ID970 STA REPORT

LIVERPOOL REVIEWS AND IMPLEMENTATION GROUP (LRiG)

Atezolizumab for treating locally advanced or metastatic non-small cell lung cancer after chemotherapy [ID970]

Confidential until published

This report was commissioned by the NIHR HTA Programme as project number 16/56/11

Completed 27 April 2017

Produced by: Liverpool Reviews & Implementation Group (LR*i*G)

Authors: Sophie Beale, Research Associate (Decision Analyst), LR*i*G, University of Liverpool

James Mahon, Director, Coldingham Analytical Services, Berwickshire

Sarah Nevitt, Research Associate (Medical Statistician), LR*i*G, University of Liverpool

Angela Boland, Associate Director, LR*i*G, University of Liverpool

Janette Greenhalgh, Senior Research Fellow (Clinical Effectiveness), LR*i*G, University of Liverpool

Marty Richardson, Research Associate (Medical Statistician), LR*i*G, University of Liverpool

Eleanor Kotas, Information Specialist, LR*i*G, University of Liverpool

Joanne McEntee, Senior Medicines Information Pharmacist, North West Medicines Information Centre, Pharmacy Practice Unit, Liverpool

John Green, Consultant in Medical Oncology, The Clatterbridge Centre NHS Foundation Trust, Liverpool

Correspondence to: Sophie Beale, Research Associate, Liverpool Reviews and Implementation Group, University of Liverpool, Whelan Building, The Quadrangle, Brownlow Hill, Liverpool L69 3GB

Date completed: 27 April 2017

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 16/56/11

Declared competing interests of the authors: Dr Escriu received reimbursement fees from Merk, Sharp & Dohme for attending a lung cancer conference in 2016 and from AstraZeneca for attending a meeting.

Acknowledgements: The authors would like to thank Dr Carles Escriu (Consultant Medical Oncologist, The Clatterbridge Centre NHS Foundation Trust, Liverpool) and Dr Rui Duarte (LR*i*G) for their feedback on a final draft version of the report.

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

This report should be referenced as follows: Beale S, Mahon J, Nevitt S, Boland A, Greenhalgh J, Richardson M, Kotas E, McEntee J, Green J. Atezolizumab for treating locally advanced or metastatic non-small cell lung cancer after chemotherapy [ID970]: A Single Technology Appraisal. LRiG, University of Liverpool, 2017

Contributions of authors:

Sophie Beale	Critical appraisal of the company submission
James Mahon	Critical appraisal of the submitted economic evidence
Sarah Nevitt	Critical appraisal of the submitted statistical evidence
Angela Boland	Critical appraisal of the company submission
Janette Greenhalgh	Critical appraisal of the submitted clinical evidence
Marty Richardson	Critical appraisal of the submitted statistical evidence
Eleanor Kotas	Cross checking of the submission search strategies
Joanne McEntee	Critical appraisal of the company submission
John Green	Clinical advice and critical appraisal of the clinical sections of the company submission

All authors read and commented on draft versions of this report.

Table of contents

List of tables

List of figures

LIST OF ABBREVIATIONS

1º P	OAK trial primary population
2º P	OAK trial secondary population
AE	adverse event
ALK	anaplastic lymphoma kinase
BSA	body surface area
CDF	Cancer Drugs Fund
CHMP	Committee for Medicinal Products for Human Use
CI	confidence interval
CNS	central nervous system
CR	complete response
CrI	credible interval
CS	company submission
CSR	clinical study report
DIC	Deviance Information Criteria
DOR	duration of response
EAMS	Early Access to Medicines Scheme
ECOG	Eastern Cooperative Oncology Group
EGFR	epidermal growth factor receptor
EMA	European Medicines Agency
eMIT	electronic market information tool
EORTC	European Organisation for the Treatment of Cancer
EQ-5D-3L	European quality of life - 5 dimensions, 3 levels questionnaire
ERG	Evidence Review Group
ESMO	European Society for Medical Oncology
FDA	Food and Drug Administration
FE	fixed effect
FP	fractional polynomial
FAD	final appraisal determination
HR	hazard ratio
HRG	healthcare resource group
HRQoL	health-related quality of life
HTA	health technology assessment
IC	tumour-infiltrating immune cell
ICER	incremental cost effectiveness ratio
IHC	immunohistochemistry
ITC	indirect treatment comparison
ITT	Intention-to-treat
IV	intravenous
K-M	Kaplan-Meier
KRAS	Kirsten rat sarcoma
NA	not applicable
NCI CTCAE	National Cancer Institute Common Terminology Criteria for Adverse Events
NE	not evaluable
NLCA	National Lung Cancer Audit
NR	not reported
NSCLC	non-small cell lung cancer
ORR	objective response rate
OS	overall survival
PD-1	programmed cell death protein 1
PD-L1	programmed death-ligand 1
PFS	progression-free survival
PH	proportional hazards
PPS	post-progression survival
PR	partial response
PRO	patient reported outcome
PS	performance score
PSA	probabilistic sensitivity analysis
PSS	personal social services
PSSRU	Personal Social Services Research Unit
QALY	quality adjusted life year
RCT	randomised controlled trial
RE	random effect
RECIST	response evaluation criteria in solid tumours
sd	standard deviation
SD	Jansen method for assessing heterogeniety
SAE	serious adverse event
SAP	statistical analysis plan
SmPC	Summary of Product Characteristics
TC	tumour cell
ToT	time on treatment
TRAE	treatment-related adverse event
TTD	time to treatment discontinuation

1 SUMMARY

1.1 Scope of the submission

The remit of the Evidence Review Group (ERG) is to comment on the clinical and cost effectiveness evidence submitted to the National Institute for Health and Care Excellence (NICE) as part of the Single Technology Appraisal (STA) process. Clinical and economic evidence has been submitted to NICE by Roche Products Limited in support of the use of atezolizumab (Tecentriq®) for the treatment of locally advanced or metastatic non-small cell lung cancer (NSCLC) after chemotherapy.

1.2 Critique of the decision problem in the company submission

Population

The population described in the final scope issued by NICE is people with locally advanced or metastatic NSCLC whose disease has progressed after chemotherapy. This population can be considered to be the same as the population addressed in the company submission (CS).

Intervention

Atezolizumab does not currently have a UK marketing authorisation. The company made an application on 20th April 2016 and anticipates receiving the Committee for Medicinal Products for Human Use (CHMP) opinion in [redacted], with regulatory approval expected in [redacted] The application is for the treatment of adult patients with locally advanced or metastatic NSCLC) after prior chemotherapy.

Atezolizumab is a monoclonal antibody that binds to and inactivates a protein called programmed death-ligand 1 (PD-L1) on the surface of tumour cells (TCs) and tumour-infiltrating immune cells (ICs), inhibiting the binding to PD-1 and B7.1.

The treatment regimen for atezolizumab is a flat dose of 1200mg administered intravenously in a hospital setting, over a 30-minute period, every 3 weeks. It is stated within the draft Summary of Product Characteristics (SmPC) that patients should be treated with atezolizumab until loss of clinical benefit or unmanageable toxicity.

Comparators

The comparators specified in the final scope issued by NICE are docetaxel, nintedanib+docetaxel, pembrolizumab, nivolumab and best supportive care (BSC).

- Included comparators:
  - direct evidence is available for the comparison of the effectiveness of atezolizumab versus **docetaxel** (administered at a dose of 75mg/m^2 every three weeks) from the OAK and POPLAR trials
  - treatment with **nintedanib+docetaxel** is recommended by NICE as an option for treating locally advanced, metastatic or locally recurrent NSCLC of adenocarcinoma histology that has progressed after first-line chemotherapy. In the absence of direct evidence to allow a comparison of the effectiveness of treatment with atezolizumab versus nintedanib+docetaxel, the company undertook an indirect treatment comparison (ITC).
- Excluded comparators:
  - **nivolumab** was not recommended by NICE for the treatment of locally advanced or metastatic NSCLC and hence cannot be considered a standard of care
  - the company provides three reasons for excluding **pembrolizumab**. First, marketing authorisation for pembrolizumab is only for patients with PD-L1 positive NSCLC and therefore does not match the anticipated marketing authorisation for atezolizumab. Second, accurate comparisons between treatments is not possible due to the differences between tests used in clinical studies to select patients. Third, pembrolizumab has only been recently recommended by NICE to treat patients with NSCLC and is unlikely to represent a standard of care at this time
  - **BSC** was excluded due to a clinically validated assumption that patients eligible for treatment with atezolizumab would be considered fit enough to receive other treatments.

Outcomes

Clinical evidence is presented in the CS for all five outcomes specified in the final scope issued by NICE: progression-free survival (PFS), overall survival (OS), objective response rate (ORR), adverse events (AEs) and health-related quality of life (HRQoL).

Subgroups

It is specified within the final scope issued by NICE that, if evidence allows, consideration will be given to subgroups based on biological markers. Within the CS, results have been provided from the OAK trial by baseline characteristics and for histology subgroups (squamous and non-squamous disease). Results have also been presented for patients with no measurable PD-L1 expression (TC0/IC0) and for patients with ≥1% PD-L1 expression.

Other considerations

- Agreed patient access schemes (PAS) are in place for atezolizumab and nintedanib
- The company has not identified any equality issues
- The company has presented a case for atezolizumab to be assessed against the NICE End of Life criteria.

1.3 Summary of clinical effectiveness evidence submitted by the company

The direct clinical evidence for the treatment of atezolizumab versus docetaxel was derived from the OAK and POPLAR trials.

Results from the OAK and POPLAR trials

Results from both the OAK and POPLAR trials show that treatment with atezolizumab is associated with a statistically significant and clinically meaningful improvement in median OS (4.2 months in the OAK trial and 2.9 months in the POPLAR trial) compared to docetaxel in patients with Eastern Cooperative Oncology Group performance status (ECOG PS) 0 and 1. In the OAK trial, this statistically significant gain in OS is observed regardless of histology and PD-L1 status. However, in the POPLAR trial, statistically significant improvement is observed only in the non-squamous histology subgroup and for individuals with any type of NSCLC of ≥1% PD-L1 expression. Improvement in OS with atezolizumab compared with docetaxel is also generally consistent across patient baseline characteristics in both trials. No statistically significant difference in investigator-assessed PFS was observed between atezolizumab and docetaxel arms in either trial.

Results from the company's indirect treatment comparison (ITC) suggest that the best estimate of the expected difference in OS is around 6 to 7 months for atezolizumab versus docetaxel (compared to a median OS gain of 4.2 months and 2.9 months in the OAK and POPLAR trials respectively). Results from the company's ITC suggest that the best estimate of the expected difference in OS is around 5 to 6 months for the comparison of atezolizumab versus nintedanib+docetaxel. Also, results from the company's ITC analyses suggest that there is no statistically significant difference in PFS survival for the comparison of atezolizumab versus docetaxel and for atezolizumab versus nintedanib+docetaxel.

The company has collected HRQoL outcome data using the European Organisation for Research and Treatment Cancer (EORTC) Quality of Life questionnaire, the EORTC Quality of Life in Lung Cancer questionnaire and the EQ-5D-3L questionnaire. Analyses of HRQoL data collected during the OAK trial show that there was no clinically meaningful worsening of commonly reported cancer treatment-related symptoms for patients treated with

atezolizumab, while there was a clinically meaningful worsening in alopecia and peripheral neuropathy throughout treatment for patients treated with docetaxel. In addition, patients treated with atezolizumab demonstrated prolonged time to deterioration of patient-reported chest pain compared with patients treated with docetaxel (hazard ratio [HR] 0.72, 95% confidence interval [CI]: 0.55 to 0.93).

1.4 Summary of the ERG's critique of submitted clinical effectiveness evidence

The ERG considers that the OAK and POPLAR trials were of good quality and well conducted; patient characteristics were balanced across the arms and the statistical methods were generally appropriate. However, the open-label design of these trials provides the opportunity for investigator-assessed outcomes to be biased. Also, the ERG notes that OS and PFS HRs must be interpreted with caution due to hazards not being proportional, as demonstrated by the company.

The ERG does not agree with the ITC approach taken by the company as the main network includes comparators that are not listed in the final scope issued by NICE. In addition, the ERG does not consider that the company was justified in excluding pembrolizumab from the ITC network of comparators relevant to this appraisal. During the clarification process, the ERG asked the company to undertake two further ITC analyses. However, the company undertook these using non-equivalent populations and results should be viewed with extreme caution:

- based on a (reduced) network using data from the intention-to-treat (ITT) populations of the OAK and POPLAR trials and the adenocarcinoma population from the LUME-Lung 1 trial, results suggest that the best estimate of expected difference in OS for atezolizumab versus nintedanib+docetaxel is 3.33 months (compared to 4.74 months when the analysis was carried out using LUME-Lung 1 trial total population) and is not statistically significant. However, this analysis was undertaken using non-equivalent populations and results should be viewed with caution.
- based on a (reduced) network using data from the ITT populations of the OAK, POPLAR and KEYNOTE-010 trials (the latter assessing the efficacy of pembrolizumab as a first-line treatment for metastatic NSCLC in adults whose tumours express PD-L1 with a ≥50% tumour proportion score) suggest that there is no statistically significant difference in OS or PFS for patients receiving atezolizumab when compared with pembrolizumab.

The ERG considers that the company's use of a fractional polynomial (FP) approach to conduct the ITC is appropriate. However, FP ITC results are influenced by a range of factors (e.g., comparators and population selected, type of FP model chosen and the use of FE or RE) which means that it is difficult to identify the most appropriate combination of factors to

use to generate ITC results. The ERG also considers that the expected values generated by the ITC are difficult to interpret. In addition, the ERG considers that the company's criteria for assessing the presence of heterogeneity in the analyses is inappropriate.

1.5 Summary of cost effectiveness evidence submitted by the company

The company developed a de novo partitioned survival model in Microsoft Excel to compare the cost effectiveness of treatment with atezolizumab versus docetaxel or nintedanib+docetaxel for previously treated patients with advanced NSCLC. The model comprises three mutually exclusive health states: 'on treatment', 'off treatment' and death. All patients start in the 'on treatment' state until they discontinue treatment or die. The model time horizon is set at 25 years with a 1-week cycle length. The model perspective is that of the UK NHS. Outcomes were measured in quality adjusted life years (QALYs), and both costs and QALYs were discounted at an annual rate of 3.5%, as recommended by NICE.

The OS of patients treated with atezolizumab was estimated using a mixed cure-rate model: survival for 98% of the population was a log-logistic distribution fitted to OAK trial data, while the remaining 2% were considered to have the same chance of survival as the general age-matched population. The OS models for patients receiving docetaxel and nintedanib+docetaxel were constructed by adjusting the trajectory for patients receiving atezolizumab using results from the company's FP ITC analyses. The company's base case analysis prediction is a mean of 2.22 life years gained (LYG) for patients receiving atezolizumab, 1.19 LYG for patients receiving docetaxel and 1.31 LYG for patients receiving nintedanib+docetaxel.

HRQoL data collected as part of the OAK trial using the EQ-5D-3L questionnaire were used in the company model. These data were differentiated by the time to a patient's death and by whether patients were 'on treatment' or 'off treatment' treatment. The mean EQ-5D utility scores by time to death used in the company base case for the 'on treatment' and 'off treatment' states are >30 weeks before death: 0.77 and 0.68; >15 weeks and ≤30 weeks before death: 0.71 and 0.58; >5 and ≤15 weeks before death: 0.61 and 0.43; and ≤5 weeks before death: 0.39 and 0.35.

Resource use and costs were estimated based on information from the OAK trial, published sources and clinical experts. For atezolizumab, the company provided the list price and the Department of Health PAS discount. Full list prices were used to represent the cost of the comparator drugs. The company is unaware of the PAS price for nintedanib.

Using list prices only, the company base case incremental cost effectiveness ratio (ICER) for the comparison of treatment with atezolizumab versus docetaxel is £72,356 per QALY gained; treatment with atezolizumab generates 0.748 additional QALYs at an additional cost of £53,970. For the comparison of treatment with atezolizumab versus nintedanib+docetaxel, the ICER is £56,076 per QALY gained; treatment with atezolizumab generates 0.646 additional QALYs at an additional cost of £36,209.

The company carried out a wide range of deterministic sensitivity analyses. The most influential parameters for both atezolizumab versus docetaxel and atezolizumab versus nintedanib+docetaxel are related to the cure fraction rate applied to atezolizumab, the monthly cost of atezolizumab and the discount rate used for effects.

The company's probabilistic sensitivity analysis (PSA) results show that when the cost effectiveness of treatment with atezolizumab is compared with docetaxel and nintedanib+docetaxel, there is a 1% probability of treatment with atezolizumab being cost effective at a threshold of £50,000 per QALY gained. The company carried out 12 scenario analyses and results from these demonstrate that the cost effectiveness of treatment with atezolizumab is only sensitive to the distribution chosen to extrapolate time to treatment discontinuation (TTD) with atezolizumab and then only if a log-logistic distribution is chosen.

1.6 Summary of the ERG's critique of submitted cost effectiveness evidence

The ERG considers that there are three errors in the company model that need to be corrected if the model is to produce accurate cost effectiveness results that reflect the underlying assumptions of the company base case. These errors are:

- incorrect application of discounting
- absence of age-dependent utility decrements
- incorrect use of a half-cycle correction to TTD data.

The ERG estimates that the accurate ICER, under the company base case assumptions for the comparison of the cost effectiveness of atezolizumab versus docetaxel is £77,569 per QALY gained and for the comparison of treatment with atezolizumab versus nintedanib+docetaxel it is £60,366 per QALY gained.

The ERG considers that the company's approach to modelling OS generates overly optimistic survival gains when treatment with atezolizumab is compared with docetaxel and when atezolizumab is compared with nintedanib+docetaxel. The ERG has identified three issues

with the mixed cure-rate approach taken by the company to model OS for patients receiving atezolizumab:

- use of the log-logistic function produces an implausibly long survival tail
- there is insufficient evidence for application of a cure-rate
- the value for the cure-rate used by the company was not justified by the company.

A further issue with the company's atezolizumab OS model relates to the company's assumption that treatment with atezolizumab has a lifetime protective effect. This assumption has been criticised by a previous NICE Appraisal Committee when considering the use of an immunotherapy for treating patients with previously treated advanced or metastatic NSCLC. In addition, the ERG highlights that the company's approach to modelling OS for patients receiving atezolizumab results in mortality rates that are, at some points, lower than the mortality rates of the UK general population of the same age.

The company approach to modelling of OS for docetaxel and for nintedanib+docetaxel involved adjusting the company's OS atezolizumab model using the relevant hazard rates generated by the company's FP ITC. Due to concerns relating to the company's FP ITC, including the fact that the FP ITC used to generate hazard rates involved inputs that are not relevant to this appraisal, the ERG has little confidence in the results produced by this approach. The ERG highlights that the expected docetaxel survival results produced by the company's FP ITC are optimistic when compared with median OS from the OAK trial. The ERG also highlights that the FP ITC used by the company to model OS for patients receiving nintedanib+docetaxel was not restricted to the nintedanib+docetaxel licensed population (patients with adenocarcinoma), meaning that the company's ITC results for this treatment are not relevant to this appriasal.

1.7 Summary of company's case for End of Life criteria being met

To meet the NICE End of Life criteria the company must demonstrate that:

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

The company has put forward a case that atezolizumab meets NICE's End of Life criteria based on the following points:

- the company quotes data that show median survival for patients with Stage IIIb and Stage IV NSCLC is 7.5 months and 3.4 months, respectively
- base case results generated by the company's economic model suggest that the mean difference in OS between patients treated with atezolizumab versus docetaxel or nintedanib+docetaxel is more than 3 months.

1.8 ERG commentary on End of Life criteria

The ERG agrees with the company that patients with advanced NSCLC have a life expectancy of less than 24 months.

An examination of the ERG's remodelled OS suggests that treatment with atezolizumab generates a mean survival gain of 4.7 months compared to docetaxel. However, compared to treatment with nintedanib+docetaxel, the size of the survival gain is uncertain. The company has provided evidence that suggests there is no statistically significant difference in OS for atezolizumab (total population) compared to nintedanib+docetaxel (adenocarcinoma patients only). If this result is reliable, then, for the adenocarcinoma population, atezolizumab does not meet the NICE End of Life criteria for life extension.

1.9 ERG commentary on the robustness of evidence submitted by the company

1.9.1 Strengths

Clinical evidence

- OAK and POPLAR trials were of good quality and well conducted
- EQ-5D data were collected during the OAK trial
- the ERG recognises the considerable effort made by the company to generate ITC results employing a methodology which accounts for hazards not being proportional.

Cost effectiveness evidence

- the economic model was well constructed
- the company used TTD to cost study treatments
- the company used EQ-5D utility scores by time to death
- the company carried out a comprehensive range of deterministic sensitivity and scenario analyses.

1.9.2 Weaknesses and areas of uncertainty

Clinical evidence

- the company should have included pembrolizumab as a comparator
- only investigator-assessed PFS results are available from the OAK and POPLAR trials
- the ERG considers that the company should have included full subgroup analyses of effectiveness and cost effectiveness by levels of PD-L1 expression
- the PFS and OS HRs from OAK and POPLAR trial data were calculated using a pre-specified method that relies on an assumption that hazards are proportional. However, as demonstrated by the company, this assumption does not hold and therefore OS and PF HRs must be interpreted with caution
- the company approach to the ITC is influenced by a range of factors (e.g., comparators and population selected, type of FP model chosen and the use of FE or RE) which means that it is difficult to identify the most appropriate combination of factors to use to generate ITC results
- the FP ITC results are difficult to interpret
- the company's criteria for assessing the presence of heterogeneity in the ITC analyses is inappropriate
- clinical advice to the ERG is that AEs arising from treatment with atezolizumab and other immunotherapies in patients with NSCLC require careful monitoring by a specialist clinical team with the experience to provide early recognition and management of immunotherapy-related AEs.

Cost effectiveness evidence

- the ERG identified three model construction errors: incorrect application of discounting, absence of age-dependent utility decrements and incorrect use of a half-cycle correction to TTD data
- the company's approach to modelling of OS for patients treated with atezolizumab used a mixed cure-rate model; however, there is insufficient evidence for the application of a cure-rate and the value used for the cure-rate was not justified by the company the company's approach to modelling OS for patients treated with atezolizumab is implausible as it resulted in survival rates that, at some points, were higher than that of the UK general population
- the company assumed a lifetime duration of treatment effect for atezolizumab, an approach that has been criticised by a previous NICE Appraisal Committee when assessing an immunotherapy for the treatment of patients with advanced or metastatic NSCLC
- confidence in modelling OS for patients receiving docetaxel by adjusting the OS atezolizumab model by hazard rates generated by the company's ITC is limited by the ERGs concerns relating to the company's FP ITC, including the fact that the FP ITC used to generate hazard rates involved inputs that are not relevant to this appraisal

- confidence in modelling OS for patients receiving nintedanib+docetaxel by adjusting the OS atezolizumab model by the hazard rates generated by the company's ITC is limited by concerns relating to identifying the most relevant FP ITC, including the fact

that the FP ITC used to generate hazard rates involved inputs that are not relevant to this appraisal and that the FP ITC was not limited to patients with adenocarcinoma histology.

1.10 Summary of exploratory and sensitivity analyses undertaken by the ERG

The ERG's preferred method was to model OS for both atezolizumab and docetaxel by using Kaplan-Meier (K-M) data from the OAK trial for as long as possible, and then to append exponential curves to project OS for the remainder of the model time horizon. The ERG also limited the duration of treatment effect of atezolizumab to approximately 3 years.

The FP ITC results generated by the company showed that when treatment with atezolizumab (whole population) was compared with nintedanib+docetaxel (adenocarcinoma population) using the reduced network (i.e., comparators of relevance to this appraisal) there was no statistically significant difference in expected OS between the two therapies. The ERG, therefore, undertook an analysis in which the OS of patients receiving nintedanib+docetaxel was the same as that of patients receiving atezolizumab, and the treatment effect of both interventions was limited to 3 years. Hence, the only modelled differences were therapy costs and HRQoL (utility values were adjusted for each treatment to take into account the incidence of AEs).

1.11 Cost effectiveness conclusions

Application of the ERG model amendments results in an ICER for the comparison of treatment with atezolizumab versus docetaxel of £170,497 per QALY gained.

Application of the ERG model amendments results in an ICER for the comparison of treatment with atezolizumab versus nintedanib+docetaxel of £1,170,793 per QALY gained.

2 BACKGROUND

2.1 Critique of company's description of underlying health problem

Section 3.1 of the company submission (CS[1]) includes an overview of non-small cell lung cancer (NSCLC). Section 3.2 of the CS includes a description of the effects of the disease on patients, carers and society. Key points from these sections of the CS are included as bulleted items in Box 1 and Box 2. The Evidence Review Group (ERG) considers that these points appropriately summarise the underlying health problems.

Box 1 Company overview of NSCLC

- Lung cancer is the third most common cancer in the UK, accounting for 13% of all new cancer cases (n=46,403) in 2014. It is responsible for 22% of all cancer deaths in the UK, making it the most common cause of cancer death. Around 35,900 people died of lung cancer in the UK in 2014. One in 13 men and 1 in 17 women will be diagnosed with lung cancer during their lifetime.
- Lung cancer is classified based upon its histology and can be broadly divided between small cell lung cancer and NSCLC. NSCLC represents approximately 85% of all lung cancer cases and includes several subtypes.
- Two subtypes are squamous and non-squamous histologies, with adenocarcinoma accounting for 96% of non-squamous cases.
- It has been observed the population with squamous disease suffers significantly poorer overall survival than the population with non-squamous disease.
- Early diagnosis of NSCLC is difficult as, at this stage, the disease is often asymptomatic, and symptoms of late-stage or advanced disease are non-specific. As a result, the majority of patients with lung cancer are initially diagnosed with disease that is already locally advanced or metastatic.
- NSCLC is staged according to the TNM classification, based on the primary tumour size and extent (T), regional lymph node involvement (N), and presence or absence of distant metastases (M). This information is combined to assign an overall stage of 0, I, II, III, or IV. This submission focuses on locally advanced and metastatic NSCLC, i.e. unresectable Stage IIIA, Stages IIIB and IV.
- The discovery of the EGFR mutations and rearrangements of the ALK gene have led to a paradigm shift with the advancement of targeted therapies for the 10 to 20% of patients with metastatic NSCLC whose tumours harbour these oncogenic alterations.
- Disease progression is still inevitable in the majority of patients treated with targeted therapies. Furthermore, patients without a mutation conferring sensitivity to a targeted agent are typically treated with chemotherapy, especially platinum-based chemotherapy, which is associated with modest treatment benefits and significant toxicities.
- There, therefore, remains an unmet need for new treatments which do not cause significant toxicity or a deterioration in quality of life and that improve survival for those patients who progress following targeted therapy and for patients ineligible for targeted therapy that relapse after first-line chemotherapy for whom docetaxel-based treatments are currently the most widely used.

ALK=anaplastic lymphoma kinase; EGFR=epidermal growth factor receptor; NSCLC=non-small cell lung cancer; TNM= Tumour-Node-Metastasis
Source: CS, Section 3.1

Box 2 Company overview of the effects of NSCLC on patients, carers and society

- The symptoms of lung cancer include persistent coughing (sometimes with blood present), chest pain, shortness of breath, recurrent chest infections, weight loss and tiredness. The high symptom burden in patients with advanced NSCLC has a highly negative impact on HRQoL, well-being and on family functioning.
- Due to its severe toxicity profile, chemotherapy is often associated with various complications and diminished HRQoL in patients with lung cancer. In addition, disease progression can also have a marked impact on patients' HRQoL.
- Advanced lung cancer can have a significant impact on the emotional and social wellbeing of the patient's family. The lives of patients and their families may become centred around clinic appointments, while increasing physical limitations can lead to changes in interpersonal roles and relationships, adversely affecting family relationships.
- Lung cancer is associated with a significant burden on caregivers, which can include social isolation, psychological impairment and poorer quality of life.
- Caregivers shoulder an economic burden with higher annual indirect costs with presenteeism-related impairment (impairment while working) and overall work impairment. A modelling study estimated the mean cost of providing informal care to lung cancer patients at the end of life in England and Wales was £73m, approximately one third of the total cost of care for this patient group.
- The direct costs associated with the treatment of lung cancer place a considerable burden on healthcare budgets, especially since the diagnosis, treatment and follow-up of lung cancer predominantly occurs within secondary care.
- A recent retrospective, descriptive cohort study conducted to evaluate the direct costs of hospital care in the diagnosis and management of 3,274 lung cancer patients, using routine NHS data (costs adjusted to 2013/14 prices) estimated mean cumulative costs to be £5,852 at 90 days and £10,009 at one year. The majority of costs (58.5%) were accrued within the first 90 days, with acute inpatient costs the largest contributor at one year (42.1%).

HRQoL=health-related quality of life; NSCLC=non-small cell lung cancer
Source: CS, Section 3.2

2.2 Company's overview of current service provision

The company presents an overview of the clinical care pathway in Section 3.3 of the CS. Details include a treatment algorithm outlining the existing treatment pathway for patients with advanced or metastatic NSCLC (reproduced in Figure 1). The algorithm is based on published NICE guidelines[2] and guidance[3-15] as listed in Section 3.5 of the CS. The guidelines and guidance that were identified by the company, along with additional guidance identified by the ERG, are summarised in Section 2.4.

The anticipated positioning of atezolizumab in the pathway is for patients who have progressed on a prior chemotherapy regimen. The ERG notes that two targeted therapies (erlotinib and crizotinib) are presented as comparators in the company's algorithm. However, expert advice to the company is that targeted therapy treatment options are likely to be preferred over immunotherapy in patients with confirmed epidermal growth factor receptor (EGFR) or anaplastic lymphoma kinase (ALK) mutations.

The ERG considers that the algorithm presented by the company reflects current clinical practice and would capture the treatment pathway in the event that atezolizumab were recommended by NICE for use in the NHS.

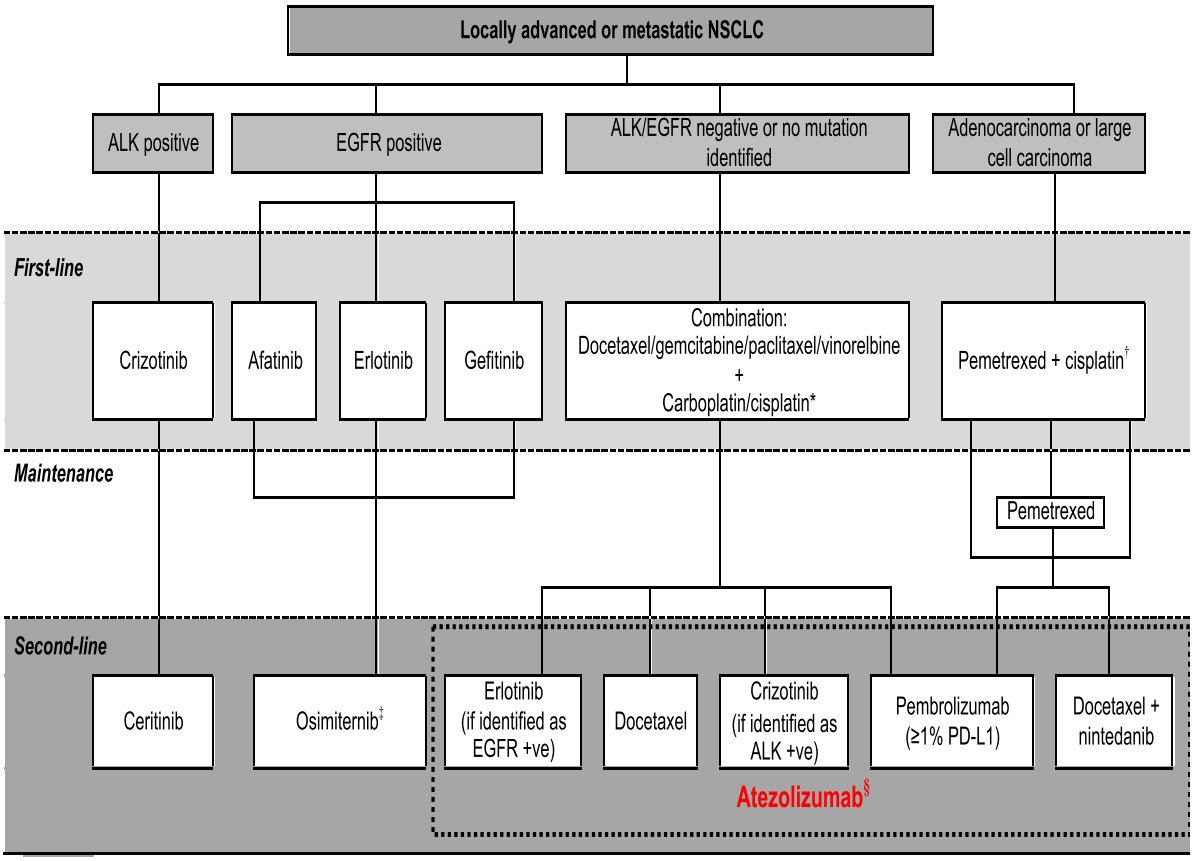


Figure 1 Treatment pathway based on NICE lung cancer clinical guideline (CG121)
§Dotted box indicates proposed position of atezolizumab based on anticipated indication
Source: CS, Figure 3

2.3 Life expectancy of people with NSCLC

The company presents information published by Cancer Research UK[16] that shows that lung cancer was the most common cause of cancer death in the UK in 2014 (approximately 35,900 deaths) and that it accounted for 22% of all cancer deaths in the UK that year. The company also provides data from a publication by Beckett et al[17] that suggest that the median survival for patients with Stage IIIb and Stage IV disease is 7.5 months and 3.4 months respectively (additional information can be found in CS, Table 9). In addition, the proportions of patients with Stage IIIb and Stage IV disease who are alive at 5 years are 7% and 3% respectively.

Table 1 Survival figures for patients with Stage IIIb/IV NSCLC and PS 0 or 1

	Chemotherapy	No chemotherapy
1-year survival	47%	25%
Median survival	11.2 months	5.3 months

PS=performance score; NSCLC=non-small cell lung cancer
Source: CS, Table 9

2.4 Summary of relevant clinical guidance and guidelines

The company provides details of relevant published guidance[3-15] and treatment guidelines[2] in Section 3.5 of the CS. NICE guidance and guidelines identified by the company and additional guidance identified by the ERG, are summarised in Table 2.

Table 2 Relevant NICE guidelines and guidance

NICE guideline or guidance	Summary of NICE recommendations
Guideline	
Lung cancer: diagnosis and management CG121[2] (2011)	• For patients with tumours of negative or unknown EGFR status and good performance status (WHO 0, 1 or a Karnofsky score of 80–100) chemotherapy should be offered; where the chemotherapy should be a combination of a single third generation drug (docetaxel, gemcitabine, paclitaxel or vinorelbine) plus a platinum drug (either carboplatin or cisplatin) • Patients who are unable to tolerate combination therapy may be offered single-agent chemotherapy with a third-generation drug
First-line treatment	
TA181[5] (2009)	• **Pemetrexed in combination with cisplatin**: if the histology of the tumour has been confirmed as adenocarcinoma or large-cell carcinoma
TA192[6] (2010)	• **Gefinitib**: patients whose tumours test positive for EGFR tyrosine kinase mutation
TA258[8] (2012)	• **Erlotinib**: patients whose tumours test positive for EGFR tyrosine kinase mutation
TA310[3] (2014)	• **Afatinib**: patients whose tumours test positive for EGFR tyrosine kinase mutation
TA406[12] (2016)	• **Crizotinib**: patients whose tumours test positive for ALK mutation
Maintenance treatment	
TA190[7] (2010)	• **Pemetrexed**: patients with other than predominantly squamous cell histology if disease has not progressed immediately following platinum-based chemotherapy in combination with gemcitabine, paclitaxel or docetaxel
TA402[13] (2016)	• **Pemetrexed**: patients with non-squamous disease whose disease has not progressed immediately after 4 cycles of pemetrexed and cisplatin induction therapy and who have an ECOG PS 0 or 1 at the start of maintenance treatment
Second-line treatment	
TA374[9] (2015)	**Erlotinib** is an option for patients who have: • had non-targeted chemotherapy because of delayed confirmation that their tumour is EGFR-TK mutation-positive. • progressed after non-targeted chemotherapy and who have tumours of unknown EGFR-TK mutation status, but only if the result of an EGFR-TK mutation diagnostic test is unobtainable because of an inadequate tissue sample or poor-quality DNA; the treating clinician considers that the tumour is very likely to be EGFR-TK mutation-positive; and there is an observed response within the first 2 cycles of treatment.
TA395[11] (2016)	• **Ceritinib**: adults with advanced ALK positive disease who have previously received crizotinib
TA347[4] (2015)	• **Nintedanib+docetaxel**: for patients with locally advanced, metastatic or locally recurrent NSCLC of adenocarcinoma histology that has progressed after first-line chemotherapy,
TA416[14] (2016)	• **Osimertinib**: patients with EGFR T790M mutation-positive disease whose disease has progressed after first-line treatment with an EGFR-TK inhibitor (only available via the CDF)
TA422[10] (2016)	• **Crizotinib:** previously treated adults with ALK positive NSCLC (after a rapid re-review by the CDF)
TA428[15] (2017)	• **Pembrolizumab:** patients with PD-L1 positive NSCLC in adults who have had at least one prior chemotherapy (and EGFR/ALK targeted treatment, if relevant) if treatment is stopped at 2 years of uninterrupted treatment and no documented disease progression

ALK=anaplastic lymphoma kinase; CDF=Cancer Drugs Fund; ECOG PS=Eastern Cooperative Oncology Group performance status; EGFR=epidermal growth factor receptor; NSCLC=non-small cell lung cancer; PD-L1=programmed death-ligand 1; TK=tyrosine kinase; WHO=World Health Organization

Other guidelines

Other relevant guidelines identified by the company (CS, Section 3.6) are:

- Lung cancer in adults: quality standards (QS17)[18]
- European Society for Medical Oncology (ESMO) Clinical Practice Guidelines, 2016[19]
- National Comprehensive Cancer Network (NCCN) Clinical Practice Guidelines.[20]

2.5 Innovation

The company states (CS, pp37-38) that atezolizumab:

- differs from other (anti-PD-1) antibodies approved for the treatment of NSCLC as it results in the dual blockade of PD-1 and B7.1 while leaving the PD-1/PD-L2 interaction intact, thereby potentially preserving peripheral immune homeostasis
- is anticipated to be approved for all locally advanced or metastatic NSCLC patients with prior chemotherapy, regardless of PD-L1 expression status.

The company considers that treatment with atezolizumab addresses a significant unmet need and represents a clinically significant innovative therapeutic option, which will provide significant positive impact on patients' lives.

The company highlights that a number of anti-PD-1 antibodies (e.g., pembrolizumab and nivolumab) are currently under development, along with anti-PD-L1 antibodies (e.g., atezolizumab) in a range of adult cancers. At the present time, the comparative efficacy of the two classes is unknown although the adverse events profiles are broadly similar and include immune-related effects on endocrine, neurological and pulmonary function. The role of the biomarker PD-L1 assessed by immunohistochemistry remains under development.

The ERG notes that atezolizumab is the first PD-L1 antibody to be assessed by NICE for the treatment of NSCLC (pembrolizumab and nivolumab are both PD-1 antibodies).

2.6 Number of patients eligible for treatment with atezolizumab

The company estimates that in England and Wales, approximately [redacted] patients will be eligible for treatment with atezolizumab in 2018. The company's method for calculating this number is described in the CS (Table 101, p228) and relies heavily on assumptions.

In 2014, the manufacturer of nintedanib estimated that 703 patients with locally advanced or metastatic adenocarcinoma would be eligible for second-line treatment with nintedanib+docetaxel and that there would be no population growth between 2014 and 2018.[21] The ERG, therefore, considers that the company estimate of [redacted] patients may be too high, even for the whole population.

3 CRITIQUE OF COMPANY'S DEFINITION OF DECISION PROBLEM

A summary of the ERG's comparison of the decision problem outlined in the final scope issued by NICE and that addressed within the CS is presented in Table 3. Each parameter in Table 3 is discussed in more detail in the text following the table (Section 3.1 to Section 3.7).

