

University of Groningen

## Durvalumab for the Treatment of Locally Advanced, Unresectable, Stage III Non-Small Cell Lung Cancer

Witlox, Willem J. A.; van Asselt, Antoinette D. I.; Wolff, Robert; Armstrong, Nigel; Worthy, Gill; Chalker, Annette; Buksnys, Titas; Stirk, Lisa; Kleijnen, Jos; Joore, Manuela A.

*Published in:*  
Pharmacoeconomics

*DOI:*  
[10.1007/s40273-019-00870-w](https://doi.org/10.1007/s40273-019-00870-w)

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*  
Publisher's PDF, also known as Version of record

*Publication date:*  
2019

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Witlox, W. J. A., van Asselt, A. D. I., Wolff, R., Armstrong, N., Worthy, G., Chalker, A., ... Grimm, S. E. (2019). Durvalumab for the Treatment of Locally Advanced, Unresectable, Stage III Non-Small Cell Lung Cancer: An Evidence Review Group Perspective of a NICE Single Technology Appraisal. *Pharmacoeconomics*. <https://doi.org/10.1007/s40273-019-00870-w>

### Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

### Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*



# Durvalumab for the Treatment of Locally Advanced, Unresectable, Stage III Non-Small Cell Lung Cancer: An Evidence Review Group Perspective of a NICE Single Technology Appraisal

Willem J. A. Witlox<sup>1</sup> · Antoinette D. I. van Asselt<sup>1,2</sup> · Robert Wolff<sup>3</sup> · Nigel Armstrong<sup>3</sup> · Gill Worthy<sup>3</sup> · Annette Chalker<sup>3</sup> · Titas Buksnys<sup>3</sup> · Lisa Stirk<sup>3</sup> · Jos Kleijnen<sup>3,4</sup> · Manuela A. Joore<sup>1,4</sup> · Sabine E. Grimm<sup>1</sup>

© The Author(s) 2019

## Abstract

As part of the Single Technology Appraisal (STA) process, the National Institute for Health and Care Excellence (NICE) invited the manufacturer (AstraZeneca) of durvalumab (IMFINZI™) to submit evidence for the clinical and cost effectiveness of durvalumab for the treatment of patients with locally advanced, unresectable, stage III non-small cell lung cancer whose tumours express programmed death-ligand 1 (PD-L1) on  $\geq 1\%$  of tumour cells and whose disease has not progressed after platinum-based chemoradiation therapy. Kleijnen Systematic Reviews Ltd, in collaboration with Maastricht University Medical Centre, was commissioned to act as the independent Evidence Review Group (ERG). This paper summarises the company submission (CS), presents the ERG's critical review on the clinical- and cost-effectiveness evidence in the CS, highlights the key methodological considerations, and describes the development of the NICE guidance by the Appraisal Committee. The CS included a systematic review that identified one randomised controlled trial, comparing durvalumab with SoC. Participants with tumours expressing PD-L1 on  $\geq 1\%$  of tumour cells accounted for approximately 40% of the total participants. In this subgroup, a benefit in progression-free survival (PFS) [hazard ratio (HR) 0.44, 95% confidence interval (CI) 0.31–0.63] and overall survival (HR 0.54, 95% CI 0.35–0.81) was reported. Adverse events were comparable between both treatments, but more serious adverse events were reported for durvalumab (64/213 [30%] vs. 18/90 [20%]). The ERG's concerns regarding the economic analysis included a likely overestimation of PFS for the durvalumab arm, the choice of timepoint for treatment waning, as well as the way treatment waning was incorporated in the model, and potential overestimation of utility values without applying an age- or treatment-related decrement. The revised ERG base-case resulted in a deterministic incremental cost-effectiveness ratio of £50,238 per quality-adjusted life-year gained, with substantial remaining uncertainty. NICE recommended durvalumab as an option for use within the Cancer Drugs Fund only in a subpopulation (concurrent platinum-based chemoradiation therapy) with a commercially managed access agreement in place.

