in collaboration with:

ERASMUS UNIVERSITEIT ROTTERDAM
INSTITUTE OF HEALTH POLICY & MANAGEMENT

Maastricht University

Durvalumab for treatment of locally advanced, unresectable, Stage III non-small cell lung cancer in adults whose tumours express PD-L1 on ≥1% of tumour cells and whose disease has not progressed after platinum-based chemo-radiation therapy

Produced by	Kleijnen Systematic Reviews Ltd. in collaboration with Erasmus University Rotterdam (EUR) and Maastricht University
Authors	Robert Wolff, Deputy Director, Kleijnen Systematic Reviews Ltd, UK Sabine Grimm, Health Economist, Maastricht UMC Thea van Asselt, Health Economist, Maastricht UMC Willem Witlox, Health Economist, Maastricht UMC Nigel Armstrong, Health Economics Manager, KSR Ltd Gill Worthy, Statistician, KSR Ltd Annette Chalker, Systematic Reviewer, KSR Ltd Titas Buksnys, Health Economist, KSR Ltd Lisa Stirk, Information Specialist, KSR Ltd Manuela Joore, Health Economist, Maastricht UMC Jos Kleijnen, Director, KSR Ltd; Professor of Systematic Reviews in Health Care, Maastricht University
Correspondence to	Robert Wolff, Kleijnen Systematic Reviews Unit 6, Escrick Business Park Riccall Road, Escrick York, UK YO19 6FD
Date completed	08/11/2018 updated 20/02/2019

Source of funding: This report was commissioned by the NIHR HTA Programme as project number STA 17/109/12.

Declared competing interests of the authors
None.

Acknowledgements
None.

Commercial in confidence (CiC) data are highlighted in blue throughout the report.

Academic in confidence (AiC) data are highlighted in yellow throughout 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:
Wolff R, Grimm S, van Asselt, ADI, Witlox W, Armstrong N, Worthy G, Chalker A, Buksnys T, Stirk L, Joore MA, Kleijnen J. Durvalumab for treatment of locally advanced, unresectable, Stage III, non-small cell lung cancer in adults whose tumours express PD-L1 on ≥1% of tumour cells and whose disease has not progressed after platinum-based chemo-radiation therapy: A Single Technology Appraisal. York: Kleijnen Systematic Reviews Ltd, 2018.

Contributions of authors
Robert Wolff acted as project lead and systematic reviewer on this assessment, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Sabine Grimm acted as health economic project lead, critiqued the company's economic evaluation and contributed to the writing of the report. Thea van Asselt, Willem Witlox, Titas Buksnys, and Nigel Armstrong acted as health economists on this assessment, critiqued the company's economic evaluation and contributed to the writing of the report. Annette Chalker and Titas Buksnys acted as a systematic reviewer, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Gill Worthy acted as statistician, critiqued the analyses in the company's submission and contributed to the writing of the report. Lisa Stirk critiqued the search methods in the submission and contributed to the writing of the report. Manuela Joore acted as health economist on this assessment, critiqued the company's economic evaluation, contributed to the writing of the report and provided general guidance. Jos Kleijnen critiqued the company's definition of the decision problem and their description of the underlying health problem and current service provision, contributed to the writing of the report and supervised the project.

Abbreviations

1-L	First-line
2-L	Second-line
AACR	American Association for Cancer Research
AE	Adverse event
AIC	Akaike information criterion
AiC	Academic in confidence
AJCC	American Joint Committee on Cancer
ASCO	American Society of Clinical Oncology
AUC	Area under the curve
BIC	Bayesian information criterion
BICR	Blinded independent central review
BSC	Best supportive care
BTOC	British Thoracic Oncology Group
CADTH	Canadian Agency for Drugs and Technologies in Health
CDSR	Cochrane Database of Systematic Reviews
CHMP	Committee for Medicinal Products for Human Use
CI	Confidence interval
CiC	Commercial in confidence
CRF	Case report form
CRT	Chemoradiation therapy
CS	Company submission
CSR	Clinical study report
CTCAE	Common Terminology Criteria for Adverse Events
CT	Computed Tomography
CTx	Chemotherapy
DARE	Database of Abstracts of Reviews of Effects
DCO	Data cut-off
DCR	Disease control rate
DoR	Duration of response
DSA	Deterministic sensitivity analysis
EAP	Early access program
EGFR	Epidermal growth factor receptor
ELCC	European Lung Cancer Conferences
EMA	European Medicines Agency
EORTC	European Organisation for Research and Treatment of Cancer
EORTC QLQ-C30	European Organisation for Research and Treatment of Cancer 30-item core quality of life questionnaire
EORTC	European Organisation for Research and Treatment of Cancer quality of life questionnaire and lung cancer module
eMIT	Electronic Market Information Tool
EPAR	European public assessment report
EQ-5D	European Quality of Life-5 Dimensions
EQ-5D-3L	European Quality of Life-5 Dimensions, three-level scale
EQ-5D-5L	European Quality of Life-5 Dimensions, five-level scale
ERG	Evidence Review Group
ESMO	European Society for Medical Oncology
EUR	Erasmus University Rotterdam
EuroQOL SC	EuroQOL self-classifier
EuroQOL VAS	EuroQOL visual analogue scale
FE	Fixing errors
FV	Fixing violations
HR	Hazard ratio
HRG	Healthcare Resource Group

HRQoL	Health-related quality of life
HTA	Health technology assessment
ICER	Incremental cost effectiveness ratio
IDMC	Independent data monitoring committee
IO	Immuno-oncological
ITT	Intention-to-treat
IV	Intravenous
IVRS	Interactive voice response system
IWRS	Interactive web response system
kg	Kilogram
KM	Kaplan–Meier
KSR	Kleijnen Systematic Reviews
LY	Life year
LYG	Life years gained
MedDRA	Medical Dictionary for Regulatory Activities
mg	Milligram
MIMS	Monthly Index of Medical Specialities
MJ	Matters of judgements
N/A	Not applicable
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
NSCLC	Non-small cell lung cancer
ORR	Objective response rate
OS	Overall survival
OS24	Overall survival after 24 months
PD	Progressed disease
PD-L1	Programmed death-ligand 1
PF	Progression-free
PFS	Progression-free survival
PFS12	Progression-free survival after 12 months
PFS18	Progression-free survival after 18 months
PFS2	Time from randomisation to second progression or death
PPS	Post-progression survival
PRESS	Peer review of electronic search strategies
PSA	Probabilistic sensitivity analyses
PSS	Personal Social Services
PSSRU	Personal Social Services Research Unit
Q2W	Every two weeks
QALY(s)	Quality-adjusted life year(s)
RCR	Royal College of Radiologists
RCT	Randomised controlled trial
RECIST	Response evaluation criteria in solid tumours
RT	Radiotherapy
SAP	Statistical analysis plan
SIGN	Scottish Intercollegiate Guidelines Network
SLR	Systematic literature review
SmPC	Summary of product characteristics
SoC	Standard of care
STA	Single technology appraisal
UK	United Kingdom
UMC	University Medical Centre
TC	Tumour cell
TFTS	Time to first or subsequent therapy or death
TNM	Tumour-node-metastasis

TSD	Technical support document
TSST	Time to second subsequent therapy or death
TTDM	Time to death or distant metastasis
TTF	Time to treatment failure
TTP	Time to progression
UK	United Kingdom
US	United States
WCLC	World Conference on Lung Cancer
WHO	World Health Organization

Table of Contents

Table of Tables

Table of Figures

1. Summary

1.1 Critique of the decision problem in the company's submission

The population defined in the company submission (CS) is adults with locally-advanced, unresectable, stage III non-small cell lung cancer (NSCLC) whose tumours express programmed death-ligand 1 (PD-L1) on ≥1% of tumour cells (TCs) and whose disease has not progressed following platinum-based chemoradiation therapy (CRT). Compared to the National Institute for Health and Care Excellence (NICE) scope, the population is narrower, i.e. only includes patients in the relevant population whose tumours expressed PD-L1.

The intervention (durvalumab 10mg/kg every two weeks intravenously), comparator (standard of care) and outcomes are defined in line with the NICE scope.

1.2 Summary of the key issues in the clinical effectiveness evidence

The CS comprised of a systematic review of the evidence for durvalumab for the treatment of locally advanced unresectable, stage III NSCLC in adults whose tumours express PD-L1 on ≥1% of tumour cells and whose disease has not progressed after platinum-based CRT. The CS and response to clarification provided sufficient details for the Evidence Review Group (ERG) to appraise the literature searches. A good range of databases and conference proceedings were searched. Of concern to the ERG was the restrictive population search, which combined NSCLC terms with disease stage and chemoradiation therapy search terms, and did not include intervention terms as an additional facet. However, this is unlikely to have greatly affected the recall of results.

The CS presented direct evidence from one randomised controlled trial (RCT), PACIFIC, which compared durvalumab to standard of care in adults with locally-advanced, unresectable, stage III NSCLC whose tumours express PD-L1 on ≥1% of TCs and whose disease has not progressed following platinum-based CRT. The population of participants receiving durvalumab represents approx. 67% of the overall population included in PACIFIC. It should be noted that randomisation was not stratified based on PD-L1 status. While reported baseline characteristics, such as age, histology, or smoking status, were balanced between the durvalumab and placebo groups, there are potential problems linked to overinterpretation of subgroup analyses which might impact on the findings.

The PACIFIC trial included only eight patients from the United Kingdom (UK). Another concern to the ERG was the applicability of durvalumab to a population receiving different types of CRT cycles. The CS notes that in the PACIFIC trial concurrent CRT was received prior to beginning treatment with durvalumab. However, the clinical expert highlighted that *"most UK patients receive sequential rather than overlapping treatment"* while *"the efficacy and safety of durvalumab in the PACIFIC study was evaluated after overlapping, rather than sequential, CRT"*. The response to request for clarification suggested the cohort in the PACIFIC trial is generalisable to UK patients with locally-advanced, unresectable, stage III, NSCLC. It also suggested that survival rates might be lower amongst patients treated with sequential CRT approaches than overlapping. However, more pertinently, as the company admitted in the response to clarification, the effectiveness of durvalumab in following sequential therapy remains unknown, i.e. *"…clinicians would expect to see some benefit of durvalumab treatment after sequential CRT, although the magnitude of this remains uncertain in the absence of robust clinical evidence"*. These issues impact on the certainty regarding these findings and might limit the applicability of any findings to UK clinical practice.

The CS reported a progression-free survival (PFS) benefit with durvalumab when compared to placebo in the PD-L1 <1% and PD-L1 ≥1%, and unknown PD-L1 expression groups. Patients in the PD-L1

≥1% and unknown expression groups receiving durvalumab observed an overall survival (OS) benefit. The CS also reported the statistically significant and clinically meaningful PFS and OS benefits in the PD-L1≥1% group. However, it should be noted that these results come from an interim cut-off, i.e. not from the final analysis. Durvalumab treated patients also observed statistically-significant improvements in key secondary endpoints when compared to placebo.

Based on the PACIFIC data there appears to be a benefit in both PFS and OS for durvalumab patients compared with placebo patients, however, the data are immature and there remains substantial uncertainty about the comparative effectiveness.

Common adverse events were reported in both the durvalumab and placebo groups. The common AEs in the durvalumab patients included cough, fatigue, and radiation pneumonitis, whereas patients in the placebo group also included dyspnoea. Overall, more serious adverse events were reported for durvalumab compared to placebo (64/213 (30%) vs. 18/90 (20%)).

1.3 Summary of the key issues in the cost effectiveness evidence

Individual searches were undertaken for economic, cost and resource use and health-related quality of life (HRQoL) evidence. The company submission and response to clarification provided sufficient details for the ERG to appraise the literature searches and the 2018 update searches. A good range of databases and additional resources were searched. None of the included cost effectiveness studies were conducted from the UK perspective.

The company submission was largely in line with the NICE reference case. The modelled population, however, was narrower than that in the scope, but in line with the anticipated marketing authorisation (focussing on the subgroup with PD-L1 tumour expression ≥1%).

The company developed a de novo semi-Markov cohort state transition model. The model comprised of three health states, i.e. progression-free (PF), progressed disease (PD) and death. The company considered these health states to capture the most important clinical aspects in the treatment of stage III NSCLC patients, namely the time spent in PF and the time spent alive. The company estimated PFS, time-to-progression (TTP) and post-progression survival (PPS) to inform transitions between health states. Given the immaturity of the survival data in the PACIFIC subpopulation, the ERG had concerns about the appropriateness of the semi-Markov approach and questioned its superiority over a partitioned survival model approach. Therefore, the ERG would have liked to see both approaches appropriately explored. The company claimed that the semi-Markov approach largely avoided crossing of PFS and OS curves. However, relying on PPS to estimate survival instead of using OS drew on even fewer patients for extrapolation and potentially introduced additional bias (selection bias by relying on early progressors, with more progressions in the placebo arm than in the durvalumab arm). The magnitude and direction of any bias are unclear.

In line with its anticipated marketing authorisation, durvalumab was considered in the cost effectiveness model for the treatment of locally-advanced, unresectable, stage III NSCLC patients whose tumours express PD-L1 on ≥1% of TCs and whose disease has not progressed after ≥2 cycles of platinum-based CRT. This was a subgroup from the final scope issued by NICE, which considered the same population regardless of their PD-L1 status. However, the generalisability of PACIFIC to the United Kingdom setting was questioned, because patients in PACIFIC largely received overlapping CRT, whilst sequential CRT is standard practice in the United Kingdom. The direction and magnitude of any potential bias stemming from this could not be assessed.

Durvalumab was considered within the economic evaluation as per the anticipated licensed indication in NSCLC. Durvalumab was, in line with the dosage used in PACIFIC, modelled with a posology of 10mg/kg administered as an intravenous infusion over 60 minutes every two weeks, until disease progression or unacceptable toxicity, or a maximum of 12 months. The comparator in the economic model was described as active follow-up or standard of care (SoC), which applies up to disease progression. The intervention was implemented as per its marketing authorisation and dosage.

The analysis took an NHS and Personal Social Services (PSS) perspective. Discount rates of 3.5% were applied to both costs and benefits. The model cycle length was two weeks in the first year and four weeks thereafter with a lifetime time horizon (40 years). A half-cycle correction was applied, except to treatment and treatment administration costs.

Perspective, time horizon and discounting were in line with the NICE reference case, however, in the absence of any justification for not applying the half-cycle correction to treatment and treatment administration costs, the ERG considered this inconsistent with the calculation of resource use and other model calculations, which lowered the ICER.

The main source of evidence on treatment effectiveness used for intervention and comparators was the PACIFIC study. Only data from the subgroup of PD-L1 ≥1% patients (according to the anticipated marketing authorisation) and from the March data cut were used in the model. The ERG had concerns about the model being based on treatment effectiveness estimates derived from a post-hoc subgroup analysis and post-hoc TTP and PPS analyses, as well as in a population with mostly prior overlapping CRT instead of sequential CRT, but any bias introduced by this remained unclear.

Parametric survival curves were fitted to patient level data from PACFIC data on PFS, but instead of using the OS data from PACIFIC, the company performed survival analyses on the outcomes TTP and PPS, as explained below. The probability of remaining in the progression free (PF) state was estimated using PFS data by fitting independent parametric survival models. Based on statistical goodness of fit, the generalised gamma was selected to model PFS for both durvalumab and placebo. The Gompertz distribution was used in scenario analysis and the log-normal distribution was not used, despite the log-normal making a better fit than the Gompertz in both arms. The main concern of the ERG was that it considered durvalumab PFS to be probably over-estimated in the model, due to a model choice (generalised gamma) that resulted in PFS estimates above those observed in PACIFIC at three years, with the PACIFIC estimate notably being derived from small numbers of patients at risk. This model choice probably caused ICERs to be lower than with other model choices. It is noteworthy that any modelling choice for modelling PFS is associated with high levels of uncertainty, given the immaturity of the data, and that different PFS model choices have a large impact on the ICERs. Alternative modelling methods, such as spline-based models, remained unexplored in the CS.

The PFS curve for durvalumab was altered in the long run to reflect a potential treatment waning effect caused by stopping treatment at a maximum of 12 months. From a chosen cut-off point, which was set to 10 years in the company's base-case, a hazard ratio of one was applied to the placebo curve to model durvalumab PFS. The ERG considers this choice of time-point as highly uncertain, not appropriately validated, and potentially late, further adding to the likely over-estimation of durvalumab PFS. Furthermore, the implementation of the treatment waning effect could cause counter-intuitive results.

The probability of patients moving from the PF state to the progressed disease (PD) health state was determined by survival analysis of TTP data (PFS data with deaths treated as censored) from PACIFIC. The generalised gamma distribution was chosen in the base-case, based on best statistical fit (Akaike

information criterion; AIC and Bayesian information criterion; BIC) and to align with extrapolation of PFS.

The probability of patients moving from PD to death was estimated using survival analysis of pooled PPS data from both treatment arms in PACIFIC (choice of exponential distribution based on best statistical fit). The effectiveness of subsequent treatments was captured in the PPS to the extent that patients in the PACIFIC study received subsequent treatments. In a scenario analysis, an alternative method for extrapolating PPS was used, where PPS was informed by published data from the KEYNOTE-024 study, data from the pembrolizumab arm used for those patients in PACIFIC who received immuno-oncological (IO) treatment, and data from the chemotherapy arm used for those not receiving IO treatment. The ERG noted the uncertainty in PPS introduced by immature PPS data from PACIFIC, uncertainty about subsequent treatments and potential bias in extrapolating PPS in the light of even smaller number of patients and immature data, rather than OS. Exploratory analyses showed that any impact of this on the ICER was probably relatively small, with the main treatment benefit of durvalumab extending PFS.

The main source of evidence on treatment adverse events used for durvalumab and SoC was the PACIFIC study. Adverse events (AEs) that were of grade 3/4 and had a frequency of ≥2% in either arm of the PACIFIC study were included in the model in terms of their costs and not their impact on HRQoL. AEs were modelled as a per-cycle occurrence while patients are on treatment. Whilst AEs causally related to treatment were mostly higher for the durvalumab arm than in the placebo arm in PACIFIC, incidence of AEs in the model between treatments was comparable. It was unclear how this discrepancy occurred, likely lowering ICERs of durvalumab versus SoC. Exploratory analyses however showed that any bias caused by this would be limited.

EQ-5D-5L data were collected in PACIFIC and mapped to 3L utility scores using the crosswalk mapping algorithm as per the NICE position statement. A mixed effects model with only progression as a covariate was used to estimate utility values for the PF (0.819) and PD (0.776) health states. The ERG considered utility values for both health states to be potentially over-estimated, being comparable to those in the general population and not adjusted by general population utility estimates. The high PF utility value produces lower ICERs for durvalumab, whilst the high PD utility value produces higher ICERs for durvalumab versus SoC. Although the mapped utility scores from PACIFIC were higher in the placebo arm as compared to the durvalumab arm at almost all measurement moments, treatment was found to be statistically insignificant in the mixed effects model and therefore, equal utilities were assumed for durvalumab and SoC. The ERG was concerned that by excluding treatment as a factor in the mixed effects model, and at the same time including disutilities of a limited set of AEs only in a sensitivity analysis, the true impact of treatment with durvalumab and adverse events was not appropriately captured in the model. The exclusion of treatment as a covariate in the utility mixed effects model resulted in lower ICERs. No adverse event related disutilities were taken into account.

Costs in the model included costs for PD-L1 testing, costs associated with treatment, costs associated with disease management and patient observation, and costs associated with end of life care. Unit costs were based on the National Health Service (NHS) reference costs, Personal Social Services Research Unit (PSSRU), Monthly Index of Medical Specialities (MIMS), and the electronic Market Information Tool (eMIT). Treatment cost per durvalumab infusion was calculated based on average body weight in PACIFIC, with treatment duration taken from PACIFIC Kaplan-Meier (KM) data. No drug wastage, i.e. perfect vial sharing, was assumed. The model assumed zero acquisition and administration costs for SoC. Once patients progressed in the model, a one-off cost for subsequent treatments was accrued. This cost was informed by the type of treatment, the required treatment dose, the dosing schedule, the unit

drug cost at list prices, and the duration of treatment. Resource use for the PF state was modelled in accordance with European Society for Medical Oncology guidelines, and resource use for the PD health state was derived from NICE Technology Appraisal 531 in the metastatic setting. The frequency of occurrence of included AEs was combined with a one-off cost per AE to obtain a total per-cycle cost for each arm. The ERG considered the assumption of perfect vial sharing to be unrealistic in this setting, given the limited number of patients in England and Wales that would be eligible for treatment with durvalumab. This assumption caused the ICER of durvalumab against SoC to be lowered.

Total deterministic life years (LYs) and quality-adjusted life years (QALYs) gained were larger in the durvalumab arm compared to the SoC arm. Incremental QALYs (2.93) were mainly driven by QALY gains in the PF health state. The revised (in response to clarification letter an error was corrected) deterministic incremental cost effectiveness ratio (ICER) amounted to £19,366 per QALY gained. Compared with the deterministic results, the probabilistic sensitivity analysis (PSA) with 1,000 iterations showed lower incremental QALYs and higher incremental costs, which resulted in an increased ICER (£21,601 per QALY gained). Some deterministic sensitivity analyses (DSA) and scenario analyses significantly affected the ICER.

At the clarification stage, the ERG identified several errors in the company's base-case and scenario analyses, including several settings in the controls sheet that were not functioning, and incorrect results of scenario analyses, which were corrected by the company. The ERG was still unable to reproduce one of the company's scenarios added in response to the clarification letter and found an error in another.

Face and internal validity checks were performed by the company and a third-party provider, as well as an expert in the field. Cross validity checks were not performed. OS predictions from the model were validated against PACIFIC, other sources and expert opinion. No firm conclusion could be drawn from the external validation exercise performed by the company using alternative data sources, due to differences in population.

1.4 Summary of the ERG's preferred assumptions and resulting ICER

The ERG made various adjustments to the company's base-case, including the fixing of errors, violations and amending the model according to its preferred assumptions (matters of judgement).

1.4.1 Fixing errors

1. Correction of age calculations
2. Correction of nivolumab and pembrolizumab vial sharing calculations
3. Correction of probabilistic utility decrements for progression and treatment

1.4.2 Fixing violations

4. Applying the half-cycle correction also to treatment and administration costs
5. Assumption of no vial sharing
6. Excluding patient characteristics from the PSA

1.4.3 Matters of judgment

7. Use of the lognormal instead of the generalised gamma distribution for modelling durvalumab PFS (and also TTP, as per company's default setting)
8. Treatment waning effect after five-year cut-off instead of 10-year cut-off
9. Applying an age-related utility decrement
10. Including treatment as a covariate in the utility mixed effects model

The probabilistic ERG base-case ICER of durvalumab versus SoC was £52,353 per QALY gained (based on 5,000 iterations). This was higher than the deterministic ERG base-case ICER of £50,238 per QALY gained. This difference was also observed in the company base-case results, and was likely caused by the skewedness of distributions used for modelling PFS.

Table 1.1: ICER resulting from ERG's preferred assumption

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
Deterministic ERG base-case					
Durvalumab	[redacted]	[redacted]	[redacted]	1.32	£50,238
SoC	[redacted]	[redacted]			
Probabilistic ERG base-case					
Durvalumab	[redacted]	[redacted]	[redacted]	1.25	£52,353
SoC	[redacted]	[redacted]			
ERG = Evidence Review Group = ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year; SoC = standard of care					

1.5 Summary of exploratory and sensitivity analyses undertaken by the ERG

Deterministic scenario analyses were performed to examine the potential impact of alternative assumptions on the cost effectiveness estimates. Alternative assumptions regarding PFS and treatment waning effects had the most significant impact. The scenario in which PFS distributions were changed into generalised gamma for both arms (as per the company's base-case) resulted in an ICER of £29,302 per QALY gained, whereas applying lognormal distributions for PFS in both arms drove up the ICER to £52,300 per QALY gained. Applying the company's KEYNOTE-024 PPS scenario, but with errors corrected, resulted in an ICER of £59,131 per QALY gained. The scenario exploring a treatment waning effect with three-year cut-off and using the lognormal distribution for both durvalumab and SoC PFS increased the ICER the most (to £64,531 per QALY gained). All other scenarios had a relatively modest impact (<£5,000) on the ERG base-case ICER.

In conclusion, given that the ERG base-case ICER was estimated to be substantially above £40,000 per QALY gained, and only one scenario resulting in ICERs slightly below £30,000 per QALY gained, and the large uncertainty induced by mainly the immature survival data, uncertainty around the cost effectiveness of durvalumab is substantial.

Table 1.2: Exploratory analyses undertaken by the ERG

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
ERG base-case					
Durvalumab	[redacted]	[redacted]	[redacted]	1.32	£50,238
SoC	[redacted]	[redacted]			
ERG base-case, no treatment waning effect (0)					
Durvalumab	[redacted]	[redacted]	[redacted]	1.10	£60,928
SoC	[redacted]	[redacted]			
Alternative PFS distributions both arms, generalised gamma (1)					
Durvalumab	[redacted]	[redacted]	[redacted]	2.19	£29,302

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
SoC	[redacted]	[redacted]			
Alternative PFS distributions both arms, lognormal (2)					
Durvalumab	[redacted]	[redacted]	[redacted]	1.27	£52,300
SoC	[redacted]	[redacted]			
Treatment waning at 3 years, PFS as ERG base-case (3a)					
Durvalumab	[redacted]	[redacted]	[redacted]	1.35	£48,766
SoC	[redacted]	[redacted]			
Treatment waning at 3 years, PFS as scenario 2 (3b)					
Durvalumab	[redacted]	[redacted]	[redacted]	1.04	£64,531
SoC	[redacted]	[redacted]			
Treatment waning at 7 years, PFS as ERG base-case (4a)					
Durvalumab	[redacted]	[redacted]	[redacted]	1.25	£52,833
SoC	[redacted]	[redacted]			
Treatment waning at 7 years, PFS as scenario 2 (4b)					
Durvalumab	[redacted]	[redacted]	[redacted]	1.41	£47,000
SoC	[redacted]	[redacted]			
PACIFIC PPS, but generalised gamma (5)					
Durvalumab	[redacted]	[redacted]	[redacted]	1.33	£49,868
SoC	[redacted]	[redacted]			
Company's KEYNOTE-024 PPS scenario, with errors corrected (6)					
Durvalumab	[redacted]	[redacted]	[redacted]	1.10	£59,131
SoC	[redacted]	[redacted]			
Adverse events with amended incidence and including impact on HRQoL (7)					
Durvalumab	[redacted]	[redacted]	[redacted]	1.32	£50,288
SoC	[redacted]	[redacted]			
Alternative PF utility score (8)					
Durvalumab	[redacted]	[redacted]	[redacted]	1.42	£46,539
SoC	[redacted]	[redacted]			
Alternative PF and PD utility scores (9)					
Durvalumab	[redacted]	[redacted]	[redacted]	1.28	£51,587
SoC	[redacted]	[redacted]			
Vial sharing possible at 30% (10)					
Durvalumab	[redacted]	[redacted]	[redacted]	1.32	£49,350
SoC	[redacted]	[redacted]			

ERG = Evidence Review Group; HRQoL = health-related quality of life; ICER = incremental cost effectiveness ratio; PD = progressed disease; PF = progression-free; PFS = progression-free survival; PPS = post-progression survival; QALY = quality-adjusted life year; SoC = standard of care

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

1.6.1 Strengths

The company's submission provided sufficient details for the ERG to appraise the database searches, which were generally transparent and reproducible. An adequate number of databases were searched and a good range of additional searches were conducted for grey literature.

Study design limits to identify economic evaluations, utility data, and costs and resource use data were applied. A good range of intervention terms for durvalumab and comparators were used in the cost effectiveness strategy.

The company submission was largely in line with the NICE reference case. Utility scores were estimated using a mixed effects model based on observed EQ-5D data in the PACIFIC study.

The model was, in general, well-built and transparent. Apart from their base-case, the company provided opportunities for exploratory analyses using alternative data derived from clinical trials in similar populations.

1.6.2 Weaknesses and areas of uncertainty

The population facet for each search conducted included a limited use of synonyms, and therefore may have missed relevant literature. Given the small number of references retrieved from the search, study design filters were not essential, and may have been unnecessarily restrictive.

The population included in the PACIFIC trial is narrower than in the NICE scope and the ERG identified additional issues which might potentially limit the applicability of study results, see Section 1.1.