Table 3 Comparison between NICE scope and company decision problem

Final scope issued by NICE Parameter and specification	**Summary of a comparison between the decision problem stated in the NICE scope and addressed in the CS**
Population People with locally advanced or metastatic NSCLC whose disease has progressed after chemotherapy	Adult patients with locally advanced or metastatic NSCLC after prior (platinum containing) chemotherapy The company recognises that targeted therapies are likely to be the preferred second-line option in patients with confirmed or suspected mutations and these are excluded from the company's economic analysis (CS, Figure 3, footnote)
Intervention Atezolizumab	Atezolizumab
Comparators • Docetaxel • Nintedanib with docetaxel (for people with adenocarcinoma histology) • Nivolumab (subject to ongoing NICE appraisal) • Pembrolizumab (PD-L1-expressing tumours) • Best supportive care	**Direct evidence** • Docetaxel The OAK[22,23] and POPLAR[24,25] trials were designed to compare the clinical effectiveness of atezolizumab versus docetaxel **Indirect evidence** • Nintedanib+docetaxel The company used an indirect comparison to compare the effectiveness of treatment with atezolizumab versus nintedanib+docetaxel in the ITT populations of the OAK and POPLAR trials versus the whole population participating in the LUME-Lung 1[26] trial (nintedanib+docetaxel vs docetaxel). However, nintedanib+docetaxel is only recommended by NICE for the treatment of patients with adenocarcinoma histology **No evidence** • Nivolumab At the time of submission, nivolumab had not been recommended by NICE for the treatment of any patients with NSCLC and, therefore, could not be considered a standard of care • Pembrolizumab The company considered that: ○ the marketing authorisation for pembrolizumab is only for people with PD-L1 positive NSCLC, i.e. a sub-set of the population described in the anticipated marketing authorisation for atezolizumab ○ accurate comparison between the two treatments is not possible as different, non-comparable, PD-L1 expression tests have been used in the pembrolizumab and atezolizumab studies ○ at the time of the submission, pembrolizumab had only recently been recommended by NICE and, therefore, it was unlikely that, at this time, it would represent a standard of care • Best supportive care Clinical advice to the company is that patients eligible for treatment with atezolizumab would be considered fit enough for other treatment The ERG agrees with the company's arguments for not including nivolumab and best supportive care as comparators but considers that pembrolizumab is a relevant comparator

Final scope issued by NICE Parameter and specification	**Summary of a comparison between the decision problem stated in the NICE scope and addressed in the CS**
Outcomes OS PFS ORR AEs HRQoL	The company has presented results for all outcomes detailed in the final scope issued by NICE
Economic analysis The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per QALY	Cost effectiveness has been assessed using ICERs per QALY gained
If appropriate, the appraisal should include consideration of the costs and implications of additional testing for biological markers, but will not make recommendations on specific diagnostic tests or devices	Not applicable – the anticipated marketing authorisation for atezolizumab is the whole population of patients with NSCLC
The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared	The model time horizon is 25 years
Costs will be considered from an NHS and Personal Social Services perspective	Costs have been considered from an NHS perspective
The availability of any patient access schemes (PAS) for the intervention or comparator technologies will be taken into account	Details relating to the PAS for atezolizumab have been provided in a confidential appendix that formed part of the CS. The PAS for nintedanib is confidential and, therefore, not known to the company. However, the ERG has re-run the company's base case analysis using the PAS price for nintedanib (see confidential appendix of this ERG report for results)
Subgroups to be considered If the evidence allows, consideration will be given to subgroups based on biological markers	The company states that clinical benefit is observed in all subgroups of patients with NSCLC who are treated with atezolizumab and that, as such, no analyses have been conducted on restricted populations
Special considerations None identified	None identified

AE=adverse event; CS=company submission; ERG=evidence review group; HRQoL=health-related quality of life; ICER=incremental cost effectiveness ratio; ITT=intention to treat; NICE=National Institute for Health and Care Excellence; NSCLC=non-small cell lung cancer; ORR=objective response rate; OS=overall survival; PAS=patient access scheme; PD-L1=programmed death ligand 1; PFS=progression-free survival; PSS=Personal Social Services; QALY=quality adjusted life year
Source: CS, Table 1 and ERG assessment

3.1 Population

The population described in the final scope issued by NICE is people with locally advanced or metastatic NSCLC whose disease has progressed after chemotherapy. The population discussed in the CS is the population recruited to the OAK[22,23] trial and POPLAR[24,25] trial. The ERG notes that these two populations are identical, except that the recruitment criteria for the trials specify that the population should be adults who had received a maximum of two previous chemotherapies and that prior chemotherapy should have been platinum containing.

Clinical advice to the ERG is that the clinical evidence submitted by the company is relevant to NHS patients with NSCLC whose disease has progressed following chemotherapy, except that the OAK and POPLAR trial populations are younger and fitter than those likely to be treated in the NHS.

3.2 Intervention

The intervention specified in the final scope issued by NICE, and discussed in the CS, is atezolizumab. Atezolizumab does not currently have a UK marketing authorisation. The company made an application on 20th April 2016 and anticipates receiving the Committee for Medicinal Products for Human Use (CHMP) opinion [redacted] The application is for the treatment of adult patients with locally advanced or metastatic NSCLC after prior chemotherapy.

Atezolizumab is a monoclonal antibody that binds to and inactivates a protein called PD-L1 on the surface of tumour cells (TCs) and tumour-infiltrating immune cells (ICs), inhibiting the binding to PD-1 and B7.1.[27] The company explains that atezolizumab differs from anti-PD-1 antibodies already approved for the treatment of NSCLC as it results in the dual blockade of PD-1 and B7.1 while leaving the PD-1/programmed death-ligand 2 (PD-L2) interaction intact, thereby potentially preserving peripheral immune homeostasis.[28]

The company has presented evidence for the effectiveness of atezolizumab from two trials (the OAK trial and the POPLAR trial). The randomisation protocol for both trials included PD-L1 status as a stratification factor. However, the company states that phase I data did not demonstrate a clear relationship between PD-L1 expression and response to atezolizumab (CS, p16). In view of the pathway that is blocked by treatment with atezolizumab, as part of the clarification process, the ERG asked the company to explain why atezolizumab might be effective in treating tumours that do not express PD-L1. The company provided three possible reasons (Box 3).

Box 3 Possible reasons why atezolizumab might be effective in treating tumours that are PD-L1 negative

The first is the biological hypothesis that atezolizumab increases anticancer immunity through enhanced priming of new anticancer immune responses.[23] PD-L1 is expressed on T cells and antigen presenting cells (APCs) present in the lymph nodes. Here it binds to B7.1, which is also expressed on T cells and APCs; as with PD-1 to PD-L1 interactions, this interaction can downregulate T cell activity and subsequent immune responses. Inhibition of this interaction in the lymph node environment may therefore prevent this downregulation and stimulate an immune response in tumours that are PD-L1 negative.[29,30]

The second reason is that a PD-L1 negative tumour is defined as PD-L1 expression on less than 1% expression of tumour cells (TCs) and tumour-infiltrating immune cells (ICs), i.e. TC0 and IC0. Consequently, there could still be low levels of PD-L1 expression within the tumour environment that are sufficient to induce anti-tumoural immune responses following treatment with atezolizumab.

Finally, PD-L1 expression in tumours may be heterogeneous and variable over time in a subset of tumours. This means that a biopsies taken from different areas of a tumour may show different levels of PD-L1 expression, or that the PD-L1 expression level may have changed since the biopsy was taken and may not reflect the current PD-L1 status.[31-34]

APC=antigen presenting cells; IC=immune cells PD-1=programmed death-1; PD-L1=programmed death-ligand 1; T=tumour; TC=tumour cells
Source: Company clarification letter response

The ERG highlights that the company has provided OAK trial results comparing OS for patients treated with atezolizumab versus docetaxel (CS, Section 4.7) for patients with ≥1% (TC1/2/3 or IC1/2/3) PD-L1 expression (HR 0.74, 95% CI: 0.58 to 0.93; p=0.0102). Furthermore, OAK trial OS results by level of PD-L1 expression are in the public domain.[23] The ERG, therefore, considers that the company should have presented clinical and cost effectiveness results within the CS, or justified their absence.

Atezolizumab is currently being assessed by NICE for the treatment of locally advanced or metastatic urothelial carcinoma[35] (company submission: 18 January 2017) and is already available in the UK for patients with this condition under the Early Access to Medicines Scheme (EAMS). In October 2016 the US Food and Drug Administration (FDA)[36] approved atezolizumab for the treatment of patients with metastatic NSCLC whose disease progressed during or following platinum-containing chemotherapy. Regulatory approval for this indication has also been received in Kuwait and South Korea.

The treatment regimen for atezolizumab is a flat dose of 1200mg intravenous infusion administered in a hospital setting every 3 weeks (Q3W). The initial dose must be administered over 60 minutes. If the first infusion is tolerated, all subsequent infusions may be administered over 30 minutes. It is stated within the draft Summary of Product Characteristics[1] (SmPC) that patients should be treated with atezolizumab until loss of clinical benefit or unmanageable toxicity.

3.3 Comparators

The comparators specified in the final scope issued by NICE are docetaxel, nintedanib+docetaxel, pembrolizumab, nivolumab and best supportive care (BSC).

3.3.1 Included comparators

Docetaxel

Direct evidence is available for the comparison of the effectiveness of atezolizumab versus docetaxel from the OAK and POPLAR trials. The company states that docetaxel monotherapy is regarded as the standard of care in the NHS for patients with locally advanced or metastatic NSCLC in whom relapse has occurred after previous chemotherapy. The ERG, however, notes that nintedanib+docetaxel is a standard of care for the subgroup of patients with NSCLC of adenocarcinoma histology.

Docetaxel is indicated for the treatment of patients with locally advanced or metastatic NSCLC after failure of prior chemotherapy.[37] Within the OAK and POPLAR trials, docetaxel 75mg/m^2 is administered intravenously on day 1 of each 21-day cycle until disease progression per standard RECIST v1.1 or unacceptable toxicity. Clinical advice to the ERG is that, within the NHS, patients typically only receive between four and six cycles of treatment.

Nintedanib+docetaxel

Treatment with nintedanib+docetaxel is recommended by NICE as an option for treating locally advanced, metastatic or locally recurrent NSCLC of adenocarcinoma histology that has progressed after first-line chemotherapy.[38]

In the absence of direct evidence to allow a comparison of the effectiveness of treatment with atezolizumab versus nintedanib+docetaxel, the company undertook an indirect treatment comparison (ITC). The company states that, to allow for a like-with-like comparison versus atezolizumab according to its anticipated licence, the ITC used data from the intention to treat (ITT) populations of the OAK, POPLAR and LUME-Lung-1[26] trials. However, the ERG notes that treatment with nintedanib+docetaxel is only licensed[39] (and recommended by NICE[40]) for the treatment of patients with adenocarcinoma and that results from the LUME-Lung 1 trial show that treatment with nintedanib+docetaxel results in better PFS and OS outcomes for the population with adenocarcinoma histology than for the whole LUME-Lung 1 trial population. The ERG, therefore, considers that the relevance of the indirect comparison undertaken by the company is limited as the comparison undertaken by the company underestimates the efficacy of nintedanib+docetaxel in the NHS patient population for which it is recommended. A full critique of the company's ITC can be found in Section 4.6 of this report.

3.3.2 Excluded comparators

Nivolumab

The ERG agrees with the company that nivolumab is not a relevant comparator. This is because, at the time the CS was sent to NICE, nivolumab had not been recommended by NICE as a treatment for the population under consideration in this appraisal. However, the ERG notes that at the time of submitting the ERG report to NICE (April 2017), two STAs considering the use of nivolumab for the treatment of locally advanced or metastatic NSCLC were on-going, one for patients with squamous disease (ID811)[41] and the other for patients with non-squamous disease (ID900).[42]

Pembrolizumab

The ERG notes that, in the CS (Figure 2) the company has placed pembrolizumab in the same position in the treatment pathway as atezolizumab. In addition, the company highlights that the European marketing authorisation for pembrolizumab is similar to the anticipated European marketing authorisation for atezolizumab in that both are targeted at adults with locally advanced or metastatic NSCLC who have received at least one prior chemotherapy treatment. However, pembrolizumab is only recommended by NICE[15] for the treatment of people with PD-L1 positive NSCLC. The company considers that this discrepancy means that a comparison of the effectiveness of treatment with atezolizumab versus pembrolizumab would not be meaningful as the relative clinical benefits of treatment with pembrolizumab would be overestimated. The ERG considers that, as results from the OAK trial (CS, p81) show that treatment with atezolizumab versus docetaxel was associated with a similar improvement in OS in the ITT population (hazard ratio [HR] 0.73, 95% CI: 062 to 0.87; p=0.0003) and in patients with ≥1% (TC1/2/3 or IC1/2/3) PD-L1 expression (HR 0.74, 95% CI: 0.58 to 0.93; p=0.0102), the company's argument is not compelling.

The company also highlights that, in studies of the effectiveness of atezolizumab and pembrolizumab, the tools used to assess PD-L1 expression differ significantly, both in how expression is measured (atezolizumab: TC and IC, pembrolizumab: TC only) and also in terms of which patients are considered positive expressors. The company considers that this means that even a subgroup analysis of PD-L1 positive patients would not be appropriate. As part of the clarification process the ERG requested an explanation from the company as to how the two tests differ. The response provided by the company focused on the differences in terms of detection antibody, immunohistochemistry (IHC) platform, cell types scored (TC and IC versus TC) and cut-off points. However, the company states (CS, p49) that analyses of data from the OAK trial showed a statistically significant and clinically meaningful improvement in OS when treatment with atezolizumab was compared with docetaxel in patients with ≥1% PD-

L1 expression (HR 0.74, 95% CI: 0.58 to 0.93, p=0.012) and, in the clarification response, the company explained that TC1/2/3 or IC1/2/3 was defined as PD-L1 expression of 1% or more of TCs or ICs. It, therefore, appears that the company considers that it is possible to compare the output measures from the two tests used to determine level of PD-L1 expression

In addition, the company highlights that pembrolizumab has only recently been recommended by NICE[15] for use in patients with NSCLC (guidance issued 11th January 2017) and, therefore, considers it unlikely to represent a standard of care at this time (February 2017). The ERG considers that while there may not have been wide use of pembrolizumab within the NHS at the time of the CS, it is likely to have become an established option by the time the final appraisal determination (FAD) for this appraisal of atezolizumab is published. The ERG, therefore, does not find this line of argument compelling.

The ERG considers that it is difficult accept the company's argument that a difference in marketing authorisations/study populations is a barrier to undertaking a comparison between atezolizumab and pembrolizumab. In addition, the ERG highlights that the company has included nintedanib+docetaxel as a comparator even though this treatment is only recommended by NICE for the population with adenocarcinoma histology.

The ERG considers that pembrolizumab is an appropriate comparator, but only for the population for which it is currently recommended by NICE, i.e., patients whose tumours express PD-L1 (≥1%) and who have had at least one prior chemotherapy regimen (and targeted treatment if they have an EGFR- or ALK-positive tumour).

Best supportive care

Clinical advice to the company is supported by clinical advice to the ERG, namely that patients who are eligible for treatment with atezolizumab would be fit enough for other treatments and, therefore, BSC is not an appropriate comparator.

Erlotinib and crizotinib

Erlotinib and crizotinib were not included in the final scope issued by NICE but they are included in the company's treatment pathway algorithm (CS, p46). However, clinical advice to the company is that targeted therapy treatment options are likely to be preferred over immunotherapy in patients with confirmed EGFR or ALK mutations. The ERG considers that as the prevalence of EGFR and ALK oncogenic alterations are low (EGFR: 10% to 28%[43] of patients with NSCLC; ALK: approximately 3.4%[12] of the non-squamous population), and as EGFR and ALK testing is now routinely carried out in the NHS, very few patients are likely to receive erlotinib or crizotinib after prior chemotherapy. In addition, there is no evidence to allow

a direct comparison of the effectiveness of atezolizumab with either erlotinib or crizotinib. The ERG, therefore, agrees with the company that it was appropriate not to include either erlotinib or crizotinib as comparators.

3.4 Outcomes

Clinical evidence from the OAK and POPLAR trials is reported for all five outcomes specified in the final scope issued by NICE: (investigator assessed) progression-free survival (PFS), overall survival (OS), objective response rate (ORR), i.e. proportion of patients achieving best overall response of partial response (PR) or complete response (CR), adverse events (AEs) of treatment and health-related quality of life (HRQoL).

The ERG notes that duration of response (DOR; interval between first documented objective response [CR or PR] and first documented progressive disease) was also a secondary endpoint of both the OAK and POPLAR trials.

3.5 Economic analysis

As specified in the final scope issued by NICE, the cost effectiveness of treatments was expressed in terms of the incremental cost per quality adjusted life year (QALY) gained. Outcomes were assessed over a 25-year time-period (equivalent to a lifetime horizon) and costs were considered from an NHS perspective.

3.6 Subgroups

It is specified within the final scope issued by NICE that, if evidence allows, consideration will be given to subgroups based on biological markers. Within the CS, the company has provided OS results from the OAK trial by demographic (sex and age) and baseline prognostic characteristics:

- Eastern Cooperative Oncology Group (ECOG) PS
- prior lines of chemotherapy
- smoking history
- prior metastases (OAK trial: central nervous system metastases, POPLAR trial: liver metastases and bone metastases)
- mutational status (Kirsten rat sarcoma [KRAS] and EGFR)
- histology (squamous and non-squamous).

Efficacy was also evaluated by level of PD-L1 expression but results are only presented in the CS from the OAK trial for the comparison of the effectiveness of treatment with atezolizumab versus docetaxel for patients with no measurable PD-L1 expression (TC0/IC0) and for patients with ≥1% PD-L1 expression (CS, p16). However, results have been presented for four PD-L1

subgroups (TC3 or IC3, TC2/3 or IC2/3, TC1/2/3 or IC1/2/3 and TC0 and IC0) in a published paper.[23]

3.7 Other considerations

The company did not identify any equity or equality issues. Details relating to the patient access scheme (PAS) for atezolizumab have been provided by the company in a confidential appendix that formed part of the CS. A PAS is also in place for nintedanib. Both PAS prices are confidential and, therefore, the PAS price for nintedanib is not known to the company. However, the ERG has re-run the company's base case analysis using the PAS price for nintedanib (see confidential appendix to this ERG report for results).

4 CLINICAL EFFECTIVENESS

4.1 Systematic review methods

The company carried out a systematic search of the literature in June 2016 to identify phase II-IV randomised controlled trials (RCTs) investigating the efficacy and safety of pharmacological interventions for second- and further-line treatment for locally advanced or metastatic NSCLC. The purpose of the review was to identify studies to include in the company's ITC, which was conducted to support pricing and reimbursement submissions across all markets, and included comparators not listed in the final scope issued by NICE.

The data sources searched and the time spans for the searches are provided in Table 4, while a summary of, and ERG comments on, the review methods used by the company are presented in (Table 5).

Table 4 Data sources for the clinical systematic review

Search strategy component	Source	Search date range	
		Start	End
Electronic database searches	EMBASE	1988	Not provided
	MEDLINE	1946	
	MEDLINE In-Process	1946	
	Cochrane Central Library of Controlled Trials (CENTRAL)	January 2012	June 2016
	Cochrane Database of Systematic Reviews (CDSR)		
Congress proceedings	American Society of Clinical Oncology (ASCO) European Society for Medical Oncology (ESMO) International Association for the Study of Lung Cancer (IASLC)/World Conference on Lung Cancer (WCLC) International Lung Cancer Congress (ILCC) European Lung Cancer Conference (ELCC) British Thoracic Oncology Group (BTOG)	1 January 2013	17 June 2016
Clinical trial registries	ClinicalTrials.gov	1 January 2012	21 July 2016
	WHO's meta-registry 'International Clinical Trials Registry Platform Search Portal' (ICTRP)		
	EU Clinical Trial Registry	1 January 2012	30 August 2016

Source: CS, pp50-51

Table 5 Summary of, and ERG comment on, company systematic review methods

Review method	Results	ERG comment
Searching		
Sources searched: • Electronic databases • Congress proceedings • Clinical trial registries	22,502 studies	• The searches were completed in summer 2016, meaning that there is a risk that some relevant studies will not have been included in the search results • The search terms were relevant but could have been expanded regarding the search terms relating to cancer • The searches only included population terms and not indication terms
Formal eligibility criteria		
Two analysts independently assessed study eligibility based on the criteria presented in Table 10 of the CS (pp52-54)	303 studies	• Use of two independent assessors improves the quality of reviews • The high number of results from the initial search tests the concentration of reviewers
Additional eligibility criteria		
1. Although no language restrictions were included in the search strategies, the company excluded Asian language publications at the data extraction stage due to the extra complexity associated with translating these articles and the limited relevant additional data these would provide 2. Based on input from clinical experts, studies that compared investigational interventions and interventions that have not yet been labelled/approved for treating NSCLC in Europe or the USA were excluded 3. Based on further expert input, erlotinib combination arms were also excluded	38 Asian articles were excluded, leaving 265 publications from 184 different studies 49 RCTs 19 RCTs reporting 16 active treatments	• The need to employ two further sets of eligibility criteria highlights the very un-focused nature of the original searches undertaken by the company
Quality assessment		
The company conducted a quality assessment exercise using the minimum criteria recommended by NICE in the company submission template. The company applied guidance from the Cochrane Handbook of Systematic Reviews to assess each of the criteria. The results of the company assessment of the OAK and the POPLAR trials are presented in the CS. The results of the assessment of the RCTs included in the company's ITC are presented in Appendix 4 of the CS.		

CS=company submission; ERG=Evidence Review Group; NSCLC=non-small cell lung cancer; RCT=randomised controlled trial
Source: CS, pp55-56 and pp96-97

4.1.1 Evidence synthesis

The company presents direct evidence to support the clinical efficacy of atezolizumab from two RCTs (the OAK trial and the POPLAR trial). The CS includes a narrative description of both of these trials. No evidence synthesis was undertaken.

4.2 ERG critique of direct clinical effectiveness evidence

4.2.1 Identified trials

Key trials: the OAK and POPLAR trials

The company presents evidence for the clinical effectiveness of atezolizumab from the OAK (phase III) and POPLAR (phase II) trials. Both are open-label multicentre RCTs that were designed to investigate the efficacy and safety of atezolizumab versus docetaxel in patients with locally advanced or metastatic NSCLC whose disease had progressed during or following a platinum-containing regimen. Patients were randomised to receive either: atezolizumab 1200mg Q3W or docetaxel 75mg/m^2 Q3W. Details relevant to the OAK and POPLAR trials are reported in the CS, in the trial clinical study reports (CSRs[22,24]) and in published papers.[23,25] Details of these trials have also been presented at a number of conferences.[44-48]

Other trials

The clinical development programme of atezolizumab in NSCLC included two single-arm phase II studies, BIRCH[49] (study GO28754) and FIR[50] (study GO28625). The company states that these trials have not been discussed in the CS as the patient populations in both trials had PD-L1 positive disease and, therefore, are not relevant to this appraisal.

The ERG is not aware of any trials that directly compare the clinical effectiveness of atezolizumab with any of the comparators, other than docetaxel, as per the final scope issued by NICE.

4.2.2 Key characteristics of the OAK and POPLAR trials

The key characteristics of the OAK and POPLAR trials are provided in the CS (pp58-77) and are summarised in Table 6.

Eligibility criteria for entry into the OAK and POPLAR trials were provided by the company (CS, pp61-63). Clinical advice to the ERG is that the eligibility criteria are reasonable. The OAK and POPLAR trials were conducted internationally (in 31 and 13 countries respectively). The OAK trial included eight UK sites (31 patients) and the POPLAR trial included four UK sites (11 patients). Patients were randomly assigned (1:1) to receive either atezolizumab or docetaxel using an interactive voice or web response system. Randomisation was stratified by previous lines of chemotherapy (one versus two) and histology (non-squamous versus squamous). In addition, randomisation was stratified by PD-L1 IC status (four categories: IC0, IC1, IC2, and IC3). The ERG notes that an exploratory objective of the OAK and POPLAR trials was the evaluation of the relationship between PD-L1 expression and measures of efficacy.

Table 6 Key characteristics of the OAK and POPLAR trials

	OAK trial	POPLAR trial
Location	International (31 countries, 194 centres, including 8 in the UK [31 patients])	International (13 countries, 61 centres, including 4 in the UK [11 patients])
Design	Randomised (1:1), phase III, open-label	Randomised (1:1), phase II, open-label
Population	Primary population: a total of 825 patients were randomised, 425 to the atezolizumab arm and 425 to the docetaxel arm Secondary population: following the interim analysis of data from the POPLAR trial, the population size was increased to ensure at least 220 patients with PD-L1 TC3 or IC3 (assuming a 20% prevalence) were enrolled. In total, 1225 patients were randomised (614 to the atezolizumab arm and 611 to the docetaxel arm)	A total of 287 patients were randomised, 143 patients to the docetaxel arm and 144 patients to the atezolizumab arm
Intervention	Atezolizumab (1200mg Q3W) was administered as long as patients experienced a clinical benefit as assessed by an investigator in the absence of unacceptable toxicity or symptomatic deterioration attributed to disease progression	Atezolizumab (1200mg Q3W) was administered as long as patients experienced a clinical benefit as assessed by an investigator in the absence of unacceptable toxicity or symptomatic deterioration attributed to disease progression
Comparator	Docetaxel (75mg/m^2 Q3W) was administered until disease progression or unacceptable toxicity	Docetaxel (75mg/m^2 Q3W) was administered until disease progression or unacceptable toxicity
Primary outcome	Co-primary: • OS in the ITT population • OS in patients with ≥1% PD-L1 expression (TC1/2/3, IC1/2/3)	OS
Secondary outcomes	PFS, ORR and DOR	PFS, ORR and DOR
Safety endpoints	Safety and tolerability of treatment with atezolizumab compared with docetaxel	Safety and tolerability of treatment with atezolizumab compared with docetaxel
Patient reported outcomes	Data collected using: • EQ-5D-3L tool • EORTC-QLC-C30 and its lung cancer module (LC13)	Data collected using EORTC-QLC-C30 and its lung cancer module (LC13)
Duration of study	• First patient randomised: 11 March 2014 • Last patient randomised in the primary population: 28 November 2014 • Last patient randomised in the secondary population: 29 April 2015	• First patient randomised: 5 August 2013 • Last patient randomised: 31 March 2014
Data analyses	Primary analysis: clinical cut-off 7 July 2016	Interim analysis: clinical cut-off 30 January 2015 Primary analysis: clinical cut-off 8 May 2015 Updated efficacy analysis (OS and DOR): clinical cut-off 1 December 2015
Median duration of follow-up (primary analysis)	Atezolizumab: 21.4 months Docetaxel: 21.3 months	Atezolizumab: 14.8 months Docetaxel: 15.7 months

DOR=duration of response; EORTC-QLC-C30=European Organisation for Research and Treatment of Cancer-Quality of Life Questionnaire-Core 30; EQ-5D-3L=EuroQol-5 Dimensions-3 Levels; IC=immune cell; ITT=intention to treat; OS=overall survival; ORR=objective response rate; PD-L1=programmed death-ligand 1; PFS=progression-free survival; PR=partial response; Q3W=every 3 weeks; TC=tumour cell
Source: CS, Section 4.3

4.2.3 Characteristics of patients enrolled in the OAK and POPLAR trials

The key baseline characteristics of patients included in the OAK and POPLAR trials are listed in Table 7. The company reports (CS, p75) that the patients enrolled in the OAK trial were predominately white (70%), male (61%) with a median age of 64 years (range 33.0-85.0 years) and an ECOG PS of 1 (63%). In addition, the majority of patients had a history of tobacco use: 67% were previous smokers and 15.0% were current smokers. Similarly, the company reports (CS, p79) that the patients enrolled in the POPLAR trial were predominantly white (78.7%), male (58.9%) with a median age of 62 years (range 36-84 years) and an ECOG PS of 1 (68.0%). The majority of patients in this trial also had a history of tobacco use: 64.5% were previous smokers and 16.0% were current smokers.

The ERG considers that patients' baseline characteristics are generally well balanced across the treatment arms. In addition, clinical advice to the ERG is that the patients recruited to the two trials can be considered to be broadly representative of patients with advanced NSCLC, treated in the NHS, albeit slightly younger and fitter.

Table 7 Demographic and baseline characteristics (ITT populations)

	OAK trial		POPLAR trial	
	Atezolizumab	**Docetaxel**	**Atezolizumab**	**Docetaxel**
N	425	425	144	143
Male n (%)	261 (61)	259 (61)	93 (64.6)	76 (53.1)
Mean months from initial diagnosis to randomisation (sd)	21.04 (21.45)	20.06 (23.0)	16.96 (15.52)	20.27 (19.66)
Age				
Age, years, median (range)	63.0 (33.0 to 82.0)	64.0 (34.0 to 85.0)	62.0	62.0
<65 years n (%)	235 (55)	218 (51)	87 (60.4)	87 (60.8)
≥65 years n (%)	190 (45)	207 (49)	57 (39.6)	56 (39.2)
ECOG PS n (%)			n=142	n=142
0	155 (36)	160 (38)	46 (32.4)	45 (31.7)
1	270 (64)	265 (62)	96 (67.6)	97 (68.3)
Histology				
Non-squamous	313 (74)	315 (74)	95 (66.0)	95 (66.4)
Squamous	112 (26)	110 (26)	49 (34.0)	48 (33.6)
Current disease status (%)				
Locally advanced	29 (7)	19 (5)	8 (5.6)	5 (3.5)
Metastatic	396 (93)	406 (95)	136 (94.4)	138 (96.5)
Number of prior therapies n (%)				
1	320 (75)	320 (75)	93 (64.6)	96 (67.1)
2	105 (25)	105 (25)	51 (35.4)	47 (32.9)
Smoking status n (%)				
Never	84 (20)	72 (17)	27 (18.8)	29 (20.3)
Current	59 (14)	67 (16)	25 (17.4)	21 (14.7)
Previous	282 (66)	286 (67)	92 (63.9)	93 (65.0)
Metastases				
Number of metastatic sites at enrolment, mean (sd)	2.89 (1.43)	2.97 (1.32)	2.97 (1.38)	3.1 (1.39)
Confirmed metastases at enrolment n (%)				
Liver	83 (20)	94 (22)	33 (22.9)	33 (23.1)
Bone	135 (32)	133 (31)	35 (24.3)	46 (32.2)
Brain	38 (9)	47 (11)	8 (5.6)	15 (10.5)
Lung	386 (91)	391 (92)	132 (91.7)	125 (87.4)
Pleural effusion	84 (20)	96 (23)	41 (28.5)	27 (18.9)
Lymph nodes	277 (65)	291 (66)		
PD-L1 expression				
TC3 or IC3, n (%)	72 (16.9)	65 (15.3)	24 (16.7)	23 (16.1)
TC2/3 or IC2/3, n (%)	129 (30.4)	136 (32.0)	50 (34.7)	55 (38.5)
TC1/2/3 or IC1/2/3, n (%)	241 (56.7)	222 (52.2)	93 (64.6)	102 (71.3)

ECOG PS=Eastern Cooperative Oncology Group performance score; IC=immune cell; PD-L1=programmed death-ligand 1; sd=standard deviation; TC=tumour cell
Source: CS, Table 24 and Table 26

4.2.4 Statistical approach adopted in the OAK and POPLAR trials

In this section, the ERG provides a description and critique of the statistical approaches used to analyse data collected during the OAK and POPLAR trials that relate to the outcomes stipulated in the final scope issued by NICE. Information relevant to the statistical approach taken by the company has been extracted from the CSRs, the trial protocols,[51,52] the trial statistical analysis plans (TSAPs),[53,54] which included a modification plan for OAK and the CS.

Determination of sample size and trial design

The original sample size of the OAK trial was calculated as 850 patients in an ITT population so that approximately 255 PD-L1 IC2/3 patients and 425 PD-L1 IC1/2/3 patients would be enrolled. Following interim analysis of the POPLAR trial (clinical cut-off date 30th January 2015) and additional data from PCD4989g[55] (phase I trial) and FIR[50] (single-arm, phase II trial), the TSAP was modified according to a pre-specified modification plan (OAK TSAP, p24 and Appendix 4). The sample size of the OAK trial was increased to approximately 1100 patients (up to a maximum of 1300) to ensure at least 220 patients with PD-L1 TC3 or IC3 status, assuming a 20% prevalence of the TC3 or IC3 subgroup, were recruited. The final enrolment in the OAK trial was 1225 patients.

Subsequently, the results from the primary analysis of the POPLAR trial (POPLAR CSR, Table 21, data cut-off 8th May 2015) showed that clinical efficacy was observed in all PD-L1 subgroups including patients with PD-L1 negative NSCLC. Therefore, assuming that the OAK trial would also show clinical efficacy in all defined subgroups based on PD-L1 expression (herein referred to as PD-L1 subgroups), the OAK trial would be fully statistically powered for OS evaluation in an ITT population with fewer than 1225 patients.

Therefore, prior to unblinding the data, the OAK trial TSAP was modified again on 28th January 2016 (OAK TSAP, p24 and Appendix 4) to conduct the analysis of OS in the OAK trial on the primary population (1°P) of the first 850 randomised patients. The data cut-off of the 1°P would occur when approximately 595 deaths had occurred, which would correspond to an estimated 384 deaths in the TC 1/2/3 or IC 1/2/3 subgroup (TSAP, p14).

If the null hypothesis for OS in the 1°P was rejected, an analysis of the secondary population (2°P) of 1225 randomised ITT patients would be performed at the secondary analysis time (OAK TSAP, Appendix 4; Section 1.1). To control the type I error rate in the evaluation of OS in the 1°P and 2°P, alpha was split between the ITT population and the TC1/2/3 or IC1/2/3 subgroup (OAK TSAP, Appendix 4; Table 2).

Treatment crossover from the docetaxel arm to the atezolizumab arm was not originally permitted in the OAK trial; however, it was subsequently allowed following the primary analysis time (CS, Section 4.7). No efficacy results for the 2°P are available at the time of writing this ERG report and the company informed the ERG that [redacted] The ERG notes that any crossover between treatment arms will potentially confound the results of the planned analysis of the 2°P.

The ERG is satisfied that the modification plan for the sample size calculation was pre-specified (final date: 21st November 2013). The ERG is also satisfied that the modifications were made before the date of primary analysis of the OAK trial (data cut-off: 7th July 2016) and were, therefore, unlikely to have been driven by the results of the trial.

The POPLAR trial was designed to enrol a minimum of 54 patients with PD-L1 IHC 2/3 NSCLC, with a maximum of 300 patients enrolled in the case that the prevalence of PD-L1 IHC 2/3 NSCLC was lower than 18%. The trial was expected to enrol 285 patients, including 55 patients with PD-L1 IHC 2/3 NSCLC. The ERG is satisfied that this sample size calculation was pre-specified (POPLAR TSAP, p7-8).

Design assumptions of the OAK and POPLAR trials are summarised in Table 8.

Table 8 OAK and POPLAR trial design assumptions

OAK trial	POPLAR trial
Event times are exponentially distributed A 7.5% 24-month dropout rate assumed for both treatment arms Over 95% power for the primary analysis of OS in the ITT population Median survival of 10 months in the docetaxel arm 65% prevalence rate for TC1/2/3 or IC1/2/3	Event times are exponentially distributed Median PFS in the control arm is 3 months Median OS in the control arm is 8 months Patients are enrolled over 8 months

IC=tumour-infiltrating immune cell; ITT=intention-to-treat; OS=overall survival, PFS=progression-free survival; TC=tumour cell
Source: CS, adapted from Section 4.4

Outcomes and analysis approach in the OAK and POPLAR trials

The primary therapeutic aims of the OAK and POPLAR trials were to reduce tumour burden, delay disease progression and ultimately prolong life. Therefore, the primary endpoint of the two trials was OS, selected to explore the impact of treatment with atezolizumab in reaching these aims. In the OAK trial, OS was measured as a co-primary endpoint in both the ITT population and in patients with ≥1% PD-L1 expression.

The company argues that PFS is a less suitable endpoint to assess the activity of immunotherapies, but includes PFS as a secondary outcome in both trials.

Definitions and methods of statistical analysis for OS and PFS are provided in Table 9. The ERG is satisfied that the analysis method for each of these efficacy outcomes was pre-specified in the TSAPs, and that all results are reported fully in the CSRs.

The ERG notes that the log-rank and Cox regression methodology employed for the analysis of OS and PFS in both the OAK and POPLAR trials require the assumption of proportional hazards (PH) for the interpretation of estimated log-rank p-values and HRs. The company demonstrates that the PH assumption does not hold for OS and PFS in either the OAK trial (CS, Figure 20 and Figure 21) or the POPLAR trial (company response to ERG clarification letter, reproduced in Section 10.1]). This violation of the PH assumption is taken into account in the statistical approach used in the ITC and also in the approach to cost effectiveness analysis (CS, Section 5.3).

The ERG notes that methodology for the analysis of OS and PFS was pre-defined in both TSAPs before the data cut-off dates in the OAK and POPLAR trials, and that violation of the PH assumption could not have been known when the TSAPs were written. However, use of HRs to summarise treatment effect of OS and PFS is not appropriate in the absence of PH; therefore, the HRs reported in the Section 4 of the CS must be interpreted with caution.

Table 9 Definition and analysis method for key efficacy outcomes (OAK and POPLAR trials)

Outcome	Outcome definition	Censoring definition	Statistical analysis[a]
Primary efficacy outcome			
OS	Time from the date of randomisation to the date of death due to any cause OAK trial: further defined in patients with a ≥1% PD-L1 expression (TC1/2/3 or IC1/2/3)	Date patient last known to be alive or at date of randomisation (plus 1 day for those without baseline information)	OAK trial: K-M methodology, log-rank test, and Cox regression, stratified in the 1°P and TC1/2/3 or IC1/2/3 subgroup POPLAR trial: K-M methodology and stratified log-rank test for ITT, unstratified log-rank test for biomarker subsets, Cox regression, stratified for ITT and unstratified for biomarker subsets
Secondary efficacy outcome			
PFS	Interval between date of randomisation and date of first documented PD per RECIST v1.1 or death	Last tumour assessment for those without PD and alive or at date of randomisation (plus 1 day for those without post-baseline assessments)	OAK trial: K-M methodology, Cox regression, stratified in the 1°P ITT population and TC1/2/3 or IC1/2/3 subgroup, unstratified for all other subgroups POPLAR trial: K-M methodology, Cox regression, stratified for ITT and unstratified for biomarker subsets

[a] The stratification factors used in analysis in both trials were those used in randomisation i.e., tumour PD-L1 status (four categories of PD-L1 IC expression), the number of prior lines of therapy (1 vs 2), and histology (non-squamous vs squamous)
1°P=primary population; IC=tumour-infiltrating immune cell; ITT=intention-to-treat; K-M=Kaplan-Meier; OS=overall survival, PD=progressive disease; PD-L1=programmed death-ligand 1; PFS=progression-free survival; RECIST=response evaluation criteria in solid tumours; TC=tumour cell
Source: CS, adapted from Table 19, Table 21 and Table 22

Objective response rate (ORR) and DOR were also secondary efficacy outcomes in both trials. For completeness, the definitions and methods of analysis of ORR and DOR are described in Section 10.1). Patient-reported outcomes (PROs) and safety endpoints were also measured in both trials. Further details of these outcomes are described in Section 4.4 and Section 4.5 respectively.

All primary and secondary outcomes measured in both trials were investigator-assessed. In their response to the clarification letter, the company confirmed that no blinded independent central review of any endpoints had been explored in either the OAK or POPLAR trials.

Analysis populations

The populations used for analyses of different outcomes of the OAK and POPLAR trials are summarised in Table 10. The ERG is satisfied that these populations were pre-defined in the TSAPs, except for the analysis populations of the PROs which are specified in the CS but not explicitly mentioned in the TSAPs. The ERG also notes a differently defined analysis population for ORR in the OAK trial in the TSAP (the analysis population of ORR will be all randomised patients with measurable disease at baseline) compared to the CS (see Table 10). The ERG is satisfied that all results are reported within the CSRs for the relevant population of each outcome.

Table 10 OAK and POPLAR trial analysis populations

Analysis	Population
Efficacy	Efficacy outcomes (OS, PFS, ORR and DOR) were analysed in the randomised (ITT) populations OAK trial: Two ITT populations were defined as follows: The primary population (1°P) was defined as the first 850 ITT patients, regardless of whether they received any trial drug The secondary population (2°P) was defined as all 1225 randomised ITT patients POPLAR trial: The ITT population was defined as all randomised patients, regardless of PD-L1 expression and whether they received any trial drug
PROs	OAK trial: The PRO evaluable population was defined as patients in the ITT population who had a non-missing baseline PRO assessment and at least one on-trial non-missing post-baseline PRO assessment POPLAR trial: The PRO evaluable population was defined as patients with a baseline PRO assessment and at least one post-treatment PRO assessment
Safety	OAK trial: Primary safety analyses were based on all 1225 randomised patients who received any dose of a trial drug during the treatment period POPLAR trial: Primary safety analyses were based on all randomised patients who received any dose of a trial drug during the treatment period

DOR=duration of response; ITT=intention-to-treat; ORR=objective response rate OS=overall survival, PD-L1=programmed death-ligand 1; PFS=progression-free survival; PRO=patient-reported outcome.
Source: CS, Section 4.4, OAK protocol and POPLAR protocol

Additional ERG assessments of the statistical approach of the OAK and POPLAR trials

A summary of the additional checks made by the ERG in relation to the statistical approach used by the company to analyse data from the OAK and POPLAR trials is provided in Table 11. Having carried out these checks, the ERG is satisfied with the statistical approach employed by the company.