## 1 Introduction

Durvalumab (trade name IMFINZI™) was appraised within the National Institute for Health and Care Excellence (NICE) single technology appraisal (STA) process. Health technologies must be shown to be clinically effective and to represent a cost-effective use of National Health Service (NHS) resources in order to be recommended by NICE. Within the STA process, the company (AstraZeneca) provided NICE with a written submission and a mathematical health economic model, summarising the company's estimates of the clinical- and cost-effectiveness of durvalumab for the treatment of locally advanced, unresectable, stage III non-small cell lung cancer (NSCLC) in adults whose tumours express programmed death-ligand 1 (PD-L1) on  $\geq 1\%$  of tumour cells and whose disease has not progressed after platinum-based

✉ Willem J. A. Witlox  
willem.witlox@mumc.nl

<sup>1</sup> Department of Clinical Epidemiology and Medical Technology Assessment, Maastricht University Medical Centre+ (MUMC+), PO Box 5800, 6202 AZ Maastricht, The Netherlands

<sup>2</sup> Department of Epidemiology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

<sup>3</sup> Kleijnen Systematic Reviews Ltd, York, UK

<sup>4</sup> Care and Public Health Research Institute (CAPHRI), Maastricht University, Maastricht, The Netherlands

### Key Points for Decision Makers

There is uncertainty regarding the generalisability of the trial data to the UK setting, and generalisability issues should be a key consideration in National Institute for Health and Care Excellence decision making.

The use of interim analyses may result in assumptions regarding longer-term effectiveness and treatment waning, for which there is currently a lack of guidance, thereby causing increased uncertainty and the risk of biased estimates.

When a trial that can provide answers to the most pressing questions is still ongoing and there is doubt over the cost effectiveness of a new technology, reimbursement within the Cancer Drugs Fund should be carefully considered and the value of further data collection should be formally assessed.

chemoradiation therapy. This company submission (CS) was reviewed by an Evidence Review Group (ERG) independent of NICE [1]. The ERG, Kleijnen Systematic Reviews in collaboration with Maastricht University Medical Centre, produced an ERG report [1]. After consideration of the evidence submitted by the company, and the ERG report, the NICE Appraisal Committee (AC) issued guidance on whether or not to recommend the technology by means of the Final Appraisal Determination (FAD), which is open for appeal. This paper presents a summary of the ERG report and the development of the NICE guidance. Furthermore, it highlights important methodological issues that were identified which may help in future decision making.

Full details of all relevant appraisal documents (including the appraisal scope, CS, ERG report, consultee submissions, appraisal consultation document [ACD], FAD, and comments from consultees) can be found on the NICE website [1].

## 2 The Decision Problem

The CS defined the population as “adults with locally advanced, unresectable, stage III NSCLC whose tumours express programmed death-ligand 1 (PD-L1) on  $\geq 1\%$  of tumour cells (TCs) and whose disease has not progressed following platinum-based chemoradiation therapy (CRT)” [1]. In the CS, the company emphasises the prevalence of lung cancer as being the third most common cancer in the UK. Lung cancer was identified as being the main cause of cancer-related death [1]. Compared with the NICE scope,

the population was narrower, i.e. only included patients in the relevant population whose tumours expressed PD-L1 [1]. The intervention (durvalumab 10 mg/kg every 2 weeks intravenously), comparator (standard of care [SoC]) and outcomes are defined in line with the NICE scope [1].

## 3 Independent Evidence Review Group (ERG) Review

The ERG reviewed the clinical- and cost-effectiveness evidence of durvalumab for this indication. As part of the STA process, the ERG and NICE had the opportunity to ask for clarification on specific issues in the CS, in response to which the company provided additional information [1]. The ERG also produced an ERG base-case to assess the impact of alternative assumptions and parameter values on the model results, by modifying the health economic model submitted by the company. Sections 3.1–3.6 summarise the evidence presented in the CS, as well as the review of the ERG.

### 3.1 Clinical Effectiveness Evidence Submitted by the Company

The CS reported a systematic review (SR) of the evidence for durvalumab for the treatment of locally advanced, unresectable, stage III NSCLC in adults whose tumours express PD-L1 on  $\geq 1\%$  of tumour cells and whose disease has not progressed after platinum-based CRT. The SR identified one randomised controlled trial (RCT), PACIFIC, which compared durvalumab with SoC in adults in the population defined in the decision problem [2]. The population of participants receiving durvalumab represented approximately 67% of the overall population included in PACIFIC.

