CONFIDENTIAL UNTIL PUBLISHED

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

Atezolizumab in combination for treating advanced non-squamous non-small-cell lung cancer

Produced by Southampton Health Technology Assessments Centre (SHTAC)

Authors Mr Olu Onyimadu, Research Fellow, SHTAC
Professor Joanne Lord, Professorial Fellow, SHTAC
Dr Joanna Picot, Senior Research Fellow, SHTAC
Dr Keith Cooper, Senior Research Fellow, SHTAC
Mr David A. Scott, Director, Diligent Agile Synthesis Limited
Dr Jonathan Shepherd, Principal Research Fellow, SHTAC

Correspondence to Dr Jonathan Shepherd
Southampton Health Technology Assessments Centre (SHTAC)
Wessex Institute
Alpha House
Enterprise Road, University of Southampton Science Park
Southampton SO16 7NS
www.southampton.ac.uk/shtac

Date completed 8th November 2018

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 127302.

Declared competing interests of the authors

None

Acknowledgements

We would like to thank Dr Judith Cave, Consultant Medical Oncologist, Southampton University Hospitals NHS Trust, who offered clinical advice and comments on the draft report. Dr Cave declares no competing interests. We would also like to thank: Karen Welch, Information Scientist, SHTAC, for appraising the literature search strategies in the company's submission, and for running searches where necessary; and Dr Jill Colquitt Senior Reviewer / Partner, Effective Evidence LLP, for providing a quality assurance review of the draft ERG 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:

Onyimadu, O., Lord, J., Picot, J., Cooper, K., Scott, D., Shepherd, J., Atezolizumab in combination for treating advanced non-squamous non-small-cell lung cancer*:* A Single Technology Appraisal*.* Southampton Health Technology Assessments Centre (SHTAC), 2018.

Contributions of authors

Olu Onyimadu critically appraised the health economic systematic review, critically appraised the economic evaluation, and drafted the report; Joanne Lord critically appraised the health economic systematic review, critically appraised the economic evaluation, and drafted the report; Joanna Picot critically appraised the clinical effectiveness systematic review, and drafted the report; Keith Cooper critically appraised the health economic systematic review, critically appraised the economic evaluation, and drafted the report; David

Scott critically appraised the network meta-analyses and drafted the report; Jonathan Shepherd critically appraised the clinical effectiveness systematic review, drafted the report, project managed the report and is the project guarantor.

Word count: 41,101

[redacted]

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AE	Adverse event
AESI	Adverse event of special interest
AIC	Akaike Information Criterion
ALK	Anaplastic lymphoma kinase
Atezo	Atezolizumab
AUC	Area under the curve
Bev	Bevacizumab
BIC	Bayesian Information Criterion
BNF	British National Formulary
BSC	Best supportive care
CARB	Carboplatin
CHMP	Committee for Medicinal Products for Human Use
CI	Confidence interval
CIS	Cisplatin
CP	Carboplatin and paclitaxel
CR	Complete response
CRD	Centre for Reviews and Dissemination
CrI	Credible interval
CS	Company submission
CSR	Clinical study report
DIC	Deviance Information Criteria
DOR	Duration of response
DSU	Decision Support Unit
ECOG PS	Eastern Cooperative Oncology Group performance status
EORTC	The European Organisation for Research and Treatment of Cancer
EPAR	European Public Assessment Report
EQ-5D	EuroQol 5-Dimension
ERG	Evidence Review Group
FDA	Food and Drug Administration
FE	Fixed effect
FP	Fractional polynomials
HR	Hazard ratio
HRQoL	Health related quality of life

ICER	Incremental cost effectiveness ratio
IRF	Independent review facility
IPD	Individual patient data
ITC	Indirect treatment comparison
ITT	Intent-to-treat population
IV	Intravenous
KM	Kaplan-Meier
LYG	Life years gained
MAIC	Matching adjusted indirect comparisons
MRI	Magnetic resonance imaging
NE	Not evaluable
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
NLCA	National Lung Cancer Audit
NMA	Network meta-analysis
NR	Not reported
NSCLC	Non-small cell lung cancer
ORR	Objective response rate
OS	Overall survival
PAC	Paclitaxel
PAS	Patient access scheme
PD	Progressive disease
PEM	Pemetrexed
PEMB	Pembrolizumab
PF	Progression free
PFS	Progression-free survival
PH	Proportional hazards
PLAC	Placebo
PR	Partial response
PRO	Patient-reported outcomes
PSA	Probabilistic sensitivity analysis
PSS	Personal social services
PSSRU	Personal Social Services Research Unit
QALY	Quality adjusted life year

QLQ	Quality of life questionnaire
QoL	Quality of life
RCT	Randomised controlled trial
RE	Random effects
RECIST	Response Evaluation Criteria in Solid Tumours
SAE	Serious adverse event
SD	Standard deviation
SmPC	Summary of Product Characteristics
STA	Single Technology Appraisal
TA	Technology appraisal
TEAE	Treatment emergent adverse events
TKI	Tyrosine kinase inhibitor
TTD	Time to treatment discontinuation
TTO	Time trade-off
WT	Wild-type

SUMMARY

Scope of the company submission

The company submission (CS) assesses the clinical effectiveness and cost effectiveness of atezolizumab (Atezo) in combination with bevacizumab (Bev), carboplatin and paclitaxel (CP) as a first-line treatment for adult patients with metastatic non-squamous, non-small cell lung cancer (NSCLC). The anticipated marketing authorisation for Atezo+Bev+CP covers all patients with first-line metastatic non-squamous NSCLC, regardless of level of programmed death-ligand 1 PD-L1 expression (an immune checkpoint protein). The scope of the CS is narrower than the anticipated marketing authorisation, focusing on two patient subgroups:

- patients with low or negative PD-L1 expression (tumour proportion score 0–49%, TC/IC 0,1,2).
- patients ineligible for, intolerant to or who have progressed on targeted therapy for epidermal growth factor receptor (EGFR) or anaplastic lymphoma kinase (ALK) tumour mutations.

Thus, patients with high PD-L1 expression, who currently would be eligible to receive pembrolizumab (NICE TA531), are not included in the CS, a deviation from the NICE scope. No cost-effectiveness comparison is made by the company with pembrolizumab in PD-L1 high expression patients. The company is therefore not seeking NHS reimbursement for treatment with atezolizumab in this patient sub-group.

[redacted] Expert clinical opinion to the Evidence Review Group (ERG) concurs with this assertion.

The CS omits the comparison of Atezo+Bev+CP to chemotherapy (docetaxel, gemcitabine, paclitaxel or vinorelbine) in combination with a platinum drug (carboplatin or cisplatin), with or without pemetrexed maintenance treatment (included in the NICE scope). Instead, Atezo+Bev+CP is compared to pemetrexed in combination with cisplatin / carboplatin, with or without pemetrexed maintenance treatment (also included in the NICE scope). The justification for the focus on this comparison is that it is the most commonly-used UK chemotherapy, based on clinical expert opinion sought by the company and UK market share data. Expert clinical advice to the ERG concurs with this assertion, but notes that in England pemetrexed should only be given in combination with cisplatin (based on NICE guidance, TA181). (though the ERG has identified recent audit data showing that some patients receive pemetrexed in combination with carboplatin). Patients who cannot tolerate cisplatin would therefore be treated with a carboplatin-based regimen (i.e. docetaxel,

gemcitabine, paclitaxel or vinorelbine in combination with carboplatin), followed by maintenance treatment with pemetrexed. This treatment regimen is not included in the CS.

Summary of submitted clinical effectiveness evidence

The company conducted a systematic review to identify and select clinical effectiveness evidence. The ERG considers that the literature searches are of good quality and are fit for purpose. The searches and inclusion criteria were designed to identify relevant randomised controlled trials (RCTs) of atezolizumab and comparator treatments for potential inclusion in a network meta-analysis (NMA).

The systematic review identified one RCT of atezolizumab, the IMpower150 RCT. This trial evaluated atezolizumab in combination with carboplatin plus paclitaxel with or without bevacizumab. A further six RCTs of comparator treatments were identified for possible inclusion in the NMA.

IMpower150 is a Phase III, open-label RCT which enrolled adult chemotherapy-naive patients with stage IV non-squamous NSCLC. Eligible patients were randomised in a 1:1:1 ratio to one of three treatment arms:

- Arm A (n=402): Atezolizumab + carboplatin + paclitaxel (Atezo+CP) induction (four or six 21-day cycles) followed by atezolizumab maintenance (21-day cycles)
- Arm B (n=400): Atezolizumab + bevacizumab + carboplatin + paclitaxel (Atezo+Bev+CP) induction (four or six 21-day cycles) followed by atezolizumab + bevacizumab maintenance (21-day cycles)
- Arm C (n=400): Bevacizumab + carboplatin + paclitaxel (Bev+CP) induction (four or six 21-day cycles) followed by bevacizumab maintenance (21-day cycles).

The marketing authorisation applied for covers Atezo+Bev+CP (Arm B) only and therefore data for Atezo+CP (Arm A) were not included in the CS. The Atezo+CP regimen was included in the NICE scope but as it is not included in the marketing authorisation NICE will not be able to include it in its guidance.

Trial results are available for two data cuts: 15th September 2017 (final progression free survival (PFS), interim overall survival (OS)), and 22nd January 2018 (updated PFS; second interim OS). The most recent data from the second interim OS results are based on 422 deaths across the two treatment arms relevant to this appraisal, with a median follow-up of approximately 20 months. Median OS has been reached in both treatment arms, however, final OS data is expected in [redacted].

The sample size calculation for the trial was performed for the co-primary endpoints of OS and PFS in two trial subgroups, the intention to treat wild type (ITT-WT) patient population (that is, all patients except those with EGFR / ALK mutations) (for the outcomes of OS and PFS) and the Teff-high WT population (for the outcome of PFS only). The latter refers to patients who had a specific T-effector (Teff) gene signature, excluding patients with an activating EGFR mutation or ALK translocation. The CS reports that, since the Atezo+Bev+CP combination demonstrated a clinically meaningful improvement in outcomes regardless of Teff gene signature status, this biomarker was not deemed to be clinically relevant and therefore these data did not impact the anticipated marketing authorisation. This population is not reported in detail in the CS, and is not mentioned in the NICE scope. Therefore, it is not discussed in this ERG report.

The analyses of clinical effectiveness and cost effectiveness in the CS are based on the ITT population and not the ITT-WT or the Teff-high WT populations. The statistical power calculation is thus based on a sample size for a WT population that is slightly smaller than the sample size for the ITT analyses presented in the CS. The ERG notes that the ITT-WT population comprises 87% (n=1040/1202) of the ITT population, so the difference in the size of these two populations is relatively small.

[redacted]

Results of the IMpower150 trial

In the ITT population at the 22nd January 2018 clinical cut-off date (minimum follow up 13.5 months, median follow-up approximately 20 months) median investigator-assessed PFS was longer in the Atezo+Bev+CP group (8.4 months, 95% CI 8.0 to 9.9) than in the Bev+CP group (6.8 months, 95% CI 6.0 to 7.0). The stratified hazard ratio was 0.59 (95% CI 0.50 to 0.69). At an earlier data cut (September 2017) independent review facility (IRF) PFS results were similar to the investigator-assessed PFS.

At the most recent data cut-off (22 January 2018) 192 deaths (48.0%) had been observed in the Atezo+Bev+CP group and 230 (57.5% deaths) in the Bev+CP group. The stratified HR for OS was 0.76 (95% CI 0.63 to 0.93) indicating that among the ITT population, patients in the Atezo+Bev+CP arm had a 24% relative reduction in the risk of death in comparison with the Bev+CP arm. The median survival of 19.8 months (95% CI 17.4 to 24.2) in the Atezo+Bev+CP arm was 4.9 months longer than the Bev+CP arm (median OS 14.9 months, 95% CI 13.4 to 17.1).

The CS also reports outcomes of response and duration of response which were also in favour of the Atezo+Bev+CP arm. Treatment with both Atezo+Bev+CP and Bev+CP was reported by patients to lead to worsening peripheral neuropathy and alopecia. A clinically meaningful improvement in cough was reported by patients in both trial arms. For other measures outcomes were deemed not to be clinically meaningful and were comparable between treatment arms.

In terms of safety, the total number of adverse events was higher in the Atezo+Bev+CP group (n=6419) compared with the Bev+CP group (n=4630). However, the proportion of patients with at least one adverse event or one treatment-related adverse event was similar between groups (patients with at least one adverse event: Atezo+Bev+CP 98.2% vs Bev+CP 99.0%; patients with at least one treatment-related adverse event Atezo+Bev+CP 94.1% vs Bev+CP 95.7%). The proportion of patients experiencing treatment-related Grade 3-4 adverse events, serious adverse events and treatment-related serious adverse event were all higher in the Atezo+Bev+CP arm compared with the Bev+CP arm.

Subgroup results of the IMpower150 trial

PFS results for the subgroup of patients with low or negative PD-L1 expression favoured Atezo+Bev+CP compared to Bev+CP, though the difference between treatments was not as strongly in favour of the Atezo+Bev+CP group as it was in the total ITT population (unstratified HR 0.66, 95% CI 0.56 to 0.79 vs. unstratified HR 0.58 95% CI 0.50 to 0.68 respectively). In comparison to the ITT population (unstratified HR) the unstratified hazard ratio for the low or negative PD-L1 expression subgroup indicates slightly worse overall survival than in the ITT group (0.80 versus 0.77) with a slightly wider confidence interval which at the upper boundary extends to 0.99 therefore falling short of the line of no effect (1.0) (95% CI 0.65 to 0.99 in the low or negative PD-L1subgroup versus 0.63 to 0.93 in the ITT population).

Median investigator assessed PFS in the EGFR/ALK+ population was longer in the Atezo+Bev+CP group (10.0 months compared to 6.1 months in the Bev+CP group). The unstratified hazard ratio indicates a difference in favour of the Atezo+Bev+CP group that is slightly better than in the total ITT population (unstratified HR 0.55, 95% CI 0.34 to 0.90 vs. unstratified HR 0.58 95% CI 0.50 to 0.68 respectively). In terms of OS, median survival has not been reached in the Atezo+Bev+CP group. There is therefore more uncertainty associated with the hazard ratio for OS and the upper bound of the confidence interval crosses the line of no effect (unstratified HR EGFR/ALK subgroup 0.54, 95% CI 0.29 to

1.03), p=0.0578 compared with ITT unstratified HR 0.77, (95% CI 0.63 to 0.93). This subgroup analysis should be interpreted cautiously due to the low number of patients included (n=104/800, 13%).

Network meta-analysis

The CS reports two indirect comparisons of atezolizumab with other treatments:

1. A network meta-analysis (NMA) comparing Atezo+Bev+CP versus pemetrexed-based chemotherapy
2. A matched adjusted indirect comparison (MAIC) comparing Atezo+Bev+CP versus pembrolizumab in patients with high PD-L1 expression.

The MAIC is not used to inform the economic evaluation as the company are not seeking NHS reimbursement for pembrolizumab in patients with high PD-L1 expression. We therefore do not provide a critical appraisal of the MAIC in this report.

A total of seven RCTs were included in the NMA, including the IMpower150 trial of atezolizumab. The structure of the OS and the PFS network is identical. Atezo+Bev+CP is compared with two pemetrexed-based regimens:

- Pemetrexed in combination with carboplatin or cisplatin, followed by pemetrexed maintenance (PEM+CARB/CIS then PEM maintenance)
- Pemetrexed in combination with cisplatin followed by placebo maintenance with best supportive care (PEM+CIS then PLAC main + BSC).

The eligibility criteria in the included trials and patient characteristics at baseline were similar with some exceptions, particularly relating to one trial (the PARAMOUNT trial) in which patients were randomised to maintenance treatment only if they had responded to induction therapy. This is in contrast to the other included trials, which required patients to have either no prior treatment for Stage III and/or IV non-squamous NSCLC. To address this, the company reports a scenario analysis in which this trial is omitted.

The company uses a fractional polynomial approach for indirect comparison estimates of OS and PFS. Unlike traditional NMA methods which assume a constant HR over time, a fractional polynomial model aims to better capture variations in the HR over time through fitting a range of polynomial models to the data. The company's justification for using the fractional polynomial approach was based on the assertion that chemotherapy and immunotherapy have different mechanisms of action leading to different survival kinetics. Patients treated with the former demonstrate early survival benefits, whilst those treated with

the latter show a delayed but more sustained survival benefit. Expert clinical advice to the ERG concurs with this assertion. The ERG therefore agrees that the use of a fractional polynomial methodology is reasonable in this appraisal.

Two orders of fractional polynomial model were considered for inclusion: first-order, and second order. A first order model with a P=0 would be equivalent to a Weibull model, and a first order model with P=1 would correspond to a Gompertz model. The best fitting fractional polynomial model chosen for OS and PFS was the fixed effect first order model with P1=0 (Weibull). This model was used in the ITT NMA as well as the subgroup and sensitivity analyses, for methodological consistency. Based on the information provided the ERG considers that the methods used to implement the fractional polynomial model are appropriate.

[redacted]

Summary of submitted cost effectiveness evidence

The company's submission includes a review of published cost-effectiveness evidence and a new economic model developed for this appraisal. The model estimates the cost-effectiveness of Atezo+Bev+CP for people with metastatic non-squamous NSCLC in comparison to pemetrexed + cisplatin (with or without pemetrexed maintenance).

Review of published economic evidence

The company conducted a systematic search for published cost-effectiveness evidence for first-line treatment of NSCLC. They reported that out of 66 economic evaluations with full publications in English, ten used data derived from the UK, of which seven were NICE technology appraisals. None of the UK economic evaluations related to the NICE decision problem for this appraisal.

Description of the company's economic model

The submitted model includes analyses for three populations:

- ITT (the IMpower150 trial population);
- untreated PD-L1 low or negative; and
- EGFR/ALK positive after targeted treatment.

For each population, treatment with Atezo+Bev+CP is compared to pemetrexed + cisplatin (with or without pemetrexed maintenance).

The model uses a partitioned survival approach with three health states (pre-progression, post-progression and death) to estimate costs and QALYs over a 20-year time horizon. Patients enter the model in the pre-progression state at the start of treatment. Rates of progression and mortality are determined by PFS and OS curves for each treatment

- Baseline PFS and OS curves for Atezo+Bev+CP are estimated from IMPower150 clinical trial data and extrapolated using parametric survival modelling.
- For the comparators, PFS and OS curves are estimated by applying time-varying hazard ratios from the company's fractional polynomial NMA to the baseline Atezo+Bev+CP curves.

The company estimated time to treatment discontinuation (TTD) for atezolizumab and bevacizumab in the Atez+Bev+CP intervention from IMPower150 data. In their base case, they assume a maximum of two years treatment with Atezo+Bev+CP with persistence of the survival advantage (relative to the comparator with pemetexed maintenance) for a further three years. Treatment with pemetrexed maintenance is assumed to persist until progression (without a stopping rule or cap on effectiveness) and other treatments are of fixed duration. Subsequent treatments are not modelled explicitly but a cost is added to post-progression state to reflect an average mix of second and subsequent treatments.

The company fitted parametric curves to OS, PFS and TTD data from the Atezo+Bev+CP arm of IMPower150 (exponential, Weibull, log-normal, log-logistic, generalised gamma and Gompertz). They also considered a piecewise approach, with Kaplan-Meier (KM) data up until 20% of patients remain at risk (n=80) and then extrapolation with the six parametric functions. The choice of curves was based on statistical and visual fit to KM data and expert opinion on the plausibility of the extrapolations.

Table 1 Company base case survival curves for Atezo+Bev+CP

Population	OS	PFS	TTD (Atezo and Bev)
ITT	Exponential (log-logistic as a "plausible alternative")	KM + log-logistic	KM + exponential
PD-L1 low/ negative			
EGFR/ALK positive		Log-normal	Exponential
Stopping rule			Maximum 2 years
Effect cap	5 years from baseline mortality rate equal to PEM+CIS/CARBO (with maintenance)		

These curves were adjusted for the two pemetrexed-based comparators (with/without maintenance) using hazard ratios from the fractional polynomial NMA. In their base case, the company used the fixed effects first-order fractional polynomial with P1=0 (Weibull), for the ITT population, and separately for the PD-L1 and EGFR/ALK subgroups.

For their base case, the company used health utilities estimated from EQ-5D-3L data collected in the IMPower150 trial with a proximity to death approach: utility estimates for <35, 34-75, 74- 210 and >211 days before death. The same utilities were applied to all populations and treatment arms and did not include disutility associated with adverse events.

The model includes resource costs associated with drug acquisition, drug administration, subsequent treatment (docetaxel, nivolumab, pembrolizumab and atezolizumab monotherapy), follow-up and monitoring, adverse events and terminal care. These were in line with previous TAs (including TA531).

Company's base case results

The company base case results are shown in Table 1, Table 2 and Table 3 for the ITT, PD-L1 negative/low and EGFR/ALK positive populations, respectively. These results include patient access scheme (PAS) price discounts for atezolizumab and bevacizumab but list prices for comparators and subsequent treatments. We show results with all available price discounts in the confidential addendum to this report.

Table 2 Company base case results, ITT population (PAS for Atezo and Bev, list price for all other treatments) – deterministic (Clarification Response Table 35)

Treatment	Total		Incremental analysis ICER (£/QALY)	Pairwise ICERs vs. comparator (£/QALY)
	Costs (£)	QALYs		
PEM+plat	[redacted]	[redacted]		£16,419
PEM+plat+PEM maint	[redacted]	[redacted]	£35,985	Dominant
Atezo+Bev+CP	[redacted]	[redacted]	Dominant	-

Table 3 Company base case results, PD-L1 negative/low population (PAS for Atezo and Bev, list price for all other treatments) – deterministic (Clarification Response Table 36)

Treatment	Total		Incremental analysis ICER (£/QALY)	Pairwise ICERs vs. comparator (£/QALY)
	Costs (£)	QALYs		
PEM+plat	[redacted]	[redacted]		£13,424
PEM+plat+PEM maint	[redacted]	[redacted]	£38,943	Dominant
Atezo+Bev+CP	[redacted]	[redacted]	Dominant	-

Table 4 Company base case results, EGFR/ALK positive population (PAS for Atezo and Bev, list price for all other treatments) – deterministic (Clarification Response Table 37)

Treatment	Total		Incremental analysis ICER (£/QALY)	Pairwise ICERs vs. comparator (£/QALY)
	Costs (£)	QALYs		
PEM+plat	[redacted]	[redacted]		£14,552
PEM+plat+PEM maint	[redacted]	[redacted]	£31,523	£7,014
Atezo+Bev+CP	[redacted]	[redacted]	£7,014	-

Commentary on the robustness of submitted evidence

Strengths

- The ERG considers that the company's systematic literature review of clinical effectiveness evidence is of a good standard, with comprehensive literature searches, inclusion screening, data extraction and critical appraisal.
- Overall, the ERG believes the IMpower150 RCT has been well conducted but, as an open label trial, the outcomes are susceptible to performance bias and detection bias.
- The model structure is appropriate for NSCLC and correctly implemented.
- The economic analysis complies with methodological criteria in the NICE reference case (although the decision problem does not match that in the scope, see below).

- The company's base case assumption of a two-year stopping rule for atezolizumab and bevacizumab in the Atez+Bev+CP intervention is consistent with previous NICE guidance for atezolizumab (TA520 and TA525) and other immunotherapies (e.g. TA531). The assumption that pemetrexed maintenance therapy continues until progression is also consistent with committee conclusions in TA402.
- The company assumption of a three-year cap on survival benefits (after the maximum 2-year treatment) for Atez+Bev+CP is reasonably cautious and consistent with previous guidance (e.g. TA520).
- The company's approach to extrapolating OS, PFS and TTD curves is good. They consider a range of baseline extrapolations from the Atezo+Bev+CP arm of the IMpower150 trial, including fully parametric and piecewise (KM with parametric tail). For the piecewise approach, the KM is used up to the time when 20% of trial patients remain at risk, which results in long-term extrapolations that are consistent with the fully parametric curves.
- The company's choice of survival curve extrapolations for PFS and TTD is reasonable and appropriate.
- The company's approach to estimating health state utility values is reasonable and consistent with previous NICE technology appraisals. The use of IMPower150 utility data is preferable to other estimates of utility in this population and we agree that the 'proximity to death' approach has more face validity than the pre/post progression analysis.

Weaknesses and areas of uncertainty

- Median OS has been reached in the IMpower150 RCT, however, final OS data are not yet available.
- The comparators used for the EGFR/ALK positive (pemetrexed + cisplatin with or without pemetrexed maintenance) do no match the NICE scope which includes pembrolizumab and docetaxel. The company has also omitted chemotherapy with carboplatin comparators for the untreated PD-L1 low/negative subgroup which may reflect current practice for patients who cannot tolerate cisplatin.
- There is significant uncertainty over the extrapolation of OS and ICERs are quite sensitive to this uncertainty. We agree that the company's choice of an exponential OS curve for the atezolizumab combination in their base case has a good fit to the trial data and gives clinically plausible extrapolations of survival at five and ten years. We consider that the Weibull distribution is also plausible, with more conservative

survival predictions. The log-logistic gives over optimistic long-term predictions (around 10% survival at 10 years).

- We do not consider the company's assumption of a persistent survival advantage for pemetrexed maintenance throughout the time horizon to be realistic. This is not consistent with committee conclusions in TA402 and is likely to have overestimated the long-term survival gain for Atez+Bev+CP and for the pemetrexed maintenance comparator in comparison with pemetrexed with platinum induction alone. This implies that the ICER for Atez+Bev+CP relative to PEM+CIS without maintenance is likely to be underestimated.
- Cost-effectiveness estimates for the PD-L1 low/negative based on the subgroup analysis of IMpowe150 data are quite similar to the ITT results, and reasonably robust. However, estimates for the EGFR/ALK positive subgroup are much more uncertain as they are based on a small subgroup from the trial (n=41).
- The ERG considers that the utility impact of differences in the incidence of treatment related adverse events between treatments have not been fully captured in the company's base case analysis. It is unclear whether patients treated with Atezo+Bev+ CP have the same health state utility whilst on treatment as those treated with pemetrexed + platinum (with or without pemetrexed maintenance).
- There are some minor discrepancies to some of the cost estimates, which have not been updated correctly.

Summary of additional work undertaken by the ERG

We corrected minor discrepancies in the in the model and re-ran the company's analyses. Changes to the results were minimal.

In addition, we ran the model for an ERG base case, including preferred assumptions and parameters (see table below). This included changes to: the company's baseline Atez+Bev+CP OS curve (Weibull instead of exponential); relative treatment effects excluding the PARAMOUNT trial (which restricts the results to a comparison with pemetrexed maintenance); and inclusion of disutility for adverse events. We also present selected scenario analyses around the ERG base case to reflect key uncertainties.

Table 5 ERG base case and ERG scenarios

	Subgroup	Company base case	ERG base case	ERG scenarios
Baseline OS	All	Exponential	Weibull	• Exponential • Log-logistic
Baseline PFS	ITT & PD-L1 low/-ve	KM + log-logistic	KM + log-logistic	• KM + exponential • KM + Weibull
	EGFR/ALK +ve	Log-normal	Log-normal	• Exponential • Weibull
NMA (OS & PFS)	ITT	FP (FE) ITT	ITT FP excluding PARAMOUNT (FE)	• ITT FP (RE) • ITT PH • Subgroup specific
	PD-L1 low/-ve	FP (FE) PD-L1 low/-ve		
	EGFR/ALK +ve	FP (FE) EGFR/ALK +ve		
TTD	All	KM + exponential for atezo and bev	KM + exponential for atezo and bev	• Bev until progression
		PEM follows PFS	PEM follows PFS	
Stopping rule and effect cap	All	2 year treatment + 3 year OS effects	2 year treatment + 3 year OS effects	• 2 years for OS • 5 years for OS • 3 years for PFS • no stopping rule or effect cap
Utilities	All	IMPower150 EQ-5D time-from-death with no treatment effect	IMPower150 EQ-5D time-from-death + disutility per grade 3+ treatment related AE	• IMpower150 EQ-5D health state model • No AE disutility
Subsequent treatments	All	UK scenario (CS Tab 34)	UK scenario (CS Tab 34)	Exclude nivolumab

FE Fixed effect; FP Fractional polynomial; KM Kaplan Meier; NMA network meta-analysis; RE Random effects

Results from the ERG ITT base case are shown below (with PAS for atezo and bev, list price for other treatments). Scenario analyses are presented in section 4.4.2 below. These indicate that the model is most sensitive t: extrapolations of overall survival and treatment duration, the use of a stopping rule for atezolizumab and bevacizumab as part of Atezo+Bev+CP and the costs of subsequent treatments. Results with all available PAS discounts are shown in an addendum to this report.

Table 6 ERG base case for ITT population (PAS for atezolizumab and bevacizumab and list price for comparators and subsequent treatments)

Technologies	Total costs (£)	Total QALYs	ICER (£) fully incremental analysis	ICER (£) pairwise; Atezo+Bev+CP vs comparator
PEM+platinum w PEM maint	[redacted]	[redacted]		Dominant
Atezo+Bev+CP	[redacted]	[redacted]	Dominant	

1 Introduction to ERG Report

This report is a critique of the company's submission (CS) to NICE from Roche Product Limited on the clinical effectiveness and cost effectiveness of atezolizumab in combination for non-squamous non-small-cell lung cancer. It identifies the strengths and weakness 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 manufacturer by the ERG via NICE on 28th September 2018. A response from the company via NICE was received by the ERG on 15th October and this can be seen in the NICE committee papers for this appraisal.

2 BACKGROUND

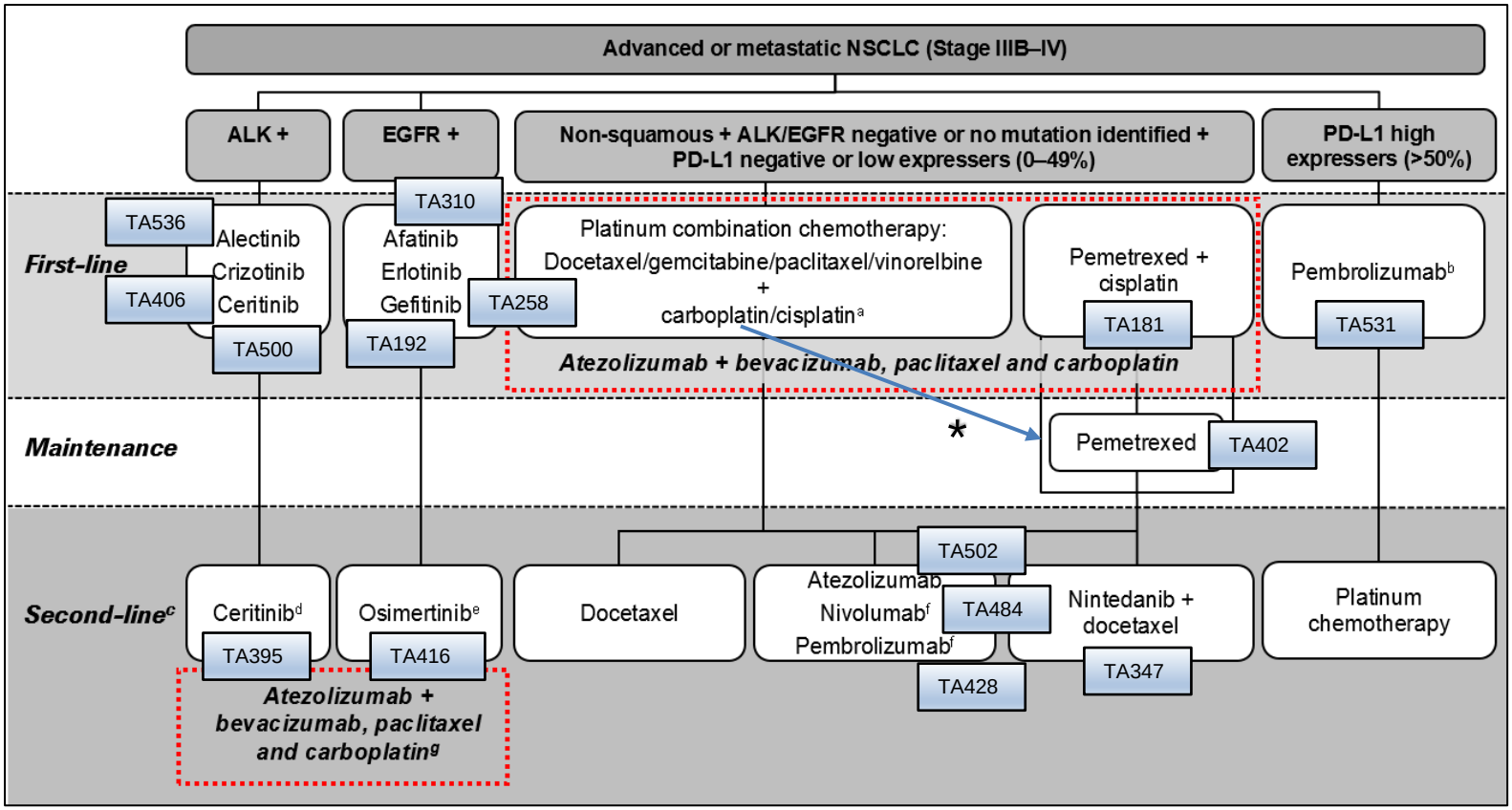
2.1 Critique of company's description of underlying health problem and overview of current service provision

The CS provides a clear and accurate overview of the disease and current service provision, citing relevant NICE guidance and clinical guidelines (including NICE clinical guideline 121[1]). Figure 1 shows the care pathway, reproduced from the CS. Expert clinical advice to the ERG is that the pathway is reflective of current clinical practice, though the advice given was that some patients may be unable to tolerate cisplatin (which is used in combination with pemetrexed), and therefore would therefore commence treatment with platinum combination chemotherapy containing carboplatin, and then receive pemetrexed maintenance therapy.

The CS briefly mentions the key factors that influence choice of treatment for NSCLC, including:

- The presence of epidermal growth factor receptor (EGFR) and anaplastic lymphoma kinase (ALK) mutations.
- Programmed Death Ligand 1 (PD-L1) expression (commonly categorised as high expression, or low or negative expression). High expression is defined by a tumour proportion score (TPS) > 50%, TC/IC3. Low or negative expression is defined as a tumour proportion score 0-49%, TC/IC 0,1,2.
- Patient-specific factors such as age, comorbidities, and personal preferences.

Expert clinical advice to the ERG notes that patients with EGFR/ALK mutations tend to have better survival, and respond better to pemetrexed chemotherapy. Patients with high PD-L1 expression are also more likely to respond to pemetrexed chemotherapy. These patient subgroups are included in this company submission and are referred to extensively in this ERG report.



[a]Single-agent chemotherapy with a third-generation drug offered if a platinum combination cannot be tolerated; [b]PD-L1 expression ≥50% TPS; [c]Patients who progress following non-targeted therapy may receive an ALK or EGFR tyrosine kinase inhibitor as a second-line treatment if an actionable mutation is identified or suspected; [d]Certinib after crizotinib failure; not suitable after first-line alectinib; [e]EGFR T790M mutation-positive only; [f]PD-L1 positive patients only; [g]Atezolizumab+bevacizumab, paclitaxel and carboplatin would be available as a second-line treatment option for patients who progress on targeted therapy (after exhausting all available options) and are ineligible for osimertinib, i.e. non T790Mpatients

Adapted from CS Figure 1. TA = Technology appraisal. * The asterisk and arrow reflect the clinical advice to the ERG that patients who are unable to tolerate cisplatin would commence treatment with platinum combination chemotherapy containing carboplatin, and then receive pemetrexed maintenance therapy.