Table 11 ERG assessment of statistical approach used to analyse trial data

Component	Statistical approach with ERG comments	
	OAK trial	**POPLAR trial**
Protocol amendments	Protocol amendments are provided within the CSR (p86-87) The largest amendments to the protocol were related to the sample size and statistical testing procedure of the trial (as outlined in 'Determination of sample size and analysis populations' above). This amendment was made on 28th January 2016, before the date of the primary analysis (data cut-off 7th July 2016) All other protocol amendments and rationale for amendments are outlined in detail. All amendments were made before the date of primary analysis, and so were unlikely to have been driven by the results of the trial	Protocol amendments are provided within the CSR (p91-93) All protocol amendments and rationale for amendments are outlined in detail. All amendments were made before the date of primary analysis, and so were unlikely to have been driven by the results of the trial
Subgroup analyses for OS	Subgroup analyses of OS performed in the primary population based on (CS, Section 4.4 and CSR p81): • Demographic characteristics (e.g. age, sex, ethnicity) • Baseline prognostic characteristics (e.g. PD-L1 expression subgroups, ECOG performance status, prior lines of chemotherapy, histology, smoking history) Summaries of OS, including the unstratified HR estimated from a Cox proportional hazards model and K-M estimates of median survival time, were produced separately for each level of the categorical variables The ERG is satisfied that subgroups were pre-defined in the TSAP (p20)	Subgroup analyses of OS performed based on (CS, Section 4.4 and CSR p83): • Demographic characteristics (e.g. age, sex, ethnicity) • Baseline prognostic characteristics (e.g. PD-L1 expression subgroups, ECOG performance status, prior lines of chemotherapy, histology, smoking history) Summaries of OS, including the unstratified HR estimated from a Cox proportional hazards model and Kaplan-Meier estimates of median survival time, were produced separately for each level of the categorical variables The ERG is satisfied that subgroups were pre-defined in the TSAP (p19)
Sensitivity analyses for OS	Two sensitivity analyses of OS were presented in the CSR for non-protocol anti-cancer therapy (p152-155) The ERG is satisfied that these sensitivity analyses were pre-defined in the TSAP (p20)	No sensitivity analyses were pre-specified in TSAP or presented in the CS or CSR for any analyses

Component	Statistical approach with ERG comments	
	OAK trial	**POPLAR trial**
Analysis of AEs	Many different summaries of AEs are provided in the CSR. Protocol defined adverse events of special interest (AESIs) (CS, Table 20) were summarised separately. AEs, SAEs, AESIs, and imAEs are summarised by treatment arm and grade (per NCI CTCAE v4.0). AEs, SAEs, severe AEs (Grade ≥ 3), AESIs, imAEs, and AEs leading to trial drug discontinuation or interruption are summarised separately. Additionally, AE summaries are provided by PD-L1 expression subgroups (TC3 or IC3, TC2/3 or IC2/3, TC1/2/3 or IC1/2/3, TC0 and IC0) for each treatment arm A complete list of the different summary tables is provided on p180-241 of the CSR. Further details of AEs are presented in Section 4.7 of this report The ERG is satisfied that the methodology used to analyse the AEs is appropriate and was pre-specified in the TSAP (p21-22)	Many different summaries of AEs are provided in the CSR; AEs were summarised by treatment arms and overall in incidence tables by NCI CTCAE grade, seriousness, relationship to trial drug, AEs leading to death, trial drug discontinuation, and dose modification/interruption Protocol defined AESIs (CS, Table 20) were summarised separately. A complete list of the different summary tables is provided on p195 to 226 of the CSR. The ERG is satisfied that the methodology used to analyse the AEs is appropriate and was pre-specified in the TSAP (p20)
Analysis of PROs	Analysis of PROs is presented in the CSR (p85-86) Global health status/ HRQoL, functioning, treatment-related symptoms and lung cancer symptoms were assessed by the EORTC QLQ-C30 and LC13. All the scales and single-item measures were linearly transformed so that each score ranged from 0 to 100. Summary statistics (mean, sd, median, range and mean change from baseline (and 95% CI) of linearly transformed scores) are reported for all the items and subscales of the EORTC QLQ-C30 and the EORTC QLQ-LC13 questionnaires Time to confirmed symptomatic deterioration was also measured, defined as the time from baseline to the first time the patient's score showed a ≥ 10-point increase above baseline in any of the lung cancer symptom scores and analysed by K-M and Cox regression methodology The ERG is satisfied that the methodology used to analyse PROs is appropriate and was pre-specified in the TSAP (p23)	Analysis of PROs is presented in the CSR (p77, p90) HRQoL and lung cancer symptoms were assessed using the EORTC QLQ-C30 and QLQ-LC13 questionnaires. All the scales and single-item measures were linearly transformed so that each score ranged from 0 to 100. Summary statistics (mean, sd, median, range and mean change from baseline (and 95% CI) of linearly transformed scores) are reported for all the items and subscales of the EORTC QLQ-C30 and the EORTC QLQ-LC13 questionnaires The ERG is satisfied that the methodology used to analyse PROs is appropriate and was pre-specified in the TSAP (p21-22)

AE=adverse event; AESI=adverse events of special interest; CI=confidence interval CS=company submission; CSR=clinical trial report; EORTC=European Organisation for the Research and Treatment of Cancer; ERG=Evidence Review Group; HRQoL=health-related quality of life; IC=tumour-infiltrating immune cell; K-M=Kaplan-Meier; NCI CTCAE=National Cancer Institute common terminology criteria for adverse events; OS=overall survival; PD-L1=programmed death-ligand 1; PFS=progression-free survival; PRO=patient-reported outcome; QLQ-C30=quality of life questionnaire in cancer (30 items); QLQ-LC13=quality of life questionnaire in lung cancer (13 items); SAE=serious adverse events; sd=standard deviation; TC= tumour cell; TSAP=trial statistical analysis plan; imAEs=immune-mediated adverse events
Source: adapted from the CS, OAK CSR, POPLAR CSR, OAK TSAP, POPLAR TSAP, the company's response to the ERG clarification letter, and ERG comment

4.2.5 Risk of bias assessment for the OAK and POPLAR trials

The ERG considers that the risk of bias for the OAK and POPLAR trials is low for the of the criteria in

Table 12. However, the open-label design provides the opportunity for investigator-assessed outcomes to be biased.

An additional possible source of bias is the fact that, in both the OAK and POPLAR trials, the treatment stopping rules for patients receiving atezolizumab and docetaxel differed. Treatment with atezolizumab was administered as long as patients experienced a clinical benefit (as assessed by an investigator) in the absence of unacceptable toxicity or symptomatic deterioration attributed to disease progression. Treatment with docetaxel was continued until disease progression or unacceptable toxicity.

Effectiveness evidence is immature. However, although a further analysis of OAK trial data is planned, the ERG notes that results from this analysis may be difficult to interpret as, although crossover from the docetaxel arm to the atezolizumab arm was not originally permitted in the OAK trial, it was allowed following the primary analysis of the primary population (n=850).

Table 12 Risk of bias assessment of the OAK and POPLAR trials

Study question	Company assessment		ERG comment
	OAK trial	POPLAR trial	
Was randomisation carried out appropriately?	Yes	Yes	Agree
Was the concealment of treatment allocation adequate?	N/A (open-label study)	N/A (open-label study)	Disagree that this question is N/A Patients were randomised via IVRS and therefore treatment allocation was concealed
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	Yes	Yes	Agree
Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	N/A (open-label study)	N/A (open-label study)	Disagree that this question is N/A The open-label nature of the trials provides an opportunity for subjective results to be biased
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	No	No	Agree
Is there any evidence to suggest that the authors measured more outcomes than they reported?	N/A (full data available)	N/A (full data available)	Only limited details by PD-L1 status are presented in the CS
Did the analysis include an intent-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes	Yes	Agree

ERG=Evidence Review Group; IVRS=interactive voice response system; N/A=not applicable; CS=company submission; PD-L1=programmed death ligand 1
Source: CS, Table 27

4.3 Results from the OAK and POPLAR trials

All of the data from the OAK trial presented in this section correspond to the data cut-off date of 7th July 2016 which was the primary analysis time in the 1°P (see Section 4.2.4 for definitions of populations).

The data presented in the CS from the POPLAR trial correspond to the data cut-off date of 1st December 2015, which was the date of an updated efficacy analysis. The POPLAR trial data presented in the CSR correspond to a data cut-off date of 8th May 2015, which was the date of the primary analysis. Unless otherwise stated, POPLAR trial results presented in this section are those from the updated analysis presented in the CS.

As outlined in Section 4.2.4 of this report, the assumption of PH for the outcomes of OS and PFS were demonstrated not to be valid for either the OAK or POPLAR trials. Therefore, the ERG notes that HRs reported in this section must be interpreted with caution.

4.3.1 Participant flow in OAK and POPLAR

OAK

Within the 1°P of the OAK trial, a total of 850 participants were randomised; 425 to each treatment arm. The participant flow in the 1°P of the OAK trial is presented in Figure 6 of the CS and reasons for treatment discontinuation are provided in Table 23 of the CS. The ERG noted an error in Table 23 of the CS and, in their response to the clarification letter, the company clarified that the number of participants withdrawn from treatment with atezolizumab was 364; 316 discontinued treatment due to progressive disease, 36 due to an AE, 9 due to withdrawal by patient, 2 due to physician decision, and 1 due to "other" (the participant was randomised to receive docetaxel, but received atezolizumab in error).

The median duration of survival follow-up was 21.4 months (range 0.1 to 27.1 months) in the atezolizumab arm and 21.3 months (range 0 to 26.9 months) in the docetaxel arm at the time of primary analysis. The minimum length of follow-up in both treatment arms was 19 months (duration from last patient randomised date to clinical cut-off date).

POPLAR

A total of 287 participants were randomised; 144 participants to the atezolizumab arm and 143 participants to the docetaxel arm. The participant flow in the POPLAR trial is presented in Figure 7 of the CS and reasons for treatment discontinuation are provided in Table 25 of the CS.

The median duration of survival follow-up was 14.8 months (range 0.2 to 19.6 months) in the atezolizumab arm and 15.7 months (range 0 to 18.7 months) in the docetaxel arm at the time of primary analysis. The updated analysis provided an additional 7 months of follow-up and, at this time, the minimum length of follow-up was 20 months (duration from last patient randomised date to clinical cut-off date for updated analysis).

4.3.2 Primary efficacy outcome: overall survival

The primary outcome of both trials was OS. At the time of the primary analysis, in the OAK trial, 569 out of 850 randomised participants in the 1°P had died and, at the time of updated analysis in the POPLAR trial, 200 out of 287 randomised participants had died. OS results from the OAK and POPLAR trials are presented in

Table 13.

Table 13 OS results from the OAK and POPLAR trials

Outcome	Atezolizumab	Docetaxel
OAK trial		
Number of participants analysed	425	425
Median OS, months (95% CI)	13.8 (11.8 to 15.7)	9.6 (8.6 to 11.2)
HR (95% CI) - stratified in the 1°P and TC1/2/3 or IC1/2/3 subgroup	0.73 (0.62 to 0.87, log-rank p-value=0.0003)	
POPLAR trial		
Number of participants analysed	144	143
Median OS, months (95% CI)	12.6 (9.7 to 16.0)	9.7 (8.6 to 12.0)
HR (95% CI) – stratified for ITT population	0.69 (0.52 to 0.92, log-rank p-value=0.011)	

1°P=primary population; CI=confidence interval; IC=tumour-infiltrating immune cell; HR=hazard ratio; ITT=intention-to-treat; OS=overall survival; PD-L1=programmed death-ligand 1; TC=tumour cell
Source: CS, adapted from Section 4.7, Figure 8 and Figure 11

Results of earlier interim and primary analyses of data from the POPLAR trial are presented for completeness in (Section 10.2)

In the OAK trial, compared to docetaxel, treatment with atezolizumab was associated with a statistically significant and clinically meaningful improvement in OS in the 1°P (stratified HR 0.73, 95% CI: 0.62 to 0.87; log-rank p-value=0.0003). From Kaplan-Meier (K-M) data (CS, Figure 8), the company considers that the curves separate at around 3 months, and the benefit for atezolizumab over docetaxel is maintained thereafter.

In the POPLAR trial, compared to docetaxel, treatment with atezolizumab was also associated with a statistically significant and clinically meaningful improvement in OS in the ITT population (stratified HR 0.69, 95% CI: 0.52 to 0.92; log-rank p-value=0.011). From K-M data (CS, Figure 11), the company considers that the curves separate at around 3 months, and the benefit for atezolizumab over docetaxel is maintained thereafter, with further separation of the K-M curves at around 9 months and increased benefit shown with atezolizumab compared to docetaxel with extended follow-up (CS, Figure 12).

Overall survival according to histology and to PD-L1 status in the OAK and POPLAR trials

In the OAK trial, a statistically significant improvement in OS with treatment with atezolizumab compared to docetaxel was observed regardless of histology; however, a longer median OS was observed in atezolizumab treated patients with non-squamous NSCLC (15.6 months) compared to patients with squamous NSCLC (8.9 months). This pattern was also observed in the docetaxel arm (non-squamous: 11.2 months, squamous: 7.7 months). In line with clinical advice to the ERG, the company states that this result reflects the inherently worse prognosis of patients with squamous cancers.

Results of the OAK trial also showed statistically significant improvements in OS with treatment with atezolizumab compared to docetaxel for people regardless of PD-L1 status; for individuals with NSCLC of ≥1% PD-L1 expression (TC 1/2/3 or IC 1/2/3; stratified HR 0.74, 95% CI: 0.58 to 0.93; p=0.0102) and for individuals with NSCLC of no measurable PD-L1 expression (TC0/IC0; HR 0.75, 95% CI: 0.59 to 0.96; p=0.0205). Results for OS according to histology and according to PD-L1 status subgroups are presented in Table 14. Additionally, results for OS according to individual TC or IC expression levels are presented in Figure 18 of the CSR for the OAK trial.

Primary and updated analysis results for OS in the POPLAR trial, according to histology and according to PD-L1 status are presented in Section 10.3. Results of the POPLAR trial showed a statistically significant improvement in OS with atezolizumab compared to docetaxel in patients with non-squamous NSCLC but no statistically significant difference between treatment arms for patients with squamous NSCLC, despite separation of survival curves over time. Results from the POPLAR trial also show a statistically significant improvement for individuals with NSCLC of ≥1% PD-L1 expression but not for individuals with NSCLC of no measurable PD-L1 expression.

Table 14 OS results in the OAK trial according to histology and to PD-L1 status

Outcome	Atezolizumab	Docetaxel
ITT population		
Number of participants analysed	425	425
Median OS, months (95% CI)	13.8 (11.8 to 15.7)	9.6 (8.6 to 11.2)
HR (95% CI) - stratified in the 1°P and TC1/2/3 or IC1/2/3 subgroup	0.73 (0.62 to 0.87, log-rank p-value=0.0003)	
Histology: non-squamous NSCLC		
Number of participants analysed	313	315
Median OS, months (95% CI)	15.6 (13.3 to 17.6)	11.2 (9.3 to 12.6)
HR (95% CI) – unstratified	0.73 (0.60 to 0.89, log-rank p-value=0.0015)[a]	
Histology: squamous NSCLC		
Number of participants analysed	112	110
Median OS, months (95% CI)	8.9 (7.4 to 12.8)	7.7 (6.3 to 8.9)
HR (95% CI) – unstratified	0.73 (0.54 to 0.98, log-rank p-value=0.0383)[a]	
PD-L1 subgroup: TC3 or IC3		
Number of participants analysed	[redacted]	[redacted]
Median OS, months (95% CI)	[redacted]	[redacted]
HR (95% CI) – unstratified	[redacted]	
PD-L1 subgroup: TC2/3 or IC2/3		
Number of participants analysed	[redacted]	[redacted]
Median OS, months (95% CI)	[redacted]	[redacted]
HR (95% CI) – unstratified	[redacted]	
PD-L1 subgroup: TC2/3 or IC2/3 excluding TC3 or IC3		
Number of participants analysed	[redacted]	[redacted]
Median OS, months (95% CI)	[redacted]	[redacted]
HR (95% CI) – unstratified	[redacted]	
PD-L1 subgroup: TC1/2/3 or IC1/2/3		
Number of participants analysed	[redacted]	[redacted]
Median OS, months (95% CI)	[redacted]	[redacted]
HR (95% CI) – stratified for PD-L1 status	[redacted]	
PD-L1 subgroup: TC1/2/3 or IC1/2/3 excluding TC2/3 or IC2/3		
Number of participants analysed	[redacted]	[redacted]
Median OS, months (95% CI)	[redacted]	[redacted]
HR (95% CI) – unstratified	[redacted]	
PD-L1 subgroup: TC0/IC0		
Number of participants analysed	[redacted]	[redacted]
Median OS, months (95% CI)	[redacted]	[redacted]
HR (95% CI) – unstratified	[redacted]	

[a] The company state that p-values for histology subgroups are presented for descriptive purposes only

1°P=primary population; CI=confidence interval; HR=hazard ratio; IC=tumour-infiltrating immune cell; ITT=intention-to-treat; NE=not evaluable; NR=not reported; NSCLC=non-small cell lung cancer; OS=overall survival; PD-L1=programmed death-ligand 1; TC=tumour cell

Source: CS, adapted from Section 4.7, Section 4.8, Figure 8 and Figure 14; OAK trial CSR, adapted from Table 23, Table 33, Table 34 and Figure 14

Overall survival according to baseline characteristics in the OAK and POPLAR trials

Subgroup analyses were also performed in both the OAK and POPLAR trials based on demographic characteristics (e.g. age, sex, ethnicity) and baseline prognostic characteristics (e.g. PD-L1 expression subgroups, ECOG PS, prior lines of chemotherapy, histology, smoking history).

Figure 15 and Figure 18 of the CS show results of OS according to these demographic and baseline characteristics in the OAK and POPLAR trials respectively. The company states that improvement in OS with atezolizumab compared with docetaxel is consistent across baseline characteristics and highlights the result for participants with central nervous system (CNS) metastases in the OAK trial (HR 0.54, 95% CI 0.31 to 0.94; OAK CSR, Figure 20). The ERG generally agrees with the company's interpretation of these subgroup analyses but notes that improvement with atezolizumab does not seem to be consistent in the subgroup of patients with positive NSCLC EGFR mutation in the OAK trial (HR 1.24, 95% CI 0.71 to 2.18; OAK CSR, Figure 20). The ERG also notes that the findings of these subgroup analyses should be treated with caution, due to small numbers of patients included in some of the subgroups, such as CNS metastases, EGFR and KRAS mutation subgroups, leading to wide CIs around subgroup-specific HRs.

4.3.3 Secondary efficacy outcome: progression-free survival

A secondary efficacy outcome of both the OAK and POPLAR trials was investigator-assessed PFS per response evaluation criteria in solid tumours (RECIST) v.1.1. OAK and POPLAR trial PFS results are presented in Table 15.

Table 15 Investigator-assessed PFS results in the OAK and POPLAR trials

Outcome	Atezolizumab	Docetaxel
OAK trial		
Number of participants analysed	425	425
Median PFS, months (95% CI)	2.8 (2.6 to 3.0)	4.0 (3.3 to 4.2)
HR (95% CI) - stratified in the 1°P and TC1/2/3 or IC1/2/3 subgroup	0.95 (0.82 to 1.10, log-rank p-value=0.4928)	
POPLAR trial		
Number of participants analysed	144	143
Median PFS, months (95% CI)	2.7 (2.0 to 4.1)	3.4 (2.8 to 4.1)
HR (95% CI) – stratified for ITT population	0.92 (0.71 to 1.20, log-rank p-value=0.556)	

1°P=primary population; CI=confidence interval; HR=hazard ratio; IC=tumour-infiltrating immune cell; ITT=intention-to-treat; PD-L1=programmed death-ligand 1; PFS=progression-free survival; TC=tumour cell
Source: CS adapted from Section 4.7, Figure 9 and Figure 13

Results from earlier interim and primary analyses of the POPLAR trial are presented for completeness in Section 10.4.

There was no statistically significant difference in investigator-assessed PFS between atezolizumab and docetaxel in the OAK trial (Table 15) or in any of the analyses of the POPLAR trial (Section 10.3). The company states that this is consistent with the known profiles and mechanism of action of immunotherapies.

The ERG notes numerical inconsistency in the investigator-assessed PFS results from the OAK trial. Specifically, the stratified HR and log-rank p-value indicate no statistically significant difference between treatments, but the 95% CIs of median PFS do not overlap, suggesting a potentially longer median PFS for patients treated with docetaxel (4.0 [95% CI: 3.3 to 4.2] months) compared to atezolizumab (2.8 [95% CI: 2.6 to 3.0] months). The ERG suggests that the apparent inconsistency between these results may be due to the use of a HR to summarise the relative treatment effect, when the PH assumption required for the interpretation of this effect measure is violated (CS, Figure 34).

The company also states that late separation of K-M curves in the updated analysis of the POPLAR trial reflects the prolonged responses seen in some atezolizumab recipients.[46]

4.3.4 Other secondary efficacy outcomes

The company reported ORR and DOR as additional secondary efficacy outcomes. These are described in Section 10.2 and Section 10.5 of this report.

4.3.5 Subsequent therapies in the OAK and POPLAR trials

OAK

The proportion of patients receiving a non-protocol anti-cancer therapy was similar in the two treatment arms (48.5% of patients randomised to atezolizumab and 45.2% of patients randomised to docetaxel; CS, Table 28). The proportion of patients receiving a subsequent cancer immunotherapy was 5% in the atezolizumab arm and 17% in the docetaxel arm. Further information about subsequent therapies is presented in Table 28 of the CS.

POPLAR

The proportion of patients receiving a non-protocol anti-cancer therapy was similar in the two treatment arms (40.3% of patients randomised to atezolizumab and 41.3% of patients randomised to docetaxel, POPLAR CSR; Table 24). No patients randomised to atezolizumab, and 5% of patients in the docetaxel arm received a subsequent cancer immunotherapy. Further information of subsequent therapies is presented in Table 24 of the POPLAR CSR.

4.4 Health-related quality of life

Three patient reported outcome questionnaires were used in the OAK and POPLAR trials to collect data on the impact of treatment with atezolizumab and docetaxel on patients' disease-related symptoms and HRQoL:

- The EQ-5D-3L[56] questionnaire
- The European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire – Core 30 items (EORTC QLQ-C30[57])
- The European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire – Lung Cancer 13 items (EORTC QLQ-LC13[58])

The company reports that completion rates in the OAK trial were consistently high over the course of treatment. Key findings from data collected using the EORTC QLC-C30/LC13 questionnaires are summarised in Section 4.4.1. The ERG describes how utility results generated from analyses of data collected using the EQ-5D-3L[56] questionnaire are used in the company model in Section 5.5.1 of this ERG report.

4.4.1 EORTC QLQ-30/LC13 results

A summary of the main EORTC-QLQC-C30/LC13[57,58] results presented by the company in the CS is provided in Table 16.

Table 16 EORTC-QLQ-C30/LC13 results

Measure	OAK trial		POPLAR trial
	Atezolizumab	Docetaxel	
Average global health status and functioning scores (i.e. physical, role, social, emotional and cognitive)	No clinically meaningful deterioration over time for both treatment arms		Patients in the atezolizumab arm did not demonstrate any clinically meaningful change (improvement or decline) on any of the subscales assessed, while patients in the docetaxel arm had a meaningful increase in alopecia
Commonly reported cancer treatment-related symptoms	Patients in the atezolizumab arm did not show clinically meaningful worsening of symptoms, while patients in the docetaxel arm demonstrated a clinically meaningful worsening in alopecia and peripheral neuropathy throughout treatment		
Chest pain: time to deterioration of patient-reported chest pain	Patients treated with atezolizumab demonstrated prolonged time to deterioration compared with patients treated with docetaxel (HR 0.72, 95% CI: 0.55 to 0.93)		There was no difference between the arms in the time to deterioration of lung cancer symptoms
Chest pain: baseline			
No chest pain	57.7%	60.6%	N/A
Other categories (not at all, a little, quite a bit, very much)	Similar proportions		N/A
Chest pain: at radiographic disease progression			
Asymptomatic	66.4%	54.2%	N/A
Clinically meaningful worsening in chest pain severity (≥10 point increase from baseline)	11.4%	25.4%	N/A

CI=confidence interval; HR=hazard ratio; N/A=not applicable
Source: CS, pp84-86 and p89

4.5 Adverse events reported in the OAK and POPLAR trials

Clinical advice to the ERG is that AEs arising from treatment with immunotherapy in patients with NSCLC require careful monitoring by a specialist clinical team with the experience to provide early recognition and management of immunotherapy-related AEs.

Details of the AEs experienced by patients participating in the OAK trial are presented in Section 4.12 of the CS (pp116-124). The ERG notes that the safety evaluable population of the OAK trial includes all patients randomised to the trial who received a dose of study drug (N=1187). The 1°P includes only the first 850 patients randomised to the OAK trial.

Details of AEs experienced by patients in the POPLAR trial are not presented in the CS, but are available in the CSR and the published report.[25] The company has focussed on the AE data from the OAK trial (as this trial is bigger than the POPLAR trial) and reports that the rates

of AEs experienced by patients participating in the POPLAR trial were similar to those experienced by patients in the OAK trial (Table 17). The ERG agrees with the company that the rates and types of AEs experienced by patients in both trials are similar.

The ERG notes that, with the exception of serious adverse events (SAEs), the incidence rates of all categories of AEs, in the OAK and POPLAR trials, are lower in the atezolizumab arms than in the docetaxel arms (Table 17).

Table 17 Overview of adverse events (OAK and POPLAR trials)

Category of event	OAK trial		POPLAR trial	
	Atezolizumab N=609 n (%)	Docetaxel N=578 n (%)	Atezolizumab N=142 n (%)	Docetaxel N=135 n (%)
Patients with at least one event	573 (94)	555 (96)	136 (96)	130 (96)
Treatment-related AEs	390 (64)	496 (86)	95 (67)	119 (88)
Grade 3 and 4 AEs	227 (37)	310 (54)	57 (40)	71 (53)
Treatment-related Grade 3 and 4 AEs	90 (15)	247 (43)	16 (11)	52 (39)
Grade 5 AEs	10 (2)	14 (2)	6 (4)	5 (4)
Treatment-related deaths	0 (0)	1 (0.2)	1 (1)	3 (2)
Serious AEs	194 (32)	181 (31)	50 (35)	46 (34)
AEs leading to withdrawal from treatment	46 (8)	108 (19)	11 (8)	30 (22)
AE leading to dose modification/interruption	152 (25)	210 (36)	34 (24)	44 (33)

AE=adverse event
Source: CS, Table 43

Study drug exposure in the OAK trial

The company reports (CS, p117) that the median duration of treatment for patients in the atezolizumab arm was longer than that for patients in the docetaxel arm (3.4 months and 2.1 months respectively). In addition, the median number of treatment cycles for patients in the atezolizumab arm was higher than that for patients in the docetaxel arm, six and four cycles respectively (Table 18).

Table 18 Study drug exposure in the OAK trial

	Atezolizumab (n=609)	Docetaxel (n=578)
Median treatment duration, months (range)	3.4 (0-26)	2.1 (0-23)
Treatment duration, n (%)		
0 to ≤3 months	294 (48)	351 (61)
>3 to ≤6 months	113 (19)	162 (28)
>6 to ≤12 months	77 (13)	51 (9)
>12 months	125 (21)	14 (2.4)
Median number of doses (range)	6.0 (1-38)	4.0 (1-30)

Source: CS, Table 44

Adverse events of any grade and any cause in the OAK trial

The majority of patients in the atezolizumab and docetaxel arms of the OAK trial experienced at least one AE of any grade (94% and 96%). In Table 19 of the CS, the company has provided details of the specific AEs (any grade) experienced by ≥20% of patients in the OAK trial. The ERG notes that the rates of anaemia and alopecia are substantially lower in the cohort of patients who were treated with atezolizumab than in the cohort of patients treated with docetaxel.

Table 19 Adverse events (any grade) reported by ≥20% of patients (OAK trial)

Adverse event	Atezolizumab (n=609) n (%)	Docetaxel (n=578) n (%)
Nausea	108 (18)	131 (23)
Diarrhoea	94 (15)	141 (24)
Fatigue	163 (27)	205 (36)
Decreased appetite	143 (24)	136 (24)
Cough	141 (23)	105 (18)
Anaemia	70 (12)	136 (24)
Alopecia	3 (0.5)	202 (35)

Source: adapted from the CS, Table 45

The ERG requested (via the clarification process) details of all-cause AEs of any grade that were reported in ≥10% of patients; however, the company's clarification response included only details of treatment-related AEs of any grade that were reported in ≥10% of patients. The ERG notes that data for AEs reported by ≥10% of patients in the OAK trial are available in Table S5 of the supplementary appendix of the published report[23] of the OAK trial. The ERG is satisfied that the AEs reported in the supplementary appendix do not signal any safety concerns.

The AEs (any grade, any cause) for which there was a ≥5% difference in incidence between either of the arms of the OAK trial are illustrated in Figure 29 of the CS (p118). Adverse events

reported by a higher proportion of patients in the atezolizumab arm compared to the docetaxel arm include musculoskeletal pain (10.5% versus 4.3%) and pruritus (8.2% versus 3.1%). The company reports (CS, p118) that the majority (94%) of the musculoskeletal and pruritus events were Grade 1 or Grade 2 events and that there were no AEs for which the incidence in the atezolizumab arm was ≥10% higher than the incidence in the docetaxel arm.

Treatment-related adverse events in the OAK trial

More patients in the docetaxel arm (86%) than in the atezolizumab arm (64%) of the OAK trial experienced at least one treatment-related AE (Table 17). The company has provided details of the treatment-related AEs reported in ≥10% of patients (Table 20). The ERG notes that the incidence rates for all treatment-related AEs listed in Table 20 are higher in the docetaxel arm than in the atezolizumab arm.

The company reports (CS, p119) that compared with patients in the docetaxel arm, a lower proportion of patients treated with atezolizumab experienced Grade 3 or 4 treatment-related AEs (42.7% versus 14.8%). The Grade 3 and Grade 4 events experienced by ≥10% of patients treated with docetaxel included: fatigue, asthenia, nausea, diarrhoea, stomatitis, alopecia, anaemia, neutropenia, febrile neutropenia, peripheral neuropathy, decreased appetite, and myalgia. The single Grade 3 and 4 event reported in the atezolizumab arm was fatigue.

Table 20 Treatment-related AEs (any grade) in ≥10% of patients (OAK trial)

Adverse event	Atezolizumab (n=609) n (%)	Docetaxel (n=578) n (%)
Alopecia	3 (0.5)	198 (34)
Fatigue	87 (14)	177 (31)
Decreased appetite	52 (9)	116 (20)
Anaemia	24 (4)	114 (20)
Nausea	53 (9)	112 (19)
Diarrhoea	47 (8)	109 (19)
Asthenia	51 (8)	96 (17)
Neutropenia	7 (1)	85 (15)
Myalgia	21 (3)	81 (14)
Febrile neutropenia	0	61 (11)
Stomatitis	13 (2)	59 (10)
Peripheral neuropathy	6 (1)	58 (10)

Source: CS, Table 46

Serious adverse events in the OAK trial

Similar rates of SAEs were reported between patients treated with atezolizumab and patients treated with docetaxel (31.9% and 31.3%). The SAEs considered by the trial investigators to be treatment-related are listed in Table 21. The ERG notes that fewer patients treated with atezolizumab experienced a treatment-related SAE (10.3% versus 17.6%).

The company also reports (CS, p123) the number of deaths due to AEs (rather than progressive disease) that occurred during the 30 days following the last study treatment. In the atezolizumab arm, 10 of 62 deaths (16.1%) were due to AEs rather than progressive disease. In the docetaxel arm, 14 of 42 deaths (33.3%) were due to AEs rather than progressive disease.

Table 21 Treatment-related SAEs reported by ≥2% of patients (OAK trial)

Adverse event	Atezolizumab (n=609) n (%)	Docetaxel (n=578) n (%)
Total number of patients with at least one event	63 (10.3)	102 (17.6)
Febrile neutropenia	0	36 (6.2)
Pneumonia	2 (0.3)	11 (1.9)
Diarrhoea	0	6 (1.0)
Pyrexia	3 (0.5)	5 (0.9)
Neutrophil count decreased	0	5 (0.9)
Anaemia	0	4 (0.7)
Pleural effusion	1 (0.2)	3 (0.5)
Vomiting	0	3 (0.5)
Dehydration	0	3 (0.5)
Neutropenia	0	3 (0.5)
Lung infection	0	3 (0.5)
Colitis	1 (0.2)	2 (0.3)
Acute kidney injury	1 (0.2)	2 (0.3)
Lower respiratory tract infection	0	2 (0.3)
Neutropenic sepsis	0	2 (0.3)
Urinary tract infection	0	2 (0.3)
Asthenia	0	2 (0.3)
Syncope	0	2 (0.3)
Pneumonitis	6 (1.0)	1 (0.2)
Hypersensitivity	3 (0.5)	0
Meningitis	3 (0.5)	0

Source: CS, Table 47

Adverse events of special interest in the OAK trial

The protocol-defined AEs of special interest (AESIs) in the OAK trial are summarised in (Table 22). The AESIs were events considered by the company to be potentially auto-immune mediated.

The company reports that the majority of AESIs in both trial arms were of Grade 1 or Grade 2 severity and that a higher rate of dermatological, hepatic and endocrine events was reported by patients treated with atezolizumab than by patients treated with docetaxel. A higher rate of neurological events was reported by patients treated with docetaxel than by patients treated with atezolizumab. The company considers the AESI profiles of atezolizumab and docetaxel are consistent with the mechanisms of action of the two drugs.

Clinical advice to the ERG is that [redacted] are difficult to treat and potentially dangerous.

Table 22 Summary of OAK trial adverse events of special interest

Adverse event	Atezolizumab (n=609) n (%)	Docetaxel (n=578) n (%)
Any AE	184 (30.2)	132 (22.8)
Grade 1	87 (14.3)	82 (14.2)
Grade 2	66 (10.8)	36 (6.2)
Grade 3	28 (4.6)	14 (2.4)
Grade 4	3 (0.5)	0
[redacted]		
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]

[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]

ALT=alanine aminotransferase; AST=aspartate transaminase
Source: CS, Table 48

4.6 ERG summary and critique of the indirect evidence

The company states that the ITC was conducted to support pricing and reimbursement across a range of countries and, therefore, included several comparators (afatinib, dacomitinib, erlotinib, gefitinib, paclitaxel and pemetrexed) that are not relevant to the final scope of the present appraisal. Results presented in the CS are restricted to the relevant comparators within the UK.

The ERG has concerns about the company's approach to the ITC as the final efficacy results are adjusted for all comparators within the network, including comparators that are not included in the final scope issued by NICE (see Section 3.3 of this ERG report for further details). The ERG, therefore, as part of the clarification process, asked the company to repeat the ITC using a reduced network that included only the relevant drugs (atezolizumab 1200mg, docetaxel 75mg/m^2, nintedanib 200mg+docetaxel 75mg/m^2).

4.6.1 Trials identified for inclusion in the indirect treatment comparison

The company conducted a systematic search (see Section 4.1 of this report for further details) to identify phase II-IV RCTs investigating the efficacy and safety of pharmacological interventions for second- and further-line treatments for locally advanced/metastatic NSCLC.

Three trials (OAK, POPLAR and LUME-Lung 1 trials) that included comparators relevant to this appraisal were identified and included in the reduced network ITC. These three trials form a network (Figure 2), which enable an ITC of atezolizumab, docetaxel and nintedanib+docetaxel for the outcomes of OS and PFS to be carried out.

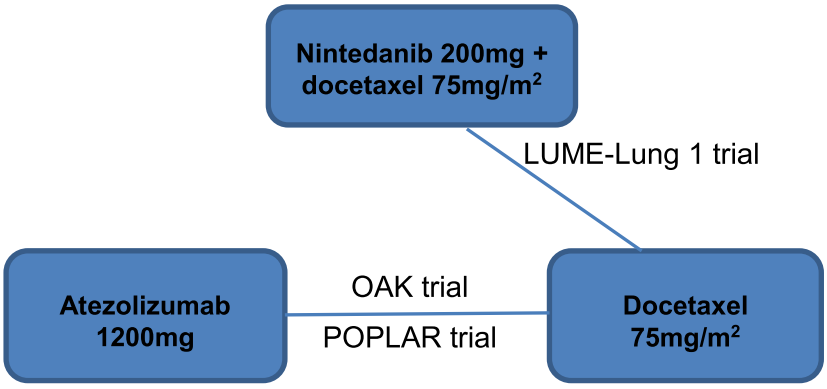


Figure 2 Network plots for ITCs of OS and PFS

For the OAK and POPLAR trials, characteristics are summarised in Table 6 of this report and an assessment of the risk of bias is provided in Section 4.2.5. Design characteristics and a risk of bias assessment for the LUME-Lung 1 trial are available in Section 10.6 and detailed

inclusion and exclusion criteria for recruitment to the LUME-Lung 1 trial can be found in Appendix 4 of the CS.

In summary, the ERG considers that design characteristics and eligibility criteria of the LUME-Lung 1 trial are similar to those of the OAK and POPLAR trials. The principal differences in design between the trials are that the LUME-Lung 1 trial is double-blinded, with nintedanib or matching placebo added to docetaxel, and the primary efficacy outcome of the LUME-Lung 1 trial is centrally assessed PFS; OAK and POPLAR are open-label trials, both with the primary efficacy outcome of OS. The company judged the LUME-Lung 1 trial to be at low risk for all domains of bias considered; the ERG agrees with this assessment.

Patient demographic and baseline characteristics of the three trials are summarised in Table 23. Further demographic and baseline characteristics of the OAK and POPLAR trials are summarised in Table 7. The ERG notes that the 'evaluable population' of the direct clinical effectiveness evidence for the OAK trial is the 1°P (the first 850 randomised participants) while the evaluable population of the OAK trial for the ITC is the 2°P (i.e. all 1225 randomised participants). Therefore, the baseline characteristics in Table 7 and in Table 23 for the OAK trial relate to different populations.

The comparability and representativeness of the OAK and POPLAR populations analysed for the direct clinical effectiveness evidence is discussed in Section 4.2 of this report. The evaluable populations of the OAK and POPLAR trials included within the ITC also have similar baseline characteristics, are balanced across treatment arms and generally are representative of patients with advanced NSCLC who are likely to be treated in the NHS, with the caveat that patients in the trials were slightly younger and fitter than NHS patients.

