the average body surface area of patients in the IMvigor 210 study is representative of the model cohorts. The CS states that due to data constraints, dose modifications and treatment breaks are not assumed for atezolizumab or any of the comparators. As none of the comparators are licensed for use in metastatic urothelial carcinoma in the UK, the CS uses information from four sources. These sources are discussed in CS section 5.5.4 and listed in CS Table 65. This table is reproduced below in Table 30.

The cost of atezolizumab in the CS model is the proposed company cost stated in the CS. For the comparators (gemcitabine + carboplatin, docetaxel, and paclitaxel) the CS uses the costs stated in eMit (2015)[77] for the base-case analysis, and then estimates non-weighted averages from published list prices for scenario analysis. The ERG notes that while cost-effectiveness analysis results for the scenario analysis of the CS are given in CS Tables 93 and 94, and CS section 5.8.3, the sources of the above-mentioned non-weighted averages are not explicitly listed in the CS. Given the paucity of data, we believe the assumptions applied by company for estimating drug dose and cost to be reasonable.

Administration costs for all comparators are sourced from the National Schedule of Reference Costs - Year 2015-16 - NHS trusts and NHS foundation trusts. They are reported in Table 31 (CS Table 68). We note that an error has been made regarding the stated sources in the CS (2014-2015 instead of 2015-2016). The CS assumes the same administrative costs for atezolizumab as for docetaxel. No rationale is given for this assumption, but the ERG's clinical expert advisor suggested this is reasonable.

Table 30 Dose and drug costs for intervention and comparators (CS Table 65)

First-line	Dose	Source	List price	eMit price
Gemcitabine	1000mg/m^2 IV over 30 mins Day 1 and 8 of each 21 day cycle for maximum 6 cycles	SmPC, Guideline, phase III trial dose	200mg vial £31.60	200mg vial £3.99
Carboplatin	400mg /m^2 IV over 15 to 60 mins Day 1 of each 21 day cycle for maximum 6 cycles	SmPC,	50mg vial £21.74	50mg vial £3.57
Atezolizumab	1200mg IV over 60 mins for first infusion, thereafter 30 mins Day 1 of each 21 day cycle	Draft SmPC	1200mg vial £3807.69	n/a
Second-line	**Dose**	**Source**	**List price**	**eMit price**
Paclitaxel	80 mg/m^2 IV over 60 mins Weekly	Guideline, expert clinical advice	30mg vial £99.12 150mg vial £442.28	30mg vial £3.41 150mg vial £11.50
Docetaxel	75 mg/m^2 IV over 60 mins Day 1 of each 21-day cycle	SmPC, phase III trial	140mg vial £900.00	140mg vial £17.77
BSC	n/a	n/a	n/a	n/a
Atezolizumab	1200mg IV over 60 mins for first infusion, thereafter 30 mins Day 1 of each 21 day cycle	Draft SmPC	1200mg vial £3807.69	n/a

BSC: best supportive care; eMit: pharmaceutical electronic market information tool; IV: intravenous; SmPC: summary of product characteristics; n/a: not applicable

Table 31 Drug administration costs (CS Table 68)

Drug	Type of administration		NHS reference code	Cost per administration	Source
Atezolizumab	Deliver simple Parenteral Chemotherapy at first attendance	Outpatient Setting	SB12Z	£199	NHS reference costs 2015-16[78]
Docetaxel	Deliver simple Parenteral Chemotherapy at first attendance	Outpatient setting	SB12Z	£199	NHS reference costs 2015-16[78]
Paclitaxel	Deliver simple Parenteral Chemotherapy at first attendance	Outpatient setting	SB14Z	£304	NHS reference costs 2015-16[78]
Gemcitabine and carboplatin	Deliver simple Parenteral Chemotherapy at first attendance	Outpatient setting	SB13Z	£265	NHS reference costs 2015-16[78]

The company obtained the types and rates of adverse events for atezolizumab from IMvigor 210 (these are summarised above in section 3.3.7). Adverse event rates for the comparators were obtained from comparator studies that were included in the network meta-analysis of overall survival, but are not reported in the CS. The ERG noted some discrepancies in the adverse event data within the model (e.g. for second-line docetaxel, adverse events were taken from Chouieri et al.[50] only, not also from Kim et al.;[48, 49] and the adverse event rate for best supportive care was set to zero, although Bellmunt et al.[21, 45] reported a rate >0). The CS does not discuss these issues, although the ERG believes they are relatively unimportant compared to other sources of uncertainty in the company's analysis.

Details of adverse event costing are given in CS Table 70. Note that there are discrepancies between CS Table 70 and the company's model. For instance, while renal failure is listed in CS Table 70 as having a cost £310.00, it was omitted in the company's model. Leucopenia is said to cost £362.22 in CS Table 70 while in the model it is set at £362.66. The NICE appraisal[75]

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 B3).

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.

4.3.8 Model validation

4.3.8.1 Internal consistency

The CS reports (CS section 5.10.1) that clinical experts were consulted to validate key aspects of the model including methodological and clinical assumptions. The assumptions included the model structure and health states, the prediction model, overall survival and progression-free survival extrapolation, utility values and resource use. The CS reports that internal quality control was completed for the two models by an external consultancy (ICON). The models were internally validated by checking formulas, cell references and model functionality. The models were 'pressure tested' by using extreme values and comparing these results with the expected outcomes.

The economic models are coded in Microsoft Excel and are fully executable and user-friendly. We have not undertaken a comprehensive check of all cells in the models; internal consistency checks have been performed and random checking of the models has been done for some of the key equations in the models. We have performed a detailed checking of all model inputs reported in the CS (white box testing); changing the parameter values produced intuitive results (black box testing) and from random checking the 'wiring' of the model appears to be accurate. Through our checking of the models, we have not identified any errors, except for some errors in the reporting of costs (as discussed in section 4.3.7).

4.3.8.2 External consistency

The CS has not compared the results from their modelling to other external models.

The ERG compared the costs and QALYs for best supportive care for the current submission to the previous submission for vinflunine. The results are shown in Table 32 below.

The costs for best supportive care in the previous vinflunine appraisal were almost double those for the current appraisal, largely as a result of differences in health state costs. The QALYs were less than half for best supportive care in the vinflunine appraisal compared to the current submission, due to the utility values for post-progression in the vinflunine submission being substantially lower than the current submission. The life years for best supportive care were lower for the vinflunine appraisal compared to the current submission, which may be due to a different distribution being chosen that had a shorter extrapolated 'tail'.

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.

The CS summarises the results of the PSA by presenting these as ICERs in CS Tables 90 and 91. The ICER for first-line atezolizumab compared to gemcitabine + carboplatin is £47,593 per QALY gained and £129,333 per QALY for the second-line comparison to paclitaxel. The CS urges caution in the interpretation of the PSA results and states that they are unlikely to be reliable due to the high level of uncertainty in the fractional polynomial model.

The CS comments that the first-line base-case ICER is below the acceptable willingness to pay threshold for a treatment considered under the end-of-life criteria. The base case ICER based on the proposed list price of atezolizumab in second-line metastatic urothelial carcinoma treatment is above the acceptable threshold for all comparators.

4.3.10 Assessment of Uncertainty

One-way sensitivity analyses

The company varied the following parameters in deterministic sensitivity analyses: cost of atezolizumab, on-treatment utility (atezolizumab), on-treatment utility (comparator), off-treatment utility, off-treatment care costs (atezolizumab) and off-treatment care costs (comparator). The parameter values used in the analyses and rationale for their choice are shown in Table 35. Results of the analyses are displayed in Figure 15 to Figure 18.

Table 35 Parameter values for univariate sensitivity analysis

Parameter	Base case value	Lower value	Higher value	Rationale for value range
Monthly cost of atezolizumab	£5500	+ 50%	- 50%	
Atezolizumab on-treatment utility	0.750	0.653	1	Lower value: Prior NICE mUC appraisals Higher value: Maximum utility value
Comparator on- treatment utility	0.750	0.653	1	Lower value: Prior NICE mUC appraisals Higher value: Maximum utility value
Off-treatment utility	0.71	0.5	1	Lower value: 50% of possible utility value Higher value: 100% of possible utility value
Atezolizumab off-treatment supportive care costs	£146.79	+50%	-50%	
Comparator off- treatment supportive care costs	£146.79	+50%	-50%	

mUC metastatic urothelial carcinoma

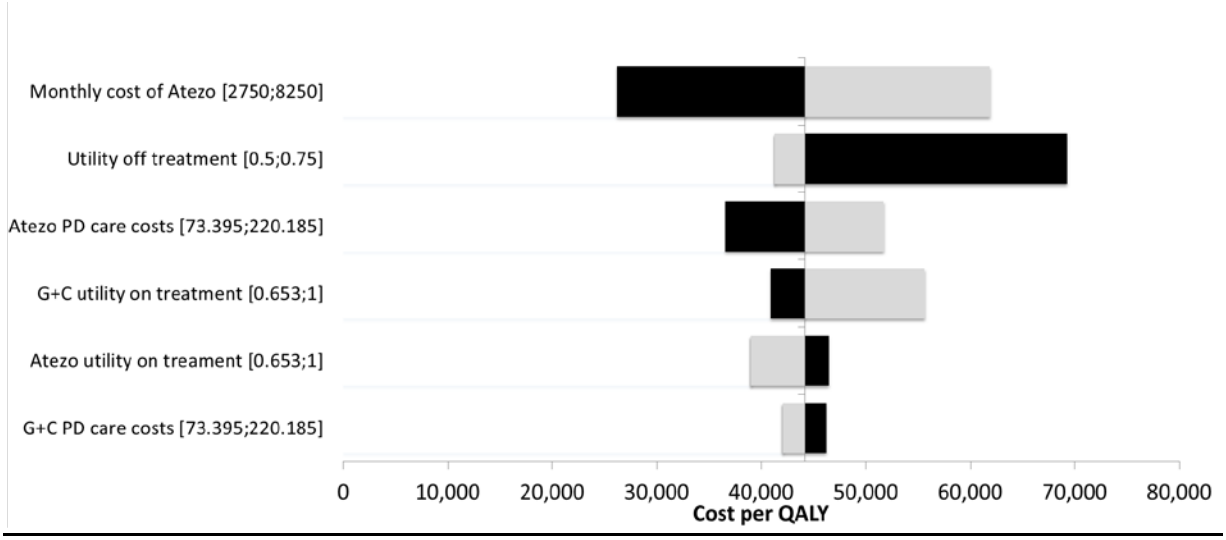


Figure 15 Univariate sensitivity analysis for comparison of first-line atezolizumab to gemcitabine + carboplatin (dark bar = lower value; light bar = higher value)

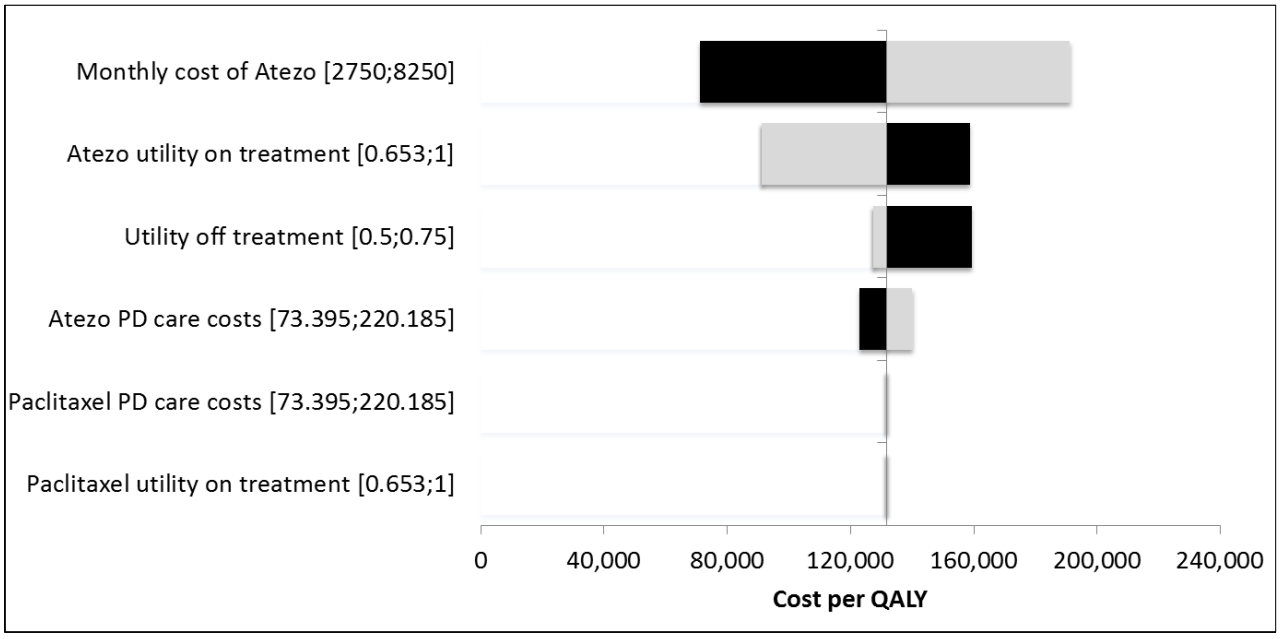


Figure 16 Univariate sensitivity analysis for comparison of second-line atezolizumab to docetaxel (dark bar = lower value; light bar = higher value)