A substantial source of uncertainty lies in the generalisability of PACIFIC data to the UK setting, as PACIFIC pertains predominantly to prior overlapping CRT, whereas in clinical practice in the UK, mostly sequential CRT is applied. In addition, the PD-L1≥1% subgroup and TTP and PPS analyses were performed post-hoc. Furthermore, main results come from interim analysis, e.g. according to the response to the request for clarification the company *"will conduct a final analysis of OS once a maturity of 70% has been reached. OS is an event-driven endpoint; therefore, the timing of this analysis is uncertain"* which will be reached *"when 491 OS events have occurred"*. The current maturity in the relevant subgroup is 33.0% for durvalumab and 49.5% for placebo.

A main limitation was the immaturity of survival data in the PACIFIC subpopulation, and the inherent uncertainty in PFS and PPS extrapolations. The ERG particularly considers durvalumab PFS to be overestimated, even more so because the company chose to incorporate treatment waning only at 10 years. Given the immaturity of survival data, the ERG also has concerns over the appropriateness of the semi-Markov model structure, but the company did not provide an opportunity to explore a partitioned survival approach. Alternative modelling methods, such as spline-based models, remained unexplored in the CS.

Lastly, the utility scores used in the model do not seem representative of the patient population. The ERG considers the utilities for both (progression-free and progressed disease) health states to be an overestimate.

2. Background

2.1 Introduction

In this report, the ERG provides a review of the evidence submitted by AstraZeneca in support of durvalumab, trade name IMFINZI™, for the treatment of adults with locally-advanced, unresectable, stage III NSCLC whose tumours express PD-L1 on ≥1% of TCs and whose disease has not progressed following platinum-based CRT.

2.2 Background and underlying health problem

In the CS,[1] the company emphasises the prevalence of lung cancer as being the third most common cancer in the UK.[2] Lung cancer was identified as being the main cause of cancer-related death.[3]

The company describes the progression of the stages of lung cancer, through the use of the Tumour-Node-Metastasis (TNM) system according to the American Joint Committee on Cancer (AJCC).[4] This system determines the overall cancer stage in accordance with the size of the primary tumour, the regional lymph node involvement, and the presence or absence of distant metastases. The company has made stage III NSCLC the focus of the submission due to the disease's representation of a highly-heterogeneous disease stage as well as stage III occurring before the progression to metastatic stages allowing for the treatment intent to be curative.[1] The company highlights the classification of the stages across a patient population in the UK, with 20% of patients in England and Wales having stage III at the time of diagnosis.[5]

The CS identifies the symptoms experienced by patients within Stage III as including a persistent or worsening cough, difficulty breathing, pain experienced while breathing, an altered voice, and chest pain.[1] However, this burden of symptoms increases once the disease progresses to the metastatic stages. This disease progression places patients outside of the time frame to be treated with curative intent. The CS states that the increased experience of a high symptom burden also places the patient in a position to experience a decrease in HRQoL, particularly once the patient progresses to stage IV.

The CS highlights the treatment pathways according to the NICE guidelines, with surgical measures, based on suitability and fitness, being the first choice.[6] However, these guidelines emphasise if the patient is suitable for surgery, neo-adjuvant chemotherapy is not recommended, unless for the purpose of a clinical trial.[1] The CS includes a comparison of patients with stage I and stage II NSCLC, and the patients with stage III NSCLC who receive treatment with curative intent. According to the National Lung Cancer Audit (NLCA), *"81% of patients diagnosed with stage I–II (…) and a World Health Organization (WHO) performance status (PS) of 0–2 received curative-intent treatment"*.[7] In results not specific to the UK, 68%-92% of patients with stage I and 53%-60% of patients with stage II remained alive at five years.[8] This differs by the different classifications of stage III patients of which, if identified as either as stage IIIA or stage IIIB, 40% and 16% received treatment with curative intent, respectively. The CS states that 13% of Stage III patients in England and Wales had surgery.[7]

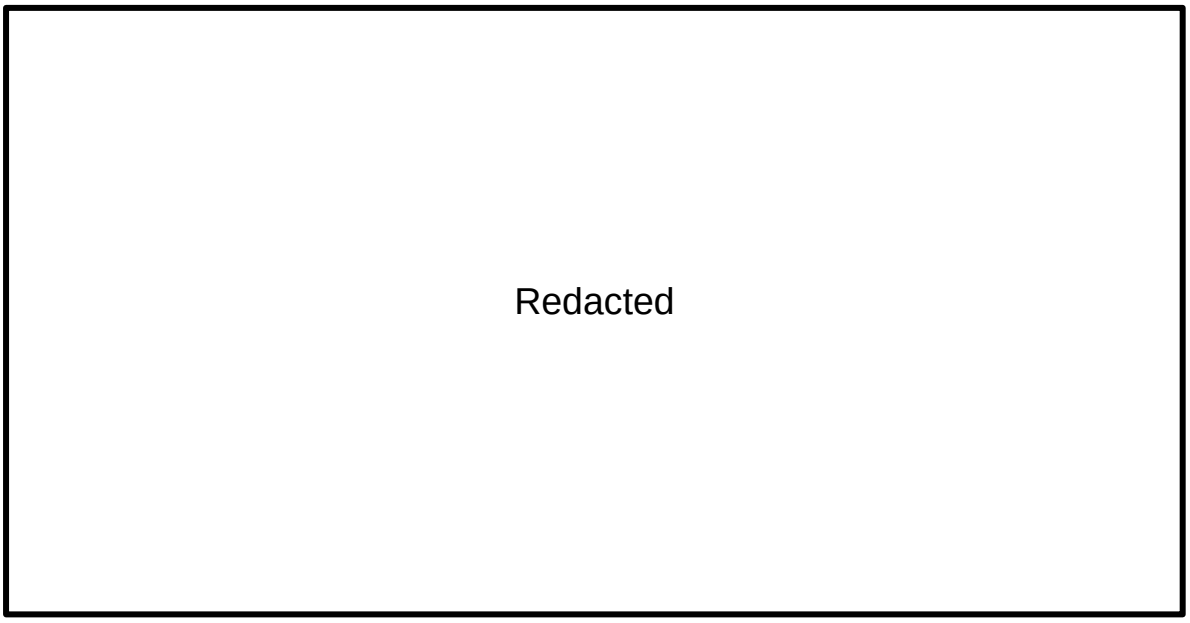
If surgery is not feasible for Stage III NSCLC patients, CRT is the standard of care.[6] Combinations of radiotherapy and chemotherapy provide better outcomes relative to radiotherapy alone.[9, 10] The CS notes that national guidelines, such as the Royal College of Radiologists and the British Thoracic Society, and regional guidelines, such as the London Cancer Alliance, which are used for the support of Stage III NSCLC patients in the UK, are in agreement with guidelines from NICE and the European Society for Medical Oncology (ESMO).[11-13] However, no new treatments have been approved for unresectable stage III NSCLC patients.[6, 14] This allows, as the company highlights, for active surveillance and best supportive care (BSC) to take place.[1] The CS emphasises that in the absence of active treatments, most

unresectable stage III NSCLC patients will experience disease progression following the completion of CRT. The CS identified several targeted therapies that have been evaluated as part of consolidation or maintenance upon completion of CRT. However, these were found to have either a moderate efficacy, while others were deemed unacceptable for integration.[15-17] During this absence of effective active treatments, disease progression within a year was experienced between 59.6% and 62.1% of Stage III patients upon completion of CRT.[18]

The CS reports that one-third of patients develop brain metastases, which can then result in poor outcomes with patients having a median overall survival (OS) of roughly four months.[19] If patients remain disease-free for a period of >12 months, they are treated with first-line (1-L) systemic drug therapies, otherwise second-line (2-L) drug therapies are utilised. Upon receiving insights from clinical experts in the UK, it is determined about 18% of Stage III patients receive further therapy after CRT. The UK clinical experts also revealed roughly 7% of Stage III patients are treated with a targeted therapy, while nearly 30% of patients receive an anti-PD-1/PD-L1 therapy.[20] The CS emphasises that during the metastatic stages the current treatment intent is palliative which identifies an unmet need for a curative treatment strategy that promotes the initial benefits achieved from CRT.[1]

The CS highlights the use of anti-PD-1/PD-L1 immunotherapy between the completion of CRT and before disease progression, which allows for *"T-cells to be reinvigorated at a time when the volume of tumour burden is low"*.[1] Anti-PD-1/PD-L1 antibodies have been reported to augment the stimulation of the immune effects from radiotherapy, resulting in an improvement of disease control.[21]

Figure 2.1: Treatment of stage III NSCLC



Source: Based on Figure 5 of the CS[1]
Footnote: ** Assumes that 95% of patients will not have experienced disease progression within six weeks or 42 days of completing CRT.
BSC = best supportive care; CRT = chemoradiation therapy; CTx = chemotherapy; NSCLC = non-small cell lung cancer; RT = radiotherapy

ERG comment: The ERG has no specific comments on the background presented in the CS.

However, it is noteworthy that no active treatment has been approved for patients with unresectable stage III NSCLC, providing a justification for the use of standard of care as comparator. The ERG also wants to direct attention to the relatively small patient population considered appropriate for treatment with durvalumab, as shown in Figure 2.1.

3. Critique of company's definition of decision problem

Table 3.1: Statement of the decision problem (as presented by the company)

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
Population	Adults with locally advanced, unresectable non-small cell lung cancer (NSCLC) whose disease has not progressed after platinum-based chemoradiation therapy (CRT)	Adults with locally-advanced, unresectable, stage III NSCLC whose tumours express PD-L1 on ≥1% of tumour cells (TCs) and whose disease has not progressed following platinum-based CRT	The submission will focus on locally advanced (stage III), unresectable NSCLC patients, whose tumours express PD-L1 on ≥ 1% of TCs, to reflect the opinion adopted by the Committee for Medicinal Products for Human Use (CHMP) of the European Medicines Agency (EMA)[§], and the anticipated Marketing Authorisation for durvalumab in this indication
Intervention	Durvalumab	Durvalumab (10 mg/kg every two weeks [Q2W] via intravenous [IV] infusion)	N/A
Comparator(s)	Best supportive care	Best supportive care (referred to as "active follow-up" throughout)	N/A

Outcomes	• Overall survival (OS)
	• Progression-free survival (PFS)
	• Response rates
	• Health-related quality of life (HRQL)
	• Adverse effects of treatment
		• PFS (primary endpoint)
		  • Secondary endpoints: proportion of patients alive and progression free at 12 and 18 months (PFS12 and PFS18)
		  • Supportive summary analysis: time to first subsequent therapy or death (TFST)
		• PFS2[*]
		  • Supportive summary analysis: time to second subsequent therapy or death (TSST)
		• Post-progression survival (PPS; *post-hoc* analysis)
			Time from randomisation to second progression or death (PFS2) and time to death or distant metastasis (TTDM) endpoints are relevant given the earlier disease setting (stage III) relative to previous immunotherapy appraisals in NSCLC (stage IV metastatic setting). They provide important information about the benefits of treatment beyond delaying disease progression:
			• PFS2 is an intermediate endpoint between PFS and OS and reflects real-life treatment decisions and patient experience. Its use is recommended by the EMA to capture potential negative impacts on next-line therapy and to demonstrate that any potential tolerability concerns are outweighed by treatment benefit.[22]

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope

		• OS (primary endpoint)
		  • Secondary analysis: proportion of patients alive at 24 months (OS24)
		  • *Post-hoc* analysis: impact of subsequent immunotherapy use
		• Response rates
		• TTDM*
		• HRQL (EORTC QLQ-C30 and EORTC QLQ-LC13)
		• Adverse effects of treatment
			• TTDM captures the value of maintaining local control and delaying progression to more-advanced metastatic disease stage

Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. 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. Costs will be considered from a National Health Service (NHS) and Personal Social Services perspective.	As per National Institute for Health and Care Excellence (NICE) reference case. A lifetime time horizon is appropriate in this setting to capture all differences in costs or outcomes between the technologies being compared.	N/A

Source: Based on Table 1 of the CS[1]

CHMP = Committee for Medicinal Products for Human Use; CRT = chemoradiation therapy; CS = company submission; EMA = European Medicines Agency; EORTC = European Organisation for Research and Treatment of Cancer; EORTC QLQ-C30 = European Organisation for Research and Treatment of Cancer 30-item core quality of life questionnaire; EORTC QLQ-LC13 = European Organisation for Research and Treatment of Cancer quality of life questionnaire and lung cancer module; HRQL =

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
health-related quality of life; IV = intravenous; N/A = not applicable; NHS = National Health Service; NICE = National Institute for Health and Care Excellence; NSCLC = non-small cell lung cancer; OS = overall survival; OS24 = overall survival after 24 months; PD-L1 = programmed cell death-ligand 1; PFS = progression-free survival; PFS12 = proportion of patients alive and progression free at 12 months; PFS18 = proportion of patients alive and progression free at 18 months; PFS2 = time from randomisation to second progression or death; PPS = post-progression survival; Q2W = every two weeks; TC = tumour cell; TFST = time to first or subsequent therapy or death; TSST = time to second subsequent therapy or death; TTDM = time to death or distant metastasis Footnotes: * Different from draft scope; § On 26 July 2018, the CHMP adopted a positive opinion, recommending the granting of a marketing authorisation for the medicinal product durvalumab (IMFINZI™) as monotherapy for the treatment of locally advanced, unresectable NSCLC in adults whose tumours express PD-L1 on ≥ 1% of TCs and whose disease has not progressed following platinum-based CRT.[23]			

3.1 Population

The ERG identified three issues which might limit the applicability of any findings presented in the CS:

- The population defined in the CS is adults with locally-advanced, unresectable, stage III NSCLC whose tumours express PD-L1 on ≥1% of TCs and whose disease has not progressed following platinum-based CRT. Compared to the NICE scope, this definition is narrower due to the incorporation of opinions expressed by the CHMP and the EMA, i.e. included patients in the relevant population whose tumours expressed PD-L1.
- As detailed in Section B.2.3 of the CS, it is important to mention that only eight UK patients of the PACIFIC trial (the main trial identified for clinical effectiveness) were included in the trial and according to the response to request for clarification *"it was not considered appropriate to present analyses where there were <20 events in a subgroup, as this sample size is too small for meaningful analyses / interpretation of data"*.[1, 24] Therefore, outcomes data on these eight UK patients in PACIFIC were not analysed separately. As stated in clarification letter, [redacted], however, analyses of these data are not available at the moment.[24]
- Clinical expert Dr Susan Harden stated that *"most UK patients receive sequential rather than overlapping treatment"* while *"the efficacy and safety of durvalumab in the PACIFIC study was evaluated after overlapping, rather than sequential, CRT"*.[24] This issue is discussed in Section B.1.3 of the CS.[1]

3.2 Intervention

The intervention (durvalumab 10mg/kg Q2W via IV infusion is in line with the scope. However, concomitant treatments were used in the PACIFIC trial. This issue is addressed in Section B.1.1 of the CS.[1]

In July 2018, the CHMP recommended the granting of a marketing authorisation for the medical product IMFINZI™.[23] The final summary of product characteristics (SmPC) and European public assessment report (EPAR) are not available at the present time (October 2018).

3.3 Comparators

The NICE scope listed only one comparator, namely best supportive care (BSC). In the CS BSC was also defined as *"active follow-up"* and *"standard-of-care"*. These definitions were used interchangeably. Since there are no active treatment options after CRT in unresectable Stage III patients whose disease has not progressed, the comparator described in the company's clarification letter as *"surveillance every six months for two years with a visit including history, physical examination and—preferably contrast-enhanced—volume chest CT scan at least at 12 and 24 months is recommended, and thereafter an annual visit including history, physical examination and chest CT scan in order to detect second primary tumours"* match the comparator described in the final scope.[24]

3.4 Outcomes

All of the outcomes defined in the NICE scope have been addressed in the CS.

Several measures have been included for PFS and HRQoL, as detailed in Table 3.1. Furthermore, an additional outcome, TTDM, was included.

3.5 Other relevant factors

The company describes the economic analysis as per the NICE reference case. However, the company also describe a lifetime time horizon as being appropriate for the setting.

Durvalumab is available in the UK under an Early Access Program (EAP).

4. Clinical effectiveness

4.1 Critique of the methods of the review(s)

The systematic literature review in the CS, which was used to find clinical trial data on the efficacy and safety of durvalumab when compared to active follow-up in locally-advanced, unresectable, stage III NSCLC in patients whose disease has not progressed upon completion of CRT, identified only one trial, the PACIFIC study.

4.1.1 Searches

The following paragraphs contain summaries and critiques of the searches related to clinical effectiveness presented in the company submission. The Canadian Agency for Drugs and Technologies in Health (CADTH) evidence-based checklist for the Peer Review of Electronic Search Strategies (PRESS), was used to inform this critique.[25] The submission was checked against the single technology appraisal (STA) template for company/sponsor submission of evidence.[26] The ERG has presented only the major limitations of the search strategies in the main report. Further minor comments can be found in Appendix 1.

Appendix D.1 of the CS states that MEDLINE, MEDLINE in Process, Embase, Cochrane Central Register of Controlled Trials, CDSR and DARE were searched for the identification of published clinical trial data on the efficacy and safety of durvalumab.[27] The search strategy was reported in detail in Appendix D.1.[27] Searches were conducted on 24 January 2018 using the OvidSP interface from 2002, and limited to English language studies only. Results were limited to RCTs, using search terms based on the Scottish Intercollegiate Guidelines Network (SIGN) RCT search filters.

Searches were conducted and reported for conference proceedings from 2014-2017 for the following conferences: American Society of Clinical Oncology (ASCO), ESMO, European Lung Cancer Conferences (ELCC), World Conference on Lung Cancer (WCLC) and American Association for Cancer Research (AACR).

No additional search methods, such as clinical trials register searches, handsearching or reference checking were reported.

ERG comment:

- The selection of databases searched was adequate, and searches were clearly reported. The database name, host, date range and date searched were provided.
- In response to clarification the company confirmed that a single search was conducted across MEDLINE, MEDLINE in Process, Embase, Cochrane Central Register of Controlled Trials, CDSR and DARE, using the OvidSP platform. This approach has limitations when using subject heading terms which could affect recall of results. While the ERG noted the inclusion of separate trials filters designed specifically for MEDLINE and Embase, only MEDLINE subject heading terms (MeSH) were used in the population facet of the search strategy. Although simultaneous searching of Embase should automatically identify and search for equivalent Embase subject heading terms (Emtree), it is not clear if this is the case for all potentially useful Emtree terms. Given the possible limitations of this approach, the ERG considered it preferable to search each database separately, or at least to ensure inclusion of both Emtree and MeSH terms in all facets of the search strategy. Reporting individual searches is also good practice in order to clarify the numbers identified on each database.

- Of concern to the ERG was that the search terms used for the population facet of the strategy were limited. The strategy combined NSCLC search terms with both disease stage and chemoradiation therapy (CRT) search terms, resulting in a very focussed strategy which may have missed relevant studies. Although the population is clearly defined in the scope, it is not possible to be sure that the search terms in the strategy will necessarily be included in the title and/or abstract of relevant references. In addition, only one MeSH term was used for NSCLC, and few synonyms were used for disease stage or CRT. Additional synonyms and subject heading terms could have been added to the strategy for NSCLC, disease stage and CRT, and use of these terms could have increased the retrieval of potentially relevant records.
- Durvalumab and comparator terms were not included in the database search strategy, although they were included in the conference searches. In response to clarification, the company stated that *"Durvalumab and comparator terms were not included in the database search strategy because we wanted to capture all possible treatments investigated in the post-chemoradiation therapy (CRT) setting"*.[24] The ERG believed that the addition of intervention and comparator terms to the database strategy as a separate facet (i.e. not combined with the other elements of the search) could have broadened the search to identify other potentially relevant studies. Given the company's awareness of relevant literature in the field and additional search methods however, this is unlikely to have greatly affected the recall of results.
- The trials filter used in the search of all included databases was unnecessary for the search of CENTRAL which contains only controlled trials. For the searches of CDSR and DARE, the trials filters will have removed all records, as these databases contain only systematic reviews. The use of a trials filter for these databases therefore risks removing potentially relevant records.
- The ERG was concerned that limiting the clinical effectiveness searches to English language may have introduced potential language bias. Current best practice states that 'Whenever possible review authors should attempt to identify and assess for eligibility all possibly relevant reports of trials irrespective of language of publication'.[28]
- Additional search methods, such as clinical trials register searches, handsearching or reference checking might have been useful to identify additional relevant studies and grey literature.

4.1.2 Inclusion criteria

The CS provided a table illustrating the inclusion and exclusion criteria for the systematic review in order to ensure decisions were consistent (Table 4.1). The inclusion screening made distinctions between level 1 (primary) and level 2 (secondary) screening. Level 1 screening utilised a broad set of inclusion criteria in order to identify trials in which at least one CRT regimen was concurrent in unresectable, stage III NSCLC patients. During the level 2 screening, the definition used for level 1 is expanded upon to include the comparison of the outcomes of durvalumab and active follow-up, BSC, or observation. After applying the criteria, one RCT was found to be appropriate for inclusion in the systematic review. However, in the CS, the company appeared to use terms such as "best supportive care," "active follow-up," "standard of care," "placebo," and "active surveillance" interchangeably. The company amends this in their response to clarification by indicating the terms "active follow-up" and "standard of care" were meant to be used interchangeably in the CS, whereas the term "placebo" was used to refer to the control arm in the PACIFIC trial, see Table 4.2.[24]

In the CS, the outcomes used in the PACIFIC trial, time to progression (TTP) and post progression survival (PPS), were not pre-specified.[1] Upon response for clarification, the company defined TTP as the time from randomisation until the date of the first objective disease progression.[24] The company elaborates further by indicating the use of the TTP definition in this manner was consistent with to the

definition in the EMA guideline.[22] In the response for clarification, the company defines PPS as the time from objective disease progression until censoring or death due to any cause.[24] However, due to PPS not being used in regulatory approvals, there is no definition available from the EMA.

Table 4.1: Detailed inclusion/exclusion criteria for the systematic literature review

Criteria	Inclusion	Exclusion
Population	**Level 1 screening** Unresectable stage III NSCLC (≥80% of the trial population) **Level 2 screening** Unresectable stage III NSCLC patients whose disease has not progressed after completing CRT	Patient populations that do not meet the adjacent inclusion criteria (Note: clinical trials that investigated the efficacy and safety of CRT regimens in unresectable stage III NSCLC patients were initially included [at level 1] for full-text review, to ensure no relevant publications were incorrectly discarded; see **Error! Reference source not found.**)
Interventions and comparators	**Level 1 screening** CRT, including cisplatin or carboplatin in combination with: etoposide, vinblastine, vinorelbine, paclitaxel, docetaxel, or pemetrexed **Level 2 screening** CRT, as per above, followed by either durvalumab or observation / BSC only	Studies that do not meet inclusion criteria specified at each level of screening
Outcomes	**Level 1 and 2 screening:** Overall survival (OS); including hazard ratio, median, landmark survival rates PFS; including hazard ratio, median, landmark survival rates, time to progression (TTP) based on criteria reported in the relevant publication Time to death or distant metastasis (TTDM) Time to treatment failure (TTF) Time to disease progression or death on subsequent therapy (PFS2) Objective response rate (ORR), disease control rate (DCR), duration of response (DoR) based on criteria reported in the relevant publication Overall treatment discontinuation and discontinuation due to efficacy and safety reasons, respectively	Studies that do not report on any of the outcomes listed in the adjacent inclusion criteria

Criteria	Inclusion	Exclusion
	Rates of overall and treatment related grade 3–5 adverse events (AEs)	
Study design	**Level 1 screening** Clinical trials evaluating two or more CRT regimens, involving at least one concurrent regimen (e.g. head-to-head trials of concurrent CRT regimens or comparisons of concurrent and sequential protocols)* CTx used in CRT regimens were as per the inclusion criteria listed in the "intervention and comparators" section Clinical trials evaluating post-CRT maintenance / consolidation therapies **Level 2 screening** Clinical trials evaluating durvalumab or observation / BSC in unresectable stage III NSCLC patients whose disease has not progressed after completing CRT Outcomes should have been measured from randomisation (following confirmation of response / stable disease after concurrent CRT)	**Level 1 screening** Clinical trials that included CTx regimens not specified in the "intervention and comparators" section Observational studies Cases reports or editorial comments Note: studies that evaluated concurrent CRT regimens were initially included for full-text review, to ensure no relevant articles were incorrectly discarded **Level 2 screening** Clinical trials that did not meet the specified level 2 inclusion criteria Note: clinical trials where it was not possible to evaluate outcomes of interest from randomisation to durvalumab or BSC were also excluded (e.g. clinical trials that reported outcomes from initiation of CRT)
Language	Abstracts and / or full-text articles published in English	References published in any language other than English
Countries of interest	No restriction	No restriction
Date	2002 to January 24, 2018	References published outside of this date limit

Source: Table 3 of the CS appendices[27]

BSC = best supportive care; CRT = chemoradiation therapy; CS = company submission; CTx = chemotherapy; DCR = disease control rate; DoR = duration of response; NSCLC = non-small cell lung cancer; PFS = progression-free survival; PFS2 = time to disease progression or death on subsequent therapy; ORR = objective response rate; OS = overall survival; SLR = systematic literature review; TTDM = time to death or distant metastasis; TTF = time to treatment failure; TTP = time to progression

Footnotes: * trials involving at least one concurrent CRT regimen were selected to align with the SoC in unresectable Stage III NSCLC setting and the PACIFIC study population, which only included patients who had not experienced disease progression after ≥2 cycles of overlapping (i.e. concurrent) CRT

Table 4.2: Comparator terminology

Comparator Terms	Company Definitions
"Active follow-up" or "Standard-of-care"	Includes surveillance visits, history, physical examination, and chest CTs every six months for the first two years and annually thereafter to detect second primary tumours. Terms are used interchangeably.
Placebo	Refers to the control arm of the PACIFIC clinical trial or other trials.
Source: Based on response to request for clarification[24] CT = computed tomography	

ERG comment: The definition of PPS is similar to the overall survival (OS) endpoint, except it is calculated from the point of first objective disease progression, not randomisation.

4.1.3 Critique of data extraction

According to the appendices of the CS, data extraction was restricted to full publications and health technology assessments (HTAs) that were conducted from a UK perspective.[27] This resulted in one full publication and 20 HTAs being considered for data extraction. The studies selected for data extraction, were assessed by two reviewers to determine if pre-defined inclusion/exclusion criteria were met. A third-party member was involved in order to resolve any discrepancies. The data extraction was checked by a second reviewer in order to identify any inconsistencies.

ERG comment: The ERG has no further comments on this matter.

4.1.4 Quality assessment

The quality of the PACIFIC study was assessed by the company and presented in the appendices of the CS.[27] The elements that were considered in the quality assessment were appropriate randomisation, adequate concealed treatment allocation, the presence of unexpected imbalances in drop-outs between groups, any evidence suggesting the authors measured more outcomes than they reported, the inclusion of an appropriate intention-to-treat analysis, and the use of appropriate methods to account for missing data. Table 4.3 provides an overview of the quality assessment of the PACIFIC study.