Table 23 Characteristics of evaluable patients

Parameter	OAK		POPLAR		LUME-Lung 1	
	Atezolizumab	Docetaxel	Atezolizumab	Docetaxel	Nintedanib+ docetaxel	Placebo+ docetaxel
Evaluable, n	613	612	144	143	655	659
Demographics						
Male, n (%)	378 (61.7)	379 (61.9)	93 (64.6)	76 (53.1)	476 (72.7)	479 (72.7)
Median age	63	64	62	62	60	60
Age <65 years, n (%)	335 (54.7)	326 (53.3)	87 (60.4)	87 (60.8)	200 (30.5)	214 (32.5)
Age ≥65 years, n (%)	278 (45.3)	286 (46.7)	57 (39.6)	56 (39.2)	455 (69.5)	445 (67.5)
Median (range) months since initial diagnosis	13.9 (2.4–285.6)	12.8 (2.3–215.0)	12.3 (2.3–114.3)	13.5 (3.4–115.7)	8.8 (5.4–13.6)[a]	8.6 (5.4–13.6)[a]
Ethnicity, n (%)						
White	438 (71.5)	432 (70.6)	110 (76.4)	116 (81.1)	533 (81.4)	530 (80.4)
Black	11 (1.8)	16 (2.6)	3 (2.1)	4 (2.8)	4 (0.6)	5 (0.8)
Asian	124 (20.2)	125 (20.4)	23 (16.0)	13 (9.1)	116 (17.7)	123 (18.7)
Other or unknown	11 (1.8)	12 (2.0)	8 (5.5)	10 (7.0)	2 (0.3)	1 (0.2)
Smoking history, n (%)						
Current or former smoker	501 (81.7)	516 (84.3)	117 (81.3)	114 (79.7)	490 (74.8)	498 (75.6)
Current smoker	83 (13.5)	107 (17.5)	25 (17.4)	21 (14.7)	-	-
Former smoker	418 (68.2)	409 (66.8)	92 (63.9)	93 (65.1)	-	-
Never smoked	112 (18.3)	96 (15.7)	27 (18.7)	29 (20.3)	165 (25.2)	161 (24.4)
Disease stage at initial diagnosis, n (%)						
< IIIB or unknown	183 (29.9)	151 (24.7)	31 (21.5)	39 (27.3)	108 (16.5)	105 (15.9)
IIIB	68 (11.1)	87 (14.2)	25 (17.4)	14 (9.8)	148 (22.6)	146 (22.2)
IV	362 (59.0)	374 (61.1)	88 (61.1)	90 (62.9)	399 (60.9)	408 (61.9)
Histology, n (%)						
Adenocarcinoma	-	-	-	-	322 (49.2)	336 (51.0)
Large cell	-	-	-	-	25 (3.8)	16 (2.4)
Squamous cell	161 (26.3)	160 (26.1)	49 (34.0)	48 (33.6)	276 (42.1)	279 (42.3)
Other[b]	452 (73.7)	425 (69.4)	95 (66.0)	95 (66.4)	28 (4.3)	23 (3.5)
ECOG PS, n (%)						
0	221 (36.1)	234 (38.2)	46 (32.4)	45 (31.7)	187 (28.5)	189 (28.7)
1	392 (63.9)	378 (61.8)	96 (67.6)	97 (68.3)	467 (71.3)	470 (71.3)

[a] Interquartile range rather than range reported for LUME-Lung 1
[b] 'Other' defined as non-squamous in the OAK and POPLAR trials.
ECOG PS=Eastern Cooperative Oncology Group performance status
Source: CS Appendix 4, adapted from Tables 9 to 11; POPLAR CSR, Table 11 and Table 19; OAK CSR, p1586-90; Reck et al[26]

The baseline characteristics of the patients in the LUME-Lung 1 trial are also well balanced across the treatment arms. When comparing the common docetaxel control arm of the three trials the ERG notes that:

- slightly more male patients are included in the LUME-Lung 1 trial compared to the OAK and POPLAR trials (72.7% compared to 61.9% and 53.1% respectively)
- substantially more patients over the age of 65 are included in LUME-Lung 1 compared to the OAK and POPLAR trials (67.5% compared to 46.7% and 39.21% respectively)
- slightly fewer white/Caucasian patients are included in the OAK trial than the POPLAR and LUME-Lung 1 trials (70.6% compared to 81.1% and 80.4% respectively)
- slightly more current or former smokers are included in the OAK trial than the POPLAR and LUME-Lung 1 trials (84.3% compared to 79.7% and 75.6% respectively)
- the median months since initial NSCLC diagnosis was shorter in the LUME-Lung 1 trial compared to the OAK and POPLAR trials (8.6 compared to 12.8 and 13.5 respectively)
- more patients with squamous cell histology are included in the LUME-Lung 1 trial compared to the OAK and POPLAR trials (42.3% compared to 26.1% and 33.6% respectively)
- slightly fewer patients with ECOG PS of 0 are included in the LUME-Lung 1 trial compared to the OAK and POPLAR trials (28.7% compared to 38.2% and 31.7% respectively)
- disease stage at initial diagnosis was similar across all three trials.

The ERG notes that these observed differences in design and characteristics of the three trials included in the ITC should be taken into account when interpreting numerical results. However, the ERG does not consider that the majority of the observed differences would violate the assumption of transitivity required for the inclusion of these three trials in the same network.

The ERG notes that, in Europe, nintedanib+docetaxel is licensed for the treatment of patients with NSCLC adenocarcinoma. The company states that this is not consistent with the anticipated marketing authorisation for atezolizumab. To conduct a 'like-with-like' comparison, the company conducted the ITC using data from the 'total population' of the LUME-Lung 1 trial (i.e. all participants regardless of histology) and the evaluable ITT populations of the OAK and POPLAR trials (i.e. including patients with non-squamous and squamous histology).

The ERG questioned whether it was appropriate to include data in the ITC from patients in the LUME-Lung-1 trial that were not specified in the licensed population for treatment with nintedanib+docetaxel. The ERG asked the company to repeat the ITC using only the three relevant trials (reduced network) and for adenocarcinoma subgroups or non-squamous subgroups only. These results are discussed in Section 4.9.4.

4.6.2 Methodological approach to the indirect comparison

The company performed ITCs on OS and PFS as time-to-event outcomes, and OS at 12 months, ORR and TRAEs as binary outcomes. ITCs of ORR and TRAE did not contribute to the company's cost effectiveness analyses; therefore, the methods and results presented in this ERG report relate only to the ITCs of OS and PFS as time-to-event outcomes. Methodology and results for ITCs of binary outcomes are available in Appendix 4 of the CS.

The company demonstrated that the PH assumption does not hold for OS and PFS in either the OAK trial (CS, Figure 20 and Figure 21) or the POPLAR trial (see Section 10.1). The company therefore used an ITC methodology that does not rely on the PH assumption, namely one using fractional polynomial (FP) models under a Bayesian framework in WinBUGS statistical software.[59] Specifically, the company employed the method of network meta-analysis of FPs, developed by Jansen.[60]

Under the assumption of PH, the HR is represented as a single parameter (i.e. a number) that is assumed to be constant over time. This alternative approach using FPs is designed to model the hazard function with multiple parameters as a function of time, allowing the HR to change over time in the presence of non-PH. FP models of any 'order' can be fitted to time-to-event data to capture the shape of the hazard functions; 1st order FP models model time as a function with one additional parameter, 2nd order FP models model time as a function with two additional parameters, and so on. However, as the order of the FP model increases, so too does the statistical complexity required to fit the model. Therefore, the company restricted their analysis to 1st and 2nd order FP models only; a range that the company considered broad enough to model the hazard function shapes of the given example.

Fixed effects (FE) FP models were fitted in the first instance, with random effects (RE) FP models fitted subsequently, if data allowed. Five FP models were considered; two 1st order FP models (equivalent to Weibull and Gompertz models) and three 2nd order FP models, herein referred to as models 2nd order (1), 2nd order (2) and 2nd order (3). Under the Bayesian framework, uninformative prior distributions, as outlined by Jansen,[60] were used in all analyses. Further methodological details including the statistical code of the FP models are available in Appendix 4 of the CS.

ITCs were conducted with individual participant data from the OAK and POPLAR trials and survival proportions across monthly time intervals were extracted from digitalised K-M curves. A 5-year time horizon for OS and a 2.5-year time horizon for PFS were used for presenting the time-dependent results of the ITC (expected difference in survival and functional HRs).

The ERG is satisfied that the company has applied the methods described by Jansen[60] appropriately (comparing the statistical code outlined in Appendix 4 of the CS to the template statistical code provided in the Appendix of the Jansen paper[60]) and that the restriction of analyses to 1st and 2nd order FP models was justified.

The ERG notes that due to the lack of a closed loop within the network (Figure 2), the HRs (modelled as FPs) generated by the ITCs are based on indirect evidence only. Subsequently, ITC methodological assumptions of consistency of direct and indirect evidence cannot be investigated statistically. The ERG considers that the unknown validity of this consistency assumption should be taken into account when interpreting numerical results, particularly for the indirect comparison between atezolizumab and nintedanib+docetaxel, where no direct evidence exists.

The ERG notes that the company fitted five FE FP models but, within the CS, only provided numerical results for the best fitting model for OS and PFS. The best fitting model was judged by the company according to the Deviance Information Criteria (DIC) statistic and visual inspection of fitted HR functions. Furthermore, the company repeated the ITC with RE but only for the best fitting FE FP model and interpreted the presence of heterogeneity as a difference in the DIC of the FE and RE models if greater than five points. The company cites two references to support this interpretation.[61,62] The ERG agrees that these references suggest that DIC (along with the residual deviance statistic) can be used to compare the fit of FE and RE models, but cannot find any mention within these references of using a difference in DIC of at least five points to indicate that heterogeneity is present in analyses.

The ERG considers that the DIC is a measure of model fit rather than of statistical heterogeneity and that choices between FE and RE models within an ITC should be made taking into account consistency of trial designs, populations and evidence sources,[61] rather than solely on model fit. The ERG also notes that the methods employed by the company allow for estimation of a heterogeneity parameter (referred to as SD by Jansen[60]) for all RE models, and considers that this SD is a more appropriate measure of statistical heterogeneity than the DIC statistic.

As part of the clarification process, the ERG asked the company to provide ITC results for all five FP models for the network outlined in Figure 2, fitted with both FE and RE and, for the RE FP models, estimates of SD. These results are reported in Section 4.6.3.

4.6.3 Results from the company's indirect treatment comparisons using a reduced network and total trial population data

As noted at the start of Section 4.6, the ERG, as part of the clarification process, asked the company to repeat the ITCs presented in the CS using a reduced network of relevant comparators (outlined in Figure 2). All results reported in this section relate to this reduced network, not to the results relating to the larger network, as reported within the CS.

Overall survival

Results from all FP models fitted to the reduced network (outlined in Figure 2) are shown in Table 24, Figure 3 and Figure 4. The company provided further survivor plot figures as a measure of the visual fit of the survival curves from the FE FP models; these plots are available in Section 10.7.1.

In the original ITC analysis presented in the CS, the company disregards the 2nd order models based on the fitted curves showing a survival 'plateau'; in other words, the curves flattened and a proportion of participants did not experience the event during the time horizon.

From visual inspection of the survivor plots of the reduced network (see Section 10.7.1) the ERG notes that the 2nd order (2) and 2nd order (3) models do begin to flatten at around 24 months, becoming completely flat at around 48 months but the 1st order models and 2nd order model (1) all appear to be of a visually similar shape, tending toward zero by the end of the 5-year time horizon (i.e. all participants will have experienced an event after 5 years).

In the clarification response letter, the company states that for the FE FP, the Weibull model is still the best fitting model in the reduced network, according to DIC and the fitted curves. The HR functions for the Weibull FE model were provided graphically by the company and are available in Section 10.7.1.

Table 24 OS results of FP models, model fit and heterogeneity

FP model	Expected survival difference in months (95% CrI)		DIC	SD (95% CrI)
	Atezolizumab vs docetaxel	Atezolizumab vs nintedanib+docetaxel		
Weibull, FE	5.71 (3.49 to 8.03)	4.74 (2.13 to 7.60)	910.4255	NA
Weibull, RE	5.79 (-8.05 to 25.82)	4.82 (-26.37 to 28.66)	911.4979	0.368 (0.013 to 1.872)
Gompertz, FE	6.82 (3.98 to 9.77)	6.01 (2.69 to 9.26)	934.1241	NA
Gompertz, RE	6.94 (-8.69 to 27.53)	5.93 (-25.12 to 27.70)	935.3138	0.373 (0.012 to 1.838)
2nd order (1)[a], FE	6.44 (3.55 to 9.55)	5.71 (2.09 to 9.32)	837.1486	NA
2nd order (1)[a], RE	6.63 (-8.06 to 27.42)	5.56 (-25.12 to 31.03)	838.3337	0.384 (0.012 to 1.824)
2nd order (2)[b], FE	6.72 (3.54 to 10.12)	6.01 (2.06 to 9.97)	837.6918	NA
2nd order (2)[b], RE	7.15 (-9.01 to 31.12)	6.05 (-25.96 to 34.04)	839.1147	0.379 (0.011 to 1.851)
2nd order (3)[c], FE	6.79 (3.33 to 10.15)	5.84 (1.47 to 9.97)	853.9698	NA
2nd order (3)[c], RE	7.15 (-9.01 to 31.12)	6.05 (-25.96 to 34.04)	854.9049	0.365 (0.010 to 1.858)

[a] 2nd order model (1) corresponds to a model of the form: log hazard=beta0+beta1(log t)+beta2(log t)^2; CS, page 108
[b] 2nd order model (2) corresponds to a model of the form: log hazard=beta0+beta1(log t)+beta2(t); CS, page 108
[c] 2nd order model (3) corresponds to a model of the form: log hazard=beta0+beta1(t)+beta2(t*log t), CS, page 108
CrI=credible interval; FE=fixed effects; FP=fractional polynomial; DIC=deviance information criterion; NA=not applicable OS=overall survival; RE=random effects; SD=standard deviation of the heterogeneity parameter
Source: Company response to ERG clarification letter, adapted from Figure 18, Figure 20, Figure 22, Figure 24, Figure 26, Figure 38, Figure 39, Figure 40, Figure 41, Figure 42, Table 2

The ERG suggests that according to the model fit criteria defined by the company, 2nd order model (1) could also be deemed to be the best fitting model, but notes that any judgement of model fit is subjective and that numerical results of all FE FP models are similar (Table 24, Figure 3 and Figure 4).

Table 24, Figure 3 and Figure 4 show expected difference in survival (months) according to all FP models fitted in the FE and RE ITCs. The ERG notes that across all ten models fitted, the expected difference in survival is very similar, ranging between 5.7 and 7.2 months for atezolizumab compared to docetaxel and between 4.7 and 6.1 months for atezolizumab compared to nintedanib+docetaxel. The ERG also notes that heterogeneity seems to be present in all RE FP models according to SD as defined by Jansen (i.e., SD>0).[60] The SD range is estimated to be 0.36 to 0.39 across the five RE models (Table 24). Also, compared to the FE models, the resulting 95% CrI of the expected survival difference is substantially larger and crosses the line of no effect for all five RE models (Figure 3 and Figure 4).

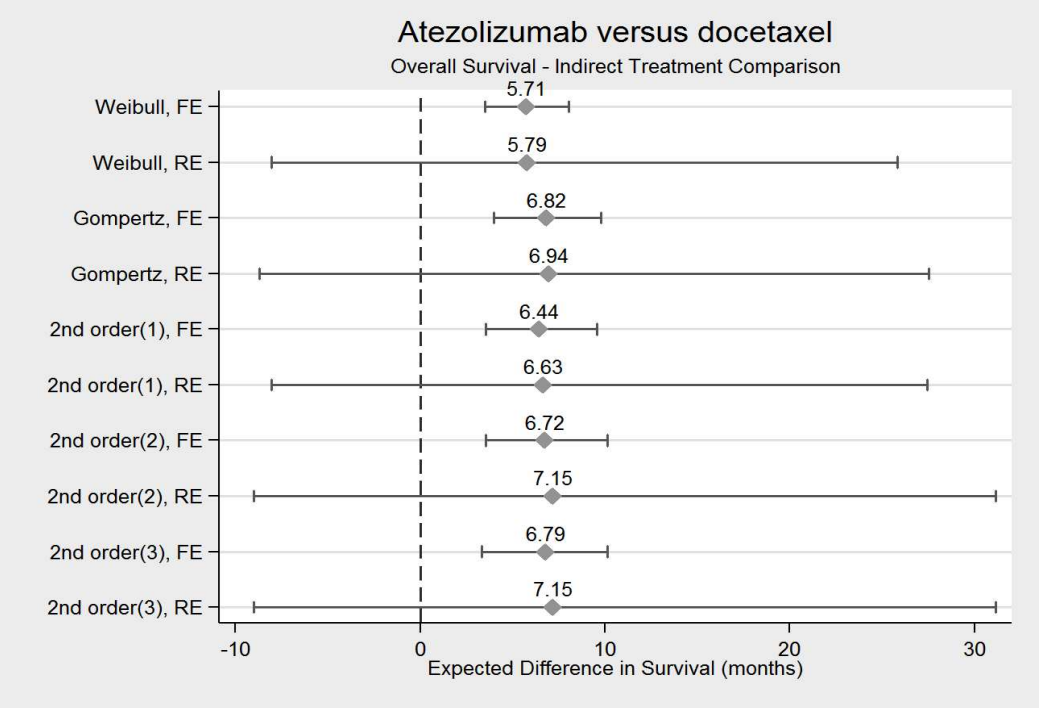


Figure 3 Results of FE and RE FP models, expected difference in OS (months) and 95% CrI for atezolizumab compared to docetaxel

See Table 10 for definitions of 2nd order models (1) (2) and (3)
CrI=credible interval; FE=fixed effects; FP=fractional polynomial; OS=overall survival; RE=random effects
Source: Company response to ERG clarification letter, adapted from Figure 18, Figure 20, Figure 22, Figure 24, Figure 26, Figure 38, Figure 39, Figure 40, Figure 41, Figure 42

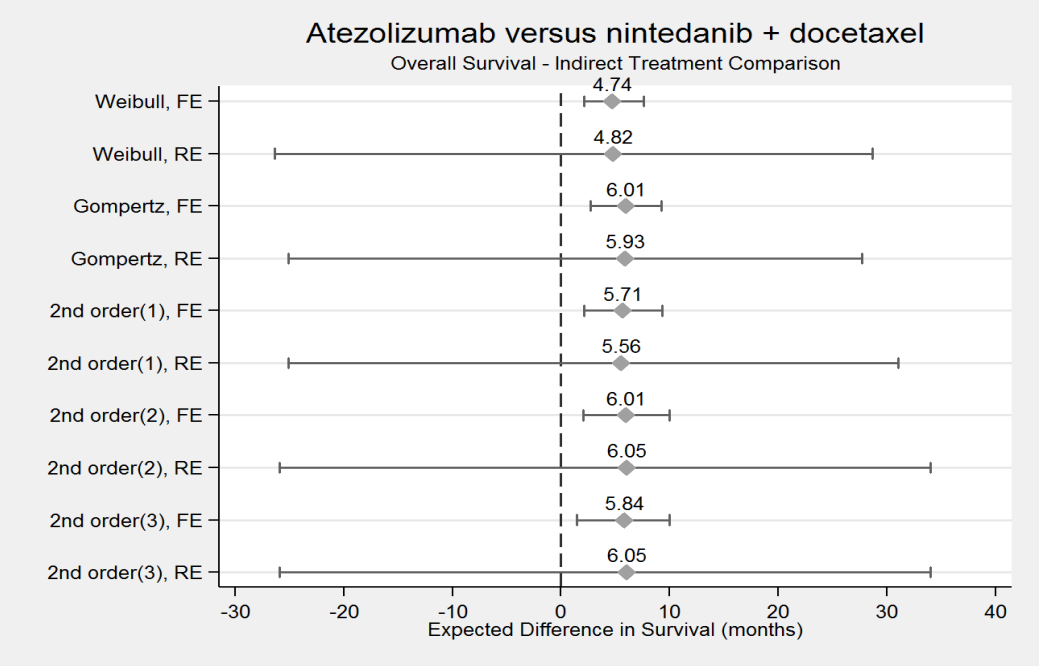


Figure 4 Results of FE and RE FP models, expected difference in OS (months) and 95% CrI for atezolizumab compared to nintedanib+docetaxel

See Table 24 for definitions of 2nd order models (1) (2) and (3)
CrI=credible interval; FE=fixed effects; FP=fractional polynomial; OS=overall survival; RE=random effects
Source: company response to ERG clarification letter, adapted from Figure 18, Figure 20, Figure 22, Figure 24, Figure 26, Figure 38, Figure 39, Figure 40, Figure 41, Figure 42

The ERG was not provided with survivor plots for the RE FP models so was unable to visually inspect the fit of survival curves. Therefore, the ERG can only judge the fit RE models based on the DIC alone, which the ERG considers to be similar for each RE FP model fitted.

In the clarification response letter, the company states that an assessment of heterogeneity is difficult for a network as small as the reduced network and such a small network will result in RE models with wide 95% CrIs. The company considers that the model with the lowest DIC (FE or RE) depicts the best fit to the data and that the presence of heterogeneity is indicated by a difference between the model DIC scores of greater than five points. The ERG questions how, if there is no statistical heterogeneity present in the network as defined by the company (i.e. a difference in DIC of less than five points), the same model fitted with FE and RE can show a wide range of credible results. For example, when atezolizumab is compared with docetaxel, the Weibull FE model generates an expected difference in survival of 5.71 (95% CrI: 3.49 to 8.03) months while Weibull RE model generated an expected difference in survival of 5.79 (95% CrI: -8.05 to 25.82) months (Table 24).

The results suggest that the best estimate of the expected difference in OS is around 6 to 7 months for atezolizumab versus docetaxel (compared to a median OS gain of 4.2 months and 2.9 months in the OAK and POPLAR trials from direct evidence) and around 5 to 6 months for atezolizumab versus nintedanib+docetaxel. The ERG notes that the precision and, therefore, the reliability of the ITC estimates are influenced by the choice of FP model and by potential statistical heterogeneity in the network which has not been acknowledged by the company, or accounted for in any ITC analyses.

Progression-free survival

Results from all FP models fitted to the reduced network (outlined in Figure 2) are shown in Table 25, Figure 5 and Figure 6. The company provided further survivor plot figures as a measure of the visual fit of the survival curves from the FE FP models; these plots are available in Section 10.7.2.

In the original ITC analysis described in the CS, the company disregards the 2nd order models based on visual inspection of the fitted curves. From visual inspection of the survivor plots of the reduced network (Section 10.7.2) the ERG notes that all models seem to 'plateau' at around 12 months, with the extent of the plateau being more prominent in the 2nd order models than the 1st order models.

In the clarification response letter, the company states that, within the reduced network for the FE FP, the Weibull model is the best fitting model; however, within the CS, the Gompertz

model was judged to be the best fitting model (CS, Table 34). The company provided HR functions for the Weibull FE model graphically and are available in Section 10.7.2).

Table 25 PFS results of ITC FP models, model fit and heterogeneity

FP model	Expected survival difference in months (95% CrI)		DIC	SD (95% CrI)
	Atezolizumab vs docetaxel	Atezolizumab vs nintedanib+docetaxel		
Weibull, FE	0.64 (0.01 to 1.32)	-0.30 (-1.39 to 0.70)	1123.198	NA
Weibull, RE	0.64 (-3.36 to 10.80)	-0.31 (-11.21 to 9.59)	1124.86	0.328 (0.010 to 1.832)
Gompertz, FE	0.53 (-0.12 to 1.27)	-0.43 (-1.61 to 0.66)	1157.567	NA
Gompertz, RE	0.50 (-3.30 to 9.53)	-0.42 (-12.32 to 9.03)	1159.318	0.313 (0.010 to 1.837)
2nd order (1)[a], FE	0.72 (-0.20 to 1.65)	0.77 (-1.18 to 2.15)	874.2588	NA
2nd order (1)[a], RE	0.79 (-3.83 to 11.19)	0.76 (-11.20 to 11.39)	875.9772	0.320 (0.008 to 1.838)
2nd order (2)[b], FE	0.80 (-0.12 to 1.75)	0.79 (-1.74 to 2.25)	974.306	NA
2nd order (2)[b], RE	0.78 (-4.19 to 12.53)	0.70 (-12.38 to 11.99)	975.6571	0.340 (0.012 to 1.849)
2nd order (3)[c], FE	0.88 (-0.08 to 1.90)	0.54 (-2.34 to 2.36)	1056.233	NA
2nd order (3)[c], RE	0.86 (-4.19 to 11.87)	0.49 (-13.03 to 12.35)	1057.436	0.308 (0.010 to 1.868)

[a] 2nd order model (1) corresponds to a model of the form: log hazard=beta0+beta1(log t)+beta2(log t)^2; CS, page 1118
[b] 2nd order model (2) corresponds to a model of the form: log hazard=beta0+beta1(log t)+beta2(t); CS, page 111
[c] 2nd order model (3) corresponds to a model of the form: log hazard=beta0+beta1(t)+beta2(t*log t), CS, page 111
CrI=credible interval; FE=fixed effects; FP=fractional polynomial; DIC=deviance information criterion; NA=not applicable, PFS=progression-free survival; RE=random effects; SD=standard deviation of the heterogeneity parameter
Source: company response to ERG clarification letter, adapted from Figure 28, Figure 30, Figure 32, Figure 34, Figure 36, Figure 43, Figure 44, Figure 45, Figure 46, Figure 47, Table 3.

The ERG suggests that, according to the model fit criteria defined by the company, the 2nd order models with lower DIC values could be deemed to fit survival data better than the 1st order models, but notes that any judgement of model fit is subjective and that numerical results of all FE FP models are similar (Table 25, Figure 5 and Figure 6).

Table 25, Figure 5 and Figure 6 show expected difference in survival (months) according to all FP models using FE and RE. The ERG notes that, across all ten of the fitted models, the expected difference in PFS is similar, and not statistically significant for all except one result (Weibull FE model for atezolizumab compared to docetaxel, 0.64 [0.01 to 1.32] months).

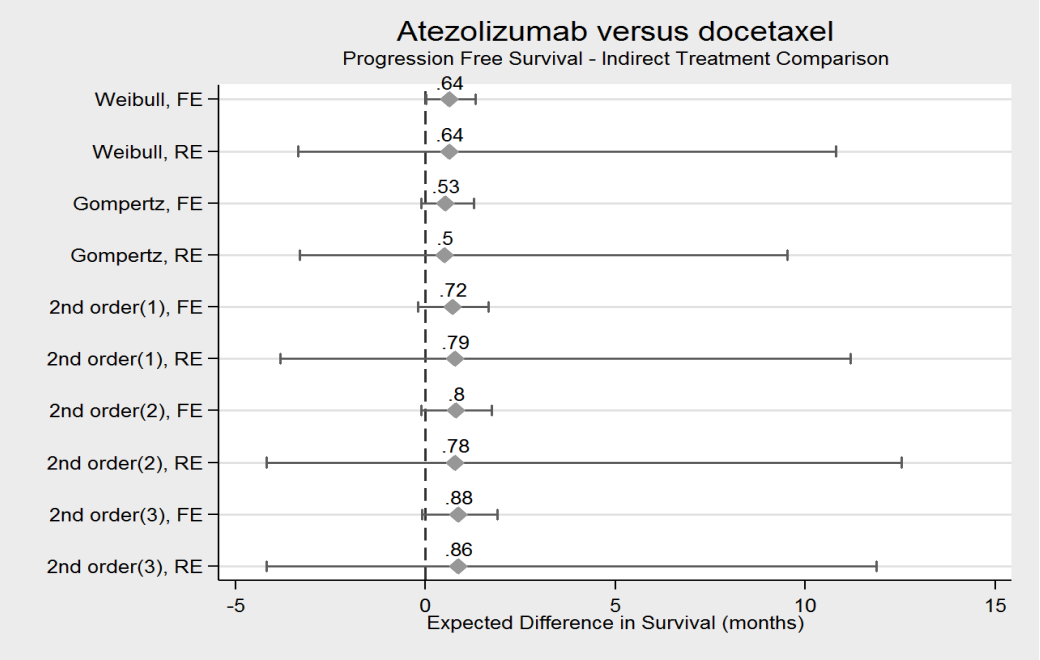


Figure 5 Results of FP models with FE and RE, expected difference in PFS (months) and 95% CrI for atezolizumab compared to docetaxel

See Table 11 for definitions of 2nd order models (1) (2) and (3)
CrI=credible interval; FE=fixed effects; FP=fractional polynomial; PFS=progression-free survival; RE=random effects
Source: company response to ERG clarification letter, adapted from Figure 28, Figure 30, Figure 32, Figure 34, Figure 36, Figure 43, Figure 44, Figure 45, Figure 46, Figure 47

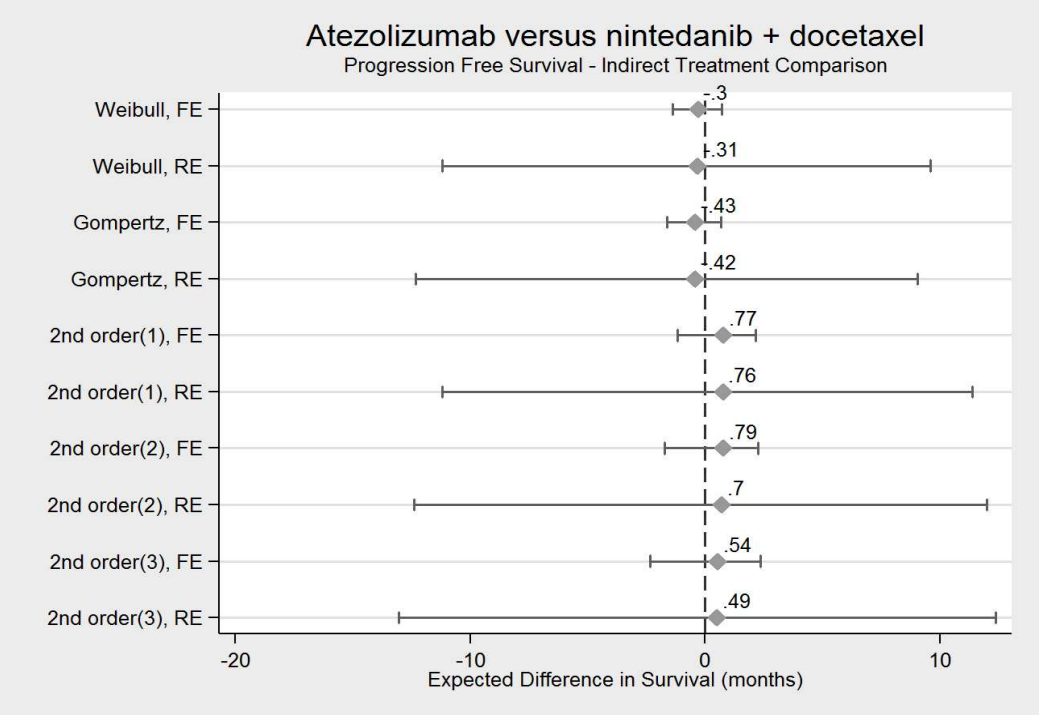


Figure 6 Results of FP models with FE and RE, expected difference in PFS (months) and 95% CrI for atezolizumab compared to nintedanib+docetaxel

See Table 11 for definitions of 2nd order models (1) (2) and (3)
CrI=credible interval; FE=fixed effects; FP=fractional polynomial; PFS=progression-free survival; RE=random effects
Source: company response to ERG clarification letter, adapted from Figure 28, Figure 30, Figure 32, Figure 34, Figure 36, Figure 43, Figure 44, Figure 45, Figure 46, Figure 47

The ERG notes also that heterogeneity seems to be present in all RE FP models with SD (as defined by Jansen[60]) estimated to range from 0.31 to 0.34 across the five RE models (Table 25) and the resulting 95% CrI of the expected survival difference being substantially larger for all RE models compared to the FE models (Figure 5 and Figure 6).

The company did not provide survivor plots for the RE FP models and so the ERG was unable to visually inspect the fit of survival curves. Therefore, the ERG can only judge the fit of FE and RE models based on the DIC alone, which appears to be similar for each FP model fitted with FE and each fitted with RE (Table 25). As discussed for the ITC results for OS, the ERG questions how the same model fitted with FE and RE can show such a different range of credible results if there is no statistical heterogeneity present in the network as defined by the company (i.e. a difference in DIC of less than five points).

The ITC results consistently suggest that there are no statistically significant differences in expected PFS when comparing atezolizumab versus docetaxel (in line with the results of the OAK and POPLAR trials) and when comparing atezolizumab versus nintedanib+docetaxel. The ERG considers that the precision and reliability of the expected differences in PFS are influenced by the choice of FP model and by potential heterogeneity in the network, which has not been acknowledged by the company or accounted for in any of the ITC analyses.

4.6.4 Results from additional indirect comparisons requested by the ERG

As part of the clarification process, the ERG asked the company to perform additional ITCs. These are described in this section. The ERG assumes that the methodology the company applied to undertake these additional ITCs is the same as the methodology outlined in Section 4.6.2.

The ERG suggests that the results of these additional ITCs should be interpreted with caution, due to concerns regarding heterogeneity and the impact upon the reliability of FP results within the network as discussed in Section 4.6.3.

Adenocarcinoma histology

As outlined in Section 4.6.1 the ERG requested that the company repeat the ITC for the reduced network of relevant comparators (outlined in Figure 2) in the adenocarcinoma subgroups of the three trials as nintedanib+docetaxel is licensed only for participants with adenocarcinoma histology.

In the clarification response letter, the company states that the TSAPs for the OAK and POPLAR trials did not include subgroups according to the presence of adenocarcinoma and therefore did not provide results for the ITC requested by the ERG. The ERG anticipated that adenocarcinoma subgroups may not have been defined in the OAK and POPLAR trials and therefore, if this were the case, requested alternatively that the company repeat the ITCs using data from the non-squamous subgroups of the OAK and POPLAR trials and the adenocarcinoma subgroup of the LUME-Lung 1 trial. The company, however, provided results for a comparison between atezolizumab within its intended licensed population (total OAK and POPLAR trial populations) with nintedanib+docetaxel (in the subgroup of patients with adenocarcinoma histology). The ERG notes that these results are, therefore, derived from comparing non-equivalent populations and thus should also be treated with extreme caution.

The company applied the Weibull FE FP model; results for OS and PFS are provided in Table 26 and plots of HR functions provided by the company are provided in Section 10.7.3.

Table 26 Expected survival differences: atezolizumab (total population) and nintedanib+docetaxel (adenocarcinoma subgroup)

Expected survival difference in months (95% CrI)*		
Outcome	**Atezolizumab vs docetaxel**	**Atezolizumab vs nintedanib+docetaxel**
OS	5.84 (3.68 to 8.07)	3.33 (-0.16 to 6.74)
PFS	0.68 (-0.04 to 1.46)	-0.07 (-1.76 to 1.28)

*Results came from the 'best fitting' Weibull FE FP model
CrI=credible interval; FE=fixed effects; FP=fractional polynomial; PFS=progression-free survival; OS=overall survival
Source: company response to ERG clarification letter, adapted from Figure 10, Figure 12

The ERG notes that when restricting the ITC to the adenocarcinoma subgroup for nintedanib+docetaxel, when comparing atezolizumab to nintedanib+docetaxel, the expected OS difference is reduced from around 4.74 months (see Table 24) to 3.33 months and the result is no longer statistically significant. The expected PFS difference when comparing atezolizumab to nintedanib+docetaxel is similar to the results showed in Table 25. The ERG also notes that OS and PFS results for the comparison of atezolizumab versus docetaxel are similar to those shown in Table 24 and Table 25.

The company states that using the 'total population' of the LUME-Lung 1 trial to conduct a 'like for like' comparison between atezolizumab and nintedanib+docetaxel is 'not anticipated to significantly affect overall results' (CS, Section 4.10). However, this statement is not supported by the additional results provided by the company (Table 26) which show that, when restricting the ITC to the adenocarcinoma subgroup of the LUME-Lung 1 trial, treatment with atezolizumab no longer shows a statistically significant difference in OS compared to nintedanib+docetaxel. The ERG considers that the results of an ITC conducted within the

adenocarcinoma subgroups of the OAK, POPLAR and LUME-Lung 1 trials are needed to fully appreciate the impact of the choice of trial population on comparative efficacy.

Inclusion of pembrolizumab in the network

The ITCs presented in the CS included comparators to atezolizumab that were not considered in the final scope issued by NICE. The ERG notes that pembrolizumab, which was specified in the final scope issued by NICE for this appraisal, was included within the ITCs for OS and PFS. However, the company did not present the ITC results for the comparison of atezolizumab versus pembrolizumab, as the company does not consider pembrolizumab to be a relevant comparator for this appraisal (see Section 3.3.2). As further outlined in Section 3.3.2, the ERG considers that pembrolizumab is an appropriate comparator but only for the population for which it is currently recommended by NICE (treatment of patients with PD-L1 ≥1% NSCLC after chemotherapy). Therefore, as part of the clarification process, the ERG asked the company to carry out an ITC for the network outlined in Figure 7, for OS and PFS. This network includes three trials, the OAK and POPLAR trials and the KEYNOTE-010 trial.[63] Further details of the design and participant characteristics of the KEYNOTE-010 trial can be found within the primary reference[63] and within Appendix 4 of the CS. Overall, the ERG considered the characteristics of the OAK, POPLAR and KEYNOTE-010 trials to be broadly similar and therefore suitable for inclusion in the same ITC.

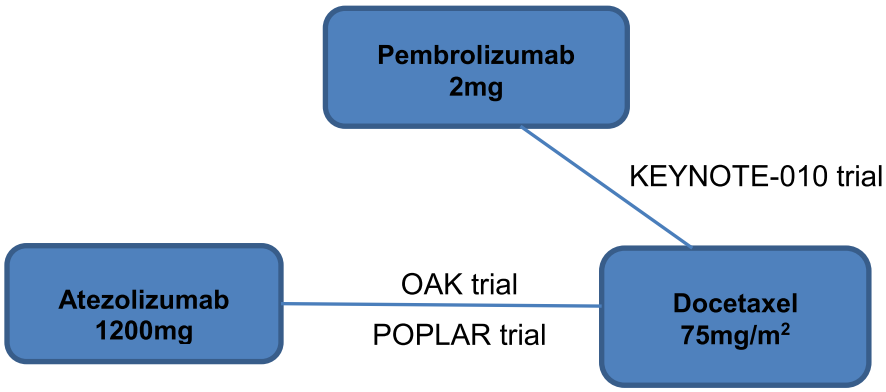


Figure 7 Network plots for ITCs of OS and PFS including pembrolizumab

The company provided results of the additional ITC as requested, applying the Weibull FE FP model (see Section 4.6.3 for further details of model fit); results for OS and PFS are provided in Table 26 and plots of HR functions provided by the company are provided in Section 10.4). Expected survival differences are shown in Table 27.

Table 27 Expected survival differences including pembrolizumab

Expected survival difference in months (95% CrI)*		
Outcome	**Atezolizumab vs docetaxel**	**Atezolizumab vs pembrolizumab**
OS	5.79 (3.63 to 8.05)	-0.24 (-5.38 to 4.44)
PFS	1.17 (0.29 to 2.03)	-0.30 (-2.17 to 1.40)

*Results came from the 'best fitting' Weibull FE FP model
CrI=credible interval; FE=fixed effects; FP=fractional polynomial; PFS=progression-free survival; OS=overall survival
Source: company response to ERG clarification letter, adapted from Figure 14, Figure 16

Results for the comparison of atezolizumab versus docetaxel are similar to those shown in Table 24 for OS and the expected difference in PFS is slightly greater than the differences shown in Table 25. The ERG notes that there is no statistically significant difference between atezolizumab and pembrolizumab in terms of OS or PFS.

In the clarification response letter, the company states that the ITC of atezolizumab in its licensed indication versus pembrolizumab in its licensed indication (PD-L1 positive), compares two non-equivalent populations, and hence there is a risk the relative clinical benefits of pembrolizumab are overestimated. The company emphasises that this analysis should not be considered as a robust and true reflection of the comparative efficacy of pembrolizumab versus atezolizumab. The ERG agrees with this statement and advocates extreme caution when interpreting comparative results of atezolizumab versus pembrolizumab.

The ERG considers that robust analysis approaches are important but should not come at the expense of making inappropriate comparisons, such as including data for patients not specified in the licensed population for treatment with nintedanib+docetaxel.

The ERG considers that the approach to the ITC is influenced by a range of factors (e.g., comparators and population selected, type of FP model chosen and the use of FE or RE). This means that it is difficult to identify the most appropriate combination of factors to use to generate and interpret ITC results.

4.7 Additional work on clinical effectiveness undertaken by ERG

OS data from the OAK trial, for PD-L1 subgroups, were published in January 2017.[23] The ERG has reproduced these results for information (Table 28).

Table 28 OS in the ITT population and PD-L1 subgroups

Population	n (%)	Median OS (months)		HR (95% CI)
		Atezolizumab	Docetaxel	
ITT	850 (100)	13.8	9.6	0.73 (0.62 to 0.87)
TC3 or IC3	137 (16)	20.5	8.9	0.41 (0.27 to 0.64)
TC2/3 or IC2/3	265 (31)	16.3	10.8	0.67 (0.49 to 0.90)
TC1/2/3 or IC1/2/3	463 (54)	15.7	10.3	0.74 (0.58 to 0.93)
TC0 and IC0	379 (45)	12.6	8.9	0.75 (0.59 to 0.96)

CI=confidence interval; HR=hazard ratio; IC=immune cell; ITT=intention to treat; OS=overall survival; TC=tumour cell
Source: Rittmeyer[23]

4.8 Conclusions of the clinical effectiveness section

Discrepancies between the decision and the final scope issued by NICE

The ERG considers that the submitted evidence largely reflects the decision problem defined in the final scope issued by NICE. However, there are a number of exceptions:

- Comparators:
  - the comparison of the efficacy of treatment with atezolizumab versus nintedanib+docetaxel should have been carried out using data from the population for which nintedanib+docetaxel is licensed, i.e. patients with adenocarcinoma rather than the whole trial population
  - the company should have compared the efficacy of treatment with atezolizumab versus pembrolizumab for the population for which pembrolizumab is licensed and recommended by NICE (people with PD-L1 positive NSCLC).
- Subgroups:
  - it is specified within the final scope issued by NICE that, if evidence allows, consideration will be given to subgroups based on biological markers. As analyses by level of PD-L1 expression are specified in the protocols for the OAK and POPLAR trials, full results from both trials (rather than just by no measurable PD-L1 expression and ≥1% PD-L1 expression from the OAK trial) should have been provided in the CS.