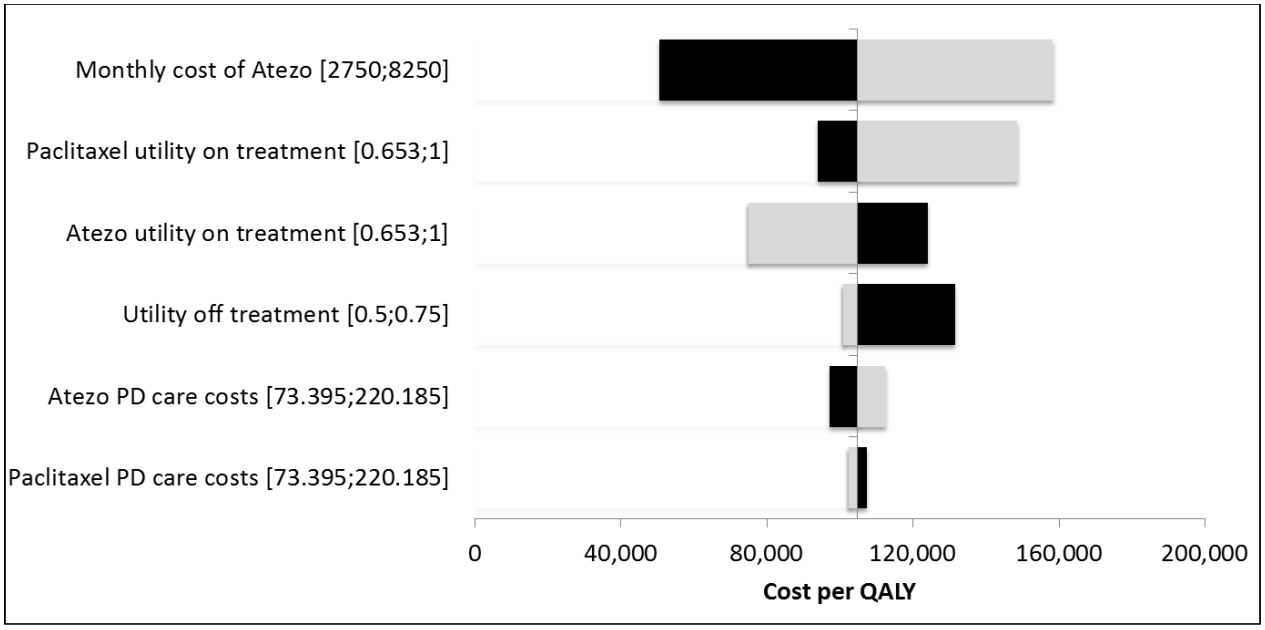


Figure 17 Univariate sensitivity analysis for comparison of second-line atezolizumab to paclitaxel (dark bar = lower value; light bar = higher value)

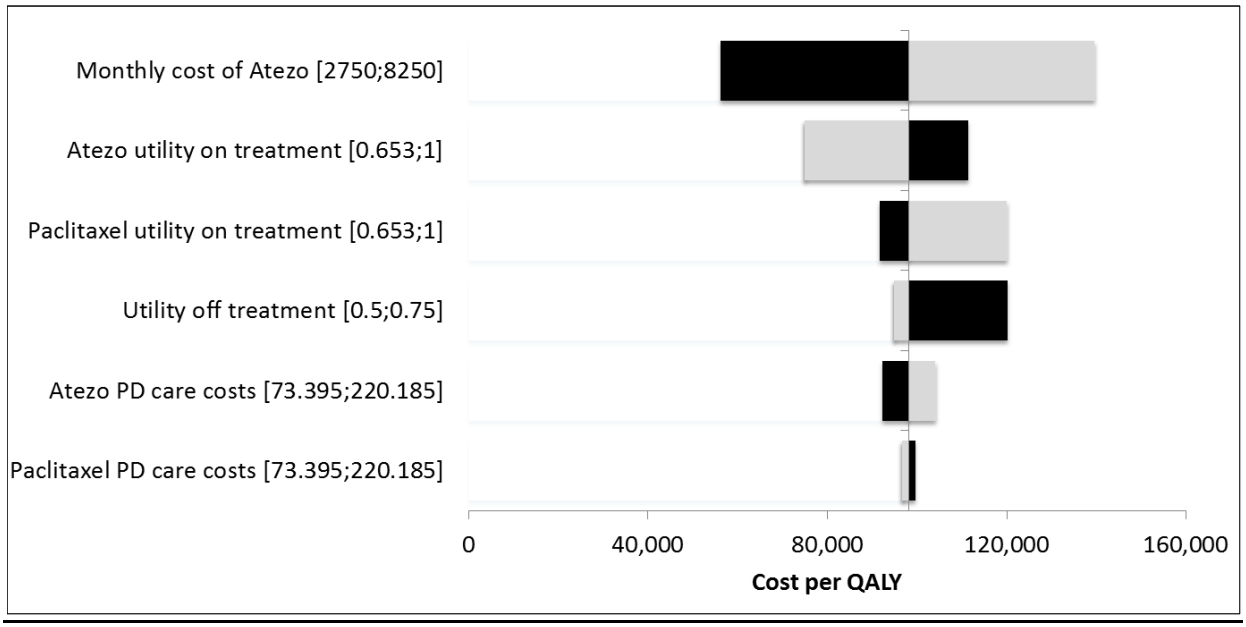


Figure 18 Univariate sensitivity analysis for comparison of second-line atezolizumab to best supportive care (dark bar = lower value; light bar = higher value)

The ERG notes that some of the input parameters have been varied in the sensitivity analyses and others have been varied in the scenario analyses. Some parameters have not been varied

in either analysis, such as alternative overall survival distributions. We note that the on-treatment utility and the treatment supportive costs for atezolizumab and its comparators have been varied independently. However, we consider that these parameters will be highly correlated between treatments.

The main drivers of the first-line economic analysis results are the price of atezolizumab and the utility of patients in the progressed disease state. The CS states that the ICER remains below the end-of-life willingness to pay threshold in the majority of scenarios explored. For the second-line results, the ICER is most sensitive to the price of atezolizumab. The ERG notes that the parametric survival functions for overall survival have not been varied in either the sensitivity analyses or the scenario analyses and these are also drivers of the first-line and second-line economic analysis results.

Scenario Analyses

The company conducted scenario analyses to assess uncertainty around structural assumptions and changes to input parameters for the model. The following scenarios were explored for parameter changes to: drug costs for comparators; alternative overall survival cure-rates; alternative progression-free survival parametric distributions; progression-free survival as a proxy for treatment duration for atezolizumab; on-treatment utilities; off-treatment utilities; time horizons of 10 years; and cost and effects discount rates.

Results are shown below in Table 36 and Table 37 for first-line comparisons (CS Table 93) and second-line comparisons (CS Table 94). The results are most sensitive to changes to assumptions around the treatment duration, the time horizon and off-treatment utility.

The ERG notes that there are no scenario analyses varying the distributions used for overall survival. The ERG investigated the effect of varying these parameters as reported in section 4.4 below.

Table 36 Scenario analysis results for first-line atezolizumab vs gemcitabine + carboplatin

Scenario	Parameter	Value	ICER vs gemcitabine + carboplatin
Base case	Comparator price	eMIT drug prices	£44,158
		List prices	£41,309
Base case	Cure rate	0%	
		1%	£44,026
		2%	£43,891
		3%	£43,754
Base case	Distribution PFS	Gamma	£44,158
		Log-normal	£44,075
		Log-logistic	£44,139
Base case	Comparator relative effect PFS	Equal to atezolizumab	
Base case	Treatment duration assumption	Actual treatment duration	£44,158
		Until progression	£64,365
Base case	Time horizon	20	£44,158
		10	£58,992
		15	£48,563
Base case	On-treatment utility (all products)	0.750	£44,158
	Atezolizumab on-treatment utility	0.800	£43,028
	GEM + CAR on-treatment utility	0.653	£40,884
Base case	Off-treatment utility	0.710	£44,158
		0.500	£69,252
		0.750	£41,307
Base case	Discount rate – effects and costs	3.5% for both	£44,158
	Discount rate - costs	1.5% (3.5% for effects)	£46,807
	Discount rate – effects	1.5% (3.5% for costs)	£37,859
	Discount rate – effects and costs	1.5% for both	£40,130

CAR: carboplatin; eMit: pharmaceutical electronic market information tool: GEM: gemcitabine; ICER: incremental cost-effectiveness ratio

Table 37 Scenario analysis results for second-line atezolizumab vs docetaxel, paclitaxel or best supportive care

Scenario	Parameter	Value	ICER vs docetaxel	ICER vs paclitaxel	ICER vs BSC
Base case	Comparator price	eMIT drug prices	£131,579	£104,850	£98,208
		List prices	£108,819	£72,477	£98,208
Base case	Cure rate	0%	£131,579	£104,850	£98,208
		1%	£126,277	£101,507	£95,403
		2%	£121,364	£98,369	£92,708
		3%	£116,805	£95,430	£90,115
Base case	Distribution PFS	Gamma	£131,579	£104,850	£98,208
		Log-normal	£131,509	£108,757	£97,819
		Log-logistic	£131,427	£109,624	£97,581
Base case	Comparator relative effect PFS	Equal to atezolizumab	£131,579	£104,850	£98,208
		FP	£132,250	£99,996	£98,273
Base case	Treatment duration assumption	Actual treatment duration	£131,579	£104,850	£98,208
		Until progression	£102,982	£78,727	£78,028
Base case	Time horizon	20	£131,579	£104,850	£98,208
		10	£158,410	£119,719	£109,318
		15	£139,012	£109,279	£101,541
Base case	On-treatment utility (all products)	0.750	£131,579	£104,850	£98,208
	Atezolizumab on-treatment utility	0.800	£120,864	£97,100	£92,507
	Comparator on-treatment utility	0.653	£117,567	£94,104	£91,738
Base case	Off-treatment utility	0.710	£131,579	£104,850	£98,208
		0.500	£159,492	£131,530	£120,299
		0.750	£127,334	£100,949	£94,889
Base case	Discount rate – effects and costs	3.5% for both	£131,579	£104,850	£98,208
	Discount rate - costs	1.5% (3.5% for effects)	£136,976	£108,999	£102,067
	Discount rate – effects	1.5% (3.5% for costs)	£116,599	£95,227	£89,962
	Discount rate – effects and costs	1.5% for both	£121,382	£98,995	£93,497

BSC: best supportive care; eMit: pharmaceutical electronic market information tool; ICER: incremental cost-effectiveness ratio; PFS: progression-free survival

Probabilistic Sensitivity Analyses

The company performed probabilistic sensitivity analyses using 1000 simulations. The simulation takes about 2 minutes to run. The distributions and sources to estimate parameters are reported in CS Table 71 (CS section 5.6). The analyses were based on the proposed list price of atezolizumab, and the eMIT drug prices for the comparators. Patient age, discount rate, time horizon and costs for the atezolizumab and the comparator treatments were not varied in the analyses. Utility values were varied using the beta distribution; the parametric survival curves were varied using the multivariate normal distribution; and costs were varied by the log-normal distribution. The ERG considers that the distributions used in the PSA were appropriate. We note that the on-treatment utilities for atezolizumab and the comparators have been varied independently and the treatment supportive costs for atezolizumab and its comparators have also been varied independently. However, we consider that the on-treatment utilities will be highly correlated between treatments and in the same way the supportive care costs will be highly correlated between treatments.

The results of the first-line and second-line PSA are presented in Table 38 and Table 39. The probability of first-line atezolizumab being cost-effective is 10.9% and 53.9% at willingness to pay thresholds of £30,000 and £50,000 per QALY respectively. The probability of second-line atezolizumab being cost-effective is 0% and 0% at willingness to pay thresholds of £30,000 and £50,000 per QALY respectively

The results for the PSA differ from those presented for the deterministic base case, with the PSA ICERs for atezolizumab about 10-20% higher than for the deterministic results. The first-line and second-line cost effectiveness acceptability curves are shown in Figure 19 and Figure 20. The probability of first-line atezolizumab being cost-effective is 10.9% and 53.9% at a willingness to pay thresholds of £30,000 and £50,000 per QALY respectively. The probability of second-line atezolizumab being cost-effective is 0% and 0% at willingness to pay thresholds of £30,000 and £50,000 per QALY respectively

Table 38 Probabilistic sensitivity analysis results for first-line treatment

	Costs	QALYs	ICER (£/QALY)
Atezolizumab	£82,893	2.775	
Gemcitabine + carboplatin	£20,605	1.467	£47,593

ICER: incremental cost-effectiveness ratio; QALY: quality-adjusted life year

.