Table 4.3: Quality assessment results for PACIFIC

Study question	How is the question addressed in the study?	Risk of bias
Was randomisation carried out appropriately?	Yes Treatments were assigned using the randomisation scheme in the IVRS / IWRS. One randomisation list was produced for each of the randomisation strata. A blocked randomisation was generated, and all study centres used the same list to minimise any imbalance in the number of patients assigned to each treatment group.	Low

Study question	How is the question addressed in the study?	Risk of bias
Was the concealment of treatment allocation adequate?	Yes The PACIFIC study was conducted in a double-blind manner. The reconstituted durvalumab solution and its matching placebo were identical in colour; IV bags used for administration were identical in size. The study drug was blinded using an opaque sleeve, fastened with tamper-evident tape over the IV bag.	Low*
Were the groups similar at the outset of the study in terms of prognostic factors?	Yes Patients were stratified at randomisation based on their age (<65 versus ≥65 years), gender, and smoking history (current or former smoker versus never smoked). Patients randomised to durvalumab and placebo groups were well balanced in terms of demographics, baseline disease characteristics (including PD-L1 expression and *EGFR* mutational status), and prior anti-cancer therapy (including best response to previous concurrent CRT).	Low
Were the care providers, participants and outcome assessors blind to treatment allocation?	Yes The PACIFIC study was conducted in a double-blind manner. The patient, the Investigator and study centre staff were blinded to study drug allocation. Only the study centre pharmacist was unblinded and prepared the durvalumab infusion or placebo for a patient, as specified by the randomisation scheme and IVRS. No member of the extended study team at AstraZeneca/MedImmune, at the investigational centres, or any Contract Research Organisation handling data had access to the randomisation scheme until the time of the final data analysis (exceptions noted in the Clinical Study Protocol[29]. Investigators were only unblinded to treatment allocation in cases of medical emergency. Note: the IDMC were provided with unblinded data for their review but AstraZeneca/MedImmune and Quintiles staff and Investigators involved in the study remained blinded.	Low

Study question	How is the question addressed in the study?	Risk of bias
Were there any unexpected imbalances in drop-outs between groups?	No At the most-recent data cut-off (interim OS analysis), 22 patients (4.6%) in the durvalumab group and 14 patients (5.9%) in the placebo group (had terminated the study by choice. One patient in the durvalumab group and no patients in the placebo group were lost to follow-up. The primary reason for study termination was death The number and reasons for discontinuations from treatment did not raise any concerns about the conduct of the study. More patients in the placebo group discontinued treatment due to worsening of the condition under investigation (49.6%, versus 31.3% in the durvalumab group), as expected given the study hypothesis.	Low
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No Full documentation relating to the PACIFIC clinical trial methodology, analyses, and outcomes are included in the CS	Low
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes Efficacy and HRQL analyses were performed on the ITT population; standard censoring methods used to account for missing data. Note: safety analyses were performed on the Safety Analysis Set, which included all patients all patients who received at least one dose of randomised study drug and for whom any post-dose data were available	Low

Source: Based on Table 7 of the CS appendices[27]
CRT = chemoradiation therapy; CS = company submission; EGFR = epidermal growth factor receptor; HRQL = health-related quality of life; IDMC = independent data monitoring committee; ITT = intention-to-treat; IV = intravenous; IVRS = interactive voice response system; IWRS = interactive web response system; OS = overall survival; PD-L1 = programmed death ligand 1

ERG comment: In the quality assessment of the PACIFIC trial, presented in Table 7 of the CS appendices, the company identifies the PACIFIC trial as having a low risk of bias for concealment of treatment allocation.[27] The response does not describe how concealment of allocation was concealed, i.e. this question should be rated as unclear. However, describing the randomisation, the company describes that IVRS/IWRS were used which are acceptable methods of concealment of allocation.

4.1.5 Evidence synthesis

The analysis utilised in the CS was done in accordance with a comprehensive Statistical Analysis Plan (SAP). Three interim analyses were utilised.

Only one RCT, PACIFIC, was identified. Therefore, no evidence synthesis was done.

4.2 Critique of trials of the technology of interest, the company's analysis and interpretation (and any standard meta-analyses of these)

4.2.1 Overview of the direct evidence in the submission

The CS states that the phase III PACIFIC RCT is the only study in which a direct comparison was made focusing on the clinical effectiveness of durvalumab 10mg every two week (Q2W) versus active follow-up in locally-advanced, unresectable Stage III NSCLC patients whose disease has not progressed following CRT.

The data supporting this submission is from the PACIFIC study, which is a randomised, double-blind, placebo-controlled, multicentre, international study. The main features of the PACIFIC study are summarised in Table 4.4.

The CS noted that most participants in the PACIFIC trial received two or more overlapping, or concurrent, cycles of CRT.[1] However, according to the clinical expert cited in the CS, sequential CRT is the method of treatment most often received for patients in the UK and is identified as the standard of care. While the company acknowledges this difference, they state the PACIFIC patient population is broadly general to UK patients with locally-advanced, unresectable, stage III NSCLC patients who receive curative-intent CRT treatment.

Table 4.4: Quality assessment results for PACIFIC

Trial name	**PACIFIC trial**
Population	Patients with locally-advanced, unresectable, Stage III NSCLC whose disease has not progressed following two or more overlapping cycles of definitive, platinum-based CRT.
Intervention	Durvalumab (n=476)
Comparator	Placebo (n=237)
Outcomes	• PFS* -PFS12, PFS18, TFST • OS* -OS24 • Adverse effects of treatment* • Response rates • PPS* • HRQL -EQ-5D* -EORTC • Time to treatment discontinuation* • TTDM • PFS2* -TSST
Study design	PACIFIC is an ongoing, randomised, double-blind, placebo-controlled, multi-centre, international, phase III study.
Duration of trial and trial phases	Randomisation completed as late as 42 days after last radiation dose.
	Durvalumab 10 mg/kg Q2W and Placebo Q2W received for up to 12 months.
	Re-treatment for patients who experienced disease control at the end of 12 months of treatment but progressed during follow-up.

Trial name	PACIFIC trial
Settings and locations where the data were collected	235 study centres in in 26 countries: Australia, Belgium, Canada, Chile, France, Germany, Greece, Hungary, Israel, Italy, Japan, Mexico, Netherlands, Peru, Poland, Singapore, Slovakia, South Africa, South Korea, Spain, Taiwan, Thailand, Turkey, UK, United States (US), and Vietnam.
Source: Table 3 and Figure 6 of the CS[1] Footnote: * included in economic model CRT = chemoradiation therapy; CS = company submission; EORTC = European Organisation for Research and Treatment of Cancer; EQ-5D = European Quality of Life-5 dimensions; HRQL = health-related quality of life; OS = overall survival; OS24 = proportion of patients alive at 24 months; PFS = progression free survival; PFS12 = proportion of patients alive and progression free at 12 months; PFS18 = Proportion of patients alive and progression free at 18 months; PFS2 = time to second progression or death; PPS = post-progression survival; Q2W= every 2 weeks; TFST = time to first or subsequent therapy or death; TSST = time to second subsequent therapy or death; TTDM = time to death or distant metastasis; US = United States	

ERG comment: According to the response to request for clarification, *"the efficacy and safety of durvalumab in locally-advanced, unresectable, Stage III NSCLC whose disease had not progressed following sequential CRT protocols was not investigated in the PACIFIC study, and as such, robust evidence from randomised clinical trial(s) is missing in this setting"*.[24] Therefore, most participants in the PACIFIC trial received two or more concurrent cycles of CRT. However, most UK patients, according to a clinical expert, receive sequential cycles of CRT which was not included in the evaluation of durvalumab.[1] Survival rates are lower amongst patients treated with sequential CRT approaches than overlapping and this should be also taken into consideration.[24]

In the response to request for clarification, the company provided some information about clinical experts' thoughts on rationale for using durvalumab after sequential CRT.[24] Given pre-clinical data, *"clinicians would expect to see some benefits of durvalumab treatment after sequential CRT, although the magnitude of this remains uncertain in the absence of robust clinical evidence"*.[24]

4.2.2 Participants in the PACIFIC trial

In the PACIFIC study, in order to be included patients had to be adults who had histologically- or cytologically-confirmed unresectable Stage III NSCLC. The patients also had to receive at least two overlapping cycles of CRT without disease progression upon completion. In order to be included in the PACIFIC study, the last received radiation dose had to have been completed 42 days prior to the first dose of study treatment. Further inclusion criteria included the patients to have had an estimated life expectancy. Nine hundred and eighty-three patients were enrolled from 235 centres, of which 713 ITT patients were randomised to receive either durvalumab or placebo. Of the 713 ITT patients, 76.4% had biopsies available for PD-L1 analysis, which was later determined to be 303 patients had ≥1% PD-L1 expression. The table below indicates the demographics of the patients included in the PACIFIC study.

The mean age of participants in the PACIFIC study in both the durvalumab ITT group and the durvalumab PD-L1≥1% groups was 63.0 years. In the PD-L1≥1% group was comprised of 67.9% males and 32.1% females, whereas the placebo group was comprised of 71.4% males and 28.6% females. In the PD-L1≥1% group, 68.9% of the group were identified as being white, whereas in the placebo group, 65.9% identified as being white. The durvalumab and placebo groups within the identified PD-L1≥1% group appeared well balanced among the factor focusing on smoking status. The durvalumab group had 18.4% who identified as being current smokers, 72.2% identified as being former smokers, and 9.4% identified as never smoked before. The placebo group had 14.3% who identified as being current

smokers, 78.0% identified as being former smokers, and 7.7% had never smoked before. Of both the placebo and durvalumab group 99.7% received chemotherapy concurrent with radiotherapy, see Table 4.5.

Table 4.5: Patient demographics, baseline disease characteristics, and prior anti-cancer therapies

Characteristic	ITT			PD-L1 ≥1% group		
	Durvalumab (n=476)	Placebo (n=237)	Total (n=713)	Durvalumab (n=212)	Placebo (n=91)	Total (n=303)
Demographics						
Age, mean (SD)	63.0 (8.7)	62.6 (9.6)	62.9 (9.0)	63.0 (8.4)	63.1 (8.8)	63.1 (8.5)
Age, median (range) [years]	64 (31−84)	64 (23−90)	64 (23−90)	64 (36−83)	64 (41−90)	64 (36−90)
Age groups (years), n (%)						
<50	30 (6.3)	22 (9.3)	52 (7.3)	12 (5.7)	6 (6.6)	18 (5.9)
≥50−<65	231 (48.5)	108	339	104 (49.1)	45 (49.5)	149 (49.2)
≥65−<75	178 (37.4)	88 (37.1)	266	81 (38.2)	34 (37.4)	115 (38.0)
≥75	37 (7.8)	19 (8.0)	56 (7.9)	15 (7.1)	6 (6.6)	21 (6.9)
Sex, n (%)						
Male	334 (70.2)	166	500	144 (67.9)	65 (71.4)	209 (69.0)
Female	142 (29.8)	71 (30.0)	213	68 (32.1)	26 (28.6)	94 (31.0)
Race **Race, n (%)**						
White	337 (70.8)	157	494	146 (68.9)	60 (65.9)	206 (68.0)
Black /	12 (2.5)	2 (0.8)	14 (2.0)	8 (3.8)	1 (1.1)	9 (3.0)
Asian	120 (25.2)	72 (30.4)	192	58 (27.4)	27 (29.7)	85 (28.1)
Native	1 (0.2)	1 (0.4)	2 (0.3)	0	1 (1.1)	1 (0.3)
American	4 (0.8)	5 (2.1)	9 (1.3)	0	2 (2.2)	2 (0.7)
Other	1 (0.2)	0	1 (0.1)	0	0	0
Missing	1 (0.2)	0	1 (0.1)	0	0	0
Weight, mean	71.9 (17.39)	69.4	71.1	72.6 (17.88)	67.4 (15.4)	71.1 (17.3)
Weight,	69 (34−175)	69	69	69 (34−133)	65	69 (34−133)
Weight group (kg), n (%)						
<70	243 (51.1)	124	367	107 (50.5)	54 (59.3)	161 (53.1)
≥70-≤90	174 (36.6)	93 (39.2)	267	77 (36.3)	31 (34.1)	108 (35.6)
>90	58 (12.2)	19 (8.0)	77 (10.8)	28 (13.2)	6 (6.6)	34 (11.2)
Missing	1 (0.2)	1 (0.4)	2 (0.3)	0	0	0
Smoking status, n (%)						
Current	79 (16.6)	38 (16.0)	117	39 (18.4)	13 (14.3)	52 (17.2)
Former	354 (74.4)	178	532	153 (72.2)	71 (78.0)	224 (73.9)

Characteristic	ITT			PD-L1 ≥1% group		
	Durvalumab (n=476)	Placebo (n=237)	Total (n=713)	Durvalumab (n=212)	Placebo (n=91)	Total (n=303)
Never	43 (9.0)	21 (8.9)	64 (9.0)	20 (9.4)	7 (7.7)	27 (8.9)
Disease characteristics						
Disease Stage, n (%)						
IIIA	252 (52.9)	125	377	118 (55.7)	48 (52.7)	166 (54.8)
IIIB	212 (44.5)	107	319	89 (42.0)	42 (46.2)	131 (43.2)
Other[a]	12 (2.5)	5 (2.1)	17 (2.4)	5 (2.3)	1 (1.1)	6 (2.0)
WHO performance-status score, n (%)[b]						
0	234 (49.2)	114	348	105 (49.5)	45 (49.5)	150 (49.5)
1	240 (50.4)	122	362	106 (50.0)	46 (50.5)	152 (50.2)
Not reported	2 (0.4)	1 (0.4)	3 (0.4)	1 (0.5)	0	1 (0.3)
Tumour histological type, n (%)						
Squamous	224 (47.1)	102	326	109 (51.4)	41 (45.1)	150 (49.5)
Non-	252 (52.9)	135	387	103 (48.6)	50 (54.9)	153 (50.5)
PD-L1 status, n (%)[c]						
TC <25%	187 (39.3)	105	292	97 (45.8)	47 (51.6)	144 (47.5)
TC ≥25%	115 (24.2)	44 (18.6)	159	115 (54.2)	44 (48.4)	159 (52.5)
Unknown[d]	174 (36.6)	88 (37.1)	262	N/A	N/A	N/A
EGFR mutation status, n (%)						
Positive	29 (6.1)	14 (5.9)	43 (6.0)	17 (8.0)	4 (4.4)	21 (6.9)
Negative	317 (66.6)	165	482	180 (84.9)	84 (92.3)	264 (87.1)
Unknown[d]	130 (27.3)	58 (24.5)	188	15 (7.1)	3 (3.3)	18 (5.9)
Prior anti-cancer therapy						
Previous radiotherapy, n (%)[e]						
<54 Gy	3 (0.6)	0	3 (0.4)	2 (0.9)	0	2 (0.7)
≥54 to ≤66	442 (92.9)	217	659	193 (91.0)	86 (94.5)	279 (92.1)
>66 to ≤74	30 (6.3)	19 (8.0)	49 (6.9)	17 (8.0)	5 (5.5)	22 (7.3)
Missing[f]	1 (0.2)	1 (0.4)	2 (0.3)	0	0	0
Previous chemotherapy, n (%)[g]						
Adjuvant	3 (0.6)	1 (0.4)	4 (0.6)	2 (0.9)	0	2 (0.7)
Induction	123 (25.8)	68 (28.7)	191	49 (23.1)	21 (23.1)	70 (23.1)
Concurrent with radiation therapy	475 (99.8)	236 (99.6)	711 (99.7)	211 (99.5)	91 (100.0)	302 (99.7)
Best response to previous CRT, n (%)[h]						
Complete	9 (1.9)	7 (3.0)	16 (2.2)	3 (1.4)	2 (2.2)	5 (1.7)
Partial	232 (48.7)	111	343	[redacted]	[redacted]	[redacted]

Characteristic	ITT			PD-L1 ≥1% group		
	Durvalumab (n=476)	Placebo (n=237)	Total (n=713)	Durvalumab (n=212)	Placebo (n=91)	Total (n=303)
Stable	222 (46.6)	114	336	[redacted]	[redacted]	[redacted]
Progression	2 (0.4)	0	2 (0.3)	[redacted]	[redacted]	[redacted]
Non-	9 (1.9)	4 (1.7)	13 (1.8)	[redacted]	[redacted]	[redacted]
Not	2 (0.4)	1 (0.4)	3 (0.4)	[redacted]	[redacted]	[redacted]

Source: Based on Table 4 of the CS[1]

Key: CRT = chemoradiation therapy; CS = company submission; CSR = clinical study report; DCO = data cut-off; EGFR = epidermal growth factor receptor; ITT = intention to treat; N/A = not applicable; PD-L1 = programmed cell death ligand 1; SD) standard deviation; TC) tumour cell; WHO) World Health Organization

Note: The PD-L1 subgroup has been defined using the re-scored PD-L1 data.

Footnotes: [a] Patients with other disease stages included 12 patients in the durvalumab group (four with Stage IV, four with Stage IIB, three with Stage IIA, and one with Stage IA) and five patients in the placebo group (two with Stage IIB, one with Stage IIA, and two with Stage IB); [b] WHO performance-status scores range from 0 to 5, with 0 indicating no symptoms and higher scores indicating increased disability; [c] PD-L1 status was collected before patients received CRT; [d] No sample collected or no valid test result. The *EGFR* status for 2 patients in the durvalumab group changed from unknown to negative between the 13 February 2017 and 22 March 2018 DCOs, as the results for these 2 patients were analysed after the previous DCO; [e] The decision regarding the actual dose was based on investigator or radiologist assessment of each individual patient, resulting in doses that differed from the inclusion criteria. All radiation therapy was administered concurrently with chemotherapy; [f] For the two patients with missing data, the biologically effective radiotherapy dose could not be calculated, primarily because their radiotherapy treatment planning data were neither collected nor accessible; [g] Patients may have received previous chemotherapy in more than one context; h, best response to prior therapy is based on the last therapy prior to entering the study.

ERG comment: In the PACIFIC study, randomisation was not stratified based on PD-L1 status. While reported baseline characteristics, such as age, histology, or smoking status, were balanced between the durvalumab and placebo groups, there are potential problems linked to overinterpretation of subgroup analyses which might impact on the findings.[30]

4.2.3 Efficacy outcomes

The main findings from the PACIFIC study are presented in the CS and reproduced below, see Tables 4.6 and 4.7.

Table 4.6: Key efficacy outcomes for durvalumab versus placebo from the PACIFIC RCT (ITT and PD-L1 ≥1% group; 22 March 2018 DCO)

Endpoint	ITT		PD-L1 ≥1%	
	Durvalumab (N=476)	Placebo (N=237)	Durvalumab (N=212)	Placebo (N=91)
Primary endpoints				
PFS (13 February 2017 DCO; BICR) Median (95% CI) [months]	16.8 (13.0, 18.1)	5.6 (4.6, 7.8)	17.8 (16.9, NR)	5.6 (3.6, 11.0)
HR (95% CI); P-value	0.52 (0.42, 0.65); P<0.001		0.44 (0.30, 0.64); *P*<0.0001	

Endpoint	ITT		PD-L1 ≥1%	
	Durvalumab (N=476)	Placebo (N=237)	Durvalumab (N=212)	Placebo (N=91)
OS (22 Mar 2018 DCO) Median (95% CI), [months]	NR (34.7, NR)	28.7 (22.9, NR)	NR (NR, NR)	29.1 (17.7, NR)
HR (95% CI); P-value	0.68 (0.53, 0.87); *P* =0.003		0.54 (0.35, 0.81); *P*=0.003	
Updated PFS and secondary endpoints (at the time of OS interim analysis; 22 March 2018 DCO)				
PFS (BICR) Median (95% CI) [months]	17.2 (13.1, 23.9)	5.6 (4.6, 7.7)	23.9 (17.2, NR)	5.6 (3.6, 11.0)
HR (95% CI); *P*-value	0.51 (0.41, 0.63); *P*<0.0001		0.44 (0.31,0.63); *P*<0.0001	
TFST Median (95% CI) [months]	21.0 (16.6, 25.5)	10.4 (8.3, 12.5)	25.8 (18.7, 37.8)	10.0 (7.0, 17.0)
HR (95% CI); *P*-value	0.58 (0.47, 0.72); *P* <0.0001		0.51 (0.36, 0.73); *P*=0.0002	
PFS2 Median (95% CI) [months]	28.3 (25.1, 34.7)	17.1 (14.5, 20.7)	33.8 (26.7, NR)	16.5 (10.3, 22.1)
HR (95% CI); *P*-value	0.58 (0.46, 0.73); *P* <0.0001		0.44 (0.30, 0.64); *P*<0.0001	
TSST Median (95% CI) [months]	29.3 (26.0, 34.9)	18.6 (14.8, 23.9)	34.7 (28.8, NR)	17.9 (12.7, 26.2)
HR (95% CI); *P*-value	0.63 (0.50, 0.79); *P* <0.0001		0.49 (0.33, 0.71); *P*=0.0002	
Response rate ORR, % (95% CI)	30.0 (25.8, 34.5)	17.8* (13.0, 23.7)	32.5 (26.0, 39.5)	16.5 (9.3, 26.1)
P-value	*P* <0.001		*P*<0.005	
TTDM Median (95%CI)	28.3 (24.0, 34.9)	16.2 (12.5, 21.1)	NR (26.2, NR)	17.1 (9.2, 20.6)
HR (95% CI); *P*-value	0.53 (0.41, 0.68); *P* <0.0001		0.40 (0.26, 0.61); *P*<0.0001	

Source: Based on Table 6 of the CS[1]

BICR = blinded independent central review; CI = confidence interval; CRT = chemoradiation therapy; CS = company submission; CSR = clinical study report; DCO = data cut-off; HR = hazard ratio; ITT = intention-to-treat; NR = not reached; ORR = objective response, OS = overall survival; PD-L1 = programmed cell death ligand 1; PFS = progression-free survival; PFS2 = time to second progression or death; RECIST = Response Evaluation Criteria In Solid Tumours; TFST = time to first subsequent therapy or death; TSST = time to second subsequent therapy or death; TTDM = time to death or distant metastasis

* may reflect residual effect from prior CRT. The analysis of time to event endpoints was performed using a stratified log rank test adjusting for age at randomisation (<65 versus ≥65), sex (male versus female), and smoking history (smoker versus non-smoker), with ties handled using the Breslow approach.

Table 4.7: Semi-parametric analysis of PPS in patients with confirmed disease progression (BICR); PD-L1 ≥1% group (22 March 2018 DCO)

Post-progression survival	Durvalumab (N=86)	Placebo (N=57)
Total events, n (%)[a]	44	33
Ratio (durvalumab:placebo)	1.33	
Difference (durvalumab–placebo)	11	
Median time to event, months (95% CI)	18.6 (12.5, 26.5)	15.3 (12.5, 18.5)
Ratio (durvalumab:placebo)	1.21 (1.0, 1.4)	
Difference (durvalumab–placebo)	3.22 (0, 8)	

Source: Based on Table 13 of the CS[1]
BICR = blinded independent central review; CI = confidence interval; DCO = data cut-off; ITT = intention-to-treat; PD-L1 = programmed cell death ligand 1; PPS = post-progression survival.

ERG comment: Due to information not being presented in the initial CS, the ERG had to request further information in the request for clarification. This was needed in order to elaborate further on outcomes focusing on OS and PFS. The efficacy results reported in the CS are largely in favour of durvalumab.

However, it should be noted that some results are not yet available as PACIFIC is ongoing. According to the response to request for clarification, the company *"will conduct a final analysis of OS once a maturity of 70% has been reached. OS is an event-driven endpoint; therefore, the timing of this analysis is uncertain"* which will be reached *"when 491 OS events have occurred"*.[1, 24] A Table on page 254 of the CS appendices details the current maturity in the PD-L1≥1% subgroup[27]:

- OS: Durvalumab 33.0%, Placebo 49.5%
- PFS2: Durvalumab 39.6%, Placebo 62.6%
- PFS (BICR): Durvalumab 46.7%, Placebo 72.5%
- PPS (BICR): Durvalumab 51.2%, Placebo 57.9%

4.2.4 Adverse events (AEs)

Key AEs were identified for inclusion in the economic model, see Table 4.8. The CS noted that the incidence and severity of AEs between the durvalumab and placebo groups were comparable. The CS stated that 96.8% of patients in the durvalumab group and 94.9% of patients in the placebo group had experienced at least one AE by the latest data cut-off (DCO), durvalumab was stated to be well-tolerated and had a manageable safety profile relative to placebo. Of the patients in the durvalumab and placebo groups within the safety analysis set, 32.6% and 28.2%, respectively, experienced an AE of Common Terminology Criteria for Adverse Events (CTCAE) Grade 3 or 4. Within the PD-L1≥1% groups, 33.8% patients in the durvalumab group experienced an AE of CTCAE Grade 3 or 4, whereas this was experienced by 23.3% of patients in the placebo group. Within the safety analysis set, serious adverse events (SAE), which included events with death as an outcome, were experienced in 29.1% of patients in the durvalumab group and 23.1% of the patients in the placebo group. In the PD-L1≥1% group, this was seen in 30.0% and 20.0% of durvalumab and placebo group patients, respectively. Within the PD-L1≥1% group the CS identifies the most common AEs among the durvalumab group as being cough, fatigue, radiation pneumonitis, dyspnoea, and diarrhoea. Whereas, the most common AEs in the placebo group were identified as being cough, dyspnoea, fatigue, diarrhoea, arthralgia, and hypothyroidism. According to the CS, of the durvalumab patients the most common AEs of CTCAE Grade 3 or 4 were identified as being pneumonia, anaemia, and pneumonitis. Whereas in the placebo group, the most

common AEs of CTCAE Grade 3 or higher were found to be pneumonia, anaemia, and hypokalemia. The CS also reports the percentage of PD-L1≥1% patients whose AEs resulted in discontinuation of the study treatment as being 36 (16.9%) of durvalumab patients and five (5.6%) of placebo patients. The CS further states the investigators identified 24 patients in the durvalumab group and two patients in the placebo group whose discontinuation was deemed to be causally related to the study treatment. According to the DCO on 22 March 2018, 15 patients in the durvalumab arm and the 10 patients in the placebo arm had died during treatment or within 90 days of the last dose. In the durvalumab group, most of the deaths were attributed to cardiac arrest, whereas in the placebo group deaths were attributed to pneumonia, haemoptysis, intestinal obstruction, and radiation pneumonitis.

Table 4.9 reports the most common AEs (>5% in any treatment group) while Table 4.10 reports HRQoL.

Table 4.8: Summary of key safety events; PACIFIC safety analysis set and PD-L1 ≥1% group (22 March 2018 DCO)

AE category, n (%)[a, b]	Safety analysis set		PD-L1 ≥1% group	
	Durvalumab (n=475)	Placebo (n=234)	Durvalumab (n=213)	Placebo (N=90)
Any AE	**460 (96.8)**	**222 (94.9)**	**205 (96.2)**	**83 (92.2)**
Any AE causally related to treatment[c]	322 (67.8)	125 (53.4)	144 (67.6)	48 (53.3)
Any AE of CTCAE Grade 3 or 4	**155 (32.6)**	**66 (28.2)**	**72 (33.8)**	**21 (23.3)**
Any AE of CTCAE Grade 3 or 4, causally related to treatment[c]	59 (12.4)	11 (4.7)	26 (12.2)	4 (4.4)
Any SAE (including events with outcome of death)	**138 (29.1)**	**54 (23.1)**	**64 (30.0)**	**18 (20.0)**
Any SAE (including events with outcome of death), causally related to treatment[c]	41 (8.6)	9 (3.8)	16 (7.5)	1 (1.1)
Any AE leading to discontinuation of study treatment	**73 (15.4)**	**23 (9.8)**	**36 (16.9)**	**5 (5.6)**
Any AE leading to discontinuation of study treatment, causally related to treatment[c]	47 (9.9)	8 (3.4)	24 (11.3)	2 (2.2)
Any AE with outcome of death	**21 (4.4)**	**15 (6.4)**	**8 (3.8)**	**4 (4.4)**
Any AE with outcome of death, causally related to treatment[b]	7 (1.5)	4 (1.7)	2 (0.9)	0
Any AE leading to dose delay[d]	203 (42.2)	72 (30.8)	96 (45.1)	27 (30.0)
Any other significant AEs[e]	0	0	0	0
Immune mediated AEs[c]	166 (34.9)	39 (16.7)	73 (34.3)	16 (17.8)
Infusion reaction AEs[c]	15 (3.2)	7 (3.0)	3 (1.4)	3 (3.3)

AE category, n (%)[a, b]	Safety analysis set		PD-L1 ≥1% group	
	Durvalumab (n=475)	Placebo (n=234)	Durvalumab (n=213)	Placebo (N=90)

Source: Based on Table 18 of the CS[1]
Note: The PD-L1 subgroup has been defined using the re-scored PD-L1 data.
Footnotes: [a] Patients with multiple events in the same category are counted only once in that category. Patients with events in more than one category are counted once in each of those categories; [b] Includes AEs with an onset date on or after the date of first dose or pre-treatment AEs that increase in severity on or after the date of first dose up to and including 90 days following the date of last dose of study medication or up to and including the date of initiation of the first subsequent therapy (whichever occurs first); [c] As assessed by the Investigator. Missing responses are counted as related; [d] AEs on the AE case report form with Action taken = Drug interrupted, excluding those AEs on the dosing CRF forms only leading to infusion interruptions; [e] Significant AEs, other than SAEs and those AEs leading to discontinuation of study treatment, which are of particular clinical importance, are identified and classified as other significant AEs.
AE = adverse event; CRF = case report form; CS = company submission; CSR = clinical study report; CTCAE = Common Terminology Criteria for Adverse Events; DCO = data cut-off; PD-L1 = programmed cell death ligand 1; SAE = serious adverse event

Table 4.9: Most common AEs (>5% in any treatment group) by preferred term; PACIFIC PD-L1 ≥1% group (22 March 2018 DCO)

Preferred term[a]	Number of patients, n (%)[b, c]	
	Durvalumab (N=213)	Placebo (N=90)
Patients with any AE	**205 (96.2)**	**83 (92.2)**
Cough	71 (33.3)	24 (26.7)
Fatigue	60 (28.2)	19 (21.1)
Radiation pneumonitis[d]	47 (22.1)	10 (11.1)
Dyspnoea	46 (21.6)	23 (25.6)
Diarrhoea	43 (20.2)	14 (15.6)
Pruritus	36 (16.9)	4 (4.4)
Pneumonia	30 (14.1)	7 (7.8)
Pyrexia	29 (13.6)	6 (6.7)
Decreased appetite	28 (13.1)	9 (10.0)
Upper respiratory tract infection	28 (13.1)	8 (8.9)
Rash	27 (12.7)	7 (7.8)
Constipation	27 (12.7)	5 (5.6)
Arthralgia	27 (12.7)	14 (15.6)
Pneumonitis[d]	26 (12.2)	6 (6.7)
Hypothyroidism	26 (12.2)	1 (1.1)
Nausea	24 (11.3)	14 (15.6)
Headache	24 (11.3)	10 (11.1)
Asthenia	23 (10.8)	8 (8.9)
Back pain	22 (10.3)	10 (11.1)
Nasopharyngitis	22 (10.3)	5 (5.6)

Preferred term[a]	Number of patients, n (%)[b, c]	
	Durvalumab (N=213)	Placebo (N=90)
Productive cough	20 (9.4)	6 (6.7)
Vomiting	19 (8.9)	10 (11.1)
Hyperthyroidism	18 (8.5)	1 (1.1)
Anaemia	18 (8.5)	8 (8.9)
Dry skin	18 (8.5)	5 (5.6)
Oedema peripheral	17 (8.0)	5 (5.6)
Non-cardiac chest pain	16 (7.5)	12 (13.3)
Insomnia	15 (7.0)	4 (4.4)
Pain in extremity	15 (7.0)	4 (4.4)
Myalgia	14 (6.6)	5 (5.6)
Bronchitis	14 (6.6)	8 (8.9)
Musculoskeletal pain	14 (6.6)	5 (5.6)
Hypokalaemia	14 (6.6)	6 (6.7)
Dizziness	13 (6.1)	12 (13.3)
Musculoskeletal chest pain	13 (6.1)	7 (7.8)
Hypertension	11 (5.2)	4 (4.4)
Paraesthesia	11 (5.2)	5 (5.6)

Source: PACIFIC PD-L1 subgroup analyses; 22 March 2018 DCO, provided in response to request for clarification [24]

Notes: The PD-L1 subgroup has been defined using the re-scored PD-L1 data; [a] MedDRA version 19.1; [b] Includes adverse events with an onset date on or after the date of first dose or pre-treatment AEs that increase in severity on or after the date of first dose up to and including 90 days following the date of last dose of study medication; [c] Patients with multiple AEs are counted once for each preferred term. Included are events that were reported in at least 5% of the patients in either group; patients with multiple events only counted once in each row; includes AEs with an onset date on or after the date of first dose or pre-treatment AEs that increase in severity on or after the date of first dose up to and including 90 days following the date of last dose of study medication or up to and including the date of initiation of the first subsequent therapy (whichever occurs first); [d] Pneumonitis or radiation pneumonitis was assessed by investigators with subsequent review and adjudication by the study sponsor. In addition, pneumonitis is a grouped term that includes acute interstitial pneumonitis, interstitial lung disease, pneumonitis, and pulmonary fibrosis.