The CS reported a progression-free survival (PFS) benefit with durvalumab (median 23.9 months) when compared with placebo (median 5.6 months) in the PD-L1  $\geq 1\%$  expression group (hazard ratio [HR] 0.44, 95% confidence interval [CI] 0.31–0.63). Patients in the PD-L1  $\geq 1\%$  expression groups receiving durvalumab also observed an overall survival (OS) benefit (HR 0.54, 95% CI 0.35–0.81).

Common adverse events (AEs) 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. While any AEs are comparable between durvalumab (205/213, 96.2%) and placebo (83/90, 92.2%), more serious AEs were reported for durvalumab (64/213 [30%] vs. 18/90 [20%]).

### 3.2 Critique of Clinical Effectiveness Evidence and Interpretation

As previously highlighted, the decision problem used in the CS was narrower than the NICE scope. Participants in the relevant subgroup (tumour expressing PD-L1 on  $\geq 1\%$  of tumour cells) accounted for approximately 40% of the total participants in the PACIFIC trial; however, the 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 were potential problems linked to overinterpretation of subgroup analyses that might impact on the findings.

The ERG was concerned about 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 a request for clarification suggested the cohort in the PACIFIC trial is generalisable to UK patients with locally advanced, unresectable, stage III NSCLC, and that survival rates might be lower among patients treated with sequential CRT approaches than overlapping approaches. However, as stated in the response, 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 impacted on the certainty regarding the findings and might limit the applicability to UK clinical practice. Furthermore, the PACIFIC trial included only eight patients from the UK.

Efficacy results came from an interim cut-off in March 2018, i.e. not from the final analysis, which was planned at a ‘maturity’ of 70%, i.e. when 70% of patients had died. At the cut-off, the percentage of patients who had died was reported to be 33% in the durvalumab group and 49.5% in the placebo group.

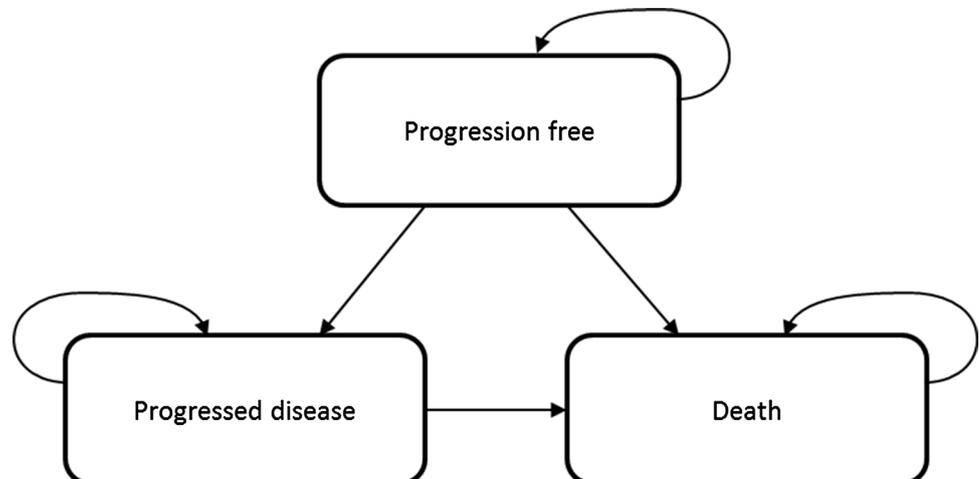
### 3.3 Cost-Effectiveness Evidence Submitted by the Company

The company conducted systematic literature reviews to identify relevant economic, cost and resource use, and health-related quality of life (HRQoL) evidence. None of the included cost-effectiveness studies were conducted from a UK perspective, and the company therefore built a de novo semi-Markov cohort state transition model. The model comprised three health states, i.e. progression-free (PF), progressed disease (PD) and death (Fig. 1). The company employed a semi-Markov approach, rather than a partitioned survival analysis approach, estimating PFS, time to progression (TTP) and post-progression survival (PPS), to inform transitions between health states [1, 3]. This avoided logical inconsistencies caused by the OS curve falling below the PFS curve.