Figure 1 Treatment pathway for advanced or metastatic NSCLC (based on NICE clinical guideline 121), showing intended position of atezolizumab-based therapy

2.2 Critique of company's definition of decision problem

There are some key differences between the decision problem and the NICE scope, as outlined in CS Table 1. The key differences are:

- **Intervention** - The combination of atezolizumab with carboplatin and paclitaxel (without bevacizumab) is not pursued in the anticipated marketing authorisation and therefore not included in the decision problem. Given that NICE only recommends treatments within their marketing authorisation this exclusion is acceptable.
- **Population** – For people without EGFR or ALK tumour mutations there is a restriction to patients with low or negative PD-L1 expression (tumour proportion score 0–49%, TC/IC 0,1,2). Patients with high PD-L1 expression, who currently would be eligible to receive pembrolizumab (NICE TA531), are not included. The justification for this is explained below under 'comparator'. Patients who have EGFR or ALK tumour mutations can have any level of PD-L1 expression, and they are included in the decision problem.
- **Comparator 1 (PD-L1 negative or low patients, ALK/EGFR negative patients)** - the decision problem omits the comparison with chemotherapy (docetaxel, gemcitabine, paclitaxel or vinorelbine) in combination with a platinum drug (carboplatin or cisplatin), with or without pemetrexed maintenance treatment. Instead, the CS focuses on the comparison with pemetrexed in combination with cisplatin / carboplatin, with or without pemetrexed maintenance treatment. The NICE scope permits this comparison but with the caveat that it applies to adenocarcinoma or large cell carcinoma only (based on NICE TA181) – the CS does not mention this caveat. The ERG notes that the histology of NSCLC is predominantly adenocarcinoma and squamous cell carcinoma, with the remainder of histology sub-types comprising large cell or undifferentiated. The justification for use of pemetrexed in this subgroup, is that pemetrexed in combination with cisplatin/carboplatin, with or without pemetrexed maintenance, is the most appropriate UK chemotherapy, based on clinical expert opinion sought by the company and UK market share data. Expert clinical advice to the ERG concurs with this assertion, but notes that in England pemetrexed should only be given in combination with cisplatin (based on NICE guidance, TA181[2]).Patients who cannot tolerate cisplatin would therefore be treated with a carboplatin-based regimen (i.e. docetaxel, gemcitabine, paclitaxel or vinorelbine in combination with carboplatin), followed by maintenance treatment with pemetrexed (illustrated by the arrow in Figure 1). In practice, however, the findings of

the National Lung Cancer Report for 2017 (for the audit period 2016)[3] show that pemetrexed is given in combination with carboplatin as well as in combination with cisplatin.

- **Comparator 2 (PD-L1 high patients)** - No cost-effectiveness comparison is made with pembrolizumab in PD-L1 high expression patients. An indirect comparison of clinical effectiveness is presented in the CS but, based on the results [redacted] and UK clinical expert advice, a cost effectiveness comparison with pembrolizumab in PD-L1 high patients is not included in the CS. The CS states that UK clinical opinion suggests that [redacted]. Expert advice to the ERG concurs with this suggestion. The company is therefore not seeking NHS reimbursement for treatment with atezolizumab in this patient sub-group.
- **Comparator 3 (EGFR/ALK positive patients)** – the CS omits the comparison with docetaxel or pembrolizumab in patients with EGFR-or ALK-positive advanced, non-squamous NSCLC previously treated with targeted therapy. Instead, the only comparison made is to pemetrexed in combination with cisplatin/carboplatin, with or without pemetrexed maintenance treatment. The NICE scope does not specify pemetrexed as a comparator for this patient subgroup. Expert clinical advice to the company and to the ERG suggests that pemetrexed can be considered an appropriate comparator for these patients.
- **Outcomes** – all outcomes in the scope are included in the decision problem. Time to treatment discontinuation is included in the decision problem, though not included in the scope. This is an input parameter for the economic model and is appropriate to the analysis.

ERG conclusion: The company's decision problem does not fully adhere to the NICE scope, in terms of relevant treatment comparisons. One key omission is comparison to first line chemotherapy regimens including docetaxel, gemcitabine, paclitaxel or vinorelbine in combination with a platinum drug (carboplatin or cisplatin) (with or without pemetrexed maintenance treatment). Whilst clinical advice to the company suggests pemetrexed in combination with cisplatin is the standard of care, clinical advice to the ERG also suggests that these chemotherapy regimens may be used in combination with carboplatin for patients who cannot tolerate cisplatin. Omission of a comparison to pembrolizumab in high expressing or positive PD-L1 patients is supported by clinical opinion.

3 CLINICAL EFFECTIVENESS

3.1 Critique of company's approach to systematic review

3.1.1 Description of company's search strategy

The CS reports literature searches for clinical effectiveness literature (Appendix D), cost-effectiveness literature (Appendix G), health-related quality of life (HRQoL) (Appendix H), and cost and healthcare resource identification, measurement and valuation (Appendix I). All searches are deemed to be fit for purpose. They are of good quality, contain a balance of descriptor terms, free text terms and suitable study design filters have been applied to identify RCT, cost, resource use & HRQoL. They are well documented and reproducible. A suitable range of databases and grey literature, including ongoing trial databases and pertinent conference proceedings, have been searched. Search results are represented in Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow charts. The clinical effectiveness searches were reasonably up to date (February 2018), however the ERG elected to run targeted searches on atezolizumab on Medline, Embase, general internet searches and www.clinicaltrials.gov to check for any recently published material during 2018. No additional sources of information were identified that were relevant to this appraisal.

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

The company conducted a systematic review to identify and select clinical effectiveness evidence. The inclusion and exclusion criteria for the systematic review are presented in CS Appendix D Table 10. Although the scope of the systematic review is not described, it is apparent from the inclusion and exclusion criteria that the remit was wider than the NICE scope for this appraisal to encompass outcomes of relevance to a network meta-analysis (NMA) and non-NMA outcomes (see section 3.1.7 of this report for a description and critique of the NMA). The population for the systematic review "Adult patients aged over 18 years with any Stage IV non-squamous NSCLC who have not received prior chemotherapy for Stage IV NSCLC" matches that of the NICE scope. Although the systematic review criteria specify stage IV disease and the NICE scope states advanced disease (which could include Stage IIIb and IIIc disease), the NICE scope also indicates that atezolizumab must be used within its marketing authorisation which is the treatment of adult patients with metastatic NSCLC. Metastatic NSCLC would normally be interpreted as stage IV disease.

Ten eligible interventions were specified (four of which included atezolizumab) and the eligible comparators were any pharmacological treatment or placebo. The specified interventions and the broad nature of the comparator are likely to have ensured that the evidence matching the decision problem would be identified (though note that, as discussed earlier, the decision problem and the inclusion criteria exclude some of the comparators specified in the NICE scope). Outcomes were divided into those to consider in the NMA and additional (non-NMA) outcomes. RCTs (phase II to IV) were eligible for inclusion and systematic reviews published in the last five years were used a source of references. Conference abstracts published in the last five years were only included if they provided additional data associated with an included full-text publication. Studies published in Chinese without a detailed abstract in English were excluded. No other restrictions are reported to limit inclusion in the systematic review.

The flow diagram (CS appendix D Figure 1), showing the flow of studies through the inclusion and exclusion screening stages, is provided but no details about how screening was achieved are presented (i.e. how many reviewers involved). In response to clarification question A16 the company stated that two reviewers independently undertook the record selection with a third reviewer involved to adjudicate any disagreements. Furthermore, the company's response to clarification question A12 suggests that the flow-diagram depicts screening against broader criteria for a "global network" (to inform HTA submissions in other countries). The specific UK network criteria for patient eligibility appear to have been applied once a broader set of trials had been identified. The two differences between the UK network and the global network were firstly, for the UK network, the proportion of patients in each study with Stage IV non-squamous NSCLC had to be at least 90% if outcomes were not reported separately for this group. Secondly, only interventions relevant to the UK as shown in CS Appendix D Table 10 were included. CS Appendix D Table 12 lists the 895 full texts documents excluded along with the reasons for exclusion.

> **ERG conclusion:** The ERG believes that the broad scope of the company's systematic review not only encompasses the proposed population and licensed indication for atezolizumab but also the need to identify evidence for the NMA.

3.1.3 Identified studies

The systematic review identified one RCT, the IMpower150 RCT, of atezolizumab in combination with carboplatin plus paclitaxel with or without bevacizumab. A further 13 RCTs, six of which included interventions that were relevant to the UK, were identified for

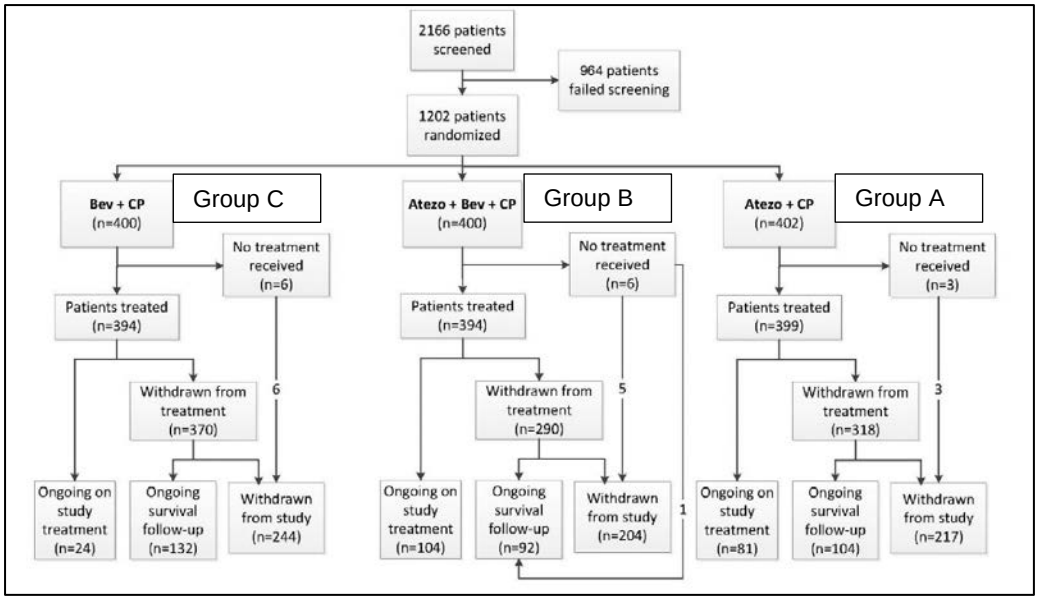
possible inclusion in the NMA. No non-randomised evidence was included in the submission.

Summary details of the IMpower150 RCT, which was sponsored by the company, are provided in the CS. IMpower150 is a Phase III, open-label RCT which enrolled adult chemotherapy-naive patients with stage IV non-squamous NSCLC. Eligible patients were randomised in a 1:1:1 ratio (stratified by gender, PD-L1 expression and the presence of liver metastases) to one of three treatment arms:

- Arm A (n=402): Atezolizumab + carboplatin + paclitaxel (Atezo+CP) induction (four or six 21-day cycles) followed by atezolizumab maintenance (21-day cycles)
- Arm B (n=400): Atezolizumab + bevacizumab + carboplatin + paclitaxel (Atezo+Bev+CP) induction (four or six 21-day cycles) followed by atezolizumab + bevacizumab maintenance (21-day cycles)
- Arm C (n=400): Bevacizumab + carboplatin + paclitaxel (Bev+CP) induction (four or six 21-day cycles) followed by bevacizumab maintenance (21-day cycles).

The marketing authorisation applied for covers Atezo+Bev+CP (Arm B) only and therefore data for Atezo+CP (Arm A) were not included in the CS and this arm is not discussed further in this report.

A flow-chart showing numbers of patients randomised, treated, withdrawn and ongoing either on study treatment or being followed up for survival at the 22nd January 2018 data cut off is reproduced below in Figure 2 (note that groups are presented from left to right in reverse order).



Reproduction of CS Appendix D Figure 19 with group labels added for clarity.

Figure 2 IMpower150 RCT flow chart

Among the ITT population of the Atezo+Bev+CP and the Bev+CP arms of the IMpower150 trial patient characteristics and baseline demographics were well balanced (Table 7). The only exception that the CS highlights (in Appendix D.1.1) is that there were more patients with an ECOG performance status of 1 in the treatment arm (59.9%) than in the control arm (54.9%). The ERG agrees that, other than this, the arms are well balanced. Furthermore, clinical advice to the ERG was that although ECOG performance status is a prognostic factor, the difference between arms is regarded as small and not clinically important.

Results are also presented in the CS for the ITT-WT (wild-type) and EGFR/ALK+ populations (see section 3.1.6 of this report for an explanation of the different analysis populations used in the CS) so the company were asked to provide the baseline characteristics for these populations (clarification question A3). The ITT-WT population baseline characteristics are very similar to those of the ITT population, which is not surprising because the ITT-WT population is 87% of the ITT population. Baseline characteristics of the arms of the ITT-WT population are well balanced. The EGFR/ALK+ population, which is small (n=104 in total across arms B and C of the trial), differs from the ITT population not only in terms of EGFR mutation status and EML-4-ALK rearrangement status, as expected, but additionally the proportion of male participants is lower (approximately 50% compared with approximately 60% in the ITT population), a greater proportion of Asian participants (approximately 35% compared with 13%, respectively) and lower proportion of white participants (approximately 62% compared with 82%, respectively). Activating EGFR mutation and ALK translocations are known oncogenic driver mutations in

NSCLC (i.e. they are responsible for the initiation and maintenance) therefore, the observed greater proportion of participants in the EGFR/ALK+ population who had never smoked (48% compared to 2%, respectively) is not unexpected. There are some imbalances between the trial arms of the EGFR/ALK+ subgroup but these may well be due to the smaller participants numbers and selection bias associated with the non-random nature of this subgroup. Of particular note is the lower proportion of participants in the Atezo+Bev+CP arm with liver metastases at baseline in comparison with the Bev+CP group (12.2% versus 15.9%) (Liver metastases are reported in the CS as being associated with limited therapeutic benefit with checkpoint-inhibitor monotherapy i.e. therapies such as atezolizumab that block immune system checkpoint proteins thus allowing the immune system to kill cancer cells better). There were also imbalances in PD-L1 status between arms. Given these imbalances caution is required in the interpretation of the results of the EGFR/ALK positive subgroup.

The IMpower150 RCT is still ongoing. No other ongoing studies of atezolizumab in this indication are presented by the company (CS section B.2.11) and none were identified by the ERG.

Table 7 Summary of key patient demographics and baseline characteristics in IMpower150 (ITT population)

	Atezo+Bev+CP n=400	**Bev+CP n=400**
Mean age, years (SD) Median age, (range)	63.0 (9.5) 63.0 (31–89)	63.1 (9.3) 63.0 (31–90)
Male, n (%)	240 (60.0)	239 (59.8)
Race, White, n (%)	322 (80.5)	335 (83.8)
ECOG PS, n (%) 0 1	n=397 159 (40.1) 238 (59.9)	n=397 179 (45.1) 218 (54.9)
Smoking status, n (%) Never Current Previous	 82 (20.5) 90 (22.5) 228 (57.0)	 77 (19.3) 92 (23.0) 231 (57.8)
EGFR mutation status, n (%) Positive Negative Unknown	 34 (8.5) 353 (86.3) 10 (2.5)	 45 (11.3) 345 (86.3) 10 (2.5)
EML4-ALK rearrangement status, n (%) Positive Negative Unknown	 11 (2.8) 386 (96.5) 3 (0.8)	 20 (5.0) 376 (94.0) 4 (1.0)

Liver metastases at enrolment from IxRS, n (%)		
Yes	67 (16.8)	69 (17.3)
No	333 (83.3)	332 (82.8)
PD-L1 IHC stratification factor from IxRS, n (%)		
TC0/1/2 and IC0/1	299 (74.8)	301 (75.3)
TC0/1/2 and IC2/3	53 (13.3)	50 (12.5)
TC3 and any IC	48 (12.0)	49 (12.3)

ECOG PS, Eastern Cooperative Oncology Group performance status; EGFR, epidermal growth factor receptor; EML4-ALK, *EML4*-anaplastic lymphoma kinase; IC, tumour-infiltrating immune cell; IHC, immunohistochemistry; IxRS, Interactive Voice/Web Response System; PD-L1, programmed death-ligand-1SD, standard deviation; TC, tumour cell
Source: adapted from CS Table 5

3.1.4 Description and critique of the approach to validity assessment

The company presented an appraisal of aspects of study risk of bias both for the key IMpower150 RCT and the RCTs included in the NMA. For the IMpower150 RCT a summary of the assessment is presented in CS section 2.5, Table 8 using NICE's suggested criteria. Further details are presented in Appendix D Table 32; however, the questions in this table differ from those in CS Table 8. In particular, question 3 in CS Table 8, regarding the similarity of the groups in terms of prognostic factors, is not present in Table 32. Furthermore some questions are answered differently in CS Table 8 and CS Appendix D Table 32 (e.g. question 2 regarding whether the concealment of allocation was adequate where answers are either 'not applicable' or 'yes' depending on which table is consulted). In response to clarification question A6 the company provided a revised version of CS Table 8 and the detailed risk of bias assessment (CS Appendix D Table 32).

Neither the CS nor the published paper for the IMpower150 RCT provided sufficient details for the ERG to complete an assessment of study methods using the NICE suggested criteria. Fortunately, the company had supplied the clinical study report (CSR) for IMpower150 and the ERG used this to complete the assessment (Table 8). The opinion of the ERG and the company differed for one question and partially differed for one question. The reasons for the differences in opinion are provided in the comment rows of Table 8.

> **ERG conclusion:** Overall, the ERG believes the IMpower150 RCT has been well conducted but, as an open label trial, the outcomes are susceptible to performance bias and detection bias.

Table 8 Company and ERG assessment of trial quality

	IMpower150	
1. Was randomisation carried out appropriately?	CS:	Yes
	ERG:	Yes
Comment:		
2. Was concealment of treatment allocation adequate?	CS:	Yes
	ERG:	Yes
Comment: Study site was not a stratification factor so the probability of the next allocation will depend on previous allocations at all the other sites. Therefore, it is unlikely that the next allocation could be guessed in advance. Furthermore each study site obtained a randomization number and treatment assignment for each eligible patient from the interactive voice/Web response system (IxRS/IWRS).		
3. Were groups similar at outset in terms of prognostic factors?	CS:	Yes
	ERG:	Yes
Comment: In the ITT population there were more patients with an ECOG performance status of 1 in the treatment arm (59.9%) than in the control arm (54.9%) but clinical advice to the ERG was that this difference is not clinically important. Arms are well balanced other than this.		
4. Were care providers, participants and outcome assessors blind to treatment allocation?	CS:	N/A (open label study)
	ERG:	No
Comment: Open label study to care providers and participants aware of treatment allocation. No evidence that outcome assessors were blind to treatment allocation.		
5. Were there any unexpected imbalances in drop-outs between groups?	CS:	No
	ERG:	No
Comment:		
6. Is there any evidence that authors measured more outcomes than reported?	CS:	No
	ERG:	No
Comment: All the key clinical effectiveness outcomes are reported. Some other patient reported outcomes (PROs) are not reported in the CS e.g. EQ-5D-3L data required for economic modelling but utility scores were provided in response to clarification question A5. The IMpower150 study protocol states that [redacted]. The CSR states that [redacted].		
7. Did the analysis include an ITT analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	CS:	Yes
	ERG:	Yes for most efficacy outcomes Unclear for PRO outcomes
Comment: An ITT analysis was conducted for efficacy outcomes. For PFS and OS appropriate censoring methods are described. It is not clear how missing data were accounted for in the analysis of response or of PROs.		

PGIS – Patient Global Impression of Severity; PRO – Patient reported outcome; SILC – Symptoms In Lung Cancer

3.1.5 Description and critique of company's outcome selection

The outcomes selected by the company for their decision problem and the results presented in the CS match the outcomes listed in the NICE scope. In addition, the company presents evidence on time to treatment discontinuation (TTD) which is required to inform treatment duration for atezolizumab in the economic model.

Overall survival – defined as time from randomisation to death from any cause.

Progression-free survival – Investigator-assessed PFS according to RECIST v1.1. Defined as time from randomisation to first documented progressive disease or death from any cause, whichever occurred first. Although PFS was also assessed by an independent review facility (IRF) these results were not presented in the CS.

Time to treatment discontinuation – this was not defined in the CS but as treatment could continue after progression this could be longer than PFS.

Response rate – Objective Response Rate (ORR) was defined as the proportion of patients with either a complete response (CR) or partial response (PR) as judged by the investigator using RECIST v1.1 with confirmation not required. The ERG notes that the RECIST v.1.1 criteria state that " …. elimination of the requirement for response confirmation may increase the importance of central review to protect against bias, in particular in studies which are not blinded." As far as the ERG can determine no central review of response outcomes took place and hence this outcome may be at risk of bias.

In addition to ORR the CS also reports duration of response (DOR) defined as time from the first documented objective response to documented progressive disease or death from any cause whichever occurred first. Similarly, to ORR, DOR was investigator assessed using RECIST v1.1 with no confirmation required.

Adverse effects of treatment - Adverse events were assessed according to the National Cancer Institute Common Terminology Criteria for Adverse Events, version 4.0.

Health-related quality of life – time to deterioration in patient-reported lung cancer symptoms using the European Organisation for the Research and Treatment of Cancer (EORTC) Quality-of-life Questionnaire Core (QLQ-C30) and supplemental lung cancer module (QLQ-LC13). Change from baseline in patient-reported HRQoL as assessed by the EORTC QLQ-C30 and the QLQ-LC13. The CS does not describe these two questionnaires but the ERG can confirm that both the QLQ-C30 core questionnaire and the QLQ-LC13 module are validated instruments. EORTC questionnaires were administered at every cycle during the treatment period until either progressive disease (for the Bev+CP arm) or until loss of clinical benefit (for the Atezo+Bev+CP arm) and after disease progression at 3- and 6-month follow-up visits. For the EORTC QLQ-C30 Global health status, physical functioning and role functioning scales scores range from 0 to 100 with a higher score indicating better quality of life. Clinically meaningful worsening is indicated by a 10-point or greater decrease in mean score. For the EORTC QLQ-LC13 module scores (range 0-100) are produced for dyspnea, coughing, chest pain, arm/shoulder pain, pain in other parts and

pain with a lower score indicating lower symptom severity and a clinically meaningful worsening indicated by an increase in mean score of 10 points or more.

ERG conclusion: The outcomes presented by the company are appropriate.

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

In this section we describe and critique the statistical approach used in the IMpower150 trial, focusing on the outcome measures that directly inform the economic model (i.e. PFS and OS and HRQoL).

3.1.6.1 Analysis populations

Table 9 provides a summary of the patient analysis populations in the trial at the two data cut-off dates, as discussed in the following sub-sections.

Table 9 Summary of the analysis populations and data cuts in the trial

	PFS			**OS**	
Data cut-off date	September 2017	January 2018	Data cut-off date	September 2017	January 2018
Analysis status	Final	Updated	Analysis status	Interim	2nd interim
Analysis population (method of tumour assessment)			Analysis population		
ITT (INV)	✓	✓	ITT	✓	✓
ITT-WT (INV)	X	✓	ITT-WT	✓	✓
ITT-WT (IRF)	✓	X			
EGFR/ALK+ (INV)[a]	✓	✓	EGFR/ALK+[a]	✓	✓
PD-L1 low/- (INV)[b]	✓	✓	PD-L1 low/-[b]	✓	✓

INV = investigator assessed
IRF = Independent review facility assessed
[a] based on the ITT population
[b] based on ITT population and also the ITT-WT population
Grey shading indicates which populations and data cuts from the IMpower150 trial inform the economic model

The CS reports the following patient analysis populations:

- **ITT (n=1202)** - all randomised patients, regardless of receipt of the assigned treatment. This is the study population used in the analysis of clinical effectiveness and cost effectiveness in the CS. The CS reports that the anticipated marketing authorisation is based on the entire ITT population (i.e. including patients with activating EGFR mutation or ALK translocation).
- **Safety population (n=1187)** - randomised patients who received any amount of any component of study treatment. Patients were grouped according to whether any

amount of atezolizumab was received, including when atezolizumab was received in error.

- **ITT-WT (n=1040)** - intention-to-treat wild type population. This is the same as the ITT population (see below) with the exclusion of patients with activating EGFR mutation or ALK translocation.
  ([redacted]).
    - **Teff-high WT population (n=445)** –patients who had a specific T-effector (Teff) gene signature, excluding patients with an activating EGFR mutation or ALK translocation (thus, a subgroup of the ITT-WT population). The CS reports that since the Atezo+Bev+CP combination demonstrated a clinically meaningful improvement in outcomes regardless of Teff gene signature status, this this biomarker was not deemed to be clinically relevant and therefore this data did not impact the anticipated marketing authorisation. This population is not reported in detail in the CS, and is not mentioned in the NICE scope. Therefore, it is not discussed in this ERG report. Expert clinical advice to the ERG is that, at Southampton General Hospital patients are not routinely tested for this biomarker.

3.1.6.2 Sample size calculation and hypotheses

The determination of patient populations for the primary and secondary analyses has a complex background, as summarised below.

Originally, the primary endpoint analysis was to be performed on the:

- **ITT population**, and the
- **PD-L1 selected population**. (Not explicitly defined, but the ERG assumes this is patients with patients with high PD-L1 expression).

However, there was a protocol amendment during the study (March 2017) which changed the primary-analysis populations to the:

- **ITT-WT population** (i.e. excluding patients with activating EGFR mutation or ALK translocation) and the
- **Teff-high WT** populations

The analyses of PFS and OS in patients defined by their PD-L1 expression status became a secondary analysis.

The sample size calculation was therefore performed for the co-primary endpoints of OS and PFS in the:

- **ITT-WT** patient population (OS and PFS) and the
- **Teff-high WT** population (PFS only).

As mentioned above, the analysis of clinical effectiveness and cost effectiveness in the CS are based on the **ITT population** (effectively a secondary analysis following the protocol amendment), and not the ITT-WT and the Teff-high WT populations (i.e. the primary analysis following protocol amendment). The statistical power calculation is thus based on a sample size for a WT population that is smaller than the sample size for the ITT analyses presented in the CS. The ERG notes that the ITT-WT population comprises 87% (n=1040/1202) of the ITT population, so the difference in the size of these two populations is relatively small.

[redacted]

The sample size was based on the number of events required to demonstrate efficacy with regard to both PFS and OS for the comparison of the Atezo+Bev+CP vs. Bev+CP (Arm B vs Arm C). An 'alpha-spending algorithm' was employed so that if there was a significant difference between Atezo+Bev+CP and Bev+CP then the Atezo+CP arm would be compared with the Bev+CP arm (Arm A vs Arm C). The study was not designed to test a comparison between Atezo+CP and Atezo+Bev+CP (Arm A vs Arm B). Thus, it is not possible to statistically compare a atezolizumab regimen with and without bevacizumab.

The sequence of testing was as follows:

- With a two-sided significance level of 0.05, a two-sided alpha value of 0.012 was allocated to PFS (split equally into 0.006 for each primary-analysis population (the **ITT-WT population** and the **Teff-high WT population**)), and a two-sided alpha value of 0.038 was allocated to OS in the **ITT-WT population.**
- If there was a statistically significant difference in PFS between the Atezo+Bev+CP group and the Bev+CP group, the alpha value would then be recycled for the comparison of OS between the Atezo+Bev+CP group and the Bev+CP group.
- If the result of the comparison of OS between the Atezo+Bev+CP group and the Bev+CP group was significant, the remaining alpha value would be used to compare

both PFS and OS between the Atezo+CP group and the Bev+CP group (i.e. Arm A versus Arm C).

- If there was a statistically significant difference in OS between the Atezo+CP group and the Bev+CP group (A versus C), testing would be extended to the **ITT population**, including patients with EGFR or ALK mutations.

The rationale for this sequence of testing was to maximise statistical power to detect a significant benefit for the addition of atezolizumab to bevacizumab, cisplatin and paclitaxel. If the addition of atezolizumab to this regimen did not provide a significant benefit it was considered unlikely that substituting atezolizumab for bevacizumab in the Bev+CP regimen (i.e. comparing Arm A vs Arm C) would provide significant benefit.

The CS reports that the comparison of Atezo+CP to Bev+CP did not show a statistically significant survival benefit (HR=0.88, 95% CI: 0.72, 1.08; p=0.2041), thus marketing authorisation was only sought for the Atezo+Bev+CP regimen. The CS therefore does not present results for the Atezo+CP arm. Likewise, results for this arm are not reported in this ERG report.

3.1.6.3 Completeness of follow-up

Results are available for two data cuts: September 2017 (final PFS, interim OS), and 22nd January 2018 (updated PFS; second interim OS) (Table 9). The most recent data from the second interim OS results are based on 422 deaths across the two treatment arms relevant to this appraisal, with a median follow-up of approximately 20 months. Median OS has been reached in both treatment arms, however, final OS data analysis will be conducted when there are 507 deaths across the two relevant trial arms (in the ITT-WT population). Analysis is expected in [redacted]. Thus, the PFS results are based on mature data, but the OS data are not fully yet mature, and thus caution is required in the interpretation of the OS results.

3.1.6.4 Tumour progression assessment

Investigator assessed PFS results are presented in the CS. The only reporting of independent review facility (IRF) assessed PFS is for the ITT-WT population, in the trial journal publication (a secondary outcome). Independent assessment of tumour progression can sometimes differ from investigator assessment, and it is informative to conduct and report both. The company were asked to provide IRF PFS results for the ITT, ITT-WT populations and subgroup analyses (clarification question A4). The company provided the IRF-assessed PFS data from the 15 September 2017 data cut and stated that as the IRF was disbanded after the primary endpoint for PFS was met IRF-assessed data are not available for the most recent 22nd January 2018 data cut.

3.1.6.5 Subgroup analyses

The NICE scope included provision for subgroup analysis by level of PD-L1 expression if the evidence allowed. As described earlier in section 2.2, the company's decision problem, for people without EGFR or ALK tumour mutations there is a restriction to patients with low or negative PD-L1 expression (tumour proportion score 0–49%, TC/IC 0,1,2). Patients with high PD-L1 expression are not included (tumour proportion score > 50%, TC/IC3). Within the sub-population of people with EGFR or ALK tumour mutations people with any level of PD-L1 expression are included in the decision problem.

Pre-planned subgroup analyses results are available for OS and PFS by genetic mutation characteristics (e.g. EGFR/ALK status; PD-L1 status) thus including the PD-L1 status subgroups identified in the scope. Additional subgroup results for OS and PFS across a range of baseline demographic variables (e.g. age, race) and disease status (e.g. ECOG performance status; metastases site – liver, lung, lymph node, adrenal gland) are also presented. CS Appendix E reports subgroup results, based on the ITT population, and also based on the ITT-WT population (for PD-L1 status only). Results for other subgroup analyses based on the ITT-WT population are available in the trial journal article and the CSR (by default these do not include the EGFR/ALK status subgroups).

The PD-L1 subgroup analysis provides results according to high, positive, low and negative PD-L1 expression subgroups, and varying combinations of these groups. As noted above, only the PD-L1 low or negative expression subgroups (tumour proportion score 0-49%, TC/IC 0,1,2) are relevant to the company's decision problem and are one of the population groups used to inform the economic model. The low or negative expression PD-L1 subgroups comprise the majority of the randomised patients across the two trial arms relevant to the company's decision problem (n=652/800, 82% in the ITT population).

The ERG asked the company to clarify whether any statistical interaction tests were performed for the subgroups, and also whether any adjustment made for multiple testing among the subgroup analyses (clarification question A11). The company responded that interaction tests and adjustments for multiple testing were not performed therefore, as is commonly the case in clinical trials, caution is required in the interpretation of these subgroup analyses. Some of the subgroups have small sample sizes which may not be sufficiently powered to detect a statistically significant difference (evidenced by wide confidence intervals). Furthermore, the subgroups are effectively observational in nature and carry a risk of potential selection bias between the randomised trial arms (though note,

randomisation was stratified by PD-L1 status, sex and the presence of liver metastases thus the risk of selection bias on these variables is lower).

3.1.6.6 Procedures for handling missing data

Censoring criteria for the assessment of PFS, OS and tumour response are reported in CS Table 6. The PFS censoring criteria appear to be similar to those commonly used in cancer treatment clinical trials. Patients who were alive and without experiencing progressed disease at time of analysis were censored on the date of the last tumour assessment; data for patients with no post-baseline tumour assessment were censored at the date of randomisation plus 1 day. The same censoring criteria were used across the analysis populations (ITT; ITT-WT; PD-L1 status).

The ERG asked the company to clarify the choice of the censoring criteria used for assessing PFS (clarification question A10). The company responded that the criteria were based on those used by the FDA.

The CS does not state whether censoring occurred for patients receiving any subsequent therapies following discontinuation of the study treatment, or for receipt of any non-protocol specified anti-cancer therapy before a PFS or an OS event.

[redacted] (see section 3.3.1).

[redacted] (see section 3.3.2).

As stated in the trial protocol, the impact of missing scheduled tumour assessments on the primary analysis of PFS was evaluated in a sensitivity analysis, using two imputation rules:

- If a patient missed two or more scheduled tumour assessments immediately prior to the date of the PFS event, the patient was censored at the last tumour assessment prior to the first of these missed visits (see section 3.3.1).

- If a patient missed two or more tumour assessments scheduled immediately prior to the date of the PFS event the patient was counted as having progressed on the date of the first of these missing assessments.

To account for patients lost to follow-up a sensitivity analysis for OS was conducted. Patients lost to follow-up were considered as having died at the last date they were known to be alive. Loss to follow-up was very small in the trial (0.3% overall) (CS appendix D Table 31) (see section 3.3.2).

[redacted]

ERG conclusion: the procedures for handling missing PFS, OS and response data in the trial are acceptable.

3.1.6.7 Statistical tests used

The statistical tests used appear to be similar to those commonly used in cancer treatment clinical trials (CS Table 6). Kaplan-Meier methodology was used to estimate median PFS and OS and to construct survival curves. A stratified log-rank test, and stratified Cox regression were performed for the co-primary endpoints of PFS and OS in the ITT-WT population and the ITT population. The stratification factors included: sex, presence of liver metastases at baseline, and PD-L1 tumour expression (i.e. the same as the randomisation stratification factors).

The ERG asked the company to clarify whether any unstratified analyses were performed in the primary analysis (clarification question A8). The company provided the unstratified analyses for PFS and OS for the ITT, ITT-WT and the EGFR/ALK+ populations. The ERG notes that the results of the stratified and the unstratified analyses are similar.

ERG conclusion: The statistical procedures used in the IMPower150 trial are appropriate for use in cancer treatment clinical trials. However, there is a complex background to the analyses populations of the trial. The trial was statistically powered for a sub-group of this trial – the ITT-WT population (87% of the ITT population). However, the assessment of clinical effectiveness and cost effectiveness in the CS is based on the ITT population (all randomised patients), to reflect the anticipated marketing authorisation.

[redacted].

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

As only one trial of atezolizumab in this indication was included in the submission, IMPower150, a meta-analysis of atezolizumab trials was not possible. The CS provides a narrative review of the trial, with data presented in tables and text.

The CS reports two indirect comparisons of atezolizumab with other treatments:

- A network meta-analysis (NMA) comparing Atezo+Bev+CP versus pemetrexed-based chemotherapy
- A matched adjusted indirect comparison (MAIC) comparing Atezo+Bev+CP versus pembrolizumab in patients with low or negative PD-L1 expression.

The NMA is used to inform estimates of clinical effectiveness in the economic evaluation. The MAIC is not used to inform the economic evaluation as the company are not seeking NHS reimbursement for pembrolizumab in patients with high PD-L1 expression (as discussed earlier in this report, section 2.2). For this reason we do not provide a critique of the MAIC in this report or report its results.

In the following sub-sections we provide a description and critique of the NMA as used to estimate OS and PFS (see also Appendix 9.1 for a quality assessment checklist of this NMA).

3.1.7.1 Evidence networks

The inclusion criteria for the NMA is reported in CS Appendix Table 10. The ERG notes that the inclusion criteria are comprehensive and match the company's decision problem. A total of seven RCTs were included in what the CS describes as the UK network based on these criteria (an additional seven trials were eligible for a "global network" to inform HTA submissions in other countries, which had a wider set of comparators. Citations to these trials are not reported in the CS).

- In three of these trials the experimental treatment under evaluation was pembrolizumab (KEYNOTE-021; KEYNOTE-024; KEYNOTE-189).
- In a further three trials the experimental treatment was pemetrexed-based chemotherapy (ERACLE; PRONOUNCE; PARAMOUNT).
- The remaining trial was the IMpower150 trial of atezolizumab.

The ERG is not aware of any trials relevant to the decision problem that have not been included in the NMA. However, as noted earlier in this report (section 2.2), there are some discrepancies between the decision problem and the NICE scope of the appraisal. Thus, trials comparing pemetrexed with other chemotherapy regimens in the NICE scope (i.e. docetaxel, gemcitabine, paclitaxel or vinorelbine) were not included. The ERG identified a published systematic review and economic evaluation of first-line chemotherapy for locally advanced or metastatic non-small cell lung cancer[4] which reports a mixed treatment comparison for OS and PFS. Pemetrexed + a platinum drug linked to both the OS and PFS networks which contained other platinum containing doublet chemotherapies (e.g. docetaxel + platinum, gemcitabine+platinum). If evidence for chemotherapy regimens such as those reported in the published systematic review[4] had been sought and been possible to include, an indirect comparison between atezolizumab and these other chemotherapies in the NICE scope (clarification question A13) might have been possible.