Table 39 Probabilistic sensitivity analysis results for second-line treatment

	Costs	QALYs	ICER (£/QALY)
Atezolizumab	£74,165	1.26	
Docetaxel	£10,621	0.82	£143,144
Paclitaxel	£18,075	0.83	£129,333
BSC	£5,637	0.58	£101,247

ICER: incremental cost-effectiveness ratio; QALY: quality-adjusted life year

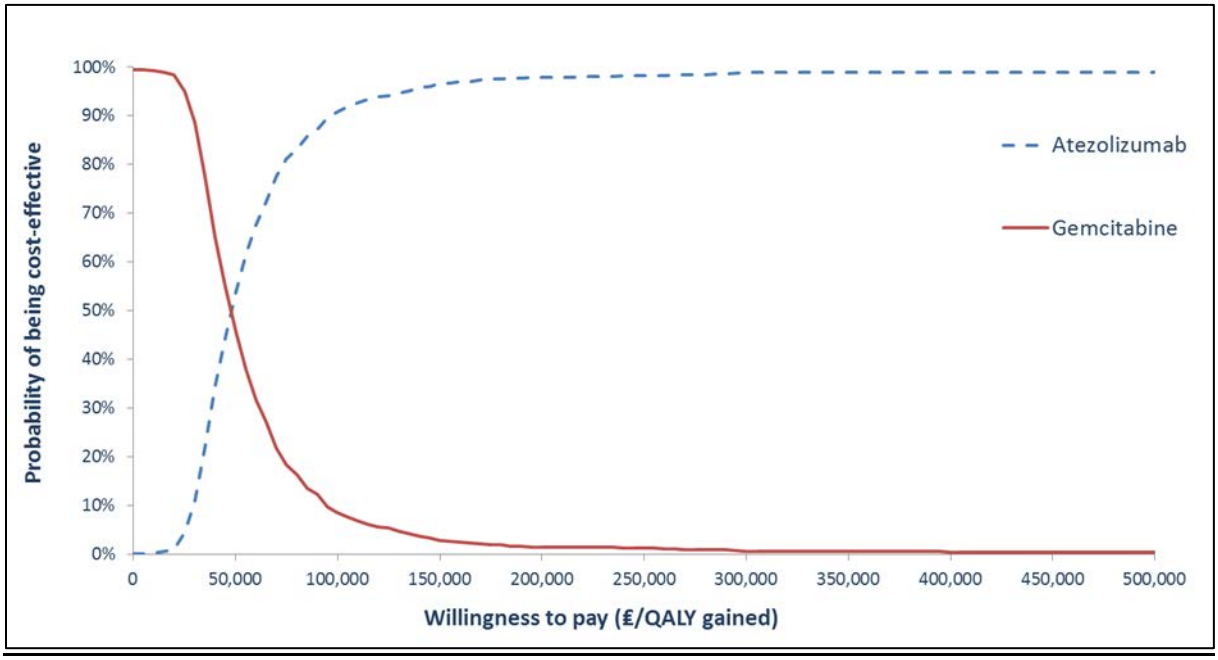


Figure 19 Cost-effectiveness acceptability curves for first-line treatment

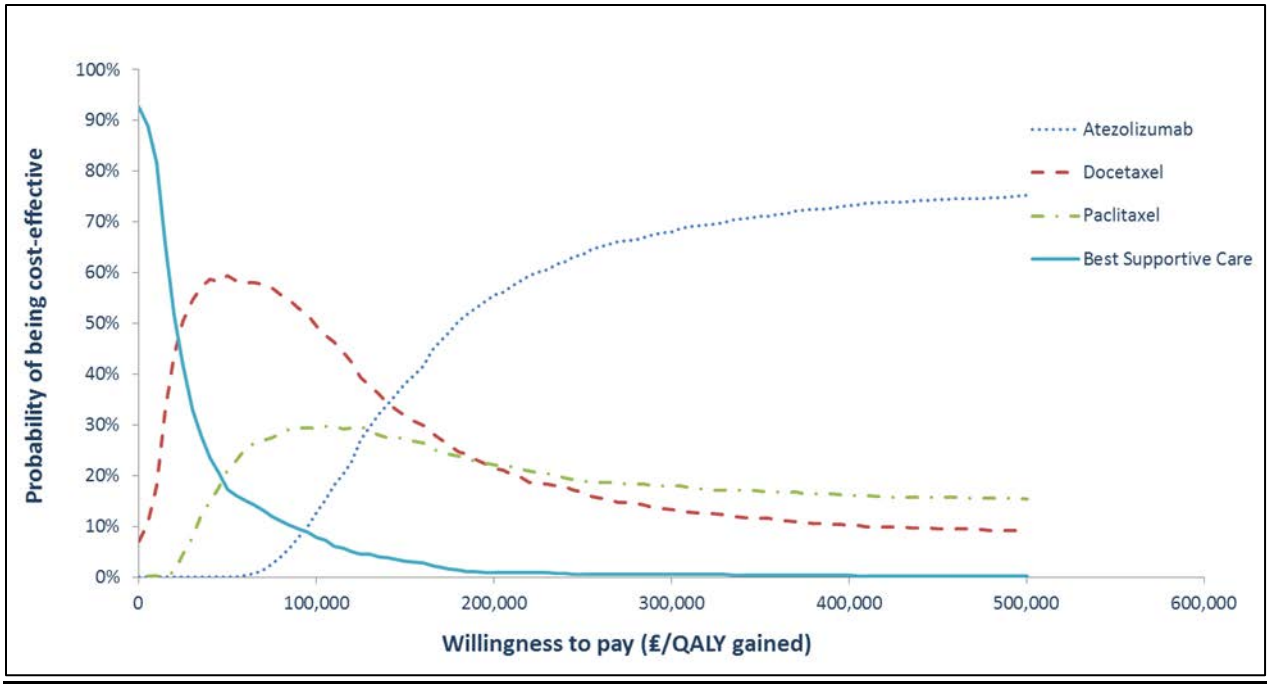


Figure 20 Cost-effectiveness acceptability curves for second-line treatment

The CS discusses the results of the PSA and states that they should be interpreted with caution, as they are unlikely to be reliable. The CS notes that there is a high level of uncertainty in the fractional polynomial model and the prediction model provides a skewed output for overall survival, which leads to an unrealistically large proportion of patients in the comparator arms surviving beyond 20 years for some of the probabilistic analyses.

4.4 Additional work undertaken by the ERG

This section details the ERG's further exploration of the issues and uncertainties raised in the review and critique of the company's cost effectiveness analyses. This consists of five additional sensitivity analyses: i) for the parametric functions for extrapolating TTD and overall survival, ii) the treatment effect and iii) assumptions for the time point at which to cap hazard ratios; iv) varying contrast estimates and varying utility values.

i) Time to treatment discontinuation / overall survival extrapolation

The CS does not contain sensitivity analyses for different parametric distributions for TTD and overall survival. These were varied by the ERG for alternative plausible parametric distributions for first-line and second-line treatment comparisons in Table 40 and Table 41. The model allows

the use of the Kaplan-Meier data for the first part of the survival curve, followed by a parametric function for the extrapolation of the tail of the curve. Changing the parametric distributions for TTD and overall survival has a significant effect on the model results. Changing both parametric functions for TTD and overall survival shows there is considerable uncertainty in the model results. For example, with the log-logistic function for TTD and the Weibull function for overall survival, the ICER increases from the base case of £44,158 to £124,485 per QALY for first-line atezolizumab compared to gemcitabine + carboplatin. For second-line comparisons, with the log-logistic function for TTD and lognormal function for overall survival, the ICER increases from the base case of £104,850 to £165,527 per QALY for atezolizumab compared to paclitaxel. As shown in Table 40, other choices of parametric distribution produce even higher ICERs.

Table 40 ERG sensitivity analyses selecting different parametric functions for extrapolating TTD and overall survival for first-line treatment

First-line		
Parameter	**Value**	**ICER (£/QALY) vs gemcitabine + carboplatin**
TTD	Base case (gamma)	£44,158
	Weibull	£42,683
	Log-logistic	£66,750
OS	Base case (cure generalised gamma)	£44,158
	Log-logistic	£51,387
	K-M + Weibull tail	£79,592
	K-M + Gompertz tail	£101,711
TTD / OS	Base case	£44,158
	TTD: log-logistic; OS: K-M + Weibull	£124,485
	TTD: log-logistic; OS: K-M + Gompertz	£159,590

ICER: incremental cost-effectiveness ratio; OS: overall survival; QALY: quality-adjusted life year; TTD: time to treatment discontinuation

Table 41 ERG sensitivity analyses selecting different parametric functions for extrapolating TTD and overall survival for second-line treatment

Second-line				
Parameter	**Value**	**ICER (£/QALY) vs docetaxel**	**ICER (£/QALY) vs paxlitaxel**	**ICER (£/QALY) vs BSC**
TTD	Base case (gamma)	£131,579	£104,850	£98,208
	Weibull	£119,025	£93,370	£89,322
	Log-logistic	£180,213	£149,491	£133,035
OS	Base case (cure generalised gamma)	£131,579	£104,850	£98,208
	Lognormal	£172,146	£131,214	£120,612
	Log-logistic	£149,321	£117,785	£110,144
	K-M + Weibull tail	£287,175	£176,090	£153,806
	K-M + Gompertz tail	£310,246	£182,347	£158,396
TTD / OS	Base case	£131,579	£104,850	£98,208
	TTD log-logistic; OS lognormal	£211,180	£165,527	£147,261
	TTD log-logistic; OS K-M + Weibull tail	£302,826	£187,599	£162,359
	TTD log-logistic; OS K-M + Gompertz tail	£324,116	£192,246	£165,707

ICER: incremental cost-effectiveness ratio; OS: overall survival; QALY: quality-adjusted life year; TTD: time to treatment discontinuation

ii) Treatment effect

The CS does not contain sensitivity analyses varying the treatment effect of atezolizumab. The ERG varied the treatment effect according to the lower and upper bounds of the contrast estimates for overall survival. The contrast estimates consist of two parameters: intercept and slope from the fractional polynomial model and bounds have been provided for both these parameters. It is unclear which values should be used when varying the contrast estimates, so the intercept parameter values have been varied only and the slope parameter kept constant. The effect of varying these parameters is shown in Table 42. The sensitivity analyses show that the ICER varies substantially at the lower and upper bounds. For the first-line comparison, the ICER varies between £33.432 and £191,793 per QALY gained for atezolizumab compared to gemcitabine + carboplatin. For second-line comparisons, atezolizumab is dominated by its comparator using the intercept lower bound (i.e. atezolizumab is more expensive and less

effective than its comparators). Using the intercept upper bound, the ICER for atezolizumab is £87,990 versus docetaxel, £68,427 versus paclitaxel and £79,017 versus best supportive care. For comparison, we have also included a sensitivity analysis for first-line treatment using the upper and lower confidence interval for the hazard ratio assuming proportional hazards. Using these values, there is a much smaller variation in ICER than for the analysis with the fractional polynomial contrast estimates.

Table 42 ERG sensitivity analyses comparing atezolizumab vs comparators for treatment effect

Parameter	First-line	ICER (£/QALY)		
		vs gemcitabine + carboplatin		
Treatment effect, OS	Fractional polynomial	£44,158		
	Fractional polynomial (Intercept lower bound)	£191,793		
	Fractional polynomial (Intercept higher bound)	£33,432		
	Proportional hazard, HR = 0.62	£46,562		
	HR = 0.47	£36,488		
	HR = 0.82	£87,898		
	Second-line	**vs docetaxel**	**vs paxlitaxel**	**vs BSC**
	Base case	£131,579	£104,850	£98,208
	Fractional polynomial (Intercept lower bound)	Dominated [a]	Dominated [a]	Dominated [a]
	Fractional polynomial (Intercept higher bound)	£87,990	£68,427	£79,017

BSC: best supportive care; HR: hazard ratio; ICER: incremental cost-effectiveness ratio; OS: overall survival; QALY: quality-adjusted life year
[a] Atezolizumab is more expensive and less effective than its comparators

iii) Capping of hazard ratios

As discussed in section 4.3.5, the ERG has some concerns around the parameter estimates derived from the network meta-analysis using the fractional polynomial model approach. The company caps the hazard ratio at different time points for first-line and second-line comparisons. The ERG investigated changing the time point at which the hazard ratios are capped and reducing the contrast estimate slope parameter so that it is no longer necessary to cap the hazard ratios.