The model adopted the perspective of the NHS and Personal Social Services. The model time horizon was 40 years, with a cycle length of 2 weeks in the first year and 4 weeks thereafter. A half-cycle correction was applied, but not to treatment and treatment administration costs. All costs and quality-adjusted life-years (QALYs) were discounted at a rate of 3.5% per year.

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  $\geq 1\%$  of TCs and whose disease has not progressed after two or more cycles of platinum-based CRT. This was a subgroup

**Fig. 1** Company’s model structure for locally advanced, unresectable, stage III non-small cell lung cancer in adults whose tumours express programmed death-ligand 1 on  $\geq 1\%$  of tumour cells and whose disease has not progressed after platinum-based chemoradiation therapy



from the final scope issued by NICE, which considered the same population regardless of their PD-L1 status. In line with the dosage used in the PACIFIC trial, durvalumab was modelled with a posology of 10 mg/kg administered as an intravenous infusion over 60 min every 2 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 SoC, which applies up to disease progression. The company provided a more comprehensive definition of SoC in its response to the clarification letter, i.e. “*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*” [1].

The main source of evidence on treatment effectiveness used for intervention and comparators was the PACIFIC study [4]. Only data from the subgroup of PD-L1  $\geq 1\%$  patients (according to the anticipated marketing authorisation) and from the March 2018 interim analysis data cut-off were used in the model. Parametric survival curves were fitted to patient-level data from PACIFIC on PFS, but instead of using the OS data from PACIFIC, the company performed survival analyses on TTP and PPS. The probability of remaining in the 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 PFS curve for durvalumab was altered 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, an HR of 1 was applied to the placebo curve to model durvalumab PFS.

Survival analysis of PFS data was used to determine the proportion of patients remaining in the progression-free health state. Instead of calculating the proportion of progressed patients based on OS and PFS, the probability of patients moving from PF to the PD health state was determined by survival analysis of TTP data (PFS data with deaths treated as censored) from PACIFIC. 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 PACIFIC received subsequent treatments. An alternative method for extrapolating PPS was used in a scenario analysis, where PPS was based on published data from the KEYNOTE-024 study, where data from the pembrolizumab arm was used for those patients in PACIFIC who received immuno-oncological (IO) treatment, and data from

the chemotherapy arm was used for those not receiving IO treatment [5].

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 [6]. A mixed-effects model with a covariate on progression status was used to estimate utility values for PF (0.810) and PD (0.776). No AE-related disutilities were taken into account.

Costs in the model included costs for PD-L1 testing, costs associated with treatment, disease management and patient observation, and costs associated with end-of-life care. Unit costs were based on the NHS reference costs [7], Personal Social Services Research Unit (PSSRU) [8], Monthly Index of Medical Specialities (MIMS) [9], and the electronic Market Information Tool (eMIT) [10]. Treatment costs per durvalumab infusion were calculated based on the average body weight for patients in PACIFIC, with treatment duration taken from PACIFIC Kaplan–Meier (KM) data. The model assumed no drug wastage (i.e. perfect vial sharing) and 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 health state was modelled in accordance with European Society for Medical Oncology (ESMO) guidelines [11], and resource use for PD was derived from NICE Technology Appraisal 531 in the metastatic setting [12]. 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. All costs were inflated to the 2017 price level, if necessary.

In the company’s deterministic base-case analysis, total life-years (LYs) and QALYs gained, as well as total costs, were higher in the durvalumab arm compared with the SoC arm. Incremental QALYs were mainly driven by QALY gains in the PF health state, whereas incremental costs mainly resulted from higher treatment costs. After the company corrected a few errors and violations identified by the ERG, the deterministic (probabilistic, based on 1000 iterations) incremental cost-effectiveness ratio (ICER) amounted to £19,366 (£21,601) per QALY gained. Some deterministic sensitivity analyses and scenario analyses significantly affected the ICER, of which applying a 10-year time horizon and using a 3-year treatment waning cut-off were the most notable.

### 3.4 Critique of Cost-Effectiveness Evidence and Interpretation

The CS and response to clarification provided sufficient detail for the ERG to appraise the literature searches and the 2018 update searches, which were deemed appropriate.

The ERG had concerns about the model being based on post-hoc analyses: first, treatment effectiveness estimates were derived from a post-hoc subgroup analysis and second, post-hoc TTP and PPS analyses were performed. Furthermore, analyses were based on a population with mostly prior overlapping CRT instead of sequential CRT. Any bias introduced by these analyses remained unclear.