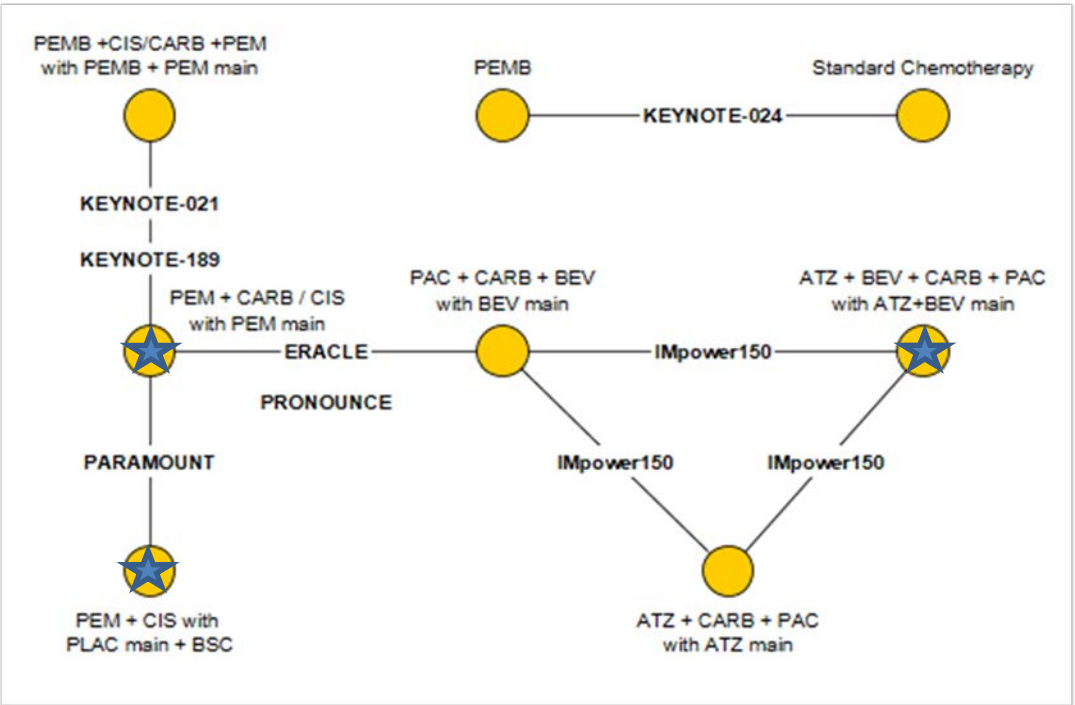
Further, the company argues that the only way to include KEYNOTE-024 in the network "is to assume that all [5] chemotherapy arms are equivalent". They then exclude the study from the network as it "does not share a common treatment arm with any of the other trials". However, we infer that the company, by excluding the docetaxel, gemcitabine, paclitaxel or vinorelbine regimens, did not attempt to connect this study to main network.

A feasibility assessment was undertaken by the company to determine the suitability of the included trials to inform a connected network of evidence for each outcome relevant to the decision problem (see CS Appendix Table 13). Following this assessment, networks were considered feasible for the outcome measures OS, PFS, ORR and adverse events leading to discontinuation. Of these, OS and PFS directly inform the company's economic model, and are the focus in this ERG report.

The structure of the OS and the PFS network is identical, as illustrated in Figure 4. As can be seen, Atezo+Bev+CP is compared with two pemetrexed-based regimens (denoted by a star symbol in the figure):

- Pemetrexed in combination with carboplatin or cisplatin, followed by pemetrexed maintenance (PEM+CARB/CIS then PEM maintenance)
- Pemetrexed in combination with cisplatin followed by placebo maintenance with best supportive care (PEM+CIS then PLAC main + BSC).

Expert clinical advice to the ERG is that, of these two regimens, pemetrexed followed by pemetrexed maintenance therapy is the standard of care in England.



ATZ - Atezolizumab BEV – bevacizumab; BSC – best supportive care; CARB – carboplatin; CIS – cisplatin; MAIN – maintenance; PAC – paclitaxel; PEM-pemetrexed; PEMB – pembrolizumab; PLAC – placebo.

★ denotes relevant NMA treatment comparisons

Source: CS Appendix Figure 2 and Figure 4

Figure 3 Network of studies informing the NMA, for OS and PFS

The network contains only one closed evidence loop, formed of the three arms of the IMpower150 trial. Therefore, there are no relevant treatment comparisons in which both indirect and direct evidence is available hence no assessment of consistency is required.

As noted above, the KEYNOTE-024 trial was not able to be connected to the network, as the control arm contains a mixture of five chemotherapy regimens, thus it is not included in the NMA. It is, however, included in the company's MAIC comparing pembrolizumab to chemotherapy in patients with high PD-L1 expression (not discussed here).

3.1.7.2 Clinical heterogeneity assessment

The eligibility criteria in the included trials and patient characteristics at baseline were similar (CS Appendix Table 15, Table 16 and Table 17), with some exceptions:

- **Induction therapy.** The PARAMOUNT trial had a different study design, with patients randomised to maintenance treatment only if they had responded to induction therapy (Only 539 of the 900 patients included in the induction phase were

randomised). This is in contrast to the other included trials, which required patients to have either no prior treatment for Stage III and/or IV non-squamous NSCLC, no previous systemic treatment, or to be chemotherapy or treatment naïve. Thus, all patients randomised in the PARAMOUNT trial had demonstrated a response to induction treatment, whilst the patients in the other trials would likely comprise a mixture of responders and non-responders to any induction treatment, and thus overall would be less likely to respond to treatment. The ERG agrees with the company's assertion. The CS considers that inclusion of the trial biases results in favour of pemetrexed plus platinum plus pemetrexed maintenance in the base case NMA and in the economic model. To address this, the company reports a scenario analysis in which this trial is omitted (CS Figure 18 and 19), the results of which were more favourable for Atezo+Bev+CP on OS and PFS (NB. the PARAMOUNT trial was the only study which included PEM+CIS then PLAC main + BSC, thus omission of this trial effectively removes this comparator from the analysis).

- **Liver metastases.** Only the IMpower150 trial reported the percentage of patients with liver metastases at baseline (16%-17%). Liver metastases reported in the CS as being associated with limited therapeutic benefit with checkpoint-inhibitor monotherapy, therefore the lack of reporting of this characteristic in the comparator trials creates uncertainty about whether this is a source of clinical heterogeneity in the network.
- **EGFR and ALK mutations**. Only the IMpower150 trial reported inclusion of patients with EGFR and ALK mutations (11.25% of patients in the Atezo+Bev+CP arm). The KEYNOTE trials excluded these patients, and the remaining pemetrexed trials did not report inclusion of any such patients. The base case NMA uses the ITT population of the IMpower150 trial (thus including EGFR and ALK positive patients from IMpower150). A sub-group NMA analysis uses outcome data for EGFR and ALK positive patients from IMpower150 compared to ITT data for the pemetrexed trials (for which it was not reported whether EGFR and ALK positive patients were included). The CS makes the assumption that EGFR and ALK status are not effect modifiers for pemetrexed based regimens. However, expert clinical advice to the ERG did not agree with this assumption. Caution is also advised given that the EGFR and ALK subgroup contains a small percentage of patients (13%).
- **ECOG performance status**. The CS notes substantial variation between (and within) trials in ECOG performance status 0-1. Expert clinical advice to the ERG is that differences in the proportions of patients with either an ECOG performance status of 0 or 1 are unlikely to be clinically significant.

- **Sex.** There was variation in the proportion of males across the trials (37% to 78%). It is unclear what impact this might have on the results of the NMA.

  ERG conclusion: There is some potential clinical heterogeneity in the NMA, primarily associated with the patients in the PARAMOUNT trial who were more likely to respond to treatment than in the other trials. In our base case analysis we exclude the PARAMOUNT trial but we retain it in a scenario analysis (see section 4.4 of this report).

3.1.7.3 Critical appraisal of trials included in the first-line treatment NMA

CS Appendix table 32 provides the company's risk of bias assessment of the trials included in the NMA.

The CS does not provide a narrative summary or discussion of this risk of bias assessment. The ERG has performed an independent risk of bias assessment of the trials included in the NMA (using only the key reference for studies other than IMpower150) which is presented in Appendix 8.2. The ERG's observations broadly concur with the company's assessment with most differences being due to the ERG assessing blinding and missing data separately for different outcomes. The chief risk of bias is that many of the studies were open-label or there was insufficient information in the primary publication for the ERG to determine if blinding was in place.

3.1.7.4 Statistical NMA methods used

The company uses a fractional polynomial approach[5] for indirect comparison estimates of OS and PFS (NB. for the outcomes of ORR and adverse events leading to discontinuation they use a generalised linear modelling approach). Unlike traditional NMA methods which assume a constant HR over time, a fractional polynomial model aims to better 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). Credible interval curves can be plotted alongside the log-hazard function curves.

The company's justification for using the fractional polynomial approach was based on the assertion that chemotherapy and immunotherapy have different mechanisms of action leading to different survival kinetics. Patients treated with the former demonstrate early survival benefits, whilst those treated with the latter show a delayed but more sustained survival benefit. Expert clinical advice to the ERG concurs with this assertion. Furthermore, a fractional polynomial approach was also used in the company submission for the NICE

appraisals of atezolizumab for treating locally advanced or metastatic NSCLC after chemotherapy (TA520)[6], atezolizumab for treating metastatic urothelial bladder cancer (TA525)[7], and a number of other recent NICE appraisals of cancer treatments.

The CS cites evidence of the difference in survival kinetics in the IMpower150 trial log-cumulative hazard plots (CS section B.3.3) where the curves cross. The ERG observes that the IMpower150 log-cumulative hazard curves do indeed cross (CS Figure 32 and 33), though we also note that the Kaplan-Meier PFS and OS curves appear generally parallel (CS Figures 3 and Figure 4, respectively). Furthermore, the CS does not state whether the proportional hazards assumption holds in the comparator trials in the NMA. A factor which will influence the uncertainty around proportionality of hazards in the trials is the maturity status of the survival data. The more mature the data the less uncertainty there is in the interpretation of proportional hazards. In response to a clarification question (question A27) the company stated that for OS, data maturity ranged from 33% to 72%, and that maturity in some trials was insufficient for median OS to be reached. For PFS, data follow-up were described as reasonably mature (over 50%) in all studies except KEYNOTE-021. The company explains that the time horizon for calculating expected survival was restricted to reduce the influence of extrapolations based on immature data.

On balance, given the expert clinical opinion and previous use in technology appraisals of a fractional polynomial model to differentiate between immunotherapy and chemotherapy we agree the use of fractional polynomial methodology is reasonable. As an alternative to the fractional polynomial time-varying hazards estimation, the company reported a fractional polynomial model approximating an exponential model (i.e. assuming a proportional hazards).

> **ERG conclusion:** The company's clinical rationale for assuming time-varying hazards between treatments is clinically justified. The use of a fractional polynomial model that approximates a proportional hazards exponential model is an informative alternative approach.

3.1.7.4.1 Model fitting

Two orders of fractional polynomial model were considered for inclusion: first-order, and second order. The exponent (power level) for each order were chosen from the following set P1=0, P1=1. A first order model with a P=0 would be equivalent to a Weibull model, and a first order model with P=1 would correspond to a Gompertz model. For the second order

model the following exponents were considered: P1=0 P2=0; P1=1 P2=0; P1=1; P2=1. (There is an apparent typo on page 134 which suggests P1=0 P2=1 but this is inconsistent with the rest of the CS.)

The ERG notes that only a relatively narrow range of powers (P1 and P2 in the range 0 to 1) were considered in the company's analysis. The CS states that the models used covered a broad range of hazard ratio shapes, and this was judged to be sufficiently broad to capture the variation in hazards observed in the data. However, we note U-shaped curves are not represented in the selection of hazard ratios presented. Further, the CS concludes that their exclusion of higher order polynomials or further exponents is consistent with previous NICE submissions however, the reference supplied (CS appendix reference 28) is unrelated to the issue of fractional polynomial models and appears to have been cited in error. Nevertheless, the ERG notes that the hazard ratio plots for OS and PFS provided by the company for the fractional polynomial models tested (clarification question A18) do encompass a variety of shapes and are likely to capture a broad range of survival estimations. The ERG therefore agrees with the company's choice of powers.

Fixed effect versions of the five fractional polynomial models and the exponential model were fitted and evaluated for the ITT analysis for both OS and PFS.

To select the most appropriate fractional polynomial model from the first and second order models considered, the company used the deviance information criterion (DIC) to compare goodness-of-fit. The DIC is commonly used to compare the fit of Bayesian statistical models with the smallest DIC indicative of best fit. The DIC values are reported in CS appendix Table 29. The company also visually inspected the hazard curves (CS appendix Figure 11 and 13) and survival curves (CS appendix Figure 12 and Figure 14), and considered the clinical plausibility of the extrapolated survival curves.

The best fitting fractional polynomial model chosen for OS and PFS was the fixed effects model with P1=0 (Weibull). This model was used in the ITT NMA as well as the subgroup and sensitivity analyses, for methodological consistency. For completeness, the ERG would have preferred the range of fractional polynomial models rerun for the subgroup and sensitivity analysis given the different population makeup. Whilst the second order models had lower DIC values (indicating better fit) the company observed that they were not clinically plausible due to unrealistically high survival times. This could also be seen as an

argument in favour of experimenting with other exponents or higher order fractional polynomial models.

> **ERG conclusion:** Having inspected the hazard ratio plots supplied by the company (clarification question A18) the ERG agrees with the company that the second order models are not clinically appropriate, and we note that they are associated with greater uncertainty due to wider credible intervals. Amongst the two first order models tested the ERG agrees with the company's choice of the P1=0 (Weibull) model. This model had a lower DIC value than the P1=1 model, and we include the P1=0 (Weibull) in our own base case analysis (see section 4.4 of this report).

3.1.7.4.2 Outcome data used in NMA

The OS and PFS survival data are reported in CS Table 20 and Table 21. However, these are not the data which input into the NMA.

Individual patient data (IPD) were available for the IMpower150 trial, combined with data reconstructed from the Kaplan-Meier curves (using the Guyot method[8]) from all other studies. It is the binary data (deaths/progression, at risk) extracted from these sources in monthly time periods which populated the NMA model and was reported in vector format in response to clarification question A21. It has not been validated. Furthermore, the company did not state whether the data reconstructed from the Kaplan-Meier graphs was validated against the reported hazard ratios.

The ERG presumes that the company has used the most recent data cuts available for the trials (NB. As discussed above, the company commented on the maturity status of the survival data in the trials in response to clarification question A27). In response to clarification question A28 the company reported that most of the trials included in the NMA used independent review committee assessment of PFS (using the RECIST criteria). The IMpower150 trial and the PARAMOUNT trial used investigator assessed PFS. Since PFS results can differ according to whether investigator-assessed or independent review committee-assessed it would be preferable to use one or the other (or both, separately) in the NMA. Since independent review is frequently more conservative the ERG would have expected a scenario analysis using the earlier (September 2017) data cut for IMpower150 which reported both investigator and independent assessment.

3.1.7.4.3 Choice between random effects and fixed effect models

As stated above, the NMA base case results are reported based on a fixed effect model. In response to a clarification question (A20) the company reported that they only fitted a random effects version of the best fitting fixed effects model (first order FP, P1=0 model) for OS and PFS. The justification the CS provides for using fixed effect model rather than random effects was that the "small differences in DIC indicated a low level of detectable heterogeneity". This is not strictly correct, DIC is simply a measure of relative model fit. It does not indicate heterogeneity or an absence thereof. Nevertheless, DIC is similar across both models and the use of a published informative prior with the random effects is indicative of there not being sufficient data to use a vague prior.

The ERG concurs that the random effects DIC values are similar to the fixed effects DIC values (CS appendix Table 29), but regards this as not a wholly sufficient justification for choice of effect models. Consideration should be given to clinical heterogeneity, and as noted above, there was notable clinical heterogeneity with regard to inclusion of the PARAMOUNT trial. We believe the analysis incorporating PARAMOUNT should use the random effects model, but otherwise the trials are sufficiently similar in terms of ECOG 0-1, disease stage and histology to justify fixed effects (see Section 3.1.7.2).

The results are similar between the two effects models (as would be expected), with the random effects model producing slightly larger credible intervals (clarification question A18, Figure 1 and Figure 2).

> **ERG conclusion:** In principle a random effects model is preferable in the presence of heterogeneity. We retain the fixed effect model in our base case analysis (which omits the PARAMOUNT trial), but we use random effects in a scenario analysis (which includes PARAMOUNT) (see section 4.4 of this report).

3.1.7.4.4 Bayesian modelling methods

The model code was written in Just Another Gibbs Sampler (JAGS) and run via R. JAGS code for selected FP models was provided in response to clarification question A17. The code was validated against published code.[9] However, the code used to approximate a proportional hazards exponential model was not provided so could not be validated. Random effects code was only provided for a first order fractional polynomial as this was the only random effects model run.

Uninformative priors were appropriately used for the fixed effects models. Informative priors as calculated by Turner (2015) were correctly implemented used for the random effects model (checked against Turner Table IV). Given the few trials available and the essentially star-shaped network, informative priors were necessary to estimate between-trial heterogeneity. No random effects model using a vague prior was reported.

A burn-in of 50,000 iterations to ensure convergence was followed by a further 50,000 for estimation, thinned by a factor of 50. Three chains were run giving a total of 3,000 iterations for parameter estimation. Trace plots and Gelman-Rubin statistics were inspected to ensure convergence and the CS model fitting process was reported to be externally validated by an independent statistician. The ERG also ran the P1=0 deterministic model and reported very similar results to the CS.

The NMA output parameters used in the economic model are not reported in the CS documentation. They are reported in the "NMA Raw Inputs" worksheet of the model which only includes the parameters for the proportional hazards and P1=0 fractional polynomial models. The HRs at each timepoint are reported in response to clarification question A17. Coda outputs from JAGS were used in the probabilistic model.

As there were no loops besides those constituted by multi-arm trials, an evaluation of network internal consistency was therefore not required.

ERG conclusion:
Based on the information provided the ERG considers that the methods used to implement the fractional polynomial model are appropriate and correspond to the methods specified in the original methodological texts.[5]

3.2 Summary statement of company's approach

Table 10 provides the ERG's response to a quality assessment checklist for systematic reviews. As can be seen, all of the pre-requisites were met. For example, record selection was independently undertaken by two reviewers with a recourse to a third reviewer for any disagreements.

Table 10 Quality assessment (CRD criteria) of 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

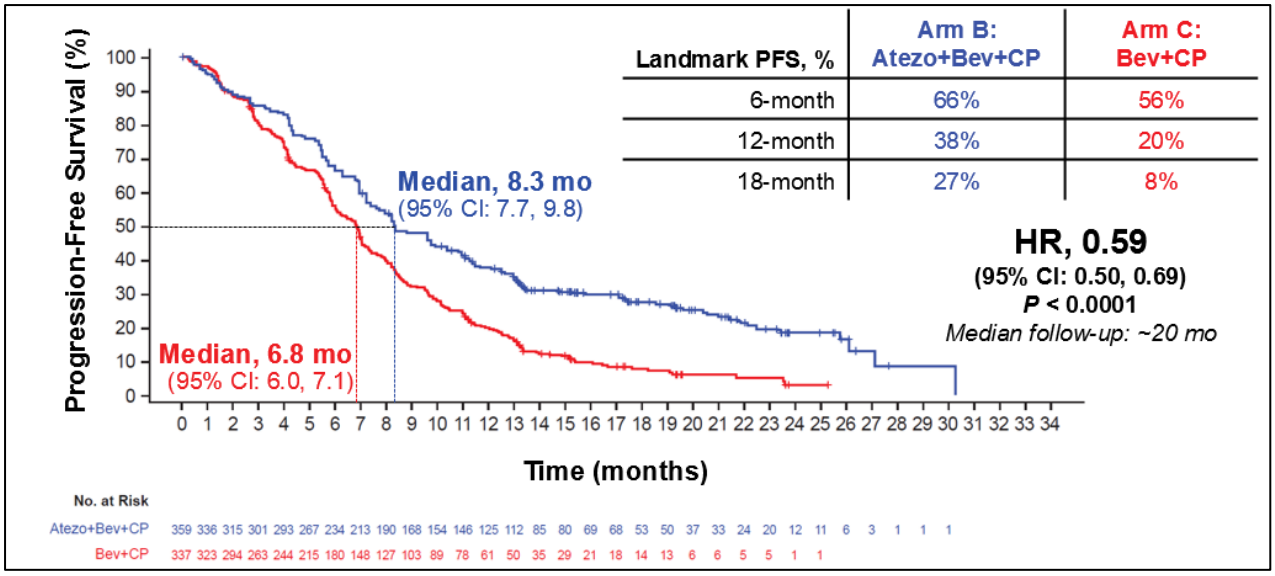
2. Is there evidence of a substantial effort to search for all relevant research? i.e. all studies identified	Yes
3. Is the validity of included studies adequately assessed?	Yes
4. Is sufficient detail of the individual studies presented?	Yes
5. Are the primary studies summarised appropriately?	Yes

3.3 Summary of submitted evidence

In the following subsections we summarise the results of the IMpower150 RCT for the most recent data cut available (22nd January 2018) as presented in the CS (see Table 9 earlier for a summary of the endpoints at the different data cuts). The anticipated marketing authorisation is based on the ITT population, (i.e. including the patients with EGFR mutant and ALK-positive NSCLC) and consequently the CS presents data for the ITT population. Data from the ITT population (as well as the PD-L1 and EGFR/ALK+ subpopulations) are also included in the economic model. Therefore, the focus in the following subsections is on the ITT population despite the fact that the co-primary endpoints of the IMpower150 RCT were analysed in the ITT-WT population.

3.3.1 PFS in the ITT population

Investigator-assessed PFS in the ITT population was a secondary outcome of the IMpower150 RCT. At the 22nd January 2018 clinical cut-off date (minimum follow up 13.5 months, median follow-up approximately 20 months) median PFS was longer in the Atezo+Bev+CP group (8.4 months, 95% CI 8.0 to 9.9) than in the Bev+CP group (6.8 months, 95% CI 6.0 to 7.0) (Figure 4). The stratified hazard ratio was 0.59 (95% CI 0.50 to 0.69).



Source: reproduction of CS Figure 3

Figure 4 KM curve – investigator-assessed PFS in the ITT population (clinical cut-off date 22 January 2018)

Upon request the company provided the independent review facility (IRF) PFS results (clarification question A4) for the September 2017 data cut. The IRF was disbanded after this time so this comparison is not possible for the later data cut of January 2018. These results are reproduced in Table 11. As can be seen, the results were similar between the two methods of assessment.

Table 11 Comparison of independent review facility and investigator-assessed PFS in the ITT population (Clinical cut off date September 15, 2017)

PFS assessor	IRF		Investigator	
	Atezo+Bev+CP n=400	Bev+CP n=400	Atezo+Bev+CP n=400	Bev+CP n=400
Patients with event, n (%)	269 (67.3)	296 (74.0)	66.8%	82.8%
Median PFS, months (95% CI)	8.5 (8.1, 9.7)	7.0 (6.1, 7.8)	8.3 (7.9, 9.8)	6.8 (6.0, 7.1)
Stratified HR (95% CI)	0.67 (0.56, 0.79)		0.61 (0.52, 0.72)	
p value	p<0.0001		p value not reported	

HR, hazard ratio; IRF- Independent Review Facility; PFS, progression-free survival

Table compiled by ERG from data presented in the responses to clarification questions, the CS and the published paper for the IMpower150 RCT.

As noted in section 3.1.6.6, a sensitivity analysis was conducted on the September 15 2017 data cut

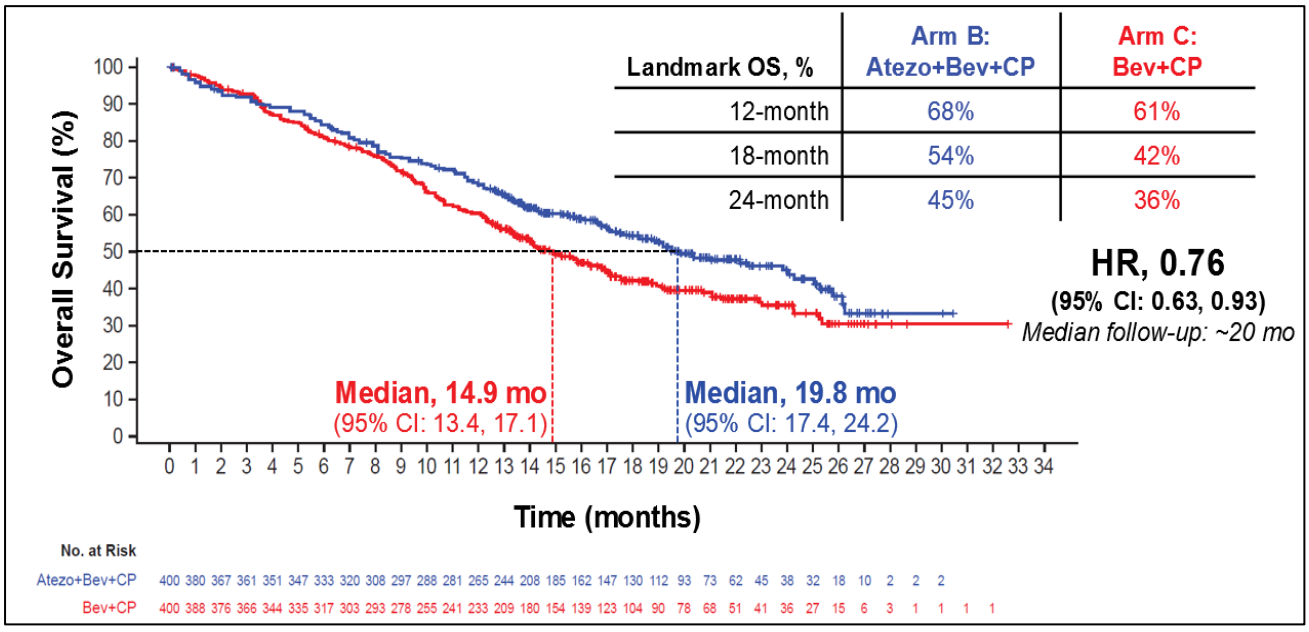
[redacted]

[redacted]. The results of a further sensitivity analysis, conducted to assess the impact of missing tumour assessments, were reported for one of the two imputation rules described in section 3.1.6.6. When patients who missed two or more scheduled tumour assessments immediately prior to the date of the PFS event were censored at the last tumour assessment prior to the first of the missed visits the results were consistent with those of the overall ITT-WT population.

ERG conclusion: Treatment with Atezo+Bev+CP leads to an improvement in PFS in the ITT population in comparison to Bev+CP.

3.3.2 OS in the ITT population

At the most recent data cut-off (22 January 2018) 192 deaths (48.0%) had been observed in the Atezo+Bev+CP group and 230 deaths (57.5%) in the Bev+CP group. As Figure 5 shows, the stratified HR was 0.76 (95% CI 0.63 to 0.93) indicating that among the ITT population, patients in the Atezo+Bev+CP arm had a 24% relative reduction in the risk of death in comparison with the Bev+CP arm. The median survival of 19.8 months (95% CI 17.4 to 24.2) in the Atezo+Bev+CP arm was 4.9 months longer than the Bev+CP arm (median OS 14.9 months, 95% CI 13.4 to 17.1).



Source: reproduction of CS Figure 4

Figure 5 KM curve –OS in the ITT population (Clinical cut-off date 22 January 2018)

As noted in section 3.1.6.6, [redacted].

ERG conclusion: Treatment with Atezo+Bev+CP leads to an improvement in OS in the ITT population in comparison to Bev+CP.

3.3.3 Response rate

Objective response (shown as 'Responders' in Table 12) was defined as all those with either a complete response (CR) or a partial response (PR).

Table 12 Summary of response in the ITT population (Clinical cut-off date 22 January 2018)

	Atezo+Bev+CP n=397	Bev+CP n=393
Responders, n (%)	224 (56.4)[a]	158 (40.2)[a]
Odds ratio (95% CI)	1.94 (1.46, 2.58)	
Complete response, n (%) (95% CI)	11 (2.8) (1.4, 4.9)	3 (0.8) (0.2, 2.2)
Partial response, n (%) (95% CI)	213 (53.7) (48.6, 58.6)	155 (39.4) (34.6, 44.5)
Stable disease, n (%) (95% CI)	111 (28.0) (23.6, 32.7)	160 (40.7) (35.8, 45.8)
Progressive disease, n (%) (95% CI)	23 (5.8) (3.7, 8.6)	38 (9.7) (6.9, 13.0)
Missing or unevaluable, n (%)	39 (9.8)	37 (9.4)

Reproduced from CS Table 11

[a] CS Table 11 has an error in this row. The correct figures were supplied by the company (clarification question A7).

In the ITT population, a higher proportion of patients in the Atezo+Bev+CP arm (56.4%, 95% CI 51.4 to 61.4) had a confirmed objective response compared with the Bev+CP arm (40.2%, 95% CI 35.3 to 45.2). The odds ratio was in favour of the Atezo+Bev+CP arm (OR =1.94, 95% CI 1.46 to 2.48).

ERG conclusion: A greater proportion of patients obtain an objective response after treatment with Atezo+Bev+CP in comparison to those treated with Bev+CP.

3.3.4 Duration of response

Of the 224 patients in the Atezo+Bev+CP arm with a confirmed objective response 136 (60.7%) had an event (either death or disease progression) with 88 (39.3%) still with an ongoing response at the clinical cut-off date (Table 13). In contrast, in the Bev+CP arm a higher proportion of those with a confirmed objective response experienced an event (88.6%) with just 11.4% of patients with an ongoing response at the clinical cut-off date. The median duration of response in the Atezo+Bev+CP arm was 11.5 months (95% CI 8.9 to 15.7) compared with 6.0 months (95% CI 5.5 to 6.9) in the Bev+CP arm (stratified HR 0.41, 95% CI 0.32 to 0.53; p<0.0001).

Table 13 Duration of confirmed response in the ITT population (Clinical cut-off date 22 January 2018)

	Atezo+Bev+CP n=224	**Bev+CP n=158**
Patients with event, n (%)	136 (60.7)	140 (88.6)
Patients with ongoing response at CCOD, n (%)	88 (39.3)	18 (11.4)
Median DOR, months (95% CI)	11.5 (8.9, 15.7)	6.0 (5.5, 6.9)
Stratified HR (95% CI) p value	0.41 (0.32, 0.53) p<0.0001	

Reproduced from CS Table 12

ERG conclusion: A greater proportion of participants had an ongoing confirmed objective response in the Atezo+Bev+CP arm than in the Bev+CP arm at the 22 January 2018 data cut off.

3.3.5 Summary of health related quality of life

Within the clinical effectiveness section of the CS (Section B.2.6) the only patient reported outcomes reported were those obtained from the EORTC QLQ-C30, which have been presented in a conference abstract.[10] These data are from the September 15th 2017 data

cut (time of final PFS analysis). EQ-5D-3L data, which are used within the economic model (see CS Section B.3.4.1), are not reported in the clinical effectiveness section of the CS. However, the company did supply these on request (clarification question B3).

3.3.5.1 EQ-5D-3L

The company supplied an Excel spreadsheet containing EQ-5D health status data in response to clarification question A5. However, no interpretation of these data was provided by the company. The spreadsheet reports UK index values with numbers of patients, mean, standard error and 95% confidence intervals for each time point. The ERG observes that the number of patients declines over time, but whether this is due to deaths, missing assessments, fewer patients with follow-up to the longer time points, or a combination of these reasons is not explained. At day 1 of the first cycle EQ-5D values in the two treatments groups were almost identical (Table 14).

Table 14 EQ-5D values on day one of the first cycle

Treatment group	Number of patients	Mean EQ-5D value	Standard Error	95% CI
Atezo+Bev+CP	359	0.699	0.014	0.671 to 0.727
Bev+CP	353	0.697	0.014	0.669 to 0.724

Note the ERG have limited the data to three decimal places

3.3.5.2 EORTC QLQ-C30

EORTC QLQ-C30 results were reported as mean change from baseline with a 10-point score change or more being used as the threshold value for clinical meaningful change. The data were interpreted only up to [redacted] Cycle 13 (39 weeks) for the Bev+CP arm because this was the point where approximately 25% or less of the evaluable population remained.

At baseline scores across the different domains of the EORTC QLQ-C30 (global health status, physical functioning and disease burden symptom scores) were comparable between treatment arms. During treatment, average global health status and physical functioning scores numerically worsened but did not cross the threshold for clinically meaningful worsening. Once chemotherapy was completed, scores numerically improved but again did not cross the threshold for clinical meaningful improvement from baseline.

Treatment related symptoms of peripheral neuropathy and alopecia worsened initially in both treatment arms (≥ 30-point mean increase from baseline for peripheral neuropathy; ≥60-point mean increase from baseline for alopecia) but over time this effect was observed to attenuate (data not presented in the CS). No clinically meaningful worsening was observed for a range of other treatment-related symptoms including fatigue, constipation, diarrhea, nausea/vomiting, haemoptysis, dysphagia and sore mouth for the period that data were interpretable (cycle 18 in the Atezo+Bev+CP arm; cycle 13 in the Bev+CP arm).

For lung cancer symptoms both the time taken to deterioration and mean changes from baseline in scores were reported. The time-to-deterioration in each of the individual lung cancer symptoms included (cough, dyspnoea single-item, dyspnoea multi-item, chest pain and pain in arm/shoulder) did not differ between treatment arms. In the ITT population median time-to-deterioration was not reached in any arm for any of the symptom scores. The mean changes from baseline in the patient-reported symptom scores decreased (improved) numerically in all treatment arms to cycle 13 but a clinically meaningful improvement was only observed for coughing scores (i.e. mean scores decreased by 10 points or more from baseline).

> **ERG conclusion**: Treatment with both Atezo+Bev+CP and Bev+CP was reported by patients to lead to worsening peripheral neuropathy and alopecia. A clinically meaningful improvement in cough was reported by patients in both trial arms. For other measures outcomes were deemed not to be clinically meaningful and were comparable between treatment arms.

3.3.6 Sub-group analyses results

The decision problem focuses on patients with a low or negative PD-L1 expression (tumour proportion score 0-49%, TC/IC 0,1,2). The ERG notes that a subgroup of participants with low or negative PD-L1 expression can be drawn from both the ITT and ITT-WT populations as shown in Table 15 (i.e. it is possible for patients to have a low or negative PD-L1 expression and be EGFR/ALK+).

Table 15: Patients with low or negative PD-L1 expression in the ITT, ITT-WT and EGFR/ALK+ groups of participants in the IMpower150 RCT

	ITT N=800		ITT-WT n=696		EGFR/ALK+ n=104	
	Atezo+Bev+CP n=400	Bev+CP n=400	Atezo+Bev+CP n=359	Bev+CP n=337	Atezo+Bev+CP n=41	Bev+CP n=63
PD-L1 low or negative sub-population	n=325	n=327	n=288	n=272	n=37	n=55

In this section we focus on

- subgroup results for patients with low or negative PD-L1 expression (tumour proportion score 0-49%, TC/IC 0,1,2) drawn from the ITT population, and
- the EGFR/ALK+ subgroup.

This is because these two patient subgroups are considered in the economic model alongside the ITT population. However, the ERG notes that whereas the hazard ratios for the ITT population reported in the CS are stratified hazard ratios, only unstratified hazard ratios are reported for the subgroups. We also briefly report the company's data for other subgroups based on baseline characteristics.

3.3.6.1 Subgroup of patients with low or negative PD-L1 expression

The subgroup of patients with low or negative PD-L1 expression represented 652/800 (81.5%) of the ITT population and randomisation was stratified by PD-L1 expression status. Investigator assessed PFS in the subgroup of patients with low or negative PD-L1 expression from the ITT population was numerically in favour of the Atezo+Bev+CP group (76.9% PFS events compared to 89.3% PFS events in the Bev+CP group) (Table 16). However, as the unstratified hazard ratio shows, the difference was not as strongly in favour of the Atezo+Bev+CP group as it was in the total ITT population (unstratified HR 0.66, 95% CI 0.56 to 0.79 vs. unstratified HR 0.58 95% CI 0.50 to 0.68 respectively).

Overall survival in the subgroup of patients with low or negative PD-L1 expression from the ITT population was also numerically in favour of the Atezo+Bev+CP group (49.2% OS events compared to 58.1% OS events in the Bev+CP group). In comparison to the ITT population (unstratified HR) the unstratified hazard ratio for the low or negative PD-L1 expression subgroup indicates slightly worse overall survival than in the ITT group with a

slightly wider confidence interval which at the upper boundary extends to 0.99 therefore falling short of the line of no effect (1.0).