The effects of changing the time point at which the hazard ratios are capped are shown in Table 43. The time points were varied so that they are the same for first-line and second-line comparisons. The results show that for the second-line comparison of atezolizumab versus docetaxel there is a large impact on the ICER, which increases to £310,395 per QALY.

Table 43 ERG sensitivity analyses varying the time until hazard ratios are capped

Parameter	First-line	ICER (£/QALY)		
		vs gemcitabine + carboplatin		
Time to cap hazard ratios	8 months (base case)	£44,158		
	21.16 months	£35,764		
	Second-line	**vs docetaxel**	**vs paxlitaxel**	**vs BSC**
	21.16 months (base case)	£131,579	£104,850	£98,208
	8 months	£310,395	£107,514	£97,397

BSC: best supportive care; ICER: incremental cost-effectiveness ratio; QALY: quality-adjusted life year

iv) Reducing the slope parameter for the contrast estimates

The effect of reducing the slope contrast estimate so that capping the hazard ratios is no longer needed is shown in Table 44. The time to cap the hazard ratio was increased to 20 years (i.e. at the end of the model duration). As for the preceding analysis, the largest effect of varying the slope parameter is for the second-line comparison between atezolizumab and docetaxel, with the ICER increasing to £193,686 per QALY.

Table 44 ERG sensitivity analyses varying the slope parameter

Parameter	First-line	ICER (£/QALY)		
		vs gemcitabine + carboplatin		
Slope parameter estimate	0.051 (base case)	£44,158		
	0.01	£47,505		
	Second-line	**vs docetaxel**	**vs paxlitaxel**	**vs BSC**
	0.044 (base case)	£131,579	£104,850	£98,208
	0.02	£193,686	£101,835	£99,417

BSC: best supportive care; ICER: incremental cost-effectiveness ratio; QALY: quality-adjusted life year

v) Utility values

The assumptions used by the company for health state utility values differed from the advice received by the ERG from their clinical expert. We considered that patients on-treatment with

atezolizumab would have a higher HRQoL than those on gemcitabine + carboplatin, docetaxel or paclitaxel. The CS and the ERG's assumption for the utility values for the on-treatment and off-treatment utility values for the pre-progression health state are shown in Table 45. The results of the sensitivity analyses using the ERG's assumption for the utility values are shown in Table 46. The ICER decreases slightly for the analyses for atezolizumab compared to gemcitabine + carboplatin (first-line), docetaxel and paxlitaxel (second-line) and increases slightly for atezolizumab compared to best supportive care (second-line).

Table 45 Pre-progression utility values used in the CS and the ERG analysis

	CS Pre-progression utility		**ERG pre-progression utility values**	
	Atezolizumab	Comparators	Atezolizumab	Comparators
On-treatment	0.75	0.75	0.75	0.71
Off-treatment	0.71	0.75	0.75	0.75

Table 46 ERG sensitivity analyses with changes to the assumptions for pre-progression health state utility values

Parameter	**First-line**	**ICER (£/QALY)**		
		vs gemcitabine + carboplatin		
Utility values	Base case	£44,158		
	ERG assumption	£43,317		
	Second-line	**vs docetaxel**	**vs paxlitaxel**	**vs BSC**
	Base case	£131,579	£104,850	£98,208
	ERG assumption	£127,528	£101,654	£99,409

BSC: best supportive care; ICER: incremental cost-effectiveness ratio; QALY: quality-adjusted life year

ERG base case analysis

Table 47 lists the assumptions used for the ERG base case, along with their justifications. The first-line treatment results for the ERG base case are shown in Table 48 and the second-line treatment results in Table 49. The ERG considers this presents the most representative analysis of the available evidence for atezolizumab for first- and second-line treatment compared to its comparators.

Table 47 Assumptions for the ERG base case analysis

Treatment line	Parameter	Value	Justification
First- and second-line	Utility	As shown in Table 45	Clinical expert advice to ERG
First-line	OS	K-M + exponential tail	Best fit for atezolizumab and gemcitabine + carboplatin
	TTD	Weibull	Best fit according to AIC and/ BIC
Second-line	OS	KM + Weibull tail	Best fit for atezolizumab and BSC
	TTD	Log-logistic	Best fit according to AIC and BIC

BSC: best supportive care; ICER: incremental cost-effectiveness ratio; K-M: Kaplan-Meier; OS: overall survival; QALY: quality-adjusted life year; TTD: time to treatment discontinuation; AIC Akaike Information Criteria; BIC Bayesian Information Criteria

Table 48 ERG first-line base case analysis results

	Costs	Incremental costs	QALYs	Incremental QALYs	ICER (£/QALY)
Atezolizumab	£60,650		1.32		
Gemcitabine + carboplatin	£12,469	£48,181	0.81	0.51	£93,948

ICER: incremental cost-effectiveness ratio; QALY: quality-adjusted life year

The ERG base case ICER for first-line atezolizunab compared to gemcitabine + carboplatin is £93,948 per QALY gained. The overall survival curves for first-line treatment for the observed trial data compared with the company's fitted curves and the ERG's base case are shown in Figure 21.

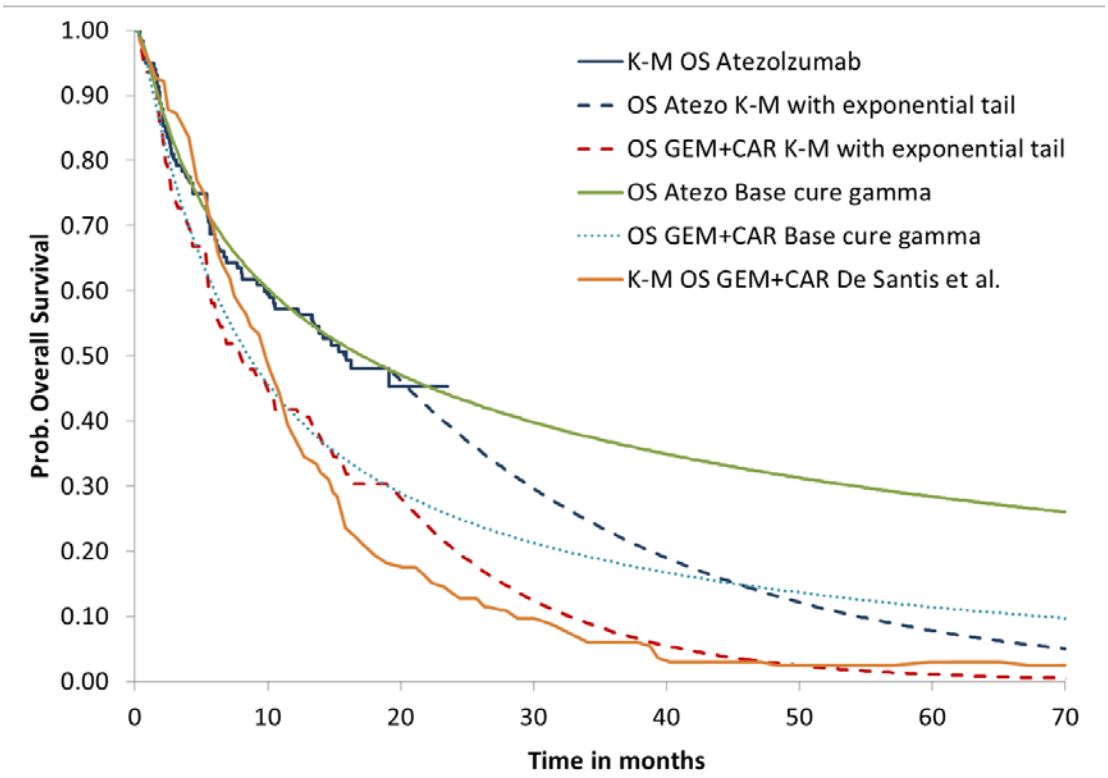


Figure 21 Overall survival curves for first-line treatment for observed trial data compared with company's fitted curves and ERG's base case

Table 49 ERG second-line base case analysis results

	Costs	Incremental costs	QALYs	Incremental QALYs	ICER (£/QALY)
Atezolizumab	£66,254		0.84		
Docetaxel	£8,196	£58,059	0.64	0.20	£288,247
Paclitaxel	£13,615	£52,640	0.55	0.29	£180,901
BSC	£4,090	£62,164	0.47	0.37	£166,805

BSC: best supportive care; ICER: incremental cost-effectiveness ratio; QALY: quality-adjusted life year

The ERG base case ICER for second-line atezolizumab compared to docetaxel, paclitaxel and best supportive care is £288,247, £180,901 and £166,805 per QALY gained respectively. The overall survival curves for second-line treatment for atezolizumab compared to best supportive care for the observed trial data compared with the company's fitted curves and the ERG's base case are shown in Figure 22.

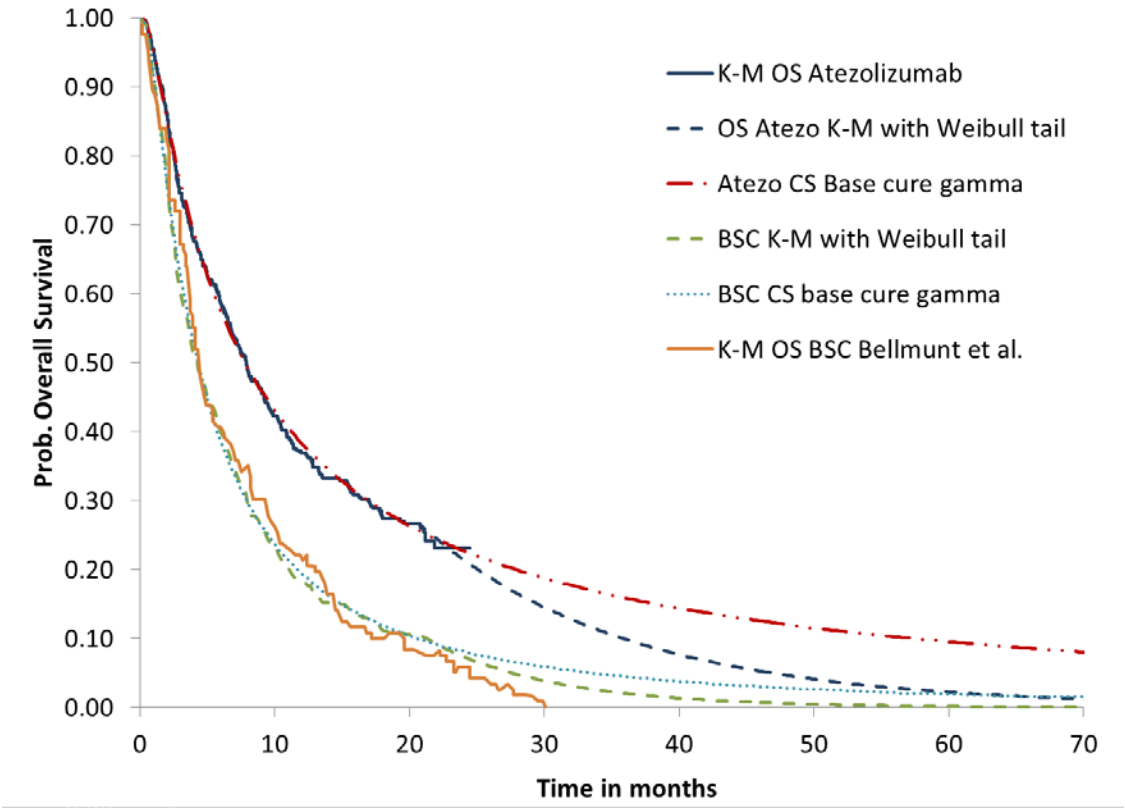


Figure 22 Overall survival curves for second-line treatment for observed trial data for atezolizumab and best supportive care compared with company's fitted curves and ERG's base case

4.5 Conclusions on cost effectiveness

The company used a model structure commonly used for economic models of cancer treatment with health states for progression-free survival, progression and death. The ERG considers the model structure to be appropriate for the decision problem.

The company used methods that are consistent with NICE methodological guidelines. The population differs from that specified from the NICE scope as the second-line treatment combines two populations: people whose disease has progressed after platinum-based chemotherapy and people for whom cisplatin-based chemotherapy is unsuitable; and those whose disease has progressed after platinum-based therapy. The comparators differ from those specified in the NICE scope as the CS does not include retreatment with first-line platinum-based chemotherapy for patients who have progressed.

The core clinical evidence for atezolizumab was from single-arm studies and there are no direct head-to-head studies between atezolizumab and its comparators. There is a weak evidence base for the comparator treatment with most studies including small number of patients. The clinical data for atezolizumab is from the phase II single-arm iMvigor 210 study.