Direct clinical evidence

The direct clinical effectiveness evidence for the treatment of atezolizumab versus docetaxel was derived from the OAK and POPLAR trials. The ERG highlights the following points:

- both these trials were of good quality and were both well conducted; patient characteristics were balanced across the groups and statistical methods were generally appropriate. However, the open-label design provides the opportunity for investigator-assessed outcomes to be biased
- the studies included some UK sites and clinical advice to the ERG is that patients recruited to these trials are broadly similar to those treated within the NHS, with the caveat that patients in the trials were slightly younger and fitter than NHS patients
- within the OAK and POPLAR trials, docetaxel is administered intravenously until disease progression or unacceptable toxicity. However, clinical advice to the ERG is that, within the NHS, patients typically only receive between four and six cycles of treatment

- in both the OAK and POPLAR trials, the treatment stopping rules for patients receiving atezolizumab and docetaxel differed: treatment with atezolizumab was administered as long as patients experienced a clinical benefit (as assessed by an investigator) in the absence of unacceptable toxicity or symptomatic deterioration attributed to disease progression and treatment with docetaxel was continued until disease progression or unacceptable toxicity
- the ERG agrees with the company that the AE data from the OAK trial are consistent with the known AE profile of atezolizumab and that no new safety concerns have been highlighted. In addition, treatment with atezolizumab is well-tolerated in comparison to treatment with docetaxel
- the ERG notes that OS and PFS HRs calculated from OAK and POPLAR trial data must be interpreted with caution due to non-PH (as demonstrated by the company). However, the ERG acknowledges that the methodology requiring the PH assumption was pre-specified and the company could not have known at the time this methodology was proposed that the PHs assumption would be violated
- results from both the OAK and POPLAR trials show that treatment with atezolizumab is associated with a statistically significant and clinically meaningful improvement in median OS (4.2 months in the OAK trial and 2.9 months in the POPLAR trial) compared to docetaxel in patients with ECOG PS 0 and 1
- in the OAK trial, this statistically significant gain in OS is observed regardless of histology and PD-L1 status. However, in the POPLAR trial, statistically significant improvement is observed only in the non-squamous histology subgroup and for individuals with NSCLC of ≥1% PD-L1 expression
- improvement in OS with atezolizumab compared with docetaxel is also generally consistent across baseline characteristics in both trials
- no statistically significant difference in investigator-assessed PFS was observed between atezolizumab and docetaxel groups in either trial.

Indirect clinical evidence

The ERG considers that the company applied the ITC methodology using FP models appropriately but does not agree with the company's criteria of using the DIC statistic (a measure of model fit) for assessing the presence of heterogeneity in the analyses. The ERG does not support the ITC approach taken by the company as:

- the main network includes comparators that are not listed in the final scope issued by NICE
- when considering the relative efficacy of atezolizumab versus nintedanib+docetaxel, the company compared effectiveness relating to the whole LUME-Lung 1 trial population, rather than considering the relevant population, i.e. the population for which nintedanib+docetaxel is licensed (patients with adenocarcinoma)
- the company was not justified in excluding pembrolizumab from the ITC network of comparators relevant to this appraisal.

The ERG asked the company to provide ITC results from a reduced network of comparators comprising those listed in the final scope issued by NICE. Based on this (reduced) network,

i.e., using data from the total populations of the OAK, POPLAR and LUME-Lung 1 trials, the company's FP ITC results suggest that:

- the company's best estimate of expected difference in OS is around 6 to 7 months for atezolizumab versus docetaxel (compared to median OS gains of 4.2 months and 2.9 months from the OAK and POPLAR trials respectively)
- the company's best estimate of expected difference in OS is around 5 to 6 months for atezolizumab versus nintedanib+docetaxel
- there appears to be no significant difference in PFS when comparing atezolizumab to docetaxel and when comparing atezolizumab to nintedanib+docetaxel.

The ERG also asked the company to undertake two further subgroup analyses. However, the company undertook these using non-equivalent populations and results should be viewed with extreme caution:

- based on a (reduced) network using data the ITT populations from the OAK and POPLAR trials and the adenocarcinoma population from the LUME-Lung 1 trial, the company's FP ITC results suggest that the company's best estimate of expected difference in OS for atezolizumab versus nintedanib+docetaxel is 3.33 months (compared to 4.74 months when the analysis was carried out using LUME-Lung 1 trial total population) and is no longer statistically significant

- based on a (reduced) network using data the ITT populations from the OAK and POPLAR trials, and the KEYNOTE-010 trial (a trial assessing the efficacy of pembrolizumab as a first-line treatment for metastatic NSCLC in adults whose tumours express PD-L1 with a ≥50% tumour proportion score) the company found no statistically significant difference in OS or PFS when comparing atezolizumab (total population') versus pembrolizumab (PD-L1 positive NSCLC patients).

The ERG highlights that the precision and reliability of all additional results are influenced by the choice of FP model and greatly influenced by potential statistical heterogeneity in the network which has not been acknowledged by the company or accounted for in any ITC analyses. In summary, the ERG considers that the approach to the ITC is influenced by a range of factors (e.g., comparators and population selected, type of FP model chosen and the use of FE or RE). This means that it is difficult to identify the most appropriate combination of factors to use to generate ITC results. Furthermore, the ERG considers that the expected survival results generated by the FP ITC are difficult to interpret.

5 COST EFFECTIVENESS

5.1 Introduction

This section provides a structured critique of the economic evidence submitted by the company in support of the use of atezolizumab for treating locally advanced or metastatic NSCLC after prior chemotherapy. The two key components of the economic evidence presented in the CS are (i) a systematic review of the relevant literature and (ii) a report of the company's de novo economic evaluation. The company has provided an electronic copy of their economic model, which was developed in Microsoft Excel.

5.2 Objective of the company's cost effectiveness review

The company's systematic review was carried out to identify cost effectiveness evidence for atezolizumab for the treatment of locally advanced or metastatic NSCLC after prior chemotherapy. The stated aim of the review was to identify economic evidence from all lines of metastatic NSCLC to support the development of cost effectiveness models for atezolizumab. Electronic searches were carried out on 4th September 2016 and hand searches were performed on 21st November 2016. The databases searched and the initial date span for each search are summarised in Table 29.

Table 29 Details of searches for the company's economic systematic review

Database	Platform	Date span of search	Date searched
Embase	Embase.com	From database inception (1974) to 3-Sep-2016 (updated daily)	04-Sep-2016
Medline	Embase.com	From database inception (1966) to 3-Sep-2016 (updated daily)	04-Sep-2016
Medline InProcess & e-publications ahead of print	PubMed search interface	From database inception to 17-Nov-2016	04-Sep-2016 initially & weekly alerts received to cut-off date of 18-Nov-2016
NHS Economic Evaluation Database (NHS EED)	Cochrane library	From database inception to 31st March 2015 (database closed)	04-Sep-2016
Health Technology Assessment Database (HTA)	Cochrane library	From database inception to July 2016 (updated monthly)	04-Sep-2016

Source: CS, Table 51

5.2.1 Eligibility criteria used in study selection

The inclusion criteria that were used to select studies are shown in Table 30; more detailed exclusion criteria are presented in Table 52 of the CS. The ERG is satisfied that these criteria

are relevant to the aim of the company's systematic review but notes that they are not focussed on the specific decision problem set out in the final scope issued by NICE. As the inclusion/exclusion criteria are broad, the ERG is satisfied that use of these criteria is acceptable.

Table 30 Inclusion criteria

Characteristic	Inclusion criteria
Population	• Adult patients (16 years+) • Locally advanced or metastatic NSCLC, second/subsequent line
Interventions / comparators	• Licensed and unlicensed pharmacological interventions used in the second/subsequent line within the metastatic setting, compared to each other or to placebo or standard of care • Companion tests + pharmacological agent, if the objective is to assess the pharmacological agent primarily (tagged)
Outcomes	• Evaluation includes both costs and effectiveness/utility measures (need not necessarily report an incremental cost-effectiveness ratio) • Sub-outcomes of interest are: cost components, health states, interim/proxy efficacy measures, safety endpoints
Study design	• Economic evaluations (cost-effectiveness analysis, cost-utility analysis) • Economic evaluations alongside a clinical trial • Health technology assessments
Country	• EMEA countries, USA, Canada, Australia and New Zealand
Perspective	• Payer, societal
Time horizon	• Unlimited
Date limits	• Unlimited
Child citation	• Citation linked to another paper but with unique data
Language	• Any foreign language paper with an English abstract if sufficient information is present in the English abstract to ensure the eligibility criteria are met

Source: CS, Table 52

5.2.2 Included and excluded studies

The company did not identify any studies of atezolizumab in its systematic review. The company presented summary details of 11 studies and related risk of bias assessments (CS, Table 53 and Appendix 7 respectively) and three NICE appraisals (CS, Table 54) that were considered to be relevant to the decision problem; none of these publications included atezolizumab as an intervention or a comparator.

The ERG notes that the company conducted a systematic review from a global perspective (excluding Asia and South America) to support the health technology assessment process for countries including, and beyond, the UK. However, only data from 11 studies relevant to the decision problem, meeting the NICE Reference Case[64] and relevant to UK decision-making were extracted and reported in the CS.

5.2.3 Findings from cost effectiveness review

The company did not report any findings from the cost effectiveness review.

5.3 ERG critique of the company's literature review

The company reports full details of the searches used to identify cost effectiveness evidence in Section 5.1 and Appendix 5 of the CS. These searches included a cost effectiveness filter. The company used population terms but did not include any indication terms; the ERG considers this approach to be appropriate. The ERG notes that the search terms used to describe the population of interest in the economic literature searches were more comprehensive than the terms that were used in the main clinical searches.

The company also searched for HRQoL data and full details of these searches are reported in Appendix 9 of the CS. The searches included a HRQoL filter, broad population search terms and covered the same time period as the cost effectiveness searches. The ERG notes that the company could have used simpler search strings.

The ERG notes that the company went to great lengths to identify relevant economic studies of atezolizumab. However, despite a wide focus, broad inclusion/exclusion criteria and summary descriptions of potentially relevant studies, no studies of atezolizumab were identified for inclusion in the review. The ERG is satisfied that no relevant studies were missed during the review process.

5.4 NICE Reference Case checklist

Table 31 NICE Reference case checklist completed by ERG

Attribute	Reference case	Does the de novo economic evaluation match the reference case?
Decision problem	The scope developed by NICE	Yes
Comparator(s)	As listed in the scope developed by NICE	No, but the company provides justification as to why this is the case
Perspective costs	All direct health effects, whether for patients or, when relevant, carers	Yes
Perspective benefits	NHS and PSS	Partial - patient related direct health effects are considered. No PSS costs have been considered
Form of economic evaluation	Cost utility analysis with fully incremental analysis	Yes
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	Yes
Synthesis of evidence on outcomes	Based on systematic review	Yes
Outcome measure	Health effects should be expressed in QALYs.	Yes
Health states for QALY	Standardised and validated instrument. The EQ-5D is the preferred measure of health-related quality of life in adults	Yes
Benefit valuation	Reported directly by patients and/or carers	Yes
Source of preference data for valuation of changes in HRQoL	Representative sample of the UK population	No. However, UK valuations of data collected during the OAK trial were requested during the clarification process and these were similar to values used in the company model
Discount rate	The same annual rate for both costs and health effects (currently 3.5%)	Yes
Equity	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	Yes
Sensitivity analysis	Probabilistic sensitivity analysis	Yes

EQ-5D=EuroQol-5 dimension; QALY=quality adjusted life year; HRQoL=health-related quality of life; PSS=personal social services

5.4.1 Drummond checklist

Table 32 Critical appraisal checklist completed by the ERG

Question*	Critical appraisal	ERG comment
Was a well-defined question posed in answerable form?	Yes	
Was a comprehensive description of the competing alternatives given?	Yes	
Was the effectiveness of the programme or services established?	Partially	The data come from a RCT but modelling of survival was required. This appears to have resulted in an over-estimate of the effectiveness of the intervention
Were all the important and relevant costs and consequences for each alternative identified?	Yes	Some of the assumptions in the model were unsupported by data
Were costs and consequences measured accurately in appropriate physical units?	Yes	
Were the cost and consequences valued credibly?	Yes	
Were costs and consequences adjusted for differential timing?	Partially	There was an error in the calculation of treatment costs for all arms due to an error in applying a half-cycle correction. There was also an error in the discounting algorithm
Was an incremental analysis of costs and consequences of alternatives performed?	Yes	
Was allowance made for uncertainty in the estimates of costs and consequences?	Yes	
Did the presentation and discussion of study results include all issues of concern to users?	Yes	

*Questions from the Drummond 10-point checklist[65]

5.4.2 Model structure

Overview of the model

The company states that the model is designed to compare the cost effectiveness of atezolizumab versus docetaxel and atezolizumab versus nintedanib+docetaxel in patients with locally advanced or metastatic NSCLC who have progressed during or after prior chemotherapy. The model inputs (efficacy, safety and tolerability) were based on the results of the phase III OAK trial that compared the effectiveness of atezolizumab versus docetaxel. Clinical effectiveness data from the FP ITC were used to inform the comparison of atezolizumab versus docetaxel and atezolizumab versus nintedanib+docetaxel. Results are reported in terms of incremental cost per life year (LY) gained and incremental cost per quality adjusted life year (QALY) gained.

Model structure

The cost effectiveness model presented by the company is a partitioned survival model. The model structure (as shown in Figure 8) is slightly different to the type of model usually submitted to NICE as part of appraisals of interventions to treat metastatic cancer as it comprises three mutually exclusive health states: 'on treatment', 'off treatment' and death (rather than PFS, progressed disease [PD] and death). The company considers that this structure is better suited to the appraisal of atezolizumab than traditionally structured models, as patients receiving atezolizumab (an immunotherapy) are permitted to continue treatment with atezolizumab for some time after disease progression. However, the company explains that the comparators are not bound by this structure. For example, nintedanib+docetaxel, treatment duration, supportive care costs and utilities are all determined through the traditional PFS/PD/Death model; and so too are the supportive care costs associated with treatment with docetaxel. This is in comparison to atezolizumab where 'on treatment' drug costs and utility benefits are determined using a time to treatment discontinuation (TTD) approach.

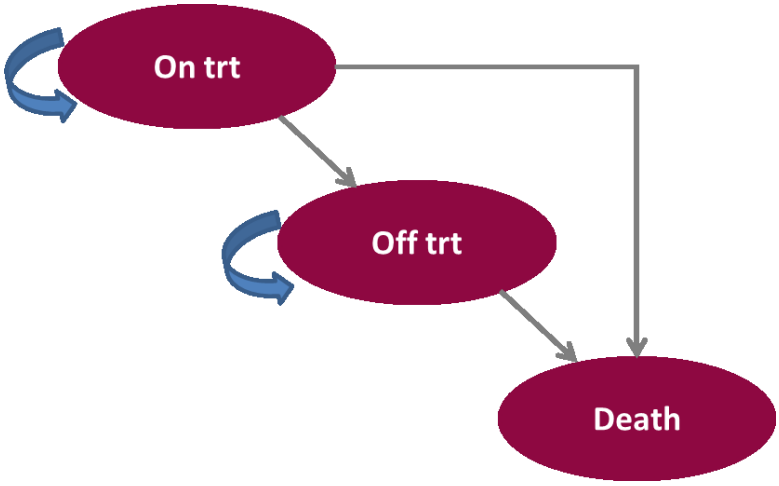


Figure 8 Area under the curve model structure

Trt=treatment
Source:CS, Figure 32

5.4.3 Population

The patient population in the company model is patients with locally advanced or metastatic NSCLC who have progressed during or after prior chemotherapy. The baseline characteristics of the modelled population reflect the characteristics of the patients in the OAK trial.

5.4.4 Interventions and comparators

In the base case, the intervention is atezolizumab, and the comparators are docetaxel and nintedanib+docetaxel. Atezolizumab is implemented in the model as per the anticipated licensed dosing regimen, i.e. fixed dose of 1200mg concentrate solution for IV infusion, administered over 60 minutes for the first infusion and, if well tolerated, as a 30 minute IV infusion every 3 weeks. Atezolizumab is administered beyond progression if the patient is considered to be continuing to receive benefit from treatment. The total drug cost per cycle is estimated to be £3,807.69 (CS, Table 65).

Docetaxel is administered at a dose of 75kg/m^2 every 3 weeks. The weighted average body surface area (BSA) for men and women from the OAK trial was used to estimate the average cost per dose of docetaxel per patient. In the model, full vial sharing is assumed for the administration of docetaxel and the maximum treatment duration is six cycles. The total drug cost per cycle is estimated to be £34.39 (CS, Table 65).

Nintedanib is administered orally (twice daily) as a 200mg soft capsule. In the model, the maximum treatment duration of docetaxel is six cycles. There is a PAS in place for nintedanib. The list price cost per cycle for nintedanib is estimated to be £1,434.07 (CS, Table 65).

Subsequent treatment

The economic model includes costs of subsequent treatment for patients who have progressed during or after initial treatment (see Table 33). At 25 months' follow-up, approximately 13% of patients were still receiving atezolizumab; this means there is no complete dataset of post-discontinuation treatments. The company states (CS, p182) that "…so as not to bias the analysis by giving a falsely low subsequent treatment cost to atezolizumab, an average has been taken by pooling the arms". As per the OAK trial, 45% of all patients were assumed to receive subsequent pharmacological treatment and 55% went on to receive radiotherapy. In line with clinical opinion, the company removes the costs of third-line immunotherapy from the base case analysis and considers the use of radiotherapy as a third-line treatment in a scenario analysis.

Table 33 Cost of subsequent treatment (drug and radiotherapy)

Cost and duration of subsequent drug and radiotherapy treatments	
Average time on subsequent drug treatment	13.59 weeks
Average cost	£1,987.06
Average number of subsequent radiotherapy doses per patient	20.58
Average cost	£1,353.08
Total cost of subsequent treatment	£3,340.14

Source: CS, Section 5.5.2.1

5.4.5 Perspective, time horizon and discounting

The company states that the economic evaluation was undertaken from the perspective of the NHS and Personal Social Services. The time horizon was set at 25 years and, in line with the NICE Guide to the Methods of Technology Appraisal,[64] both costs and outcomes were discounted at 3.5% per annum.

5.4.6 Treatment effectiveness and extrapolation

The primary data source for the company model was the OAK trial. The follow-up period over which trial data were available was shorter than the time horizon of the economic model. Therefore, modelling of OS, PFS and TTD data from OAK trial was required.

Overall survival

The company considered that the survival data available for immunotherapy agents suggest that it is plausible that some patients experience a sustained response. To model this sustained response the company constructed a mixed cure-rate model. The concept is that there is a subgroup of patients with stable disease for whom the risk of death attributable to cancer is equivalent to the risk of death from other causes. Thus, there are two populations,

those with a low risk of death and those with a high risk of death and OS is represented as an average of the two different risks for these two populations.

Following examination of data from the OAK trial, the POPLAR trial data and the NLCA, and consultation with clinicians, the company determined that 2% of patients are likely to be in the low risk of death group, i.e. have a risk of death equivalent to the age-adjusted general population mortality rate.

The company modelled the risk of death for the remaining 98% of the population based on data from the OAK trial. Standard parametric curves were fitted to OAK trial data and the company determined, based on visual assessment, the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC), that the log-logistic distribution was the most appropriate fit.

To construct OS curves for the comparator agents, the curve constructed to represent OS for the population receiving atezolizumab was adjusted using the time-dependent FP log HRs over the span of the extrapolation.

Time to treatment discontinuation

Within the company model TTD K-M data for patients treated with atezolizumab are used directly until the point where 15% of patients are still at risk of an event. From this point, for the duration of the remaining time horizon, the company has used a parametric distribution to represent the proportion of patients still receiving their initial treatment. Based on visual inspection and AIC and BIC, the company utilised a Gamma distribution in the base case analysis.

When modelling TTD for patients receiving docetaxel, OAK trial TTD K-M data were used directly in the model with a maximum treatment duration of six cycles used for costing purposes which is stated in the CS as being consistent with NHS clinical practice in England.

TTD trial data for nintedanib+docetaxel were not available to the company. The company's approach to representing TTD for patients receiving nintedanib+docetaxel was to adjust their representation of PFS for patients receiving atezolizumab using the relevant ITC FP HR. The company modelled PFS for patients receiving atezolizumab using the same methodology as used to construct their TTD model. That is, using OAK trial PFS K-M data directly in the model until 15% of patients were still at risk, at which point the 'best fitting' parametric distribution, which, in this case was, again, considered to be a gamma distribution, was fitted. Again docetaxel was limited to six cycles for costing purposes but nintedanib was administered until progression.

5.4.7 Health-related quality of life

HRQoL data were collected as part of OAK trial using the EQ-5D 3L tool.[56] The values of the utility estimates used in the company model are based on: health states and time to death. Given that patients experience both health state disutility and end of life disutility, the company considers that the two options are complimentary. To capture HRQoL as appropriately as possible, the company divided utilities into four categories reflecting time to death. These values were applied in addition to the 'on treatment' and 'off treatment' health states. A summary of the utility values used in the model is shown in Table 34.

Table 34 Summary of health states utility values – NICE Reference Case

Time period	On treatment	Off treatment
≤ 5 weeks before death	0.39	0.35
> 5 and ≤ 15 weeks before death	0.61	0.43
> 15 and ≤ 30 weeks before death	0.71	0.58
> 30 weeks before death	0.77	0.68

Source: CS, Table 60

The ERG notes that, within the company model, utility scores for all patients were not adjusted over time using an annual utility decrement (i.e., no age-related utility estimates were used in the model), nor is it clear if the company calculated utility values using the UK valuation set.

Impact of adverse events on health-related quality of life

The company took into account the impact of AEs on HRQoL by including a HRQoL decrement for all Grade 3 to Grade 5 AEs that occurred in ≥2% of patients in either treatment arm of the OAK trial. The disutility value per episode for each of the AEs listed in the model (as shown in Table 35) was sourced from studies identified in a systematic review carried out by the company to identify HRQoL evidence describing patients with metastatic NSCLC (CS, Section 5.4.2). Disutilities are applied to each treatment arm whilst patients are still receiving treatment.

Table 35 Adverse event disutilities

Adverse event	Disutility	Source
Anaemia	-0.07346	Nafees 2008[66]
Fatigue	-0.07346	
Febrile neutropenia	-0.09002	

Neutropenia	-0.08973	
Leukopenia	-0.08973	Assumed equal to neutropenia Ramucirumab for previously treated locally advanced or metastatic non-small-cell lung cancer [TA403][67] Lung cancer (non-small-cell, non-squamous, metastatic, after treatment) - nivolumab [ID900][42]
Neutropenic sepsis	-0.09002	Assumed equivalent to febrile neutropenia
Neutrophil count decreased	0	Assumption Ramucirumab for previously treated locally advanced or metastatic non-small-cell lung cancer [TA403][67] Lung cancer (non-small-cell, non-squamous, metastatic, after treatment) - nivolumab [ID900][42]
Pneumonia	-0.008	Marti et al (2013)[68]
Respiratory tract infection	-0.096	Assumption adapted from Hunter 2015[69]
White blood cell count decreased	-0.05	Nintedanib for previously treated locally advanced, metastatic, or locally recurrent non-small-cell lung cancer [TA347][4]

Source: CS, Table 62

5.4.8 Resources and costs

Drug costs

Atezolizumab is administered at a fixed dose of 1200mg over 60 minutes for the first IV infusion and, if well tolerated, as a 30-minute IV infusion every 3 weeks. The expected list price of a 20ml vial (dose per vial is 1200mg) is £3,807.69. The company base case incorporates a PAS discount of [redacted], which reduces the cost per administration to [redacted]. The PAS application is currently under review by the Department of Health.

Drug costs for docetaxel were taken from the electronic Medicines Information Tool (eMIT).[70] Drug costs for nintedanib+docetaxel were taken from the British National Formulary (BNF)[71] and eMIT[70] respectively. A PAS for nintedanib does exist; however, the company is unaware of the value of this PAS price.

The drug acquisition cost and drug cost per treatment cycle used in the company model are provided in Table 36.

Table 36 Drug acquisition cost and drug cost per treatment cycle

Drug	Vial/pack concentration	Vial/pack volume	Dose per vial/pack	Cost per vial/pack	Total drug cost per cycle

Atezolizumab *	1200mg/ml	20 ml	1200 mg	£3,807.69	£3,807.69
Docetaxel	20 mg/ml	7 ml	140 mg	£17.77	£34.39 (75mg/m^2*BSA =137.75mg)
	20 mg/ml	1 ml	20 mg	£4.92	
Nintedanib*	100 mg	120	12000	£2151.10	£1,434.07 (200mg twice daily on day 2-21 treatment cycle)
	150 mg	60	9000	£2151.10	

*List price
Source: CS, Table 64 and Table 65

Administration costs

The costs of administering the intervention and comparator drugs are shown in Table 37.

Table 37 Drug administration costs

Drug	Type of administration		NHS Reference Cost code	Cost per administration	Source
Atezolizumab	Deliver simple parenteral CTX at first attendance	Outpatient Setting	SB12Z (outpatient)	£198.94	NHS Reference Costs 2015-16, Department of Health[72]
Docetaxel	Deliver simple parenteral CTX at first attendance	Outpatient setting	SB12Z (outpatient)	£198.94	
Nintedanib (pre-docetaxel discontinuation) – base case	Deliver simple parenteral CTX at first attendance	Outpatient setting	SB12Z (outpatient)	£198.94	
	12 minutes pharmacist time every 30 days	Hospital pharmacist (band 6); radiographer cost/hour		£46 per hour= £9.20 per administration	PSSRU 2016[73]
Nintedanib (post-docetaxel discontinuation) – base case	12 minutes pharmacist time every 30 days	Hospital pharmacist (band 6); radiographer cost/hour		£46 per hour= £9.20 per administration	
Nintedanib (pre-docetaxel discontinuation) – scenario analysis	Deliver simple parenteral CTX at first attendance	Outpatient setting	SB12Z (outpatient)	£198.94	NHS Reference Costs 2015-16, Department of Health[72]
	Deliver exclusively oral CTX	Outpatient setting	SB11Z	£183.50	
Nintedanib (post-docetaxel discontinuation) – scenario	Deliver exclusively oral CTX	Outpatient setting	SB11Z	£183.50	

CTX=chemotherapy
Source: CS, Table 70

Monitoring and disease management costs

The costs of patient monitoring and disease management were applied to ‘on treatment’ and ‘off treatment’ health states. The company states (CS, Section 5.5.2.3) that the types of resource and frequency of use are derived from previous technology appraisals validated by UK clinicians. Full details of the monitoring costs, ‘on treatment’ health state resource use, ‘off treatment health state resource use’, unit costs for ‘on treatment’ and ‘off treatment’ health

states and terminal care/end of life resource use are reported in detail in the CS (Table 71 to 76).

In summary, the total cost per week for the 'on treatment' health state was £128.25, whilst the total cost per week for the 'off treatment' health state was £120.12. A one-off terminal care/end of life cost was applied to patients in the 'Death' state and this cost was assumed to be equal for all treatments. The total cost of end of life care used in the model was £3,679.37.

Cost of adverse events

The company model includes all Grade ≥3 AEs experienced by ≥2% of patients in either arm of the OAK trial, based on data from the first 850 randomised patients who received any dose of the study drug at the time of the primary analysis (n=823). In addition, the company included a Grade 5 AE, despite the low incidence of the event. Also, the company included neutropenic sepsis as clinical advice suggested that this was an appropriate approach to take given that febrile neutropenia and neutropenic sepsis are terms that are often used interchangeably.

Based on the list of AEs compiled from the OAK trial, the corresponding rates for nintedanib+docetaxel were sourced directly from the LUME-Lung 1 trial.

The weekly rate of occurrence for each AE is implemented in the model through the overall probability of any patient experiencing the event in any given cycle. This is calculated by using 'number of AE occurrence' divided by the total time (weeks) at risk, which is the sum of time on treatment for each patient in the trial. The probability of any patient experiencing the event is then multiplied by the average management costs of the AE to obtain an AE cost per patient per week. The AE costs were applied to each treatment arm whilst patients were still receiving treatment.

The costs of treating AEs are per episode. Where possible, NHS Reference Costs (2015/16)[72] were used to cost AEs. Where there were gaps in the data, costs were sourced from prior NICE submissions in NSCLC and inflated to the appropriate costing year (Table 38). Full details of this costing exercise are presented in the CS (Table 79). UK clinicians validated the assumptions around the costs of treating each AE.

Table 38 Adverse event costs

Adverse event	Unit cost used in the company model	Source	Range of AE costs used in previous appraisals

Anaemia	£1,313.09	HRG 2015/16 (SA04H)[74]	£978 to £1,313.09
Fatigue	£3,802.59	Lung cancer (non-small-cell, squamous, metastatic)-nivolumab (after chemotherapy) [ID811][41]	£2,317.20 to £3,015.13
Febrile neutropenia	£5,612.78		£2,339 to £7,331.78
Neutropenic sepsis	£5,612.78		£2,339 to £7,331.78
Leukopenia	£362.66		£354.72 to £362.66
Neutropenia	£362.66		£179.83 to ££560.08
Neutrophil count decreased	0	Lung cancer (non-small-cell, non-squamous, metastatic, after treatment)-nivolumab [ID900][42]	£0 to £179.83
Pneumonia	£2,783.99	HRG 2015/16 (DZ11T)[74]	£1,822.85 to £2,783.99
Respiratory tract infection	£3,515.13	HRG 2015/16 (DZ27P)[74]	£3,734.17 to £3,515.13
White blood cell count decreased	£432.47	Nivolumab (ID900,[42] ID811[41])	£423 to £560.08

AE=adverse event; HRG=healthcare resource group
Source: CS, Table 79

5.4.9 Cost effectiveness results (based on list price of atezolizumab)

Total costs, LYs gained, QALYs gained and the incremental cost effectiveness ratio (ICER) per QALY gained for the cost effectiveness comparison of treatment with atezolizumab versus docetaxel and versus nintedanib+docetaxel are shown in Table 39 and Table 40 respectively.

Treatment with atezolizumab generates 0.75 additional QALYs versus docetaxel at an additional cost of £53,970. The company base case ICER for the comparison of treatment with atezolizumab versus docetaxel is £72,356.07 per QALY gained.

Table 39 Base case results (atezolizumab versus docetaxel, list price)

Technologies	Total			Incremental			ICER per QALY gained
	Costs	LYs	QALYs	Costs	LYs	QALYs	
Atezolizumab	£73,911	2.20	1.47				
Docetaxel	£19,941	1.19	0.73	£53,970	0.10	0.75	£72,356.07

LYs=life years; QALYs=quality adjusted life years; ICER=incremental cost effectiveness ratio
Source: CS, Table 83

Treatment with atezolizumab generates 0.65 additional QALYs versus nintedanib+docetaxel at an additional cost of £36,209. The company base case ICER for the comparison of treatment with atezolizumab versus nintedanib+docetaxel is £56,076.16 per QALY gained.

Table 40 Base case results (atezolizumab versus nintedanib+docetaxel, list price)

Technologies	Total			Incremental			ICER per QALY gained
	Costs	LYs	QALYs	Costs	LYs	QALYs	
Atezolizumab	£73,911	2.20	1.47				
Nintedanib+docetaxel	£37,702	1.31	0.83	£36,209	0.91	0.65	£56,076.16

LYs=life years; QALYs=quality adjusted life years; ICER=incremental cost effectiveness ratio
Source: CS, Table 83

5.4.10 Deterministic univariate sensitivity analyses

The company carried out a wide range of univariate sensitivity analyses for the base case comparison of treatment with atezolizumab versus docetaxel and versus nintedanib+docetaxel. For each of the comparisons, the same three most influential parameters were apparent: the cure fraction rate, monthly cost of atezolizumab and the discount rate. Results from the analyses involving the ten parameters which, when varied, had the most influence on the company's base case results are displayed in the CS in Tornado diagrams for atezolizumab versus docetaxel and atezolizumab versus nintedanib+docetaxel and reproduced as Figure 9 (CS, Figure 61) and Figure 10 (CS Figure 62), respectively.

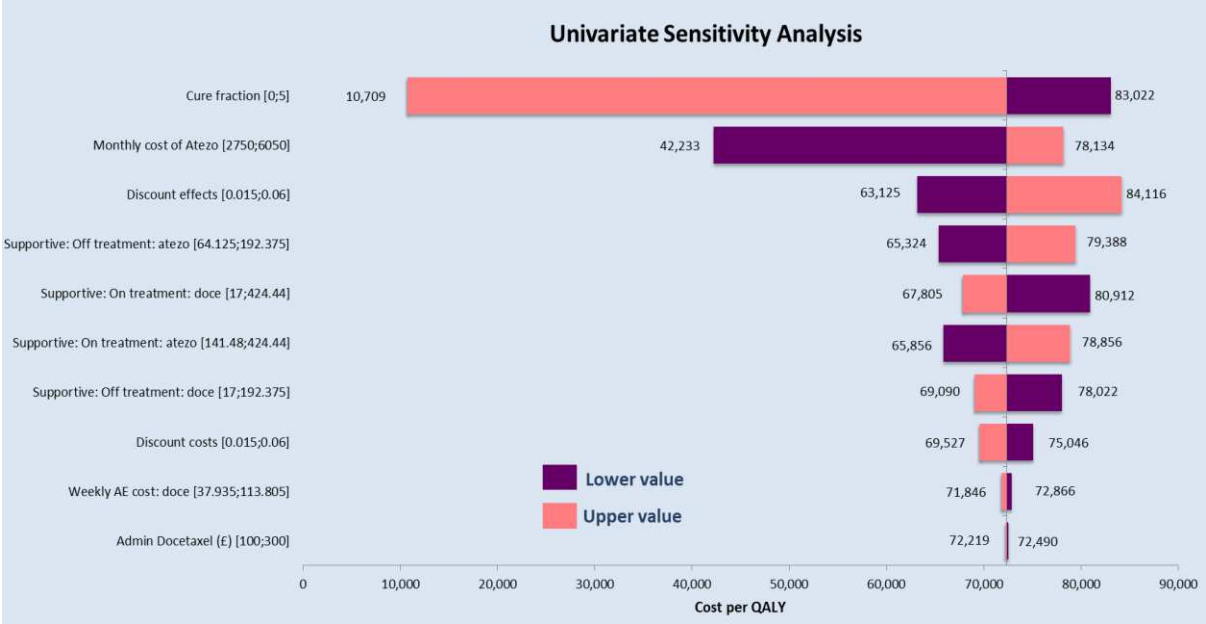


Figure 9 Univariate sensitivity analysis (atezolizumab versus docetaxel, list price)

Source: CS, Figure 61

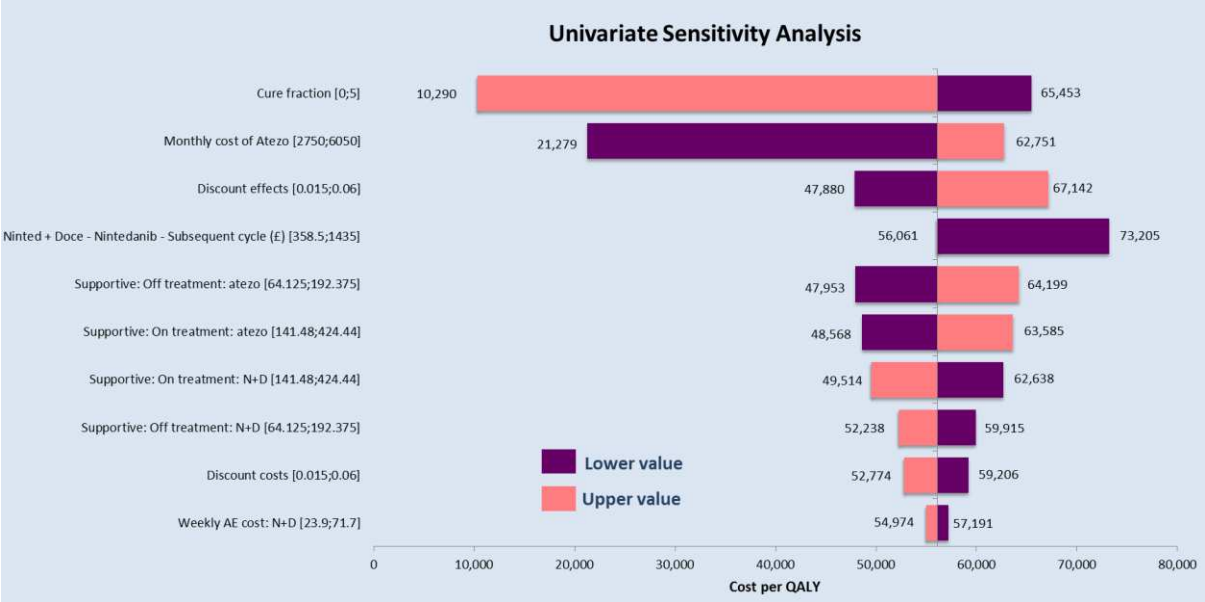


Figure 10 Univariate sensitivity analysis (atezolizumab versus nintedanib+docetaxel, list price)

Source: CS, Figure 62

Probabilistic sensitivity analysis

The company undertook a probabilistic sensitivity analysis (PSA) to assess the uncertainty surrounding the parameter values used in the model. Results from this analysis are displayed in Table 41 and show ICERs per QALY gained that are slightly higher than the ICERs in the deterministic analysis. The PSA involved running the company model 1000 times. The scatterplot of PSA results and the cost effectiveness acceptability curve (CEAC) are presented in Figure 11 and Figure 12 respectively). Examination of the CEAC shows that the chance of atezolizumab being cost effective versus docetaxel (and versus nintedanib+docetaxel) at a threshold of £50,000 per QALY gained is approximately 45% (and 1%).

Table 41 PSA results compared to base-case analysis (list price)

Treatment	Costs		QALYs		ICERs (vs docetaxel)		ICERs (vs nintedanib+docetaxel)	
	Base case	PSA	Base case	PSA	Base case	PSA	Base case	PSA
Docetaxel	£19,941	£20,880	0.73	0.74	-	-	-	-
Nintedanib+ docetaxel	£37,702	£38,676	0.83	0.84	Extendedly dominated	Extendedly dominated	-	-
Atezolizumab	£73,911	£73,033	1.47	1.47	£72,356	£73,934	£56,076	£57,777

QALYs=quality adjusted life years; ICERs=incremental cost effectiveness ratios; PSA=probabilistic sensitivity analysis
Source: CS, Table 93

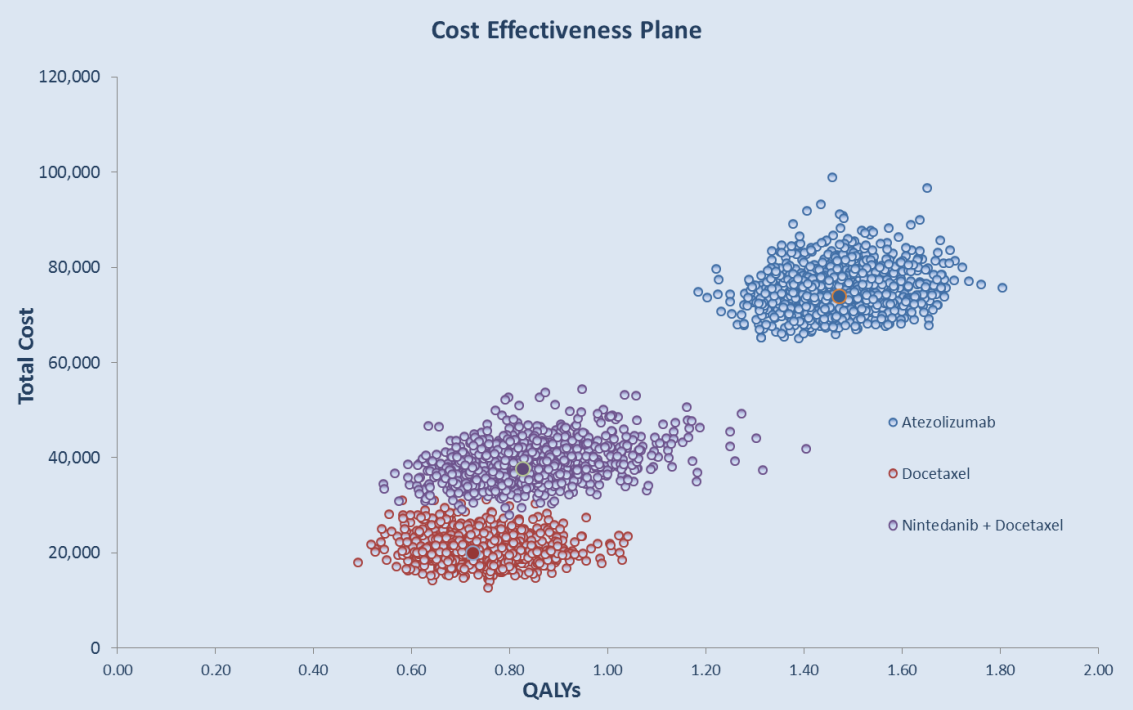


Figure 11 Scatterplot of PSA results for cost effectiveness plane

Source: CS, Figure 59

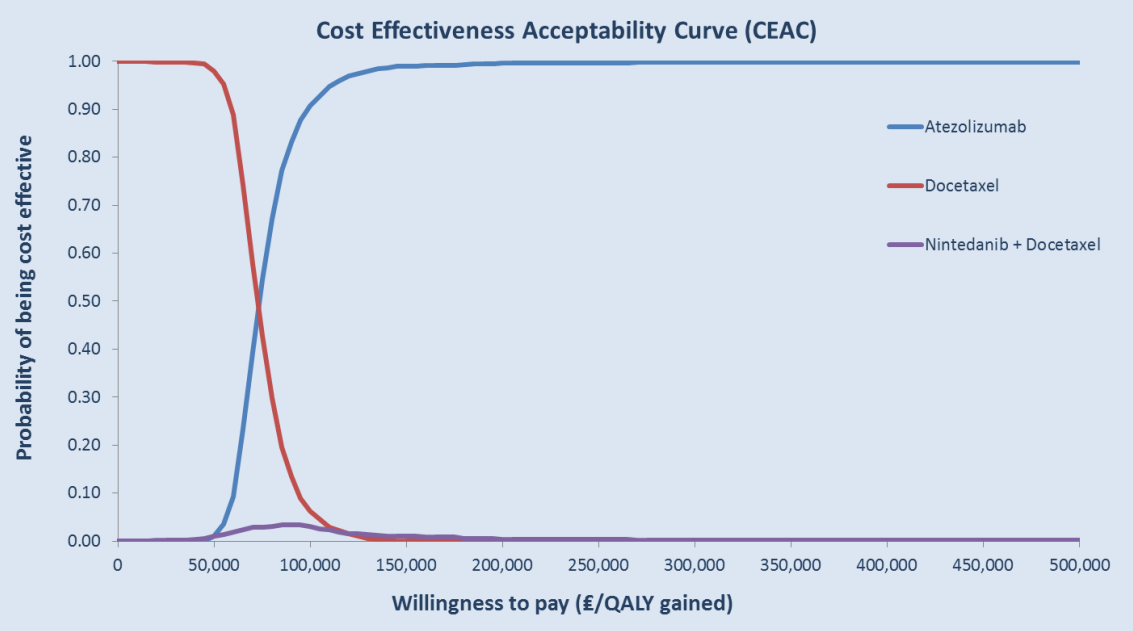


Figure 12 Cost effectiveness acceptability curve

Source: CS, Figure 60

5.4.11 Model validation and face validity check

The company reports that the model approach and inputs were validated by a number of external health economists and clinical experts on two separate occasions to ensure that the model reflected UK clinical practice. In particular, resource use, health state methodologies,

OS projections and extrapolation techniques were checked and verified. In addition, external consultants carried out internal quality control and validation of the model.