AE = adverse event; DCO = data cut-off; MedDRA = Medical Dictionary for Regulatory Activities; PD-L1 = programmed cell death ligand 1.

Table 4.10: Health-related quality of life of the PACIFIC trial

HRQoL parameter	Partial correlation[a]	Statistical significance
SF-36 Domains		
Physical functioning	-0.276	*P*=0.009
Bodily pain	-0.255	*P*=0.016
Mental health	-0.208	*P*=0.064
SF-36 Summed scores		
Physical component	-0.275	*P*=0.015
EuroQoL parameters		
EuroQoL SC	-0.236	*P*=0.027
EuroQoL VAS	-0.220	*P*=0.038

Source: Based on Table 34 of the CS appendices[27]
Footnotes: [a] A negative correlation coefficient indicates that the presence of metastasis worsens HRQL / utility, while a positive value indicates improvement.
CS = company submission; EuroQOL SC = EuroQOL self-classifier; EuroQOL VAS = EuroQOL visual analogue scale; HRQL = quality of life; SF-36 = 36-item Short Form health survey.

ERG comment: As detailed before, more adverse events are reported for participant treated with durvalumab compared to the placebo arm, see Table 4.9. However, as detailed in Table 4.8, this does include serious adverse events.

4.2.5 Ongoing trials

The CS mentions ongoing phases and phases due to commence in late 2018 of the PACIFIC trial. Such phases include PACIFIC-R, PACIFIC-5, and PACIFIC-6.[1] Pacific–R is a planned retrospective real-world study that will include a large group of patients with locally-advanced, unresectable, Stage III NSCLC who had been included in the EAP and treated with durvalumab. PACIFIC-5 is similar to PACIFIC in that it is also a Phase III, randomised, double-blind, placebo-controlled, multicentre study, which is assessing the efficacy and safety of durvalumab in patients with locally advanced, unresectable, Stage III NSCLC. However, PACIFIC-5 will recruit mainly recruit patients from China and will use a fixed dose of 1500mg every four weeks (Q4W) through an IV fusion rather than using a weight-based dosing system. PACIFIC-6 is a Phase II, open-label multi-centre international safety study focusing on 1500mg of durvalumab Q4W upon completion of sequential CRT. [redacted]

4.3 Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison

The company submission did not present an indirect comparison and/or multiple treatment comparison.

4.4 Critique of the indirect comparison and/or multiple treatment comparison

The company submission did not present an indirect comparison and/or multiple treatment comparison.

4.5 Additional work on clinical effectiveness undertaken by the ERG

No additional work on clinical effectiveness was undertaken by the ERG.

4.6 Conclusions of the clinical effectiveness section

The CS comprised of a systematic review of the evidence for durvalumab for the treatment of locally advanced unresectable, Stage III NSCLC in adults whose tumours express PD-L1 on ≥1% of tumour cells and whose disease has not progressed after platinum-based CRT. The presented evidence included one RCT, the PACIFIC trial.[1]

The PACIFIC trial included patients with confirmed PD-L1 expression on ≥1% of TCs. After randomisation procedures, this resulted in 476 patients in the durvalumab group and 237 patients in the placebo group. However, when focusing on only the patients with confirmed PD-L1 expression on ≥1% of TCs, there were 212 patients in the durvalumab group and 91 patients in the placebo group. While the PACIFIC trial met a multicentre, international design, only eight patients were seeking treatment in the UK. Due to the trial being identified as ongoing, some results are not yet available.

Based on the PACIFIC data there appears to be a benefit in both PFS and OS for durvalumab patients compared with placebo patients, however, the data are immature and there remains substantial uncertainty about the comparative effectiveness.

Common adverse events were reported in both the durvalumab and placebo groups. The common AEs in the durvalumab patients included cough, fatigue, and radiation pneumonitis, whereas patients in the placebo group also included dyspnoea. Overall, more serious adverse events were reported for durvalumab compared to placebo (64/213 (30%) vs. 18/90 (20%)).

Final results for PACIFIC will be published at a later date.

5. Cost effectiveness

5.1 ERG comment on company's review of cost effectiveness evidence

This section pertains mainly to the review of cost effectiveness analysis studies. However, the search section (5.1.1) also contains summaries and critiques of other searches related to cost effectiveness presented in the company submission. Therefore, the following section includes searches for the cost effectiveness analysis review, measurement and evaluation of health effects as well as for cost and healthcare resource identification, measurement and valuation.

5.1.1 Searches performed for cost effectiveness section

This section contains summaries and critiques of other searches related to cost effectiveness presented in the company submission. Therefore, the following section includes the searches for the cost effectiveness analysis review, health-related quality of life and for cost and healthcare resource identification, measurement and valuation. The Canadian Agency for Drugs and Technologies in Health (CADTH) evidence-based checklist for the Peer Review of Electronic Search Strategies (PRESS), was used to inform this critique.[25] The submission was checked against the single technology appraisal (STA) template for company/sponsor submission of evidence.[26] The ERG has presented only the major limitations of the search strategies in the main report. Further minor comments can be found in Appendix 1.

Sections B.3.1, B.3.4 and B.3.5 of the CS state that systematic literature reviews were undertaken to identify studies reporting economic evaluations, health state utility data and cost and resource use data in adults with locally-advanced, unresectable, Stage III NSLC. The 2018 update searches extended the scope to include advanced metastatic Stage IV NSCLC with no restriction to patients treated with CRT.

Search strategies were reported in detail in Appendix G, H and I, and in the response to clarification. MEDLINE, MEDLINE In Process, Embase, EconLit, the HTA database and the NHS Economic Evaluation Database were listed as the databases searched. All databases were searched on 24/25 October 2016, with update searches conducted on 5 March 2018. Searches were limited from 2005 for the cost effectiveness and resource identification strategies, but no date limit was applied to the health-related quality of life strategies. No language limitations were applied in any searches.

Electronic searches were supplemented with hand searching reference lists of included publications and additional websites recommended by NICE, including the cost effectiveness analysis (CEA Registry) for the cost effectiveness searches. Searches were conducted and reported for conference proceedings for the following conferences: ISPOR International and European Congress, European Society for Medical Oncology (ESMO), American Society of Clinical Oncology (ASCO) and the British Thoracic Oncology Group (BTOG).

ERG comment:

- The selection of databases searched was adequate and searches were clearly reported and reproducible. The database name, host, date range and date searched were provided. A good range of additional resources were included.
- In response to clarification, the company confirmed that 2016 Embase searches were jointly conducted for EMBASE and MEDLINE through the EMBASE.com platform. Embase.com searches were conducted as a single search simultaneously over both the Embase and MEDLINE individual databases. As the strategy used contained both MEDLINE and Embase subject heading terms, the ERG confirmed that this should be sufficient to retrieve potentially

relevant records, however the ERG was unable to assess the Embase.com searches in detail, due to lack of access to that host.

- A good range of intervention terms for durvalumab and comparators were used in the cost effectiveness strategy.
- Study design limits to identify economic evaluations, health state utility data and cost and healthcare resource data were applied. The study design filters were not referenced, so it was unclear whether the filters used were published objectively-derived filters. The filters contained a combination of subject heading terms (MeSH and Emtree) and free text terms, and the ERG deemed them to be adequate. The economic evaluation and cost facets used in the 2016 NHS EED and EconLit searches were unnecessary, however, given that these databases only contain economics literature. These limits were not applied to either database in the 2018 update searches.

5.1.2 Inclusion/exclusion criteria used in the study selection

In- and exclusion criteria for the original review on cost effectiveness studies, utilities and costs and resource use are presented in Table 8 of Appendix G, Table 17 of Appendix H and Table 20 of Appendix I of the CS, respectively.[27] To extend the scope of the review, an update was conducted in March 2017 of which the in- and exclusion criteria can be found in Table 9 of Appendix G (cost effectiveness studies), Table 18 of Appendix H (utility studies) and Table 21 of Appendix I (cost/ resource use studies).[27] Extending the scope of the review included a broader patient population (including advanced metastatic disease) and a broader range of interventions (such as immunotherapies, including nivolumab and pembrolizumab), study designs and outcomes.

ERG comment: The ERG agrees that the eligibility criteria are suitable to fulfil the company's objective to identify cost effectiveness studies.

5.1.3 Included/excluded studies in the cost effectiveness review

In total, three unique cost effectiveness studies met the pre-defined eligibility criteria in the original search (references not provided in the CS). However, none of these studies were conducted from a UK perspective. The updated searches related to cost effectiveness studies resulted in one full publication[31] and 20 HTA submissions (only 19 references provided in the CS).[32-50]

The original search yielded one utility study[51], and the updated search resulted in another 52 eligible utility studies (48 studies reporting utility data[32, 33, 38, 41-50, 52-80] and four studies reporting mapping algorithms[81-84]). Of all potentially relevant full publications identified by the original search for costs and resource, none reported UK-related costs or resource use data. The updated search resulted in five studies[85-89] reporting UK specific cost and resource use data.

ERG comment: The rationale for excluding cost effectiveness studies after full paper reviewing are considered appropriate given the defined in- and exclusion criteria.

5.1.4 Conclusions of the cost effectiveness review

The CS provided an overview of the included cost effectiveness, utility and resource use and costs studies, but no specific conclusion was formulated.

ERG comment: The company submission and response to clarification provided sufficient details for the ERG to appraise the literature searches and the 2018 update searches. A good range of databases and additional resources were searched.

Eligibility criteria were suitable for the SLR performed.

5.2 Summary and critique of company's submitted economic evaluation by the ERG

Table 5.1: Summary of the company's economic evaluation (with signposts to CS)

	Approach	Source/Justification	Signpost (location in CS)
Model	State transition (semi-Markov) model	Partitioned survival analysis produced logical inconsistencies	B.3.2
States and events	Progression-free, progressed disease, dead	Progression is a clinically important and patient-relevant endpoint	B.3.2
Comparators	Standard of Care		B.3.2
Population	Locally advanced, unresectable, Stage III NSCLC patients whose tumours express PD-L1 on ≥1% of tumour cells	In line with anticipated marketing authorisation	B.3.2
Treatment effectiveness	Estimated based on PFS and OS data from PACIFIC		B.3.3
Adverse events	Accounted for in terms of their costs (not HRQoL), based on frequency and impact and derived from PACIFC	Utility data from PACIFIC was assumed to include impact of AEs on HRQoL	B.3.3
Health related QoL	Utilities were estimated for progression-free and progressed disease states based on EQ-5D-5L data collected in PACIFIC and mapped to EQ-5D-3L using the NICE recommended cross-walk. A mixed effects model was used to estimate utilities per health state.	In line with NICE reference case	B.3.4
Resource utilisation and costs	Drug acquisition and administration costs, costs associated with treatment-related adverse events, with disease management and patient observation and end of life care were included, based on multiple sources.	Unit prices were based on the National Health Service (NHS) reference prices, Personal Social Services Research Unit (PSSRU), Monthly Index of Medical Specialities (MIMS), and electronic Market Information Tool (eMIT), ESMO guidelines and clinical expert opinion and TA531.	B.3.5
Discount rates	Discount of 3.5% for utilities and costs	Consistent with NICE reference case	CS Table 29
Subgroups	No subgroups		
Sensitivity analysis	DSA, PSA and scenario analyses were performed.		B.3.8

	Approach	Source/Justification	Signpost (location in CS)
Source: CS[1]			
AE = adverse events; CS = company submission; DSA = deterministic sensitivity analysis; ESMO = European Society for Medical Oncology; eMIT = electronic market information tool; EQ-5D-3L/5L = EuroQol Five-Dimension Questionnaire three level / five level version; HRQoL = health-related quality of life; MIMS = Monthly Index of Medical Specialities; NICE = National Institute for Health and Care Excellence; NSCLC = non-small cell lung cancer; OS = overall survival; PD-L1 = program death-ligand 1; PFS = progression-free survival; PSA = probabilistic sensitivity analysis; PSSRU = Personal Social Services Research Unit			

5.2.1 NICE reference case checklist (TABLE ONLY)

Table 5.2: NICE reference case checklist

Elements of the economic evaluation	Reference Case	Included in submission	Comment on whether *de novo* evaluation meets requirements of NICE reference case
Population	As per NICE scope	Only PD-L1 tumour expression ≥1% subgroup	In line with anticipated marketing authorisation
Comparator(s)	Therapies routinely used in the National Health Service (NHS), including technologies regarded as current best practice	Yes	
Type of economic evaluation	Cost effectiveness analysis	Yes	
Perspective on costs	NHS and Personal Social Services (PSS)	Yes	
Perspective on outcomes	All health effects on individuals	Partly	HRQoL impact of AEs excluded
Time horizon	Sufficient to capture differences in costs and outcomes	Yes	
Synthesis of evidence in outcomes	Systematic review (SLR)	Yes	
Measure of health effects	Quality adjusted life years (QALYs)	Yes	
Source of data for measurement HRQoL	Described using a standardised and validated instrument	Yes	
Source of preference data for valuation of changes in HRQoL	Time-trade off or standard gamble	Yes	
Discount rate	An annual rate of 3.5% on both costs and health effects	Yes	

Elements of the economic evaluation	Reference Case	Included in submission	Comment on whether *de novo* evaluation meets requirements of NICE reference case
Equity weighting	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	Yes	
Sensitivity analysis	Probabilistic modelling	Partly	Patient characteristics included in PSA

AEs = Adverse events; HRQoL = health-related quality of life; NHS = National Health Service; NICE = National Institute for Health and Care Excellence; PD-L1 = program death-ligand 1; PSA = probabilistic sensitivity analysis; PSS = Personal Social Services; QALY = quality-adjusted life year; SLR = systematic literature review

5.2.2 Model structure

The company developed a de novo semi-Markov cohort state transition model. The model comprised of three health states, i.e. progression free (PF), progressed disease (PD) and death. The company considered these health states to capture the most important clinical aspects in the treatment of Stage III NSCLC patients, namely the time spent in PF and the time spent alive. The company stated that disease progression impacts on patients' HRQoL, worsens symptoms, removes the possibility of cure, and was therefore also considered to be a clinically important and patient-relevant endpoint. The company's semi-Markov model used estimates of PFS, time-to-progression (TTP) and post-progression survival (PPS) to inform transitions between health states. The company considered this approach was most appropriate as there was limited evidence of the heterogenic effects of individual patient characteristics on disease course and survival (thereby ruling out an individual patient level model), and a three health state approach has been adopted in several other decision models to estimate the cost effectiveness of immunotherapies in advanced metastatic NSCLC.[32, 33, 38, 39, 41, 43, 46, 48, 50, 68]

Partitioned survival analysis was considered as an alternative by the company, however, this approach was not chosen for two reasons:

- All clinically-plausible OS and PFS curves produced logical inconsistencies where the curves crossed.
- Evidence from the PACIFIC trial suggests that prolongation of PFS is the main benefit of durvalumab and PPS is similar between both arms. Therefore, the company concluded that the data lends itself better to deriving OS from PFS and PPS data (semi-Markov approach) than independently extrapolating data for PFS and OS (as with the partitioned survival approach).

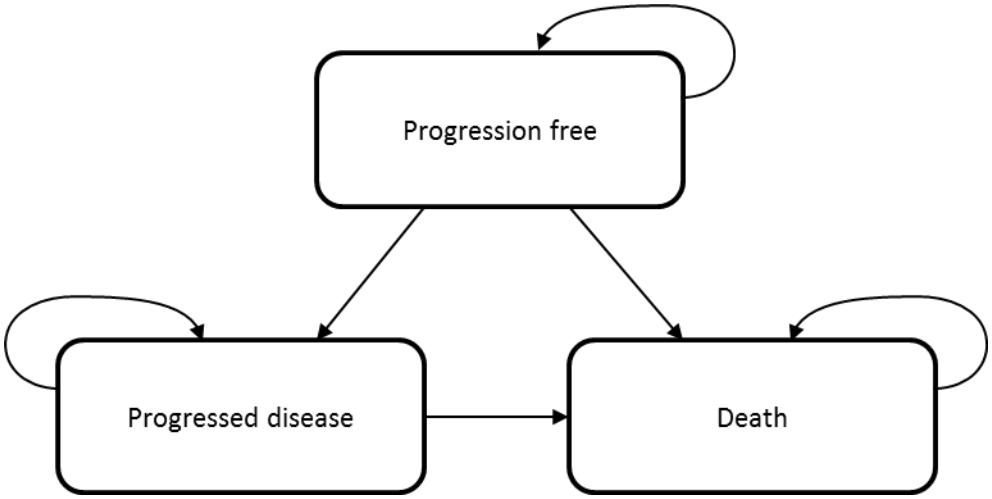
The company claimed to have conducted a partitioned survival analysis as a validation exercise, which was not included as an option in the model.

All patients entered the model in the PF health state. From there, after each cycle, they could remain progression free (modelled using PFS data), or transition to the PD (modelled using TTP data) or death states (modelled using PFS and TTP data).

Patients who experienced disease progression (i.e. local progression and/or metastatic disease) entered the progressed disease health state. The company pooled local progression and progression to

metastatic disease in the same health state. Patients could remain in the progressed disease state, or transition to death (modelled using PPS data for both arms pooled).

Figure 5.1: Model structure



Source: Based on Figure 25 of the CS[1]

ERG comment: The main concern of the ERG relates to the chosen modelling approach. The company's argument to use a semi-Markov approach over a partitioned survival analysis approach was based on the fact that the OS curve could fall below the PFS curve in partitioned survival analysis, therefore being prone to logical inconsistencies. However, in their current approach using a semi-Markov model, the ERG also observed early crossing of TTP and PFS curves as a result of extrapolating the data using the generalised gamma distribution. Furthermore, the ERG considers that the approach may be introducing bias. Survival data in PACIFIC are immature, and whilst the company is correct in pointing out that this issue persists regardless of model choice (OS or PPS), modelling PPS instead of OS is necessarily based on smaller sample sizes used for long-term extrapolation, thereby exacerbating uncertainty. Furthermore, the ERG was concerned that the PPS analysis was potentially biased because groups were no longer balanced. In Technical Support Document (TSD) 19 it is stated: *"Only those patients who have experienced a progression event within the trial follow-up will inform estimation of PPS. This could introduce bias in the extrapolation period if patients who progress within the trial are not representative of those who progress later"*.[90] More specifically, this analysis used data from early progressors, who may have different survival to patients with later progression. Moreover, the PPS data included more patients treated with placebo (who progress earlier), introducing additional bias. The ERG therefore considers that using PPS data instead of OS data may exacerbate the issue of the immaturity of the survival data. The ERG requested results of a partitioned survival analysis to assess any potential differences in results in both approaches, but this was not provided (as the company did not provide survival curves estimated using PFS and OS data from PACIFIC). The magnitude and direction of any bias are unclear.

5.2.3 Population

In line with its anticipated marketing authorisation, durvalumab was considered in the cost effectiveness model for the treatment of locally-advanced, unresectable, Stage III NSCLC patients whose tumours express PD-L1 on ≥1% of TCs and whose disease has not progressed after ≥2 overlapping cycles of platinum-based CRT. This was a subgroup from the final scope issued by NICE, which considered the same population regardless of their PD-L1 status.

Key patient baseline characteristics as applied in the base-case analysis can be found in Table 5.3 below.

Table 5.3: Key baseline patient characteristics of the PD-L1≥1% subgroup as applied in the CS base-case model

Variable	Value	Reference
Patient age (years)	63.1	Table 4, CS[1]
Body weight (kilograms)	71.1	PACIFIC study[91]
Patient body surface area (m^2)	1.83	KEYNOTE-024[92], TA447[68]
% male	69	Table 4, CS[1]
Source: Based on Table 57 of the CS appendices[27] CS = company submission; PD-L1 = program death-ligand 1; TA = technology appraisal		

ERG comment: The main concerns of the ERG relate to: a) modelling a subgroup of the population that was in the final scope issued by NICE, and b) the timing at which the modelled population received CRT.

a) The patient characteristics of the modelled population were comparable to the patient characteristics of the PACIFIC trial. However, in the current submission only a subgroup from the final scope issued by NICE (locally-advanced, unresectable, Stage III NSCLC patients whose tumours express PD-L1 on ≥1% of TCs and whose disease has not progressed after ≥2 cycles of platinum-based CRT) was used to address the decision problem. Nevertheless, the chosen population was in line with its anticipated marketing authorisation and therefore considered appropriate by the ERG.
b) Although sequential CRT is standard practice in the UK[1], the population in PACIFIC and therefore in the model largely received ≥2 overlapping cycles of platinum-based CRT. The potential bias introduced by this is unclear.

5.2.4 Interventions and comparators

Durvalumab was considered within the economic evaluation as per the anticipated licensed indication in NSCLC. Durvalumab was, in line with the dosage used in PACIFIC, modelled with a posology of 10mg/kg administered as an intravenous (IV) infusion over 60 minutes every two weeks (Q2W), until disease progression or unacceptable toxicity, or a maximum of 12 months.

The comparator in the economic model was described as active follow-up or SoC, which applied up to disease progression. The company provided a more comprehensive definition of SoC in its response to the clarification letter as *"surveillance every six months for two years with a visit including history, physical examination and—preferably contrast-enhanced—volume chest CT scan at least at 12 and 24 months is recommended, and thereafter an annual visit including history, physical examination and chest CT scan in order to detect second primary tumours"*.[24]

Patients who experienced disease progression in the model received further treatment and/or end-of-life care, and could be treated with immunotherapy if they met the required criteria. The company stated that subsequent therapies were included in the model if they were used in more than 3% of patients in the PACIFIC study. The list of included subsequent (immuno)therapies and the proportion of patients who received each therapy are shown in Table 41 of the CS.[1] The included immunotherapies were nivolumab, pembrolizumab and re-treatment with durvalumab, and other subsequent therapies were radiotherapy, docetaxel, erlotinib, carboplatin, pemetrexed, gemcitabine, cisplatin, paclitaxel and afatinib.

ERG comment: The intervention was implemented as per its marketing authorisation and dosage.

5.2.5 Perspective, time horizon and discounting

The analysis took an NHS and Personal Social Services (PSS) perspective. Discount rates of 3.5% were applied to both costs and benefits. The model cycle length was two weeks in the first year and four weeks thereafter with a lifetime time horizon (40 years). A half-cycle correction was applied, but not to treatment or treatment administration costs.

ERG comment: This was mostly in line with the NICE reference case, however, in the absence of any justification for not applying the half-cycle correction to treatment and treatment administration costs, the ERG considered this inconsistent with the calculation of resource use and other model components and amended this in its base-case.

5.2.6 Treatment effectiveness and extrapolation

The main source of evidence on treatment effectiveness used for intervention and comparators was the PACIFIC study,[91] a phase III RCT evaluating the efficacy and safety of durvalumab against placebo in all locally-advanced, unresectable, Stage III NSCLC patients regardless of PD-L1 expression levels on tumour cells. Only data from the subgroup of PD-L1 ≥1% patients (according to the anticipated marketing authorisation) and from the March data cut were used in the model. Scenarios were performed to model post-progression survival using alternative data sources namely START[93, 94] and KEYNOTE-024[92] to inform survival with subsequent treatments.

Parametric survival curves were fitted to patient level data from the (post-hoc) PACIFIC PD-L1 ≥1% subgroup and used to extrapolate survival beyond study follow-up. Instead of using the OS data from PACIFIC, the company performed survival analyses on the outcomes TTP and PPS, both post-hoc analyses. Survival analysis was also performed on the pre-specified outcome progression-free survival (PFS), however, in this subgroup this was also a post-hoc analysis.

PFS data were used to determine the number of patients staying in the alive and progression-free health state. TTP data were used to determine the number of patients transitioning to the progressed disease health state. Pre-progression mortality was calculated from PFS and TTP. Post-progression mortality was estimated separately, using PPS data.

Progression-free survival

The probability of remaining in the PF state was estimated using PFS data by fitting independent parametric survival models. The company explored whether the proportional hazard assumption was justified and found that the durvalumab and placebo curves on a log cumulative hazard plot were parallel, suggesting proportional hazards. However, the best fitting curve with this assumption showed bad visual fit to the control arm and the company therefore used independently fitted survival models. Based on statistical goodness of fit, the generalised gamma was selected to model PFS for both durvalumab and placebo. For the durvalumab arm, all other parametric models had a worse statistical fit; for the placebo arm, the log-normal distribution was relatively close. The Gompertz distribution was used in scenario analysis and the log-normal distribution was not used, despite the log-normal making a better fit than the Gompertz in both arms.

The company attempted to validate the PFS extrapolation for SoC in the model against data from the PACIFIC study, other historical RCTs and UK clinical expert opinion. The company stated that their PFS extrapolation for SoC was in line with all these data sources, although it did over-estimate PFS for durvalumab and SoC as observed in PACIFIC (CS Tables 32 and 33).[1]

The PFS curve for durvalumab was altered in the long run to reflect a potential treatment waning effect caused by stopping treatment at a maximum of 12 months. From a chosen cut-off point, which was set to 10 years in the company's base-case, a hazard ratio of one was applied to the placebo curve to model durvalumab PFS. Alternative cut-off points of five years, three years and no cut-off were explored in scenario analyses. The former two significantly drove up the incremental cost effectiveness ratios (ICERs), while the latter only decreased it marginally.

Time to progression

The rate of movement of patients moving from the PF state to the progressed disease (PD) health state was determined by survival analysis of TTP data (PFS data with deaths treated as censored) from PACIFIC. The generalised gamma distribution was chosen in the base-case, based on best statistical fit (AIC and BIC) and to align with extrapolation of PFS.