The company comparison between atezolizumab and its comparator uses contrast estimates from the company's network meta-analysis that used a fractional polynomial model approach. The ERG has identified a number of methodological issues with the company's network meta-analysis that cast doubt on the validity of the results of the analyses. However, we note that, in general, the key driver of the model is the choice of parametric function used to extrapolate overall survival and TTD. We also note that the company has not fully explored the uncertainty around overall survival and TTD through the use of sensitivity analyses. Further, the company has chosen parametric functions for overall survival and TTD that are most favourable to atezolizumab. The ERG considers that other parametric functions are also plausible and these result in atezolizumab being much less cost-effectiveness than reported in the CS base case.

5 END OF LIFE

According to the NICE criteria for End of life, the following criteria should be satisfied:

- The treatment is indicated for patients with a short life expectancy, normally less than 24 months and;
- There is sufficient evidence to indicate that the treatment offers an extension to life, normally of at least an additional 3 months, compared to current NHS treatment.

The company has considered the criteria for end of life. The CS states that median survival with or without treatment with systemic therapy is between 8-15 months.

The company considers that the mean overall survival results better reflect the outcomes of patients and the mean results are more than 3 months for atezolizumab, when taking results from the economic analysis, as shown in Table 50.

Table 50 Mean and median survival for atezolizumab compared to comparators (CS section 4.13.3)

		Mean	Median
First-line	Atezolizumab	55.3 months	17.1 months
	Gemcitabine + carboplatin	25.1 months	8.5 months
Second-line	Atezolizumab	22.7 months	7.9 months
	Docetaxel	12.9 months	7.6 months
	Paclitaxel	12.2 months	5.3 months
	BSC	9.4 months	4.4 months

The ERG notes that if the median overall survival results are used for both end-of-life criteria, atezolizumab in second-line would not meet the criteria for extension of life as it does not extend overall survival by more than 3 months. If the mean overall survival results are used for both end-of-life criteria, atezolizumab does not meet the criteria for a short life expectancy as the mean overall survival survival for gemcitabine + carboplatin is greater than 2 years. Therefore we consider it is uncertain whether both first-line and second-line atezolizumab has met the end-of-life criteria.

6 INNOVATION

The company makes the case for innovation in CS section 2.5. They state that as the first immunotherapy for locally advanced or metastatic urothelial carcinoma, atezolizumab represents a 'new paradigm' in treatment and is a clinically significant innovative therapeutic option. The ERG notes that a NICE appraisal is currently in development for another immunotherapy for urothelial cancer, pembrolizumab (ID1019). The CS summarises recent advances in conventional chemotherapy that have resulted in gains in progression-free survival but not overall survival, or improvements in tolerability only. It asserts that in contrast, atezolizumab exploits evolutionary mechanisms that can maintain responses in some patients.

Atezolizumab has been granted 'breakthrough therapy designation' by the US FDA in 2014 (granted to potential new drugs where early clinical evidence suggests substantial improvement compared with existing therapies) and 'Promising Innovative Medicine' by the Medicines and Healthcare Products Regulatory Authority in 2016. It was considered under the early access to medicines scheme (EAMS), which aims to give patients with life threatening or seriously debilitating conditions access to medicines that do not yet have a marketing authorisation when

there is a clear unmet medical need. A positive EAMS scientific opinion was issued by the MHRA in January 2017:

'*Atezolizumab has been shown to slow the progression of cancer and increase patient survival in a condition where other treatments currently have poor results (about 20% of patients alive after 12 months). With regard to the medicine's side effects, the most frequent were mild to moderate in severity and less frequent than with chemotherapy. Advanced cancer of the bladder and urinary system is a fatal condition and currently few therapies are available with low efficacy*'

The MHRA also noted that the effects of atezolizumab have not been compared to those of current treatments in the same study, and that the company has committed to provide further data when they become available.[66]

7 DISCUSSION

7.1 Summary of clinical effectiveness issues

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.

Limitations

There are methodological weaknesses in the company's network meta-analysis and in the simulated treatment comparison which supports it, as discussed in detail in section 3.1.7. The company acknowledges that the results of the analysis are limited by lack of studies. Hazard ratios for overall survival gave implausible results when included in the economic model without adjustment, whilst hazard ratios for progression-free survival also gave implausible results and were not used in the economic analysis. Results of the meta-analysis are not discussed by the company as evidence for the clinical effectiveness of atezolizumab.

Uncertainties

The company has not provided any 'reality checks' to gauge whether their analysis results might be reasonable or subject to bias. Uncertainties arising at different steps of the analysis are not discussed or propagated through to the final results so the cumulative impact of small errors and inconsistencies identified by the ERG is unclear.

The CS acknowledges the complexity of the fractional polynomial model approach (section 4.10.10) and the very limited evidence base to which it could be applied (CS section 4.10.11.1) which suggests that the fractional polynomial method may not have been the most appropriate approach to use. Other possible approaches for analysing the data (e.g. using an accelerated failure time model) were not considered.

Given that fractional polynomial network meta-analysis is a relatively complex method that involves numerous computational steps, it is important that the analysis approach is reported clearly and as fully as possible. The company's description of the methods is rather limited and it is possible that some methodological issues might have gone undiscovered by the ERG (several aspects of the methodology were only revealed indirectly in clarification responses).

7.2 Summary of cost effectiveness issues

The CS includes evidence on the cost effectiveness of atezolizumab for patients with advanced or metastatic urothelial carcinoma. Treatment with atezolizumab is compared to gemcitabine + carboplatin for 1st line treatment and compared to docetaxel, paclitaxel, and best supportive care for 2nd line treatment. The model structure adopted is generally appropriate and consistent with the clinical disease pathway. The model contains health states of progression-free, progressed disease and death and uses survival curves for progression-free survival and overall survival, based upon clinical evidence. The clinical evidence comprises of single-arm studies which leads to considerable uncertainty. The CS acknowledges the uncertainty around the model results and the weak evidence base for the comparator trials and states that much of this uncertainty will be resolved through on-going phase III trials. On this basis, the company proposes that atezolizumab be made available for patients via the Cancer Drugs fund.
The CS base case for first-line atezolizumab compared to gemcitabine + carboplatin is £44,158 per QALY gained. The ICERs for second-line atezolizumab are £131,579 versus docetaxel, £104,850 versus paclitaxel and £98,208 versus best supportive care. The CS included

deterministic sensitivity analyses for selected input parameters and scenario analyses. However, the CS does not include sensitivity analyses varying the parametric survival curves chosen for overall survival and TTD and these are shown to have a large impact on model results. The company's probabilistic sensitivity analyses showed that the probability of first-line atezolizumab being cost-effective is 10.9% and 53.9% at willingness to pay thresholds of £30,000 and £50,000 per QALY respectively. The probability of second-line atezolizumab being cost-effective is 0% and 0% at willingness to pay thresholds of £30,000 and £50,000 per QALY respectively.

The ERG conducted sensitivity analyses evaluating alternative parametric survival functions for overall survival and TTD, different assumptions for utility estimates and varying the treatment effect of atezolizumab. The ERG's alternative base case analysis for first-line atezolizumab compared to gemcitabine + carboplatin is £93,948 per QALY and for second-line atezolizumab compared to docetaxel, paclitaxel and best supportive care is £288,247, £180,901 and £166,805 per QALY respectively. However, the ERG considers there is considerable uncertainty in the model results.

8 REFERENCES

1. CRUK (Cancer Research UK). Types of bladder cancer: CRUK; 2017 [Available from: http://about-cancer.cancerresearchuk.org/about-cancer/bladder-cancer/types-stages-grades/types].
2. Rouprêt M, Babjuk M, Böhle A, Burger M, Compérat E, Cowan N, et al. Guidelines on urothelial carcinomas of the upper urinary tract. European Association of Urology, 2015.
3. Macmillan Cancer Support. Staging and grading for early bladder cancer 2017 [Available from: http://www.macmillan.org.uk/cancerinformation/cancertypes/bladder/symptomsdiagnosis/gradingstaging/earlygradingstaging.aspx].
4. Macmillan Cancer Support. Grading and staging for invasive and advanced bladder cancer 2017 [Available from: http://www.macmillan.org.uk/cancerinformation/cancertypes/bladder/symptomsdiagnosis/gradingstaging/invasivegradingstaging.aspx].

5. Witjes J, Compérat E, Cowan N, De Santis M, Gakis G, James N, et al. Guidelnes on muscle-invasive and metastatic bladder cancer. EAU (European Association of Urology); 2015. p. 1-60.
6. CRUK (Cancer Research UK). Locally advanced bladder cancer: CRUK; 2017 [Available from: http://www.cancerresearchuk.org/about-cancer/bladder-cancer/treatment/locally-advanced].
7. CRUK (Cancer Research UK). Bladder cancer risks and causes: CRUK; 2017 [Available from: http://about-cancer.cancerresearchuk.org/about-cancer/bladder-cancer/risks-causes].
8. CRUK (Cancer Research UK). Bladder cancer statistics: CRUK; 2017 [Available from: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bladder-cancer].
9. Witjes J, Lebret T, Compérat E, Cowan N, De Santis M, Bruins H, et al. Updated 2016 EAU guidelines on muscle-invasive and metastatic bladder cancer. Eur Urol. 2017;71(3):462-75.
10. CRUK (Cancer Research UK). Bladder cancer incidence statistics: CRUK; 2017 [Available from: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bladder-cancer/incidence#heading-Zero].
11. Kaufman A, Shipley W, Feldman A. Bladder cancer. Lancet. 2009;374:239-49.
12. Howlader N, Noone A, Krapcho M, Neyman N, Aminou R, Waldron W, et al. SEER cancer statistics review, bladder section, 1975-2008 2011 [Available from: http://seer.cancer.gov/csr/1975_2008].
13. CRUK (Cancer Research UK). Bladder cancer mortality statistics: CRUK; 2017 [Available from: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bladder-cancer/mortality].
14. CRUK (Cancer Research UK). Bladder cancer survival statistics: CRUK; 2017 [Available from: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bladder-cancer/survival].
15. NICE (National Institute for Health and Care Excellence). Bladder cancer: diagnosis and management. NICE guideline ng2: NICE; 2015 [1-56]. Available from: http://www.nice.org.uk/guidance/ng2].
16. Sharma S, Ksheersagar P, Sharma P. Diagnosis and treatment of bladder cancer. Am Fam Physician. 2009;80(7):717-23.

17. de Vos F, de Wit R. Choosing chemotherapy in patients with advanced urothelial cell cancer who are unfit to receive cisplatin-based chemotherapy. Ther Adv Med Oncol. 2010;2(6):381-8.
18. American Cancer Society. Bladder cancer 2015 [Available from: http://www.cancer.org/cancer/bladdercancer/detailedguide/bladder-cancer-key-statistics].
19. De Santis M, Bellmunt J, Mead G, Kerst J, Leahy M, Maroto P, et al. Randomized phase II/III trial assessing gemcitabine/carboplatin and methotrexate/carboplatin/vinblastine in patients with advanced urothelial cancer "unfit" for cisplatin-based chemotherapy: phase II—results of EORTC study 30986. J Clin Oncol. 2009;27(33):5634-9.
20. Bellmunt J, Orsola A, Leow J, Wiegel T, De Santis M, Horwich A. Bladder cancer: EMSO Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2014;25:iii40-iii8.
21. Bellmunt J, Fougeray R, Rosenberg J, von der Maase H, Schultz F, 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;24:1466-72.
22. NICE (National Institute for Health and Care Excellence). Vinflunine for the treatment of advanced or metastatic transitional cell carcinoma of the urothelial tract: NICE; 2013 [Available from: https://www.nice.org.uk/guidance/ta272].
23. Lamb B, Vasdev N, Jalil R, Mcmenemin R, Hughes S, Payne H, et al. Second-line chemotherapy for advanced bladder cancer: A survey of current UK practice. Urol Oncol. 2014;32:52e.11-52.e17.
24. Food and Drug Administration. Atezolizumab for urothelial carcinoma 2016 [Available from: https://www.fda.gov/drugs/informationondrugs/approveddrugs/ucm501878.htm].
25. Chou R, Selph SS, Buckley DI, Gustafson KS, Griffin JC, Grusing SE, et al. Treatment of muscle-invasive bladder cancer: A systematic review. Cancer. 2016;122(6):842-51.
26. Giannatempo P, Pond GR, Sonpavde G, Raggi D, Naik G, Galsky MD, et al. The impact of adding taxanes to gemcitabine and platinum chemotherapy for the first-line therapy of advanced or metastatic urothelial cancer: A systematic review and meta-analysis. Eur Urol. 2016;69(4):624-33.
27. Necchi A, Pond GR, Raggi D, Giannatempo P, Vogelzang NJ, Grivas P, et al. Efficacy and safety of gemcitabine plus either taxane or carboplatin in the first-line setting of metastatic urothelial carcinoma: A systematic review and meta-analysis. Clin Genitourin Cancer. 2017;15(1):23-30.e2.