The ERG had concerns about the appropriateness of the semi-Markov approach and questioned its superiority over a partitioned survival model approach, given the immaturity of the survival data in the PACIFIC subpopulation. PPS data were naturally available for fewer patients than OS data, which meant that extrapolation from PPS drew on small sample sizes. Additional bias was potentially introduced because data maturity had not been reached and the PPS analysis therefore relied on early progressors, with more progressions having occurred in the placebo arm than in the durvalumab arm (see also NICE Decision Support Unit [DSU], Technical Support Document [TSD] 19) [13]. Unfortunately, the company did not provide the partitioned survival analysis approach.

The main concern of the ERG in relation to the company's choice of PFS modelling (generalised gamma) was that it likely overestimated PFS, compared with PFS estimates observed in PACIFIC at 3 years. This model choice caused ICERs to be lower than with other model choices. It is noteworthy that due to the PACIFIC estimate being based on small numbers of patients at risk (many censoring events), any choice for modelling PFS was associated with high levels of uncertainty, and that different PFS model choices had a large impact on the ICERs. Alternative modelling methods, such as spline-based models, remained unexplored in the CS. The Gompertz distribution was used in a scenario analysis and the company did not explore the log-normal distribution, despite it making a better fit than the Gompertz distribution in both arms.

The ERG considered the choice of the 10-year time point for modelling treatment waning as highly uncertain, not appropriately validated, and potentially late, further adding to the likely overestimation of durvalumab PFS. Furthermore, setting the hazards for durvalumab and SoC to be equal after the chosen treatment duration cut-off in the company's model could cause counter-intuitive results in some instances (with the ICER decreasing when the treatment effect duration decreased). This was because the hazard of an event in the SoC arm could be lower than the hazard in the durvalumab arm when differential PFS distributions were chosen (so the earlier treatment effect cut-off lowered the hazard of progression in the durvalumab arm). The ERG considered that this method of modelling treatment waning was therefore inappropriate. However, alternative assumptions were also clinically implausible (e.g. a sudden drop in the number of patients not progressed or died). The ERG

concluded that modelling treatment waning may introduce bias and that further research is necessary to establish good practices for modelling treatment waning, particularly when the intervention and comparator curves are modelled separately. Ideally, longer-term data would be available to inform treatment effectiveness.

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 numbers of patients and immature data, compared with 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.

AEs were modelled as a per-cycle occurrence while patients were receiving treatment. While AEs causally related to treatment were mostly higher for the durvalumab arm than in the placebo arm in PACIFIC, the incidence of AEs in the model between treatments was comparable. It was unclear how this discrepancy occurred, but it likely lowered the ICERs of durvalumab versus SoC; however, exploratory analyses showed that any bias caused by this was likely small.

The utility values for PF and PD derived from PACIFIC data were higher than in studies identified from the SR, which could potentially partly be explained by fewer patients in PACIFIC having metastatic disease than in the comparator studies. However, utility values for both the PF and PD health states were comparable with those in the general population, which the ERG considered to be potentially an overestimate. Utility values were potentially overestimated and not adjusted by general population utility estimates. The high PF utility value produced lower ICERs, while the high PD utility value produced higher ICERs for durvalumab versus SoC. The ERG was also 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, as well as adverse events, was not appropriately captured in the model.

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 who would be eligible for treatment with durvalumab. This assumption caused the ICER of durvalumab against SoC to be lower compared with the ICER when the perfect vial sharing assumption was relaxed.

Compared with the company's deterministic base-case results, probabilistic incremental QALYs were lower. Furthermore, the company included patient characteristics in their PSA, and the 1000 iterations used in the PSA did not achieve stable results. Finally, the ERG was concerned about the external validation exercised by the company because

no firm conclusions could be drawn. The company stated that modelled OS broadly matched survival of the available sources, but this was a subjective observation. In addition, it was unclear whether the survival differences between the model and previous technology appraisals was caused by differences in population or by poor external validity of the current model.