Post-progression survival

The rate of movement of patients moving from progressed disease to death was estimated using PPS data from PACIFIC, on which survival analysis was performed. The data was only 54% mature in the PD-L1 ≥1% group (CS Figure 30[1]). The analysis was not stratified by treatment arm, but instead pooled across both arms. This was implemented using tunnel states to reflect that patients entered the PF state at different time points. The company assessed that hazards were fairly constant over time based on the log-cumulative hazard plot. The exponential distribution was chosen to model PPS based on best statistical fit (AIC and BIC). The effectiveness of subsequent treatments was captured in the PPS to the extent that patients in the PACIFIC study received subsequent treatments, with chemotherapy being the most commonly used treatment modality in both durvalumab and placebo arms in the PD-L1 ≥1% group; and immunotherapy and palliative-intent radiotherapy also being commonly used in patients who experienced disease progression after treatment with placebo, whilst less frequent treatment with immunotherapies after durvalumab treatment was expected.[24]

In scenario analysis, an alternative method for extrapolating PPS was used. In this scenario, instead of using survival data from PACIFIC, PPS was informed by published data from the KEYNOTE-024 study[92], where data from the pembrolizumab arm were used for those receiving IO treatment, and data from the KEYNOTE-024 chemotherapy arm were used for those not receiving IO treatment. Published data from the START study[93, 94] would be used for predicting survival of non-metastatic patients that did not receive IO treatment, if proportions of (non-)metastatic patients were taken into account in the model (they were not in the revised base-case submitted in response to the request for clarification).[24] A weighted PPS curve was then generated. Log-logistic curves were used to extrapolate survival from KEYNOTE-024 and START, and the company claimed that this was based on best statistical fit. This analysis was changed significantly in response to clarification question B14[24]; partly, it appeared, because the proportions of patients with metastatic disease used in the model were erroneous, and partly, because the proportions of patients receiving IO treatment were estimated based on all progressors and not only those with metastatic progression. Whilst in the earlier analysis in the CS[1], patients in the progressed disease health state were split into those with advanced metastatic disease ([redacted] in durvalumab arm and [redacted] in placebo arm, based on corrected numbers provided in response to the clarification letter[24] Table 15) and those with locally-advanced disease, to reflect patient proportions eligible for subsequent IO treatment, this distinction was no longer made in the revised model based on response to request for clarification[24], and all progressed patients were deemed eligible, based on expert feedback indicating that IO treatment would be given to patients with both metastatic and local progression.[24] The proportions of progressed patients receiving IO treatment in PACIFIC were also corrected to 20% in the durvalumab arm and 39% in the placebo arm.[24]

General population mortality

General population background mortality was implemented in the model when it exceeded the predicted PFS and PPS curves. Furthermore, the company ensured that overall deaths in the model never fell below general population mortality by adapting the transition probability from the PF state to the PD state. In the case of overall deaths in the model falling below general population mortality, instead of the TTP curve, the complement of the PFS curve and general background mortality were used. This came into effect at 2.3 years in the durvalumab arm and at 5.5 years in the SoC arm in the model.

Patient age in the model was based on the age distribution as observed in the PACIFIC trial. Because mortality for each single year age cohort was calculated separately in the base case, the mean age increase reflected different mortality rates by age groups and therefore was not linear. Since the younger patients were more likely to remain alive as compared to older patients, the average age increase in each cycle was less than the exact cycle length.

ERG comment: The main concerns of the ERG relate to: a) potential indirectness caused by the model being based on treatment effectiveness estimates derived from a post-hoc subgroup analysis and post-hoc TTP and PPS analyses, as well as in a population with prior overlapping CRT instead of sequential CRT, and immaturity of survival data; b) durvalumab PFS being potentially over-estimated in the model; c) the end of the KM curves used to extrapolate survival being based on small numbers of patients at risk; d) the implementation and choice of time-point when treatment waning kicks in; e) the implementation of general population mortality; f) crossing progression and survival curves; and g) the uncertainty introduced by immature PPS data, uncertainty about subsequent treatments and methods of extrapolation.

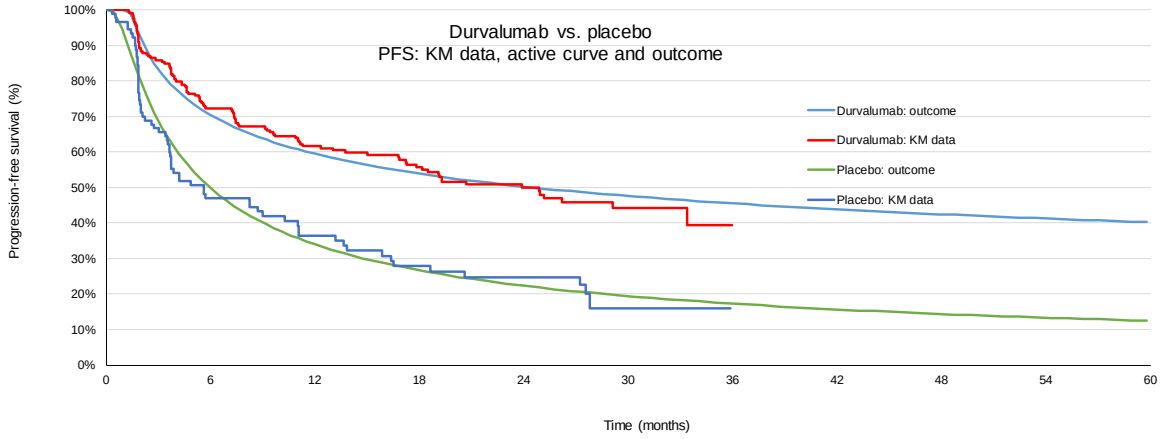
a) The treatment effectiveness in this submission was largely informed by post-hoc analyses performed on the PACIFIC study that may introduce bias in the cost effectiveness model. The subgroup analysis in patients with PD-L1 ≥1% tumour expression was a post-hoc analysis. TTP and PPS analyses were also not pre-specified. Another issue was that the PACIFIC trial may not be generalisable to the UK setting, as it included a majority of patients that received prior overlapping, instead of sequential, CRT. Furthermore, survival data were immature, with only 55% of maturity reached for PFS and 54% of maturity reached for PPS in the combined treatment arms (CS Figure 30[1]). The direction and magnitude of any potential bias stemming from this could not be assessed.
b) The ERG considers the choice of parametric model for estimating PFS (generalised gamma) to likely result in an over-estimate of durvalumab PFS in PACIFIC (see Figure 5.2) and considers all extrapolations to suffer from substantial uncertainty. This is evidenced by the vastly different PFS predictions when different models are used (CS Table 33), where even at five years into the model time horizon, PFS for SoC ranges between 15% and 1%.[1] It is noteworthy that PFS is the model aspect with the most significant impact on the ICERs. For example, in the company's analysis using one model with treatment as a factor (unstratified analysis implemented in the company's model submitted in response to the request for clarification[24]), the ICER with all other company's settings in place increases to £86,332 per QALY gained, highlighting not only the uncertainty associated with PFS, but also the impact of any modelling assumptions around this outcome on the ICER.

   The company acknowledged the potential over-estimate resulting from using the generalised gamma distribution stating *"At three years, the generalised gamma and Gompertz curves may overestimate PFS, although the data from PACIFIC is only based on one patient at this point and so caution should be taken when making comparisons"*.[24] The ERG was surprised that given this uncertainty, the only other tested model was the Gompertz model, which provided very similarly

high PFS estimates at the end of the trial period, whilst the second-best fitting and potentially more realistic lognormal distribution was omitted.

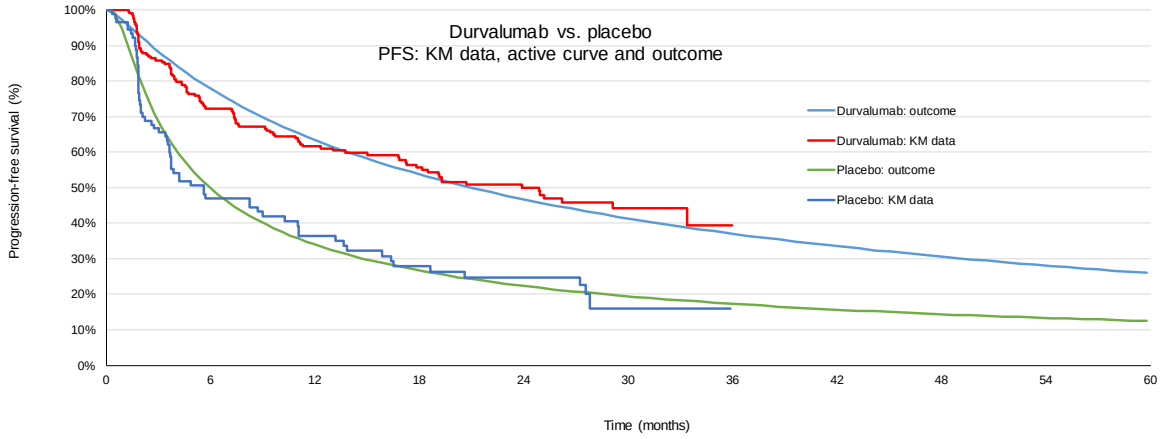
The ERG considers that alternative models with better external validity for extrapolating PFS should be considered in the analysis, particularly given the many censoring events at the end of the Kaplan-Meier curve that result in very small patient numbers at risk. The company claimed to have explored spline-based models, but these and the reasons for which they were discarded, were not reported. Especially given the apparent non-linearity in the log cumulative hazard plots shown in Figure 31 of the CS[1], the ERG considers that such spline-based analyses may potentially be informative.

Figure 5.2: PFS using generalised gamma for durvalumab and SoC (CS base-case)



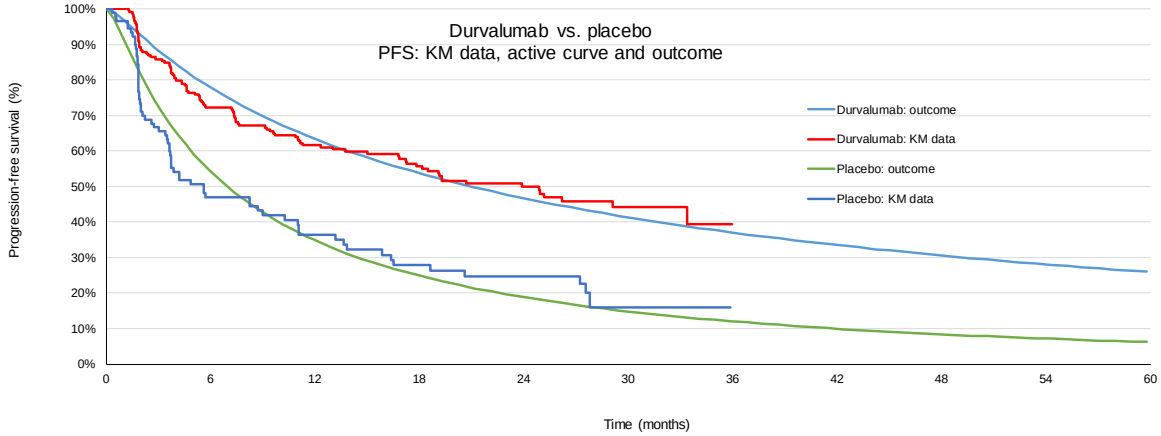
Source: Adapted from revised model[24], background mortality and treatment waning excluded

Figure 5.3: PFS using lognormal for durvalumab and generalised gamma for SoC (ERG base-case)



Source: Adapted from revised model[24], background mortality and treatment waning excluded

Figure 5.4: PFS using lognormal for durvalumab and SoC (ERG scenario)



Source: Adapted from revised model[24], background mortality and treatment waning excluded

Amongst the available fitted curves, the lognormal distribution made the second-best statistical fit based on AIC and BIC criteria for both the durvalumab and placebo arms (Table 31 of the CS), and predicted PFS below (and closer to, in case of durvalumab) that observed in PACIFIC at three years.[1] However, for SoC it would significantly under-estimate five-year PFS as observed in START and supported by expert opinion (6% in the model versus 15% in START). The ERG acknowledges that NICE DSU TSD 14 recommends the use of the same 'type' of model for individual treatment arms to avoid drastically different shapes of survival curves and recommends justification for using different model types per treatment arm by *"using clinical expert judgement, biological plausibility, and robust statistical analysis"* if different model types seem appropriate.[95] In this case, given the above arguments of external validity of modelled durvalumab survival with PACIFIC, and the match of SoC PFS extrapolations with START and clinical expectation, as well as the fact that durvalumab is a treatment with curative intent, the ERG considers there to be arguments for differential distributions per treatment. An additional argument against the choice of the generalised gamma for modelling durvalumab PFS is the potential lack of face validity of virtually no patients progressing or dying in the post-trial follow-up period when the generalised gamma was chosen (see Figure 5.2 durvalumab arm months 36 to 60). In the ERG base-case, the lognormal distribution was therefore used for durvalumab PFS, and the generalised gamma for SoC PFS (see Figure 5.3). ERG scenarios explore the use of a) the generalised gamma for both durvalumab and SoC (as per the company's base-case), and b) the lognormal distribution for both (see Figure 5.4). In each of these analyses, the distributions for modelling TTP are automatically selected based on the choice for PFS, as was done by the company. However, it is noteworthy that any of these choices for modelling PFS are associated with high levels of uncertainty, given the immaturity of the data.

c) The ERG was concerned that the small patient numbers at risk at the end of the KM curves for PFS and PPS potentially biased any extrapolation. In some studies, it has been recommended to truncate KM curves where patient numbers at risk are low.[96] As the company pointed out in response to clarification question B8[24], the NICE DSU TSD 14 only recommends such exclusion of data points when it can be clearly demonstrated that certain points are erroneous outliers.[95] In this case, the ERG considers that this condition is potentially fulfilled: upon examination of Figure 32 of the CS, it appeared that the KM curve after 28 months resulted in a PFS estimate of approximately 16% of patients in the placebo arm, which is based on one patient at risk (Table 32 of the CS).[1] The company, in response to clarification question B8, provided an analysis excluding data points where patient numbers at risk decrease below 5% for both PFS and PPS and showed that the impact on

the ICER was minimal. This satisfied the ERG's concern about small patient numbers at risk having an undue influence on the extrapolated survival curves.

d) The ERG considers the implementation and timing of the treatment waning effect assumed for PFS and TTP a major source of uncertainty. As the company acknowledge, future OS data from PACIFIC will become available that could help assess the long-term survival benefit of durvalumab. In the meantime, any economic modelling has to rely on assumptions that are not supported by data. It is therefore vital that a range of ICERs based on the different possible timings of a treatment waning effect be considered. Whilst the company have provided different scenarios, their choice for the base-case is the most optimistic of those tested (apart from no treatment waning), with a treatment waning effect only starting at 10 years after treatment initiation. In contrast to the company's statements, the OS estimates obtained using this cut-off could not be validated by expert opinion (OS only estimated for SoC at 10 years by clinical expert, see CS Table 35), or modelled OS from other appraisals in the metastatic setting (modelled OS with durvalumab 27% at 15 years compared with modelled OS with other IO treatments of 0-3%; based on CS Table 36).[1] It is the ERG's opinion (acknowledging the lack of evidence) that the five year cut-off would be more realistic than the 10 year cut-off, still resulting in durvalumab OS at 15 years of 20% (using the company's model settings). Although the ERG is unsure of its applicability to this setting, it should be noted that a five-year cut-off was accepted by NICE in TA520 (Atezolizumab for treating locally advanced or metastatic non-small-cell lung cancer after chemotherapy).[50] The ERG base-case considers a five-year cut-off point. Seven- and three-year cut-offs are tested in scenarios.

   With regards to implementation of the treatment waning effect, setting the hazard ratio to one at the chosen cut-off can cause counter-intuitive results, if the per-period hazard in the comparator arm is below that of the hazard in the intervention arm. This can occur depending on the chosen cut-off and for example when patients in the comparator arm have high rates of progression or dying in the beginning periods, such that the few remaining patients alive and not progressed in later periods then have very low rates of progression or dying. This occurs in the ERG scenarios (3) and (4), where a shorter cut-off decreases the ICER and a longer cut-off increases it. Figure 5.3 illustrates that the hazards in the extrapolated PFS for SoC are below those of durvalumab in the ERG base-case. The ERG therefore considers the method for implementing treatment waning to be flawed, but acknowledges that there is a lack of guidance related to alternative modelling methods and that there is precedent for the company's method as was highlighted in response to clarification question B9b.[24] The ERG considered alternative methods for modelling treatment waning including setting the durvalumab PFS curve equal to that of SoC at the chosen cut-off, resulting in a sudden drop of patients not progressed or died (which the company had pointed out to lack realism in response to clarification question B9b[24]). In the absence of any supporting evidence for either approach, to explore the impact of different treatment waning cut-off points, the ERG kept the company's modelling method and added two scenarios where both alternative cut-off points were explored with the lognormal distribution used for modelling PFS in both the durvalumab and SoC arms. Furthermore, the impact of not modelling any treatment waning effect on the ERG base-case was explored in a scenario.

e) The company's way of ensuring that mortality in the model never falls below general background mortality may favour durvalumab, but is likely to have a minor impact. The company applied a fix in the transition to the progressed disease health state to avoid that the "*difference between the hazards derived from PFS and TTP survival functions [...] was smaller than the general population mortality*".[24] In short, the company attempted to ensure that overall mortality from the PF state was never below that of the general population by artificially lowering the number of progressors (rather than lowering the number of people remaining in the PF state). Since this fix applies sooner for

durvalumab than for SoC, this was likely a non-conservative way of modelling transitions away from the PF state. However, this is unlikely to be influential and the ERG considers this to be acceptable.

f) Another small concern relates to crossing progression and survival curves in the company's model. In this case, this was an artefact of fitting survival models to relatively similar KM estimates. Parametric survival models with non-monotonic hazards may cause the progression curve (proportion of patients who have not progressed, with deaths censored) to drop quicker than the PFS curve (proportion of patients who have not progressed or died), resulting in a probability of progression that can exceed that of progressing or dying. The company claimed in their response to the clarification letter that this did not occur, but according to their own model check described in this response, this is not correct:[24] using the company's base-case generalised gamma distribution, this only occurred in cycle four in the durvalumab arm, and using the ERG's preferred lognormal distribution, this occurred only in cycle two in the durvalumab arm. The company's adjustment for background mortality meant that this did not result in negative patient numbers in any states, and since it only occurred in one cycle, the ERG considers this issue to be likely acceptable, with the caveat that more flexible models, such as spline models, should have been explored.

g) PPS data are immature and there is substantial uncertainty about post-progression survival. Additional bias in the extrapolation of PPS may be caused by inclusion of early progressors but not late progressors, and more progressors from the placebo arm than the durvalumab arm (see Section 5.2.2 ERG comment and NICE DSU TSD 19[90]). The company explored alternative ways of modelling PPS using data from other studies in a scenario. The generalisability of this analysis to this setting is unclear: for example, KEYNOTE-024 included metastatic patients. Apart from this, the ERG noted an error in the selection of the survival distribution applied in this scenario. The company claimed to have used the curve with the best statistical fit, but used the log-logistic distribution instead (third best statistical fit), which biased model outcomes in favour of durvalumab, compared with using the distribution with the best statistical fit in KEYNOTE-024 (lognormal). Furthermore, there was an inconsistency in the proportions of patients receiving IO treatment in the durvalumab arm, which did not appear to include those patients that were re-treated with durvalumab (7%, as used in the cost estimates of the model).

   To address some of the uncertainty in PPS, the ERG explored alternative assumptions around PPS: selecting an alternative model in the company's base-case modelling of PPS using PACIFIC (generalised gamma) with the second best statistical fit; and using the scenario to inform PPS based on the KEYNOTE-024 data, but with the distributions that exhibited the best statistical fit (lognormal instead of log-logistic) and a corrected estimate of patients receiving subsequent IO treatment in the durvalumab arm.

5.2.7 Adverse events

The main source of evidence on treatment adverse events used for durvalumab and SoC was the PACIFIC study. Adverse events (AEs) were included in the model in terms of their costs and not their impact on health-related quality of life (HRQoL). AEs were included if they had a frequency of ≥2% in either arm in the PACIFIC study (PD-L1 ≥1% group) and a severity of grade 3/4, or if they were judged to have a sizable impact on either costs or HRQoL (see CS Table 37 for a list of included AEs).[1] AEs were modelled as a per-cycle occurrence while patients were on treatment. No detail was provided on how the total treatment years per arm (183.6 and 66.2 years, for durvalumab and placebo, respectively) were derived from PACIFIC.

ERG comment: The main concerns of the ERG relate to: a) under-estimation of impact of AEs when treated with durvalumab; and b) the lack of justification for the total treatment years per arm to derive the incidence of AEs.

a) AEs were selected for inclusion in the model based on frequency (occurrence ≥2% in PACIFIC), severity (grade 3/4) and impact on costs (CS Table 37).[1] Incidence in the model for the selected AEs was comparable between treatment arms. However, in Table 18 (reproduced in Table 4.8) of the CS, percentages of 'Any AE of CTCAE grade 3 or 4, causally related to treatment', 'Any SAE (including events with outcome of death), causally related to treatment', and 'Any AE leading to discontinuation of study treatment, causally related to treatment' were mostly higher for durvalumab as compared to placebo in the PD-L1 ≥1% group.[1] The company did not provide an explanation for how this discrepancy between the AEs listed in Table 18 and the AEs used in the model occurred.[1, 24] The ERG was therefore concerned that the impact of AEs associated with durvalumab treatment in the model may be under-estimated and explored this in scenario analysis. This scenario used the numbers of events from 'Any AE of CTCAE grade 3 or 4, causally related to treatment' in Table 18 and total treatment years for durvalumab and placebo to calculate revised two-weekly (per cycle) incidences for all grade 3/4 AEs per treatment arm. Combined with the unweighted average utility decrements and costs for the AEs that were included in the company's base-case model, the ERG derived one-off costs and utility decrements per cycle that reflected the amended incidence. These were then used together with amended AE utility decrements as detailed in Section 5.2.8 of this report.
b) It was unclear how the total treatment years per arm were derived. The shorter duration in place for the placebo arm resulted in a higher incidence of AEs and may bias model outcomes in favour of durvalumab. However, the impact of assuming the same value for total treatment years on the ICER was only small.

5.2.8 Health-related quality of life

Utility values were estimated for the following health states: PF and PD. EQ-5D-5L data were collected in PACIFIC and, in alignment with the NICE position statement[97], the crosswalk mapping algorithm by van Hout et al.[98] was used to obtain EQ-5D-3L utility scores. These utility scores were subsequently used to model utility values for PF and PD health states. A variety of mixed effects models including different covariates were constructed and tested. The covariates included treatment, age, health state (pre- or post-progression), time to death, and treatment disposition (on or off treatment), but ultimately only progression was used as a covariate in the base-case analysis. As a consequence, utility values were equal across treatment arms and an age-related utility decrement was not incorporated.

Utility values resulting from the mixed effects model based on EQ-5D-3L data from the PACIFIC trial were 0.810 for PF and 0.776 for PD. These utility values were compared to the utility values in the studies identified in the SLR. Although the company stated in the CS that there was broad consistency, the utility values derived from PACIFIC data were higher than in these other studies. However, patients in PACIFIC had less metastatic disease than in the comparator studies. The PACIFIC utility values were also higher than in the general population.[99] A summary of all utility values used in the model is provided in Table 5.4.

Table 5.4: Health state utility values

State	Utility value	Reference	Justification
Progression free	0.810	PACIFIC data [100]	SLR did not identify suitable utility scores

State	Utility value	Reference	Justification
Progressed disease	0.776	PACIFIC data [100]	SLR did not identify suitable utility scores
Source: Based on Table 39 of the CS [1] CS = company submission; SLR = systematic literature review			

Health-related quality of life data identified in the review

According to the CS, the SLR identified three key studies reporting UK relevant utility values. Out of these, the company considered only one to be possibly suitable to provide an alternative value for utility after progression.[101] The other studies were deemed not suitable since they concerned a metastatic setting only, while the PD state in the CS is a combination of local and metastatic disease progression.

Adverse event related disutility values

In the base-case analysis, no adverse event related disutilities were taken into account. The company justified this claiming that the impact of adverse events on QoL was assumed to be reflected in the EQ-5D data as observed. In a scenario analysis, a disutility value was applied for grade 3/4 AEs. See Table 5.5 for details on the disutilities.

Table 5.5: Adverse event related disutility values

Adverse event	Disutility value (per 2-week cycle)	Reference	Justification
Pneumonia	-0.037	Nafees et al. 2008[57]	
Anaemia	-0.043	KEYNOTE-010 trial as per TA428[33]	
Hypertension	-0.110	Nafees et al. 2008[57]	
Pneumonitis	-0.037		Assumed equal to pneumonia
Endocrinopathy	0.000	Clinical opinion (no reference provided in CS)	
Hypokalaemia	-0.110	Nafees et al. 2008[57]	Assumed equal to fatigue
Haemoptysis	-0.037		Assumed equal to pneumonia
Radiation pneumonitis	-0.037		Assumed equal to pneumonia
Source: Based on Table 56 of the CS appendices[27] CS = company submission; TA = technology appraisal			

ERG comment: The main concerns of the ERG relate to: a) the high utility value of the PF health state which is also constant with age; b) the modest utility decrement for progressed disease; and c) utility scores for durvalumab and SoC being equal, without consideration for treatment or AEs.

a) The utility value for the PF health state was 0.810 which is comparable to the utility reported for the general population (0.80 for age category 55-64).[99] Utility scores equally high as in the general population seem quite unlikely in patients with locally advanced NSCLC. The company justified the use of the 0.810 for PF by stating that general population scores were based on EQ-5D-3L data (where PACIFIC used EQ-5D-5L) and population scores may also be outdated. In clarification question B16[24], the ERG argued that there are more recent population norms which

have not shown a significant increase (i.e. 0.81 and 0.802 for the relevant age category[102]). The difference between 3L and mapped 5L scores of the EQ-5D remains, but was recently shown to be only minor.[103] In addition, utility scores in the base-case model did not decrease with age, since age was not a significant factor in the mixed effects model. However, the mixed effects model only included two age categories (<65 and ≥65) and the ERG does not consider the absence of a significant effect in the short run of the trial to sufficiently support an assumption of utility values being constant over a lifetime time horizon. In summary, utility values for PF were remarkably high and remained high for the full-time horizon of the model. A high utility score for PF lowers the ICER, as in the model patients on durvalumab progressed later than patients receiving SoC. The ERG base-case incorporated an age-related decrement. The ERG also proposed a lower (start) utility score for PF, i.e. 0.73, taken from Ara and Brazier,[102]for people from the general population aged 65-70 with a history of cancer. Although this lower utility value may have better face validity, it does not fully apply to the population in the scope, and therefore it was only incorporated in a scenario.

b) The utility decrement for progressing to PD was -0.034, which could be considered quite modest given the information from the literature review performed by the company as provided in Table 38 of the CS[1], which shows the decrement for progressed disease to vary from -0.4[101] to -0.18[57]. The low decrement that resulted from the mixed effects model could partly be due to the fact that EQ-5D-5L data was only collected up to 30 days after progression. The company confirmed that HRQoL is likely to continue to decline further but also states that their approach was a conservative one since patients in SoC progressed earlier and a high utility value for PD would overestimate QALYs. The ERG agrees with this, but argued that a larger utility decrement would be more reflective of clinical reality. In line with findings by Chouaid et al.[101] in a Stage III/IV NSCLC population, the ERG explored a scenario (applied in addition to the lowered PF utility of 0.73 scenario mentioned above) considering a utility value of 0.67 for first-line progressed disease. Since this scenario was only performed in addition to the scenario with lowered PF utility of 0.73, it implied a decrement for progression of 0.06.

c) Although the mapped utility scores from PACIFIC were higher in the placebo as compared to the durvalumab arm at almost all measurement moments, treatment was found to be statistically insignificant in the mixed effects model and therefore, equal utilities were assumed for durvalumab and SoC. However, the company did not apply utility decrements for AEs in the base-case model as these were assumed to be incorporated in the utilities as observed. When applying utility decrements for AEs in a sensitivity analysis, the company only included these for a selected set of AEs (see also ERG comment in Section 5.2.7). In response to clarification question B18[24], the company provided results of alternative analyses using separate utility values for durvalumab and SoC, both as a factor in the mixed effects model and as the observed average EQ-5D-5L utility scores, which showed increased ICERs (£20,172 and £20,261, respectively). The ERG is concerned that by excluding treatment as a factor in the mixed effects model, and at the same time including disutilities of a limited set of AEs only in a sensitivity analysis, the true impact of adverse events was not appropriately captured in the base-case model or in the scenario. Given the fact that OS data are not fully mature (38% maturity at time of primary analysis), quality of life becomes all the more important, and therefore it is paramount to take AEs into account as accurately as possible. Also grade 1 and 2 AEs will have an impact on the patient's quality of life, but these less severe events were excluded from the analysis. For this reason, the ERG base-case included treatment as a factor in the mixed effects model.