28. Raggi D, Miceli R, Sonpavde G, Giannatempo P, Mariani L, Galsky MD, et al. Second-line single-agent versus doublet chemotherapy as salvage therapy for metastatic urothelial cancer: A systematic review and meta-analysis. Ann Oncol. 2016;27(1):49-61.
29. Sonpavde G, Pond GR, Choueiri TK, Mullane S, Niegisch G, Albers P, et al. Single-agent taxane versus taxane-containing combination chemotherapy as salvage therapy for advanced urothelial carcinoma. Eur Urol. 2016;69(4):634-41.
30. Wu X-J, Zhi Y, He P, Zhou X-Z, Zheng J, Chen Z-W, et al. Comparison of single agent versus combined chemotherapy in previously treated patients with advanced urothelial carcinoma: a meta-analysis. Onco Targets Ther. 2016;15(9):1535-43.
31. NCCC (National Collaborating Centre for Cancer). Bladder cancer: diagnosis and management. NICE Guideline 2 Evidence Review: NCCC; 2015 [Available from: https://www.nice.org.uk/guidance/ng2/evidence/evidence-review-487511101].
32. Bellmunt J. Treatment of metastatic urothelial cancer of the bladder and urinary tract: Wolters Kluwer UpToDate®; 2017 [Available from: http://www.uptodate.com/contents/treatment-of-metastatic-urothelial-cancer-of-the-bladder-and-urinary-tract].
33. Yafi F, North S, Kassouf W. First- and second-line therapy for metastatic urothelial carcinoma of the bladder. Curr Oncol. 2011;18(1):e25-e34.
34. Meluch A, Greco F, Burris H, O'Rourke T, Ortega G, Steis R, et al. Paclitaxel and gemcitabine chemotherapy for advanced transitional-cell carcinoma of the urothelial tract: a phase II trial of the Minnie Pearl Cancer Research Network. J Clin Oncol. 2001;19(12):3018-24.
35. Ko Y, Canil C, Mukherjee S, Winquist E, Elser C, Eisen A, et al. Nanoparticle albumin-bound paclitaxel for second-line treatment of metastatic urothelial carcinoma: a single-group, multicentre, phase 2 study. Lancet Oncol. 2013;14:769-76.
36. Vaishampayan U, Faulkner J, Small E, Redman B, Keiser W, Petrylak D, et al. Phase II trial of carboplatin and paclitaxel in cisplatin-pretreated advanced transitional cell carcinoma: a Southwest Oncology Group study. Cancer. 2005;104:1627-32.
37. Vaughn D, Srinivas S, Stadler W, Pili R, Petrylak DP, Sternberg C, et al. Vinflunine in platinum-pretreated patients with locally advanced or metastatic urothelial carcinoma: results of a large phase 2 study. Cancer. 2009;115:4110-7.
38. Lee J, Park H, Lim J, Kwon G, Kim C, Song J, et al. Phase II study of a polyethoxylated castor oil-free, polymeric micelle formulation of paclitaxel for patients with advanced

urothelial cancer previously treated with gemcitabine and platinum. J Clin Oncol. 2011;29 suppl 7: abstr 272.

39. Lee J-L, Ahn J-H, Park S, LIm H, Kwon J, Ahn S, et al. Phase II study of a cremophor-free, polymeric micelle formulation of paclitaxel for patients with advanced urothelial cancer previously treated with gemcitabine and platinum. Invest New Drugs. 2012;30:1984-90.
40. Balar A, Galsky M, Rosenberg J, Powles T, Petrylak D, Bellmunt J, et al. Atezolizumab as first-line treatment in cisplatin-ineligible patients with locally advanced and metastatic urothelial carcinoma: single-arm, multicentre, phase 2 trial. Lancet. 2017;389:67-76.
41. Rosenberg J, Hoffman-Censits J, Powles T, van der Heijden M, Balar A, Necchi A, et al. Atezolizumab in patients with locally advanced and metastatic urothelial carcinoma who have progressed following treatment with platinum-based chemotherapy: a single-arm, multicentre, phase 2 trial. Lancet. 2016;387:1909.
42. Bamias A, Lainakis G, Kastritis E, Antoniou N, Alivizatos G, Koureas A, et al. Biweekly carboplatin/gemcitabine in patients with advanced urothelial cancer who are unfit for cisplatin-based chemotherapy: report of efficacy, quality of life and geriatric assessment. Oncology. 2007;73:290-7.
43. De Santis M, Bellmunt J, Mead G, Kerst J, Leahy M, Daugaard G, et al. Randomized phase II/III trial comparing gemcitabine/carboplatin (GC) and methotrexate/ carboplatin/vinblastine (MCAVI) in patients (pts) with advanced urothelial cancer (UC) unfit for cisplatin based chemotherapy (CHT): Phase III results of EORTC study 30986. J Clin Oncol. 2010;28: 18s suppl: abstr LBA4519.
44. De Santis M, Bellmunt J, Mead G, Kerst J, Leahy M, Maroto P, et al. Randomized phase II/III trial assessing gemcitabine/carboplatin and methotrexate/carboplatin/vinblastine in patients with advanced urothelial cancer who are unfit for cisplatin-based chemotherapy: EORTC study 30986. J Clin Oncol. 2012;30(2):191-9.
45. Bellmunt J, Théodore C, Demkov T, Komyakov B, Sengelov L, Daugaard G, et al. Phase III trial of influnine 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.
46. Noguchi M, Matsumoto K, Uemura H, Arai G, Eto M, Naito S, et al. Randomized phase II study of personalized peptide vaccination in patients with advanced bladder cancer progressing after chemotherapy. J Clin Oncol. 2014;32: 5 suppl; abstr 3092.

47. Noguchi M, Matsumoto K, Uemura H, Arai G, Eto M, Naito S, et al. An open-label, randomized phase II trial of personalized peptide vaccination in patients with bladder cancer that progressed after platinum-based chemotherapy. Clin Cancer Res. 2016;22(1):54-60.
48. Kim Y, Choi M, Hong J, Maeng C, Lee S, Hwang I, et al. A phase II trial of weekly docetaxel for second line treatment of urothelial carcinoma. J Clin Oncol. 2013;31 suppl: abstr e15613.
49. Kim Y, Lee S, Park S, Park S, Hwang I, Lee S, et al. A phase II study of weekly docetaxel as second-line chemotherapy in patients with metastatic urothelial carcinoma. Clin Genitourin Cancer. 2016;14(1):76-81.
50. Choueiri T, Ross R, Jacobus S, Vaishampayan U, Yu E, Quinn D, 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.
51. Albers P, Siener R, Hartlein M, Fallahi M, Haeutle D, Perabo F, et al. Gemcitabine monotherapy as second-line treatment in cisplatin-refactory transitional cell carcinoma - prognostic factors for response and improvement of quality of life. Onkologie. 2002;25:47-52.
52. Kouno T, Ando M, Yonemori K, Matsumoto K, Shimizu C, Katsumata N, et al. Weekly paclitaxel and carboplatin against advanced transitional cell cancer after failure of a platinum-based regimen. Eur Urol. 2007;52:1115-22.
53. NIH (National Institutes of Health). Quality assessment tool for case series studies: NIH; 2014 [Available from: https://www.nhlbi.nih.gov/health-pro/guidelines/in-develop/cardiovascular-risk-reduction/tools/case_series].
54. NIH (National Institutes of Health). Quality assessment tool for before-after (pre-post) studies with no control group: NIH; 2014 [Available from: https://www.nhlbi.nih.gov/health-pro/guidelines/in-develop/cardiovascular-risk-reduction/tools/before-after].
55. Therasse P, Arbuck S, Eisenhauer E, Wanders J, Kaplan R, Rubinstein L, et al. New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. J Natl Cancer Inst. 2000;92(3):205-16.
56. Lencioni R, Llovet J. Modified RECIST (mRECIST) assessment for hepatocellular carcinoma. Semin Liver Dis. 2010;30(1):52-60.

57. Llovet J, Di Bisceglie A, Bruix J, Kramer B, Lencioni R, Zhu A, et al. Design and endpoints of clinical trials in hepatocellular carcinoma. J Natl Cancer Inst. 2008;100(10):698-711.
58. EMEA (European Agency for the Evaluation of Medicinal Products). Points to consider on multiplicity issues in clinical trials: EMEA; 2002 [Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003640.pdf].
59. 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.
60. Phillippo D, Ades A, Dias S, Palmer S, Abrams K, Welton N. NICE DSU Technical Support Document 18: Methods for population-adjusted indirect comparisons in submisions to NICE. Sheffield: Decision Support Unit, ScHARR, University of Sheffield, 2016.
61. ESMO (European Society for Medical Oncology). Performance scales: Karnofsky & ECOG scores: ESMO; 2017 [Available from: http://oncologypro.esmo.org/Guidelines-Practice/Practice-Tools/Performance-Scales].
62. Sterne J, White I, Carlin J, Spratt M, Royston P, Kenward M, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ. 2009;338:b2393:1-5.
63. Jansen J. Network meta-analysis of survival data with fractional polynomials. BMC Med Res Methodol. 2011;11:61:1-14.
64. Turner R, Jackson D, Wei Y, Thompson S, Higgins J. Predictive distributions for between-study heterogeneity and simple methods for their application in Bayesian meta-analysis. Stat Med. 2015;34:984-98.
65. Jansen J, Trikalinos T, Cappelleri J, Daw J, Andes S, Eldessouki R, et al. Indirect treatment comparison/network meta-analysis study questionnaire to assess relevance and credibility to inform health care decision making: An ISPOR-AMCP-NPC Good Practice Task Force report. Value Health. 2014;17:157-73.
66. MHRA (Medicines and Healthcare products Regulatory Agency). Early access to medicines scheme (EAMS) scientific opinion: Atezolizumab for locally advanced or metastatic urothelial carcinoma: MHRA; 2017 [Available from: https://www.gov.uk/government/publications/early-access-to-medicines-scheme-eams-scientific-opinion-atezolizumab-for-locally-advanced-or-metastatic-urothelial-carcinoma].

67. PBAC (Pharmaceutical benefits advisory committee). Vinflunine public summary document: PBAC; 2015 [Available from: https://www.pbs.gov.au/industry/listing/elements/pbac-meetings/psd/2015-11/files/vinflunine-psd-november-2015.pdf].
68. NICE (National Institute for Health and Care Excellence). Guide to the methods of technology appraisal 2013: NICE; 2013 [Available from: https://www.nice.org.uk/process/pmg9/resources/guide-to-the-methods-of-technology-appraisal-2013-pdf-2007975843781].
69. Bellmunt J, de Wit R, Vaughn D, et al. KEYNOTE-045: open-label, phase III study of pembrolizumab versus investigator'schoice of paclitaxel, docetaxel, or vinflunine for previously treated advanced urothelial cancer. Society for Immunotherapy of Cancer Annual Meeting; November 9-13; National Harbor, MD2016.
70. Soga N, Onishi T, Arima K, Sugimura Y. Paclitaxel carboplatin chemotherapy as a second-line chemotherapy for advanced platinum resistant urothelial cancer in Japanese cases. Int J Urol. 2007;14(9):828-32.
71. Rowen D, Brazier J, Young T, Gaugris S, Craig BM, King MT, et al. Deriving a preference-based measure for cancer using the EORTC QLQ-C30. Value Health. 2011;14(5):721-31.
72. McKenzie L, Van Der Pol M. Mapping the EORTC QLQ C-30 onto the EQ-5D instrument: the potential to estimate QALYs without generic preference data. Value Health. 2009;12(1):167-71.
73. Seal B, Shermock K, Asche C, Chang J, Kreilick C, Yao J, et al. Evaluating the cost of treating bladder cancer with and without metastases. Value Health. 2015;18(3):A201.
74. Huillard O, Goldwasser F, Morin L. Aggressiveness of care at the end of life in patients with localized and advanced bladder cancer. J Clin Oncol 2016.
75. NICE (National Institute for Health and Care Excellence). Lung cancer (non-small-cell, non-squamous, metastatic, after treatment) - nivolumab [ID900]: NICE; 2017 [Available from: https://www.nice.org.uk/guidance/indevelopment/gid-tag524].
76. NICE (National Institute for Health and Care Excellence). Pembrolizumab for treating PD-L1-positive non-small-cell lung cancer after chemotherapy: NICE; 2017 [Available from: https://www.nice.org.uk/guidance/ta428].
77. Department of Health. Drugs and pharmaceutical electronic market information (eMit): Department of Health, England, UK; 2015 [Available from:

https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit].