CS Appendix E Figure 28 (OS) and Figure 29 (PFS) present results for a variety of comparisons between other PD-L1 expression subgroups from the ITT population. Point estimates for PFS and OS hazard ratios were all in favour of Atezo+Bev+CP but for OS in some cases the upper bound of the 95% confidence interval crossed the line of no effect.

Table 16 PFS and OS in the subgroup of patients from the ITT population with low or negative PD-L1 expression

	ITT population			**Low or negative PD-L1 expression[a] (from ITT population)**	
PFS	**Atezo+Bev+CP n=400**	**Bev+CP n=400**		**Atezo+Bev+CP n=325**	**Bev+CP n=327**
Patients with event, n (%)	291 (72.8)	355 (88.8)		250 (76.9)	292 (89.3)
Median PFS, months (95% CI)	8.4 (8.0, 9.9)	6.8 (6.0, 7.0)		8.2 (NR)	6.8 (NR)
Un-stratified HR (95% CI)	0.58 (0.50, 0.68)			0.66 (0.56 to 0.79)	
OS	**Atezo+Bev+CP n=400**	**Bev+CP n=400**		**Atezo+Bev+CP n=325**	**Bev+CP n=327**
Patients with event, n (%)	192 (48.0)	230 (57.5)		160 (49.2)	190 (58.1)
Median OS, months (95% CI)	19.8 (17.4 to 24.2)	14.9 (13.4 to 17.1)		19.1 (NR)	14.9 (NR)
Unstratified HR (95% CI)	0.77 (0.63 to 0.93)			0.80 (0.65 to 0.99)	

[a] tumour proportion score 0-49%, TC/IC 0,1,2

3.3.6.2 Subgroup analysis EGFR/ALK+ patients

The proportion of patients with an EGFR mutation or who were ALK-positive was only 13% of the ITT population (104/800). Investigator assessed PFS in the EGFR/ALK+ population was longer in the Atezo+Bev+CP group (10.0 months compared to 6.1 months in the Bev+CP group) (Table 17). The unstratified hazard ratio indicates a difference in favour of the Atezo+Bev+CP group that is slightly better than in the total ITT population (unstratified HR 0.55, 95% CI 0.34 to 0.90 vs. unstratified HR 0.58 95% CI 0.50 to 0.68 respectively).

Among the EGFR/ALK+ population overall survival was also numerically in favour of the Atezo+Bev+CP group (31.7% OS events compared to 52.4% OS events in the Bev+CP group) but median survival had not been reached the Atezo+Bev+CP group. There is therefore more uncertainty associated with the hazard ratio for overall survival and the upper bound of the confidence interval crosses the line of no effect (unstratified HR EGFR/ALK subgroup 0.54, 95% CI 0.29 to 1.03), p=0.0578 compared with ITT unstratified HR 0.77, (95% CI 0.63 to 0.93).

As the numbers of patients in the two arms of the trial that are under consideration (n=41 and n=63) and as the trial was not stratified by EGFR/ALK+ status these subgroup results should be interpreted cautiously.

Table 17 PFS and OS in the subgroup of patients from the ITT population with an EGFR mutation or who were ALK-positive

	ITT population			**EGFR/ALK+ subgroup**	
PFS	**Atezo+Bev+CP n=400**	**Bev+CP n=400**		**Atezo+Bev+CP n=41**	**Bev+CP n=63**
Patients with event, n (%)	291 (72.8)	355 (88.8)		28 (68.3)	57 (90.5)
Median PFS, months (95% CI)	8.4 (8.0, 9.9)	6.8 (6.0, 7.0)		10.0 (7.9 to 15.2)	6.1 (5.6 to 8.4)
Unstratified HR (95% CI)	0.58 (0.50, 0.68)			0.55 (0.34 to 0.90), p=0.0167	
OS	**Atezo+Bev+CP n=400**	**Bev+CP n=400**		**Atezo+Bev+CP n=41**	**Bev+CP n=63**
Patients with event, n (%)	192 (48.0)	230 (57.5)		13 (31.7)	33 (52.4)
Median OS, months (95% CI)	19.8 (17.4 to 24.2)	14.9 (13.4 to 17.1)		NE (17.0 to NE)	17.5 (10.4 to NE)
Unstratified HR (95% CI)	0.77 (0.63 to 0.93)			0.54 (0.29 to 1.03), p=0.0578	

NE= not estimable

3.3.6.3 Other sub-group analyses results

In addition to the subgroup of low or negative PD-L1 expression and the subgroup of EGFR/ALK+ patients reported above there was one further pre-planned subgroup in patients with liver metastases at baseline. Other subgroup analyses by baseline risk factors (sex, TC/IC stratification factor, age group, race, baseline ECOG, tobacco use history, lung metastasis at enrolment, lymph node metastasis at enrolment, adrenal gland metastasis at

enrolment, intended number of induction treatment cycles, EML4-ALK rearrangement status and KRAS mutation status) for overall survival and progression-free survival in the ITT population are presented in CS Appendix E. These were not pre-planned.

The proportion of patients with liver metastases at enrolment was 17%. Liver metastases are known to confer a poor prognosis but in this small subgroup Atezo+Bev+CP treatment still led to a PFS and OS benefit [unstratified HRs: PFS 0.52 (95% CI 0.33 to 0.82); OS 0.41 (95% CI 0.26 to 0.62)].

Across other subgroups analyses by baseline risk factors in the ITT population a PFS benefit was observed in many. However, some subgroups were small and the results uncertain as indicated by wide confidence intervals (e.g. for the subgroup of six participants aged 85 years or over and the 15 participants of Black or African American race). For OS, although a benefit was observed in many subgroups with central OS estimates ranging between 0.47 and 1.06 the upper limit of the 95% confidence interval reaches or crosses 1 for more than half of the subgroups. The results by baseline risk factors should be interpreted cautiously because, other than the PD-L1 expression, EGFR/ALK genetic alteration and liver metastases at baseline subgroups, they were not preplanned, patient numbers are small in some groups and in response to clarification question 11 the company confirmed that no interaction tests or multiplicity adjustment were performed in the subgroup analyses.

> **ERG conclusion**: The PFS and OS benefit for Atezo+Bev+CP versus Bev+CP was maintained across the pre-planned subgroups. The results for the posthoc subgroup analyses are more uncertain due to wide confidence intervals.

3.3.7 Network meta-analysis results

The CS presents forest plots and hazard ratio plots for the fractional polynomial NMA comparing Atezo+Bev+CP with the two comparators included in the decision problem:

A. Pemetrexed in combination with carboplatin or cisplatin, followed by pemetrexed maintenance (PEM+CARB/CIS then PEM maintenance). In the company's economic evaluation this is referred to as 'pemetrexed plus platinum plus pemetrexed maintenance'.
B. Pemetrexed in combination with cisplatin followed by placebo maintenance with best supportive care (PEM+CIS then PLAC main + BSC). In the company's economic evaluation this is referred to as 'pemetrexed plus platinum'.

Below we briefly summarise the results. For full details please see CS section B.2.9 and CS Appendix D. Additional results can be found in Appendix A of the company's response to clarification questions. We summarise results for the ITT population, the EGFR/ALK positive subpopulation, and the PD-L1 low / negative subpopulation. See section 4.2.4.1.1 of this report for further information on how these populations were used to inform the fitting of baseline survival curves for atezolizumab in the economic model.

3.3.7.1 Overall survival

In the **ITT population**, as Figure 6 shows, Atezo+Bev+CP had a statistically significantly longer expected survival relative to comparison B, PEM+CIS then PLAC main + BSC, but not relative to comparison A, PEM+CARB/CIS then PEM maintenance. For the latter the credible interval crossed zero (indicating no statistically significant difference between treatments).

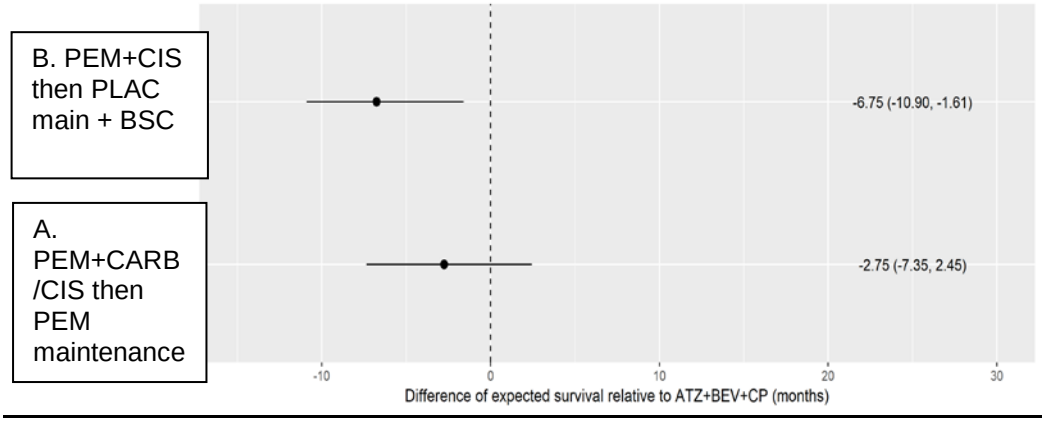
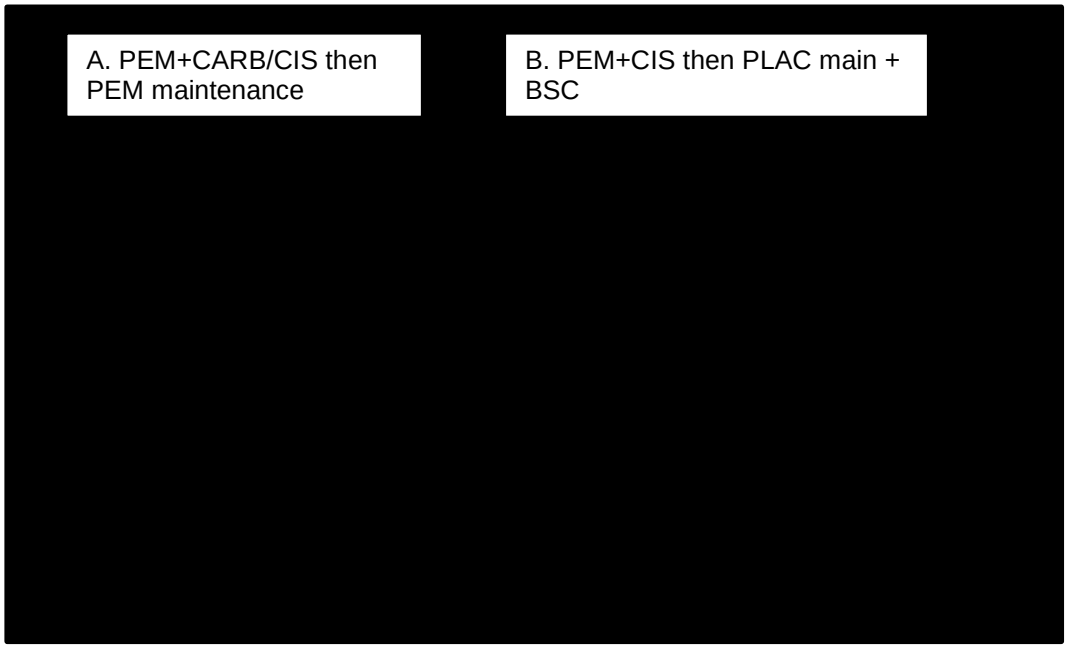


Figure 6 Forest plot of the expected mean OS difference relative to Atezo+Bev+CP (time horizon 60 months)
Reproduced from CS figure 10

- [redacted]

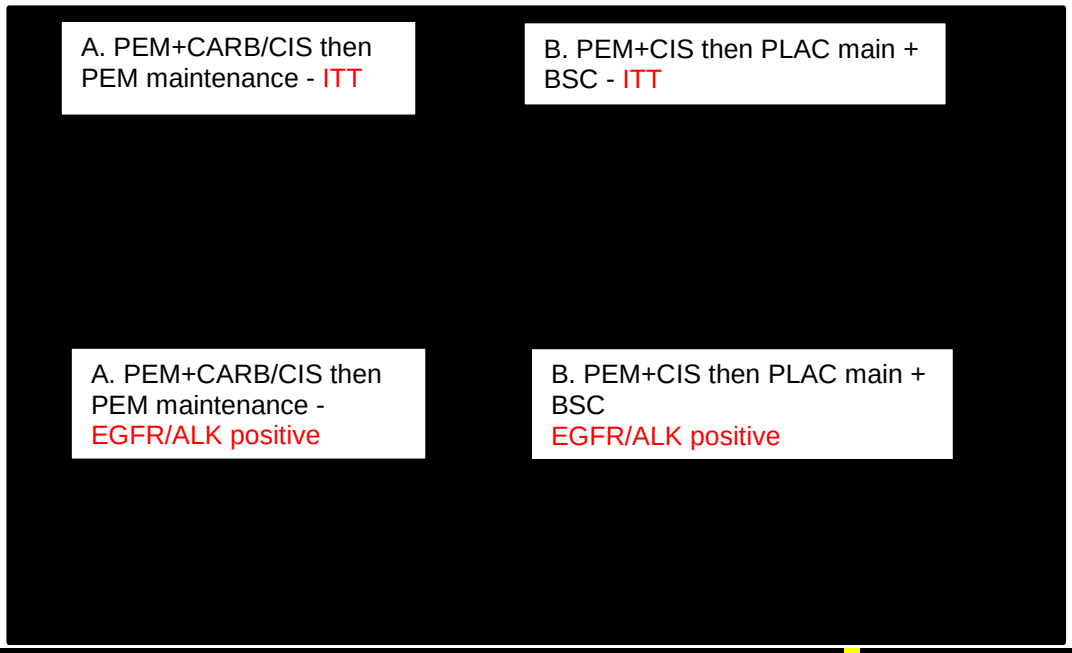
[redacted]7[redacted]

[redacted]



EGFR/ALK positive subgroup

[redacted]



[redacted] 8 [redacted]

PD-L1 low or negative subgroup (CS Figure 16):

- [redacted]
- [redacted]

[redacted]

3.3.7.2 Progression free survival

In the **ITT population**, the PFS results statistically favoured Atezo+Bev+CP compared to both comparator treatments. As Figure 9 shows, there was a statistically significantly longer expected PFS relative to PEM+CIS then PLAC main + BSC, and to PEM+CARB/CIS then PEM maintenance. The gain in PFS was greater compared to PEM+CIS then PLAC main + BSC.

[redacted]

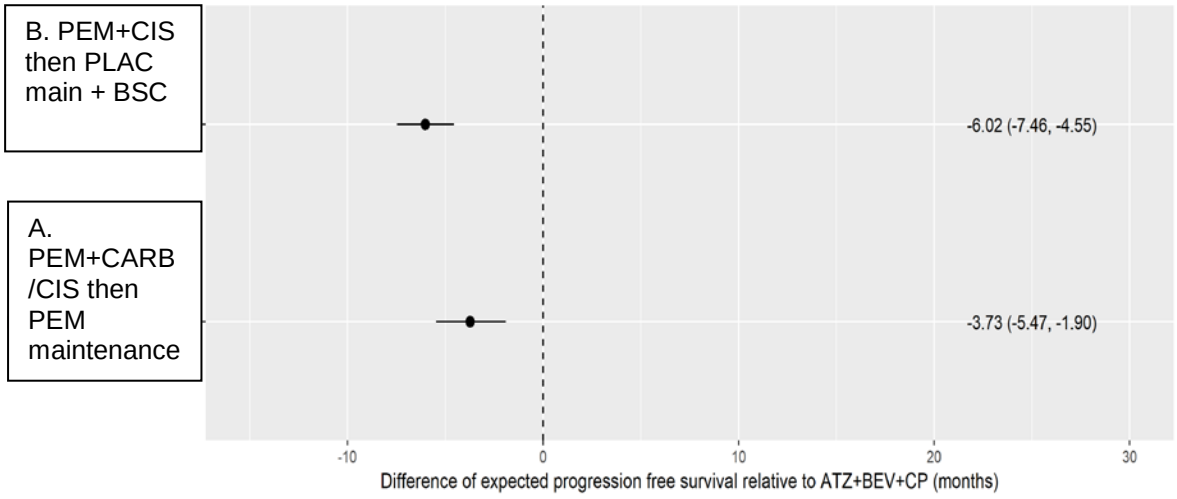
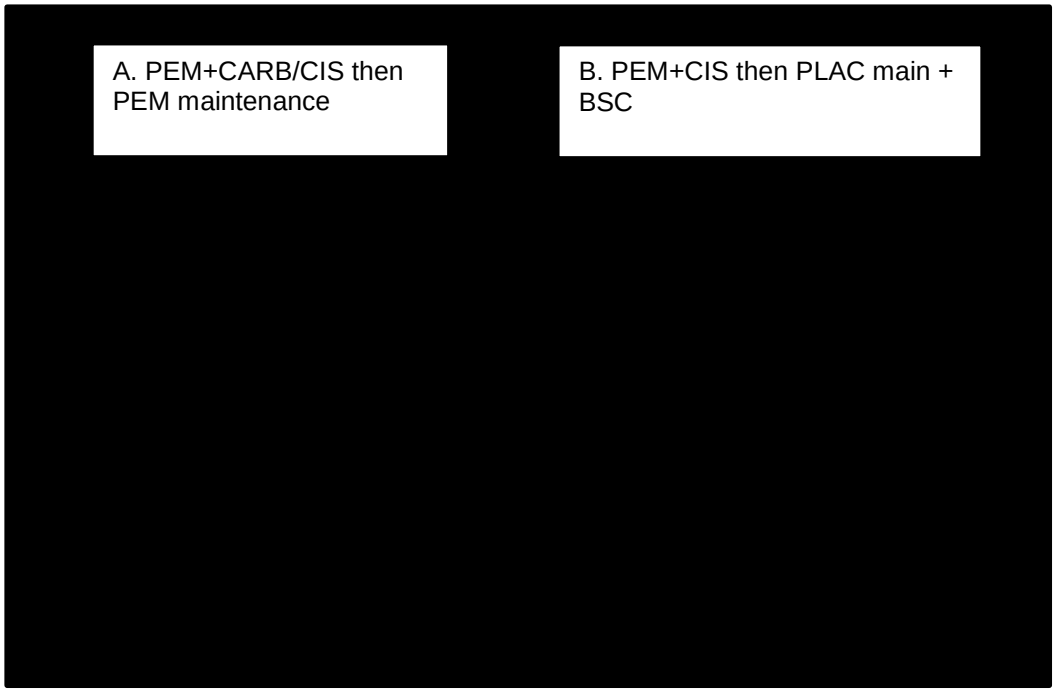


Figure 9 Forest plot of the expected PFS difference relative to Atezo+Bev+CP (time horizon 30 months)
Reproduced from CS Figure 12

The time-varying HR plots ([redacted], and CS Figure 13) show similar results to the forest plots:

- [redacted]
- [redacted]

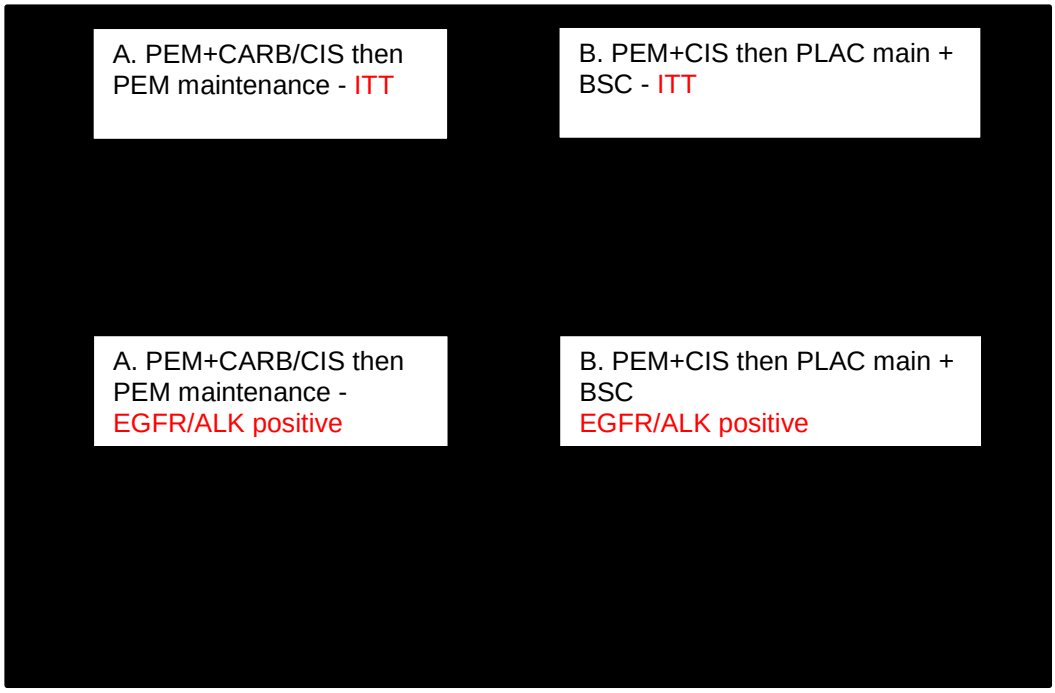


[redacted]

EGFR/ALK positive subgroup:

[redacted]

- [redacted]
- [redacted]



[redacted]

PD-L1 low or negative subgroup:

- [redacted]
- [redacted]

[redacted]

ERG conclusion: [redacted]

[redacted]

[redacted]

3.3.7.3 NMA s ensitivity analyses

The scenario analysis excluding the PARAMOUNT trial improved the OS and PFS survival estimates in favour of Atezo+Bev+CP compared to PEM+CARB/CIS then PEM maintenance (the comparison to PEM+CIS then PLAC main + BSC was no longer possible with the omission of this trial) (CS Figure 18, 19, 20, 21).

The scenario analysis using a proportional hazards model (exponential fractional polynomial) showed more favourable results in favour of Atezo+Bev+CP compared to the two pemetrexed comparator regimens than was the case under the best fitting fractional polynomial model (CS Figure 22, Figure 23, Figure 24, Figure 25). It should be acknowledged, however, that the proportional hazards assumption cannot necessarily be applied to these trial data (as discussed earlier, section 3.1.7).

3.3.8 Summary of adverse events

Information on adverse events comes from the safety population of the IMpower150 trial. The safety population included all treated patients who received any amount of any component of study treatment. Patients were grouped according to whether they received any amount of atezolizumab or not. Note however that there is a minor inconsistency in the CS. CS Appendix D Figure 19 (Patient disposition in IMpower150 at the time of the updated analysis) shows 394 treated patients in both the Atezo+Bev+CP and Bev+CP arms of the trial but CS Tables 17 to 22 show only 393 patients in the safety population for the Atezo+Bev+CP group and 394 in the Bev+CP group.

The CS presents an overview of the safety profile of Atezo+Bev+CP compared with Bev+CP which is reproduced below in Table 18. The total number of adverse events was higher in the Atezo+Bev+CP group (n=6419) compared with the Bev+CP group (n=4630). However, the proportion of patients with at least one adverse event or one treatment-related adverse event was similar between groups (patients with at least one adverse event: Atezo+Bev+CP 98.2% vs Bev+CP 99.0%; patients with at least one treatment-related adverse event Atezo+Bev+CP 94.1% vs Bev+CP 95.7%). As Table 18 shows, the proportion of patients experiencing treatment-related Grade 3-4 adverse events, serious adverse events and treatment-related serious adverse event were all higher in the Atezo+Bev+CP arm compared with Bev+CP arm.

Additional details regarding the types of adverse event, types of treatment-related grade 3-4 adverse events, grade 5 adverse events and serious adverse events are summarised in the CS with key information provided below.

Table 18 Overview of the safety profile of Atezo+Bev+CP compared with Bev+CP (Clinical cut-off date 22 January 2018)

n, (%)	Atezo+Bev+CP n=393	Bev+CP n=394
Total number of events	6419	4630
Total number of patients with at least one:		
Adverse event	386 (98.2)	390 (99.0)
Treatment-related AE	370 (94.1)	377 (95.7)
Grade 3–4 AE	250 (63.6)	230 (58.4)
Treatment-related Grade 3–4 AE	223 (56.7)	191 (48.5)
Grade 5 AE	24 (6.1)	21 (5.3)
Treatment-related Grade 5 AE	11 (2.8)	9 (2.3)
Serious AE	174 (44.3)	135 (34.3)
Treatment-related serious AE	103 (26.2)	78 (19.8)
AE leading to withdrawal from any treatment	133 (33.8)	98 (24.9)
AE leading to any dose modification/interruption	246 (62.6)	188 (47.7)

Reproduced from CS Table 17

Among the total number of patients who experienced at least one adverse event there were some events, shown in Table 19, where there was a difference of at least 5% between treatment arms. With the exception of epistaxis (more commonly known as nosebleed) which was experienced by a greater proportion of patients in the Bev+CP arm, the remaining types of adverse event in Table 19 were experienced by a greater proportion of patients who received Atezo+Bev+CP. However, the CS states that the majority of the common adverse events were of Grade 1 or 2 and were generalised symptoms and events that are consistent with events known to be associated with the Bev+CP chemotherapy backbone.

Table 19 Common adverse events with a difference of at least 5% between treatment arms (Clinical cut-off date 22 January 2018)

n, (%)	Atezo+Bev+CP n=393	Bev+CP n=394
Total number of patients with at least one AE	386 (98.2)	390 (99.0)
Gastrointestinal disorders		

Nausea	154 (39.2)	125 (31.7)
Constipation	117 (29.8)	92 (23.4)
Diarrhoea	126 (32.1)	97 (24.6)
Stomatitis	51 (13.0)	25 (6.3)
General disorders and administration site conditions		
Fatigue	130 (33.1)	107 (27.2)
Pyrexia	73 (18.6)	34 (8.6)
Nervous system disorders		
Peripheral neuropathy	93 (23.7)	68 (17.3)
Skin and subcutaneous tissue disorders		
Rash	65 (16.5)	26 (6.6)
Pruritus	50 (12.7)	24 (6.1)
Respiratory, thoracic and mediastinal disorders		
Epistaxis	66 (16.8)	87 (22.1)
Metabolism and nutrition disorders		
Decreased appetite	113 (28.8)	83 (21.1)
Hypomagnesaemia	51 (13.0)	23 (5.8)
Hypokalaemia	37 (9.4)	16 (4.1)
Endocrine disorders		
Hypothyroidism	45 (11.5)	11 (2.8)

Reproduced from CS Table 19

Treatment-related adverse events were also comparable between treatment arms (Atezo+Bev+CP 94.1%; Bev+CP 95.7%). The CS provides a summary of the treatment-related Grade 3-4 adverse events that occurred with an incidence of at least 2% (CS Table 20). The most commonly experienced grade 3-4 treatment-related adverse event in both groups was neutropenia (Atezo+Bev+CP 14%; Bev+CP 11.2%). Grade ≥ 3 adverse events with an incidence of ≥2% are included in the economic model (see sections 4.2.4.5 and 4.2.6.6 of this report).

A higher proportion of people in the Atezo+Bev+CP group experienced a serious adverse event than in the Bev+CP group (44.3% versus 34.3% respectively) with the most common serious adverse event being febrile neutropenia (Atezo+Bev+CP 6.4%; Bev+CP 3.8%).

At the 22nd January 2018 clinical cut-off date there had been 189 deaths in the Atezo+Bev+CP arm and 226 in the Bev+CP arm (Table 20). Of these, 81% (153/189) in the Atezo+Bev+CP arm and 87% (197/226) in the Bev+CP arm were due to progressive disease. Approximately 10.8% of deaths were due to adverse events [Atezo+Bev+CP 12.7% (24/189); Bev+CP 9.3% (21/226)] and 4.8 % due to other reasons.

Table 20 Fatal adverse events and causes (Clinical cut-off date 22 January 2018)

n, (%)	Atezo+Bev+CP n=393	Bev+CP n=394
All deaths	189 (48.1)	226 (57.4)
Adverse event	24 (6.1)	21 (5.3)
Progressive disease	153 (38.9)	197 (50.0)
Other[a]	12 (3.1)	8 (2.0)

Reproduced from CS Table 21

[a] Includes fatal events that are unrelated to study treatment and occur outside the reporting period

Of the grade 5 adverse events (i.e. deaths due to adverse events) fewer than half were judged to be related to any study treatment (Table 21). Among all the grade 5 adverse events the most commonly reported (at least 3 patients) were haemoptysis, pneumonia and febrile neutropenia.

Table 21 Grade 5 AEs [most commonly reported (at least 3 patients), clinical cut-off date 22 January 2018]

n, (%)	Atezo+Bev+CP n=393	Bev+CP n=394
Any adverse event, grade 5	24 (6.1)	21 (5.3)
Haemoptysis	[redacted]	[redacted]
Pneumonia	[redacted]	[redacted]
Febrile neutropenia	[redacted]	[redacted]
Grade 5 events related to any study treatment	11 (2.8)	9 (2.3)

Some of the numbers in this table were sourced from the clinical study report (CIC marked)

Adverse events of special interest were pre-defined in the protocol. They were based on the mechanism of action of atezolizumab and known adverse events associated with other immune-modulating treatments. As Table 22 shows, the majority of adverse events of special interest were of grade 1 or 2 in severity, in the Atezo+Bev+CP arm 12.5% were grade 3-4 in comparison to 3.3% in the Bev+CP arm, and there were no grade 5 events in either arm. Adverse events reported with at least a 2% difference between study arm are summarised in Table 22.

Table 22 Summary of selected adverse events of special interest to atezolizumab (Clinical cut-off date 22 January 2018)

n, (%)	Atezo+Bev+CP n=393	Bev+CP n=394
Total number of patients with at least one AESI	206 (52.4)	112 (28.4)
Total number of patients with at least one:		
Treatment-related AESI	182 (46.3)	70 (17.8)
Grade 3–4 AESI	49 (12.5)	13 (3.3)
Treatment-related Grade 3–4 AESI	42 (10.7)	8 (2.0)
Grade 5 AESI	0	0
Treatment-related Grade 5 AESI	0	0
Serious AESI	25 (6.4)	4 (1.0)
Treatment-related AESI	22 (5.6)	2 (0.5)
AESI leading to withdrawal from any treatment	26 (6.6)	3 (0.8)
AESI leading to any dose modification/interruption	51 (13.0)	16 (4.1)
Patients with at least one (incidence ≥2%)		
Immune-related rash	117 (29.8)	53 (13.5)
Immune-related hepatitis (diagnosis)	54 (13.7)	29 (7.4)
Immune-related hepatitis (laboratory abnormality)	48 (12.2)	29 (7.4)
Immune-related hypothyroidism	56 (14.2)	18 (4.6)
Infusion-related reactions	14 (3.6)	12 (3.0)
Immune-related pneumonitis	13 (3.3)	5 (1.3)
Immune-related hyperthyroidism	16 (4.1)	5 (1.3)
Immune-related colitis	11 (2.8)	2 (0.5)

Reproduced from CS Table 22
AESI, adverse event of special interest;

4 COST EFFECTIVENESS

4.1 ERG comment on company's review of cost-effectiveness evidence

The company presents summary results from a systematic search for economic evaluations of first-line treatments for non-squamous NSCLC (CS section B.3.1 and Appendix G). As stated earlier in this report (section 3.1.1) we regard their search strategy to be comprehensive.

The review identified 66 economic evaluations with full publications in English (CS Appendix G Table 37). Out of these studies, ten used UK data, of which seven were NICE technology appraisals. None of the 66 studies included atezolizumab. Three studies related to NICE appraisals of comparators specified in the scope: TA181 for pemetrexed with cisplatin,[2] TA190 for pemetrexed maintenance after platinum-based chemotherapy[11] and TA447 for pembrolizumab (updated in TA531[12]). We note that NICE has also published guidance on pemetrexed maintenance after pemetrexed and platinum induction in this population (TA402).[13]

Methods and results of the previous NICE appraisals for comparators in the NICE scope for atezolizumab are briefly summarised in Table 23 below. These are a potential source for cross-validation of results from the submitted model (see validation section 4.3.3 below), although none of the results are directly comparable:

- **TA181** had a shorter time horizon, a different source of effectiveness evidence (the JMDB trial) and different model structure.
- **TA402** outcomes relate to a selected population without disease progression after four cycles of PEM+CIS induction therapy and exclude costs and QALYs accrued during the induction period. The evidence base for pemetrexed maintenance in the current submission is also broader; including data from the KEYNOTE, ERACLE and PRONOUNCE trials as well as PARAMOUNT.
- **TA531** used similar methods to the current submission but results relate to a blended Standard of Care (SOC) comparator and a subgroup with high PD-L1.

Table 23 NICE technology appraisals for comparators

Study	Model	Intervention/ comparator	Population	Submitted base case for companies in TA [a]		
				Cost	QALYs	ICERs (£ / QALY)
NICE 2009 **TA181**	Markov (response; stable; PD; death) 6 years	**PEM + CIS** / • GEM+CIS (JMDB) • GEM+CARBO & DOC+CIS (ITC)	Untreated advanced NSCLC	GEM+CIS £10,310 PEM+CIS £11,674 Incr. £1,364	GEM+CIS 0.57 PEM+CIS 0.61 Incr. 0.04	Company: £33,065 **'Most plausible': £17,000 to £25,000** (for adenocarcinoma or LCC subgroup)
NICE 2010 **TA190**	Trial-based analysis (not progressed, progressed, terminal): 6 years	**PEM maintenance** / BSC (placebo) (JMEN trial)	Advanced NS NSCLC **after platinum-based chemotherapy**	BSC £8,318 PEM £17,455 Incr. £9,137	BSC 0.70 PEM 0.97 Incr. 0.27	Company: £33,732 **'Most plausible': £47,000 to £51,000**
NICE 2016 **TA402** (CDF review of TA309)	Markov (PF, PD, death) 16 years	**PEM maintenance** / BSC (placebo) (PARAMOUNT trial)	Advanced NS NSCLC **after PEM+CIS**	At CDF review**:** BSC £9,344 PEM £24,272 *Incr. £14,927*	At CDF review**:** BSC 0.91 PEM 1.12 *Incr. 0.21*	Company: £70,538 (list price) **'Most plausible':** Confidential with CAA
NICE 2018 **TA531** (update of TA447)	Partitioned survival (PF, PD, death) 20 years	**PEMB** / SOC (platinum-based chemo regimen with or without PEM maintenance) (KEYNOTE-024)	Untreated metastatic NSCLC with high PD-L1 (and not EGFR/ALK+)	PEMB £72,353 SOC £43,364 *Incr. £28,989*	PEMB 2.31 SOC 1.35 *Incr. 0.96*	Company: £30,244 **'Most plausible': £30,000 to £50,000**

[a] Results reported in company base case and 'most plausible' ICER from committee conclusions

BSC best supportive care; CAA Commercial Access Agreement; CDF Cancer Drugs Fund; CIS cisplatin; Incr. incremental; PD progressed disease; PF progression free; PEM pemetrexed; SOC standard of care;

4.2 Summary and critique of the company's submitted economic evaluation

Sections B.3.2 to B.3.11 of the CS report on the methods and results of a new economic model developed by the company for this appraisal.

4.2.1 NICE reference case checklist

The ERG assessment of whether the submitted economic evaluation complies with NICE reference case requirements is shown in Table 24. We consider that the company's analysis broadly conforms to the reference case, except that the modelled decision problem differs from the NICE scope. We discuss these differences in the following section.

Table 24 NICE reference case requirements

NICE reference case	Included in submission	ERG comment
Decision problem: The scope developed by NICE	No	The company's economic evaluation does not address the full population and comparators stipulated by NICE. In particular, people with high PD-L1 expression who would be eligible for pembrolizumab are excluded. See CS B.3.2.1 and section 4.2.2.1 below.
Comparator(s): As listed in the scope developed by NICE	No	The company economic analysis omits comparators in the scope, CS B.3.2.3. In particular, we note that none of the comparators specified for EGFR/ALK positive patients are modelled, see 4.2.2.2 below.
Perspective on outcomes: All direct health effects, whether for patients or, when relevant, carers	Yes	
Perspective on costs: NHS and PSS	Yes	CS Table 24
Type of economic evaluation: Cost utility analysis with fully incremental analysis	Yes	The CS does not include a full incremental analysis, but this was provided in response to clarification question B4 (Clarification Response Appendix D)
Time horizon: Long enough to reflect all important differences in costs or outcomes the between technologies being compared	Yes	CS Table 24
Synthesis of evidence on health effects: Based on systematic review	Yes	CS Section 2.9
Measuring and valuing health effects: Health effect should be expressed in QALYs. The EQ-5D is the preferred measure of health-related quality of life in adults	Yes	CS Table 24, CS Table 30

Source of data for measurement of health-related quality of life: Reported directly by patients and/or carers	Yes	
Source of preference data for valuation of changes in health-related quality of life: 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	
Evidence on resource use and costs: Costs should relate to NHS and PSS resources and be valued using prices relevant to the NHS and PSS	Yes	CS Table 24
Discounting: The same annual rate for costs and health effects (currently 3.5%)	Yes	CS Table 24

4.2.2 Modelled decision problem

See section 2.2 above for the ERG summary and critique of the company's decision problem. We summarise key differences between the scope and the modelled population and comparators below.