5.2.9 Resources and costs

The cost categories included in the model were costs for PD-L1 testing, costs associated with treatment (drug acquisition costs including subsequent therapies, drug administration costs including subsequent therapies, costs associated with treatment-related AEs), costs associated with disease management and patient observation, and costs associated with end of life care.

Unit costs were based on the National Health Service (NHS) reference costs[104], Personal Social Services Research Unit (PSSRU)[105], Monthly Index of Medical Specialities (MIMS)[106], and the electronic Market Information Tool (eMIT)[107].

Resource use and costs data identified in the SLR

According to Appendix I of the CS[27], the SLR performed in October 2016 (with an update in March 2018) identified 115 publications of which five reported UK specific cost/resource data. The company stated that none of the eligible UK studies were precisely aligned with the population of interest for this appraisal and none reported cost or resource use information relevant for the economic model.

Treatment costs

The average cost per infusion of durvalumab was calculated by multiplying the cost per mg (£4.93) by the average body weight in the PD-L1 ≥1% group as observed in the PACIFIC study (71.1 kg) and dosage (10mg/kg). The base-case analysis assumed no wastage (perfect vial sharing), which was explored in a scenario analysis. Duration of treatment in the durvalumab arm was according to Kaplan-Meier (KM) data from the PACIFIC study. Total mean treatment costs using these numbers amounted to [redacted] (see Table 5.6). The model assumed zero acquisition costs for SoC as concomitant treatment use was similar in durvalumab and placebo arms of the PACIFIC study.

Treatment administration cost was, in the absence of a specific tariff for durvalumab administration, based on NHS reference cost code SB12Z (cost of administering simple chemotherapy)[104] at £241.07 per cycle. For SoC there were no administration costs.

PD-L1 testing costs were calculated as a cost per eligible patient. As per information in Table 41 of the CS[1], 1.89 patients would need to undergo a PD-L1 test in order to identify one patient eligible for treatment with durvalumab. That is, of the patients in PACIFIC for whom a PD-L1 test was performed (76.4% since PD-L1 testing was not mandated for inclusion), 56% was eligible for treatment with durvalumab. Corrected for 5% of patients who would have progressed in the meantime (based on clinical expert opinion[108]), final eligibility would be 53%. Therefore, the unit price of a PD-L1 test (£40.50 as reported in NICE TA531 [32]) was multiplied with 1.89 to obtain the cost for PD-L1 testing per eligible patient of £76.68.

Table 5.6: Treatment acquisition costs

Item	Durvalumab	Justification
Dosing per administration	10mg/kg	Draft SmPC [109]
Frequency of administration	Q2W	Draft SmPC [109]
Total dose per administration	711 mg	Mean patient weight in PACIFIC PD-L1≥1% group: 71.1 kg * 10 mg/kg
Treatment cost per 120 mg vial	£592	Anticipated list price
Treatment cost per 500 mg vial	£2,466	Anticipated list price
Treatment cost per cycle (Q2W)	£3,507	711*(£2,466/500)
Total mean treatment cost	[redacted]	£3,507*(30/14)*([redacted]
Administration cost per cycle (Q2W)	241,07	Total HRGs SB12Z[104] Same source as approved NICE TAs[110]

Source: Based on Table 40 of the CS[1]

CS = company submission; HRG = Healthcare Resource Group; Q2W = every two weeks; PD-L1 = programmed death-ligand 1; TA = technology appraisal; SmPC = Summary of Product Characteristics

Costs of subsequent treatments

Upon disease progression, be it local or metastatic, patients in the model could go on to receive further treatment. Immunotherapy was an option if patients met the required criteria. Subsequent therapies were included in the model if they were used in more than 3% of patients in either arm of the PACIFIC study. Once patients progressed in the model, a one-off cost for subsequent treatments was accrued. This cost was informed by the type of treatment, the required treatment dose, the dosing schedule, the unit drug cost at list prices, and the duration of treatment (see Table 5.7). The average cost of subsequent treatment was determined using the distribution of patients across the various treatments as observed in the PACIFIC study, resulting in a one-off total subsequent treatment cost of [redacted] for durvalumab and [redacted] for SoC. Duration of treatment could be manually adjusted in the economic model. The model also allowed for selecting the START trial[93] as a source for distribution of patients across subsequent treatments, thereby excluding immunotherapy.

Table 5.7: Costs of subsequent treatments

Subsequent treatment	% of progressed patients in durvalumab arm who received treatment*	% of progressed patients in placebo arm who received treatment*	Dose	Duration of treatment (weeks)	One off cost per patient applied in the model	Reference unit prices
Immunotherapy						
Nivolumab	15%	32%	240 mg	26.33	£37,832	MIMS (Opdivo)[106]
Pembrolizumab	5%	7%	200 mg	21.40	£39,241	MIMS[106]
Durvalumab	7%	0%	10 mg/kg	17.11	£32,056	AstraZeneca, anticipated list price
Other commonly-used subsequent therapies						
Radiotherapy	36%	35%	N/A	N/A	£2,802	NHS reference costs[104]
Docetaxel	22%	7%	75 mg/m^2	14.35	£1,274	eMIT 2018[107]
Erlotinib	6%	11%	150 mg	47.83	£18,209	MIMS (Tarceva)[106]
Carboplatin	34%	30%	AUC; 500 mg dose assumed	14.35	£1,253	eMIT 2018[107]
Pemetrexed	21%	12%	500 mg/m^2	14.35	£8,155	MIMS (Alimta)[106]
Gemcitabine	21%	18%	1000 mg/m^2	14.35	£2,449	eMIT 2018[107]
Cisplatin	9%	11%	75 mg/m^2	14.35	£1,212	eMIT 2018[107]
Paclitaxel	9%	14%	200 mg/m^2	14.35	£1,268	eMIT 2018[107]
Afatinib	6%	5%	40 mg	47.83	£16,935	MIMS (Giotrif)[106]

Source: Based on Tables 42-45 of the CS[1]

* Based on PACIFIC

Note: percentages can add up to more than 100% due to use of combination treatments and multiple lines of treatment

AUC = area under the curve; CS = company submission; eMIT = electronic Market Information Tool; MIMS = Monthly Index of Medical Specialities; NHS = National Health Service

Health state costs

Resource use associated with patient observation and disease management in both PF and PD health states was applied to all patients based on their treatment arm, treatment status and disease progression status. For the PD health state, the model used resource use and costs identified and accepted in TA531.[32] It should be noted that TA531 concerned an all metastatic population, in contrast to the current

CS. For the PF health state, patients in SoC were assumed to receive care according to ESMO guidelines[111] and clinical expert opinion.[112] Patients on treatment in the durvalumab arm were assumed to receive a scan every two months as well as a blood test every visit. No costs of outpatient visits or clinical nurse specialist visits were included for patients on durvalumab, as these were expected to be captured in the cost of administering durvalumab. After discontinuation of durvalumab, SoC costs were applied to the durvalumab arm as well. For both treatment arms, observation and management costs were assumed to be reduced to £0 after five years (i.e. after four years off-treatment for durvalumab arm), which was confirmed by clinical expert opinion in response to clarification question B19.[24] See also Table 5.8.

Table 5.8: Health state related costs per month

Health state	Monthly costs durvalumab on-treatment	Monthly costs durvalumab off-treatment	Monthly costs SoC (off-treatment)	Reference resource use
PF				
Year 1	£62.28	£103.18	£103.18	Draft SmPC [109], Clinical expert opinion [112] and ESMO guidelines [111]
Year 2	NA	£68.37	£68.37	Clinical expert opinion [112] and ESMO guidelines [111]
Year 3	NA	£34.82	£34.82	Clinical expert opinion [112] and ESMO guidelines [111]
Year 4	NA	£34.82	£34.82	Clinical expert opinion [112] and ESMO guidelines [111]
Year 5	NA	0	£34.82	Clinical expert opinion [112] and ESMO guidelines [111]
Year 5+	NA	0	£0	Clinical expert opinion [112] and ESMO guidelines [111]
PD				
	NA	£304.94	£304.94	TA531 [32]and Big Lung trial [113]

Source: Based on Table 48 of the CS[1]
CS = company submission; ESMO = European Society for Medical Oncology; NA = not applicable; PD = progressed disease; PF = progression-free; SmPC = Summary of Product Characteristics; SoC = standard of care; TA = technology appraisal

Adverse event related costs and costs of terminal (end of life) care

The frequency of occurrence of grade 3/4 AEs was combined with a one-off cost per AE to obtain a total per-cycle cost for each arm: £6.88 for durvalumab and £9.20 for SoC. This total cost was applied to the duration that patients were on assigned treatment (see Table 5.9).

Furthermore, a one-off cost of £3,577 was applied in the model when a patient died, to reflect the costs of terminal care (see Table 5.9). This cost, according to the CS, was based on values accepted in a NICE multiple technology appraisal for erlotinib and gefitinib (TA374 [114]). Table 51 of the CS[1] also refers to TA531[32], the Marie Curie report[115] and NICE clinical guidance (CG) 81[116] for this.

Table 5.9: Adverse event related costs and costs of terminal care

Adverse event	Costs	Per cycle cost durvalumab	Per cycle cost SoC	Reference
Anaemia	£753.02	£0.96	£1.80	NHS reference costs 2016-2017[104]
Hypertension	£388.81	£0.00	£0.00	NHS reference costs 2016-2017[104]
Haemoptysis	£391.98	£0.00	£0.46	NHS reference costs 2016-2017[104]
Hypokalaemia	£151.69	£0.13	£0.46	NHS reference costs 2016-2017[104]
Pneumonia	£1,851.16	£4.79	£5.56	NHS reference costs 2016-2017[104]
Pneumonitis	£391.98	£0.50	£0.23	NHS reference costs 2016-2017[104]
Radiation pneumonitis	£391.98	£0.41	£0.70	Assumed equal to pneumonitis as no HRG available
Endocrinopathy	£443.46	£0.09	£0.00	NHS reference costs 2016-2017[104]
Total per cycle AE costs				
		£6.88	£9.20	
Total cost of terminal care				
One-off	£3,577.18			TA374 [114], TA351 [32], Marie Curie report [115], NICE CG81 [116]

Source: Based on Tables 49, 50, and 51of the CS [1]
CG = clinical guidance; CS = company submission; HRG = Healthcare Resource Group; NICE = National Institute for Health and Care Excellence; SoC = standard of care; TA = technology appraisal

ERG comment: The main concerns of the ERG relate to: a) the assumption of perfect vial sharing; b) resource use in the PD health state; and c) the criterion for inclusion of subsequent treatments in the model.

a) The assumption of perfect vial sharing that was maintained in the model is not realistic, also given the limited number of patients in England and Wales that would be eligible for treatment with durvalumab (367 annually). The company stated in their response to clarification question B22[24] that indeed, they did not expect perfect vial sharing to occur in clinical practice, but that their base-case was chosen based on recent policy initiatives put in place by NHS-E for IOs.[117] The ERG has looked into these policy initiatives documents and did not find information that, at this time, directly or indirectly supported the assumption of perfect vial sharing. When perfect vial sharing is so clearly not feasible in clinical practice, it should not be considered as base-case. The ERG base-case therefore assumed that there is no vial sharing, with the possibility of 30% vial-sharing in a scenario. The ERG noted an error in the implementation of vial wastage for nivolumab and pembrolizumab, which affected the company's vial sharing scenarios. The company had erroneously employed weight-based dosage calculations on a fixed dose. This was fixed in the revised ERG base-case.

b) The resource use in the PD health state was based on TA531[32], where these costs were estimated for patients with exclusively metastatic disease. However, only a minority (37%, as stated in section B.3.3 of the CS [1]) of the progression events in PACIFIC were metastatic, which is why the ERG considered that using TA531 resource use may have overestimated costs for PD in the economic model. In their response to clarification question B20, the company stated they believe that *"based on similar subsequent treatment use expected between local and metastatic progression, … TA531 is a reasonable source of costs for the PD state, in the absence of other data"*.[24] The ERG was not convinced by this argument. However, as the company has shown that incorporating a zero cost for local recurrence would only slightly increase the ICER, the ERG considers this issue likely to be acceptable.

c) Subsequent treatments were included in the model if they were used in more than 3% of patients in either treatment arm in the PACIFIC study. The company did not justify what this cut-off of 3% was based on. Moreover, this 3% criterion was apparently (looking at Table 42 of the CS[1]) applied to the total study population (including progression-free patients), not only to the progressed patients. That is, any subsequent therapy that would be used in 3% of progressed patients would not be included in the model since the percentage would be lower in the complete group. As the proportion of progressed patients was higher in the placebo arm, the effect of this difference would not be equal between the arms. That is, treatments that were given to 6% of the *progressed* durvalumab patients would *not* be included (since it would translate into just below 3% of all durvalumab patients), while treatments that were given to 6% of the *progressed* SoC population *would* be included (since it would translate into just over 3% of all SoC patients). Although the magnitude of the bias caused might be limited, the cost of subsequent treatment was an influential factor in the model. The ERG's predominant concern is that the selection criterion was not transparent, nor justified. No adjustments were made in the ERG base-case with regards to this issue.

5.2.10 Cost effectiveness results

In the deterministic base-case analysis, total LYs and QALYs gained were larger in the durvalumab arm compared to the SoC arm. Incremental QALYs (2.93) were mainly driven by QALY gains in the PF health state. Total costs were also higher for durvalumab than for SoC. The incremental costs ([redacted]) mainly resulted from higher treatment costs. The deterministic incremental cost effectiveness ratio (ICER) amounted to £19,366 per QALY gained (Table 5.10).

Table 5.10. Company's revised base-case results

	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
Durvalumab	[redacted]	[redacted]	[redacted]	--	--	--	--
SoC	[redacted]	[redacted]	[redacted]	[redacted]	3.60	2.93	£19,366

Source: Based on the revised base-case results in the economic model
ICER = incremental cost effectiveness ratio; LYG = life years gained; QALY = quality adjusted life year; SoC = standard of care

ERG comment: In response to the clarification letter, and as requested by the ERG, the company corrected an error that was present in the age calculations for the SoC treatment arm. The base-case

results were slightly changed by this correction (ICER increased from £19,320 to £19,366 per QALY gained), and Table 5.10 above presents the revised base-case results after correction by the company.

5.2.11 Sensitivity analyses

The company performed and presented a probabilistic sensitivity analysis (PSA) and deterministic sensitivity analysis (DSA) in order to show the uncertainty surrounding the base-case results.

Compared with the deterministic results, the PSA with 1,000 iterations showed lower incremental QALYs and higher incremental costs, which resulted in an increased ICER (£21,601) (Table 5.11). The cost effectiveness acceptability curve in the revised model showed that durvalumab approximately had a 87% and 98% probability of being cost effective at willingness to pay (WTP) thresholds of £30,000 and £50,000 respectively.

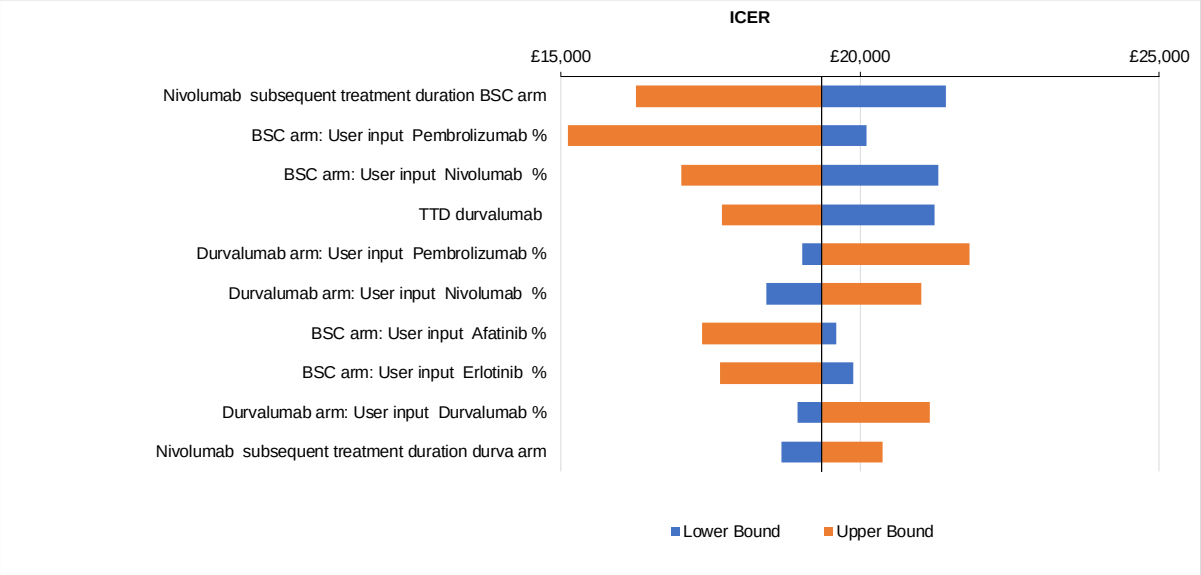
The company conducted DSAs by varying key model parameters between the upper and lower 95% CIs of the expected value used in the deterministic base-case. The ICER was most sensitive to the duration of post-progression immunotherapy use, the percentage of patients receiving subsequent immunotherapy use and the time to discontinuation of durvalumab. In none of the DSAs the ICER exceeded the WTP thresholds of either £30,000 or £50,000 (Figure 5.5).

Table 5.11: Company's revised base-case results (probabilistic, 1,000 iterations)

	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
Durvalumab	[redacted]	[redacted]	--	--	--
SoC	[redacted]	[redacted]	[redacted]	2.67	£21,601

Source: Based on the revised PSA results in the economic model.
ICER = incremental cost effectiveness ratio; QALY = quality adjusted life year; SoC =standard of care

Figure 5.5: Tornado diagram presenting the results of the deterministic sensitivity analysis (revised)



Source: Based on the revised tornado diagram presented in the model.

Scenario analyses

The company conducted several scenario analyses, which are shown in Table 5.12 below. The results showed ICERs ranging between £11,368 and £30,629 per QALY gained, excluding the scenario analyses with a 10, 20, and 30 years' time horizon. Apart from different scenarios for the time horizon, the three most influential scenarios that increased the ICER were a shorter treatment waning cut-off (three years: £30,629 and five years: £24,391), increased cost for PF health state (£24,069), and an alternative PFS distribution (independent models; Gompertz: £23,237). The three most influential scenarios that decreased the ICER were different subsequent immunotherapy durations (two years for pembrolizumab and nivolumab: £11,369) and alternative EQ-5D-5L utility values (PACIFIC, EQ-5D-5L ITT: £17,960, PACIFIC, EQ-5D-5L PD-L1 ≥ 1%: £18,162). Scenarios with shorter time horizons of 10, 20 and 30 years increased the ICER to respectively £39,161, £23,099 and £20,001.

Table 5.12: Results of the scenario analyses conducted by the company (revised)

Scenario	Values	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
Base-case	-	[redacted]	2.93	£19,366
Time horizon	10 years	[redacted]	1.43	£39,161
	20 years	[redacted]	2.45	£23,099
	30 years	[redacted]	2.84	£20,001
Alternative PFS distributions	Independent models; Gompertz	[redacted]	2.48	£23,350
	Proportional hazards; Gompertz	[redacted]	2.86	£20,078
Alternative PPS distributions	Weibull	[redacted]	2.96	£19,220
	PPS fitted by treatment arm (exponential for both arms)	[redacted]	3.14	£18,375
Parametric analyses excluding numbers of risk < 5	-	[redacted]	2.96	£19,204
Treatment waning cut-off	3 years	[redacted]	1.94	£30,629
	5 years		2.39	£24,391
	Lifetime		3.06	£18,415
Subsequent treatment	Alternative PPS curve (KEYNOTE-024), 39% subsequent immunotherapy use in SoC arm and 20% in durvalumab arm	[redacted]	2.58	£21,297
	Alternative PPS curve (KEYNOTE-024), 0% subsequent immunotherapy use in both arms	[redacted]	2.80	£22,792
	Alternative PPS curve (KEYNOTE-024), 20% subsequent immunotherapy use in both arms	[redacted]	2.75	£22,404
	Alternative PPS curve (KEYNOTE-024), + 30% subsequent immunotherapy use in SoC arm and 2% in durvalumab arm	[redacted]	2.56	£20,985

Scenario	Values	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
	Alternative PPS curve (KEYNOTE-024), + 60% subsequent immunotherapy use in SoC arm, 2% in durvalumab arm	[redacted]	2.27	£19,011
Utility approach	Time to death and progression	[redacted]	2.95	£19,280
	Inclusion of age parameter	[redacted]	2.81	£20,237
	PF utilities capped at general population levels (PF = 0.79, PD = 0.756)	[redacted]	2.86	£19,853
	Include AE dis-utilities	[redacted]	2.93	£19,365
	20% decrease in HRQoL upon progression (PF = 0.81, PD = 0.65)	[redacted]	3.00	£18,961
	PACIFIC PF EQ-5D-5L data ITT (PF = 0.874, 0.842)	[redacted]	3.16	£17,960
	PACIFIC PF EQ-5D-5L data PD-L1 ≥ 1% (PF = 0.865, 0.840)	[redacted]	3.13	£18,162
	Progression and treatment arm included (mixed model)	[redacted]	2.82	£20,172
	Mean utility scores by treatment arm (EQ-5D-3L)	[redacted]	2.80	£20,261
Vial sharing	No vial sharing	[redacted]	2.93	£23,020
	50% vial sharing	[redacted]	2.93	£21,193
Subsequent treatment costs	50% discount for all subsequent treatments, where applicable (pembrolizumab nivolumab, erlotinib, afatinib)	[redacted]	2.93	£20,744
Subsequent immunotherapy duration	Pembrolizumab and nivolumab: two years duration	[redacted]	2.93	£11,368
Increased cost for progression-free health state	Extreme scenario: cost of metastatic disease applied to stage III	[redacted]	2.93	£24,069
Lower costs for progressed disease health state	Extreme scenario: Reduction in costs of 64%	[redacted]	2.93	£19,457

Source: Based on Table 32 of the company's response to the clarification letter[24]
AE =adverse events; HRQoL = health-related quality of life; ICER = incremental cost effectiveness ratio; ITT = intention to treat; PD = progressed disease; PF = progression-free; PFS = progression-free survival; PPS = post-progression survival QALY = quality adjusted life year; SoC = standard of care

ERG comment: The ERG had minor concerns regarding a) the difference in incremental QALYs in the PSA results compared to the deterministic analysis as well as stability of PSA results, b) the inclusion of patient characteristics in the PSA, and c) errors in the company's scenario analyses.

a) Compared with the company's deterministic base-case results, probabilistic incremental QALYs were lower. The ERG agrees with the company's explanation in the response to clarification letter that

that this was likely driven by the skewedness of the generalised gamma PFS curve. Furthermore, the 1,000 iterations used in the PSA did not achieve stability of results and the ERG used 5,000 iterations.

b) The company included patient characteristics in their PSA, despite intending to exclude them (the model setting designed to exclude them did not work). Given that these parameters reflected first order uncertainty, these should not be incorporated in the PSA. This was corrected in the ERG base-case.

c) The ERG identified several errors in the company's scenario analyses, including several settings in the controls sheet that were not functioning, and incorrect results of scenario analyses in Table 55 of the CS. In response to the clarification letter, the company corrected all of these errors. However, the ERG was unable to reproduce the company's scenario in which the costs for progressed disease reduced by 64% and noted additional errors in the subsequent treatment scenarios. The ERG only presented the revised results of the scenario analyses that were corrected by the company.

5.2.12 Model validation and face validity check

Face validity and internal validity
The model was reviewed by health economists within the company who performed face validity and internal validity checks. A third-party vendor also checked the model for basic validity of model outcomes, application and sources of costs and utilities, clinical inputs, sensitivity analyses and macros. In addition, model structure and approach (partitioned survival vs. semi-Markov) was reviewed by an expert in the field who advised on most appropriate methodology.

Cross validity
No cross validity checking of the model was reported by the company.

External validity
OS predictions from the model for durvalumab and SoC were validated against other sources. OS for SoC was compared to relevant clinical trials, UK real-world data, and clinical expert opinion. OS for durvalumab was compared to OS as observed in PACIFIC (see Table 35 of CS[1]). The company concluded that modelled OS for SoC broadly matched survival from all available sources of evidence, although none of these sources provided any estimates beyond a five year time horizon (except one 10 year estimate from expert opinion by four clinical experts[118]). The company did not state anything about the comparability of modelled OS for durvalumab with PACIFIC data. From Table 35 of the CS it can be seen that from the first to the third year, modelled OS for durvalumab goes from 1% underestimation (86% predicted vs 87% observed) to 3% overestimation (63% predicted vs 60% observed).[1]

In addition, OS for both durvalumab and SoC was compared to values accepted by NICE for immunotherapies in the advanced metastatic NSCLC setting (see Table 36 of CS[1]). Modelled OS was substantially higher (in both durvalumab and SoC) than these comparator values. The company stressed that these studies concerned distinct populations and disease stages, and therefore the predicted effect could be considered in line with that seen for other immunotherapies, when accounting for the greater potential for long-term survival when treating with curative intent.

Predictive validity
No predictive validity checking was reported by the company.

ERG comment: The main concern of the ERG relates to the fact that no firm conclusion could be drawn from the external validation exercised by the company. The company stated that modelled OS broadly matched survival of the available sources, but this was a subjective observation. In addition,

there was a marked difference between model survival and survival in previous TAs, but there is no way of telling whether the differences between model predictions and other immunotherapy values in Table 35 of CS were caused by the differences in population (metastatic vs. curative intent) or whether these were partly (or largely) caused by poor external validity of the current model.[1] The ERG appreciates the fact that durvalumab is first in class and so all comparison is difficult, but the ERG also considers the company's claims that any model outcome can be considered 'in line' with previous findings not to be substantiated.

5.3 Exploratory and sensitivity analyses undertaken by the ERG

Table 5.13 summarises the main issues highlighted by the ERG in Section 5.2, indicates the expected direction of bias introduced by these issues and whether these are examined in any analyses/incorporated in the ERG base-case.

Table 5.13: Main ERG critique of company's submitted economic evaluation

Issue	Likely direction of bias introduced in ICER[a]	ERG analyses	Addressed in company analysis?
Model structure (Section 5.2.2)			
State transition model instead of partitioned survival analysis	+/-	No	Requested but not provided
Population, interventions and comparators, perspective and time horizon (Sections 5.2.3-5.2.5)			
The PD-L1 ≥ 1% population is a post-hoc subgroup of the population in the scope and the ITT population in PACIFIC	+/-	No	No
Half-cycle correction applied, but not to treatment and administration costs	+	Base-case (FV)	No
Treatment effectiveness and extrapolation (Section 5.2.6)			
Treatment effectiveness based on post-hoc subgroup (PD-L1 ≥ 1% patients), and not pre-specified TTP and PPS analysis	+/-	No	No
Evidence based on patients with mostly ≥2 overlapping cycles of platinum-based CRT, but UK practice is mostly sequential cycles	+/-	No	No
Survival evidence from PACIFIC is immature	+/-	No	No
Durvalumab PFS (extrapolated with generalised gamma) likely over-estimated compared to evidence from PACIFIC	+	Base-case (MJ), and scenarios	Scenario, but second-best fitting with better external validity unexplored
Treatment waning effect after 10-year cut-off	+	Base-case (MJ), and scenarios	Scenarios, where alternative cut-offs are explored
Age calculations performed incorrectly	+	Base-case (FE)	Addressed in response to request for clarification letter, in revised model
Uncertainty about PPS (driven by data source, modelling method and subsequent treatments)	+/-	Scenarios	Scenarios, explored using alternative data sources
Adverse events (Section 5.2.7)			
Treatment-related AEs potentially under-estimated	+	Scenario	No

Issue	Likely direction of bias introduced in ICER[a]	ERG analyses	Addressed in company analysis?
Health-related quality of life (Section 5.2.8)			
Utility scores for the PF state is likely high (0.810)	+	Scenario	No
No age-related utility decrement	+	Base-case (MJ)	Scenario uses age decrement
Utility decrement for PD state is likely small (-0.034)	-	Scenario	Yes
Utility estimates treatment-independent	+	Base-case (MJ)	Scenario in response to request for clarification[24]
Impact of AEs on HRQoL not reflected	+/-	Scenario	Scenario
Resources and costs (Section 5.2.9)			
Perfect vial-sharing assumption not appropriately justified and likely unrealistic	+	Base-case (FV), scenarios	Yes, scenarios allow for imperfect vial sharing
Resource use for PD health state based on metastatic disease	+	No	Scenario in response to request for clarification[24]
Inclusion criterion for subsequent treatments (>3% in all patients) may lead to biased inclusion per treatment arm	+/-	No	No
Cost effectiveness analyses (Sections 5.2.10 and 5.2.11)			
Patient characteristics included in PSA	+/-	Base-case (FV)	No

Footnotes: [a] Likely conservative assumptions (of the intervention versus all comparators) are indicated by '-'; while '+/-' indicates that the bias introduced by the issue is unclear to the ERG and '+' indicates that the ERG believes this issue likely induces bias in favour of the intervention versus at least one comparator.
ERG = Evidence Review Group; FE = Fixing errors; FV = fixing violations; ICER = incremental cost effectiveness ratio; MJ = matters of judgement; PD = progressed disease; PSA = probabilistic sensitivity analysis

Based on all considerations in Section 5.2 (summarised in Table 5.13), the ERG defined a new base-case. This base-case included multiple adjustments to the original base-case presented in the previous sections. These adjustments made by the ERG form the ERG base-case and were subdivided into three categories (derived from Kaltenthaler 2016[119]):

- Fixing errors (correcting the model where the company's submitted model was unequivocally wrong)
- Fixing violations (correcting the model where the ERG considered that the NICE reference case, scope or best practice had not been adhered to)
- Matters of judgement (amending the model where the ERG considers that reasonable alternative assumptions are preferred)

Fixing errors

1. Incorrect age calculations (Section 5.2.11).
   The ERG used the company's revised model in response to request for clarification[24], in which the error was corrected, as requested.
2. Incorrect vial wastage calculations for nivolumab and pembrolizumab
   The ERG corrected the error by assuming perfect vial sharing throughout for nivolumab and pembrolizumab (given that these are now offered at fixed dosages)
3. Incorrect implementation in PSA of utility decrements for progression and treatment. IN probabilistic mode the minus sign was lost, turning the decrements into increments. The ERG corrected the error.