78. Department of Health. NHS reference costs 2015-2016: Department of Health, England, UK; 2015 [Available from: https://www.gov.uk/government/publications/nhs-reference-costs-collection-guidance-for-2015-to-2016].
79. Akaza H, Naito S, Usami M, Miki T, Miyanaga N, Taniai H. Efficacy and safety of gemcitabine monotherapy in patients with transitional cell carcinoma after cisplatin-containing therapy: a Japanese experience. Jpn J Clin Oncol. 2007;37:201-6.
80. Albers P, Park S, Niegish G, Fechner G, Steiner U, Lehmann J, et al. Randomized phase III trial of 2nd line gemcitabine and paclitaxel chemotherapy in patients with advanced bladder cancer: short-term versus prolonged treatment [German Association of Urological Oncology (AUO) trial AB 20/99]. Ann Oncol. 2011;22:288-94.
81. Han K, Joung J, Kim T, Jeong I, Seo H, Chung J, et al. Methotrexate, vinblastine, doxorubicin and cisplatin combination regimen as salvage chemotherapy for patients with advanced or metastatic transitional cell carcinoma after failure of gemcitabine and cisplatin chemotherapy. Br J Cancer. 2008;98:86-90.
82. Ikeda M, Matsumoto K, Tabata K, Minamida S, Fujita T, Satoh T, et al. Combination of gemcitabine and paclitaxel is a favourable option for patients with advanced or metastatic urothelial carcinoma previously treated with cisplatin-based chemotherapy. Jpn J Clin Oncol. 2011;41:1214-20.
83. Matsumoto K, Irie A, Satoh T, Okazaki M, Iwamura M, Baba S. Gemcitabine and paclitaxel chemotherapy as a second-line treatment for advanced or metastatic urothelial carcinoma. Int J Urol. 2007;14:1000-4.
84. Srinivas S, Guardino A. A nonplatinum combination in metastatic transitional cell carcinoma. Am J Clin Oncol. 2005;28:114-8.
85. Suyama T, Ueda T, Fukasawa S, Imamura Y, Nakamura K, Miyasaka K, et al. Combination of gemcitabine and paclitaxel as second-line therapy for advanced urothelial carcinoma. Jpn J Clin Oncol. 2009;39:244-50.
86. Vaughn D, Manola J, Dreicer R, See W, Levitt R, Wilding G. Phase II study of paclitaxel plus carboplatin in patients with advanced carcinoma of the urothelium and renal dysfunction (E2896): a trial of the Eastern Cooperative Oncology Group. Cancer. 2002;95:1022-7.

9 APPENDICES

Appendix 1 ERG summary of studies which reported Kaplan-Meier curves but were excluded by the company

Study	Comparator	K-M curves reported	Required prognostic factors			
			Age >65	Sex	Liver met	ECOG PS ≥1
Akaza 2007 [79]	GEM n=44	OS	Reported	Reported	Reported	Reported
Albers 2002 [51]	GEM n=28	OS, TTP [a]	NR	NR	reported	NR
AUO trial [80]	GEM + PTX n=96	OS	NR [b]	NR	NR	NR
Han 2008 [81]	MVAC n=30	OS	NR [b]	Reported	Reported	Reported
Ikeda 2011 [82]	GEM + PTX n=24	OS	NR [b]	Reported	Reported	Reported
Ko 2013 [35]	Nab-PTX n=47	OS, PFS	NR [b]	Reported	Reported	Reported
Kouno 2007 [52]	CAR + PTX n=31	OS, PFS	NR [b]	Reported	Reported	PS >1
Matsumoto 2007 [83]	GEM + PTX n=10	OS	NR [b]	Reported	NR	Reported
Srinivas 2005 [84]	GEM + PTX n=18	OS	NR [b]	Reported	Reported	NR
Suyama 2009 [85]	GEM + PTX n=30	OS	NR [b]	Reported	Reported	NR
Vaishampayan 2005 [36]	CAR + PTX n=44	OS, PFS	NR [b]	Reported	NR	PS >1
Vaughn 2002 [86]	CAR + PTX n=37	OS [a]	NR [b]	Reported	Reported	Reported
Vaughn 2009 [37]	VFL n=151	OS, PFS	Reported	Reported	Reported	Reported [c]

CAR: carboplatin; GEM: gemcitabine; ECOG: Eastern Collaborative Oncology Group; K-M: Kaplan-Meier; met: metastases; MVAC; methotrexate, vinblastine, doxorubicin and cisplatin; Nab: nanoparticle albumin bound; NR: not reported; OS; overall survival; PFS: progression-free survival; PS: performance status; PTX: paclitaxel; TTP: time to progression; VFL: vinflunine

[a] Reported for subgroup(s) only

[b] median and range reported, not the specified cut-off proportion (the company employed a calculation to estimate the proportion aged >65 years from the median age – see section 3.1.7

[c] reported Karnofsky score, which maps directly to ECOG score[61]

Appendix 2 Summary of study PCD4989g

The CS provides supporting results from the phase I study PCD4989g (CS Section 4.11.11.3) and therefore we have summarised the characteristics of the study here (although, as noted above, this study did not meet the company's eligibility criteria). PCD4989g was a single-arm study that aimed to assess the safety and tolerability of atezolizumab, to determine the maximum tolerated dose, to evaluate the dose-limiting toxicity, and to identify a recommended phase II dose (CS section 4.11.11). According to the study protocol (provided by the company in response to clarification questions A40 and A42), PCD4989g had a broad disease scope and included patients with locally advanced or metastatic solid tumours or haematologic malignancies. A cohort of participants with locally advanced or metastatic urothelial carcinoma within the study (n=95) is relevant to the current appraisal. In clarification response A41 the company stated that 86 of these patients initially received 15 mg/kg atezolizumab intravenously every three weeks and nine received 1200 mg intravenously every three weeks but that the protocol was amended such that all 95 patients subsequently received the fixed dose of 1200 mg. The company also stated that average weight of patients was 80kg. In these patients 15 mg/kg would give on average a total dose of 1200 mg. However, the company also stated in clarification response A41 that patients received relatively less exposure at the anticipated licensed dose of 1200 mg, without stating the magnitude of the difference.

Study characteristics

At the clinical data cut-off in March 2016 the study included 95 patients with locally advanced or metastatic urothelial carcinoma, 72 of whom (75.6%) were male and 74 (77.8%) had white ethnicity. The majority of patients were ≥65 years old, with a median age of 66.0 years (range 36-89 years). Baseline characteristics of the participants are given in CS Table 40 and we have reproduced these here in Table 51.

Table 51 Baseline characteristics of participants in study PCD4989g

Baseline characteristic	Total (n=95)	
Age	Median	66.0
	Range	36–89
Gender	Male	72 (75.8%)
Baseline ECOG PS	0	37 (38.9%)

	1	58 (61.1%)
Visceral Metastases at study entry	Yes	74 (77.9%)
Liver metastases at study entry	Yes	35 (36.8%)
Haemoglobin level <10g/dL	Yes	18 (18.9%)
Prior Therapy (Adjuvant, Neoadjuvant)	0	1 (1.1%)
	1	0 (0%)
	2	17 (17.9%)
	3	15 (15.8%)
	4	14 (14.7%)
	5	17 (17.9%)
	≥6	31 (32.6%)
Prior Therapy with Platinum Based Regimen	Cisplatin-based	73 (76.8%)
	Carboplatin-based	37 (38.9%)
Time from prior chemotherapy (≤3 months)	Yes	39 (41.9%)

Results

In the bladder cancer subgroup of the PCD4989g study the median survival was 10.1 (95% CI 7.29, 16.99) months and progression free survival was 1.8 (95% CI 1.4, 3.3) months (Table 52). The corresponding results for cohort 2 of IMvigor201 are included in Table 52 for comparison.

Table 52 Survival outcomes for bladder cancer patients in study PCD4989g

Outcome (95% CI) **(RECIST v1.1; IRF assessed)**	**Imvigor 210 cohort 2** **All patients, N = 310**	**PCD4989g** **N=94 [a]**
Overall survival, median, months	7.9 (6.7–9.3)	10.1 (7.29, 16.99)
12 months survival, %	36.9% (31.4–42.3)	NR
Progression-free survival, median, months	2.1 (2.1–2.1) [b]	1.8 (1.4, 3.3)

CI: confidence interval; IRF: independent review facility; NE: not estimable
[a]as confirmed in clarification A41, not all participants received the licensed dose, results are supportive data only.
[b] ERG unclear why confidence interval as reported in the CS has zero range

In the bladder cancer subgroup of PCD4989g, 25.5% of participants achieved an objective response (Table 53) and 9.6% achieved a complete response (investigator assessment). The duration of response was 22.1 months (investigator assessment; median duration of response was not reached for independent review facility assessment). The corresponding results for cohort 2 of IMvigor201 are included in Table 53 for comparison.

Table 53 Response outcomes for bladder cancer patients in study PCD4989g

Outcome (95% CI) (RECIST v1.1; IRF assessed unless stated)	**Imvigor 210 cohort 2 All patients, N = 310**	**PCD4989g n=94[a]**
ORR, %	15.8 (11.9–20.4) [a]	25.5 (17.09, 35.57) [c]
Complete response, %	6.1% (3.7–9.4)	9.6 (4.47, 17.40) [d]
Duration of response, % with event	34.7 [b]	Not reported
Duration of response, median months	22.6	22.1 (12.12, NE) [c,d]

CI: confidence interval; IRF: independent review facility; NE: not evaluable; ORR, objective response rate
[a]ORR per immune-modified RECIST was 19.7% (95% CI 15.4–24.6).
[b]32 participants (65.3%) were ongoing at the time of the analysis.
[c]as confirmed in clarification A41, not all participants received the licensed dose; results are supportive data only.
[d]by investigator assessment, using RECIST v1.1

Appendix 3 ERG's critical appraisal of the included studies (Table 54 to Table 56)

Table 54 CS and ERG quality assessments of atezolizumab studies

		Imvigor 210	PCD4989g
Study question or objective stated?	**CS :**	Yes	Not assessed
	ERG :	Yes	Yes
Population clearly described, including case definition?	**CS:**	Balar 2017: No; Rosenberg 2016 & CSR: Yes	Not assessed
	ERG:	Yes	Yes
Were all eligible participants that met the prespecified entry criteria enrolled? (ERG additional question)	**CS:**	Not assessed	Not assessed
	ERG:	Could not determine	Could not determine
Comment: For Imvigor 210, insufficient detail provided in the publications and CSR to determine			
Were subjects comparable? [a]	**CS:**	Balar 2017: could not determine Rosenberg 2016 & CSR: No	Not assessed
	ERG:	Yes	Yes
Was the intervention clearly described?	**CS:**	Yes	Not assessed
	ERG:	Yes	Yes
Were outcome measures clearly defined, valid, reliable and implemented consistently?	**CS:**	Yes	Not assessed
	ERG:	Yes	Yes
Were outcome assessors blinded? (ERG additional question)	**CS:**	Not assessed	Not assessed
	ERG:	Not reported [b]	No
Was the sample size sufficiently large to provide confidence in the findings? (ERG additional question)	**CS:**	Not assessed	Not assessed
	ERG:	Yes	Yes
Was the length of follow-up adequate?	**CS:**	Yes	Not assessed
	ERG:	Yes (ongoing)	Yes
Were the statistical methods well described?	**CS:**	Balar 2017: No; Rosenberg 2016 & CSR: Yes	Not assessed
	ERG:	Yes	No
Were the results well described?	**CS:**	Yes	Not assessed
	ERG:	Yes	Yes

[a] ERG assessed whether the participants were comparable to the NICE scope, unclear what was assessed by the company.
[b] independent review of the responses of all patients included a blinded review of computed tomography and/or magnetic resonance imaging scans.

Table 55 CS and ERG quality assessments of first-line comparator studies

		Bamias 2007	De Santis 2012
Study question or objective stated?	**CS :**	Yes	Not assessed
	ERG :	Yes	Yes
Population clearly described, including case definition?	**CS:**	Yes	Not assessed
	ERG:	Yes	Yes
Were all eligible participants that met the prespecified entry criteria enrolled? (ERG additional question)	**CS:**	Not assessed	Not assessed
	ERG:	Could not determine	Yes
Comment: For Bamias, insufficient detail provided in the publication to determine if all participants who were potentially eligible were enrolled.			
Were subjects comparable? [a]	**CS:**	No	Not assessed
	ERG:	Could not determine	Yes
Comment: For Bamias, CS states no previous chemotherapy allowed, but also states previous neoadjuvant or adjuvant treatment was permitted provided that there was at least a 12-month treatment-free interval; no details of prior treatment given in the baseline characteristics table.			
Was the intervention clearly described?	**CS:**	Yes	Not assessed
	ERG:	Yes	Yes
Were outcome measures clearly defined, valid, reliable and implemented consistently?	**CS:**	Yes	Not assessed
	ERG:	Yes	Yes
Were outcome assessors blinded? (ERG additional question)	**CS:**	Could not determine [b]	No [b]
	ERG:	Not reported	NR
Was the sample size sufficiently large to provide confidence in the findings? (ERG additional question)	**CS:**	Not assessed	Not assessed
	ERG:	Could not determine	Yes
Comment: For Bamias, n=34, sample size determined on response rate, not survival outcomes			
Was the length of follow-up adequate?	**CS:**	Yes	Not assessed
	ERG:	Yes	Yes
Were the statistical methods well described?	**CS:**	Yes	Not assessed
	ERG:	Yes	Yes
Were the results well described?	**CS:**	Yes	Not assessed
	ERG:	Yes	Yes

[a]ERG assessed whether the participants were comparable to the NICE scope, unclear what was assessed by the company
[b]CS appendix 8.3 p. 41 Table 2, Cochrane risk of bias for RCTs.