4.2.2.1 Population and subgroups

Two populations are included in the model (CS section B.3.2.1):

- Adults with untreated metastatic non-squamous NSCLC with low or negative PD-L1 (tumour proportion score 0-49%, TC/IC 0, 1 or 2); and
- Adults with metastatic non-squamous NSCLC who are EGFR or ALK positive after targeted therapy (or who cannot have targeted therapy).

This deviates from the NICE scope in two respects. Firstly, although the scope relates to advanced NSCLC, the model is restricted to metastatic disease only. This restriction is appropriate because it follows the anticipated marketing authorisation for Atezo+Bev+CP and is consistent with clinical evidence from the IMpower150 study.

Secondly, the company excludes the subgroup with high PD-L1 expression from the untreated population. This subgroup is included in the anticipated marketing authorisation and the IMpower150 trial population, but the company states that it is not seeking

reimbursement for the high PD-L1 subgroup based on the comparison with pembrolizumab. [redacted]. Expert clinical advice to the ERG concurs with this view (see section 2.2 of this report).

Patient characteristics in the model are based on means across the Atezo+Bev+CP and Bev+CP arms of the IMpower150 trial (Table 25). The same values are used for the EGFR/ALK positive and PD-L1 low/negative populations. We understand that these characteristics are realistic for patients in clinical practice, although EGFR and ALK positive patients are more often younger and female.

Table 25 Patient characteristics used in model

Baseline characteristic	Value (ITT and subgroups)
Age (years)	63
Body weight (kg)	72
Glomerular filtration rate (GFR)	90.4
BSA (m^2)	1.81

4.2.2.2 Intervention and comparators

The model includes three combination therapies for both modelled populations (CS B.3.2.3):

- Atezolizumab with bevacizumab, carboplatin and paclitaxel
- Pemetrexed in combination with a platinum drug (cisplatin or carboplatin)
- Pemetrexed in combination with a platinum drug and pemetrexed maintenance

Intervention

The scope also includes atezolizumab with carboplatin and paclitaxel (without bevacizumab) as an intervention, but this is not covered in the anticipated marketing authorisation. On this basis, it is appropriate to omit it from the economic analysis.

Comparators for the untreated population

The model omits two scoped comparators for the untreated population: pembrolizumab and chemotherapy (docetaxel, gemcitabine, paclitaxel or vinorelbine) in combination with a platinum drug (carboplatin or cisplatin). We consider it reasonable for the company to have omitted pembrolizumab as a comparator, as they are not seeking NHS reimbursement for the high PD-L1 subgroup for whom pembrolizumab is recommended at first line (TA531).

The company argue that it is reasonable to omit chemotherapy regimens as pemetrexed in combination with a platinum drug is the standard of care, with 83% of the market share for this indication (clarification response B1). The scope specifies that the pemetrexed should be used in combination with cisplatin specifically, in line with NICE guidance (TA181).[2] However, a proportion of patients cannot tolerate cisplatin, due to its side effects. Clinical advice to the ERG is that these patients would have either pemetrexed with carboplatin, or a carboplatin chemotherapy doublet followed by pemetrexed maintenance. As noted in section 2.1 above, UK audit data does suggest that pemetrexed is sometimes given in combination with carboplatin.[3] The company also argue that data on pemetrexed with both platinum drugs were pooled in the NMA for the pembrolizumab appraisal TA531[12], as well as in the NMA for this current appraisal.

Comparators for EGFR or ALK positive people after targeted therapy

The company does not model either of the comparators specified in the scope for the EGFR/ALK positive population: the scope cited NICE TA520, which included docetaxel alone for PD-L1 negative disease and pembrolizumab for PD-L1 positive disease as comparators.[6] In their response to clarification question B1, the company stated that docetaxel and pembrolizumab are not appropriate comparators as they are only licensed and reimbursed for EGFR/ALK positive patients after targeted therapy and after treatment with chemotherapy: "effectively second-line after targeted therapy". We understand that this interpretation is correct and that docetaxel and pembrolizumab should not be considered as comparators for people with EGFR or ALK mutations.

Conversely, the company include pemetrexed-based comparators for EGFR/ALK positive patients after targeted treatment, even though this is not specified in the scope. We understand that in practice, pemetrexed combinations would be used in this population.

ERG conclusions: The decision problem addressed in the company's economic evaluation differs from that specified in the NICE scope. The restrictions to metastatic disease and to the atezolizumab combination with bevacizumab are appropriate, as they are consistent with the proposed marketing authorisation. We also consider the exclusion of the subgroup with untreated disease and high PD-L1 expression to be acceptable, as the company is not seeking NHS reimbursement for this subgroup.

The company restricts the modelled comparators to pemetrexed with platinum, with or without pemetrexed maintenance. Although this is not fully consistent with the scope, we understand that it is a reasonable representation of current practice for

most patients. We note, however, that the model does not compare against carboplatin-based chemotherapy followed by pemetrexed maintenance, which is an option for patients who cannot tolerate cisplatin. It is unclear how this omission affects the incremental cost-effectiveness results.

4.2.3 Model structure and assumptions

The company describe the key features and assumptions of their economic model in section B.3.2.2 of the CS. We reproduce their illustration of the model structure below.

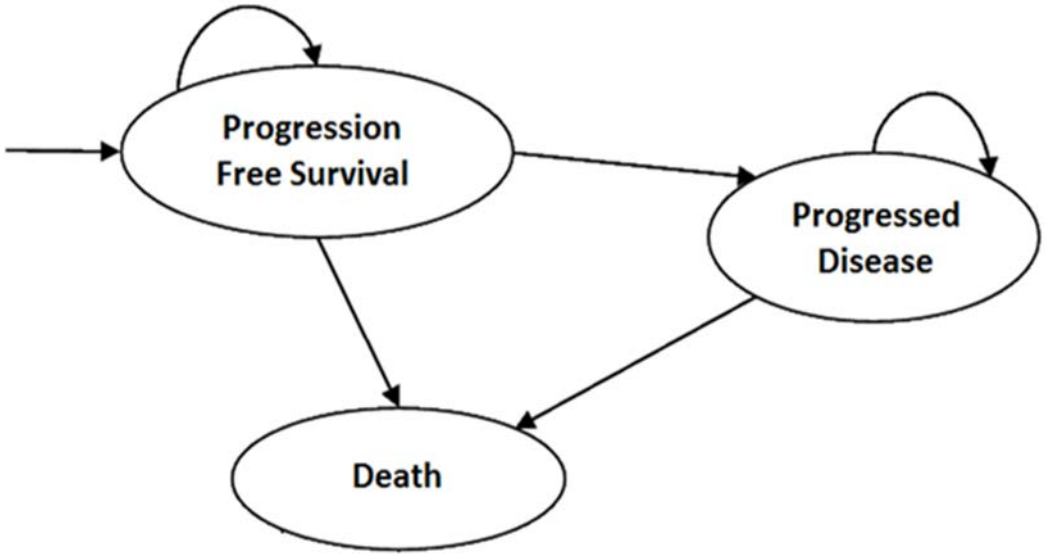


Figure 10 Economic model (reproduced from CS Figure 31)

The model follows a partitioned-survival approach with three health states: progression free (PF), progressed disease (PD) and death. The distribution of the cohort between the three states at each point in time is derived from PFS and OS curves, estimated from IMpower150 data and the NMA. All patients start in the PF state, at initiation of one of the modelled treatments. Patients move from PF to PD if their disease progresses, with the number of progressions per model cycle determined by the difference between the OS and PFS curves. Time to Treatment Discontinuation (TTD) curves estimated from trial data set the duration of each first-line medication. The model does not explicitly reflect subsequent lines of treatment, but an average cost for subsequent therapies in the PD state is included. Over time, patients transition to the absorbing state of death, with the number of deaths per cycle determined by the OS curve. The three-state partitioned-survival model is common in cancer appraisals and the ERG considers it appropriate in this case.

The company compare key features of their model with those in other related NICE appraisals in CS Table 24. We summarise the ERG view on these and other key assumptions in Table 26 below. The model has a cycle length of one week with a half-cycle correction applied to all relevant outcomes. The time horizon is 20 years. This is sufficient to reflect important cost and outcome differences between the comparators because only a small proportion of the modelled cohort are alive after 20 years. Costs and health outcomes are appropriately discounted at 3.5%.

4.2.3.1 Treatment stopping rule and duration of effect

The model also includes assumptions about the maximum duration of treatment and persistence of survival benefits for Atez+Bev+CP (section B.3.2.2 of the CS). In the base case, treatment with atezolizumab and bevacizumab stops after a maximum of two years and the effect on survival lasts for a further three years (five years from treatment initiation). The company state that these assumptions are conservative, adopted for consistency with other NICE appraisals of atezolizumab (TA520[6] and TA525[7]). They report scenarios with no stopping rule and with longer effects on survival, from 105 to 240 months (CS section B.3.8.3). It is stated in the CS that the effect cap for Atez+Bev+CP is also applied to PFS (section B.3.3.3). However, we note that although the model includes this as an option, the cap on PFS effects is not applied in the company's base case or scenario analyses.

We consider that it is appropriate to limit treatment duration as in previous atezolizumab appraisals, as this was based on clinical concerns over possible consequences of longer-term treatment (paragraph 3.13 TA520[6] and 3.11 in TA525[7]). Similar stopping rules have been applied for other immunotherapies, including pembrolizumab for untreated PD-L1 positive metastatic NSCLC (TA531).[12] In TA520, the committee assumed that the effects of second-line atezolizumab monotherapy would last for up to three years after stopping treatment but noted that the length of any continued effect was uncertain. Based on this, we agree with the company's base case cap on survival effect, but we conduct additional scenario analysis to explore decreases as well as increases in the duration of effects (see section 4.4 below).

The original submitted model only allowed pairwise comparisons of Atez+Bev+CP with the two pemetrexed-based comparators (with and without pemetrexed maintenance). In each comparison, the treatment effect cap was implemented by setting the mortality rate for the atezolizumab combination equal to the pemetrexed comparator. This led to different survival predictions for the intervention depending on the comparator (e.g. see CS Tables 46 and 47). This is counter-intuitive and prevents full incremental analysis.

Table 26 Key assumptions in company's base case (adapted from CS Table 45)

Area	Company base case assumption	ERG comment
Time horizon	20 years (from age 63 to 82 years). Sufficient to reflect differences in costs and effects between treatments. The model predicts less than 1% of patients alive at 20 years for intervention and comparators for ITT population.	We agree that the time horizon is reasonable, as the company's base case and most scenarios predict that few patients would survive to 20 years (except for log-normal and log-logistic extrapolations for Atez+Bev+CP without an effect cap).
Cycle length	One week with half-cycle correction	The cycle length is appropriate, and the half-cycle correction is correctly applied
Treatment stopping rule	Maximum of 2 years treatment with Atezo+Bev+CP. Lack of evidence for a stopping rule, but it is applied for consistency with NICE guidance for atezolizumab (TA520 and TA525). Scenario with no stopping rule.	We agree. Stopping rules for atezolizumab and other immunotherapies in NICE TAs were based on clinical concerns about possible consequences of longer-term treatment. Effect in model is to reduce costs for the intervention and hence ICERs
Duration of treatment effect	Effect of Atezo+Bev+CP on OS lasts for 5 years (3 years after maximum treatment). Lack of evidence but conservative approach following assumption in TA520. Scenarios for increased cap on OS effect, up to 20 years. In the revised model, OS effect cap for Atez+Bev+CP applied relative to pemetrexed maintenance comparator (clarification response B4). No cap on duration of survival effect for pemetrexed maintenance.	We agree with the 5-year cap on survival effect given precedent in related appraisals. We test a scenario with reduced duration of effect as well as increase. Implementation of the OS effect cap relative to the with-maintenance pemetrexed comparator is reasonable but is likely to overestimate survival for both Atez+Bev+CP and PEM+CIS with maintenance.

In response to a clarification question (B4), the company submitted a revised economic model in which the mortality rate for the atezolizumab combination was set equal to that of the with-maintenance pemetrexed comparator after five years, while maintaining the extrapolated survival advantage for pemetrexed with maintenance relative to pemetrexed without maintenance. This is not consistent with committee conclusions in TA402[13]: that patients would receive pemetrexed maintenance until progression but that there was not any

evidence for a post-progression survival benefit for pemetrexed maintenance over placebo (paragraph 4.15). The company's revised base case is therefore likely to overestimate long-term survival for both Atez+Bev+CP and the pemetrexed maintenance comparator, and hence to underestimate the ICER for Atez+Bev+CP compared with PEM+CIS without maintenance.

ERG conclusion: The three-state partitioned survival structure of the company's model is appropriate and correctly implemented. The 20-year time horizon is reasonable given model projections of survival.

We agree with company's base case assumption of a two-year maximum treatment duration for atezolizumab and bevacizumab as part of the Atez+Bev+CP intervention. This is consistent with existing guidance for atezolizumab (TA520 and TA525) and for other immunotherapies (e.g. TA531). The assumption that pemetrexed maintenance therapy continues until progression is also appropriate, given committee conclusions in TA402.

The company assumption of a three-year cap on survival effects (after the maximum 2-year treatment) for Atez+Bev+CP is reasonable. This is consistent with committee assumptions for atezolizumab at second line in TA520, although we note the high uncertainty over the persistence of survival effects after treatment is stopped. The company test scenarios with a longer duration of treatment effect (up to 20 years). We also test a scenario with a shorter duration of effect.

However, we do not consider the company's assumption of a persistent survival advantage for pemetrexed maintenance throughout the time horizon to be realistic. This is not consistent with committee conclusions in TA402 and is likely to have overestimated the long-term survival gain for Atez+Bev+CP and for the pemetrexed maintenance comparator. This implies that the ICER for Atez+Bev+CP relative to PEM+CIS without maintenance is likely to be underestimated.

We comment on the sources and assumptions for model input parameters on clinical effectiveness, utilities and resource use and costs in the following sections.

4.2.4 Clinical effectiveness

The model requires four sets of input parameters for clinical effectiveness:

- OS extrapolations for each comparator (CS section B.3.3.2)
- PFS extrapolations for each comparator (CS section B.3.3.3)
- TTD for atezolizumab and for bevacizumab as part of the Atezo+Bev+CP intervention and for pemetrexed maintenance (B.3.3.4)
- AE incidence for each comparator (CS B.3.5.3)

4.2.4.1 Overview of methods for estimating OS and PFS

The company outline their approach to estimating OS and PFS in section B.3.3.1 of the CS. This was a two-step process.

> **Step 1: Extrapolation of PFS and OS curves for Atez+Bev+CP**
> Parametric survival models were fitted to data from the Atezo+Bev+CP arm of the IMpower150 trial (January 2018 data cut, with investigator-assessed progression).
>
> **Step 2: Estimation of PFS and OS curves for comparators**
> PFS and OS curves for the pemetrexed-based comparators were obtained by applying hazard ratios from the NMA to the fitted Atezo+Bev+CP curves.

We discuss general issues related to the methods of extrapolation in this section and give a more detailed description and critique of the company's selection of OS and PFS curves in the following sections, 4.2.4.2 and 4.2.4.3 respectively.

4.2.4.1.1 Methods used to fit baseline curves for atezolizumab

- *Relevance of IMpower150 to UK population* (section 3.1.3 of this report): We consider this a suitable source of data, as the trial population is broadly reflective of patients with metastatic non-squamous NSCLC in routine practice in the UK NHS.
- *ITT versus subgroup analyses* (section 3.1.6.5): The CS reports results with baseline curves fitted to data for the ITT population and the low/negative PD-L1 and positive EGFR/ALK positive subgroups. On balance, we think that the subgroup analyses are a better source for baseline survival estimates than the ITT analysis - as they are specific to the populations of interest and exclude patients with high PD-L1 expssion for whom pembrolizumab would be a more appropriate treatment. However, the ITT analysis should be more robust as the sample is larger: 400 patients randomised to Atez+Bev+CP, of whom 325 had low or negative PD-L1 and 41 were positive for

EGFR or ALK. The EGFR/ALK subgroup in particular is very small and subject to high uncertainty. We also note that although both subgroups were pre-specified, randomisation was only stratified by PD-L1 high/low status. In ERG analysis, we therefore follow the same approach as the company and report results using baseline curves for the ITT population as well as for the separate subgroups (see section 4.4 below).

- *Separate fitting of parametric curves to one trial arm.* The company argues that this is justified because of different mechanisms of action for immunotherapies and chemotherapies and evidence from log-cumulative hazards plots that proportional hazards do not hold in IMpower150 (see section 3.1.7 above). We agree and note that the other arms in the IMpower150 are not comparators in the economic analysis.
- *Choice of parametric function*. The company selects parametric curves for OS and PFS by considering how well they fit the trial data and the plausibility of the projections; see sections 4.2.4.2 and 4.2.4.3 below.

4.2.4.1.2 Methods used to estimate relative treatment effects

See section 3.1.7.4 above for our explanation and critique of the company's NMA analyses. Key issues arising for the economic analysis are:

- *ITT versus subgroup NMA* (section 3.3.6 above). As with the baseline curves, a decision has to be made whether to use the ITT or subgroup versions of the NMA. The cost-effectiveness results for ITT, EGFR/ALK positive and PD-L1 low/negative populations reported in the CS (B.3.7) each use the corresponding subgroup for the NMA as well as for the fitted baseline curves. Expert clinical advice to the ERG suggests that people with EGFR/ALK mutations and those with high PD-L1 expression are more likely to respond to pemetrexed chemotherapy than other patients. However, analysis of the IMpower150 trial did not show any evidence of effect modification for the EGFR/ALK or PD-L1 subgroups. We consider that the ITT NMA is a more robust source for relative treatment effects than the subgroup NMAs, and so we use the ITT NMA in the ERG base case for both subgroups.
- *Inclusion of PARAMOUNT in the network* (see section 3.1.7.2 above). There is heterogeneity in the NMA, primarily associated with the patients in the PARAMOUNT trial, who were more likely to respond to pemetrexed maintenance. This is not surprising due to the different trial design: in PARAMOUNT, only patients who responded to induction treatment (539 out of 900 patients) were randomised. The company make this point, but include PARAMOUNT in their base case, as this was

the only source of evidence for the pemetrexed comparators without maintenance. However, we consider that the difference in study design is too serious as a likely source of bias for the indirect comparison. We therefore use the NMA excluding PARAMOUNT for ERG base case analyses (section 4.4). We report a scenario including PARAMOUNT to enable comparison against pemetrexed and cisplatin without maintenance (this is the only trial to include this comparison). Expert clinical advice to the ERG is that most patients would receive pemetrexed maintenance therapy.

- *Fixed or random effects NMA* (3.1.7.4.3 above). The company use a fixed effect fractional polynomial NMA model in their base case. In principle, a random effects model is preferable in the presence of heterogeneity. We use the fixed effects model in the ERG base case analysis, as this omits PARAMOUNT, which is the main source of heterogeneity, but we use random effects in the scenario analysis that includes PARAMOUNT.
- *Constant or time-varying hazard ratios* (3.1.7.4 above). The company used a time-varying fractional polynomial model in their base case, based on arguments about the different mechanisms (and hence speeds) of action for immunotherapies and chemotherapies, precedent in previous appraisals and evidence from the IMpower150 trial. We agree with these arguments but note that the comparators in IMpower150 are out of scope and that the company has not presented evidence on proportional hazards for other trials in the network. Nevertheless, we agree with the use of an fractional polynomial to allow for change in relative treatment effects over time.
- *Fractional polynomial model selection* (section 3.1.7.4.1). The ERG also agrees with the company's choice of the first order P1=0 (Weibull) model for their base case, with an alternative, exponential (i.e. proportional hazards) model in scenario analysis.

ERG conclusions: The methods used to extrapolate OS and PFS for the economic model are reasonable. This involved fitting baseline parametric survival curves for the Atez+Bev+CP arm to IMpower150 trial data and then applying time-varying hazard ratios from the NMA to estimate survival curves for the comparators.

- We consider that the baseline curves fitted to data for the PD-L1 low/negative and EGFR/ALK positive subgroups are most relevant to the decision problem: the ITT curves are subject to bias due to inclusion of patients with high PD-L1

expression. However, we also consider results with ITT baseline curves, as these are likely to be more robust.

- There is a lack of evidence of effect modification for the subgroups, so we prefer the ITT version of the NMA. We agree with the company's choice of NMA model (first order FP with P1=0), but consider that the analysis including PARAMOUNT is likely to be biased due to the exclusion of patients who did not respond to pemetrexed with platinum induction. For ERG analysis, we prefer the NMA without PARAMOUNT (fixed effects). This restricts results to the comparison including pemetrexed maintenance, but we understand that this is the most common current practice. To enable comparison against with pemetrexed without maintenance we also run a scenario including PARAMOUNT but with random effects.

4.2.4.2 Overall survival extrapolations

We show KM plots and fitted parametric curves for overall survival in the IMpower150 Atezo+Bev+CP arm ITT (n=400), PD-L1 low/negative (n=325) and EGFR positive (n=41) datasets in Figure 13, Figure 14 and Figure 15 below. The predictions for the PD-L1 low/negative subgroup are similar but slightly less favourable than for the ITT population. The prognosis is better for the EGFR positive subgroup, although these predictions are very uncertain due to the small sample size (n=41).

Goodness-of-fit to trial data

Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) statistics for the Atez+Bev+CP parametric curves are reproduced in Table 27, Table 28 and Table 29 for the ITT population, PD-L1 low/negative and EGFR positive subgroups respectively. The company state that the best-fitting function for OS is the Weibull, although all models apart from log-normal have similar AIC and BIC values (CS B.3.3.2). We consider that the Gompertz, exponential and Weibull have the best AIC/BIC statistics and a good visual fit to the KM plots for the ITT population and PD-L1 low/negative subgroup. It is difficult to differentiate the EGFR/ALK positive subgroup curves on the basis visual fit, but the exponential has the best AIC and BIC statistics.

Plausibility of survival projections

Five and ten-year survival estimates from the parametric extrapolations for Atezo+Bev+CP and modelled OS curves for the pemetrexed with platinum comparators (with and without pemetrexed maintenance) are reproduced in Table 27, Table 28 and Table 29 for the ITT, PD-L1 low/negative and EGFR positive groups.

Since immunotherapy has only been available for the last two to three years, there is uncertainty over survival expectations for the intervention arm, although there is a clinical expectation that there may be a small proportion of long-term survivors. The company report estimates of five-year survival with Atez+Bev+CP from 10 UK clinicians of between 12% and 27%, with an average of 17%. If correct, this would imply that the exponential, log-normal or log-logistic OS extrapolations are realistic for the ITT and PD-L1 low/negative subgroup. Modelled five-year survival is much higher for the EGFR/ALK positive subgroup.

With regard to the comparator arms, the company compares the model predictions against five-year survival estimates from the NICE appraisal of pembrolizumab for untreated PD-L1 positive metastatic NSCLC (TA531).[12] The TA531 committee concluded that predictions derived from the control arm of the KEYNOTE-024 rial of 8 to 11% survival at five years were plausible for standard care (chemotherapy or pemetrexed with platinum, with or without pemetrexed maintenance). The company also compare against survival estimates from a cohort of newly diagnosed patients with NSCLC from a large US database (the Flatiron database[14]): reported in CS Appendix M pages 376 to 381. This gave five-year survival estimates of 8.3% for pemetrexed with platinum and 12.3% for pemetrexed with platinum and pemetrexed maintenance, similar to the range cited in TA531.[12] Neither source is directly comparable to the IMpower150 population; KEYNOTE-024 was restricted to patients with high PD-L1 high expression and no sensitizing EGFR mutations or ALK translocations; and the Flatiron cohort included patients with stage IIIb disease as well as stage IV. Nevertheless, the target range of 8-11% with comparator treatments appears reasonable.

None of the modelled estimates for the comparator without maintenance fall within the 8-11% range for the ITT or PD-L1 low/negative subgroup. However, for the comparator with maintenance, the exponential and Weibull baseline extrapolations do. We note that with pemetrexed maintenance, the log-logistic and log-normal extrapolations appear unrealistically optimistic, particularly the 10-year extrapolations. The Gompertz and generalised gamma extrapolations for the comparators appear too pessimistic.

Company choice of OS curves

The company present scenario analysis using all six parametric baseline OS functions in Tables 62-66 of the CS (B.3.8.3). These show that expected life years, and hence QALYs and ICERs, are sensitive to the choice of baseline OS curve. The company concludes that the exponential extrapolation provides “appropriate but still conservative” survival estimates.

They use this in their base case analysis, with the log-logistic as an alternative scenario that they consider plausible.

We consider that the Weibull extrapolation also has a good fit to the trial data and produces five-year survival predictions within the plausible range for the pemetrexed combination with maintenance. We also note that a Weibull or exponential extrapolation for Atez+Bev+CP is consistent with the P1=0 FP (Weibull) estimates of relative effects, producing Weibull curves for the comparator arms.

Impact of five-year cap on treatment effect for Atez+Bev+CP

As discussed in section 4.2.3 above, the company base case includes an assumption that the survival advantage for Atez+Bev+CP over the pemetrexed comparators lasts for a maximum of five years (three years beyond the maximum duration of atezolizumab and bevacizumab treatment). The company illustrates the impact of setting the mortality rate for Atez+Bev+CP equal to that for the pemetrexed comparators without maintenance after five years in CS Figures 34 and 35. They argue that the projections for Atezo+Bev+CP without a cap on survival effect is more in line with long-term expectations for immunotherapies, with a small proportion of patients (about 2%) surviving to 10 years.

The impact of applying the five-year cap on survival effect compared with the pemetrexed combination with maintenance is illustrated in CS Figures 36 and 37. This shows a counter-intuitive reduction in survival for the Atezo+Bev+CP when the effect cap is removed. This results from a declining hazard ratio for Atez+Bev+CP compared with the pemetrexed comparator with maintenance over time in the company's preferred FP NMA model (P1=0) when the PARAMOUNT trial is included (clarification response A18 Figure 1). The company argues that this is likely to be a consequence of bias in the PARAMOUNT trial, which only included patients who responded to pemetrexed with cisplatin induction. We agree, and consider the comparison between Atez+Bev+CP and the pemetrexed with maintenance comparator to be more reliable without PARAMOUNT data.

We show the modelled survival curves under the company's base case analysis for the ITT population over a 10-year period in Figure 16. The effect of excluding the PARAMOUNT trial from the NMA is shown in Figure 17 and scenarios with Weibull and log-logistic parametric survival functions for Atez+Bev+CP in Figure 18 and Figure 19 respectively.

ERG conclusion: The company uses an exponential baseline OS curve for the atezolizumab combination in their base case. This has a good fit to the IMPower150

data and clinically plausible extrapolations of survival at five and ten years. We also consider that the Weibull distribution is plausible, with more conservative survival predictions. The log-logistic gives over optimistic long-term predictions (around 10% survival at 10 years).

Table 27 OS Atezo+Bev+CP: ITT population (five-year effect cap)

Baseline distribution	Goodness-of-fit		Atezo+Bev+CP		Pem+plat		Pem+plat with maintenance	
	AIC (rank)	BIC (rank)	5 year	10 year	5 year	10 year	5 year	10 year
Exponential	942.3(3)	946.3(1)	13%	3%	2%	0%	12%	3%
Weibull	941.7(2)	949.7(3)	10%	1%	1%	0%	9%	1%
Log-logistic	947.2(5)	955.2(5)	20%	12%	5%	1%	18%	10%
Log-normal	958.1(6)	966.0(6)	24%	15%	7%	1%	21%	13%
Gamma	942.8(4)	954.7(4)	4%	0%	0%	0%	5%	0%
Gompertz	940.7(1)	948.7(2)	3%	0%	0%	0%	4%	0%

Reproduced from CS Tables 26 and 27, and model

Table 28 OS Atezo+Bev+CP: PD-L1 low/negative (five-year effect cap)

Baseline Distribution	Goodness-of-fit		Atezo+Bev+CP		Pem+plat		Pem+plat with maintenance	
	AIC (rank)	BIC (rank)	5 year	10 year	5 year	10 year	5 year	10 year
Exponential	763.0(4)	766.8(1)	12%	4%	3%	0%	13%	4%
Weibull	760.8(2)	768.4(3)	7%	1%	1%	0%	10%	2%
Log-logistic	765.5(5)	773.1(4)	18%	11%	6%	1%	19%	12%
Log-normal	776.8(6)	784.3(6)	22%	14%	9%	2%	21%	14%
Gamma	762.1(3)	773.4(5)	3%	0%	0%	0%	6%	0%
Gompertz	760.3(1)	767.8(2)	1%	0%	0%	0%	3%	0%

Reproduced from CS Appendix N Tables 62 and 63, and model

Table 29 OS Atezo+Bev+CP: EGFR/ALK positive (five-year effect cap)

Baseline distribution	Goodness-of-fit		Atezo+Bev+CP		Pem+plat		Pem+plat with maintenance	
	AIC (rank)	BIC (rank)	5 year	10 year	5 year	10 year	5 year	10 year
Exponential	80.3(1)	82.0(1)	27%	17%	11%	3%	18%	12%
Weibull	82.3(4)	85.7(4)	26%	16%	11%	3%	18%	11%
Log-logistic	82.1(3)	85.5(3)	35%	28%	15%	9%	22%	18%
Log-normal	81.8(2)	85.2(2)	39%	33%	18%	11%	25%	21%
Gamma	83.8(6)	88.9(6)	42%	36%	20%	13%	26%	23%
Gompertz	82.3(5)	85.7(5)	27%	17%	11%	3%	18%	12%

Reproduced from CS Appendix N Tables 64 and 65, and model

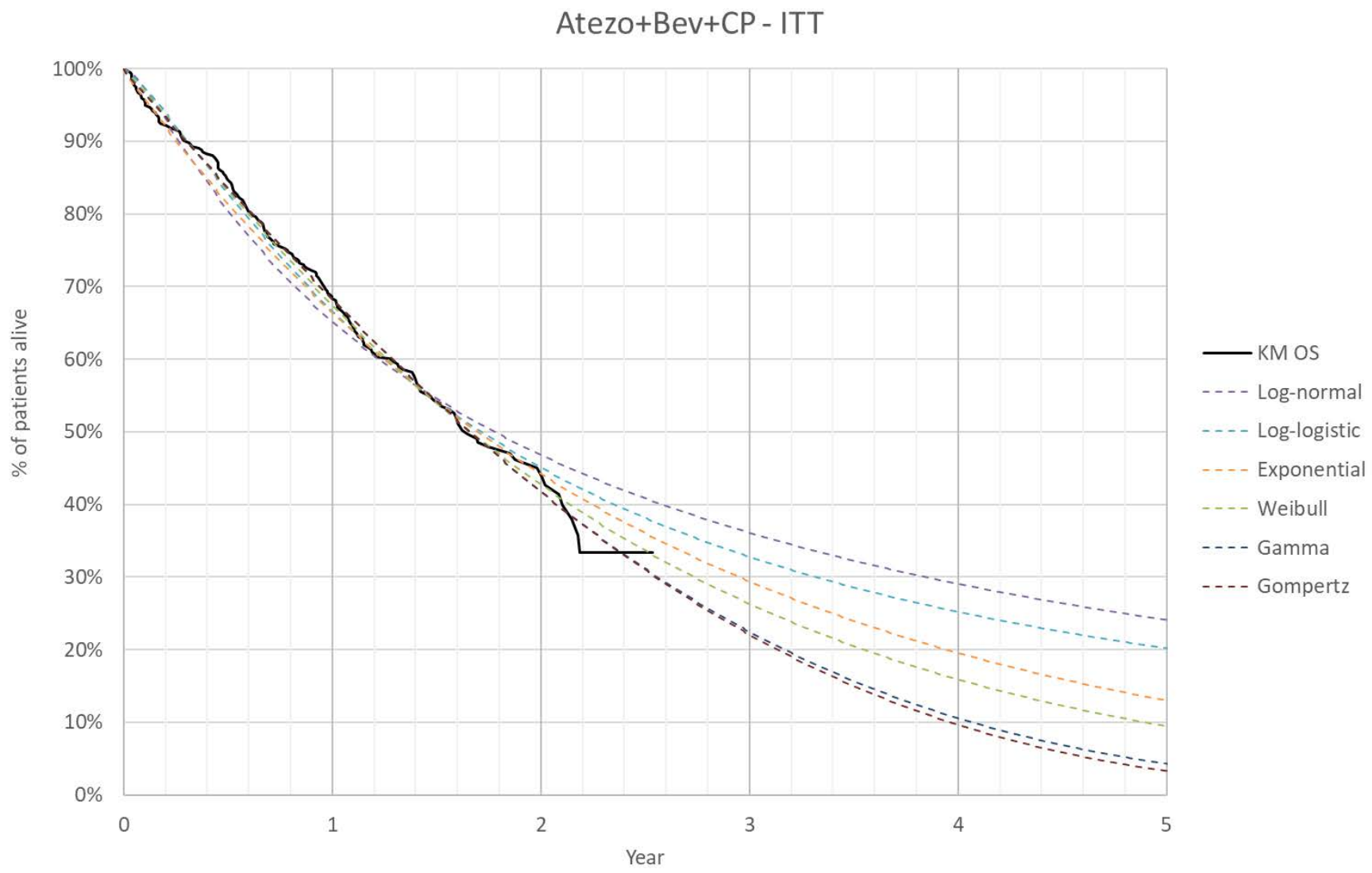


Figure 11 Overall survival curves fitted to IMpower150 Atezo+Bev+CP arm: ITT population

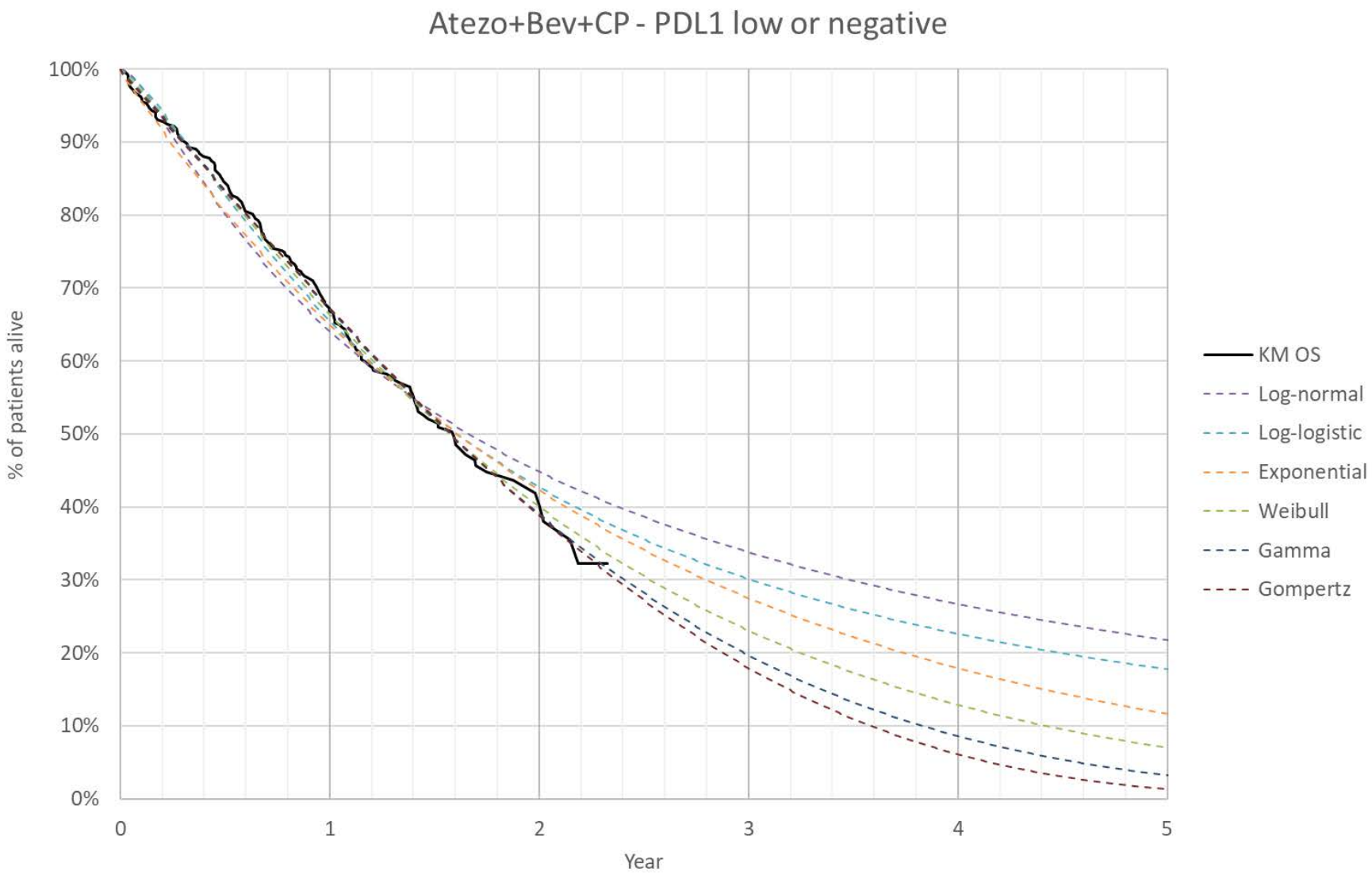


Figure 12 Overall survival curves fitted to IMpower150 Atezo+Bev+CP arm: PD-L1 low or negative subgroup