Fixing violations

4. Half-cycle correction not applied to treatment and administration costs (Section 5.2.5).
   The ERG corrected this.
5. Perfect vial sharing assumption lacks plausibility.
   The ERG assumed no vial sharing.
6. Patient characteristics included in the PSA (Section 5.2.11).
   The ERG corrected this.

Matters of judgment

7. Durvalumab PFS likely over-estimated using the generalised gamma (Section 5.2.6).
   The ERG used the lognormal instead for durvalumab PFS (and also TTP, by company's default setting).
8. Treatment waning effect after 10-year cut-off (Section 5.2.6).
   The ERG used a five-year cut-off instead.
9. No age-related utility decrement used (Section 5.2.8).
   The ERG applied an age-related utility decrement.
10. Treatment was excluded from utility mixed effects model (Section 5.2.8).
    The ERG included treatment as a covariate in the utility mixed effects model.

Table 6.1 shows how individual adjustments impact the results plus the combined effect of all abovementioned adjustments simultaneously, resulting in the (deterministic) ERG base-case. The FV and MJ ERG analyses were performed also incorporating the 'fixing error' adjustments given that the ERG considered that the 'fixing error' adjustments corrected unequivocally wrong issues.

5.3.1 ERG base-case results

The results of the deterministic ERG base-case showed that incremental costs were [redacted] and incremental QALYs were 1.32 (Table 5.14). This resulted in an ICER of £50,238, which was mainly

driven by using the lognormal distribution for extrapolation of durvalumab PFS, using a five-year cut-off for treatment waning effect, and assuming no vial sharing (see Table 6.1).

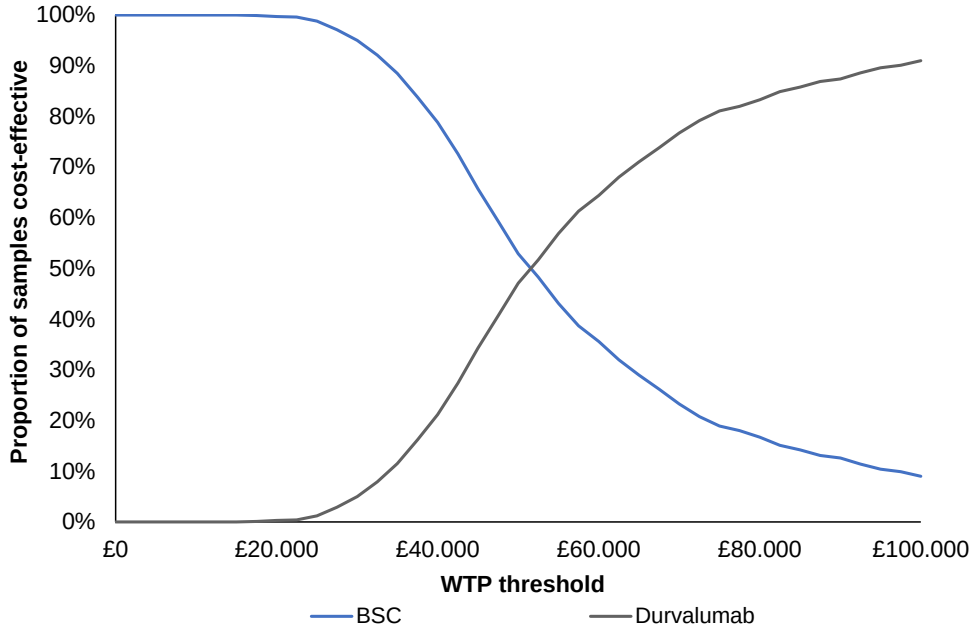
Compared with the deterministic ERG base-case results, the ERG PSA with 5,000 iterations resulted in lower incremental QALYs and slightly lower incremental costs, which resulted in an increased ICER (£52,353). The company's base-case also showed a marked difference between the deterministic and probabilistic ICERs. In their response to clarification question B24d[24] the company argued that this difference was due to the skewedness of the generalized gamma PFS curve, which caused skewed QALY results, but slightly differently so for durvalumab and SoC. At a later stage, the ERG noted an error in the model in the implementation of the utility decrements for progression and treatment, turning these into increments when running the PSA. The ERG fixed this for the ERG analyses. The company's probabilistic ICER results still contain the error but as no treatment decrement was applied in the company base-case, only the effect of the progression decrement remains. The cost effectiveness acceptability curve showed that durvalumab approximately had a 5.0% and 47.1% probability of being cost effective at willingness to pay (WTP) thresholds of £30,000 and £50,000 respectively (Figure 5.6).

Table 5.14: ERG base-case results

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
Deterministic ERG base-case					
Durvalumab	[redacted]	[redacted]	[redacted]	1.32	£50,238
SoC	[redacted]	[redacted]			
Probabilistic ERG base-case					
Durvalumab	[redacted]	[redacted]	[redacted]	1.25	£52,353
SoC	[redacted]	[redacted]			

ERG = Evidence Review Group = ICER = incremental cost effectiveness ratio; QALY = quality-adjusted life year; SoC = standard of care

Figure 5.6: ERG base-case cost effectiveness acceptability curve



5.3.2 Additional exploratory analyses performed based on the ERG base-case

Additional sensitivity analyses were performed to examine the potential impact of alternative assumptions on the cost effectiveness estimates. These were all performed using the ERG base-case. Results are presented in Table 6.2 in Section 6.

Exploratory analyses using the ERG base-case:

0. ERG base-case with no treatment waning effect
1. Alternative PFS distributions: generalised gamma for both durvalumab and SoC
2. Alternative PFS distributions: lognormal for both durvalumab and SoC
3. Treatment waning effect at three years with a) durvalumab and SoC PFS modelled using lognormal and generalised gamma and b) durvalumab and SoC PFS modelled using lognormal
4. Treatment waning effect seven years with a) durvalumab and SoC PFS modelled using lognormal and generalised gamma and b) durvalumab and SoC PFS modelled using lognormal
5. Use PPS based on PACIFIC, but using generalised gamma
6. Use PPS based on KEYNOTE-024, but correcting the following errors:
    - corrected the choice of survival distribution to the one with the best statistical fit (lognormal instead of log-logistic)
    - corrected estimate of patients receiving subsequent IO treatment in the durvalumab arm (27%)
7. Adverse events incorporated with amended incidence and including impact on HRQoL
8. Alternative utility score for PF state from literature (0.73)
9. Alternative utility scores for PF state (0.73) and PD state (0.67) from literature
10. Vial sharing possible (30%)

5.3.3 Subgroup analyses performed based on the ERG base-case

No subgroup analyses were performed.

5.4 Conclusions of the cost effectiveness section

- The company submission and response to clarification provided sufficient details for the ERG to appraise the literature searches and the 2018 update searches. A good range of databases and additional resources were searched.

- The company submission was largely in line with the NICE reference case. The modelled population, however, was narrower than that in the scope, but in line with the anticipated marketing authorisation (focussing on the subgroup with PD-L1 tumour expression ≥1%).

- Given the immaturity of the survival data in the PACIFIC subpopulation, the ERG had concerns about the appropriateness of the semi-Markov approach and its superiority over a partitioned survival model approach and would have liked to see both approaches appropriately explored.

- The ERG had concerns about the model being based on treatment effectiveness estimates derived from a post-hoc subgroup analysis and post-hoc TTP and PPS analyses, as well as in a population with mostly prior overlapping CRT instead of sequential CRT, although any bias introduced by this remained unclear.

- The main concern of the ERG was that it considered modelled long-term durvalumab PFS as highly uncertain and likely over-estimated, due to a model choice (generalised gamma) that resulted in PFS estimates above those observed in PACIFIC at three years, with the PACIFIC estimate notably being derived from small numbers of patients at risk and immature data. This model choice caused

ICERs of durvalumab versus SoC to be lower than other model choices. This issue was exacerbated by the choice of time-point at which treatment waning was modelled (10 years), which was deemed by the ERG as highly uncertain, not appropriately validated, and potentially late, additionally lowering ICERs of durvalumab versus SoC in the CS. Alternative modelling methods, such as spline-based models, remained unexplored in the CS.

- There was a discrepancy between AEs causally related to treatment in PACIFIC, which were mostly higher for the durvalumab arm than in the placebo arm, and AE incidence in the model, which was comparable between treatments, that remained unexplained, likely lowering ICERs of durvalumab versus SoC.

- The ERG considered utility values for both (progression-free and progressed disease) health states to be potentially over-estimated, being comparable to those in the general population and not adjusted by general population age utility estimates. Excluding treatment as a factor in utility estimation and excluding the HRQoL impact of AEs contributes to QALY gains being likely over-estimated. These assumptions on balance likely lowered ICERs of durvalumab versus SoC.

- The ERG considered the assumption of perfect vial sharing to be unrealistic in this setting, given the limited number of patients in England and Wales that would be eligible for treatment with durvalumab. This assumption caused the ICER of durvalumab against SoC to be lower than alternative assumptions.

- The ERG made various adjustments to the company base-case. The probabilistic ERG base-case ICER of durvalumab versus SoC was £52,353 per QALY gained (based on 5,000 iterations). The difference was likely caused by the skewedness of distributions used for modelling PFS.

- Deterministic scenario analyses were performed to examine the potential impact of alternative assumptions on the cost effectiveness estimates. Alternative assumptions regarding PFS and treatment waning effects had the most significant impact. The scenario in which PFS distributions were changed into generalised gamma for both arms (as per the company's base-case) resulted in an ICER of £29,302 per QALY gained, whereas applying lognormal distributions for PFS in both arms drove up the ICER to £52,300 per QALY gained. Applying the company's KEYNOTE-024 PPS scenario, but with errors corrected, resulted in an ICER of £59,131 per QALY gained. The scenario exploring a treatment waning effect with three-year cut-off and using the lognormal distribution for both durvalumab and SoC PFS increased the ICER the most (to £64,531 per QALY gained). All other scenarios had a relatively modest impact (<£5,000) on the ERG base-case ICER.

- In conclusion, given that the ERG base-case ICER was estimated to be substantially above £40,000 per QALY gained, and only one scenario resulting in ICERs slightly below £30,000 per QALY gained, and the large uncertainty induced by mainly the immature survival data, uncertainty around the cost effectiveness of durvalumab is substantial.

6. Impact on the ICER of additional clinical and economic analyses undertaken by the ERG

6.1 Analyses undertaken by the ERG

In Section 5.3 the ERG base-case was presented, which was based on various changes compared to the company base-case. Table 6.1 shows how individual changes impact the results plus the combined effect of all changes simultaneously. The exploratory scenario analyses are presented in Table 6.2. These are all conditional on the ERG base-case. The analyses numbers in Tables 6.1 and 6.2 correspond to the analyses numbers reported in Section 5.3. The submitted model file contains technical details on the analyses performed by the ERG (e.g. the "ERG" sheet provides an overview of the cells that were altered for each adjustment).

Table 6.1: Deterministic ERG base-case

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
CS original base-case					
Durvalumab	[redacted]	[redacted]	[redacted]	2.94	£19,320
SoC	[redacted]	[redacted]			
Fixing error (1, age calculations)					
Durvalumab	[redacted]	[redacted]	[redacted]	2.93	£19,366
SoC	[redacted]	[redacted]			
Fixing violations (3, half-cycle correction for treatment costs)					
Durvalumab	[redacted]	[redacted]	[redacted]	2.93	£20,001
SoC	[redacted]	[redacted]			
Fixing violations (4, no vial sharing) and error 2 (vial wastage)					
Durvalumab	[redacted]	[redacted]	[redacted]	2.93	£20,647
SoC	[redacted]	[redacted]			
Matter of judgement (6, lognormal for durvalumab PFS)					
Durvalumab	[redacted]	[redacted]	[redacted]	1.32	£45,878
SoC	[redacted]	[redacted]			
Matter of judgement (7, treatment waning at 5 years)					
Durvalumab	[redacted]	[redacted]	[redacted]	2.39	£24,391
SoC	[redacted]	[redacted]			
Matter of judgement (8, age-related utility decrement applied)					
Durvalumab	[redacted]	[redacted]	[redacted]	2.81	£20,237
SoC	[redacted]	[redacted]			
Matter of judgement (9, treatment included in utility model)					
Durvalumab	[redacted]	[redacted]	[redacted]	2.82	£20,172
SoC	[redacted]	[redacted]			
ERG base-case					
Durvalumab	[redacted]	[redacted]	[redacted]	1.32	£50,238
SoC	[redacted]	[redacted]			

CS = company submission; ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; PFS =progression-free survival; QALY = quality-adjusted life year; SoC = standard of care

Table 6.2: Deterministic scenario analyses conditional on ERG base-case

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
ERG base-case					
Durvalumab	[redacted]	[redacted]	[redacted]	1.32	£50,238
SoC	[redacted]	[redacted]			
ERG base-case, no treatment waning effect (0)					
Durvalumab	[redacted]	[redacted]	[redacted]	1.10	£60,928
SoC	[redacted]	[redacted]			
Alternative PFS distributions both arms, generalised gamma (1)					
Durvalumab	[redacted]	[redacted]	[redacted]	2.19	£29,302
SoC	[redacted]	[redacted]			
Alternative PFS distributions both arms, lognormal (2)					
Durvalumab	[redacted]	[redacted]	[redacted]	1.27	£52,300
SoC	[redacted]	[redacted]			
Treatment waning at 3 years, PFS as ERG base-case (3a)					
Durvalumab	[redacted]	[redacted]	[redacted]	1.35	£48,766
SoC	[redacted]	[redacted]			
Treatment waning at 3 years, PFS as scenario 2 (3b)					
Durvalumab	[redacted]	[redacted]	[redacted]	1.04	£64,531
SoC	[redacted]	[redacted]			
Treatment waning at 7 years, PFS as ERG base-case (4a)					
Durvalumab	[redacted]	[redacted]	[redacted]	1.25	£52,833
SoC	[redacted]	[redacted]			
Treatment waning at 7 years, PFS as scenario 2 (4b)					
Durvalumab	[redacted]	[redacted]	[redacted]	1.41	£47,000
SoC	[redacted]	[redacted]			
PACIFIC PPS, but generalised gamma (5)					
Durvalumab	[redacted]	[redacted]	[redacted]	1.33	£49,868
SoC	[redacted]	[redacted]			
Company's KEYNOTE-024 PPS scenario, with errors corrected (6)					
Durvalumab	[redacted]	[redacted]	[redacted]	1.10	£59,131
SoC	[redacted]	[redacted]			
Adverse events with amended incidence and including impact on HRQoL (7)					
Durvalumab	[redacted]	[redacted]	[redacted]	1.32	£50,288
SoC	[redacted]	[redacted]			
Alternative PF utility score (8)					
Durvalumab	[redacted]	[redacted]	[redacted]	1.42	£51,805
SoC	[redacted]	[redacted]			
Alternative PF and PD utility scores (9)					
Durvalumab	[redacted]	[redacted]	[redacted]	1.28	£51,587

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
SoC	[redacted]	[redacted]			
Vial sharing possible at 30% (10)					
Durvalumab	[redacted]	[redacted]	[redacted]	1.32	£49,350
SoC	[redacted]	[redacted]			

ERG = Evidence Review Group; HRQoL = health-related quality of life; ICER = incremental cost effectiveness ratio; PD = progressed disease; PF = progression-free; PFS = progression-free survival; PPS = post-progression survival; QALY = quality-adjusted life year; SoC = standard of care

7. End of life

NICE end of life considerations apply when two criteria are satisfied:

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

Table 27 of the CS summarises available data which might support the applicability of end-of-life criteria.[1] The data are summarised below.

Criterion 1: According to the CS, "in the PACIFIC study, median OS in the placebo arm was 28.7 months in the ITT population, and 29.1 months in the PD-L1 ≥1% group of patients".[1] However, the company highlighted that *"data may not reflect real-world survival outcomes in the UK cohort of locally-advanced, unresectable, Stage III NSCLC patients* …(as)*…intensive management of patients and other factors that are unique to a clinical trial setting can improve patient outcomes relative to what is known / expected in real-world settings"*.[1]

In support of this statement, the CS presented some UK-specific data:

- National Lung Cancer Audit (2016 audit period){Royal College of Physicians, 2018 [accessed 13.2.18] #399}: Average 1-year survival rate from diagnosis of stage III patients = 42.5%
- Moller et al. 2017{Moller, 2018 #56}: Patients treated with radical radiotherapy with 2-year survival probability from diagnosis <25%
- Royal College of Radiologists audit{Royal College of Radiologists (RCR), 2016 #432}: Median OS following radical radiotherapy = 22 months, 2-year survival rate = 44%; 2-year survival rate (overlapping CRT) = 46%
- Public Health England{AstraZeneca, 2018 #435}: Median OS for patients with unresected stage III who had received overlapping CRT = 20.7 months
- SOCCAR RCT{Maguire, 2014 #51}: Median OS from start of overlapping / sequential CRT = 24.3 / 18.4 months
- Expert opinion (mean of 10 responses){AstraZeneca, 2018 #402} = 22.3 months (median OS)

ERG comment: While this claim is plausible, it should be noted that for NICE committees mean values are preferable to median values when measuring OS time.{National Institute for Health and Care Excellence, 2017 [accessed 4.12.18] #436} Therefore, the extent of the possible effect is unclear, i.e. whether the reported data (including median OS) could indicate that patients would have a life expectancy of less than 24 months (mean OS). There is additional uncertainty due to the immaturity of the OS data reported in PACIFIC.

Criterion 2: The company highlights that PACIFIC found *"significantly extended OS relative to placebo in the PD-L1 ≥1% group"*, presenting two different estimates: HR 0.54, 95% CI 0.35 to 0.81, section B.2.6; and HR 0.53, 95% 0.36 to 0.77, Table 27.[1]

ERG comment: There is insufficient evidence whether the treatment offers an extension to life as no OS estimate is reported for the durvalumab arm in the relevant PD-L1 ≥1% subgroup (Table 4.6). However, it should be noted that in the whole trial population, a difference of median survival time of 12 months can be seen when comparing the lower 95% CIs (Table 4.6). However, this again is based on median survival time (when normally mean is preferable) and is unlikely to be *"sufficiently robust"*.{National Institute for Health and Care Excellence, 2013 [accessed 4.12.18] #19} Furthermore,

there is additional uncertainty due to the immaturity of the OS data reported in PACIFIC; results in the relevant subgroup might become available in future analyses.

8. Overall conclusions

8.1 Statement of principal findings

The CS comprised of a systematic review of the evidence for durvalumab for the treatment of locally advanced unresectable, Stage III NSCLC in adults whose tumours express PD-L1 on ≥1% of tumour cells and whose disease has not progressed after platinum-based CRT.

The main database searches in the CS were on the whole transparent and reproducible, and a good range of resources were used. Better use of synonyms could have been applied in some database searches to aid the retrieval of relevant references. The presented evidence included one RCT, PACIFIC.

The PACIFIC trial included patients with confirmed PD-L1 expression on ≥1% of TCs. After randomisation procedures, this resulted in 476 patients in the durvalumab group and 237 patients in the placebo group. However, when focusing on only the patients with confirmed PD-L1 expression on ≥1% of TCs, there were 212 patients in the durvalumab group and 91 patients in the placebo group. While the PACIFIC trial met a multi-centre, international design, only eight patients were seeking treatment in the UK.

Based on the PACIFIC data there appears to be a benefit in both PFS and OS for durvalumab patients compared with placebo patients, however, the data are immature and there remains substantial uncertainty about the comparative effectiveness.

Common adverse events were reported in both the durvalumab and placebo groups. The common AEs in the durvalumab patients included cough, fatigue, and radiation pneumonitis, whereas patients in the placebo group also included dyspnoea. Overall, more serious adverse events were reported for durvalumab compared to placebo (64/213 (30%) vs. 18/90 (20%)).

Due to the PACIFIC trial being ongoing, final results will be confirmed at a later date.

Economic evaluation

The ERG made various adjustments to the company's base-case. The probabilistic ERG base-case ICER of durvalumab versus SoC was ££52,353 per QALY gained (based on 5,000 iterations). This was higher than the deterministic ERG base-case ICER of £50,238 per QALY gained. The difference was likely caused by the skewedness of distributions used for modelling PFS.

Deterministic scenario analyses were performed to examine the potential impact of alternative assumptions on the cost effectiveness estimates. Alternative assumptions regarding PFS and treatment waning effects had the most significant impact. The scenario in which PFS distributions were changed into generalised gamma for both arms (as per the company's base-case) resulted in an ICER of £29,302 per QALY gained, whereas applying lognormal distributions for PFS in both arms drove up the ICER to £52,300 per QALY gained. Applying the company's KEYNOTE-024 PPS scenario, but with errors corrected, resulted in an ICER of £59,131 per QALY gained. The scenario exploring a treatment waning effect with three-year cut-off and using the lognormal distribution for both durvalumab and SoC PFS increased the ICER the most (to £64,531 per QALY gained). All other scenarios had a relatively modest impact (<£5,000) on the ERG base-case ICER.

In conclusion, given that the ERG base-case ICER was estimated to be substantially above £40,000 per QALY gained, and only one scenario resulting in ICERs slightly below £30,000 per QALY gained, and the large uncertainty induced by mainly the immature survival data, uncertainty around the cost effectiveness of durvalumab is substantial.

8.2 Strengths and limitations of the assessment

Overall, the CS reported searches were clearly reported and reproducible. The selection of databases searched was adequate and a good range of additional resources were included. Study design limits to identify economic evaluations, utility data, and costs and resource use data were applied. A good range of intervention terms for durvalumab and comparators were used in the cost effectiveness strategy.

A substantial source of uncertainty lies in the generalisability of PACIFIC data to the UK setting, as PACIFIC pertains predominantly to prior overlapping CRT, whereas in clinical practice in the UK, mostly sequential CRT is applied. In addition, the PD-L1≥1% subgroup and TTP and PPS analyses were performed post-hoc. Furthermore, main results come from interim analysis, e.g. according to the response to the request for clarification the company *"will conduct a final analysis of OS once a maturity of 70% has been reached. OS is an event-driven endpoint; therefore, the timing of this analysis is uncertain"* which will be reached *"when 491 OS events have occurred"*.[1, 24] The current maturity in the relevant subgroup is 33.0% for durvalumab and 49.5% for placebo.[27]

The company submission was largely in line with the NICE reference case. Utility scores were estimated using a mixed effects model based on observed EQ-5D data in the PACIFIC study.

The model was, in general, well-built and transparent. Apart from their base-case, the company provided opportunities for exploratory analyses using alternative data derived from clinical trials in similar populations.

A main limitation was the immaturity of survival data in the PACIFIC subpopulation, and the inherent uncertainty in PFS and PPS extrapolations. The ERG considers particularly durvalumab PFS to be overestimated, even more so because the company chose to incorporate treatment waning only at 10 years. Given the immaturity of survival data, the ERG also has concerns on the appropriateness of the semi-Markov model structure, but the company did not provide an opportunity to explore a partitioned survival approach. Alternative modelling methods, such as spline-based models, remained unexplored in the CS.

Lastly, the utility scores used in the model do not seem representative of the patient population. The ERG considers the utilities for both (progression-free and progressed disease) health states to be an overestimate.

8.3 Suggested research priorities

PACIFIC is an ongoing trial so more information will be available to reduce the uncertainties in progression-free and overall survival, and other outcomes.

9. References

[1] AstraZeneca. *Durvalumab for treatment of locally advanced, unresectable, stage III non-small cell lung cancer in adults whose tumours express PD-L1 on >=1% of tumour cells and whose disease has not progressed after platinum-based chemo-radiation therapy (ID1175): Submission to National Institute of Health and Care Excellence. Single technology appraisal (STA): Document B*: AstraZeneca, 2018. 201p.

[2] Office for National Statistics. *Cancer registration statistics, England: first release, 2016*: Office for National Statistics, 2018 [accessed 14.2.18] Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/cancerregistrationstatisticsengland/2016

[3] Office for National Statistics. *Deaths registered in England and Wales (series DR) 2016*: Office for National Statistics, 2017 [accessed 16.2.18] Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/datasets/deathsregisteredinenglandandwalesseriesdrreferencetables

[4] American Joint Committee on Cancer (AJCC). *Cancer staging manual (7th edition)*, 2009 [accessed 7.3.18] Available from: https://cancerstaging.org/references-tools/deskreferences/Pages/default.aspx

[5] Detterbeck FC, Boffa DJ, Kim AW, Tanoue LT. The eighth edition lung cancer stage classification. *Chest* 2017;151(1):193-203.

[6] National Institute for Health and Care Excellence. *Lung cancer: diagnosis and management: NICE technology appraisal guidance 121 [Internet]*. London: NICE, 2018 [accessed 17.2.18] Available from: https://www.nice.org.uk/guidance/TA121

[7] Royal College of Physicians. *National Lung Cancer Audit Annual report 2017 (for the audit period 2016)*: Royal College of Physicians, 2018 [accessed 13.2.18] Available from: https://www.hqip.org.uk/public/cms/253/625/19/1068/NLCA2017AnnualReport_web_FINAL.PDF?realName=woLvXS.pdf&v=0

[8] Goldstraw P, Chansky K, Crowley J, Rami-Porta R, Asamura H, Eberhardt WE, et al. The IASLC Lung Cancer Staging Project: Proposals for Revision of the TNM Stage Groupings in the Forthcoming (Eighth) Edition of the TNM Classification for Lung Cancer. *J Thorac Oncol* 2016;11(1):39-51.

[9] Chang JY, Kestin LL, Barriger RB, Chetty IJ, Ginsburg ME, Kumar S, et al. ACR Appropriateness Criteria(R) nonsurgical treatment for locally advanced non-small-cell lung cancer: good performance status/definitive intent. *Oncology (Williston Park)* 2014;28(8):706-10, 712, 714 passim.

[10] Eberhardt WE, De Ruysscher D, Weder W, Le Pechoux C, De Leyn P, Hoffmann H, et al. 2nd ESMO Consensus Conference in Lung Cancer: locally advanced stage III non-small-cell lung cancer. *Ann Oncol* 2015;26(8):1573-88.

[11] London Cancer Alliance (LCA). LCA lung cancer clinical guidelines. 2013 [accessed 24.7.18].

[12] Lim E, Baldwin D, Beckles M, et al. Guidelines on the radical management of patients with lung cancer. *Thorax* 2010;65(Suppl 3):iii1-27.

[13] Royal College of Radiologists (RCR). Lung cancer. In: *Radiotherapy dose fractionation*. Second ed, 2016: 50-8.

[14] Rizvi NA, Peters S. Immunotherapy for unresectable stage III non-small-cell lung cancer. *N Engl J Med* 2017;377(20):1986-1988.