Table 56 CS and ERG quality assessments of second-line comparator studies

		Bellmunt 2009	Choueiri 2012	Kim 2013, 2016	Lee 2011, 2012	Noguchi 2014, 2016
Study question or objective stated?	**CS :**	Not assessed	Not assessed	Yes	Yes	Not assessed
	ERG :	Yes	Yes	Yes	Yes	Yes
Population clearly described, including case definition?	**CS:**	Not assessed	Not assessed	Yes	Yes	Not assessed
	ERG:	Yes	Yes	Yes	Yes	Yes
Were all eligible participants that met the prespecified entry criteria enrolled? (ERG additional question)	**CS:**	Not assessed	Not assessed	Not assessed	Not assessed	Not assessed
	ERG:	CD	Yes	CD	CD	Yes
Comment: For Bellmunt, Kim and Lee, insufficient detail was provided in the publications to determine if all participants who were potentially eligible were enrolled.						
Were subjects comparable? [a]	**CS:**	Not assessed	Not assessed	No	No	Not assessed
	ERG:	Yes	Yes?	Yes?	Yes	Yes
Comment: Kim 2016: includes progression after ≤ 1 platinum-based regimens (includes 3rd line)?. Choueiri 2012 includes progression after platinum-cased regimen, 3 systemic therapies and prior paclitaxel allowed?						
Was the intervention clearly described?	**CS:**	Not assessed	Not assessed	Yes	Yes	Not assessed
	ERG:	Yes	Yes	Yes	Yes	No
Comment: Noguchi 2016 gives limited details of best supportive care						
Were outcome measures clearly defined, valid, reliable and implemented consistently?	**CS:**	Not assessed	Not assessed	Yes	Yes	Not assessed
	ERG:	Yes	Yes	Yes	Yes	Yes
Were outcome assessors blinded? (ERG additional question)	**CS:**	No[b]	Unclear [b]	No [b]	No [b]	No [b]
	ERG:	Not reported	Not reported	Not reported	Not reported	Not reported
Comment: Choueiri 2012 described as double-blind, but details not reported.						
Was the sample size sufficiently large to provide confidence in the findings? (ERG additional question)	**CS:**	Not assessed	Not assessed	Not assessed	Not assessed	Not assessed
	ERG:	Yes	Yes	CD	CD	No
Comment: Kim 2016 n=31, sample size determined on ORR not survival outcomes. Lee 2012 n=37, sample sized determined on ORR. Noguchi 2016 authors note small sample size as limitation.						
Was the length of follow-up adequate?	**CS:**	Not assessed	Not assessed	Yes	Yes	Not assessed
	ERG:	Yes	Yes	Yes	Yes	Yes
Were the statistical methods well described?	**CS:**	Not assessed	Not assessed	Yes	Yes	Not assessed
	ERG:	Yes	Yes	Yes	Yes	Yes
Were the results well described?	**CS:**	Not assessed	Not assessed	Yes	Yes	Not assessed
	ERG:	Yes	Yes	Yes	Yes	Yes

CD: could not determine

[a] ERG assessed whether the participants were comparable to the NICE scope, unclear what was assessed by the company

[b] CS appendix 8.3 p. 42 Table 3, Cochrane risk of bias for RCTs.

Appendix 4 ERG's critical appraisal of the simulated treatment comparison

Recommendation (from DSU guidance[60])	ERG appraisal
Submissions using population-adjusted analyses in an unconnected network need to provide evidence that absolute outcomes can be predicted with sufficient accuracy in relation to the relative treatment effects, and present an estimate of the likely range of residual systematic error in the "adjusted" unanchored comparison. (Guidance,[60] section 4.2.4)	The CS does not discuss whether the extent of systematic error due to imbalance in unaccounted for covariates is acceptable and no estimates are presented for the degree of likely bias. The CS does, however, note caveats around the estimates and that the outcomes of the network meta-analysis are uncertain, producing clinically implausible results.
For an unanchored indirect comparison, population adjustment methods should adjust for all effect modifiers and prognostic variables. (Guidance,[60] section 4.2.5)	It is unlikely that all effect modifiers and prognostic variables have been identified. The Cox regression models contained a maximum of four identified prognostic factors (two of which did not appear to affect model fit and some of which were estimated by imputation).
Indirect comparisons should be carried out on the linear predictor scale, with the same link functions that are usually employed for those outcomes. (Guidance,[60] section 4.2.6)	The comparisons appropriately use a transformed scale; log-hazard for time to event outcomes and a log odds scale for binary outcomes.
The target population for any treatment comparison must be explicitly stated, and population-adjusted estimates of the relative treatment effects must be generated for this target population. (Guidance,[60] section 4.2.7)	The target population is explicitly stated for the two populations in the decision problem. However, the CS does not explain whether the population adjustment would deliver treatment effect estimates for that target population (e.g. the shared effect modifier assumption is not considered).
Reporting requirements (Guidance,[60] section 4.2.8): 1. The variables available in each study should be listed, along with their distributions. 2. Evidence for effect modifier status should be given, along with the proposed size of the interaction effect and the imbalance between study populations. The resulting potential bias reduction compared with a standard indirect comparison should be considered. 3. Measures of uncertainty (e.g. confidence intervals) should be presented alongside any estimates. 4. Estimates of systematic error before and after population adjustment should be presented. 5. Estimates should be presented for the appropriate target population. 6. In order to convey some clarity about the impact of any population adjustment, a crude unadjusted difference should be presented alongside the simulated treatment comparison estimate.	1. The variables available in each study along with their distributions are not presented. 2. Evidence for effect modifier status, and the proposed size of the interaction effect, are not reported. The imbalance between study populations is noted (CS section 4.10.6). The resulting potential bias reduction compared with a standard indirect comparison is not reported. 3. Measures of uncertainty: 95% credible intervals are reported, bootstrapping and Bayesian methods were used. Uncertainty around reconstructed digitised survival curves is not reported. 4. Estimates of systematic error before and after population adjustment are not presented 5. The CS does not comment on the representativeness of the aggregate population to the true target population. 6. The CS does not provide a crude unadjusted difference alongside the STC estimate for comparison (not provided in response to clarification request A15).

Appendix 5 ERG's critical appraisal of the network meta-analysis

Criterion	ERG assessment
NMA purpose	
1. Are the NMA results used to support the evidence for the clinical effectiveness of the intervention?	**No.** The Executive summary states the results are subject to uncertainty; CS section 4.13 (Interpretation of clinical evidence) does not mention the NMA.
2. Are the NMA results used to support the evidence for the cost-effectiveness of the intervention?	**Partly.** Results were used for OS but the values were capped. The results for PFS were not used.
Evidence selection	
3. Are inclusion/exclusion criteria adequately reported?	**Partly.** Criteria are specified in several different places in the CS and not applied consistently (see section 3.1.2).
4. Is quality of the included studies assessed?	**Yes**, although there are limitations with the approach taken (see section 3.1.4), and it includes studies that are not relevant to the NMA.
Methods – statistical model	
5. Is the statistical model described?	**Yes**, but only briefly
6. Has the choice of outcome measure used in the analysis been justified?	**Not explicitly**, but the most appropriate outcome for this cancer assessment, OS, was analysed and reported. Other relevant outcomes analysed were PFS, 12-month survival and ORR but of these only PFS results are reported. These outcomes could have been used to support the clinical effectiveness conclusions but were not.
7. Has a structure of the network been provided?	**No.**
8. Is homogeneity considered?	**Yes**, but only qualitatively.
9. Are the studies homogenous in terms of patient characteristics and study design?	**No.** Below CS Table 17 the CS states that "there are a number of differences between included trials that require some caution when interpreting the results, such as: differences in patient populations including baseline risk, treatment history, differences in trial designs, particularly in regard to primary efficacy outcome(s) measurements". In response to clarification question A24 the company stated that "it was necessary to include studies of heterogeneous populations due to the lack of alternative data" but the company did not refer to any specific variables. In the summary of study heterogeneity, CS Figure 4 shows "moderate" heterogeneity for 1L. In 2L, there was "low-moderate heterogeneity" for both the BSC and docetaxel comparisons, and "moderate" heterogeneity for the paclitaxel comparison, but these categories were not explained in the CS or in the company's response to clarification question A26. The CS does not provide baseline characteristics for comparators so the ERG tabulated these (Table 10 & Table 11). There are some differences between the comparator studies (e.g. patients' age; proportions with comorbidities; performance status), and also differences when

	comparing the atezolizumab cohorts against the comparator studies (e.g. proportion with visceral metastases; performance status) (section 3.1.7).
10. If the homogeneity assumption is not satisfied, is clinical or methodological homogeneity across trials in each set involved in the indirect comparison investigated by an adequate method? (e.g. sub group analysis, sensitivity analysis, meta-regression)	**No.** The CS states that sensitivity analyses were undertaken with different priors for between-study heterogeneity, but results of these are not presented. They were provided by the company in response to clarification question A30 for 2L treatment comparisons but not for 1L comparisons.
11. Is the assumption of similarity stated?	**No.** An implicit assumption is that the studies are similar since the prediction model should have matched them on key effect modifiers and prognostic variables. However, due to uncertainties around the covariates for effect modifiers and prognostic variables (section 3.1.7) it is unclear whether the similarity assumption is likely to hold.
12. Is any of the programming code used in the statistical programme provided (for potential verification)?	**Yes,** in CS appendix 8.6
Sensitivity analysis	
13. Does the study report sensitivity analyses?	**No.** The CS states that sensitivity analyses were performed with different priors and a random effects model but does not report results. The results were provided in response to clarification question A30 for 2L treatment comparisons but not for 1L comparisons.
Results	
14. Are the results of the NMA presented?	**Partly.** Results for OS are presented (CS 4.10.11.1 and 4.10.11.2) but are not discussed. PFS results are stated to be clinically implausible and are presented separately in CS Appendix 8.5 (not discussed). 12-month OS and ORR are not presented (CS states available on request).
15. Does the study describe an assessment of the model fit?	**Yes,** model fit was compared using DIC and unspecified "additional criteria" due to the complexity of the fractional polynomial models (CS p. 85. 88, 93, appendix 8.5)
16. Has there been any discussion around the model uncertainty?	**Partly.** Uncertainty is briefly mentioned in CS section 4.10.13 but the CS does not discuss all possible sources of uncertainty or consider which would have the most impact on the results.
17. Are the point estimates of the relative treatment effects accompanied by some measure of variance such as confidence intervals?	**Partly.** Unlabelled uncertainty ranges are displayed for the predicted atezolizumab OS curves (CS Figures 8-14) and log hazard function curves (CS Figures 15-17) but not explained or discussed. Upper and lower bounds of the log-hazard function (contrast estimate slope and intercept) are provided for 2L only (CS Table 24 and clarification response A30).
Discussion - overall results	
18. Does the study discuss both conceptual and statistical heterogeneity?	Partly. The CS does not explicitly discuss the types of heterogeneity present. However, the CS states that priors were used to represent between-study heterogeneity, and in clarification response A28 the company stated that random-effects models were included to allow for between-study heterogeneity. As noted above (items 8 and 9) the CS reports some aspects of conceptual heterogeneity qualitatively.
Discussion - validity	

19. Are the results from the indirect/NMA compared, where possible, to those just using direct evidence?	**Partly.** Visual naive comparisons between survival curves can be made by inspecting CS Figures 8-14. These are not discussed in detail in the CS. However, the CS does state that for 1L the predicted atezolizumab OS K-M curves were almost identical to the original OS K-M curve from cohort 1 of Imvigor 210 (CS Figures 8 and 9). In contrast, for 2L there were differences between the predicted and observed atezolizumab OS K-M curves, which the CS points out, e.g. for CS Figure 10. The company explained in clarification response A15 that the network meta-analysis consisted only of direct comparisons. They provided results for the pairwise direct comparisons analysed separately and these concur with the network meta-analysis results. This is to be expected as the same underlying fractional polynomial model was used for both analyses.

1L: first-line; 2L: second-line; NMA: network meta-analysis; OS: overall survival; PFS: progression-free survival