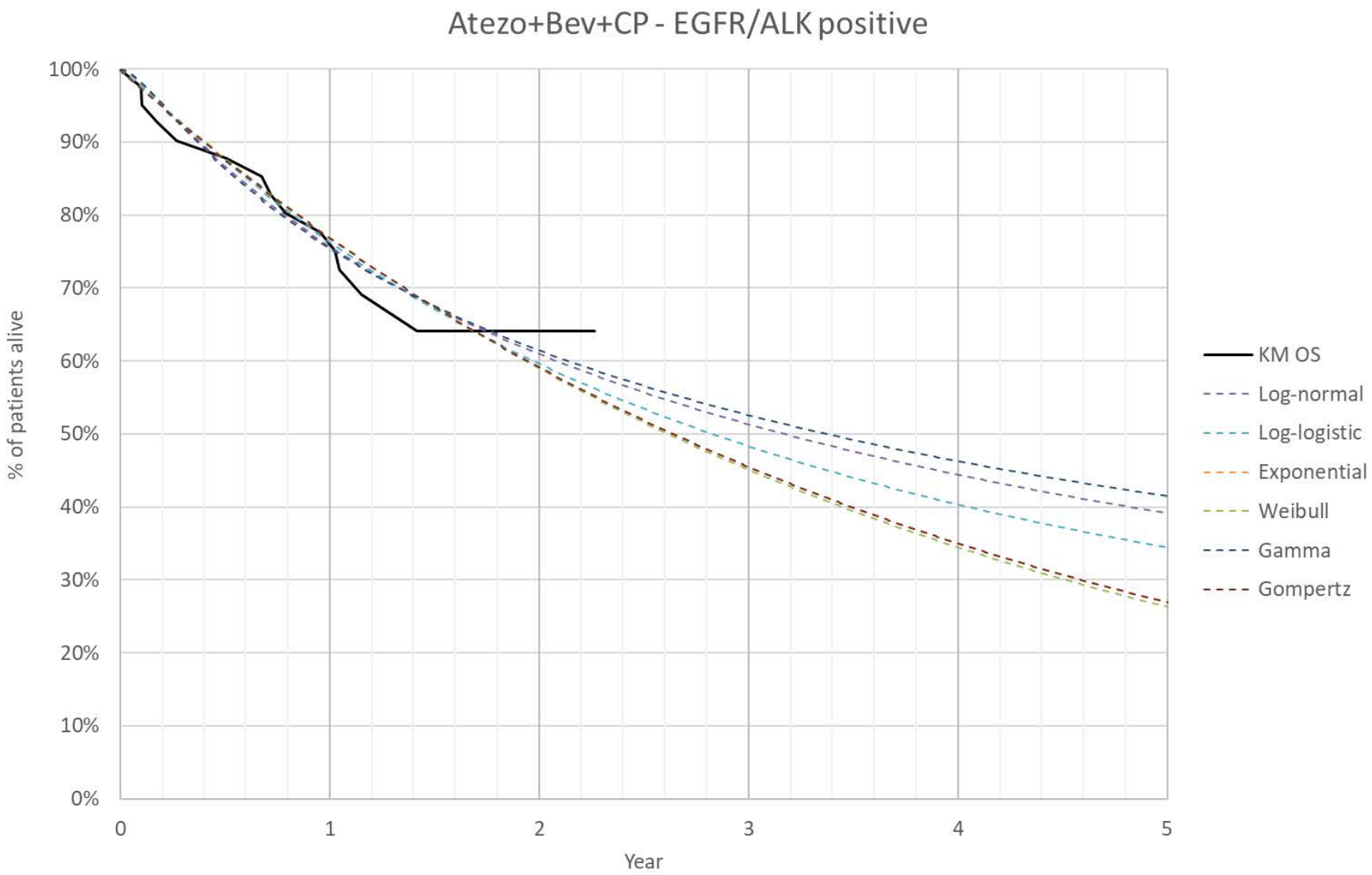


Figure 13 Overall survival curves fitted to IMpower150 Atezo+Bev+CP arm: PD-L1 low or negative subgroup

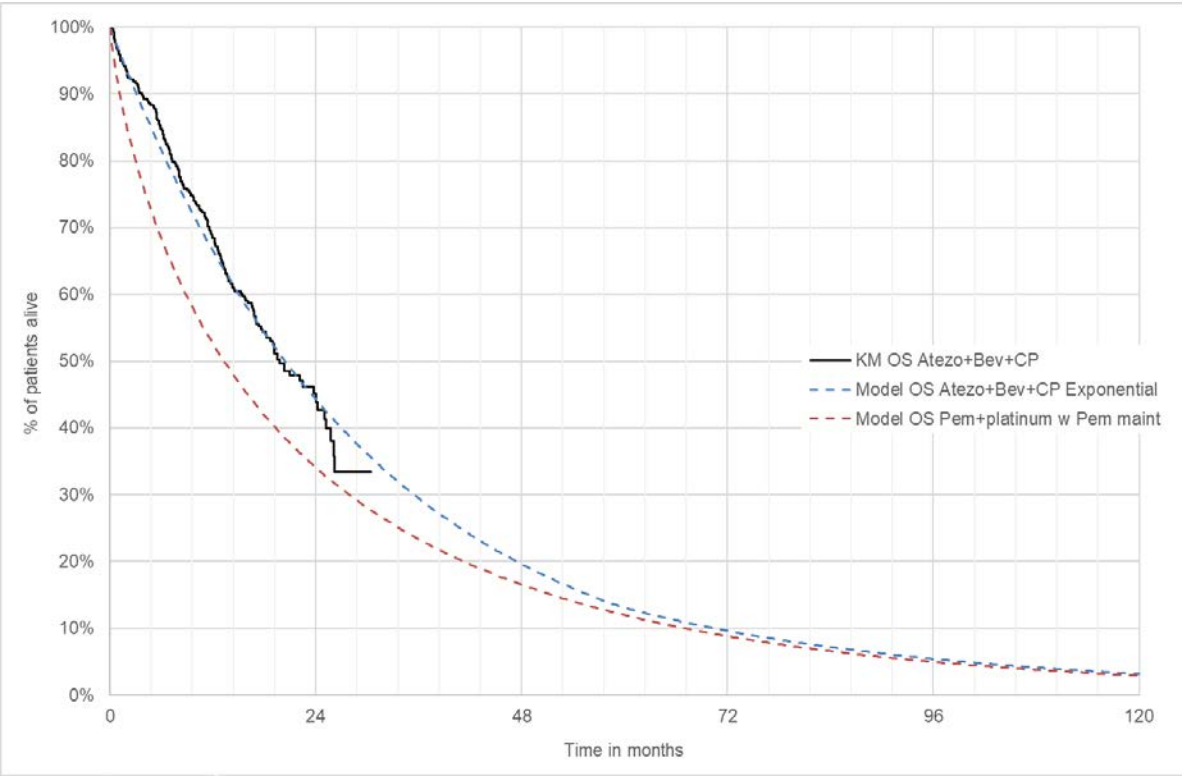


Figure 14 OS company base case ITT (exponential, five-year effect cap)

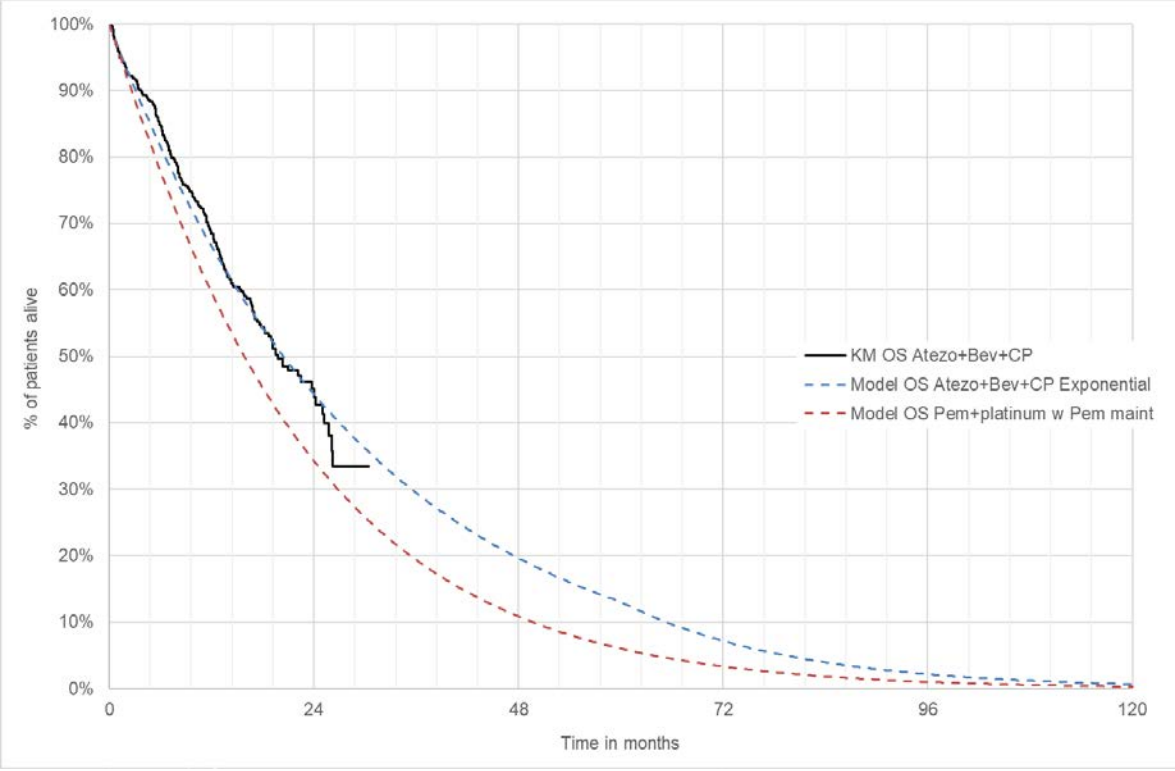


Figure 15 OS company base case ITT, NMA without PARAMOUNT trial

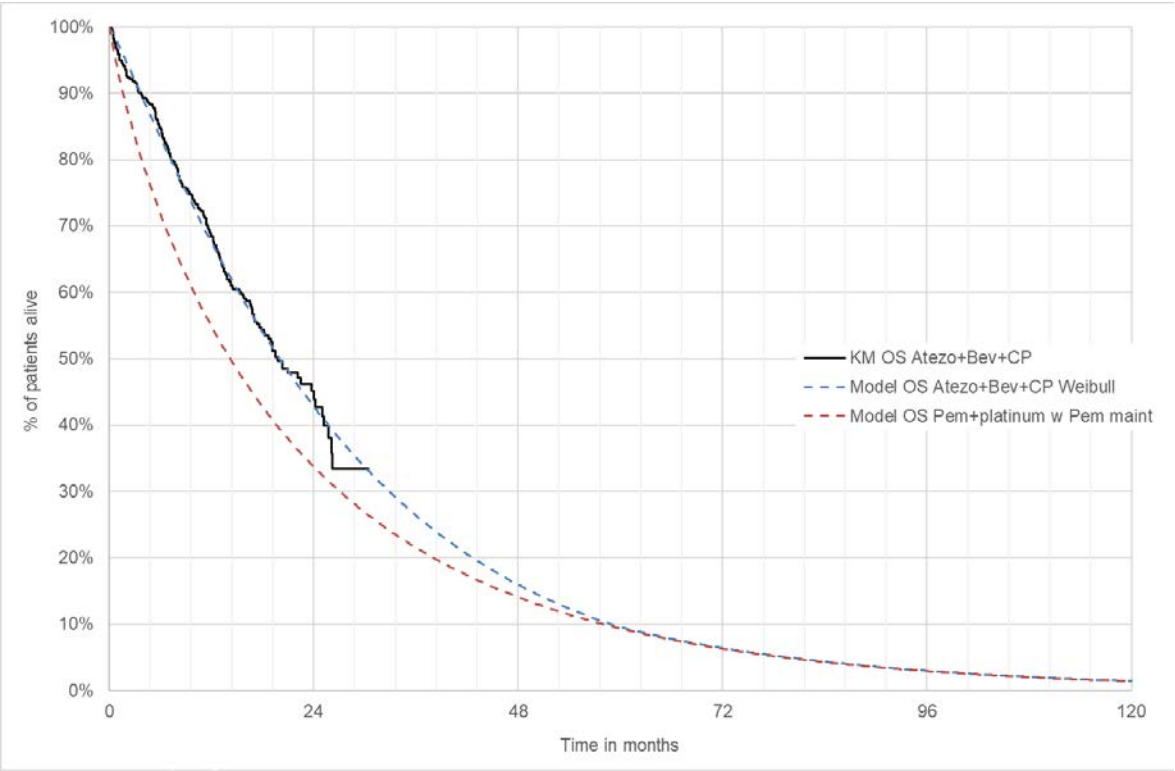


Figure 16 OS company base case ITT, with Weibull survival function

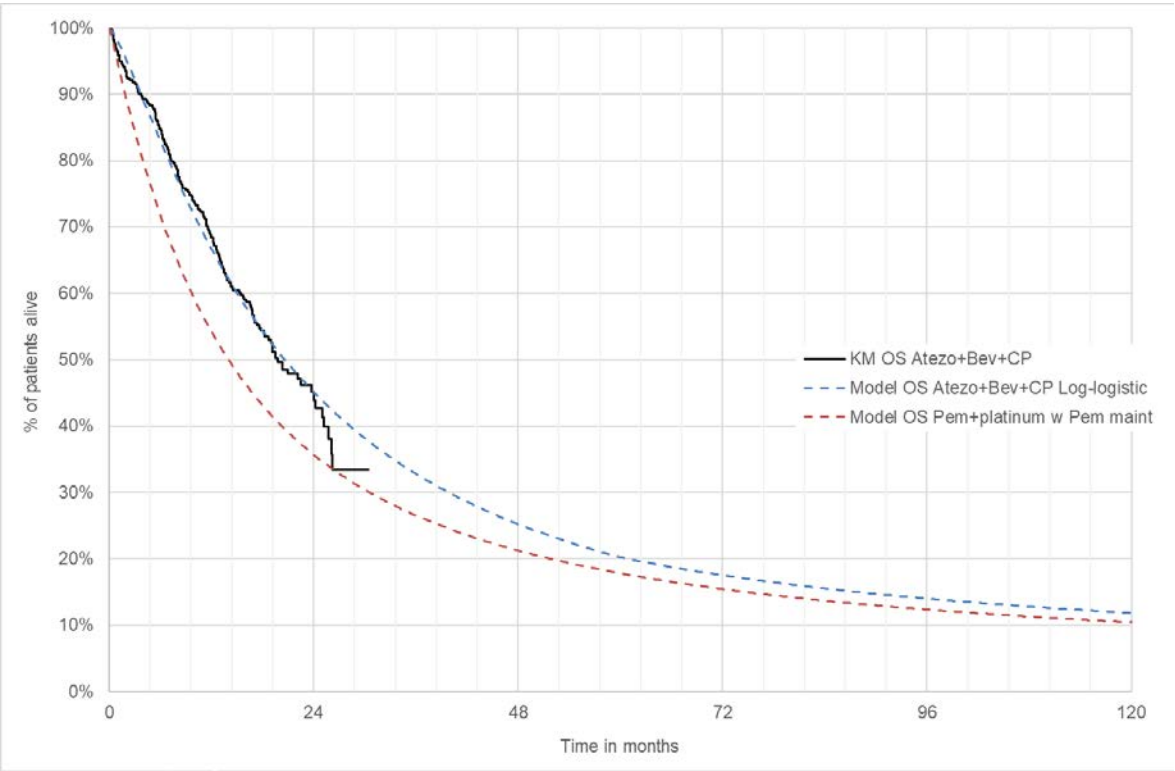


Figure 17 OS company base case ITT, with log-logistic survival function

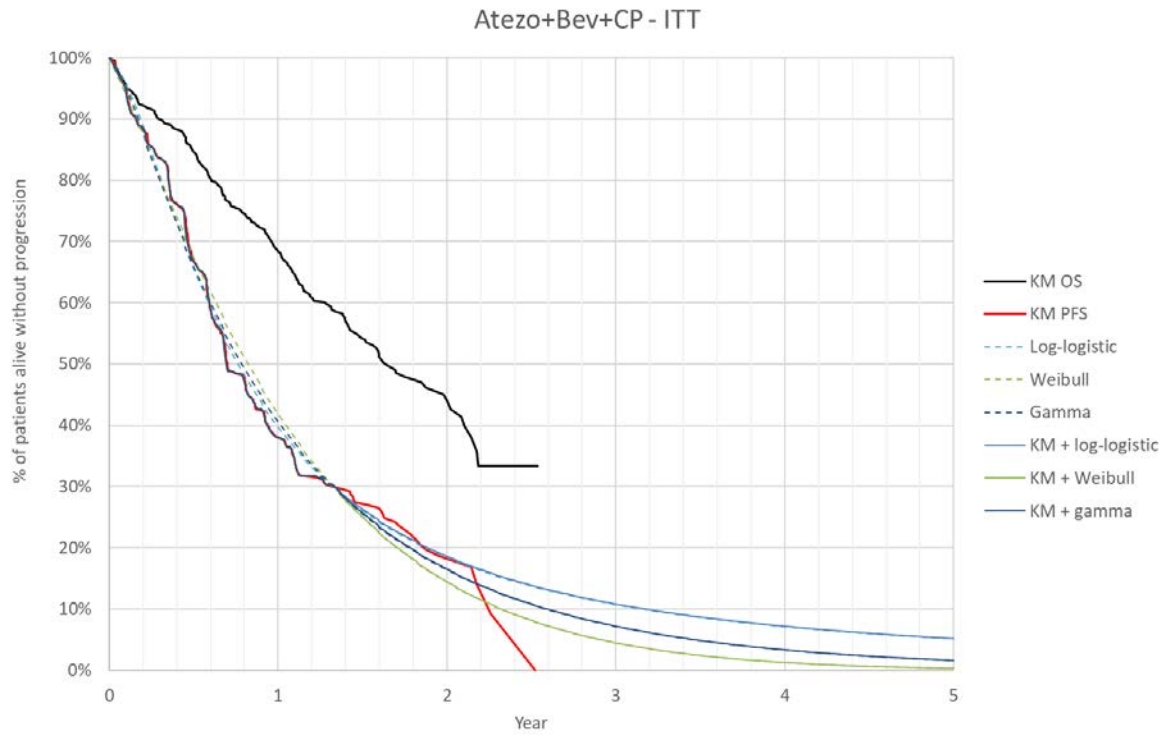
4.2.4.3 Progression free survival extrapolations

The same approach as for OS was used to fit PFS curves for Atez+Bev+CP and to estimate comparator curves based on time-varying relative effects. PFS data for the IMpower150 trial are relatively complete and cost-effectiveness results are less sensitive to different methods of extrapolating PFS (CS Tables 62 to 65 B.3.8.3).

The company report goodness-of-fit statistics for parametric Atez+Bev+CP PFS distributions in CS Table 28 for the ITT population and Tables 66 and 67 in Appendix O for the PD-L1 and EGFR/ALK subgroups. For the ITT and PD-L1 populations AIC and BIC statistics and visual fit are best for the log-logistic distribution, followed by the Weibull and generalised gamma. These three distributions provide a spread of projections, from about 2% to 5% of patients still alive and free of progression after five years.

For their ITT base case, the company use the KM curve with a log-logistic extrapolation from the point where 20% of patients remain at risk (n=81 at about 15 months). This is reasonable as the KM data are mature with a sufficient sample size, and the extrapolation from the KM is very similar to the fully parametric extrapolation (see Figure 20 below). The curves for the PD-L1 low/negative subgroup are similar but slightly less favourable. For the EGFR/ALK positive subgroup, the log-normal, exponential or log-logistic curves have the best statistical and visual fit to the KM data (see Figure 21). The company chooses a fully parametric log-normal distribution for PFS in the EGFR/ALK positive subgroup, and we consider log-normal, exponential and Weibull distributions to show a range of uncertainty around the extrapolation.

> **ERG conclusion**: The company's approach to extrapolating PFS is reasonable. A similar method was used as for OS, but the model results are much less sensitive to PFS than OS. For the atezolizumab arm, the company use the KM curve with a log-logistic extrapolation for the ITT population and PD-L1 low/negative subgroup. In the EGFR/ALK positive subgroup they used a fully-parametric log-normal distribution. We consider scenarios with exponential and Weibull extrapolations in ERG analysis.



.

Figure 18 PFS curves fitted to IMpower150 Atez+Bev+CP data: ITT

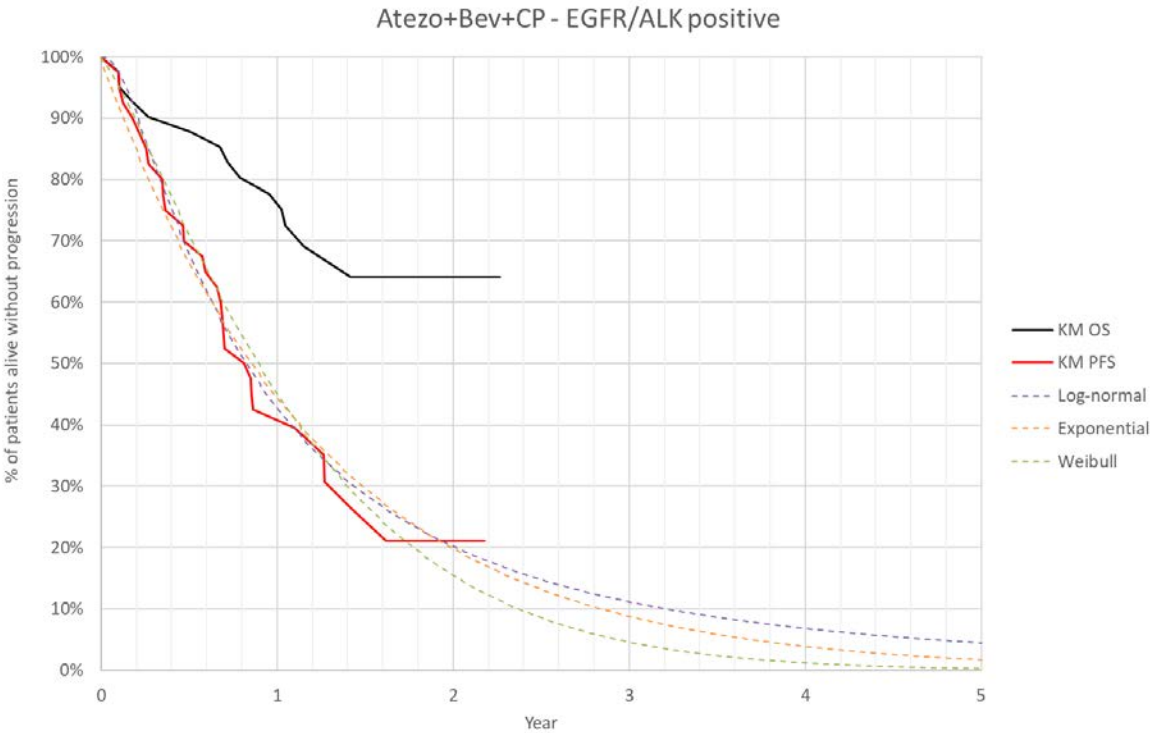


Figure 19 PFS curves fitted to IMpower150 Atez+Bev+CP data: EGFR/ALK positive

4.2.4.4 Treatment duration

TTD curves for atezolizumab and for bevacizumab (separately) were estimated by fitting parametric curves to data from the Atezo+Bev+CP arm of the IMpower150 trial. In the trial, approximately 20% of patients were still being treated with atezolizumab and 10% with bevacizumab after two years. Although the company extrapolates the TTD curves, this has little impact in the base case model due to the use of a two-year stopping rule: see section 4.2.3 above for discussion of the stopping rule and associated assumption about the duration of treatment effects.

Atezolizumab was used until loss of clinical benefit, or unmanageable toxicity. The company states that PFS is not a good surrogate for the duration of treatment with atezolizumab. However, comparison of the KM plots shows that patients tended to stop treatment before progression in the early part of the trial, with similar rates of treatment and progression free survival after about 9 months – see Figure 22. We agree with the company that the exponential curve provides the best visual fit to the KM plot for atezolizumab treatment duration. The company use a KM curve, extrapolated with an exponential curve from the point where 20% of patients remain at risk. This reasonable based on a good visual fit to the trial data.

For bevacizumab, the company notes that although the trial protocol specified that it should be administered in the Atezo+Bev+CP arm until disease progression or unacceptable toxicity, the PFS curve was not a good surrogate for bevacizumab treatment duration. This is supported Figure 23, which shows that progression free survival exceeded bevacizumab treatment duration throughout the trial.

For pemetrexed maintenance, TTD was assumed equal to PFS. This is consistent with committee conclusions in TA181.[13]

> **ERG conclusion**: The ERG agrees with the company's approach to modelling the duration of treatments with atezolizumab and bevacizumab in the Atezo+Bev+CP intervention and of pemetrexed maintenance.

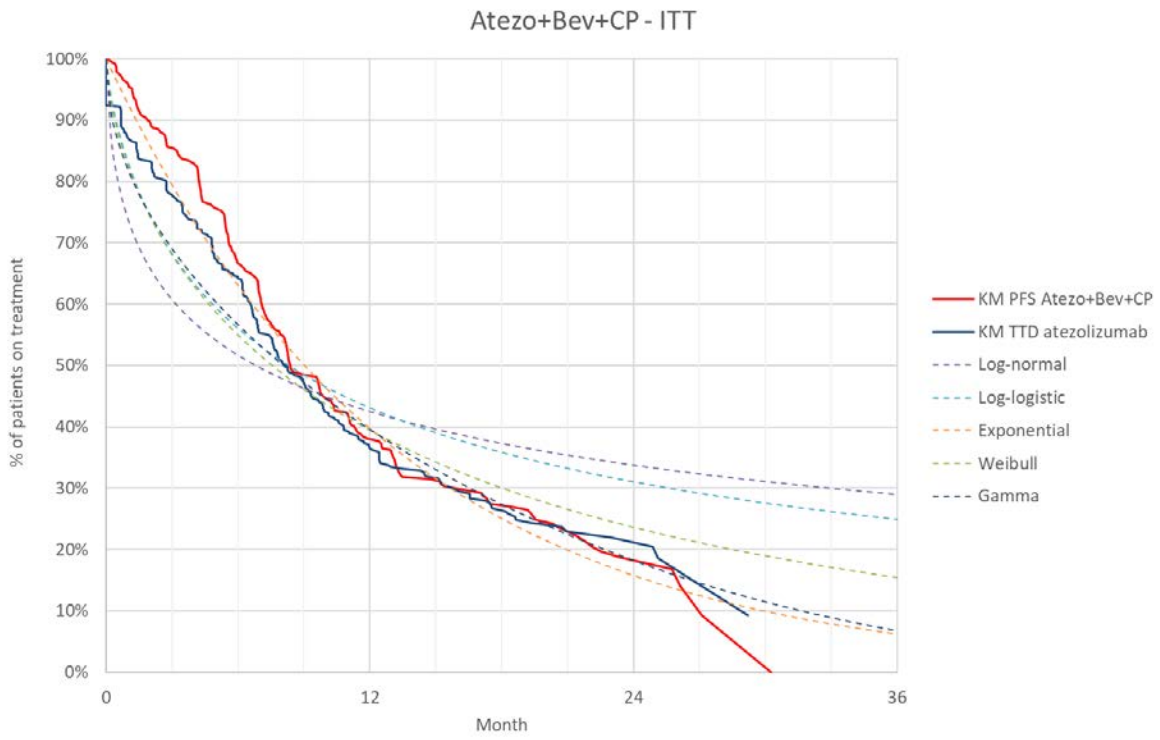


Figure 20 Duration of atezolizumab in Atez+Bev+CP arm of IMpower150: ITT

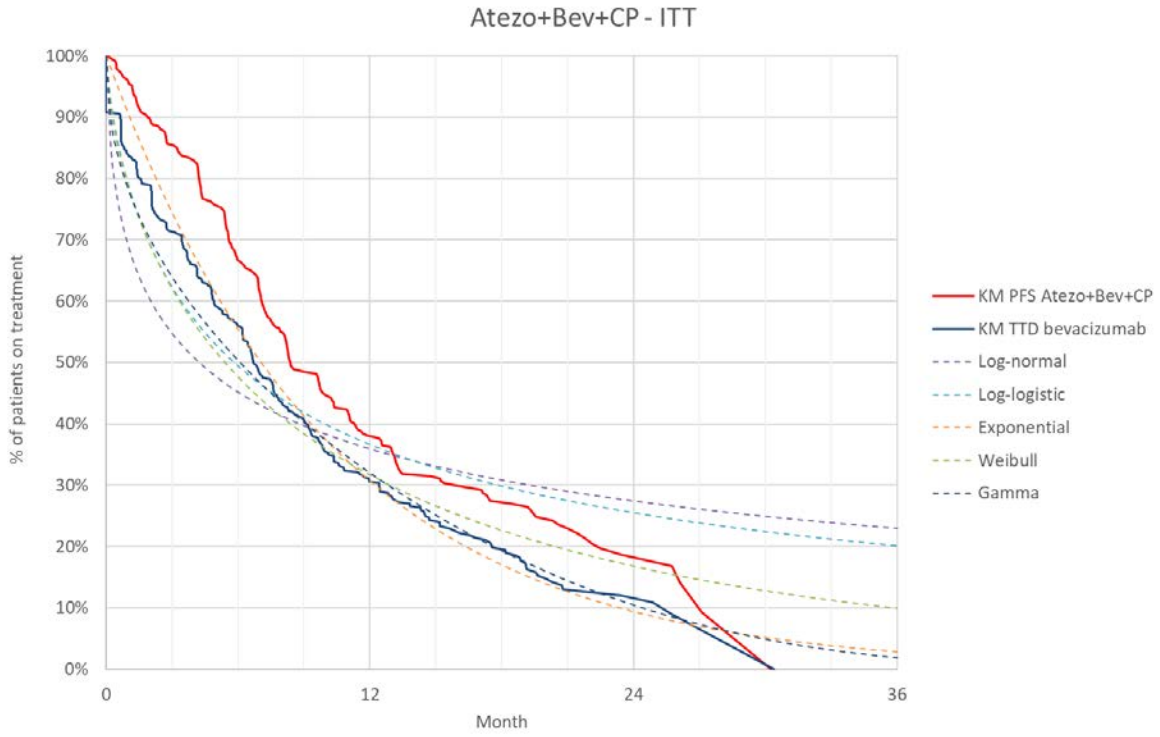


Figure 21 Duration of bevacizumab in Atez+Bev+CP arm of IMpower150: ITT

4.2.4.5 Incidence of adverse events

The base case model includes costs, but not utility loss associated with adverse events. Grade 3+ adverse events with an incidence of in 2% or more in the Atezo+Bev+CP arm of the IMPower150 trial were included in the base case analysis (see CS Table 43). The incidences of the included adverse events for the comparators were sourced from a systematic literature review. The probability of the adverse events per week calculated based on an estimated number of person weeks of follow up in the related trials (clarification question B6).

4.2.5 Health related quality of life

4.2.5.1 Company review of health-related quality of life studies

The company conducted a systematic review to identify evidence in the first-line treatment of patients with non-squamous NSCLC. The original review was completed in September 2016 with an update in February 2018. The search strategy and the inclusion criteria used in the review are detailed in Appendix H of the CS. The review identified 43 publications reporting health state utility values (HSUV) associated with first-line treatment for advanced or metastatic NSCLC, of which five reported the HSUV as graphs only and 21 presented as conference abstracts.

The company reported those studies (n=5) of most relevance to NICE, i.e. those in line with the NICE reference case where utilities were derived directly from patients using EQ-5D with the UK tariff (CS Table 71, Appendix H). Based on the company's review, the CS states that the most suitable studies that were included in the model as scenario analyses were Nafees et al.[15] and Chouaid et al[16] and that these two studies have been used in most of the economic evaluations published in NSCLC. The utility values for both these studies are shown in Table 30.

The ERG considers the company's review to be up-to date and comprehensive and we have not identified any other relevant studies for NSCLC first line treatment. We note that the study by Nafees et al.[15] was not included within the company's list of five most relevant studies as it is for patients receiving second-line treatment. Further, this study does not adhere to the NICE reference case as participants are not patients with the disease. The study has been criticised in previous appraisals for having unrealistically low utility values for patients with progressed disease. Therefore, the ERG suggests that scenario analyses using utility values from the Nafees et al study may be of limited value.

4.2.5.2 Measurement of HRQoL from the IMPower150 trial

The company used utility values in their base case analyses from the utility data collected in the IMPower150 trial. Patients in the study completed the EQ-5D-3L questionnaire and utility values were derived using the UK tariff. EQ-5D data were collected at each scheduled study visit and during survival follow-up at three and six months following disease progression (or loss of clinical benefit).

The company considered two approaches for estimating utility values: 1) proximity to death approach, ii) pre-and-post progression approach. The company used the proximity to death approach and the pre- and post-progression approach was used in scenario analyses. The utility values for both approaches are shown in Table 30. The company justifies the proximity to death approach by stating that this reflects the known decline in cancer patients' quality of life and also that this approach has been used in previous NICE appraisals in NSCLC (TA402,[13] TA428,[17] TA525[6]).

The proximity to death utilities were derived from analyses according to the time before death:

- Group 1: less than 35 days before death
- Group 2: more than 34 and less than 75 days
- Group 3: more than 74 and less than 210 days
- Group 4: more than 211 days

The analysis was based upon HRQoL estimated from those who had died at the time of clinical cut-off (52.2% of patients) for groups 1-3. Group 4 also included those patients still alive with more than 211 days follow-up. The company stated that they fitted a model to include time before death group, assessment time and treatment arm as covariates. The company considered two separate models according to treatment status: on or off treatment. However, the off treatment utilities had wide confidence intervals which overlap for different time to death groups. The company therefore decided to fit and report only utilities by time before death group according to the proximity to death.

Table 30 Summary of utility values for cost-effectiveness analysis

Category	Utility	95% CI	Reference in submission	Justification
IMpower150 utilities - Proximity to death approach – Base case				
≤ 5 weeks before death	0.52	0.49 - 0.56	Section B.3.4.1	Derived from EQ-5D data collected during IMpower150 trial. Methodology as per NICE reference case.
> 5 & ≤ 11 weeks before death	0.59	0.56 - 0.61		
> 15 & ≤ 30 weeks before death	0.70	0.68 - 0.71		
> 30 weeks before death	0.73	0.72 - 0.75		
IMpower150 utilities - Pre- and post-progression - Scenario analysis				
Pre-progression	0.71	0.70 - 0.72	Section B.3.4.1	Derived from EQ-5D data collected during IMpower150 trial.
Post-progression	0.69	0.66 - 0.72		
Pembrolizumab utilities - Proximity to death approach – US publication[18] - Scenario analysis				
≤ 5 weeks before death	0.537	0.425–0.650	Section B.3.4.3	Identified from published literature
> 5 & ≤ 15 weeks before death	0.632	0.592–0.672		
> 15 & ≤ 30 weeks before death	0.726	0.684–0.767		
> 30 weeks before death	0.805	0.767–0.843		
Utilities from Nafees et al – Scenario analysis				
Progression free	0.66*	Calculated based on utility model coefficients	Section B.3.4.3	Identified from published literature
Progressed disease	0.47*			
Utilities from Chouaid et al – Scenario analysis				
Category	**Utility**	**95% CI**	**Reference in submission**	**Justification**
Progression free	0.71*	Calculated based on utility model coefficients	Section B.3.4.3	Identified from published literature
Progressed disease	0.67*			

Table reproduced from CS Table 30
*calculated based on reported regression coefficients; CI: confidence interval

For the pre- and post-progression approach, the company fit a fixed-effects model with covariates for the pre-progression and post-progression periods. They also tested the effects of covariates for treatment arm and adverse events. The company reported that they found no difference between treatment arms and therefore pooled pre-progression mean utility, regardless of treatment arm. The company has provided a scenario analysis using utility values for the pre-progression and post-progression health states (CS Table 62-Table 65). The ERG notes that the results from these scenario analyses have only a small effect on the ICER.