[15] Center B, Petty WJ, Ayala D, Hinson WH, Lovato J, Capellari J, et al. A phase I study of gefitinib with concurrent dose-escalated weekly docetaxel and conformal three-dimensional thoracic radiation followed by consolidative docetaxel and maintenance gefitinib for patients with stage III non-small cell lung cancer. *J Thorac Oncol* 2010;5(1):69-74.

[16] Ramalingam SS, Kotsakis A, Tarhini AA, Heron DE, Smith R, Friedland D, et al. A multicenter phase II study of cetuximab in combination with chest radiotherapy and consolidation chemotherapy in patients with stage III non-small cell lung cancer. *Lung Cancer* 2013;81(3):416-21.

[17] Wozniak AJ, Moon J, Thomas CR, Jr., Kelly K, Mack PC, Gaspar LE, et al. A pilot trial of cisplatin/etoposide/radiotherapy followed by consolidation docetaxel and the combination of bevacizumab (NSC-704865) in patients with inoperable locally advanced stage iii non-small-cell lung cancer: SWOG S0533. *Clin Lung Cancer* 2015;16(5):340-7.

[18] Auperin A, Le Pechoux C, Rolland E, Curran WJ, Furuse K, Fournel P, et al. Meta-analysis of concomitant versus sequential radiochemotherapy in locally advanced non-small-cell lung cancer. *J Clin Oncol* 2010;28(13):2181-90.

[19] Greenspoon JN, Ellis PM, Pond G, Caetano S, Broomfield J, Swaminath A. Comparative survival in patients with brain metastases from non-small-cell lung cancer treated before and after implementation of radiosurgery. *Curr Oncol* 2017;24(2):e146-e151.

[20] AstraZeneca. Clinical expert opinion on subsequent therapies and survival outcomes. (DOF-IMF-005-AUG18) 29 August 2018. Data on File. [PDF provided by the Company]. 2018.

[21] Deng L, Liang H, Burnette B, Beckett M, Darga T, Weichselbaum RR, et al. Irradiation and anti-PD-L1 treatment synergistically promote antitumor immunity in mice. *J Clin Invest* 2014;124(2):687-95.

[22] European Medicines Agency (EMA). *Guideline on the evaluation of anticancer medicinal products in man*, 2017 [accessed 23.8.18] Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2017/11/WC500238764.pdf

[23] European Medicines Agency (EMA). *Summary of opinion (initial authorisation): Imfinzi; durvalumab*: European Medicines Agency, 2018 [accessed 8.8.18] Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Summary_of_opinion_-_Initial_authorisation/human/004771/WC500252664.pdf

[24] AstraZeneca. *Durvalumab for maintenance treatment of unresectable non-small-cell lung cancer after platinum-based chemoradiation - Response to request for clarification from the ERG*: AstraZeneca, 2018. 73p.

[25] McGowan J, Sampson M, Lefebvre C. An evidence based checklist for the peer review of electronic search strategies (PRESS EBC). *Evid Based Libr Inf Pract* 2010;5(1):1-6.

[26] National Institute for Health and Care Excellence. *Single Technology Appraisal: company evidence submission template [Internet]*. London: NICE, 2015 [accessed 11.6.15]. 28p. Available from: http://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisals/specification-for-company-submission-of-evidence-2015-version.docx

[27] AstraZeneca. *Durvalumab for treatment of locally advanced, unresectable, stage III non-small cell lung cancer in adults whose tumours express PD-L1 on >=1% of tumour cells and whose disease has not progressed after platinum-based chemo-radiation therapy (ID1175): Submission to National Institute of Health and Care Excellence. Single technology appraisal (STA): Document B: Appendices*: AstraZeneca, 2018. 356p.

[28] Lefebvre C, Manheimer E, Glanville J. 6.4.9 Language, date and document format restrictions (Chapter 6: Searching for studies) [Internet]. In: Higgins JPT, Green S, editors. Version 5.1.0 [updated March 2011]: The Cochrane Collaboration, 2011 [accessed 7.8.17]. Available from: http://handbook.cochrane.org/

[29] AstraZeneca. PACIFIC clinical study protocol 2016.

[30] Pocock SJ, Assmann SE, Enos LE, Kasten LE. Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: current practice and problems. *Stat Med* 2002;21(19):2917-30.

[31] Le Lay K, Myon E, Hill S, Riou-Franca L, Scott D, Sidhu M, et al. Comparative cost-minimisation of oral and intravenous chemotherapy for first-line treatment of non-small cell lung cancer in the UK NHS system. *Eur J Health Econ* 2007;8(2):145-51.

[32] National Institute for Health and Care Excellence. *Pembrolizumab for untreated PD-L1-positive metastatic non-small-cell lung cancer: NICE technology appraisal guidance 531 [Internet]*. London: NICE, 2018 [accessed 7.8.18] Available from: https://www.nice.org.uk/guidance/ta531

[33] National Institute for Health and Care Excellence. *Pembrolizumab for treating PD-L1-positive non-small-cell lung cancer after chemotherapy: NICE technology appraisal guidance 428 [Internet]*. London: NICE, 2017 [accessed 20.3.18] Available from: https://www.nice.org.uk/guidance/TA428

[34] National Institute for Health and Care Excellence. *Pemetrexed for the first-line treatment of non-small-cell lung cancer: NICE technology appraisal guidance 181 [Internet]*. London: NICE, 2009 [accessed 20.3.18] Available from: https://www.nice.org.uk/guidance/TA181

[35] Scottish Medicines Consortium (SMC). *Pemetrexed 100mg, 500mg, powder for concentrate for solution for infusion (Alimta®) SMC No. (531/09)*, 2009 Available from: https://www.scottishmedicines.org.uk/media/2146/pemetrexed_alimta_2nd_resubmission_final_january_2010_for_website.pdf

[36] National Institute for Health and Care Excellence. *Pemetrexed for the maintenance treatment of non-small-cell lung cancer: NICE technology appraisal guidance 190 [Internet]*. London: NICE, 2010 Available from: https://www.nice.org.uk/guidance/TA190

[37] Scottish Medicines Consortium (SMC). *Pemetrexed, 100mg, 500mg powder for concentrate for solution for infusion (Alimta®) SMC No. (642/10)*, 2009 Available from: https://www.scottishmedicines.org.uk/media/2149/pemetrexed_alimta.pdf

[38] National Institute for Health and Care Excellence. *Pemetrexed maintenance treatment for non-squamous non-small-cell lung cancer after pemetrexed and cisplatin: NICE technology appraisal guidance 402 [Internet]*. London: NICE, 2016 [accessed 20.3.18] Available from: https://www.nice.org.uk/guidance/TA402

[39] National Institute for Health and Care Excellence. *Pemetrexed for the treatment of non-small-cell lung cancer: NICE technology appraisal guidance 124 [Internet]*. London: NICE, 2007 Available from: https://www.nice.org.uk/guidance/TA124

[40] Scottish Medicines Consortium (SMC). *Pemetrexed 500mg vial of powder for solution for intravenous infusion (Alimta®) SMC No. (342/07)*, 2008 Available from: https://www.scottishmedicines.org.uk/media/2144/pemetrexed__alimta__2nd_resubmission_final_august_2008doc_for_website.pdf

[41] National Institute for Health and Care Excellence. *Nintedanib for previously treated locally advanced, metastatic, or locally recurrent non* *and small- cell lung cance guidance 347 [Internet]*. London: NICE, 2015 [accessed 20.3.18] Available from: https://www.nice.org.uk/guidance/TA347

[42] Scottish Medicines Consortium (SMC). *Nintedanib 100mg and 150 mg soft capsules (Vargatef®) SMC No. (1027/15)*, 2015 Available from: https://www.scottishmedicines.org.uk/media/2045/nintedanib__vargatef__final_march_2015revised_250315_for_website.pdf

[43] National Institute for Health and Care Excellence. *Ramucirumab for previously treated locally advanced or metastatic non-small-cell lung cancer: NICE technology appraisal guidance 403 [Internet]*. London: NICE, 2016 Available from: https://www.nice.org.uk/guidance/TA403

[44] Scottish Medicines Consortium (SMC). *Pembrolizumab 50mg powder for concentrate for solution for infusion and 25mg/mL concentrate for solution for infusion (Keytruda®) SMC No. (1239/17)*, 2017 Available from: https://www.scottishmedicines.org.uk/media/2142/pembrolizumab_keytruda_final_june_2017_for_website.pdf

[45] Scottish Medicines Consortium (SMC). *Pembrolizumab 50mg powder for concentrate for solution for infusion (Keytruda®) SMC No. (1204/17)*, 2016 Available from: https://www.scottishmedicines.org.uk/media/2141/pembrolizumab_keytruda_final_dec_2016_amended_020117_for_website.pdf

[46] National Institute for Health and Care Excellence. *Nivolumab for previously treated squamous non-small-cell lung cancer: NICE technology appraisal guidance 483 [Internet]*. London: NICE, 2017 [accessed 20.3.18] Available from: https://www.nice.org.uk/guidance/TA483

[47] Scottish Medicines Consortium (SMC). *Nivolumab 40mg/4mL and 100mg/10mL vials of concentrate for solution for infusion (Opdivo®) SMC No. (1144/16)*, 2016 Available from: https://www.scottishmedicines.org.uk/media/2047/nivolumab_opdivo_final_june_2016_for_website.pdf

[48] National Institute for Health and Care Excellence. *Nivolumab for previously treated non-squamous non-small-cell lung cancer: NICE technology appraisal guidance 484 [Internet]*. London: NICE, 2017 Available from: https://www.nice.org.uk/guidance/TA484

[49] Scottish Medicines Consortium (SMC). *Nivolumab 10mg/mL, concentrate for solution for infusion (Opdivo®) SMC No. (1180/16)*, 2016 Available from: https://www.scottishmedicines.org.uk/media/2052/nivolumab_opdivo_final_sept_2016_final_amended_150916_for_website.pdf

[50] National Institute for Health and Care Excellence. *Atezolizumab for treating locally advanced or metastatic non-small-cell lung cancer after chemotherapy: NICE technology appraisal guidance 520 [Internet]*. London: NICE, 2018 Available from: https://www.nice.org.uk/guidance/TA520

[51] Socinski MA, Butts CA, Mitchell P, al. e. Exploration of patient health status as measured by the generic preference-based questionnaire EQ-5D alongside the start trial of tecemotide (L-BLP25) in non-small cell lung cancer. *J Thorac Oncol* 2013;8:S1206-S7.

[52] Griebsch I, Palmer M, Fayers PM, Ellis S. Is progression-free survival associated with a better health-related quality of life in patients with lung cancer? Evidence from two randomised trials with afatinib. *BMJ OPEN* 2014;4(10):e005762.

[53] Blackhall F, Kim DW, Besse B, Nokihara H, Han JY, Wilner KD, et al. Patient-reported outcomes and quality of life in PROFILE 1007: a randomized trial of crizotinib compared with chemotherapy in previously treated patients with ALK-positive advanced non-small-cell lung cancer. *J Thorac Oncol* 2014;9(11):1625-33.

[54] Bodnar C, Ryan J, Green M. Health state utility measured by EQ-5D-5L for EGFRm T790M NSCLC patients treated with osimertinib. *Ann Oncol* 2016;27(suppl_6).

[55] Iyer S, Taylor-Stokes G, Roughley A. Symptom burden and quality of life in advanced non-small cell lung cancer patients in France and Germany. *Lung Cancer* 2013;81(2):288-93.

[56] Nafees B, Lloyd AJ, Dewilde S, Rajan N, Lorenzo M. Health state utilities in non-small cell lung cancer: An international study. *Asia Pac J Clin Oncol* 2017;13(5):e195-e203.

[57] 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.

[58] O'Kane G, Labbe C, Su S, Tse B, Tam V, Tse T, et al. P2.03b-008 The Impact of Brain Metastases and Their Treatment on Health Utility Scores in Molecular Subsets of Lung Cancer Patients. *J Thorac Oncol* 2017;12(1).

[59] Reck M, Taylor F, Penrod JR, DeRosa M, Morrissey L, Dastani H, et al. Impact of Nivolumab versus Docetaxel on Health-Related Quality of Life and Symptoms in Patients with Advanced Squamous Non-Small Cell Lung Cancer: Results from the CheckMate 017 Study. *J Thorac Oncol* 2018;13(2):194-204.

[60] Roughley A, Damonte E, Taylor-Stokes G, Rider A, Munk VC. Impact of Brain Metastases on Quality of Life and Estimated Life Expectancy in Patients with Advanced Non-Small Cell Lung Cancer. *Value Health* 2014;17(7):A650.

[61] Schuette W, Tesch H, Buttner H, Krause T, Soldatenkova V, Stoffregen C. Second-line treatment of stage III/IV non-small-cell lung cancer (NSCLC) with pemetrexed in routine clinical practice: evaluation of performance status and health-related quality of life. *BMC Cancer* 2012;12:14.

[62] Sun YN, Zhang ZY, Zeng YC, Chi F, Jin XY, Wu R. Comparative efficacy of whole-brain radiotherapy with and without elemene liposomes in patients with multiple brain metastases from non-small-cell lung carcinoma. *Curr Oncol* 2016;23(4):e377-82.

[63] van den Hout WB, Kramer GW, Noordijk EM, Leer JW. Cost-utility analysis of short- versus long-course palliative radiotherapy in patients with non-small-cell lung cancer. *J Natl Cancer Inst* 2006;98(24):1786-94.

[64] Nafees B, Stafford MR, Bhalla S, Watkins J. Pcn69 Health Utilities in the Uk for Second Line Advanced Non-Small Cell Lung Cancer (Nsclc) Following Prior Chemotherapy. *Value Health* 2006;9(6):A296-A297.

[65] Grutters JP, Joore MA, Wiegman EM, Langendijk JA, de Ruysscher D, Hochstenbag M, et al. Health-related quality of life in patients surviving non-small cell lung cancer. *Thorax* 2010;65(10):903-7.

[66] Tabberer M, Stamuli E, Walker M, Summerhayes M, Lees M. Pcn74 Utilities Associated with Non-Small Cell Lung Cancer (Nsclc): A Community Study. *Value Health* 2006;9(6).

[67] Huang M, Pellissier J, Liao J. A Trial-Based Euroqol Eq-5d Health Utility Analysis in Patients with Previously Treated Advanced Nsclc. *Value Health* 2016;19(7).

[68] National Institute for Health and Care Excellence. *Pembrolizumab for untreated PD-L1-positive metastatic non-small-cell lung cancer: NICE technology appraisal guidance 447 [Internet]*. London: NICE, 2017 [accessed 20.3.18] Available from: https://www.nice.org.uk/guidance/TA447

[69] Taylor-Stokes G, Wood R, Lees M, Chirita O. Impact of brain metastases on the humanistic burden incurred by patients with advanced non-small cell lung cancer. *J Thorac Oncol* 2017;12(11 Suppl 2):S2073.

[70] Taylor-Stokes G, Wood R, Lees M, et al. The humanistic burden of advanced non-small cell lung cancer on patients and their caregivers. *J Thorac Oncol* 2017;35(5 (Supplement 1)):A220.

[71] Wood R, Taylor-Stokes G, Lees M, Chirita O. Impact of tobacco smoking on the humanistic and financial burden of advanced non-small cell lung cancer (A-NSCLC). *J Thorac Oncol* 2017;12(11 (Supplement 2)):S1773.

[72] Chouaid C, Mitchell PLR, Agulnik J, et al. Health-related quality of life in advanced non-small cell lung cancer (NSCLC) patients. *Value Health* 2012;15(4):A227.

[73] Lewis G, Peake M, Aultman R, Gyldmark M, Morlotti L, Creeden J, et al. Cost-effectiveness of erlotinib versus docetaxel for second-line treatment of advanced non-small-cell lung cancer in the United Kingdom. *J Int Med Res* 2010;38(1):9-21.

[74] Trippoli S, Vaiani M, Lucioni C, Messori A. Quality of life and utility in patients with non-small cell lung cancer. Quality-of-life Study Group of the Master 2 Project in Pharmacoeconomics. *PharmacoEcon* 2001;19(8):855-63.

[75] Verduyn SC, Biesma B, Schramel FM, et al. Estimating quality adjusted progression free survival of first-line treatments for EGFR mutation positive non small cell lung cancer patients in The Netherlands. *Health Qual Life Outcomes* 2012;10:108.

[76] Huang M, Pelliser J, Kong F. A trial-based Euroqol EQ-5D health utility analysis in patients with previously untreated metastatic NSCLC. *Value Health* 2017;20(5):A115.

[77] Billingham L, Gaunt P, et al. Quality of life in advanced non-small cell lung cancer, effects of cisplatin dose and carboplatin in combination with gemcitabine: results from BTOG2, a British Thoracic Oncology Group phase III trial in 1363 patients. *J Thorac Oncol* 2011;6(6):S319-S20.

[78] Felip E, Blackhall F, Mok T, et al. Impact of crizotinib on patient-reported general health status compared with chemotherapy in patients with no prior systemic treatment for advanced non-squamous ALK-positive non-small cell lung cancer (NSCLC). *J Clin Oncol* 2015;33(15):A8101.

[79] Hirsh V, Yang JC, Tan E-H, et al. First-line afatinib (A) vs gefitinib (G) for patients (pts) with EGFR mutation positive (EGFRm+) NSCLC (LUX-Lung 7): patient-reported outcomes (PROs) and impact of dose modifications on efficacy and adverse events (AEs). *J Clin Oncol* 2016;34(15_suppl):9046.

[80] Tongpak P, Thongprasert S, Permsuwan U. Utility of advanced non-small cell lung cancer patients in Thailand: preliminary study. *Value Health* 2012;15(7):A657.

[81] Khan I, Morris S, Pashayan N, Matata B, Bashir Z, Maguirre J. Comparing the mapping between EQ-5D-5L, EQ-5D-3L and the EORTC-QLQ-C30 in non-small cell lung cancer patients. *Health Qual Life Outcomes* 2016;14:60.

[82] Meregaglia M, Borsoi L, Cairns J, Tarricone R. Mapping Quality of Life Scores from FACT-G, FAACT And FACIT-F Onto Preference-Based Utilities Using the 5-Level Version of EQ-5D Questionnaire. *Value Health* 2016;19(7).

[83] Meregaglia M, Borsoi L, Cairns J, Tarricone R. Mapping health-related quality of life scores from FACT-G, FAACT, and FACIT-F onto preference-based EQ-5D-5L utilities in non-small cell lung cancer cachexia. *Eur J Health Econ* 2017.

[84] Khan I, Morris S. A non-linear beta-binomial regression model for mapping EORTC QLQ- C30 to the EQ-5D-3L in lung cancer patients: a comparison with existing approaches. *Health Qual Life Outcomes* 2014;12:163.

[85] Solem CT, Penrod JR, Lees M, Macahilig C, Luo L, Verleger K, et al. Real-World Health Care Resource Utilization and Related Costs Among Patients Who Received At Least Two Lines of Treatment for Advanced Nsclc in England, The Netherlands, and Sweden. *Value Health* 2017;20(9).

[86] Pulfer A, Hodgson M, Patel A, Purchase J, Summerhayes M, Newsom-Davis T. Resource Use and Associated Costs of Managing Docetaxel-Related Neutropenic Sepsis in Patients with Advanced Non-Small Cell Lung Cancer in The UK Nhs: Reveal-Ns. *Value Health* 2017;20(9):A446-A447.

[87] Lorenzo M, Brown J, Kurosky S, Ajmera MR, Kaye JA. Resource use and Associated Costs for Metastatic Squamous Non-Small Cell Lung Cancer (SQ NSCLC) Patients: A Retrospective Chart Review in Select European Countries. *Value Health* 2016;19(7).

[88] Grossi F, Bennouna J, Havel L, Hochmair M, Almodovar T. Oral vinorelbine plus cisplatin versus pemetrexed plus cisplatin as first-line treatment of advanced non-squamous non-small-cell lung cancer: cost minimization analysis in 12 European countries. *Curr Med Res Opin* 2016;32(9):1577-84.

[89] McGuire A, Martin M, Lenz C, Sollano JA. Treatment cost of non-small cell lung cancer in three European countries: comparisons across France, Germany, and England using administrative databases. *J Med Econ* 2015;18(7):525-32.

[90] Woods B, Sideris E, Palmer S, Latimer N, Soares M. *NICE DSU Technical Support Document 19: Partitioned survival analysis for decision modelling in health care: a critical review*. Sheffield: Decision Support Unit, ScHARR, 2017. 72p. Available from: http://scharr.dept.shef.ac.uk/nicedsu/wp-content/uploads/sites/7/2017/06/Partitioned-Survival-Analysis-final-report.pdf

[91] Antonia SJ, Villegas A, Daniel D, Vicente D, Murakami S, Hui R, et al. Durvalumab after chemoradiotherapy in stage III non-small-cell lung cancer. *N Engl J Med* 2017;377(20):1919-1929.

[92] Reck M, Rodriguez-Abreu D, Robinson AG, Hui R, Csoszi T, Fulop A, et al. Pembrolizumab versus chemotherapy for PD-L1-positive non-small-cell lung cancer. *N Engl J Med* 2016;375(19):1823-1833.

[93] Butts C, Socinski MA, Mitchell PL, Thatcher N, Havel L, Krzakowski M, et al. Tecemotide (L-BLP25) versus placebo after chemoradiotherapy for stage III non-small-cell lung cancer (START): a randomised, double-blind, phase 3 trial. *Lancet Oncol* 2014;15(1):59-68.

[94] Mitchell P, Thatcher N, Socinski MA, Wasilewska-Tesluk E, Horwood K, Szczesna A, et al. Tecemotide in unresectable stage III non-small-cell lung cancer in the phase III START study: updated overall survival and biomarker analyses. *Ann Oncol* 2015;26(6):1134-42.

[95] Latimer N. *NICE DSU Technical Support Document 14: Undertaking survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data*. Sheffield: Decision Support Unit, ScHARR, 2011 (Updated: March 2013). 52p. Available from: http://scharr.dept.shef.ac.uk/nicedsu/wp-content/uploads/sites/7/2016/03/NICE-DSU-TSD-Survival-analysis.updated-March-2013.v2.pdf

[96] Tremblay G, Haines P, Briggs A. A criterion-based approach for the systematic and transparent extrapolation of clinical trial survival data. *JHEOR* 2014;2(2):147-60.

[97] National Institute for Health and Care Excellence (NICE). *Position statement on use of the EQ-5D-5L valuation set*. London: NICE, 2017. 3p. Available from: https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisal-guidance/eq5d5l_nice_position_statement.pdf

[98] van Hout B, Janssen MF, Feng YS, Kohlmann T, Busschbach J, Golicki D, et al. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. *Value Health* 2012;15(5):708-15.

[99] Kind P, Hardman G, Macran S. *UK population norms for EQ-5D: Discussion paper 172*. York: Centre for Health Economics, 1999 Available from: https://www.york.ac.uk/media/che/documents/papers/discussionpapers/CHE%20Discussion%20Paper%20172.pdf

[100] AstraZeneca. *A Phase III, randomized, double-blind, placebo-controlled, multi-center, international study of MEDI4736 as sequential therapy in patients with locally advanced, unresectable non-small cell lung cancer (Stage III) who have not progressed following definitive, platinum-based, concurrent chemoradiation therapy (PACIFIC). (Clinical Study Report: durvalumab) 21 July 2017. Data on File. [PDF provided by the Company]*, 2017

[101] Chouaid C, Agulnik J, Goker E, Herder GJ, Lester JF, Vansteenkiste J, et al. Health-related quality of life and utility in patients with advanced non-small-cell lung cancer: a prospective cross-sectional patient survey in a real-world setting. *J Thorac Oncol* 2013;8(8):997-1003.

[102] Ara R, Brazier JE. Using health state utility values from the general population to approximate baselines in decision analytic models when condition-specific data are not available. *Value Health* 2011;14(4):539-45.

[103] Mulhern B, Feng Y, Shah K, Janssen MF, Herdman M, van Hout B, et al. Comparing the UK EQ-5D-3L and English EQ-5D-5L value sets. *PharmacoEcon* 2018;36(6):699-713.

[104] National Health Service (NHS). *Reference costs 2016/17: National schedule of reference costs 1 [Internet]*, 2017 [accessed 6.7.18] Available from: https://improvement.nhs.uk/resources/reference-costs/

[105] Curtis L, Burns A. *Unit costs of health and social care 2017 [Internet]*: PSSRU, 2017 [accessed 10.1.18] Available from: https://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2017/

[106] Haymarket Media Group Ltd. *MIMS UK*, 2018 [accessed 3.1.18] Available from: https://www.mims.co.uk/

[107] Department of Health and Social Care. *eMIT National Database 2017*, 2018 (Updated: 5 January 2018) [accessed 5.1.18] Available from: https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit

[108] AstraZeneca. Disease progression after CRT in unresectable stage III NSCLC patients. (DOF-IMF-002-AUG18) 25 August 2018. Data on File. [PDF provided by the Company]. 2018.

[109] European Medicines Agency. *Summary of product characteristics (SPC): Imfinzi 50 mg/mL concentrate for solution for infusion [Pending] [Internet]*. London: EMA, 2018 [accessed 22.10.18] Available from: https://www.ema.europa.eu/medicines/human/EPAR/imfinzi

[110] National Institute for Health and Care Excellence. *Pembrolizumab for treating locally advanced or metastatic urothelial carcinoma after platinum-containing chemotherapy: NICE technology appraisal guidance 519 [Internet]*. London: NICE, 2018 Available from: https://www.nice.org.uk/guidance/TA519

[111] Postmus PE, Kerr KM, Oudkerk M, Senan S, Waller DA, Vansteenkiste J, et al. Early and locally advanced non-small-cell lung cancer (NSCLC): ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 2017;28(Suppl 4):iv1-iv21.

[112] AstraZeneca. Resource use associated with monitoring of patients. (DOF-IMF-006-AUG18) 28 August 2018. Data on File. [PDF provided by the Company]. 2018.

[113] Maslove L, Gower N, Spiro S, Rudd R, Stephens R, West P. Estimation of the additional costs of chemotherapy for patients with advanced non-small cell lung cancer. *Thorax* 2005;60(7):564-9.

[114] National Institute for Health and Care Excellence. *Erlotinib and gefitinib for treating non-small-cell lung cancer that has progressed after prior chemotherapy: NICE technology appraisal guidance 374 [Internet]*. London: NICE, 2015 [accessed 20.3.18] Available from: https://www.nice.org.uk/guidance/TA374

[115] Taylor DG, Carter S. *Valuing choice - dying at home: a case for the more equitable provision of high quality support for people who wish to die at home. An economic and social policy opinion commissioned by Marie Curie Cancer Care*: School of Pharmacy, University of London, 2004

[116] National Institute for Health and Care Excellence. *Advanced breast cancer: diagnosis and treatment: NICE clinical guidance 81 [Internet]*. London: NICE, 2009 Available from: https://www.nice.org.uk/guidance/cg81

[117] National Health Service (NHS) England. *Dose banded chemotherapy standardised products*, 2018 Available from: https://www.england.nhs.uk/commissioning/spec-services/npc-crg/group-b/b02/

[118] AstraZeneca. Clinical expert opinion on long-term survival outcomes with SoC. (DOF-IMF-004-AUG18) 25 August 2018 2018. Data on File. [PDF provided by the Company]. 2018.

[119] Kaltenthaler E, Carroll C, Hill-McManus D, Scope A, Holmes M, Rice S, et al. The use of exploratory analyses within the National Institute for Health and Care Excellence single technology appraisal process: an evaluation and qualitative analysis. *Health Technol Assess* 2016;20(26):1-48.

Appendix 1: ERG search strategies

Additional limitations of the CS searches not covered in the main body of the report:

Clinical effectiveness

- Greater use could have been made of indexing terms and free-text terms in the population facet. Recall could have been increased by truncating 'cancer*', including terms such as 'pulmonary' and 'bronchial' to the lung cancer facet, searching for specific CRT interventions and broader CRT terms, and including 'stage three' as a disease stage term.

Cost effectiveness

- Bronchial has been misspelled as "brochial" in the 2016 MEDLINE In Process and Econlit searches and the 2018 MEDLINE, MEDLINE In Process, HTA and NHSEED searches.
- There is an error in the cost effectiveness studies 2016 MEDLINE In Process strategy (line #72 should read '#9 AND #35 AND #71' not '#9 AND #35 AND #70'). This is likely to have affected the search results.
- Some of the 2016 cost effectiveness strategies (HTA, NHS EED, MEDLINE In Process, EconLit) do not include search terms for chemoradiotherapy. These terms are included in the 2016 MEDLINE and Embase database searches and all 2018 update searches

Search line numbers were omitted in some 2016 strategies. This did not affect the search results, but made it difficult to check the strategies.