5.5 Detailed critique of company's economic model

The company provided a model built in MS Excel. The ERG's assessment of the structure of the company model and the data used to populate it are provided in Section 5.5.1 to 5.5.4 of this ERG report. The ERG considers that the model is generally well constructed and welcomes the following model design choices made by the company:

- use of utilities based on time to death rather than disease state
- use of data as the basis for estimating the cost of treating patients with atezolizumab.

5.5.1 ERG corrections to company model

Health-related quality of life

The ERG considers that, as patients with metastatic NSCLC face significant HRQoL issues, the utility values used in the company model to represent the HRQoL of patients who are more than 30 weeks to death seems high (0.77). This conclusion is based on the fact that the UK population norm for people aged 63, the age of the population at the start of the model time horizon, is 0.79. As part of the clarification process, the ERG asked the company to provide utility values for all patients in the OAK trial, calculated using the UK valuation set. On examination of these results the ERG was satisfied that the utility values used by the company fairly reflect the values suggested by the EQ-5D questionnaires completed by patients who participated in the trial. However, the ERG notes that the people who completed the questionnaires were trial participants and, therefore, may not be wholly representative of all patients in NHS clinical practice who are eligible for treatment with atezolizumab.

Within the company model, different utility values are applied depending on whether patients are 'on' or 'off' treatment. The ERG is not convinced that the separation of 'on' and 'off' treatment utilities is necessary. However, as the off-treatment utility value for patients who are more than 30 weeks to death is 0.68, the ERG considers that this value likely reflects the actual utility of patients during this phase and, therefore, has not amended this aspect of the company model.

Correction C1: inaccurate application of the discount rate

Due to an algorithmic error, the company has incorrectly applied discounting in the model by starting to discount from week 1, rather than from the start of year 2. The ERG has amended this error with the result that the company's base case ICER for the comparison of atezolizumab versus docetaxel increases by £408 to £72,764 per QALY gained. This

amendment decreases the company's base case ICER for the comparison of atezolizumab versus nintedanib+docetaxel by £117 to £55,959 per QALY gained.

Correction C2: failure to apply an age-related utility decrement

The ERG considers that the company model was misspecified as it does not take into account declining utility with age. Within the company model, a patient with the same time to death at age 63 years (the start of the model time horizon) has the same utility as a patient with the same time to death at age 88 years (the end of the model time horizon). To correct this error, the ERG has incorporated age-related decrements drawn from the publication by Kind[75] (0.02 at age 65 years and 0.07 at age 74 years) to reflect the lower HRQOL that patients experience as they get older.

Applying age-related decrements increases the company's base case ICER for the comparison of atezolizumab versus docetaxel by £2,960 to £75,316 per QALY gained. This amendment also increases the company's base case ICER for the comparison of atezolizumab versus nintedanib+docetaxel by £2,532 to £58,608 per QALY gained.

Correction C3: inappropriate half-cycle correction to modelling of time on treatment

The company has applied a half-cycle correction to their modelling of time on treatment TTD for all treatment arms. As treatment is administered at the start of each cycle, rather than during it, a half-cycle correction was unnecessary for this parameter. This approach also created the implausible situation whereby 4.3% of patients in the atezolizumab arm of the model did not receive their first cycle of atezolizumab i.e., these patients stopped treatment before they even started it. The ERG has, therefore, removed the half-cycle correction applied to TTD data. This amendment increases the company's base case ICER for the comparison of atezolizumab versus docetaxel by £1,736 to £74,092 per QALY gained. For the comparison of atezolizumab versus nintedanib+docetaxel, this amendment increases the ICER by £1,873 to £57,949 per QALY gained.

The ERG notes that the data used in the model to represent TTD for patients receiving nintedanib+docetaxel have been generated from an adjustment of the PFS data from the OAK trial. The PFS estimate for nintedanib+docetaxel is, therefore, not drawn from an analysis of direct trial data and so the ERG considers that the costs of treatment with nintedanib+docetaxel within the model have a level of uncertainty that means any ICERs based upon these costs should be considered as uncertain and indicative only.

Corrected company base case

The combined effect of introducing an age-related decrement to patients' HRQoL and removing the half-cycle correction applied to TTD data increases the size of company's base case ICER for the comparison of atezolizumab versus docetaxel by £5,213 to £77,569 per QALY gained. The effect of these changes on the ICER for the comparison of treatment with atezolizumab versus nintedanib+docetaxel is to increases it by £4,290 to £60,366 per QALY gained.

5.5.2 Company's approach to modelling overall survival: atezolizumab

The company has used a mixed cure-rate model to reflect survival for patients treated with atezolizumab. The ERG considers that, within the CS, the company has failed to justify the need for the application of a 'cure rate'. Even if a case had been made, the choice of cure rate used (2%) appears to be arbitrary as it is not supported by the evidence presented in the CS. The ERG considers that cost effectiveness results generated by the company's mixed cure-rate model are an inappropriate basis for decision-making.

Company justification for application of a cure rate

The company states (CS, p161) that the rate of death of patients with cancer declines over time if patients are treated with immunotherapies, and that:

> *"Long term evidence is not available from clinical trials, and with relatively immature data from the OAK study - use of traditional parametric survival analysis which relies on the observed data for atezolizumab will fail to account for this change in mortality rate and 'flattening' of the tail of the survival curve."* (Source: CS, p160)

The company concludes that the way to account for this is to use a mixed cure-rate model. The company references TA414[76] (Cobimetinib in combination with vemurafenib for treating unresectable or metastatic BRAF V600 mutation-positive melanoma) as a previous example of the need to use a mixed cure-rate model in a cancer population receiving immunotherapy treatments.

In TA414,[76] the company described registry data from patients with advanced melanoma and explained that these data indicated that there was a subpopulation of patients who, having survived for 5 years, had a noticeably lower mortality rate than the population of patients who did not survive to 5 years. The ERG notes that the use of a mixed cure-rate model in the TA414[76] appraisal was not due to evidence of any prolonged treatment effect, or because of a lowering of the long-term mortality rate due to the effect of an immunotherapy. Rather, the

mixed cure-rate model was used because of a peculiarity in the survival trajectory of a population with malignant melanoma that could not be captured by the available trial data.

The ERG is unaware of any NSCLC registry data that suggest that a subgroup of patients exists where differential mortality rates occur once a specific survival point has been reached. The ERG considers that the company's reference to TA414[76] as a justification for applying a mixed cure-rate model to represent the survival trajectory of patients with NSCLC receiving atezolizumab is inappropriate.

The company states (CS, pp159-160) that the mixed cure-rate model is required because treatment with atezolizumab (a drug that is similar to other immunotherapies) may have a sustained effect for a subgroup of patients with Stage IV NSCLC. The company suggests that mortality rates for this subgroup are equal to population mortality rates. The ERG considers that this statement needs to be evidenced, and not simply assumed. In the absence of an evidence base, the ERG suggests that the application of a potential cure rate should be applied within the framework of a scenario analysis rather than used in the base case analysis.

In addition, the ERG notes that, for the comparison of treatment with atezolizumab versus docetaxel or nintedanib+docetaxel, application of a cure rate effectively generates a differential hazard at all time-points. The ERG considers that if there is evidence for such a difference it could be modelled by appropriately chosen distributions, based upon available trial data, and without recourse to a hypothetical cure rate.

Cure rate, OAK trial data and the National Lung Cancer Audit registry data

During the TA414[76] appraisal, the company identified a cure rate by adjusting registry data based on the characteristics of patients in the trial that provided evidence for the company's cost effectiveness analyses. The resulting extrapolation using the cure rate was then compared to survival data from a second trial, in which patients also received the study drug, to ensure it was appropriate. In the current appraisal, the company has justified the cure rate used in the model by comparing the mixed cure-rate model OS prediction against OS data from the OAK and POPLAR trials. However, when the mixed cure-rate OS model is compared with OAK trial data at 24 months, the chosen cure rate (2%) produces an overestimate of survival for patients treated with atezolizumab by 2.8% whilst underestimating OS for patients receiving docetaxel by 1.6% (recognising that the docetaxel OS curve is dependent on the atezolizumab curve although no cure rate is assumed for docetaxel). The mixed cure-rate model applied by the company, therefore, generates survival gains for patients treated with atezolizumab and docetaxel, over the first 24 months of the model time horizon, that are not supported by data from the OAK trials.

Whilst the company has presented NLCA registry data from 2006-2010 to support their choice of cure rate for patients receiving atezolizumab (CS, Table 59), the company has not undertaken any adjusted statistical analysis of the NLCA registry data. Without rebasing the data to take into account the time since diagnosis, number of prior treatments, and progression status, the company's use of the registry data as a justification of the need for, or value of, a cure rate is spurious.

The ERG, therefore, considers that the company's choice of cure rate is arbitrary; it is unsupported by the company's own trial data and cannot be verified with registry data.

Clinical opinion on 5-year survival for patients with NSCLC

To assess the clinical plausibility of any projection, the company explored potential 5-year survival rates for patients treated with atezolizumab by eliciting opinions from clinicians. In the CS (p161), the company states that unanimous clinical opinion is that a value of 10% for the 5-year OS rate of patients receiving immunotherapy "…would not be implausible". The company did not provide any context to explain how this number was elicited from clinicians. The ERG considers that the phrase "…would not be implausible" should not be interpreted as 'likely'.

To support the 5-year OS rate of 10%, the company then referenced the final appraisal determination for TA428[15] (Pembrolizumab for treating PD-L1-positive NSCLC after chemotherapy) and states that, "…under the Committee's preferred assumptions, the resulting 5-year OS estimate was 10.4%" (CS, p161). The company has acknowledged, in response to an ERG clarification question, that the quoted value was inaccurate and should have been 9.6%. However, it is not just the number that is inaccurate; it is also the statement that this value was the **Committee's** preferred assumption. During TA428,[15] this 9.6% survival rate was generated using the **company's** preferred assumptions. The Committee and the ERG for that appraisal were particularly concerned about the company's assumption that treatment with pembrolizumab would generate a lifetime treatment effect; this **company** assumption generated the 9.6% 5-year OS rate. The Committee considered a more clinically plausible duration of treatment effect would be 3 years after treatment stopped, at which point the Committee considered that the mortality hazard for patients treated with pembrolizumab would be equal to the mortality hazard for patients treated with docetaxel.

The company assumed that 10% would be a plausible 5-year OS rate for patients treated with atezolizumab. However, the company's mixed cure-rate log-logistic model used in the base case analysis leads to an estimated 12.6% of patients receiving atezolizumab being alive at 5 years. The company's mixed cure-rate log logistic model, therefore, produces more optimistic

5-year survival estimates than the 'plausible' (but not 'most likely') estimate provided as clinical advice to the company, and is higher than an estimate thought 'optimistic' by a previous NICE Appraisal Committee who had evaluated immunotherapy in a population with NSCLC.

Implausibility of long-term projection

Due to the resultant long tail that is a characteristic of any log-logistic distribution, coupled with the lifetime duration of treatment effect, results from the company model suggest that 5.6% of patients treated with atezolizumab will be alive at 10 years, and 1.4% will be alive at 25 years. The company projection generates a 5-year mortality rate of 36.9% between years 20 and 25 of the model, when patients are aged between 83 and 88 years. However, the 5-year mortality rate for all people aged between 83 and 88 years, based on UK life tables[77] provided within the company model is 39.5%. The ERG considers the company projection is implausible as it leads to a situation where treatment with atezolizumab is not just keeping people with advanced NSCLC alive for 20 years and longer after progressing on their first treatment, it is also preventing them from dying from other, non-NSCLC, causes. Therefore, the ERG considers that i) the ICERs that are generated by this approach should not be used to inform decision-making and ii) that the log-logistic distribution is a poor choice of distribution for modelling the OS trial data.

5.5.3 ERG preferred approach to modelling OS: atezolizumab versus docetaxel

Kaplan-Meier data and extrapolation of overall survival

The ERG's preferred method is to model OS for both atezolizumab and docetaxel by using K-M data from the OAK trial for as long as possible, then append curves to project OS for the remainder of the model time horizon.

The minimum period for which follow-up OS data from the OAK trial are available for all patients is 19 months (83 weeks). After this point the number of patients at risk starts to decrease rapidly through censoring. Whilst deciding the point at which K-M data are no longer robust due to censoring is a subjective judgement, the ERG considers that, given that there is limited censoring up to 83 weeks, the data up to this point may be considered to be robust.

Inspection of the OS K-M data from the OAK trial suggests that there are three clear phases in the data between trial start and week 83, with different mortality ratios between atezolizumab and docetaxel for each phase. For the first 11 weeks, the hazard rates for patients in both arms of the trial are indistinguishable (Figure 13) implying a HR of one. Between weeks 11 and 56, if the population is rebased, a clear separation of survival between patients receiving atezolizumab and patients receiving docetaxel can be seen (

Figure 14) implying a HR for atezolizumab compared to docetaxel of less than one. If the population is again rebased at weeks 56, then between weeks 56 and 83 (the point after which follow-up data cease to be available for the full trial population) the picture is unclear (

Figure 15).

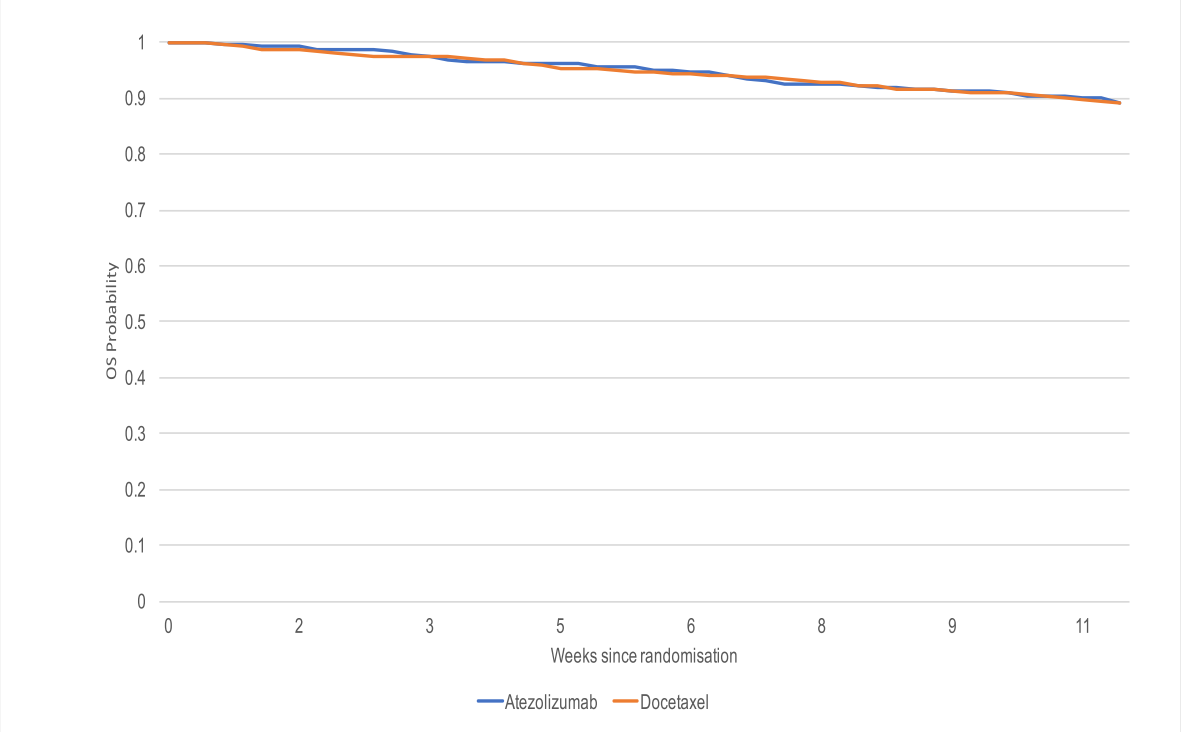


Figure 13 OS K-M data from the OAK trial for the first 11 weeks

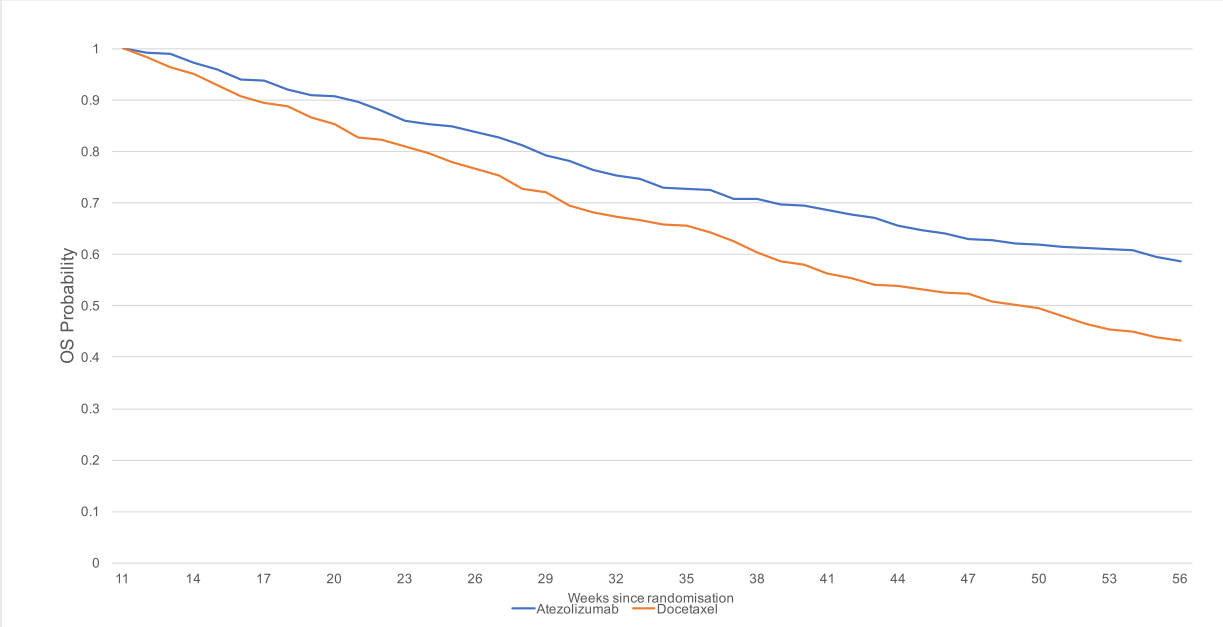


Figure 14 OS K-M data from the OAK trial, weeks 11 to 56 (rebased at week 11)

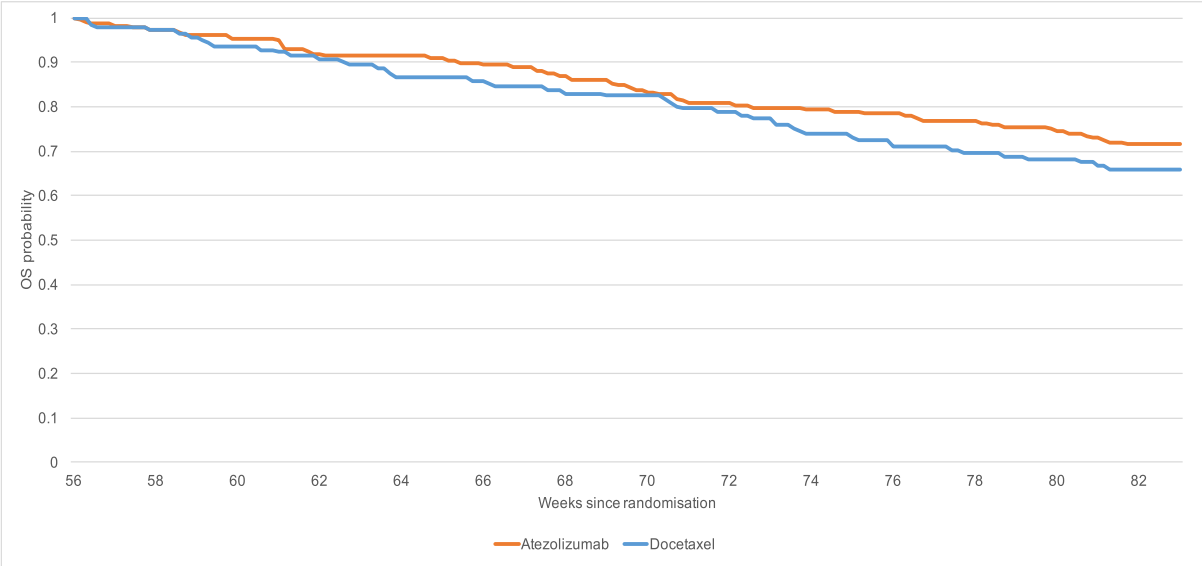


Figure 15 OS K-M data from the OAK trial, weeks 56 to 83 (rebased at week 56)

Between weeks 56 and 83 (

Figure 15) there may be some separation between the two arms of the OAK trial, but the K-M curves touch twice and visual inspection suggests that the HR between atezolizumab and docetaxel may have returned to one. Whilst the ERG is not convinced there is compelling evidence to support applying a differential hazard rate after week 56, the ERG estimated

distributions for extrapolation from week 56 onwards based on K-M data from weeks 56 to 83 separately for atezolizumab and docetaxel (see Figure 16 and Figure 17 respectively).

Inspection of the cumulative hazard plots for atezolizumab and docetaxel suggests that, between weeks 56 and 83, the cumulative hazards are linear and exponential distributions could fit both data sets over this period and could be used to extrapolate OS for both atezolizumab and docetaxel past week 56.

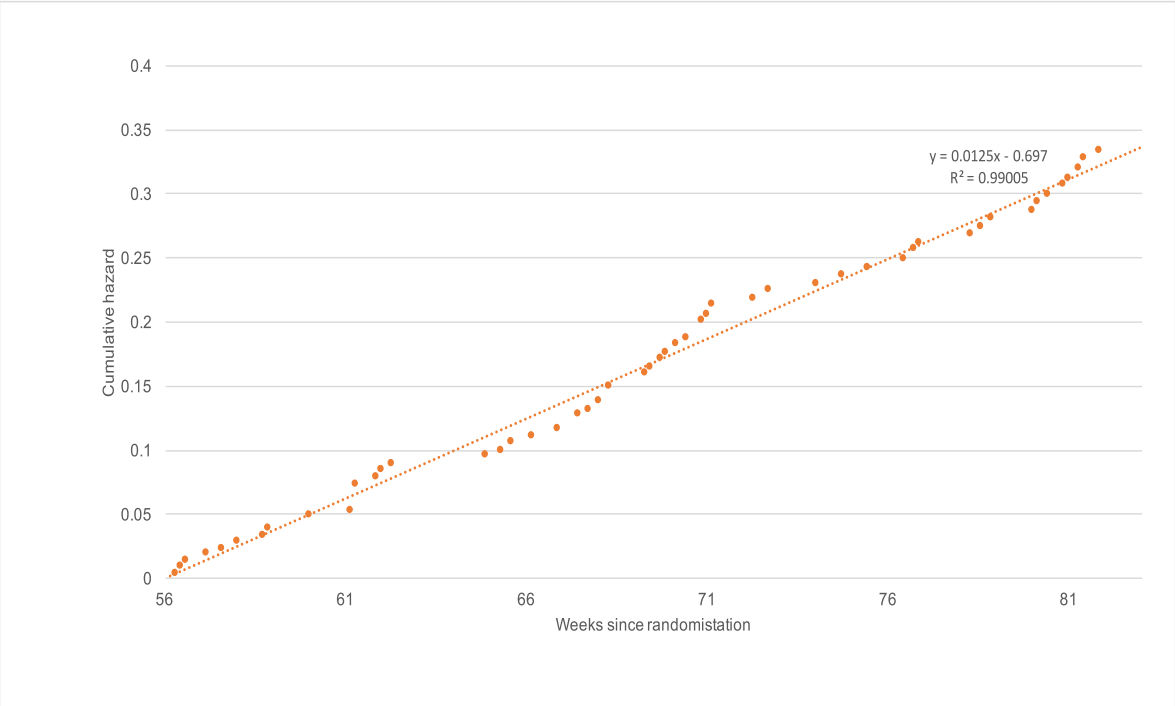


Figure 16 OS cumulative hazard plot for atezolizumab between weeks 56 and 83

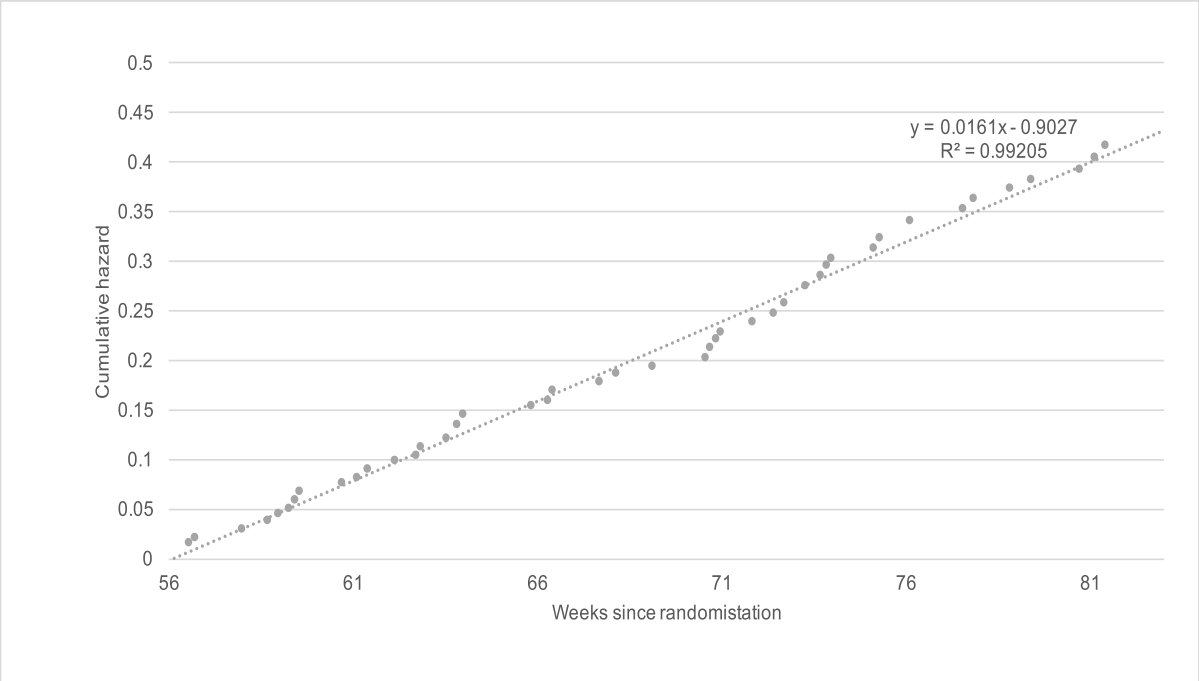


Figure 17 OS cumulative hazard plot for docetaxel between weeks 56 and 83

Duration of treatment effect

The company has assumed a lifetime duration of treatment effect for atezolizumab, this results in a lower mortality rate for patients who received atezolizumab versus docetaxel or nintedanib+docetaxel for the duration of the model. The NICE Appraisal Committee raised concerns during TA428[15] (Pembrolizumab for treating PD-L1positive NSCLC after chemotherapy) relating to the duration of treatment effect (after treatment had ended) associated with receiving an immunotherapy. Consequently, the ERG looked to cap the duration of treatment effect of atezolizumab at 3 years in line with the (TA428[15]) Committee's view on what could be considered a reasonable duration of treatment effect.

Whilst the company model allows the duration of treatment effect to be fixed, the approach that is used to stop the treatment effect in the model is simplistic. If the duration of treatment effect is set to be 'x' months in the model, then the hazard rate for atezolizumab is set to be equal to docetaxel at 'x' months after the **start** of the model. Any patients that stop atezolizumab in month 't' will have a duration of treatment effect for atezolizumab of x-t. This means the duration of treatment effect of atezolizumab in the model varies for patients and is not fixed and underestimates the true duration of treatment effect for atezolizumab of 'x' months if this is believed to exist in reality.

For example, if duration of treatment effect for atezolizumab is actually 3 years, then, in the model, setting the duration of treatment effect to 3 years would mean the duration of treatment effect of atezolizumab would be 2.5 years for a patient who stopped treatment after 6 months, but zero for a patient who is still on treatment at 3 years.

The method used in the model for dealing with duration of treatment effect for atezolizumab underestimates OS for atezolizumab if a treatment effect of 3 years actually exists and 36 months is entered into the model as the duration of treatment effect. Without restructuring the model, which is beyond the remit of the ERG, it is not possible to implement a more sophisticated approach to modelling the duration of treatment effect.

Taking the company model limitations into account but still attempting to implement a 3-year duration of treatment effect, the ERG set the *company model* duration of treatment effect to 5 years. As 8.5% of patients are predicted by the company's TTD extrapolation to be receiving atezolizumab at 2 years, this means that for those patients, if they are alive at 5 years, the duration of treatment effect will still be less than 3 years even though the duration of treatment effect is set to 5 years in the company model. However, patients who stopped treatment before 2 years and are still alive at 5 years will have a greater than 3 year treatment effect.

On balance, whilst there is no accurate way within the company model to set the duration of treatment effect for atezolizumab to 3 years, the ERG, therefore, considers that setting the *company model* duration of treatment effect to 5 years rather than 3 years probably produces more accurate ICERs per QALY gained if the real duration of treatment effect for atezolizumab is actually 3 years.

ERG remodelled OS for atezolizumab and docetaxel

The ERG's preferred OS curves for atezolizumab and docetaxel, taking into account the use of K-M data, exponential extrapolations and the ERG's preferred duration of treatment effect for atezolizumab are shown in Figure 18, with survival rates at different time points shown in Table 42.

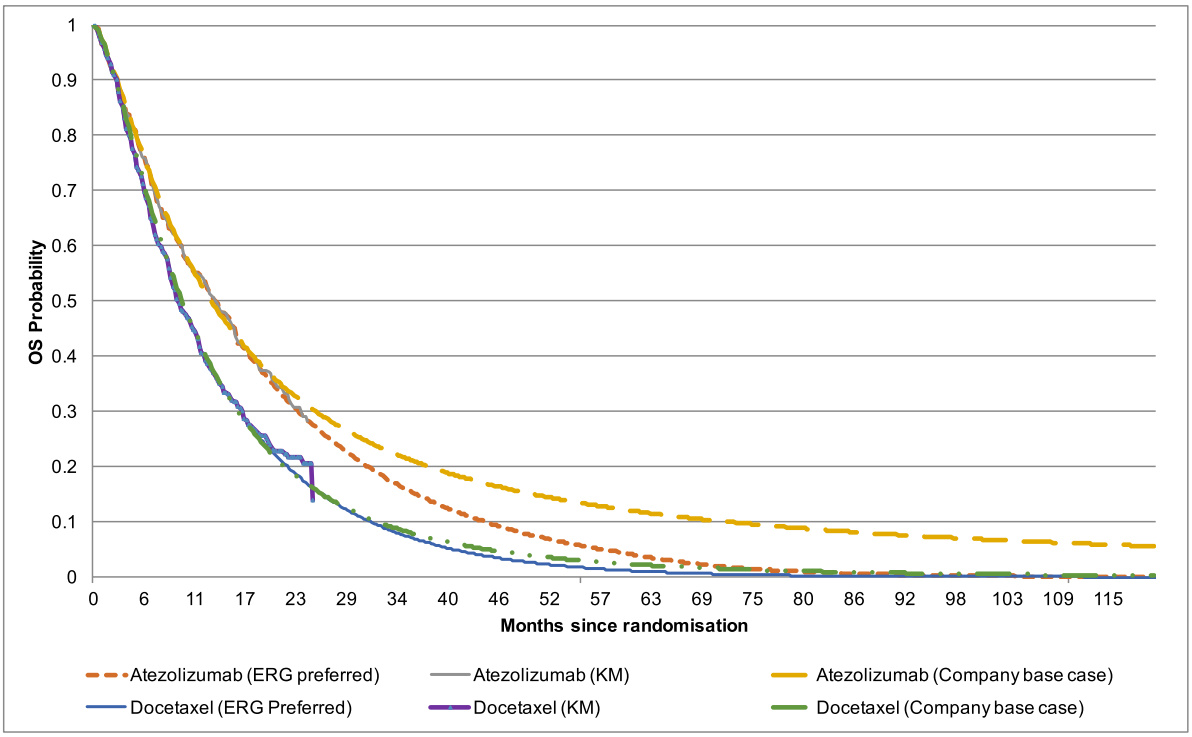


Figure 18 ERG preferred OS distributions compared to company modelled OS and K-M data

Table 42 Estimates, generated using different survival data or projections, of proportions of patients alive at different time points since randomisation

OS curve	Time since randomisation					
	1 year	2 years	5 years	10 years	20 years	25 years
Atezolizumab (K-M)	54.7%	28.1%	-	-	-	-
Atezolizumab (company base case)	53.6%	31.1%	12.2%	5.5%	2.2%	1.4%
Atezolizumab (ERG preferred)	54.7%	28.5%	4.4%	0.1%	0.0%	0.0%
Docetaxel (K-M)	41.7%	20.6%	-	-	-	-
Docetaxel (company base case)	42.5%	16.8%	2.4%	0.0%	0.0%	0.0%
Docetaxel (ERG preferred)	41.7%	17.0%	1.2%	0.0%	0.0%	0.0%

ERG=Evidence Review Group; K-M=Kaplan-Meier

Applying the ERG's preferred OS distribution and preferred duration of treatment effect to the ERG's corrected company base case reduces the incremental QALY gain from atezolizumab compared to docetaxel from 0.746 to 0.302. This increases the ERG's corrected company base case ICER by £92,928 to £170,497 per QALY gained.

5.5.4 ERG preferred approach to modelling OS: atezolizumab versus nintedanib+docetaxel

The ERG asked the company to indirectly compare atezolizumab versus nintedanib+docetaxel in the adenocarcinoma population only. However, these results were

not provided by the company. Instead, the company provided the results of atezolizumab (total population, OAK and POPLAR trials) versus nintedanib+docetaxel (adenocarcinoma population, LUME-Lung 1 trial). It may be that if only the adenocarcinoma populations were compared, then a statistically significant difference in OS would have emerged. In the absence of this analysis, the ERG considers that the company has concluded that the effectiveness of atezolizumab on OS is independent of whether a patient does or does not have adenocarcinoma. As such, the OS for patients with adenocarcinoma and treated with nintedanib+docetaxel can be compared fairly to all patients treated with atezolizumab. Consequently, as this comparison shows there is no statistically significant difference in OS, the ERG concludes that there is no justification for modelling a different OS curves for atezolizumab and nintedanib+docetaxel.

Setting OS equal for atezolizumab and nintedanib+docetaxel results in the QALY gain for atezolizumab falling to 0.027 with the ERG corrected company base case ICER for atezolizumab versus nintedanib+docetaxel increasing to £1,170,260 per QALY gained, assuming a lifetime duration of treatment effect for both treatments. If the duration of treatment effect is limited to approximately 3 years using the method described previously, the ICER would increase to £1,170,793 per QALY gained for atezolizumab versus nintedanib+docetaxel.

5.6 Conclusions of the cost effectiveness section

The ERG considers that there are three errors in the company model. These relate to discounting, age-related disutility and the half-cycle correction applied to TTD data. The ERG considers that these errors must be corrected to allow accurate estimates of the cost effectiveness of atezolizumab versus docetaxel, or atezolizumab versus nintedanib+docetaxel, under the company base case assumptions. Once these errors have been corrected, the ERG's major concerns relate to the assumptions made by the company in relation to modelling OS for patients receiving all treatments.

Treatment with atezolizumab was modelled by the company using a mixed cure-rate model that was not fully justified as being necessary, was arbitrarily specified and ultimately produced implausible projections of the mortality hazard rate associated with treatment with atezolizumab. The ERG considers that the use of a mixed cure-rate model was unnecessary and that the OS of patients receiving atezolizumab could have been modelled simply by using K-M data from the OAK trial for as long as the data are robust and then extrapolating the trial data by appending an exponential distribution. Similarly, modelling OS for patients receiving docetaxel could be carried out using the same approach.

The ERG notes that inspection of the K-M data from week 56 of the OAK trial does not necessarily justify the application of a different mortality hazard rate from this point onwards for atezolizumab and docetaxel. However, given the OAK trial was not powered to identify a difference in OS from week 56, the ERG applied different exponential distributions from week 56 for the two therapies. Nevertheless, the cost effectiveness results generated from these ERG models of OS can be interpreted as optimistically favouring treatment with atezolizumab as the ERG does not consider it implausible that the mortality hazard from week 56 may be the same for both atezolizumab and docetaxel.

The company model allowed the duration of treatment effect for atezolizumab to be fixed, albeit in a simplistic way. The ERG fixed the duration of treatment effect such that is approximately 3 years in line with the duration thought plausible for immunotherapy by the NICE Appraisal Committee assessing pembrolizumab as second-line treatment for patients with advanced or metastatic PD-L1 positive NSCLC.

The ERG does not consider that, from analysis of the clinical trial data that are currently available, there is statistically significant evidence to justify a differential OS for atezolizumab compared to nintedanib+docetaxel for the adenocarcinoma population for which nintedanib is licensed and an OS gain should not be included in the company model.

6 SUMMARY OF ADDITIONAL WORK UNDERTAKEN BY THE ERG

Details of the ERG's revisions to the company model may be found in the appendices (Section 10.8). A summary of the effects of the ERG's model amendments on the company's base case cost effectiveness results for the comparison of atezolizumab versus docetaxel and for the comparison of atezolizumab versus nintedanib+docetaxel are provided in Table 43 and Table 44 respectively.

The ERG considers the first three changes (C1 to C3) to be corrections to the company model. The changes have been implemented as a result of identifying errors or omissions and, therefore, reflect inaccuracies rather than differences of opinion between the ERG and the company. The ERG considers that making these corrections allows the generation of ICERs per QALY gained that fairly reflect the company base case assumptions. Applying the three corrections increases the size of company's base case ICER for the comparison of atezolizumab versus docetaxel by £5,213 to £77,569 per QALY gained. For the comparison of treatment with atezolizumab versus nintedanib+docetaxel, applying the three corrections increases the company's ICER by £4,290 to £60,366 per QALY gained.