### 3.5 Additional Work Undertaken by the ERG

Based on all considerations highlighted in the ERG critique, the ERG defined a new base-case in which various adjustments were made to the company's base-case. This included the correction of errors in the calculations of age and vial wastage of nivolumab and pembrolizumab as subsequent treatments and an error in the implementation of utility decrements in the PSA, the application of a half-cycle correction to treatment and administration costs, the assumption of no vial sharing and the exclusion of patient characteristics from the PSA. As matters of judgement, the ERG used a different curve to extrapolate durvalumab PFS to avoid overestimation, selected a 5-year treatment waning cut-off, applied an age-related utility decrement, and included treatment as a covariate in the utility mixed-effects model. The ERG base-case analysis resulted in larger deterministic (£50,238 per QALY gained) and probabilistic (£52,353 per QALY gained) ICERs than those in the company's base-case.

Furthermore, the ERG explored alternative PFS distributions and treatment waning effects, the use of PPS based on PACIFIC and KEYNOTE-024, AEs incorporated with amended incidence and including the impact on HRQoL, alternative utility scores for the PF and PD states and the possibility of 30% vial sharing.

### 3.6 Conclusions of the ERG Report

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.

The ERG base-case ICER was estimated to be approximately £50,000 per QALY gained, but was subject to substantial uncertainty. The modelled long-term durvalumab PFS was highly uncertain given the model choice (generalised gamma) that resulted in higher PFS estimates than observed in PACIFIC. There were concerns about the appropriateness of the semi-Markov approach given the immaturity of the survival data in PACIFIC, and the modelled population was narrower than that in the scope. Further uncertainties were related to the treatment effectiveness, potentially overestimated utility values, the assumption of perfect vial sharing, and discrepancies between AEs reported

in PACIFIC and the AE incidence in the model. Some of these uncertainties were explored in scenario analyses and mostly further increased the ICER.

In conclusion, given that the ERG base-case ICER was estimated to be approximately £50,000 per QALY gained, with none of the scenarios resulting in ICERs below £30,000 per QALY gained, and the large uncertainty induced by mainly the immature survival data, uncertainty around the cost effectiveness of durvalumab was substantial.

## 4 Key Methodological Issues

The chosen parametric model to estimate PFS (generalised gamma) in this appraisal likely overestimated durvalumab PFS in PACIFIC. NICE DSU TSD 14 recommends the use of the same type of parametric model for individual treatment arms to avoid drastically different shapes of survival curves, unless there is justification for deviating from this guidance: “*using clinical expert judgement, biological plausibility, and robust statistical analysis*” [14]. As detailed above, there were reasons to believe that differential curves to model PFS for durvalumab (log-normal) and SoC (generalised gamma) exhibited superior external validity and plausibility. Given that the number of cancer treatments with potential for cure is expected to increase, the need for guidance and best practices on when to choose differential survival curves per treatment arm may increase in parallel.

The implementation of the treatment waning effect, setting the HR to 1 at the chosen cut-off, can cause counter-intuitive results when the per-period hazard in the comparator arm is below that of the hazard in the intervention arm, which can happen, for example, when differential survival models are chosen. This can occur depending on the chosen cut-off and when rates of progression or dying in the beginning periods are high in the comparator arm, such that the few remaining patients alive and not progressed in later periods then have very low rates of progression or dying. A hazard cap was proposed by the company to solve this issue but this was not based on clinical evidence or validated methods. The ERG does not consider it entirely clinically implausible that, at some point, the proportion of patients progressed would be higher in the durvalumab arm than in the SoC arm. This could, for example, be a plausible scenario where a treatment delays progression but does not avert it more effectively than SoC. A hazard cap should therefore not be employed without being based on clinical evidence. 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 appropriate assumptions. With many appraisals currently considering treatment waning (given the uncertainty about long-term treatment effectiveness), this is a pressing

issue. Modelling treatment waning may introduce bias and further research is necessary to establish good practices for its modelling.

The approach of modelling PPS instead of OS may introduce bias. Survival data in PACIFIC are immature, and, although 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 PPS analysis was potentially biased because groups were no longer balanced. Early progressors may have a different survival time than patients with later progression [13]. In addition, the PPS was estimated using both treatment and comparator groups together, but PPS data included more patients treated with placebo (who progressed earlier), potentially introducing additional bias. Alternative standard OS analysis should be provided in addition to exploring any bias stemming from this choice of analysis and assessing the potential impact.