The ERG requested further clarification of the utility values, specifically about the repeated-measures analysis of EQ-5D (clarification question B3). The company provided information on the utility values collected in IMPower150 in Appendix C of their clarification response. These included figures showing how the utilities varied over time for each treatment arm by time before death (Appendix C, Figure 5), before progression (Appendix C, Figure 7), before progression for patients without AEs (Appendix C, Figure 8), and after progression (Appendix C, Figure 9). These figures show that, generally the utility for patients treated with Atez+Bev+CP is worse than for patients treated with Atez+CP but is similar to those treated with Bev+CP. However, as the confidence intervals overlap, the utility for the treatment arms are not statistically different. It is unclear how the health state utility values differ for patients treated with for Atez + Bev + CP differ from the comparators in the economic model (pemetrexed + C; pemetrexed + C + pemetrexed maintenance). The ERG notes that the health state utility values from the PARAMOUNT trial for patients treated with PEM+CIS/CARB + pemetrexed maintenance are [redacted] than those from IMPower150 treated with Atez+Bev+CP (PARAMOUNT: pre-progression 0.77 before randomisation for maintenance, 0.7841 during maintenance phase vs. IMPower150 pre-progression [redacted]). However, as the utility values were taken from different trials in different populations, it is unclear how meaningful these differences are.

The ERG notes that the company has not included any disutility for patients whilst on-treatment or included disutilities for adverse events. The company justified not including the disutility for adverse events because “any disutility has already been incorporated in to the base case health state utilities through the trial derived EQ-5D utilities, and incorporating an additional disutility could be considered double counting”. However, the company included disutilites for adverse events in a scenario analysis. The CS includes details of how the adverse event disutilites have been calculated in Appendix Q. The ERG has concerns with assuming that the utilities would be the same for treatment with Atezo+Bev+CP and pemetrexed plus platinum whilst the adverse

event profile for Atezo+Bev+CP is significantly worse than pemetrexed plus platinum (CS Table 43).

We also note that the scenario for including disutility for adverse events has not been conducted correctly as the same disutility has been used for all treatment arms, whereas in CS Table 70, Appendix Q, the adverse event disutility is lower for pemetrexed plus platinum than for Atezo+Bev+CP. The ERG provides an ERG analysis correcting the adverse event utilities used in section 4.4.

> **ERG conclusion:** The company's approach to estimating health state utility values is reasonable and consistent with previous NICE technology appraisals. The use of IMPower150 utility data is preferable to other estimates of utility in this population. However, the ERG considers that the differences in treatment related adverse events between treatments have not been fully captured and it is unclear whether patients treated with Atezo+Bev+CP have the same health state utility values whilst on treatment as those treated with pemetrexed + platinum (with or without pemetrexed maintenance).

4.2.6 Resource use and costs

The economic model included the following costs:

- Drug acquisition
- Drug administration
- Subsequent treatment
- Follow up monitoring and care
- Terminal care
- Adverse events

4.2.6.1 Drug acquisition

The company's base uses the list prices for all drugs, as shown in Table 31 below (CS Table 31) and the dosing schedule in Table 32 (CS Table 32). Atezolizumab and bevacizumab, carboplatin and paclitaxel are administered by intravenous infusion every three weeks. Atezolizumab and bevacizumab are administered until unacceptable toxicity or loss of clinical benefit. Carboplatin and paclitaxel are administered for four or six cycles. Pemetrexed and cisplatin are administered by intravenous infusion every three weeks for up to six cycles.

Drug costs are taken from British National Formulary[19] and eMIT[20] and the dosing schedules are taken from the IMPower150 trial[21] and the drug's summary of product characteristics. Atezolizumab and bevacizumab have an agreed confidential patient access scheme (PAS) discount (CS Table 52) and the company provides results using both the list price and the PAS price for atezolizumab and bevacizumab. There are also confidential discounts to the NHS for pemetrexed maintenance and pembrolizumab and the ERG provides results including all existing confidential discounts in a separate confidential appendix to this report.

Table 31 Drug acquisition costs

Drug	Vial/pack concentration	Vial/pack volume	Dose per vial/pack	Cost per vial/pack	Cost per mg	Source
Atezolizumab	60 mg/ml	20 ml	1200 mg	£3807.69	£3.17	BNF
Bevacizumab	25 mg/ml	4 ml	100 mg	£242.66	£2.43	BNF
Bevacizumab	25 mg/ml	16 ml	400 mg	£924.40	£2.31	BNF
Pemetrexed	100 mg powder			£160	£1.60	BNF
Pemetrexed	500 mg powder			£800	£1.60	BNF
Carboplatin	10 mg/ml	15 ml	150 mg	£6.35	£0.04	eMIT
Cisplatin	1 mg/ml	100 ml	100 mg	£10.13	£0.10	eMIT
Paclitaxel	6 mg/ml	16.7 ml	100 mg	£9.85	£0.10	eMIT

Table reproduced from CS Table 31
eMIT: 12 month period until end June 2017

The total drugs cost per combination per cycle is £6,445.89 for Atezo+Bev+CP and for pemetrexed plus cisplatin is £1471.61.

The CS states that the base case analysis assumes full vial sharing (i.e. no wastage) for the administration of all weight based therapies. The ERG notes that the model assumes that 5% of patients share vials for these treatments. The CS includes a scenario analysis where there is no vial sharing. The ERG's preference is to have no vial sharing as the base case analysis, and we remove this assumption in our corrections to the company base case (see section 4.4).

Table 32 Dosing schedule and dose per administration

Drug	Dosing per administration	Frequency of administration	Total dose	Reference for dosing
Atezolizumab	1200 mg fixed	Q3W	1200 mg	SmPC IMpower150
Bevacizumab	15 mg/kg	Q3W	1079 mg	IMpower150
Pemetrexed	500 mg/m²	Q3W	905 mg	SmPC
Carboplatin	6 mg/mL/min (AUC)	Q4W	692 mg	SmPC, IMpower150
Cisplatin	75 mg/m²	Q3W	136 mg	SmPC
Paclitaxel	200 mg/m²	Q3W	362 mg	SmPC, IMpower150
Docetaxel	75 mg/m²	Q3W	136 mg	SmPC
Nivolumab	3mg/kg	Q2W	216 mg	SmPC

Table reproduced from CS Table 32

Q3W, every three weeks; Q4W, every four weeks; AUC, area under the curve

4.2.6.2 Drug administration costs

The drug administration costs used in the economic model are shown in CS Table 38. Costs are taken from NHS reference costs 2016-17.[22] The company assumes that the administration cost for Atezo+Bev+CP (Day case cost £385.99) is higher than that used for pemetrexed + platinum (outpatient / day case cost £327.92) and pemetrexed maintenance (outpatient cost £173.99 due to the longer infusion times. The administration cost for subsequent therapies are £173.99 (outpatient cost) for the single therapies of docetaxel, atezolizumab, pembrolizumab and nivolumab. The company has based the administration costs on those used in previous NICE technology appraisal.[2 12 13]

4.2.6.3 Subsequent therapies

The company's economic model includes subsequent lines of therapy for patients with progressed disease. The company assumes that all patients treated with Atezo+Bev+CP are subsequently treated with docetaxel and patients initially treated with pemetrexed are subsequently treated with an immunotherapy or docetaxel as shown in Table 33 (CS Table 34). The CS justifies this approach by stating that it is in line with UK clinical practice and with the second-line marketing authorisation of immunotherapies and has previously been accepted by

the NICE committee in the NICE technology appraisal TA531 for pembrolizumab in first-line NSCLC.[12]

We conduct a scenario analysis excluding nivolumab as a second-line treatment, as this is currently recommended by NICE for use on the Cancer Drugs Fund rather than as part of routine commissioning (TA484).

The IMPower150 trial collected data on subsequent therapies for patients initially receiving Atezo+Bev+CP, however these data are not used in the company base case because these were not in line with current UK practice. The company provides a scenario analysis using these data for subsequent therapies from IMPower150.

The drug acquisition costs for the subsequent therapies are shown in Table 34 (CS Table 36). The ERG notes that the cost for pembrolizumab has been calculated based on patient weight assuming it is possible to buy part of a vial. However, this differs from the approach taken in the NICE technology appraisal TA428[17] for pembrolizumab therapy after chemotherapy for NSCLC. In that NICE appraisal, the company estimated the cost per patient receiving pembrolizumab, based on the KEYNOTE-010 trial where the average number of full 50mg vials received was 3.39 per patient, with a cost per treatment cycle of £4,453.13. The ERG suggests that this cost for pembrolizumab is more appropriate.

Table 33 Subsequent therapies after discontinuation - used in base case analysis

Post-discontinuation therapy	Treatments after Atezo+Bev+CP	Treatments after pemetrexed-based regimens	Duration of therapy (weeks)	Source for duration of therapy
Docetaxel	100%	15%	13.1[1]	Docetaxel SmPC
Nivolumab	0%	34%	26.52	NICE TA484
Pembrolizumab *	0%	34%	21.59	NICE TA428
Atezolizumab	0%	17%	35.80	NICE TA520

Table reproduced from CS Table 34

* Pembrolizumab is administered in second-line as per its license in this indication i.e. 2 mg/kg

[1] Value used in the model differs from that reported in CS Table 34

Table 34 Drug acquisition costs – subsequent therapies

Drug	Vial/pack concentration	Vial/pack volume	Dose per vial/pack	Cost per vial/pack	Total cost per treatment cycle[1]	Source
Atezolizumab	60 mg/ml	20 ml	1200 mg	£3807.69	£3807.69	BNF
Pembrolizumab	25 mg/ml	4 ml	100 mg	£2630.00	£3781.28	BNF
Pembrolizumab	Powder for concentrate for IV solution		50 mg	£1315.00	£3781.28	BNF
Docetaxel	20 mg/ml	7 ml	140 mg	£20.62	£20.02	eMIT
Docetaxel	20 mg/ml	1 ml	20 mg	£3.85	£20.02	eMIT
Nivolumab	10 mg/ml	4 ml	40 mg	£439.00	£2634.00	BNF

Table reproduced from CS Table 35

eMIT: 12 month period until end June 2017

[1] Values taken from company economic model

4.2.6.4 Follow up monitoring and care

The CS presents the resources used for patients with progression free and progressed disease in CS Table 40 (Table 35). The resource use was consistent with that used for the NICE technology appraisal TA531 for pembrolizumab for NSCLC[17] and the economic evaluation by Brown et al [4] on chemotherapy for NSCLC. The resources used are from the Big Lung trial[23] and a Marie Curie report,[24] which were published in 2005 and 2004 respectively. The Big Lung trial reports on a trial completed in 1999/2000. The ERG is unable to find the values reported in Brown et al[4] in the cited sources. Expert clinical advice to the ERG suggests that it may be counter-intuitive that the number of outpatient visits would be higher in the PFS state than in the progressed disease state. Furthermore, we consider that the resource use data may be out of date as they are from older studies and there have been considerable changes to the management of NSCLC since these studies were conducted. The ERG considers a better approach would have been to collect resource use data from the IMPower150 and use these data in the economic evaluation.

The unit costs for the resources used are shown in Table 36 (CS Table 41). These unit costs have been taken from NHS reference costs 2016/17,[22] PSSRU 2017[25] or from previous NICE technology appraisals and the costs have been inflated to 2016/17 using the PSSRU HCHS index.[25]

Table 35 Resource use for PFS and PD health state

Resource	PFS	PD	Unit	Source
Outpatient visit	9.61	7.91	per annum	NICE TA531
Chest Radiography	6.79	6.5	per annum	NICE TA531
CT scan (chest)	0.62	0.24	per annum	NICE TA531
CT scan (other)	0.36	0.42	per annum	NICE TA531
ECG	1.04	0.88	per annum	NICE TA531
Community nurse visit	8.7	8.7	visits (20 minutes) per patient	Appendix 1 of NICE Guideline CG121 Marie Curie report
Clinical nurse specialist	12	12	hours contact time per patient	Appendix 1 of NICE Guideline CG121
GP surgery	12	0	consultations per patient	Appendix 1 of NICE Guideline CG121
GP home visit	0	26.09	per annum (fortnightly)	Marie Curie report
Therapist visit	0	26.09	per annum (fortnightly)	Appendix 1 of NICE Guideline CG121

Table reproduced from CS Table 40
PFS, progression free state; PD, progressed disease state; GP, general practitioner; CT, computerised tomography; ECG, electrocardiogram; NICE, The National Institute for Health and Care Excellence; CG, clinical guidance

The ERG notes that the company has updated costs incorrectly using the HCHS prices index, rather than using the HCHS pay and prices index. The corrected costs are shown in Table 36 in parentheses in the unit cost column. Some categories on cost are no longer listed in PSSRU in the format reported by Brown et al.[4] The ERG has updated these costs from the latest version of PSSRU that listed these costs.

Table 36 Unit costs (PFS and PD health states)

Resource	Unit cost (ERG estimate)	Unit	Source
Outpatient follow-up visit	£136.43	per visit	NHS Reference Costs 2016-2017, Outpatient attendance data, Consultant Led, Service code 800, Clinical Oncology
Chest Radiography	£27.78 (£27.22)	per case	NICE technology appraisal TA199; (£24.04 in 2009 - inflated to 2016/17 using the PSSRU HCHS index)
CT scan (chest)	£112.07	per case	NHS Reference Costs 2016-2017, Diagnostic Imaging, Outpatient, HRG code RD24Z (two areas with contrast)

CT scan (other)	£112.07	per case	NHS Reference Costs 2016-2017, Diagnostic Imaging, Outpatient, HRG code RD24Z (two areas with contrast)
ECG	£224.99	per case	NHS Reference Costs 2016–2017, Complex ECG, HRG code EY50Z
Community nurse visit	£62.00 (£69.10[1])	per hour	PSSRU 2017 p.159: Cost per hour Band 8a
Clinical nurse specialist	£62.00 (£77.35[1])	per contact hour	PSSRU 2017 p.207: Cost per hour Band 8a
GP surgery visit	£38.00	per visit	PSSRU 2017, p.162: Cost per patient contact lasting 9.22 minutes, including direct care staff costs, including qualifications
GP home visit	£94.82 (£119.95[2])	per visit	PSSRU 2016, p.145: Cost per home visit including 11.4 minutes for consultations and 12 minutes for travel - inflated to 2016/17 using the PSSRU HCHS index
Therapist visit	£45.00	per visit	PSSRU 2017, p.177: Cost per hour for community occupational therapist, including training

Table reproduced from CS Table 41
GP, general practitioner; CT, computerised tomography; ECG, electrocardiogram; NHS, National Health Service; PSSRU, Personal Social Services Research Unit; NICE, The National Institute for Health and Care Excellence; HRG, Healthcare Resource Groups; HCHS, hospital and community health services

[1] Costs updated from PSSRU 2015 to 2016/17.
[2] Costs updated from PSSRU 2013 to 2016/17.

The total cost per week in the PFS health state is £61.80 and for the PD state £117.00. The corrected health state costs using the ERG estimates shown in the Table above is £65.53 for the PFS and £139.39 for the PD health state.

4.2.6.5 Costs of terminal care

The company's economic model includes terminal care costs reflecting the resources used by patients in various care settings. The resources have been taken from Brown et al.[4] and were originally reported in a Marie Curie report.[24] The company has updated the unit costs to 2016/17. As noted above for health state costs, the company has incorrectly updated the unit costs using the HCHS prices index, rather than the HCHS pay and prices index. The ERG has corrected the unit costs and these are shown, together with the resource use in Table 37 (CS Table 42). The total cost of terminal care used in the model is £4456.13 and the corrected ERG estimate is £4556.88.

Table 37 Resource use and unit costs for terminal care/end of life

Resource	Unit cost (ERG estimate)	Number of consumption	% of patients in each setting	Assumptions / Source
Community nurse visit	£62.00 (£69.10) per hour	28.00 hours	27%	PSSRU 2017, p.159: Cost per hour Band 8a
GP Home visit	£94.82 (£119.95) per visit	7.00 visits	27%	PSSRU 2016, p.145: Cost per home visit including 11.4 minutes for consultations and 12 minutes for travel - inflated to 2016/17 using the PSSRU HCHS index
Macmillan nurse	£41.35 (£46.07) per hour	50.00 hours	27%	Assumed to be 66.7% of community nurse cost
Drugs and equipment	£574.57 £562.73 per patient	Average drug and equipment usage	27%	Value from Brown et al study (2013) inflated to 2016/17 using the PSSRU HCHS index
Terminal care in hospital	£4003.46 (£3921.95) per episode	1 episode (9.66 days)	56%	NICE TA531, inflated to 2016/17 using the PSSRU HCHS index[25]
Terminal care in hospice	£5004.33 (£4902.44) per episode	1 episode (9.66 days)	17%	NICE TA531, assumed 25% increase on hospital inpatient care
Total cost	£4456.13 (£4556.88) per episode			

Table reproduced from CS Table 42

4.2.6.6 Adverse events

The company's economic model includes the costs for treating adverse events. Adverse event data for patients treated with Atezo+Bev+CP are taken directly from IMPower150 for grade ≥ 3 grade adverse events with an incidence of ≥2%. For the comparator treatment, the company conducted a systematic literature review (CS Appendix D). The frequency of adverse events is shown in CS Table 43. The unit costs for treating the adverse events are shown in CS Table 44 (Table 38). The unit costs are based on NHS Reference costs 2016/17. As noted for health care costs, the company has incorrectly updated some of the costs by using the HCHS prices index instead of the HCHS pay and prices index. The costs of adverse events corrected by the ERG are shown in Table 38.

The adverse event costs per patient in the economic model are £1227.68 for Atezo+Bev+CP, £272.54 for pemetrexed + platinum and £723.78 for pemetrexed + platinum + pemetrexed

maintenance. The corrected adverse event costs estimated by the ERG produce a total cost per patient of £1334.27 for Atezo+Bev+CP, £289.67 for pemetrexed + platinum and £861.56 for pemetrexed + platinum + pemetrexed maintenance

Table 38 Unit cost per adverse event used in the economic model

Adverse Event	Unit cost	ERG estimate	Reference
Anaemia	£2,748.57	£2692.61	NICE TA531[12] - inflated to 2016/17 using the PSSRU HCHS index
Asthenia	£2,914.59	£2855.25	Assumed same as fatigue
Fatigue	£2,914.59	£2855.25	Brown 2013[4] (inflated to 2016-17 using PSSRU inflation indices)
Febrile neutropenia	£7097.41	£7045.41	NICE TA531[12] - inflated to 2016/17 using the PSSRU HCHS index
Leukopenia	£376.80	£1209.92	Assumed same as neutropenia
Nausea	£1019.12	£998.38	Brown 2013[4] (inflated to 2016-17 using PSSRU inflation indices)
Neutropenia	£601.23	£1209.92	Brown 2013[4] (inflated to 2016-17 using PSSRU inflation indices)
Thrombocytopenia	£123.51	£120.99	NICE TA484,[26] NICE TA520,[6] NICE TA525[7]
White blood cell count decreased	£449.34	£440.19	NICE TA484,[26] NICE TA520,[6] NICE TA525[7]

Table reproduced from CS Table 44
[1] Costs inflated using HCHS pay and prices index, rather than HCHS prices index
[2] Brown et al (2013) assumes two episodes of hospital treatment, rather than one episode

ERG conclusion: The approach taken by the company for estimating health care resources and costs is reasonable and in line with previous NICE technology appraisals for NSCLC. There are some minor discrepancies to some of the cost estimates as they have not been updated correctly. The resources use estimates used in the model are from outdated sources and need updating. The ERG suggests that the resource use could have been taken from the IMPower150 trial, if these data were available.

4.3 Cost effectiveness results

4.3.1 Company's base case results

The CS presents results of the base case economic analysis in a pairwise format, comparing Atezo+Bev+CP to each comparator separately (CS Tables 46 – 51, B.3.7.1). In response to ERG clarification question B4, the company produced an incremental analysis comparing the two included comparators, as well as results of pairwise analysis (Clarification response Appendix D). We reproduce the results with PAS discount price discounts for atezolizumab and

bevacizumab but list prices for comparators and subsequent treatments in Table 39, Table 40 and Table 41 below. Results with all applicable PAS price discounts are presented in a separate confidential addendum to this report.

Table 39 Company base case results, ITT population (PAS for atezolizumab and bevacizumab, list prices for other treatments) – deterministic (CS Clarification response Table 35)

Treatment	Total		Incremental analysis ICER (£/QALY)	Pairwise ICERs vs. comparator (£/QALY)
	Costs (£)	QALYs		
PEM+plat	[redacted]	[redacted]		£16,419
PEM+plat+PEM maint	[redacted]	[redacted]	£35,985	Dominant
Atezo+Bev+CP	[redacted]	[redacted]	Dominant	-

Table 40 Company base case results, PD-L1 negative/low population (PAS for atezolizumab and bevacizumab, list prices for other treatments) – deterministic (CS Clarification response Table 36)

Treatment	Total		Incremental analysis ICER (£/QALY)	Pairwise ICERs vs. comparator (£/QALY)
	Costs (£)	QALYs		
PEM+plat	[redacted]	[redacted]		£13,424
PEM+plat+PEM maint	[redacted]	[redacted]	£38,943	Dominant
Atezo+Bev+CP	[redacted]	[redacted]	Dominant	-

Table 41 Company base case results, EGFR/ALK positive population (PAS for atezolizumab and bevacizumab, list prices for other treatments) – deterministic (CS Clarification response Table 37)

Treatment	Total		Incremental analysis ICER (£/QALY)	Pairwise ICERs vs. comparator (£/QALY)
	Costs (£)	QALYs		
PEM+plat	[redacted]	[redacted]		£14,552
PEM+plat+PEM maint	[redacted]	[redacted]	£31,523	£7,014
Atezo+Bev+CP	[redacted]	[redacted]	£7,014	-

The ERG found small cost differences in the total costs for comparators in the EGFR/ALK population reported in Table 41. The results when the ERG ran the company model are shown in Table 42. This does not substantively change the estimated ICERs.

Table 42 ERG rerun of company base case for the EGFR/ALK positive population (PAS for atezolizumab and bevacizumab, list prices for comparators and subsequent treatments) – deterministic

Treatment	Total		Incremental analysis ICER (£/QALY)	Pairwise ICERs vs. comparator (£/QALY)
	Costs (£)	QALYs		
PEM+plat	[redacted]	[redacted]		£14,430
PEM+plat+PEM maint	[redacted]	[redacted]	£36,206	£4,758
Atezo+Bev+CP	[redacted]	[redacted]	£4,758	-

ERG conclusion: Except for the EGFR/ALK positive population, other base case results reported in the company's clarification response were reproducible when the ERG ran the company's model.

4.3.2 Company's sensitivity analyses

The company's sensitivity analysis comprised of probabilistic sensitivity analysis (PSA), one-way sensitivity analyses and scenario analyses. The company reports these set of analysis in the CS section B.3.8 and updates them in Appendix D of the company's clarification response.

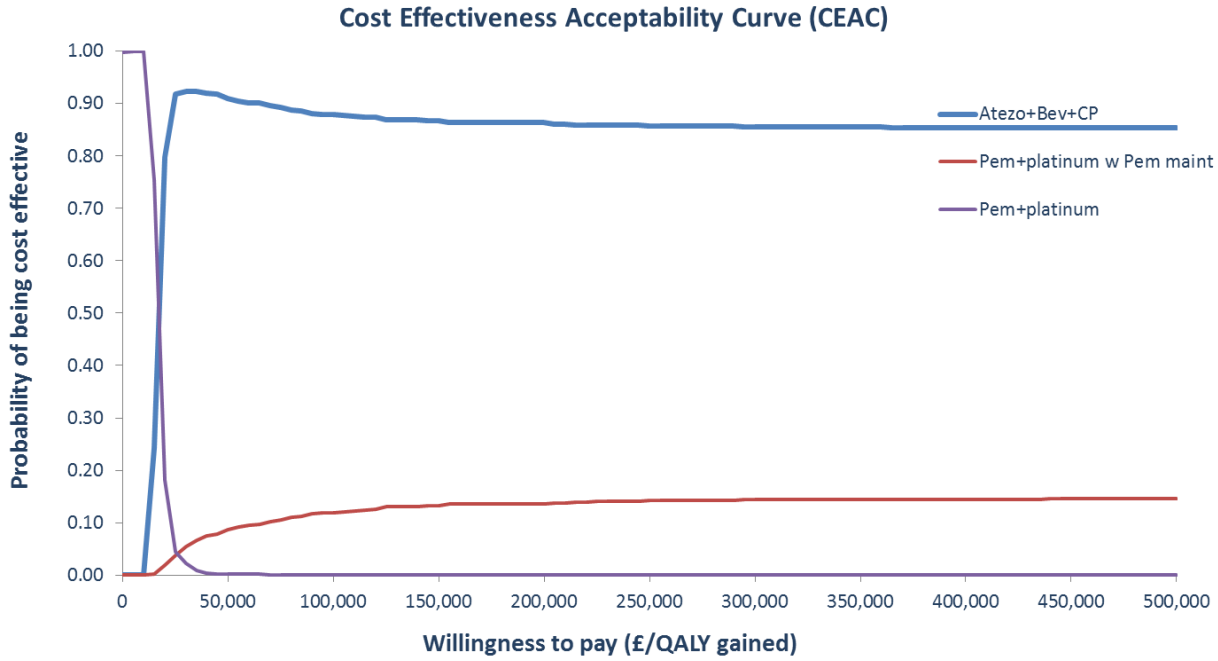
4.3.2.1 Company's probabilistic sensitivity analyses

The CS reports PSA performed on the base case analysis to assess parameter uncertainty (CS section B3.8.1) with 1000 samples.

The mean values, distributions around the means, and sources used to estimate the parameters are detailed in Appendix R of the CS. Joint uncertainty over parameters estimates used to estimate relative treatment effects on OS and PFS are sampled from the CODA output from the NMA. The company used the normal distribution for all other parameters varied in the PSA. A more standard approach is to use the gamma distribution for costs and the beta distribution for utilities. In addition, the company uses arbitrary variations for some of the input parameters of costs of +/- 5%. The ERG is of the opinion that 95% confidence intervals are more appropriate and if these CIs are not available varying by +/-25% or 30% of the base case input parameters is preferable.

Probabilistic estimates of costs, QALYs and ICERs were very similar to the mean probabilistic values (company clarification response, Appendix D, Tables 38 and 39). We reproduce the company's base case CEAC for the ITT population (with PAS discounts for the intervention only) in Figure 24. The curves are similar for the PD-L1 low/negative and EGFR/ALK positive populations.

Figure 22 Cost-Effectiveness Acceptability Curve (PAS for atezolizumab and bevacizumab, list price for comparators and subsequent treatments) – ITT population



Reproduced from the company's clarification response (Appendix D, Figure 16)

4.3.2.2 One-way sensitivity analysis

The company produced tornado plots to illustrate the effect of one-way sensitivity analysis on the ICERs in Appendix D of their clarification response (Figures 25 to 34). We reproduce the plot for the ITT population for the comparison with pemetrexed plus platinum in Figure 25 below (PAS discounts for atezolizumab and bevacizumab only). The CS states that the most influential parameters are the discount rates for costs and health outcomes, the administration cost for Atezo+Bev+CP, the utility value for the interval of >30 weeks before death and the weekly AE costs for Atezo+Bev+CP. Similar results are found for the subgroups and comparison including pemetrexed maintenance.

However, we note that the one-way deterministic sensitivity analysis does not include uncertainty over the treatment effects (either the baseline curves for Atezo+Bev+CP or relative effects versus the comparators). The company also uses arbitrary variations of +/- 5% for some of input parameters. The ERG is of the opinion that treatment effect is potentially a key driver of cost-effectiveness and should be varied according to the confidence intervals for PFS and OS.

ERG conclusion: The one-way sensitivity analyses do not capture the full uncertainty of the parameters because some parameters have only been varied by +/- 5% and the treatment effect has not been included.

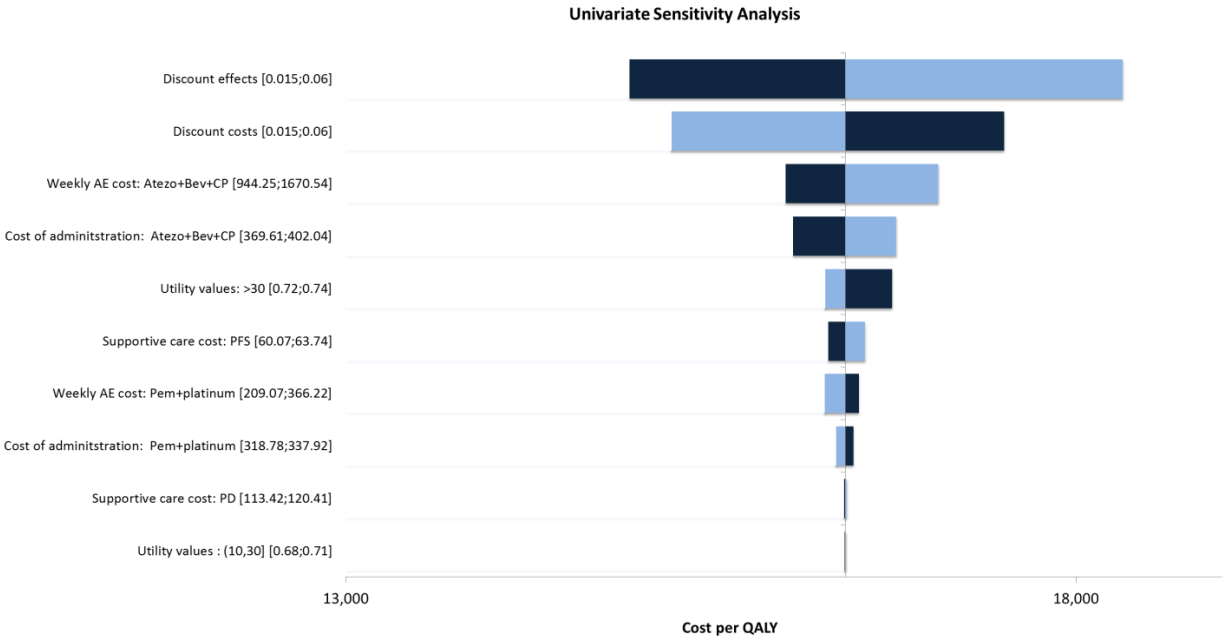


Figure 23 Tornado diagram – ITT population vs. pemetrexed plus platinum (PAS for atezolizumab and bevacizumab but not comparators or subsequent treatments)
Reproduced from Figure 27 in company's clarification question response

4.3.2.3 Scenario analyses

The company used deterministic scenario analyses to assess the impact of uncertainty around some other parameter inputs and structural assumptions: They explored the following scenarios:

- Alternative OS extrapolations
- Alternative PFS extrapolations
- Alternative TTD extrapolations
- Alternative NMA networks and models
- No treatment stopping rule for atezolizumab and bevacizumab
- Alternative time points for cap of treatment effect duration
- Alternative drug vial wastage assumptions
- Alternative utility values
- Alternative subsequent therapy approach
- Disutility for AEs

The company provided updated results in their clarification response document. We reproduce the results for the ITT population in Table 43 below (PAS discounts for atezolizumab and bevacizumab only). Other subgroups and PAS scenario analyses are in Appendix D of the company's clarification responses.

The scenario analyses show that the model results are most sensitive to changes to the choice of parametric curve for OS, treatment effect duration, treatment stopping duration, and the choice of studies to include in the NMA.

The economic model includes a macro which runs all the scenarios, with the exception of the subsequent treatment scenario. As noted in section 4.2.5 of this report, the ERG disagrees with the company's approach to estimating disutilities for the comparators. The ERG conducts a scenario including disutilities in section 4.4.

As observed with the company's base case analysis, the ERG was not able to exactly replicate the results for the scenario analyses for the EGFR/ALK positive population on the company's updated model.

> **ERG conclusion**: The company's choice of scenarios is comprehensive and informative.

Table 43 Company scenario analyses - ITT (PAS for atezolizumab and bevacizumab only)

Scenario		Pairwise ICER Atezo+Bev+CP vs	
		PEM+platinum	PEM+platinum w PEM maint
OS distribution	Exponential (base case)	£16,419	Dominant
	Weibull	£18,470	Dominant
	Log-normal	£11,840	Dominant
	Gen Gamma	£23,304	Dominant
	Log-logistic	£12,376	Dominant
	Gompertz	Does not converge	Dominant
PFS distribution	KM - Log-logistic tail (base case)	£16,419	Dominant
	Exponential	£18,324	£130
	Weibull	£18,073	Dominant
	Log-normal	£16,738	Dominant
	Gen Gamma	£17,637	Dominant
	Log-logistic	£16,418	Dominant
	Gompertz	£18,428	£612
TTD distribution	KM- Exponential tail (base case)	£16,419	Dominant
	Exponential	£17,533	Dominant
	Weibull	£15,639	Dominant
	Log-normal	£18,191	Dominant
	Gen Gamma	£14,558	Dominant
	Log-logistic	£20,885	£546
	Gompertz	Does not converge	Does not converge
Alternative NMA network	ITT (base case)	£16,419	Dominant
	ITT exclude KEYNOTE	£16,501	Dominant
	ITT exclude PARAMOUNT	Comparison not feasible - no connected network	Dominant
Alternative NMA model	NMA - Fract Poly (FE) (base case)	£16,419	Dominant
	NMA - PH	£20,028	Dominant
	NMA - Fract Poly (RE)	£16,523	Dominant
Treatment stopping rule	At 2 years (base case)	£16,419	Dominant
	No treatment stopping rule	£25,865	£12,234
Treatment effect duration	5 years (base case)	£16,419	Dominant
	105 months	£17,223	Dominant
	150 months	£17,522	Dominant
	195 months	£17,586	Dominant
	240 months (lifetime)	£17,595	Dominant
Wastage	With vial sharing (base case)	£16,419	Dominant

	No vial sharing	£16,427	Dominant
Utility values	IMpower150 (Proximity to death) (base case)	£16,419	Dominant
	IMpower150 (Pre/Post progression)	£17,090	Dominant
	Pembrolizumab utilities (US publication)	£14,960	Dominant
	Chouaid et al. 2013	£16,974	Dominant
	Nafees et al. 2008	£18,438	Dominant
Subsequent treatments	Base case	£16,419	Dominant
	IMpower150	£20,866	£1,201
AE disutility	No (base case)	£16,419	Dominant
	Yes	£16,502	Dominant

Reproduced from the company's clarification response (Table 43 and Table 44)

4.3.3 Model validation and face validity check

The company described their approach to model validation in CS section B.3.10.1. The CS states that "all outcomes of the economic model have been extensively compared to and validated against all available evidence, as well as clinical expert opinion, to assess the accuracy of the modelled survival".

For Atez+Bev+CP, the company compare the model extrapolations with 5-year survival estimates from 10 UK clinicians (4.2.4.2CS B.3.3.2). For the pemetrexed-based comparators, the company compare against predictions of 5-year survival under standard care from the pembrolizumab appraisal TA531 (8 to 11%), and against estimates from the Flatiron study[14]: 8.3% for pemetrexed with platinum and 12.3% for pemetrexed with platinum and pemetrexed maintenance (CS Appendix M). A critique of the company's selection of respective time-to-event distributions and extrapolation techniques can be found in section 4.2.2 of this ERG report.

Additional parameters validated by the company include health state inclusion, relevant comparators and resource use. To verify that these parameters were reflective of clinical practice, the company consulted UK clinical experts. The company reports that internal quality control and external validation of their economic model was conducted by a consultancy. The CS further describes other methods used to validate the model outputs such as cell-by-cell validation and pressure tests using extreme values.

The ERG checked the company's economic model for transparency and validity. The model was developed in Microsoft Excel and the visual basic codes were accessible.