The major amendments made by the ERG to the corrected company base case model relate to modelling OS. The ERG considers the company's approach to modelling OS for patients receiving atezolizumab to be insufficiently justified. The ERG considers that, not only was the approach used by the company not supported by the available OAK trial data, but that it also led to a risk of death, in the long-term, that was higher than the risk for the general population. The ERG considers that OS for atezolizumab can be more accurately and simply modelled using the OAK trial data and, once a constant hazard had been observed in the data, appending an exponential function.

In addition, the ERG also adjusted the model so as to limit the duration of treatment effect of atezolizumab to approximately 3 years from the lifetime duration of treatment effect assumed in the company base case.

In terms of modelling the survival of patients treated with docetaxel, the ERG considers that an adequate model could be created without the need for the FP ITC in a similar manner as for atezolizumab. The ERG considers that OS for atezolizumab can be more accurately and simply modelled using the OAK trial data and, once a constant hazard had been observed in the data, appending an exponential function.

Applying the ERG's preferred OS projections for patients receiving atezolizumab and docetaxel, increases the ERG's corrected company base case ICER by £87,741 to £165,310 per QALY gained (R1). While implementing the ERG's preferred projections for the atezolizumab and docetaxel arms **and** setting the treatment duration effect for atezolizumab to approximately 3 years, increases the ERG's corrected company base case ICER for the comparison of atezolizumab versus docetaxel by £92,928 to £170,497 per QALY gained (R2).

There is no statistically significant evidence to support the claim that atezolizumab generates an OS gain compared to nintedanib+docetaxel. Assuming the same OS for patients treated with atezolizumab and nintedanib+docetaxel and a lifetime duration of treatment effect, results in the ERG corrected company base case ICER for atezolizumab versus nintedanib+docetaxel increasing by £1,109,894 to £1,170,260 per QALY gained (R3). Assuming the same OS for patients treated with atezolizumab and nintedanib+docetaxel and an approximate 3 year duration of treatment effect for both treatments results in the ERG corrected company base case ICER for atezolizumab versus nintedanib+docetaxel increasing by £1,110,427 to £1,170,793 per QALY gained (R4).

Table 43 Cost effectiveness results for atezolizumab versus docetaxel with ERG revisions to company base case (list prices)

Model scenario & ERG revisions	Atezolizumab			Docetaxel			Incremental			ICER	ICER
	Cost	QALYs	Life years	Cost	QALYs	Life years	Cost	QALYs	Life years	£/QALY	Change
Company base case	**£73,911**	**1.471**	**2.224**	**£19,941**	**0.725**	**1.188**	**£53,970**	**0.746**	**1.036**	**£72,356**	-
C1) Discounting algorithms	£74,479	1.479	2.236	£20,111	0.732	1.198	£54,367	0.747	1.038	£72,764	**+£408**
C2) Age-related utility decrement	£73,911	1.437	2.224	£19,941	0.720	1.188	£53,970	0.717	1.036	£75,316	**+£2,960**
C3) TTD half-cycle correction	£75,468	1.472	2.224	£20,197	0.726	1.188	£55,271	0.746	1.036	£74,092	**+£1,736**
ERG corrected company base case (C1-C3)	**£76,046**	**1.446**	**2.236**	**£20,369**	**0.728**	**1.198**	**£55,677**	**0.718**	**1.038**	**£77,569**	**+£5,213**
R1) ERG preferred OS for atezolizumab and docetaxel	£71,525	0.998	1.544	£19,951	0.686	1.134	£51,574	0.312	0.409	£165,310	**+£92,954**
R2) ERG preferred OS for atezolizumab and docetaxel, and atezolizumab treatment duration effect set to 3 years	£71,418	0.988	1.527	£19,951	0.686	1.134	£51,467	0.302	0.393	£170,497	**+£98,141**

Costs and QALYs discounted; life years undiscounted
OS=overall survival; QALYs=quality adjusted life years; TTD=time to treatment discontinuation; ICER=incremental cost effectiveness ratio

Table 44 Cost effectiveness results for atezolizumab versus nintedanib+docetaxel with ERG revisions to company base case (list prices)

Model scenario & ERG revisions	Atezolizumab			Nintedanib+docetaxel			Incremental			ICER	ICER
	Cost	QALYs	Life years	Cost	QALYs	Life years	Cost	QALYs	Life years	£/QALY	Change
Company base case	**£73,911**	**1.471**	**2.224**	**£37,702**	**0.826**	**1.315**	**£36,209**	**0.646**	**0.910**	**£56,076**	**-**
C1) Discounting algorithms	£74,479	1.479	2.236	£37,582	0.820	1.306	£36,896	0.659	0.930	£55,959	**-£117**
C2) Age-related utility decrement	£73,911	1.437	2.224	£37,702	0.819	1.315	£36,209	0.618	0.910	£58,608	**+£2,532**
C3) TTD half-cycle correction	£75,468	1.472	2.224	£37,999	0.826	1.315	£37,470	0.647	0.910	£57,949	**+£1,873**
ERG corrected company base case (C1-C3)	**£76,046**	**1.446**	**2.236**	**£37,879**	**0.813**	**1.306**	**£38,168**	**0.632**	**0.930**	**£60,366**	**+£4,290**
R3) ERG preferred OS for atezolizumab and assumed equal for nintedanib+docetaxel	£71,525	0.998	1.544	£39,420	0.970	1,544	£32,105	0.027	0.000	£1,170,260	**+£1,114,185**
R4) ERG preferred OS for atezolizumab and assumed equal for nintedanib+docetaxel, and treatment duration effect for both set to 3 years	£71,418	0.988	1.527	£39,313	0.961	1.527	£32,105	0.027	0.000	£1,170,793	**+£1,114,718**

Costs and QALYs discounted; life years undiscounted
OS=overall survival; QALYs=quality adjusted life years; TTD=time to treatment discontinuation; ICER=incremental cost effectiveness ratio

7 END OF LIFE

The NICE End of Life criteria, and the data presented by the company to show that these have been met, are presented in Table 45.

Table 45 End of life criteria

<table>
<tr><th>NICE End of Life criteria</th><th>Data presented by the company</th></tr>
<tr><td>The treatment is indicated for patients with a short life expectancy, normally less than 24 months</td><td>The company considers this criterion to be met and quotes values from Beckett 2013[78] that show median survival for patients with Stage IIIb and Stage IV NSCLC is 7.5 months and 3.4 months, respectively (CS, Section 3.4)</td></tr>
<tr><td>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</td><td>The company considers this criterion to be met and quotes data (CS, Figure 8) from the OAK trial that show that treatment with atezolizumab is associated with a statistically significant improvement in OS compared with docetaxel in the ITT population (HR 0.73, 95% CI: 0.62 to 0.87).
The company also highlights that results from the OAK trial (CS, Section 4.7) show that median OS in the ITT population is 9.6 months (95% CI: 8.6 to 11.2) in the docetaxel arm and 13.8 months (95% CI: 11.8 to 15.7) in the atezolizumab arm
Company economic model results
Results from the company model (CS, Section 5.7) show that mean OS of patients treated with atezolizumab is >3 months versus all comparators, and median OS results are >3 months versus docetaxel:
<table>
<tr><td></td><td>Mean (months)</td><td>Median (months)</td></tr>
<tr><td>Atezolizumab</td><td>31.1</td><td>13.3</td></tr>
<tr><td>Docetaxel</td><td>14.1</td><td>9.8</td></tr>
<tr><td>Nintedanib+docetaxel</td><td>16.4</td><td>10.6</td></tr>
</table>
</td></tr>
</table>

ITT=intention to treat; HR=hazard ratio; CS=company submission; OS=overall survival
Source: CS, Table 50

Short life expectancy

The ERG agrees with the company that patients with advanced or metastatic NSCLC have a life expectancy of less than 24 months, although the survival estimates quoted by the company relate to **all** patients with Stage IIIb and Stage IV NSCLC and the population being considered in this appraisal is restricted to patients who have progressed after prior chemotherapy. However, as the K-M data from the OAK trial suggest that median life expectancy for patients receiving docetaxel is 9.6 months, the NICE End of Life criteria for short life expectancy criteria is met.

Extension to life

An examination of the ERG's remodelled OS suggests that treatment with atezolizumab generates a mean survival gain of 4.7 months compared to docetaxel. Suggesting that when the whole trial population is considered patient life expectancy is extended by more than 3 months when treatment with atezolizumab is compared with docetaxel.

However, when treatment with atezolizumab is compared with nintedanib+docetaxel, the size of the survival gain is uncertain. The company provided evidence during the clarification

process which suggests there is no statistically significant difference in OS for atezolizumab (total population) compared to nintedanib+docetaxel (adenocarcinoma patients only). If there is no statistically significant difference in OS, then, for the adenocarcinoma population, atezolizumab does not offer an extension of life of at least 3 months and so does not meet the NICE End of Life criteria for life extension.

8 OVERALL CONCLUSIONS

Discrepancies between the decision and the final scope issued by NICE

The ERG considers that the submitted evidence largely reflects the decision problem defined in the final scope issued by NICE, except that pembrolizumab was not considered as a comparator and the comparison of the efficacy of treatment with atezolizumab versus nintedanib+docetaxel should have been carried out using data from only the nintedanib+docetaxel licensed population, i.e. patients with adenocarcinoma. Furthermore, the ERG considers that full subgroup analyses based on levels of PD-L1 expression should have been undertaken.

Direct evidence

The direct clinical effectiveness evidence for the treatment of atezolizumab versus docetaxel was derived from the OAK and POPLAR trials. The ERG considers that both these trials were of good quality and were well conducted; patient characteristics were balanced across the groups, and statistical methods were generally appropriate. The ERG agrees with the company that the AE data from the OAK trial are consistent with the known AE profile of atezolizumab and that no new safety concerns have been highlighted. In terms of survival results (OS and PFS), the ERG considers that the company's median HR values from the OAK and POPLAR trials should be viewed with caution as the method used to calculate them relies on an assumption of PH which does not hold.

Results from both the OAK and POPLAR trials show that treatment with atezolizumab is associated with a statistically significant and clinically meaningful improvement in median OS (4.2 months in the OAK trial and 2.9 months in the POPLAR trial) compared to docetaxel in patients with ECOG PS 0 and 1. In the OAK trial, this statistically significant gain in OS is observed regardless of histology and PD-L1 status. However, in the POPLAR trial, statistically significant improvement is observed only in the non-squamous histology subgroup and for individuals with NSCLC of ≥1% PD-L1 expression.

Indirect evidence

The ERG considers that the company applied the ITC methodology using FP models appropriately but does not support the original ITC approach taken by the company as the:

- main network includes comparators that are not listed in the final scope issued by NICE and excludes pembrolizumab, which is listed in the final scope
- comparison of the effectiveness of atezolizumab versus nintedanib+docetaxel was carried out using data from the whole populations included in the trials and not just the population with adenocarcinoma histology (the nintedanib+docetaxel licensed population).

The ERG considers that the company's approach to the FP ITC is influenced by a range of factors (e.g., comparators and population selected, type of FP model chosen and the use of FE or RE) and that this means that it is difficult to identify the most appropriate combination of factors to use to generate ITC results. Furthermore, the ERG considers that the expected survival results generated by the ITC are difficult to interpret.

Economic evidence

The ERG considered that the mixed cure-rate model used to represent OS of patients receiving atezolizumab led to implausible OS estimates; this is due to the use of a log-logistic distribution, the cure-rate fraction, and an optimistic assumption that treatment with atezolizumab confers a lifetime effect. The ERG highlights that, at some time points, the company's OS model for atezolizumab produces survival estimates that are higher than the respective UK age-related population values.

The approach taken by the company to model OS for patients receiving docetaxel and nintedanib+docetaxel was to adjust the survival curve created to represent OS for patients receiving atezolizumab using hazard rates generated by their FP ITCs. Due to the implausibility of the company's atezolizumab OS model and methodological challenges related to the company's ITCs, the ERG considers that these OS models are unreliable.

Application of the ERG model amendments results in an ICER for the comparison of treatment with atezolizumab versus nintedanib+docetaxel of £1,170,793 per QALY gained. The ERG notes that this is a huge increase from the £56,076 per QALY gained estimated by the company and that this increase is largely due to the results of the company's ITC which showed that expected OS for patients receiving atezolizumab (total population) is not statistically significantly different from that of patients receiving nintedanib+docetaxel (adenocarcinoma population).

8.1 Implications for research

Grigg[79] et al highlight that currently published tissue studies have found PD-L1 positivity to indicate favourable, unfavourable, as well as variable correlations with histology and mutation status in NSCLC and other tumour types. Cree et al[32] identify a number of issues that they consider should be addressed to ensure that PD-L1 testing is introduced effectively into routine practice:

- relevance of tissue source and sample quality
- heterogeneity of PD-L1 expression within the tumour, between primary and metastatic lesions and over time
- impact of prior lines of treatment on PD-L1 expression

- optimal cut-offs identifying appropriate patient populations for treatment
- national and regional rates of PD-L1 positivity
- reproducibility and concordance of companion diagnostic kits and platforms
- validation of tumour infiltrating lymphocytes (LDTs) laboratory developed tests
- role of TILs and/or staining intensity in interpretation
- role of digital pathology.

The ERG considers that further research is required to address the issues identified by Cree et al.[32]

There is no direct clinical evidence to allow a comparison of the effectiveness of treatment, following chemotherapy, with atezolizumab versus nintedanib+docetaxel in patients with adenocarcinoma histology or versus pembrolizumab in patients with ≥1% PD-L1 expression. The results of head-to-head trials of atezolizumab versus these comparators in restricted patient populations would be useful.

9 REFERENCES

1. Roche Products Limited. Atezolizumab for treating locally advanced or metastatic NSCLC after chemotherapy ID970: Company submission to NICE. 2017; Available from: https://www.nice.org.uk/guidance/indevelopment/gid-ta10108/documents
2. National Institute for Health and Care Excellence (NICE). Lung cancer: diagnosis and management [CG121]. 2016; Available from: https://www.nice.org.uk/guidance/cg121/history [accessed April 2017].
3. National Institute for Health and Care Excellence (NICE). Afatinib for treating epidermal growth factor receptor mutation-positive locally advanced or metastatic non-small-cell lung cancer [TA310]. 2014; Available from: https://www.nice.org.uk/guidance/ta310/history [accessed April 2017].
4. National Institute for Health and Care Excellence (NICE). Nintedanib for previously treated locally advanced, metastatic, or locally recurrent non‐small‐cell lung cancer [TA347].2015.
5. National Institute for Health and Care Excellence (NICE). Pemetrexed for the first-line treatment of non-small cell lung cancer [TA181]. 2009; Available from: https://www.nice.org.uk/Guidance/TA181 [accessed March 2017].
6. National Institute for Health and Care Excellence (NICE). Gefitinib for the first-line treatment of locally advanced or metastatic non-small cell lung cancer [TA192]. 2010; Available from: https://www.nice.org.uk/guidance/TA192?UNLID=306327310201662222362O [accessed March 2017].
7. National Institute for Health and Care Excellence (NICE). Pemetrexed for the maintenance treatment of non-small-cell lung cancer [TA190]. 2010; Available from: https://www.nice.org.uk/guidance/TA190 [accessed March 2017].
8. National Institute for Health and Care Excellence (NICE). Erlotinib for the first-line treatment of locally advanced or metastatic EGFR-TK mutation-positive non-small cell lung cancer [TA258]. 2012; Available from: https://www.nice.org.uk/guidance/TA258 [accessed March 2017].
9. National Institute for Health and Care Excellence (NICE). Erlotinib and gefitinib for treating non-small-cell lung cancer that has progressed after prior chemotherapy [TA374]. 2015; Available from: https://www.nice.org.uk/guidance/ta374/history [accessed April 2017].
10. National Institute for Health and Care Excellence (NICE). Crizotinib for previously treated anaplastic lymphoma kinase-positive advanced non-small-cell lung cancer [TA422]. 2016; Available from: https://www.nice.org.uk/guidance/ta422/history [accessed April 2017].
11. National Institute for Health and Care Excellence (NICE). Ceritinib for previously treated anaplastic lymphoma kinase positive non-small-cell lung cancer [TA395]. 2016; Available from: https://www.nice.org.uk/guidance/ta395/history [accessed April 2017].
12. National Institute for Health and Care Excellence (NICE). Crizotinib for untreated anaplastic lymphoma kinase-positive advanced non-small cell lung cancer [TA406]. 2016; Available from: https://www.nice.org.uk/Guidance/TA406 [accessed March 2017].
13. National Institute for Health and Care Excellence (NICE). Pemetrexed maintenance treatment for non-squamous non-small-cell lung cancer after pemetrexed and cisplatin [TA402]. 2016; Available from: https://www.nice.org.uk/guidance/TA402 [accessed March 2017].
14. National Institute for Health and Care Excellence (NICE). Osimertinib for treating locally advanced or metastatic EGFR T790M mutation-positive non-small cell lung cancer [TA416]. 2016; Available from: https://www.nice.org.uk/guidance/ta416/chapter/1-recommendations [accessed March 2017].

15. National Institute for Health and Care Excellence (NICE). Pembrolizumab for treating PDL1-positive non-small-cell lung cancer after chemotherapy [TA428]. 2017; Available from: https://www.nice.org.uk/guidance/ta428/resources/pembrolizumab-for-treating-pdl1positive-nonsmallcell-lung-cancer-after-chemotherapy-82604670410437 [Accessed March 2017].
16. Cancer Research UK. Lung Cancer Incidence Statistics. 2017; Available from: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/lung-cancer [Accessed February 2017].
17. British Thoracic Society. Sharing information with lung cancer patients: guidance for healthcare professionals discussing options for patients who have lung cancer. 2013; Available from: https://www.brit-thoracic.org.uk/document-library/clinical-information/lung-cancer/sharing-information-with-lung-cancer-patients/ [accessed December 2016].
18. National Institute for Health and Care Excellence (NICE). Lung Cancer in Adults: Quality Standards (QS17). 2012; Available from: https://www.nice.org.uk/guidance/qs17/history [accessed April 2017].
19. Novello S, Barlesi F, Califano R, Cufer T, Ekman S, Levra MG*, et al.* Metastatic non-small-cell lung cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2016; 27:v1-v27.
20. National Comprehensive Cancer Network. NCCN non-small cell lung cancer clinical practice guideline, version 3.2017. 2016.
21. Boehringer Ingelheim. Single technology appraisal. Nintedanib in combination with docetaxel for the treatment of adult patients with locally advanced, metastatic or recurrent non-small cell lung cancer (NSCLC) of adenocarcinoma tumour histology after first-line chemotherapy [ID 438]. Company submission to NICE. 2014; Available from: https://www.nice.org.uk/guidance/ta347/history [accessed March 2017].
22. F. Hoffmann-La Roche Ltd. Clinical Study Report, GO28915/OAK, Report 1070445, Primary Analysis. 2016.
23. Rittmeyer A, Barlesi F, Waterkamp D, Park K, Ciardiello F, von Pawel J*, et al.* Atezolizumab versus docetaxel in patients with previously treated non-small-cell-lung cancer (OAK): a phase 3, open-label, multicentre randomised controlled trial. Lancet. 2016; 389:255–65.
24. F. Hoffmann-La Roche Ltd. Clinical Study Report, GO28753/POPLAR, Report 1065672, Primary Analysis. 2015.
25. Fehrenbacher L, Spira A, Ballinger M, Kowanetz M, Vansteenkiste J, Mazieres J*, et al.* Atezolizumab versus docetaxel for patients with previously treated non-small-cell lung cancer (POPLAR): a multicentre, open-label, phase 2 randomised controlled trial. Lancet. 2016; 387:1837-46.
26. Reck M, Kaiser R, Mellemgaard A, Douillard JY, Orlov S, Krzakowski M*, et al.* Docetaxel plus nintedanib versus docetaxel plus placebo in patients with previously treated non-small cell lung cancer (LUME-Lung 1): a phase 3, double-blind, randomised controlled trial. Lancet Oncol. 2014; 15:143-55.
27. Herbst RS, Soria JC, Kowanetz M, Fine GD, Hamid O, Gordon MS*, et al.* Predictive correlates of response to the anti-PD-L1 antibody MPDL3280A in cancer patients. Nature. 2014; 515:563-7.
28. Harshman LC, Drake CG, Choueiri TK. PD-1 blockade in renal cell carcinoma: to equilibrium and beyond. Cancer Immunol Res. 2014; 2:1132-41.
29. Butte MJ, Keir ME, Phamduy TB, Sharpe AH, Freeman GJ. Programmed death-1 ligand 1 interacts specifically with the B7-1 costimulatory molecule to inhibit T cell responses. Immunity. 2007; 27:111-22.
30. Yang J, Riella LV, Chock S, Liu T, Zhao XZ, Yuan XL*, et al.* The Novel Costimulatory Programmed Death Ligand 1/B7.1 Pathway Is Functional in Inhibiting Alloimmune Responses In Vivo. J Immunol. 2011; 187:1113-9.

31. Chaft JE, Chao B, Akerley WL, Gordon M, Antonia SJ, Callahan J*, et al.* Evaluation of PD-L1 Expression in Metachronous Tumor Samples and FDG-PET as a Predictive Biomarker in Ph2 Study (FIR) of Atezolizumab (MPDL3280A). J Thorac Oncol. 2015; 10:S176-S.
32. Cree I. PD-L1 testing for lung cancer in the UK: recognizing the challenges for implementation. Histopathology. 2016; 69:177-86.
33. Kerr KM, Hirsch FR. Programmed Death Ligand-1 Immunohistochemistry Friend or Foe? Arch Pathol Lab Med. 2016; 140:326-31.
34. Kowanetz M, Koeppen H, Boe M, Chaft JE, Rudin CM, Zou W*, et al.* Spatiotemporal Effects on Programmed Death Ligand 1 (PD-L1) Expression and Immunophenotype of Non-Small Cell Lung Cancer (NSCLC). J Thorac Oncol. 2015; 10:S199-S.
35. National Institute for Health and Care Excellence (NICE). Atezolizumab for treating metastatic urothelial bladder cancer after platinum-based chemotherapy [ID939]. 2017; Available from: https://www.nice.org.uk/guidance/indevelopment/gid-ta10111 [Accessed February 2016].
36. U.S. Food and Drug Administration (FDA). Atezolizumab (TECENTRIQ). 2016; Available from: https://www.fda.gov/Drugs/InformationOnDrugs/ApprovedDrugs/ucm525780.htm [accessed March 2017].
37. European Medicines Agency. TAXOTERE Summary of Product Characteristics. 2012 [updated 09 Nov 2016]; Available from: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Product_Information/human/000073/WC500035264.pdf.
38. National Institute for Health and Care Excellence (NICE). Final appraisal determination. Nintedanib for previously treated locally advanced, metastatic, or locally recurrent non-small-cell lung cancer. 2015; Available from: https://www.nice.org.uk/guidance/TA347/documents/lung-cancer-nonsmall-cell-nintedanib-final-appraisal-determination-document2 [accessed March 2017].
39. European Medicines Agency. Vargatef Summary of Product Characteristics. 2015 [updated 09 Nov 2016]; Available from: http://ec.europa.eu/health/documents/community-register/2014/20141121130020/anx_130020_en.pdf [accessed April 2017].
40. National Institute for Health and Care Excellence (NICE). Nintedanib for previously treated locally advanced, metastatic, or locally recurrent non‐small‐cell lung cancer [TA347]. 2015; Available from: https://www.nice.org.uk/guidance/ta347 [accessed March 2017].
41. National Institute for Health and Care Excellence (NICE). Lung cancer (non-small-cell, squamous, metastatic) - nivolumab (after chemotherapy) [ID811]. 2016 [updated 09 Nov 2016]; Available from: https://www.nice.org.uk/guidance/indevelopment/gid-tag506/ [Accessed April 2017].
42. National Institute for Health and Care Excellence (NICE). Lung cancer (non-small-cell, non-squamous, metastatic, after treatment) - nivolumab [ID900]. 2016 [updated 09 Nov 2016]; Available from: https://www.nice.org.uk/guidance/indevelopment/gid-tag524 [accessed April 2017].
43. Pao W, Chmielecki J. Rational, biologically based treatment of EGFR-mutant non-small-cell lung cancer. Nat Rev Cancer. 2010; 10:760-74.
44. Barlesi F, Park K, Ciardiello F, von Pawel J, Gadgeel S, Hida T*, et al.* Primary analysis from OAK, a randomized phase III study comparing atezolizumab with docetaxel in advanced NSCLC. Copenhagen, Denmark. 2016 [updated October 7–11]; Available from: http://www.roche.com/dam/jcr:f1e065e4-8c27-4e3f-8f48-da873f9a6f51/en/irp20161010.pdf [Accessed April 2017].

45. Gadgeel S, Ciardiello F, Rittmeyer A, Barlesi F, Cortinovis D, Barrios C*, et al.* OAK, a randomized phase III study of atezolizumab vs docetaxel in patients with advanced NSCLC: results from subgroup analyses. 2016 [updated December 4–7, 2016]; Available from: http://www.jto.org/article/S1556-0864(16)31252-7/abstract [Accessed April 2017].
46. Smith D, Vansteenkiste J, Fehrenbacher L, Park K, Mazieres J, Rittmeyer A*, et al.* Updated survival and biomarker analyses of a randomized phase II study of atezolizumab vs docetaxel in previously treated NSCLC (POPLAR). ASCO; 2016 June 3-7; Chicago.
47. Spira A, Park K, Mazieres J, Vansteenkiste J, Rittmeyer A, Ballinger M*, et al.* Efficacy, safety and predictive biomarker results from a randomized Phase II study comparing atezolizumab (MPDL3280A) vs docetaxel in 2L/3L NSCLC (POPLAR). Chicago. 2015 [updated May 29 - June 2]; Available from: http://meetinglibrary.asco.org/content/150315-156 [Accessed April 2017].
48. Vansteenkiste J, Fehrenbacher L, Spira A, Mazieres J, Park K, Smith D*, et al.* Atezolizumab monotherapy vs docetaxel in 2L/3L non-small cell lung cancer: Primary analysis for efficacy, safety and predictive biomarkers from a randomized phase II study (POPLAR). Vienna. 2015 [updated September 25-29]; Available from: http://itoc-conference.eu/files/2016/04/Rittmeyer_Achim.pdf [Acccessed April 2017].
49. Clinical trials.gov. A study of atezolizumab in participants with programmed death - ligand 1 (PD-L1) positive locally advanced or metastatic non-small-cell lung cancer (BIRCH). 2017; Available from: https://clinicaltrials.gov/ct2/show/NCT02031458?term=BIRCH+and+atezolizumab&rank=1 [accessed April 2017].
50. Clinical trials.gov. A Study of Atezolizumab in Participants With Programmed Death-Ligand 1 (PD-L1) Positive Locally Advanced or Metastatic Non-Small Cell Lung Cancer (NSCLC) [FIR]. 2017; Available from: https://clinicaltrials.gov/ct2/show/NCT01846416?cond=non+small+cell+lung+cancer&intr=Atezolizumab&rank=4 [accessed March 2017].
51. F. Hoffmann-La Roche Ltd. GO28915/OAK, Study Protocol, Version 62015.
52. F. Hoffmann-La Roche Ltd. GO28753/POPLAR, Study Protocol, Version 72013.
53. F. Hoffmann-La Roche Ltd. GO28915/OAK, Statistical Analysis Plan. 2013.
54. F. Hoffmann-La Roche Ltd. GO28753/POPLAR, Statistical Analysis Plan2015.
55. Clinical trials.gov. A Phase 1 Study of Atezolizumab (an Engineered Anti-Programmed Death-Ligand 1 [PDL1] Antibody) to Evaluate Safety, Tolerability and Pharmacokinetics in Participants With Locally Advanced or Metastatic Solid Tumors 2017; Available from: https://clinicaltrials.gov/ct2/show/NCT01375842?term=PCD4989g&rank=1 [accessed April 2017].
56. EuroQol Group. EQ-5D-3L instrument. 2015; Available from: http://www.euroqol.org/eq-5d-products.html [accessed April 2017].
57. European Organisation for Research and Treatment of Cancer (EORTC). EORTC QLQ-C30. 2016 [cited July]; Available from: http://groups.eortc.be/qol/eortc-qlq-c30 [accessed April 2017].
58. European Organisation for Research and Treatment of Cancer (EORTC). QLQ-LC13. 2016; Available from: http://groups.eortc.be/qol/sites/default/files/img/specimen_lc13_english.pdf [accessed April 2017].
59. Lunn DJ, Thomas, A., Best, N., Spiegelhalter, D. WinBUGS — a Bayesian modelling framework: concepts, structure, and extensibility. Stat Comput. 2000; 10:325-37.
60. Jansen JP. Network meta-analysis of survival data with fractional polynomials. BMC Med Res Method. 2011; 11:61.
61. Dias S, Sutton A. Evidence synthesis for decision making 2: a generalized linear modeling framework for pairwise and network meta-analysis of randomized controlled trials. Med Decis Making. 2013; 33:607-17.

62. Dias S, Welton N, Sutton A, Ades AE. NICE DSU technical support document 2: a generalised linear modelling framework for pairwise and network meta-analysis of randomised controlled trials (last updated September 2016). 2011.
63. Herbst RS, Baas P, Kim DW, Felip E, Perez-Gracia JL, Han JY*, et al.* Pembrolizumab versus docetaxel for previously treated, PD-L1-positive, advanced non-small-cell lung cancer (KEYNOTE-010): a randomised controlled trial. Lancet. 2016; 387:1540-50.
64. National Institute for Health and Care Excellence (NICE). Guide to the methods of technology appraisal. 2013; Available from: https://www.nice.org.uk/process/pmg9/chapter/foreword.
65. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. The Economic Evaluation of Health Care Programmes. Oxford: Oxford University Press; 2015.
66. Nafees B, Stafford M, Gavriel S, Bhalla S, Watkins J. Health state utilities for non small cell lung cancer. Health Qual Life Outcomes. 2008; 6:84.
67. National Institute for Health and Care Excellence (NICE). Ramucirumab for previously treated locally advanced or metastatic non-small-cell lung cancer [TA403]. 2016; Available from: https://www.nice.org.uk/guidance/ta403 [accessed April 2017].
68. Marti SG, Colantonio L, Bardach A, Galante J, Lopez A, Caporale J*, et al.* A cost-effectiveness analysis of a 10-valent pneumococcal conjugate vaccine in children in six Latin American countries. Cost Eff Resour Alloc. 2013; 11:21.
69. Hunter R. Cost-effectiveness of point-of-care c-reactive protein tests for respiratory tract infection in primary care in England. Adv Ther. 2015; 32:69-85.
70. Department of Health. Drugs and pharmaceutical electronic market information (eMit). 2011; Available from: https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit [accessed April 2017]
71. British National Formulary. MedicinesComplete Online. 2016; Available from: http://www.bnf.org/ [accessed March 2017].
72. Department of Health. NHS Reference Costs 2015-2016. 2015; Available from: https://www.gov.uk/government/publications/nhs-reference-costs-2015-to-2016 [Accessed April 2017].
73. Personal Social Services Research Unit. Unit Costs of Health and Social Care. 2016; Available from: http://www.pssru.ac.uk/project-pages/unit-costs/2016/ [accessed April 2017].
74. Department of Health. Healthcare Resource Group Costs. 2015; Available from: http://content.digital.nhs.uk/casemix/costing [accessed April 2017].
75. Kind P, Hardman G, Macran S. UK population norms for EQ-5D. Discussion paper 172. Centre for Health Economics, University of York. 1999.
76. National Institute for Health and Care Excellence (NICE). Cobimetinib in combination with vemurafenib for treating unresectable or metastatic BRAF V600 mutation-positive melanoma [TA414]. 2016; Available from: https://www.nice.org.uk/guidance/ta414 [accessed April 2017].
77. Office for National Statistics. Population estimates - summary for the UK, mid-2014. 2015; Available from: http://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates.
78. Beckett P, Callister M, Slade M, Harrison R, Fraffan J. Sharing Information with Lung Cancer Patients: Guidance for Healthcare Professionals Discussing Options for Patients who have Lung Cancer: British Toracic Society 2013.
79. Grigg C, Rizvi NA. PD-L1 biomarker testing for non-small cell lung cancer: truth or fiction? J Immunther Cancer. 2016; 4 48.

10 APPENDICES

10.1 Proportional Hazards Testing of the POPLAR trial

The company provided three diagnostic plots for OS and PFS indicating non PH between the treatment arms. The ERG agrees with the judgement of the company that the PH assumption does not hold for both OS and PFS in the POPLAR trial

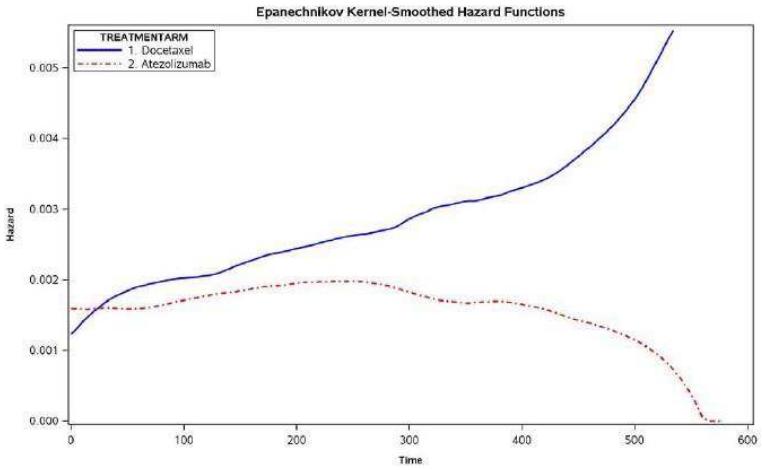
10.1.1 Overall survival

The company interpreted the three plots as follows:

For the OS hazard function plot, the two hazard curves crossed over at around 1 month and started to separate from each other around 8 months (Figure 19), the two curves of the log of negative log plots for OS overlapped at various time points and were clearly not parallel (

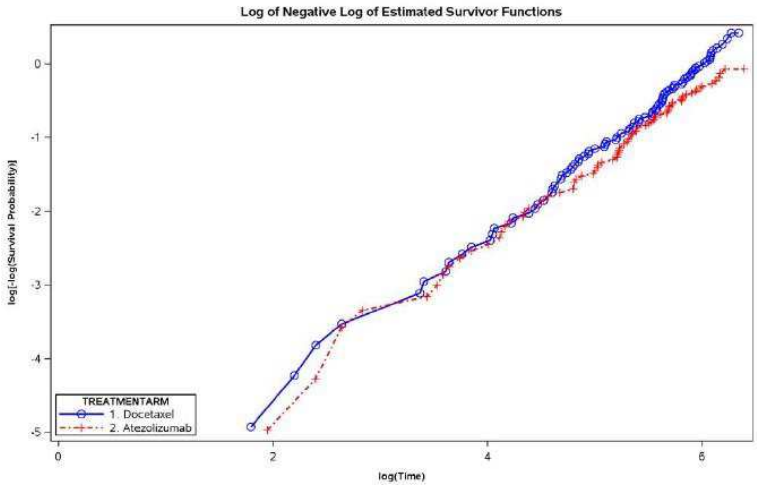
Figure 20) and for the log of survival plots, a trend of two lines passing the origin was observed, with one on top of the other, the overlap from randomization to approximately 3 months revealed a potential non-proportionality between the hazards of the two arms (Figure 21)

Figure 19 OS hazard function plot



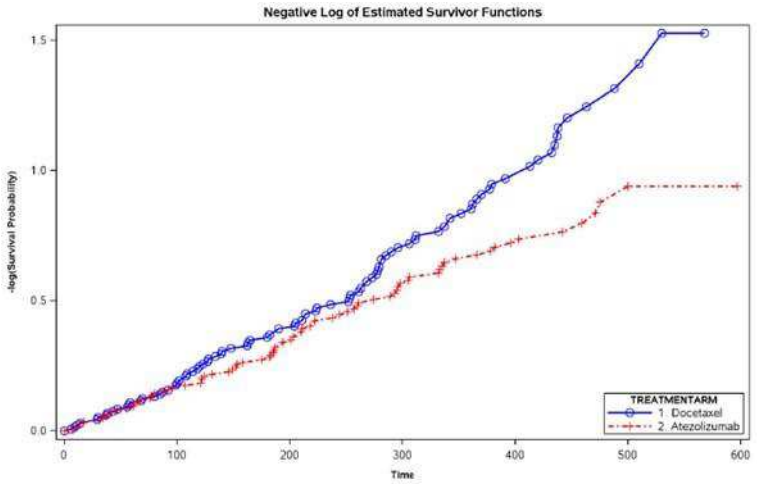
Source: Company response to ERG clarification letter

Figure 20 OS log of negative log plots



Source: Company response to clarification letter

Figure 21 OS log of survival plot



Source: Company response to ERG clarification letter

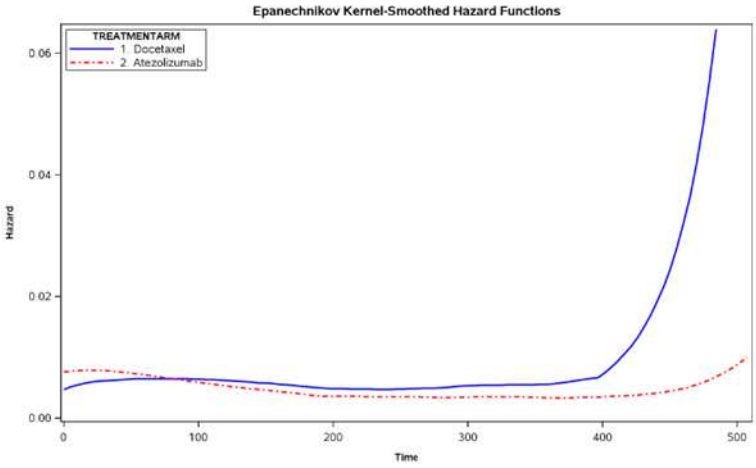
10.1.2 Progression-free survival

The company interpreted the three plots as follows:

For the PFS hazard function plot, the two curves were approximately parallel until around 13 months, corresponding to the minimum follow-up time (

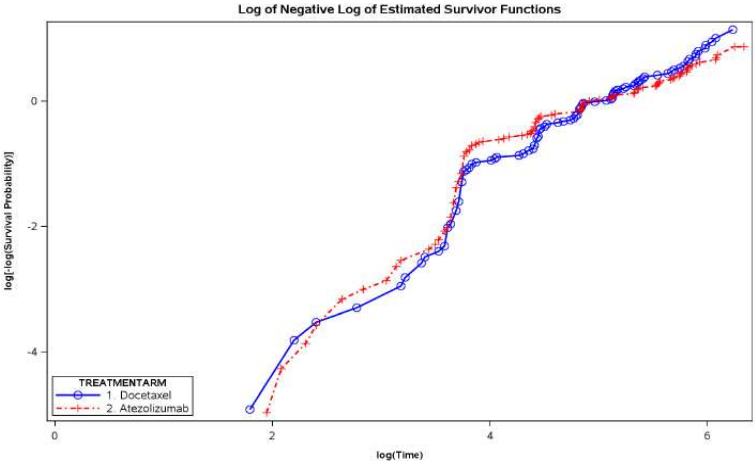
Figure 22), the two curves of the log of negative log plots for PFS overlapped (Figure 23) and the log of survival plots showed a cross-over pattern between the atezolizumab and docetaxel arms, where the crossing occurred approximately at 4-5 months (Figure 24).

Figure 22 PFS hazard function plot



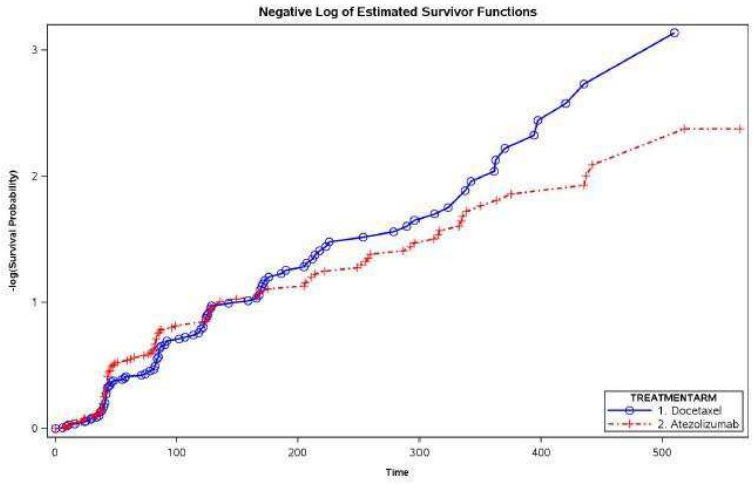
Source: Company response to ERG clarification letter

Figure 23 PFS log of negative log plots



Source: Company response to ERG clarification letter

Figure 24 PFS log of survival plot



Source: Company response to clarification letter

10.2 Additional secondary efficacy endpoints reported in the OAK and POPLAR trials

The main efficacy outcomes for the OAK and POPLAR trials were OS (primary outcome) and PFS (secondary outcome); the definitions and methods of analysis for these outcomes are presented in Table 9 of this report. ORR and DOR were also reported as secondary efficacy outcomes in the OAK and POPLAR trials, the definitions and methods of analysis for these outcomes are presented in Table 46.

Table 46 Description and method of analysis for secondary efficacy outcomes (other than time to progression and overall survival) reported in the OAK and POPLAR trials

Outcome	Outcome definition	Censoring definition	Statistical analysis
OAK			
ORR	Proportion of patients achieving confirmed best response of CR or PR per RECIST v1.1	Patients without any post baseline tumour assessments were considered non-responders	Clopper-Pearson methods for 95% CI of response rates and Mantel-Haenszel test for difference in rates
DOR	Interval between first documented objective response (CR or PR) and first documented PD or death	Date of last tumour measurement	Kaplan-Meier methodology, stratified in the 1°P and TC1/2/3 or IC1/2/3 subgroup , unstratified for all other subgroups
POPLAR			
ORR	Proportion of patients achieving confirmed best response of CR or PR per RECIST v1.1	n/a	Clopper-Pearson methods for 95% CI of response rates and Mantel-Haenszel test for difference in rates
DOR	Interval between first documented objective response (CR or PR) and first documented PD or death	Date of last tumour measurement	Kaplan-Meier methodology

1°P=primary population; CI=confidence interval; CR=complete response; DOR=duration of response; IC=tumour-infiltrating immune cell; ITT=intention-to-treat; ORR=objective response rate OS=overall survival, PD=progressive disease; PFS=progression-free survival; PR=partial response; RECIST=response evaluation criteria in solid tumours; TC=tumour cell
Source: CS, adapted from Table 21 and Table 22

The ERG is satisfied that the analysis method for each of these efficacy outcomes was pre-specified in the TSAPs, but notes a slightly different censoring definition for DOR in the OAK trial in the TSAP from the CS; the date of the first occurrence of a complete or partial response plus one day. The ERG is satisfied that all results were reported fully in the CSRs.

10.3 Additional analyses of overall survival reported in the POPLAR trial

The results for primary outcome OS for the POPLAR trial were presented from the updated analysis in the CS; these results are summarised in Section 4.3.2 of this report. OS was also analysed at two previous time points in the POPLAR trial for an interim analysis and the primary analysis and the extended follow-up in the POPLAR trial demonstrates increased benefit with atezolizumab compared to docetaxel. Results of OS in the POPLAR trial with increasing data maturity are presented in Table 47.

Table 47 OS in the POPLAR trial with increasing data maturity

Outcome	Atezolizumab (n=144)	Docetaxel (n=143)
Interim analysis[a]		
Median OS, months (95% CI)	11.4 (9.7 to NE)	9.5 (8.6 to 11.9)
HR (95% CI) – stratified for ITT population	0.77 (0.56 to 1.06, log-rank p-value=0.1145)	
Primary analysis[b]		
Median OS, months (95% CI)	12.6 (9.7 to 16.4)	9.7 (8.6 to 12.0)
HR (95% CI) – stratified for ITT population	0.73 (0.53 to 0.99, log-rank p-value=0.0404)	
Updated analysis[c]		
Median OS, months (95% CI)	12.6 (9.7 to 16.0)	9.7 (8.6 to 12.0)
HR (95% CI) – stratified for ITT population	0.69 (0.52 to 0.92, log rank p-value=0.011)	

[a] The data cut-off for the interim analysis was 30th January 2015
[b] The data cut-off date for the primary analysis was 8th May 2015
[c] The data cut-off date for the updated analysis was 1st December 2015
CI=confidence interval; ITT=intention-to-treat; NE=not evaluable; OS=overall survival
Source: CS, adapted from Section 4.7, Figure 12; POPLAR CSR adapted from Table 2, Table 21.

Overall survival results from the primary and updated analyses of the POPLAR trial according to histology and to PD-L1 status are presented in Table 48. The results presented in this table are discussed in Section 4.3.2 of this report.

Table 48 Overall survival results in the POPLAR trial according to histology and to PD-L1 status (primary and updated analyses)

Outcome	Primary analysis (data cut-off: 8st May 2015)		Updated analysis (data cut-off: 1st December 2015)	
	Atezolizumab	Docetaxel	Atezolizumab	Docetaxel
ITT population				
Number of participants analysed	144	143	144	143
Median OS, months (95% CI)	12.6 (9.7 to 16.4)	9.7 (8.6 to 12.0)	12.6 (9.7 to 16.0)	9.7 (8.6 to 12.0)
HR (95% CI) - stratified for ITT population	0.73 (0.53 to 0.99, log-rank p-value=0.0404)		0.69 (0.52 to 0.92, log-rank p-value=0.011)	
Histology: Non-squamous NSCLC				
Number of participants analysed	95	95	95	95
Median OS, months (95% CI)	15.5 (9.8 to NE)	10.9 (8.8 to 13.6)	14.8 (9.8 to 19.5)	10.9 (8.8 to 13.6)
HR (95% CI) – unstratified	0.69 (0.47 to 1.01, log-rank p-value=0.0562)		0.69 (0.49 to 0.98, log-rank p-value=0.039)	
Histology: Squamous NSCLC				
Number of participants analysed	49	48	49	48
Median OS, months (95% CI)	10.1 (6.7 to 14.5)	8.6 (5.4 to 11.6)	10.1 (6.7 to 14.5)	8.6 (5.4 to 11.6)
HR (95% CI) – unstratified	0.80 (0.49 to 1.30, log-rank p-value=0.3617)		0.66 (0.41 to 1.05, log-rank p-value=0.075)	
PD-L1 subgroup: TC3 or IC3				
Number of participants analysed	24	23	24	23
Median OS, months (95% CI)	15.5 (9.8 to NE)	11.1 (6.7 to 14.4)	NE (9.8 to NE)	11.1 (6.7 to 14.4)
HR (95% CI) – unstratified	0.49 (0.22 to 1.07, log-rank p-value=0.0684)		0.45 (0.22 to 0.95, log-rank p-value=0.033)	
PD-L1 subgroup: TC2/3 or IC2/3				
Number of participants analysed	50	55	50	55
Median OS, months (95% CI)	15.1 (8.4 to NE)	7.4 (6.0 to 12.5)	15.1 (8.4 to NE)	7.4 (6.0 to 12.5)
HR (95% CI) – unstratified	0.54 (0.33 to 0.89, log-rank p-value =0.0146)		0.50 (0.31 to 0.88, log-rank p-value=0.003)	
PD-L1 subgroup: TC2/3 or IC2/3 excluding TC3 or IC3				
Number of participants analysed	26	32	NR	NR
Median OS, months (95% CI)	9.0 (NR to NR)	6.2 (NR to NR)	NR	NR
HR (95% CI) – unstratified	0.59 (0.31 to 1.12, log-rank p-value =NR)		NR	
PD-L1 subgroup: TC1/2/3 or IC1/2/3				
Number of participants analysed	93	102	93	102
Median OS, months (95% CI)	15.5 (11.0 to NE)	9.2 (7.3 to 12.8)	15.1 (11.0 to NE)	9.2 (7.3 to 12.8)

HR (95% CI) – stratified for PD-L1 status	0.59 (0.40 to 0.85, log-rank p-value=0.0050)		0.59 (0.41 to 0.83, log-rank p-value=0.003)	
PD-L1 subgroup: TC1/2/3 or IC1/2/3 excluding TC2/3 or IC/23				
Number of participants analysed	43	47	NR	NR
Median OS, months (95% CI)	15.6 (NR to NR)	12.4 (NR to NR)	NR	NR
HR (95% CI) – unstratified	0.65 (0.37 to 1.16, log-rank p-value =NR)		NR	
PD-L1 subgroup: TC0 or IC0				
Number of participants analysed	51	41	51	41
Median OS, months (95% CI)	9.7 (6.7 to 12)	9.7 (6.8 to 12)	9.7 (6.7 to 12.0)	9.7 (8.6 to 12.0)
HR (95% CI) – unstratified	1.04 (0.62 to 1.75, log-rank p-value =0.8713)		0.88 (0.55 to 1.42, log-rank p-value=0.601)	

CI=confidence interval; IC=tumour-infiltrating immune cell; ITT=intention-to-treat; NE=not evaluable; NR=not reported NSCLC=non-small cell lung cancer; OS=overall survival; PD-L1=programmed death-ligand 1; TC=tumour cell
Source: adapted from Smith et al 2016; CS adapted from Section 4.7, Section 4.8, Figure 11 and Figure 16; POPLAR CSR, adapted from Table 23 and Table 34.

10.4 POPLAR trial: additional analyses of progression-free survival

The results for secondary outcome investigator assessed PFS for the POPLAR trial were presented from the updated analysis in the CS; these results are summarised in Section 4.3.3 of this report. PFS was also analysed at two previous time points in the POPLAR trial for an interim analysis and the primary analysis and the extended follow-up in the POPLAR trial; no statistically significant difference between treatment arms was observed at any time. Results of PFS in the POPLAR trial with increasing data maturity are presented in Table 49.

Table 49 POPLAR trial investigator-assessed PFS

Outcome	**Atezolizumab (n=144)**	**Docetaxel (n=143)**
Interim analysis[a]		
Median PFS, months (95% CI)	2.8 (2.1 to 4.1)	3.4 (2.8 to 4.1)
Hazard ratio (95% CI) – stratified for ITT population	0.98 (0.74 to 1.28, log-rank p-value=0.8606)	
Primary analysis[b]		
Median PFS, months (95% CI)	2.7 (2.0 to 4.1)	3.0 (2.8 to 4.1)
Hazard ratio (95% CI) – stratified for ITT population	0.94 (0.72 to 1.20, log-rank p-value=0.6450)	
Updated analysis[c]		
Median PFS, months (95% CI)	2.7 (2.0 to 4.1)	3.4 (2.8 to 4.1)
Hazard ratio (95% CI) – stratified for ITT population	0.92 (0.71 to 1.20, log-rank p-value=0.5560)	

[a] The data cut-off for the interim analysis was 30th January 2015
[b] The data cut-off date for the primary analysis was 8th May 2015
[c] The data cut-off date for the updated analysis was 1st December 2015
CI=confidence interval; ITT=intention-to-treat; PFS=progression-free survival
Source: CS, adapted from Section 4.7, Figure 13; POPLAR CSR adapted from Table 25; Table 48.

10.5 OAK and POPLAR trials: additional analyses of secondary endpoints

The results of the analyses for the secondary outcomes of the OAK and POPLAR trials not reported in the main body of this ERG report are provided in Table 50. ORR and DOR determined by the investigator per RECIST v1.1 were analysed for the 1°P of the OAK trial and the ITT population of the POPLAR trial. The company confirmed that no blinded independent central review of any endpoints explored in either of the OAK or POPLAR trials.

Table 50 ORR and DOR among responders (OAK and POPLAR trials)

Outcome	OAK (primary analysis)		POPLAR (updated analysis)	
	Atezolizumab	Docetaxel	Atezolizumab	Docetaxel
ORR per RECIST v1.1				
Number of participants analysed	425	425	144	143
Responders, n (%) (95% CI)	58 (13.6) (10.53 to 17.28)	57 (13.4) (10.32 to 17.02)	22 (15.3) (9.8 to 22)	21 (14.7) (9.3 to 21.6)
Complete response, n (%) (95% CI)	6 (1.4) (0.52 to 3.05)	1 (0.2) (0.01 to 1.30)	1 (0.7) (NR to NR)	0 (0) (NR to NR)
Partial response, n (%) (95% CI)	52 (12.2) (9.27 to 15.73)	56 (13.2) (10.11 to 16.77)	21 (14.6) (NR to NR)	21 (14.7) (9.3 to 21.6)
Stable disease, n (%) (95% CI)	150 (35.3) (30.75 to 40.05)	177 (41.6) (36.92 to 46.50)	43 (30.0) (NR to NR)	NR (NR to NR)
Progressive disease, n (%) (95% CI)	187 (44.0) (39.22 to 48.86)	117 (27.5) (23.33 to 32.04)	NR (NR to NR)	NR (NR to NR)
DOR among responders				
Number of participants (responders) analysed	58	57	22	21
Patients without event, n (%)	30 (51.7)	10 (17.5)	11 (50)	3 (14)
Median duration of response, months (95% CI)	16.3 (10.0 to NE)	6.2 (4.9 to 7.6)	18.6 (11.6 to NE)	7.2 (5.6 to 12.5)
Unstratified HR, (95% CI)	0.34 (0.21 to 0.55)		0.32 (0.15 to 0.70)	

CI=confidence interval; DOR=duration of response; HR=hazard ratio; NE=not evaluable; NR= not reported; ORR=objective response rate; RECIST=response evaluation criteria in solid tumours
Source: CS adapted from Section 4.7, Table 29, Table 30, Table 31.

In the OAK trial, the proportion of patients with complete response per RECIST v1.1 was similar across the treatment arms. Limited numerical results for ORR were available in the CS for the updated analysis of the POPLAR trial; ORR results for the primary analysis of the POPLAR trial are available in Table 27 of the POPLAR CSR. The company states that the proportion of patients with confirmed response was similar in the atezolizumab and docetaxel arms and that results of updated analysis did not significantly change compared to the primary analysis.

In both the OAK and POPLAR trials, the median DOR was more than doubled in the atezolizumab arms compared with the docetaxel arms in the ITT population of the trials; OAK:

16.3 compared to 6.2 months, HR 0.34 (95% CI 0.21 to 0.55) and POPLAR 18.6 compared to 7.2 months, HR 0.32 (95% CI 0.15 to 0.70).

10.6 Additional characteristics of trials included in the indirect treatment comparison

Three trials were identified for inclusion in the ITC; design characteristics of the OAK and POPLAR trials are summarised in Table 6 and an assessment of the risk of bias of the OAK and POPLAR trials is provided in Section 4.2.5 of this report.

Design characteristics of the LUME-Lung 1 trial are summarised in Table 51 and an assessment of the risk of bias of the LUME-Lung 1 trial is provided in

Table 52.

Table 51 Key characteristics of the LUME-Lung 1 trial

Characteristic	LUME-Lung 1
Location	International (27 countries, 211 centres)
Design	Randomised (1:1), phase III, double blind, placebo controlled
Population	A total of 1314 patients were randomised and analysed as the ITT population, 655 to docetaxel plus nintedanib and 659 to docetaxel plus placebo
Intervention	Docetaxel (75mg/m^2) by intravenous infusion on day 1 plus nintedanib 200mg orally BID on days 2-21 every 3 weeks, until unacceptable side effects or disease progression.
Comparator	Docetaxel (75mg/m^2) by intravenous infusion on day 1 plus matching placebo orally BID on days 2-21 every 3 weeks, until unacceptable side effects or disease progression.
Primary outcome	PFS (defined as time from randomisation to progression or death) by central independent review
Secondary outcomes	OS, investigator assessed PFS, tumour response by central review and investigator assessment
Safety endpoints	Safety and tolerability of treatment with docetaxel plus nintedanib compared to docetaxel plus placebo
Duration of trial	Patients were enrolled between 23rd December 2008 and 9th February 2011
Data analyses	Data cut-off for analysis of the ITT population: 15th February 2013
Median duration of follow-up (primary analysis)	PFS: 7.1 months (IQR 3.8 to 11.0 months) OS: 31.7 months (IQR 27.8 to 36.1 months)

BID=twice daily; IQR=interquartile range; ITT=intention to treat; OS=overall survival; PFS=progression-free survival
Source Reck et al 2014; CS Appendix 4.

Table 52 Risk of bias assessment of the LUME-lung 1 trial

Risk of bias question	Company assessment	ERG comment
Was randomisation carried out appropriately?	Low risk of bias	Agree; web-based block randomisation (by country) via web-based interactive response system
Was the concealment of treatment allocation adequate?	Low risk of bias	Agree; treatment assigned by an interactive third party (web-based response system)
Were the groups similar at the outset of the trial in terms of prognostic factors, for example, severity of disease?	Low risk of bias	Agree; demographic and baseline characteristics were well balanced between the treatment groups
Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	Low risk of bias	Agree; blinding of care providers and participants achieved by matching placebo. Primary outcome (PFS) primarily assessed by central review.
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	Low risk of bias	Agree; Reasons for withdrawal from treatment adequately reported, all randomised participants included as the ITT population for analyses of primary and secondary outcomes.
Is there any evidence to suggest that the authors measured more outcomes than they reported?	Low risk of bias	Agree; Results for all outcomes specified in the published protocol and clinical trial registry are available
Did the analysis include an intent-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Low risk of bias	Agree; all randomised participants included as the ITT population for analyses of primary and secondary outcomes.

ITT=intention to treat; PFS=progression-free survival
Source Reck et al 2014; CS Appendix 4.

10.7 Additional survivor plots of fractional polynomial models

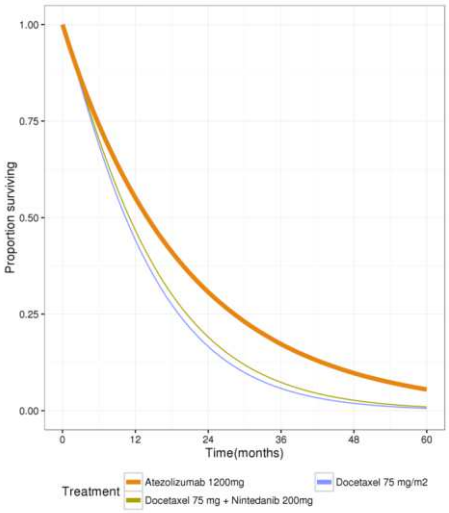
The company provided survivor plots as an output of each FE FP ITC model fitted in the response letter to ERG clarification questions. Survivor plots were not made available to the ERG by the company for each RE FP ITC model fitted.

Visual inspection of these fitted survival curves, along with the DIC statistic, was used by the company to determine the best fitting model for OS and PFS. Visual inspection of these curves conducted by the ERG is discussed in Section 4.6.3 of this report.

The company also provided graphical plots of the resulting HR functions for the best fitting model for each outcome; the Weibull FE FP model for both OS and PFS.

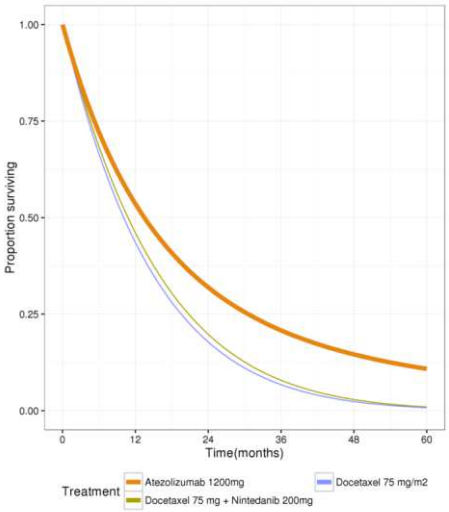
10.7.1 Overall survival

Figure 25 Survivor plot (1st order, p1=0 [Weibull])



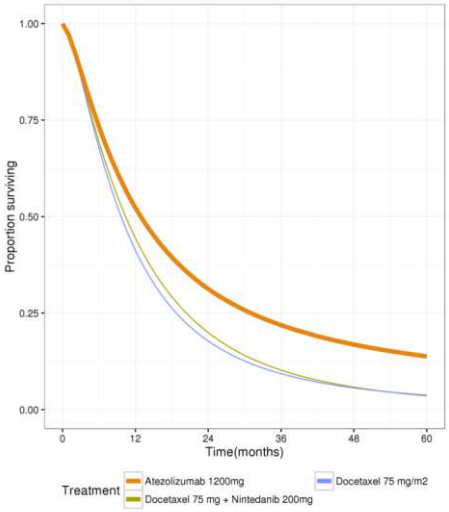
Source: Company response to clarification letter

Figure 26 Survivor plot (1st order, p1=0 [Gompertz])



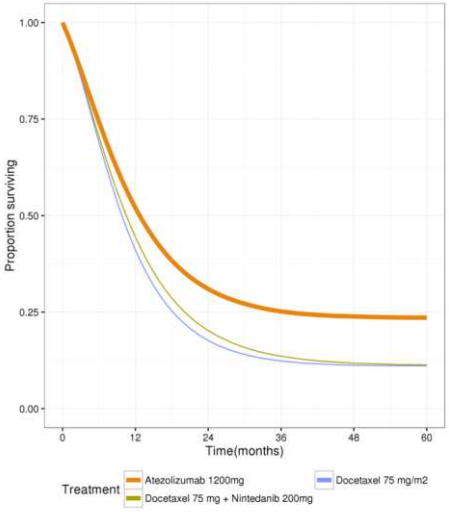
Source: Company response to clarification letter

Figure 27 Survivor plots (2nd order (1), p1=0, p2=0)



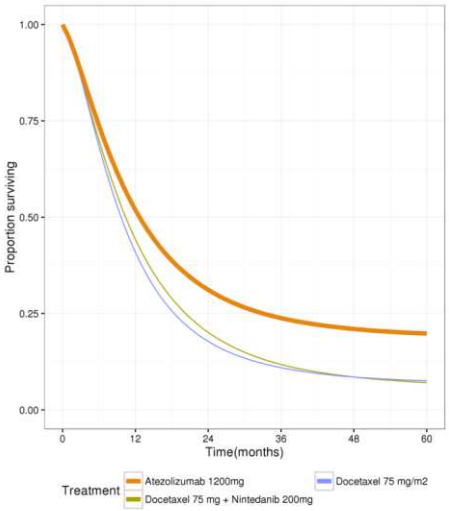
Source: Company response to clarification letter

Figure 28 Survivor plots (2nd order (2), p1=0, p2=1)



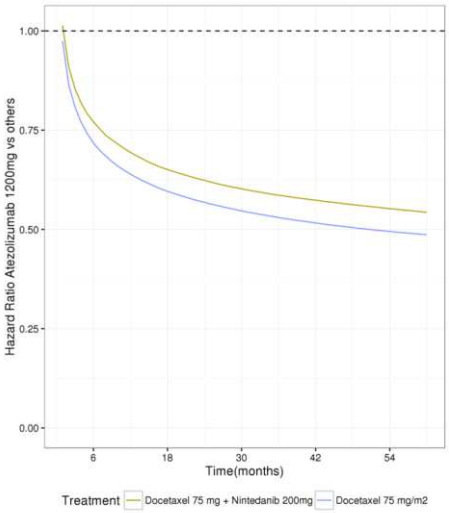
Source: Company response to clarification letter

Figure 29 Survivor plots (2nd order (3), p1=1, p2=1)



Source: Company response to clarification letter

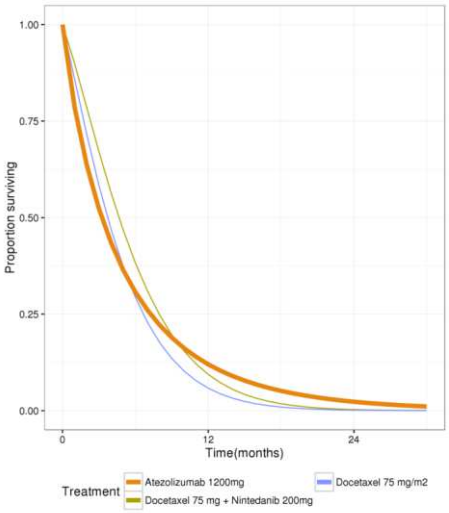
Figure 30 Hazard ratio functions of best fitting model (1st order, Weibull)



Source: Company response to clarification letter

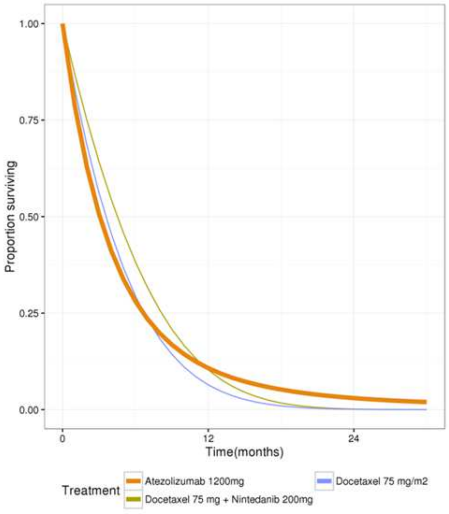
10.7.2 Progression-free survival

Figure 31 Survivor plot (1st order, p1=0 [Weibull])



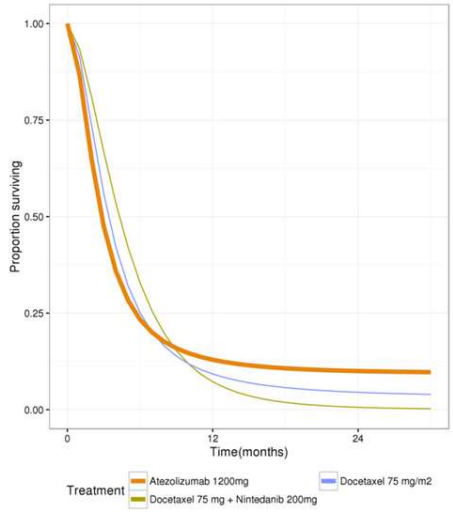
Source: Company response to clarification letter

Figure 32 Survivor plot (1st order, p1=1 [Gompertz])



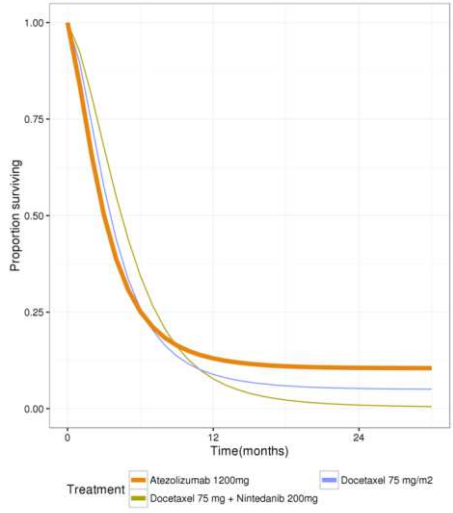
Source: Company response to clarification letter

Figure 33 Survivor plot (2nd order (1), p1=0, p2=0)



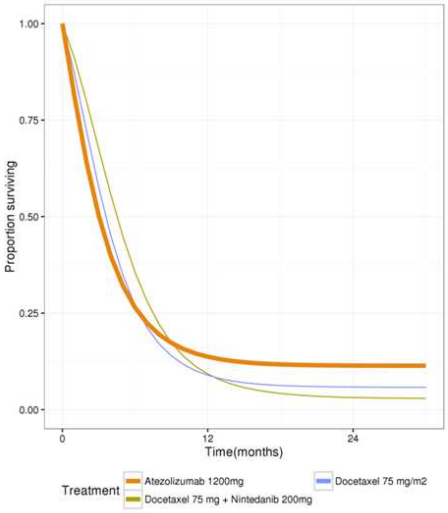
Source: Company response to clarification letter

Figure 34 Survivor plot (2nd order (2), p1=0, p2=1)



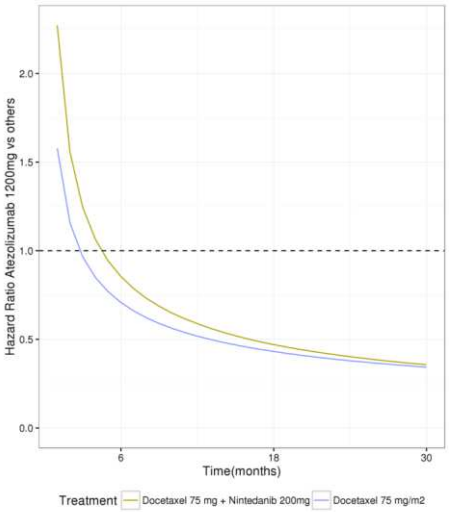
Source: Company response to clarification letter

Figure 35 Survivor plot (2nd order (3), p1=1, p2=1)



Source: Company response to clarification letter

Figure 36 Hazard ratio functions of best fitting model (1st order, Weibull)

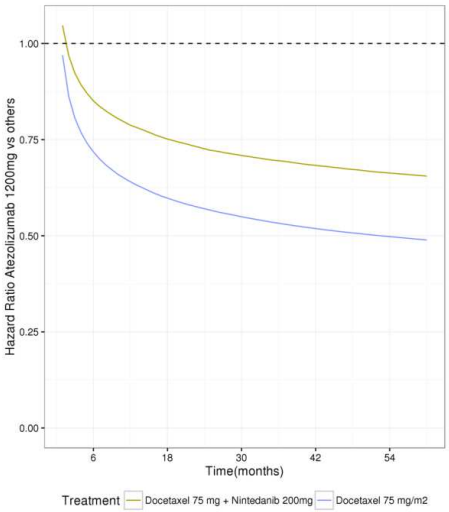


Source: Company response to clarification letter

10.7.3 Adenocarcinoma histology

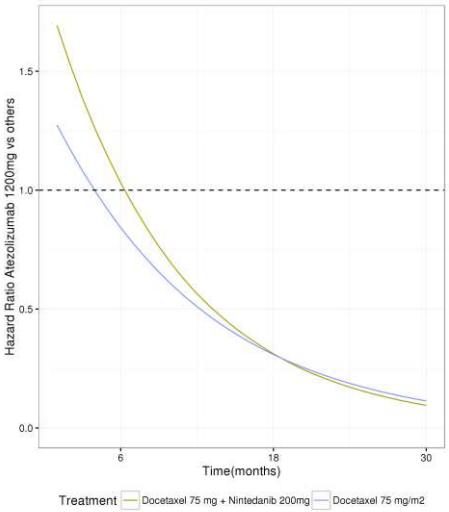
The ERG requested that the company perform additional ITCs to account for the population for which nintedanib+docetaxel is licenced for (i.e. patients with adenocarcinoma histology only) and the company provided results which compared atezolizumab within its intended licence to nintedanib+docetaxel (adenocarcinoma subgroup), see Section 4.6.4 for further discussion of this additional ITC and the results for the expected difference in OS and PFS. The company also provided graphical plots of the resulting HR functions for the best fitting model for each outcome; the Weibull FE FP model for both OS and PFS.

Figure 37 Hazard ratio functions of best fitting model (1st order, Weibull) for overall survival; atezolizumab (total population) and nintedanib+docetaxel (adenocarcinoma subgroup)



Source: Company response to clarification letter

Figure 38 Hazard ratio functions of best fitting model (1st order, Weibull) for PFS, atezolizumab (total population) and nintedanib+docetaxel (adenocarcinoma subgroup)

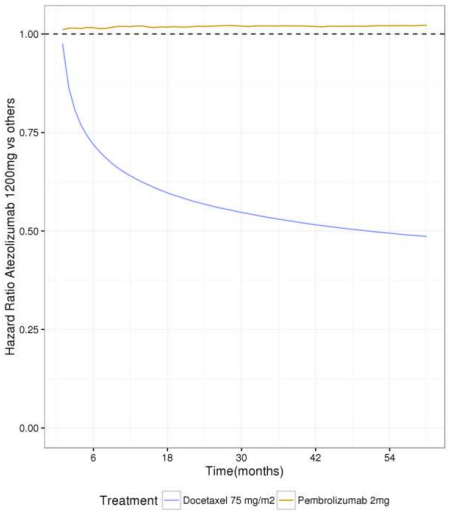


Source: Company response to clarification letter

10.7.4 Inclusion of pembrolizumab

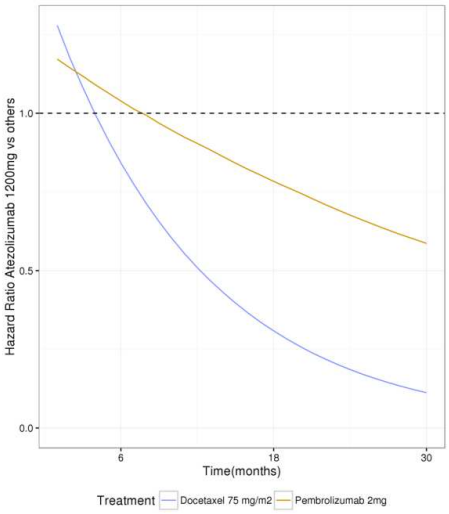
The ERG requested that the company perform additional ITCs to extend the scope of the submission to pembrolizumab, which had previously been excluded from the appraisal, see Section 4.6.4 for further discussion of this additional ITC and the results for the expected difference in OS and PFS. The company also provided graphical plots of the resulting HR functions for the best fitting model for each outcome; the Weibull FE FP model for both OS and PFS.

Figure 39 Hazard ratio functions of best fitting model (1st order, Weibull) for OS



Source: Company response to clarification letter

Figure 40 Hazard ratio functions of best fitting model (1st order, Weibull) for PFS



Source: Company response to clarification letter

10.8 ERG revisions to the company model

This appendix contains details of the changes that the ERG made to the company model. Information contained with the file named 'ID970 ERG corrected OS.xlsx' is needed to replicate the ERG's cost effectiveness results.

ERG Section 6 results table revision	Implementation instructions
C1 Discounting algorithms	In Sheets 'Atezo', 'Doce' and 'Ninted + Doce – ITC' Set value in cell **C13**= 0 Copy cell **C13** to **C14:C65** In Sheet 'Ninted + Doce – ITC' Insert formula in cell **C66** = ROUNDDOWN(F66/wk2yr,2) Copy cell **C66** to **C67:C1578**

C2 Age related utility decrement	In Sheet 'Atezo' Set formula in cell **CM118** =IF(util_optn=4,((BB118*u_Pre5On-0.02)+BC118*(u_Pre15On-0.02)+BD118*(u_Pre30On-0.02)+BE118*(u_Post30On-0.02))*(BV118*AV118),IF(util_optn=1,BV118*AV118*(u_pfs_new-0.02),AW118*(u_pfs_new-0.02))) Copy cell **CM118** to **CM119:CM586** Set formula in cell **CM587** =IF(util_optn=4,((BB587*u_Pre5On-0.07)+BC587*(u_Pre15On-0.07)+BD587*(u_Pre30On-0.07)+BE587*(u_Post30On-0.07))*(BV587*AV587),IF(util_optn=1,BV587*AV587*(u_pfs_new-0.07),AW587*(u_pfs_new-0.07))) Copy cell **CM587** to **CM588:CM1578** Set formula in cell **CO118** =IF(util_optn=4,(BB118*(u_Pre5Off-0.02)+BC118*(u_Pre15Off-0.02)+BD118*(u_Pre30Off-0.02)+BE118*(u_Post30Off-0.02))*(AY118-BV118*AV118),IF(util_optn=1,(AY118-BV118*AV118)*(u_prog-0.02),AX118*(u_prog-0.02))) Copy cell **CO118** to **CO119:CO586** Set formula in cell **CO587** =IF(util_optn=4,(BB587*(u_Pre5Off-0.07)+BC587*(u_Pre15Off-0.07)+BD587*(u_Pre30Off-0.07)+BE587*(u_Post30Off-0.07))*(AY587-BV587*AV587),IF(util_optn=1,(AY587-BV587*AV587)*(u_prog-0.07),AX587*(u_prog-0.07))) Copy cell **CO587** to **CO588:CO1578** In Sheet 'Doce' Set formula in cell **CJ118 =**IF(util_optn=4,(AZ118*(u_Pre5On-0.02)+BA118*(u_Pre15On-0.02)+BB118*(u_Pre30On-0.02)+BC118*(u_Post30On-0.02))*(BT118*AT118),IF(util_optn=1,BT118*AT118*u_pfs_com,AU118*u_pfs_com)) Copy cell **CJ118** to **CJ119:CJ586** Set formula in cell **CJ587 =** IF(util_optn=4,(AZ587*(u_Pre5On-0.07)+BA587*(u_Pre15On-0.07)+BB587*(u_Pre30On-0.07)+BC587*(u_Post30On-0.07))*(BT587*AT587),IF(util_optn=1,BT587*AT587*u_pfs_com,AU587*u_pfs_com)) Copy cell **CJ587** to **CJ588:CJ1578** Set formula in cell **CL118** =IF(util_optn=4,(AZ118*(u_Pre5Off-0.02)+BA118*(u_Pre15Off-0.02)+BB118*(u_Pre30Off-0.02)+BC118*(u_Post30Off-0.02))*(AW118-BT118*AT118),IF(util_optn=1,(AW118-BT118*AT118)*u_prog,AV118*u_prog)) Copy cell **CL118** to **CL119:CL586** Set formula in cell **CL587** =IF(util_optn=4,(AZ587*(u_Pre5Off-0.07)+BA587*(u_Pre15Off-0.07)+BB587*(u_Pre30Off-0.07)+BC587*(u_Post30Off-0.07))*(AW587-BT587*AT587),IF(util_optn=1,(AW587-BT587*AT587)*u_prog,AV587*u_prog)) Copy cell **CL587** to **CL588:CL1578** In Sheet 'Ninted + Doce – ITC'

ERG Section 6 results table revision	Implementation instructions
	Set formula in cell **AR118** =IF(util_optn=4,(U118*(u_Pre5On-0.02)+V118*(u_Pre15On-0.02)+W118*(u_Pre30On-0.02)+X118*(u_Post30On-0.02))*P118,P118*u_pfs_com3) Copy cell **AR118** to **AR119:AR586** Set formula in cell **AR587** =IF(util_optn=4,(U587*(u_Pre5On-0.07)+V587*(u_Pre15On-0.07)+W587*(u_Pre30On-0.07)+X587*(u_Post30On-0.07))*P587,P587*u_pfs_com3) Copy cell **AR587** to **AR588:AR1578** Set formula in cell **AT118** =IF(util_optn=4,(U118*(u_Pre5Off-0.02)+V118*(u_Pre15Off-0.02)+W118*(u_Pre30Off-0.02)+X118*(u_Post30Off-0.02))*Q118,Q118*u_prog) Copy cell **AT118** to **AT119:AT586** Set formula in cell **AT587** =IF(util_optn=4,(U587*(u_Pre5Off-0.07)+V587*(u_Pre15Off-0.07)+W587*(u_Pre30Off-0.07)+X587*(u_Post30Off-0.07))*Q587,Q587*u_prog) Copy cell **AT587** to **AT588:AT1578**

ERG Section 6 results table revision	Implementation instructions
C3 ToT half cycle correction	In Sheet 'Atezo' Set formula in cell **BV13 =** BU13 Copy cell **BV13** to **BV14:BV1578** In Sheet 'Doce'' Set formula in cell **BT13 =** BS13 Copy cell **BT13** to **BT14:BT1578** In Sheet 'Ninted + Doce – ITC' Set formula in cell **Z13 =** Y13 Copy cell **Z13** to **Z14:Z1578** Set formula in cell **AB13=** IF(MOD(F13,(1/cyc2wk))=0,1,0)*Y13*AA13*IF(F13=0,c_adm1_com3,c_adm_com3) Copy cell **AB13** to **AB14:AB1578** Set formula in cell **AD13=** **I**F(F13<doc_cap,IF(MOD(F13,(1/cyc2wk))=0,1,0)*Y13*AA13*IF(F13=0,c_adm1_com,c_adm_com),0) Copy cell **AD13** to **AD14:AD1578**
R1. ERG preferred OS for atezolizumab and docetaxel	Open Worksheet **'OS'** from ID970 ERG corrected OS.xlsx Copy cells A2:A1567 In sheet 'Atezo' Paste in cells AQ13:AQ1578 Open Worksheet **'OS'** from ID970 ERG corrected OS.xlsx Copy cells C2:C1567 In Sheet 'Doce' Paste in cells AO13:AO1578

ERG Section 6 results table revision	Implementation instructions
R2. ERG preferred OS for atezolizumab and docetaxel and atezolizumab treatment duration effect set to 5 years	Open Worksheet **'OS'** from ID970 ERG corrected OS.xlsx Copy cells B2:B1567 In sheet 'Atezo' Paste in cells AQ13:AQ1578 Open Worksheet **'OS'** from ID970 ERG corrected OS.xlsx Copy cells C2:C1567 In Sheet 'Doce' Paste in cells AO13:AO1578
R3 ERG preferred OS for atezolizumab applied to both atezolizumab and nintedanib+docetaxel	Open Worksheet **'OS'** from ID970 ERG corrected OS.xlsx Copy cells A2:A1567 In sheet 'Atezo' Paste in cells AQ13:AQ1578 Open Worksheet **'OS'** from ID970 ERG corrected OS.xlsx Copy cells A2:A1657 In Sheet 'Ninted + Doce – ITC' Paste in cells J13:J1578
R4 ERG preferred OS for atezolizumab applied to both atezolizumab and nintedanib+docetaxel and treatment duration effect for both set to 5 years	Open Worksheet **'OS'** from ID970 ERG corrected OS.xlsx Copy cells B2:B1567 In sheet 'Atezo' Paste in cells AQ13:AQ1578 Open Worksheet **'OS'** from ID970 ERG corrected OS.xlsx Copy cells B2:B1567 In Sheet 'Ninted + Doce – ITC' Paste in cells J13:J1578