We conducted a range of 'white box' tests to verify model inputs, calculations and outputs which consisted of:

- Cross-checking of all parameter inputs against values in the CS and cited sources;
- Checking that model outputs such as base case deterministic results and results of scenario analysis reported in the CS were reproducible by manually running the model;
- Checking individual equations and formulas within the model;
- Testing the logic of formulas in the model by substituting model inputs with a range of extreme values;
- Checking that visual basic codes did what they were designed to do.

Generally, we found the economic model to be of a good quality, with a few errors in input parameters, logic or coding. We identified a few small errors that we report and correct in section 4.4 below. However, these errors did not make any substantive difference to the results of economic analysis.

We also attempted to validate the outcomes from the company model (Table 39) against estimates from related NICE appraisals (summarised in Table 23, section 4.1 above).

- The company's ITT base case QALY estimate for pemetrexed with platinum (1.01 QALYs) is lower than that estimated in TA181 (0.61). However, this may be explained by the longer time horizon (20 years compared with 6 years in TA181).
- The QALY gain with pemetrexed-based treatment with vs. without maintenance (0.38 QALYs in the company's ITT base case) is not directly comparable with the incremental QALY gain in the pemetrexed maintenance appraisal (TA402) (0.21), because the latter does not include the induction period.
- The estimate of 1.35 QALYs attributed to standard care in the pembrolizumab appraisal (TA531) compares with 1.01 QALYs for pemetrexed with platinum and 1.39 for pemetrexed with platinum plus pemetrexed maintenance in the current company ITT base case. The standard of care comparator in TA531 includes chemotherapy regimens as well as pemetrexed-based ones, so the estimates from the current company base case are rather lower than might be expected.

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 corrections to

the model for discrepancies in costs, changes to the population used for the NMA, changes to the trials included in the NMA, using alternative assumptions for the duration of the treatment effect, changes to the parametric curves using for the survival extrapolations and inclusion of adverse event disutilities.

We firstly correct discrepancies in the model and we then run the model for our preferred base case. Table 44 details the corrections made to the company model. Our base case is explained and justified in Table 45. We conduct additional analyses by varying the ERG base case and these scenarios are shown in Table 46.

Table 44 ERG corrections to company model

Parameter	Company estimate	ERG Correction	Explanation
Vial sharing	5%	0%	No vial sharing is more appropriate
Pembrolizumab	£3781.28	£4453.13	As in TA428[17]
PFS health state cost	£61.80	£65.53	Some cost discrepancies in CS
PD health state cost	£117.00	£139.39	Some cost discrepancies in CS
Terminal care	£4456.13	£4556.88	Some cost discrepancies in CS
Adverse event cost Atezo+Bev+CP	£1227.68	£1334.27	Some cost discrepancies in CS
Adverse event cost PEM + platinum	£272.54	£289.67	Some cost discrepancies in CS
Adverse event cost PEM + platinum + PEM main	£723.78	£861.56	Some cost discrepancies in CS

Table 45 ERG additional analysis

Model aspect	Company analysis	ERG base case
Decision problem		
Population	Company reports results for ITT as well as PD-L1 and EGFR/ALK subgroups. The latter use subgroup-specific extrapolations for the atezolizumab arm survival curves and relative effects from the subgroup NMAs. • ITT curves & NMA • PD-L1 curves & NMA • EGFR/ALK curves & NMA	The ERG base case uses subgroup-specific survival curves for the atezolizumab arm for the PD-L1 & EGFR/ALK subgroups combined with relative effects from the ITT NMA, as this is more robust and there is no evidence of effect modification from the IMpower150 trial. ERG base case: • PD-L1 curves + ITT NMA • EGFR/ALK curves + ITT NMA
Intervention	Atezolizumab in combination with bevacizumab, carboplatin and paclitaxel Combination without bevacizumab not pursued in anticipated marketing authorisation	No change
Comparators	Pemetrexed in combination with cisplatin/carboplatin, with or without pemetrexed maintenance The company is not seeking reimbursement for patients eligible for pembrolizumab (PD-L1 high expressors) The company does not model other comparators that are in scope, arguing that PEM+CIS/CARBO is 'standard of care' in the UK.	No change We agree that it is acceptable to omit pembrolizumab from analysis as reimbursement is not sought for the subgroup who meet NICE TA531 criteria.[12] We consider that the platinum-based chemotherapies listed in the scope should have been included as comparators. Expert advice to the ERG is that patients who cannot tolerate cisplatin will have carboplatin-based chemotherapy, followed by pemetrexed maintenance. Model does not include any comparator specified in the scope for the EGFR/ALK subgroup. We report cost-effectiveness relative to the pemetrexed comparators for this subgroup, but note that this is out of scope.
Structure and assumptions		
Time horizon	20 years, from age 63 to 82	No change

Stopping rule	2 year maximum in the base case. Scenario with no limit on treatment duration. This aligns with stopping rules for atezolizumab after chemotherapy (TA520) and pembrolizumab (TA531).	No change
Effect duration	5 year cut off for OS (3 years after stopping), with scenario analysis from 8.75 to 20 years In the revised model this was applied by setting the mortality rate for Atezo+Bev+CP equal to that for PEM+plat with maintenance.	No change for base case, but extend the scenario analysis due to uncertainty over the duration of effects after discontinuation of immunotherapies (e.g. as noted in TA 520).
Clinical parameters		
Fitted survival curves for atezolizumab combination	ITT & PD-L1 low • OS exponential • PFS KM + log-logistic tail • TTD exponential	ERG base case: The ERG prefers the Weibull distribution for OS extrapolation (section 4.2.4.1). The choice of parametric curves for PFS and TTD are reasonable, except for PFS curves for EGFR/ALK +ve subgroup. For this curve, the ERG prefers the log-logistic distribution.
	EGFR/ALK +ve subgroup • OS exponential • PFS log-normal • TTD exponential	
	KM tails attached where 20% of patients remain at risk Parametric curves fitted separately to Atezo+Bev+CP arm of IMpower150 (Jan 2018 cut off with investigator-assessed PFS).	
Relative effects	HR from ITT NMA FP (FE) P1=0 Weibull (scenarios: PH and RE NMA models, excluding KEYNOTE, excluding PARAMOUNT)	The ERG prefers the analysis excluding the PARAMOUNT trial (due to heterogeneity), with first order Weibull, fixed effects.
AE rates	See CS Tab 43 p132	No change
Utilities		
Health state	IMpower150 EQ-5D IPD time from death analysis (IMpower150 PF/PD, Huang, Nafees, Chouaid)	No change to health state utilities, however company has not included any differences in utility between the treatments.

AE disutilities	Not included in base case. Scenario with disutility estimated from trial EQ-5D analysis	ERG prefers to include disutilities due to adverse events (see above).
Resource use and costs		
Drug acquisition	Vial sharing (no vial sharing)	ERG prefers no vial sharing (ERG correction).
Price discounts	List prices and PAS discount (atezolizumab & bevacizumab)	PAS atezolizumab PAS bevacizumab CAA nivolumab CAA pemetrexed maintenance
Drug admin	Higher costs for atezolizumab - takes longer for infusion.	No change
Subsequent treatment		No change
Health state costs		Some discrepancies in cost calculations due to incorrect updating of costs (ERG correction).
Terminal care		Some discrepancies in cost calculations due to incorrect updating of costs (ERG correction).
AE costs		Some discrepancies in cost calculations due to incorrect updating of costs (ERG correction).

CAA commercial access agreement; FE Fixed effect; FP Fractional polynomial; HR hazard ratio; KM Kaplan Meier; NMA network meta-analysis; PAS Patient access scheme; PH Proportional hazards; RE Random effects

Table 46 ERG base case and ERG scenarios

	Subgroup	Company base case	ERG base case	ERG scenarios
Baseline OS	All	Exponential	Weibull	• Exponential • Log-logistic
Baseline PFS	ITT & PD-L1 low/-ve	KM + log-logistic	KM + log-logistic	• KM + exponential • KM + Weibull
	EGFR/ALK +ve	Log-normal	Log-normal	• Exponential • Weibull
NMA (OS & PFS)	ITT	FP (FE) ITT	ITT FP excluding PARAMOUNT (FE)	• ITT FP (RE) • ITT PH • Subgroup specific
	PD-L1 low/-ve	FP (FE) PD-L1 low/-ve		
	EGFR/ALK +ve	FP (FE) EGFR/ALK +ve		
TTD	All	KM + exponential for atezo and bev	KM + exponential for atezo and bev	• Bev until progression
		PEM follows PFS	PEM follows PFS	
Stopping rule and effect cap	All	2 year treatment + 3 year OS effects	2 year treatment + 3 year OS effects	• 2 years for OS • 5 years for OS • 3 years for PFS • no stopping rule or effect cap
Utilities	All	IMPower150 EQ-5D time-from-death with no treatment effect	IMPower150 EQ-5D time-from-death + disutility per grade 3+ treatment related AE	• IMpower150 EQ-5D health state model • No AE disutility
Subsequent treatments	All	UK scenario (CS Tab 34)	UK scenario (CS Tab 34)	Exclude nivolumab

AE Adverse events; FE Fixed effect; FP Fractional polynomial; KM Kaplan Meier; NMA network meta-analysis; RE Random effects

4.4.1 ERG corrections to company base case and scenarios

The company base case results for the three populations with ERG corrections are shown in Table 47 - Table 49, with PAS price for atezolizumab and bevacizumab and list price for comparators and subsequent treatments. The ERG corrections (Table 44) only have a minor impact on the results. We show equivalent results with all available PAS discounts in a separate confidential addendum, respectively.

Table 47 ERG corrected company base case for ITT population (PAS for Atezo & Bev only) - deterministic

Technologies	Total costs (£)	Total QALYs	ICER (£) fully incremental analysis	ICER (£) pairwise; Atezo+Bev+CP vs comparator
PEM+platinum	[redacted]	[redacted]	-	£14,467
PEM+platinum w PEM maint	[redacted]	[redacted]	£37,184	Dominant
Atezo+Bev+CP	[redacted]	[redacted]	Dominant	

Table 48 ERG corrected company base case for PD-L1 low/negative population (PAS for Atezo & Bev only) - deterministic

Technologies	Total costs (£)	Total QALYs	ICER (£) fully incremental analysis	ICER (£) pairwise; Atezo+Bev+CP vs comparator
PEM+platinum	[redacted]	[redacted]	-	£11,513
PEM+platinum w PEM maint	[redacted]	[redacted]	£39,876	Dominant
Atezo+Bev+CP	[redacted]	[redacted]	Dominant	

Table 49 ERG corrected company base case for EGFR/ALK positive population (PAS for Atezo & Bev only) - deterministic

Technologies	Total costs (£)	Total QALYs	ICER (£) fully incremental analysis	ICER (£) pairwise; Atezo+Bev+CP vs comparator
PEM+platinum	[redacted]	[redacted]	-	£14,547
PEM+platinum w PEM maint	[redacted]	[redacted]	£37,024	£4,563
Atezo+Bev+CP	[redacted]	[redacted]	£4,563	

Table 50 and Table 51 show the ERG corrections to the company scenario analyses for the ITT population with PAS discounts for atezolizumab and bevacizumab only.

Table 50 ERG corrected company scenarios for ITT population, comparison with pem+plat (PAS for Atezo & Bev only) - deterministic

Scenario		Atezo+Bev+CP		Pem+platinum		ICER
		Total QALYs	Total costs	Total QALYs	Total costs	
OS distribution	Exponential (base case)	[redacted]	[redacted]	[redacted]	[redacted]	£14,467
	Weibull	[redacted]	[redacted]	[redacted]	[redacted]	£15,784
	Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	£11,728
	Gen Gamma	[redacted]	[redacted]	[redacted]	[redacted]	£19,214
	Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	£12,041
	Gompertz	Does not converge				
PFS distribution	KM with Log-logistic tail (base case)	[redacted]	[redacted]	[redacted]	[redacted]	£14,467
	Exponential	[redacted]	[redacted]	[redacted]	[redacted]	£16,766
	Weibull	[redacted]	[redacted]	[redacted]	[redacted]	£16,614
	Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	£14,803
	Gen Gamma	[redacted]	[redacted]	[redacted]	[redacted]	£16,050
	Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	£14,460
	Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	£16,958
TTD distribution	KM with Exponential tail (base case)	[redacted]	[redacted]	[redacted]	[redacted]	£14,467
	Exponential	[redacted]	[redacted]	[redacted]	[redacted]	£15,585
	Weibull	[redacted]	[redacted]	[redacted]	[redacted]	£13,687
	Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	£16,236
	Gen Gamma	[redacted]	[redacted]	[redacted]	[redacted]	£12,604
	Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	£18,936
	Gompertz	Does not converge				
Alternative NMA network	ITT (base case)	[redacted]	[redacted]	[redacted]	[redacted]	£14,467
	ITT exclude KEYNOTE	[redacted]	[redacted]	[redacted]	[redacted]	£14,596
	ITT exclude PARAMOUNT	Does not converge				
Alternative NMA model	NMA - Fract Poly (FE) (base case)	[redacted]	[redacted]	[redacted]	[redacted]	£14,467
	NMA - PH	[redacted]	[redacted]	[redacted]	[redacted]	£17,595
	NMA - Fract Poly (RE)	[redacted]	[redacted]	[redacted]	[redacted]	£14,540
	At 2 years (base case)	[redacted]	[redacted]	[redacted]	[redacted]	£14,467

Treatment stopping rule	No treatment stopping rule	[redacted]	[redacted]	[redacted]	[redacted]	£23,915
Treatment effect duration	5 years (base case)	[redacted]	[redacted]	[redacted]	[redacted]	£14,467
	105 months	[redacted]	[redacted]	[redacted]	[redacted]	£14,976
	150 months	[redacted]	[redacted]	[redacted]	[redacted]	£15,213
	195 months	[redacted]	[redacted]	[redacted]	[redacted]	£15,265
	240 months (lifetime)	[redacted]	[redacted]	[redacted]	[redacted]	£15,272
Wastage	With vial sharing (base case)	[redacted]	[redacted]	[redacted]	[redacted]	£14,467
	No vial sharing	[redacted]	[redacted]	[redacted]	[redacted]	£14,467
Utility values	IMpower150 (Proximity to death) (base case)	[redacted]	[redacted]	[redacted]	[redacted]	£14,467
	IMpower150 (Pre/Post progression)	[redacted]	[redacted]	[redacted]	[redacted]	£15,058
	Chouaid et al. 2013	[redacted]	[redacted]	[redacted]	[redacted]	£14,956
	Nafees et al. 2008	[redacted]	[redacted]	[redacted]	[redacted]	£16,246
Subsequent treatments	Base case	[redacted]	[redacted]	[redacted]	[redacted]	£14,467
	IMpower150	[redacted]	[redacted]	[redacted]	[redacted]	£21,399
AE disutility	No (base case)	[redacted]	[redacted]	[redacted]	[redacted]	£14,467
	Yes	[redacted]	[redacted]	[redacted]	[redacted]	£14,589

Table 51 ERG corrected company scenarios for ITT population, comparison with pem+plat with pem maintenance (PAS for Atezo & Bev only) - deterministic

Scenario		Atezo+Bev+CP		Pem+platinum +maintenance		ICER
		Total QALYs	Total costs	Total QALYs	Total costs	
OS distribution	Exponential (base case)	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
	Weibull	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
	Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
	Gen Gamma	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
	Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
	Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
PFS distribution	KM with Log-logistic tail (base case)	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
	Exponential	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
	Weibull	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
	Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
	Gen Gamma	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
	Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
	Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
TTD distribution	KM with Exponential tail (base case)	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
	Exponential	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
	Weibull	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
	Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
	Gen Gamma	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
	Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
	Gompertz	Does not converge				
Alternative NMA network	ITT (base case)	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
	ITT exclude KEYNOTE	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
	ITT exclude PARAMOUNT	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
Alternative NMA model	NMA - Fract Poly (FE) (base case)	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
	NMA - PH	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
	NMA - Fract Poly (RE)	[redacted]	[redacted]	[redacted]	[redacted]	Dominant

Treatment stopping rule	At 2 years (base case)	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
	No treatment stopping rule	[redacted]	[redacted]	[redacted]	[redacted]	£6,042
Treatment effect duration	5 years (base case)	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
	105 months	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
	150 months	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
	195 months	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
	240 months (lifetime)	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
Wastage	With vial sharing (base case)	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
	No vial sharing	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
Utility values	IMpower150 (Proximity to death) (base case)	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
	IMpower150 (Pre/Post progression)	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
	Chouaid et al. 2013	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
	Nafees et al. 2008	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
Subsequent treatments	Base case	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
	IMpower150	[redacted]	[redacted]	[redacted]	[redacted]	£139
AE disutility	No (base case)	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
	Yes	[redacted]	[redacted]	[redacted]	[redacted]	Dominant

4.4.2 ERG base case and scenarios

Results for the ERG base case analysis for the ITT population are shown in Table 52 (PAS for atezolizumab and bevacizumab only). This analysis uses NMA results excluding the PARAMOUNT trial, so results are only available verses the comparator with pemetrexed maintenance. Equivalent results for the PD-L1 low/negative and EGFR/ALK positive populations are shown in Table 53 and Table 54.

Table 52 ERG base case for ITT population (PAS for atezolizumab and bevacizumab and list price for comparators and subsequent treatments)

Technologies	Total costs (£)	Total QALYs	ICER (£) fully incremental analysis	ICER (£) pairwise; Atezo+Bev+CP vs comparator
PEM+platinum w PEM maint	[redacted]	[redacted]		Dominant
Atezo+Bev+CP	[redacted]	[redacted]	Dominant	

Table 53 ERG base case results for PD-L1 population (PAS for atezolizumab and bevacizumab and list price for comparators and subsequent treatments)

Technologies	Total costs (£)	Total QALYs	ICER (£) fully incremental analysis	ICER (£) pairwise; Atezo+Bev+CP vs comparator
PEM+platinum w PEM maint	[redacted]	[redacted]		Dominant
Atezo+Bev+CP	[redacted]	[redacted]	Dominant	

Table 54 ERG base case results for EGFR/ALK population (PAS for atezolizumab and bevacizumab and list price for comparators and subsequent treatments)

Technologies	Total costs (£)	Total QALYs	ICER (£) fully incremental analysis	ICER (£) pairwise; Atezo+Bev+CP vs comparator
PEM+platinum w PEM maint	[redacted]	[redacted]		Dominant
Atezo+Bev+CP	[redacted]	[redacted]	Dominant	

The results of scenarios around the ERG ITT base case are shown in Table 55. Although these analyses do not reflect agreed price discounts for pemetrexed maintenance or for some subsequent treatments, they do indicate which parameters the model is most sensitive to: extrapolations of overall surival and treatment duration, the use of a stopping rule for

atezolizumab and bevacizumab as part of Atezo+Bev+CP and the costs of subequent treatments.

Table 55 ERG scenarios for ITT (PAS for Atezolizumab and Bevacizumab and list price for comparators and subsequent treatments)

Description		Atezo+Bev+CP		PEM+platinum+PEM Maintenance		ICER
		Total QALYs	Total costs	Total QALYs	Total costs	
OS distribution	Weibull (base case)	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
	Exponential	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
	Log-logistic					Dominant
PFS distribution	KM+log-logistic (base case)					Dominant
	KM + Exponential	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
	KM+weibull					Dominant
TTD distribution	KM + Exponential Pemetrexed follows PFS (base case)					Dominant
	Bevacizumab until progression	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
Alternative NMA network/ model	ITT FP excluding PARAMOUNT (FE) (base case)	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
	ITT FP (RE)	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
	ITT Excluding PARAMOUNT + PH					Dominant
Treatment stopping rule/ treatment effect	2 years treatment + 3 years OS effect (base case)	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
	2 years OS	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
	5 years OS					Dominant
	3 years PFS					Dominant
	No stopping rule or effect cap					£8,469
Utility values	IMPower150 EQ-5D, using time from death + disutilities (base case)	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
	IMPower150 EQ-5D health states	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
AE disutility	Disutilities per grade 3+	[redacted]	[redacted]	[redacted]	[redacted]	Dominant

	treatment related AE (base case)					
	No AE disutilities	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
Subsequent treatments	Base case	[redacted]	[redacted]	[redacted]	[redacted]	Dominant
	IMpower150	[redacted]	[redacted]	[redacted]	[redacted]	£3,132
	Exclude nivolumab	[redacted]	[redacted]	[redacted]	[redacted]	£3,670

Results of the ERG analyses with all available PAS discounts are in the separate confidential addendum.

4.4.3 Conclusions on cost effectiveness

4.4.3.1 Comparators

The comparators used for the EGFR/ALK positive (pemetrexed + cisplatin with or without pemetrexed maintenance) do not match the NICE scope which includes pembrolizumab and docetaxel. The company has also omitted chemotherapy with carboplatin comparators for the untreated PD-L1 low/negative subgroup, which may reflect current practice for patients who cannot tolerate cisplatin.

4.4.3.2 Model assumptions

The model structure is appropriate for NSCLC and correctly implemented. The use of a 20-year time horizon is reasonable, given the model projections of survival. We also agree with company's base case assumptions of a 2-year stopping rule for the Atez+Bev+CP intervention and the 5-year cap on the survival benefit for this combination. These assumptions are consistent with committee assumptions in previous appraisals of atezolizumab and other immunotherapies.

4.4.3.3 Extrapolation of OS, PFS and TTD

The company's base case extrapolations for OS are reasonable. The exponential distribution for the atezolizumab combination has a good fit to the IMPower150 data and, when coupled with a five-year cap on effects relative to the pemetrexed comparator with maintenance, clinically plausible extrapolations of survival at 5 and 10 years. We consider that the Weibull distribution is also plausible and gives more conservative survival predictions. The parametric curves chosen for PFS and TTD are reasonable and appropriate.

4.4.3.4 NMA

Given concerns about potential bias due to patient selection, we think it is appropriate to exclude the PARAMOUNT study from the NMA. The company's choice of survival curves for PFS and TTD are reasonable and appropriate.

4.4.3.5 Health utility

The company's approach to health state utility values is reasonable and consistent with the NICE reference case and with previous NICE technology appraisals. However, the ERG considers that the differences in treatment related adverse events between treatments have not been fully captured and it is unclear whether patients treated with Atezo + Bev. + CP have the same health state utility values whilst on treatment as those treated with pemetrexed + platinum (with or without pemetrexed maintenance).

4.4.3.6 Health resources and costs

The approach taken by the company for estimating health care resources and costs is reasonable and in line with previous NICE technology appraisals for NSCLC. There are some minor discrepancies to some of the cost estimates as they have not been updated correctly.

5 End of life

End of life criterion 1 - "The treatment is indicated for patients with a short life expectancy, normally less than 24 months". Table 56 reports the undiscounted life years from the company's model. The estimates for pemetrexed plus platinum with pemetrexed maintenance therapy exceed 24 months. The ERG's discounted estimates for pemetrexed maintenance therapy are less than 24 months in the ITT population (Table 57).

Table 56 Company base case undiscounted life years

Absolute life years (undiscounted)	PEM+platinum	PEM+platinum with PEM maint
ITT	1.53	2.18
PD-L1	1.55	2.27
EGFR/ALK +ve	2.04	3.15

Table 57 ERG base case undiscounted life years

Absolute life Years (undiscounted)	PEM+platinum with PEM maint
ITT	1.72

End of life criterion 2 – "There is sufficient evidence to indicate that the treatment offers an extension to life, normally of at least an additional 3 months, compared with current NHS treatment". Table 58 reports the company's modelled incremental undiscounted life years gained. For all populations the estimates exceed 3 months.

Table 58 Company modelled undiscounted life years gained

Life years gained (undiscounted)	Versus Pem+platinum	Versus PEM+platinum w PEM maint
ITT	1.08	0.42
PD-L1	1.01	0.29
EGFR/ALK +ve	3.08	1.97

The ERG's modelled undiscounted life years gained estimate is also greater than 3 months in the ITT population (Table 59).

Table 59 ERG modelled undiscounted incremental life years gained

LY gained (undiscounted)	Versus PEM+platinum w PEM maint
ITT	0.46

ERG conclusion: Atezo+Bev+CP meets both of the end-of-life criteria based on the ERG's modelled estimates in the ITT population. However, it does not appear to meet all of the end of life criteria when compared to pemetrexed plus platinum with pemetrexed maintenance therapy using the company's modelled estimates.

6 Innovation

The CS provides a lengthy justification for why atezolizumab should be considered a treatment innovation for the first line treatment of metastatic NSCLC (CS section B.2.12). The justification centres on a suggested unmet need for an improvement of efficacy in first-line treatments for non-squamous metastatic NSCLC, and specifically the need for further treatment options for

patients with low or negative PD-L1 expression and in patients with an EGFR or ALK mutation who are ineligible for, intolerable to or have progressed on targeted therapy.

The biological justification for combining an immunotherapy drug such as atezolizumab with chemotherapies (i.e. bevacizumab, carboplatin/paclitaxel) is described. The ERG notes that atezolizumab is an immune checkpoint inhibitor (specifically a PD-L1 blocking antibody) where as bevacizumab inhibits angiogenesis (development of blood supply for the tumour), cisplatin stops or slows tumour growth by interfering with DNA replication and the mitotic inhibitor paclitaxel, inhibits cell division. The CS highlights the synergistic effect of atezolizumab in combination with chemotherapies to enhance anti–PD-1– dependent anti-tumour effects.

Expert clinical advice to the ERG suggests that atezolizumab can be considered a treatment innovation as, apart from pembrolizumab for PD-L1 high expressers, there is no immunotherapy option for patients in the first line advanced setting. However, the regimen would be considered a more attractive option to clinicians if it did not contain bevacizumab due to the additional cost of this drug, and potential for additional adverse effects. As discussed earlier in this report (section 3.1.6), the IMPower150 trial was not designed to compare an atezolizumab plus bevacizumab regimen to an atezolizumab regimen without bevacizumab, and the anticipated marketing authorisation is for atezolizumab in combination with bevacizumab.

7 REFERENCES

1. National Institute for Health and Care Excellence. Lung cancer: diagnosis and management (CG121), 2011.
2. National Institute for Health and Care Excellence. Pemetrexed for the first-line treatment of non-small-cell lung cancer [TA181], 2009.
3. Royal College of Physicians. National Lung Cancer Audit Annual Report 2017 (for the audit period 2016), 2018.
4. Brown T, Pilkington G, Bagust A, et al. Clinical effectiveness and cost-effectiveness of first-line chemotherapy for adult patients with locally advanced or metastatic non-small cell lung cancer: a systematic review and economic evaluation. *Health Technol Assess* 2013;17(31):1-278. doi: 10.3310/hta17310 [published Online First: 2013/07/28]
5. Jansen JP. Network meta-analysis of survival data with fractional polynomials. *BMC Medical Research Methodology* 2011;11(1):61. doi: 10.1186/1471-2288-11-61
6. National Institute for Health and Care Excellence. Atezolizumab for treating locally advanced or metastatic non-small-cell lung cancer after chemotherapy [TA520], 2018.
7. National Institute for Health and Care Excellence. Atezolizumab for treating locally advanced or metastatic urothelial carcinoma after platinum-containing chemotherapy [TA525], 2018.
8. Guyot P, Ades A, Ouwens MJ, et al. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. *BMC medical research methodology* 2012;12(1):9.
9. Dias S, Ades AE, Welton N, et al. Network Meta-Analysis for Decision Making. Chichester: John Wiley & Sons Ltd. 2018.
10. Patient-reported outcomes in the randomized, Phase III IMpower150 study of atezolizumab + chemotherapy ± bevacizumab vs chemotherapy + bevacizumab in 1L nonsquamous metastatic NSCLC. American Society of Clinical Oncology; 2018.
11. National Institute for Health and Care Excellence. Pemetrexed for the maintenance treatment of non-small-cell lung cancer [TA190], 2010.
12. National Institute for Health and Care Excellence. Pembrolizumab for untreated PD-L1-positive metastatic non-small-cell lung cancer [TA531], 2018.
13. National Institute for Health and Care Excellence. Pemetrexed maintenance treatment for non-squamous non-small-cell lung cancer after pemetrexed and cisplatin [TA402], 2016.

14. F. Hoffmann-La Roche Ltd [data on file]. Flatiron Health Data, 2018.
15. Nafees B, Stafford M, Gavriel S, et al. Health state utilities for non small cell lung cancer. *Health Qual Life Outcomes* 2008;6:84. doi: 10.1186/1477-7525-6-84
16. Chouaid C, Agulnik J, Goker E, et al. Health-related quality of life and utility in patients with advanced non-small-cell lung cancer: a prospective cross-sectional patient survey in a real-world setting. *J Thorac Oncol* 2013;8(8):997-1003. doi: 10.1097/JTO.0b013e318299243b [published Online First: 2013/06/22]
17. National Institute for Health and Care Excellence. Pembrolizumab for treating PD-L1-positive non-small-cell lung cancer after chemotherapy [TA428], 2017.
18. Huang M, Lou Y, Pellissier J, et al. Cost Effectiveness of Pembrolizumab vs. Standard-of-Care Chemotherapy as First-Line Treatment for Metastatic NSCLC that Expresses High Levels of PD-L1 in the United States. *PharmacoEconomics* 2017;35(8):831-44. doi: 10.1007/s40273-017-0527-z
19. British National Formulary. BNF Update 2017 [Available from: https://www.medicinescomplete.com/mc/bnf/current/ accessed August 2018.
20. Department of Health. Drugs and pharmaceutical electronic market information (eMit): Commercial Medicines Unit (CMU); 2018 [Available from: https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit accessed July 2018 2018.
21. Socinski M, Jotte R, Cappuzzo F, et al. Atezolizumab for First-Line Treatment of Metastatic Nonsquamous NSCLC. *New England Journal of Medicine* 2018 doi: 10.1056/NEJMoa1716948
22. Department of Health. NHS Reference Costs 2016-17 2018 [Available from: https://improvement.nhs.uk/documents/1972/1_-_Reference_costs_publication_VSnAQ5x.pdf accessed April 2017.
23. Maslove L, Gower N, Spiro S, et al. Estimation of the additional costs of chemotherapy for patients with advanced non-small cell lung cancer. *Thorax* 2005;60(7):564-9. doi: 10.1136/thx.2004.039479 [published Online First: 2005/07/05]
24. Taylor D, Carter S. Marie Curie Cancer Care; Understanding the cost of end of life care in different settings. *School of Pharmacy, University of London,* 2004
25. Curtis L, Burns A. PSSRU: Unit Costs of Health and Social Care 2017. 2017
26. National Institute for Health and Care Excellence. Nivolumab for previously treated non-squamous non-small-cell lung cancer [TA484], 2017.

8 APPENDICES

8.1 NMA Critical appraisal checklist

Checklist		Response yes/no
Does the CS present an NMA?		Yes
Are the NMA results used to support the evidence for the clinical effectiveness of the intervention		Yes
Are the NMA results used to support the evidence for the cost-effectiveness of the intervention		Yes
Homogeneity		
	1. Is homogeneity considered?	Yes
	2. Are the studies homogenous in terms of patient characteristics and study design?	Yes, with the exception of the PARAMOUNT trial
	3. Is the method used to determine the presence of statistical heterogeneity adequate? (e.g. Chi-squared test, I-squared statistic)	Yes. Pairwise meta-analyses were used where network links were informed by more than one study. The I^2 statistic was reported.
	4. 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)	Yes. Subgroup analyses were done (PD-L1 low/-ve patients, and EGFR/ALK +ve patients), and sensitivity analyses were conducted (removing particular trials thought to be heterogeneous).
Similarity		
	1. Is the assumption of similarity stated?	No
	2. Have they justified their assumption?	N/A
Consistency		
	1. Does the analysis explicitly assess consistency?	N/A. There were no closed loops in the network, apart from the 3 arms of the IMPower150 trial.
	2. Does the method described include a description of the analyses/ models/ handling of potential bias/ inconsistency/ analysis framework?	N/A
	3. Are patient or trial characteristics compared between direct and indirect evidence trials?	N/A
	4. If Q3 is yes, and inconsistency is reported, is this accounted for by not combining the direct and indirect evidence?	N/A

N/A = Not applicable

Criterion	ERG assessment
NMA purpose	
1. Are the NMA results used to support the evidence for the clinical effectiveness of the intervention?	Yes, for the indirect comparison of Atezo+Bev+CP versus pemetrexed-based chemotherapy regimens.
2. Are the NMA results used to support the evidence for the cost-effectiveness of the intervention?	Yes, as above.
Evidence selection	
3. Are inclusion/exclusion criteria adequately reported?	Yes, following clarification question (question A12). Criteria are specified in CS Appendix Table 10.
4. Is quality of the included studies assessed?	Yes. CS appendix D.1.3.
Methods – statistical model	
5. Is the statistical model described?	Yes, briefly in the CS, and in more detail in the appendix (D1.1)
6. Has the choice of outcome measure used in the analysis been justified?	Yes. A feasibility assessment was done to determine whether connected networks could be formed for a range of outcomes within the scope of the appraisal (Appendix D1.1). Networks were constructed for OS, PFS, ORR and adverse events leading to discontinuation. Of these, OS and PFS are used to inform the economic model (and are the focus of this ERG report).
7. Has a structure of the network been provided?	Yes
8. Is homogeneity considered?	Yes
9. Are the studies homogenous in terms of patient characteristics and study design?	Yes, with the exception of the PARAMOUNT trial
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)	Yes. Subgroup analyses were done (PD-L1 low/-ve patients, and EGFR/ALK +ve patients), and sensitivity analyses were conducted (removing particular trials thought to be heterogeneous).
11. Is the assumption of similarity stated?	No explicit statement is given.
12. Is any of the programming code used in the statistical programme provided (for potential verification)?	Yes, following a clarification question request (question A21).
Sensitivity analysis	
13. Does the study report sensitivity analyses?	Yes. Results scenario analysis results are presented for the FP NMAs with the exclusion of the PARAMOUNT trial and the KEYNOTE trials (021, 189) (CS B.2.9.1). A scenario analysis assuming

	proportional hazards (analogous to an exponential FP model) is reported.
Results	
14. Are the results of the NMA presented?	Yes, in CS section B.2.9, and in Appendix D1.4. Additional results for all FP models tested, and random effects FP models and were supplied on request (clarification question A18 and A20).
15. Does the study describe an assessment of the model fit?	Yes. CS appendix D1.1 (Table 29) reports DIC values for fixed effect FP models. It is also stated that fit was assessed by visual inspection of hazard curves, survival curves and validation of the clinical plausibility of the extrapolated survival curves.
16. Has there been any discussion around the model uncertainty?	Yes, CS section B.2.9.1.
17. Are the point estimates of the relative treatment effects accompanied by some measure of variance such as confidence intervals?	Yes, 95% credible intervals are illustrated in hazard ratio plots (in light grey shaded regions surrounding the hazard ratio line).
Discussion - overall results	
18. Does the study discuss both conceptual and statistical heterogeneity?	Yes
Discussion - validity	
19. Are the results from the indirect/NMA compared, where possible, to those just using direct evidence?	N/A

8.2 ERG independent assessment of risk of bias for the trials included in the clinical systematic review and in the NMA.

Author / trial ID Company & ERG assessment	Was the allocation sequence adequately generated?	Was the concealment of treatment allocation adequate?	Was knowledge of the allocated interventions adequately prevented from participants and personnel	Was knowledge of the allocated interventions adequately prevented from outcome assessors	Were incomplete outcome data adequately addressed?	Are reports of the study free of suggestion of selective outcome reporting?	Was the study apparently free of other problems that could put it at a high risk of bias?
IMpower150							
Company	Yes	Yes	N/A (open label study)	Unclear	Yes	No	Yes
ERG	Yes	Yes	No	No	Yes for PFS/OS Unclear for PRO outcomes	Yes	Yes
KEYNOTE-024							
Company	Yes	Yes	N/A (open label study)	Yes	Yes	Yes	Yes
ERG	Yes	Yes	No	Yes (central review PFS & response). Unclear (other outcomes)	Yes for PFS, OS & response.	Yes	Yes
ERACLE							
Company	Yes	Yes	No	No	No	Yes	Yes
ERG	Yes	Unclear	Unclear	Unclear	Unclear	Yes	Yes
KEYNOTE-021							
Company	Yes	Yes	No	Yes	Yes	Yes	Yes
ERG	Yes	Yes	No	Yes (central review objective response & PFS).	Yes	Yes	Yes

				Unclear (duration of response)			
KEYNOTE-189							
Company	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ERG	Yes	Yes	Yes	Yes	Yes	Yes	Yes
PARAMOUNT							
Company	Yes	Yes	Yes	Unclear	Yes	Yes	Yes
ERG	Yes	Yes	Yes	Yes (PFS, response)	Unclear	Yes	Yes
PRONOUNCE							
Company	Unclear	Unclear	No	No	Unclear	Yes	Yes
ERG	Unclear	Unclear	No	No	Unclear	Yes	Yes