## 5 National Institute for Health and Care Excellence Guidance

On 14 February 2019, NICE recommended durvalumab monotherapy for use within the Cancer Drugs Fund (CDF) as an option for treating locally advanced unresectable NSCLC in adults whose tumours express PD-L1 on  $\geq 1\%$  of tumour cells and whose disease has not progressed after platinum-based chemoradiation therapy, only if platinum-based chemoradiation was administered concurrently and the conditions in the managed access agreement for durvalumab are followed.

### 5.1 Consideration of Clinical Effectiveness

The NICE technical report considered the difference between types of prior CRT that could affect the generalisability of trial results and clinical practice in the NHS [1]. However, it stated that PACIFIC remains the best available source of evidence for this appraisal and that, based on clinical expert advice, the difference in outcomes between overlapping and sequential CRT is likely to be small.

The report also considered the suggestion of the CS that durvalumab meets the end-of-life criteria specified by NICE [1, 15]. It concluded that the criteria were not met due to the index population not meeting the short life expectancy criterion.

The report also considers the submission being based on a more restricted population than the NICE scope to be “a substantial area of uncertainty in the appraisal”.

### 5.2 Consideration of Cost Effectiveness

The AC considered that assuming a 3- to 5-year treatment effect duration is plausible but concluded that the long-term treatment effect of durvalumab after stopping treatment was highly uncertain when considering the input from clinical experts. The AC agreed that the immature PFS data from PACIFIC, and the small numbers at the end of the KM curve, introduced uncertainty in all extrapolations and therefore concluded that more mature PFS data are needed to inform long-term model predictions. The AC accepted a hazard cap that was proposed by the company to prevent the risk of progression in the durvalumab arm exceeding the risk in the standard care arm when using the log-normal and generalised gamma distributions. The AC concluded that it was acceptable to use the PFS utility value from PACIFIC for the PF health state, and, for consistency in the economic model, the use of the PD utility from PACIFIC for the PD health state was also considered appropriate. The AC also considered that using a treatment-related decrement to model adverse events was appropriate, but this decrement would not apply indefinitely. Finally, the AC concluded that the most plausible ICERs for durvalumab are uncertain and are not clearly within the range considered to be cost effective.

## 6 Conclusions

As detailed previously, there is substantial uncertainty regarding the effectiveness results that come from an interim analysis conducted in a restricted population. Furthermore, the generalisability of the results to UK clinical practice might be limited as only eight patients in the UK were included and there were differences between types of prior CRT.

The Committee did not recommend durvalumab for standard use in the NHS, but did recommend it within the CDF for a subpopulation (concurrent platinum-based chemoradiation therapy) with a commercially managed access agreement in place. This recommendation was made as the evidence from the only available trial was not generalisable to the whole UK population (which mainly receives sequential platinum-based chemoradiation therapy, not studied in the trial), the data from the trial were immature, and the trial was still ongoing. In particular, the committee was interested in further evidence on how long the treatment effect of durvalumab lasts, and how many people taking durvalumab would live without their disease progressing (‘cure’), on both of which further information is being collected in the ongoing trial. The recommendation within the CDF may be perceived as problematic given that the most plausible ICERs were above the range considered to be cost effective by NICE. Furthermore, in order to guarantee the value

of data collection within the CDF, this should be formally assessed. Reimbursement within the CDF should be considered carefully, potentially with the addition of further conditions on data to be collected and/or reimbursed.

**Acknowledgements** This summary of the ERG report was compiled after NICE issued the FAD. All authors have commented on the submitted manuscript and have given their approval for the final version to be published. The views and opinions expressed therein are those of the authors and do not necessarily reflect those of NICE or the Department of Health. Any errors are the responsibility of the authors.

**Author Contributions** All authors have commented on the submitted manuscript and have given their approval for the final version to be published. RW, AC, TB and JK critiqued the clinical effectiveness data reported by the company; LS critiqued the literature searches undertaken by the company; GW critiqued the statistical analyses performed by the company; and WW, TA, NA, TB, MJ and SG critiqued the mathematical model provided and the cost-effectiveness analyses submitted by the company. WW acts as overall guarantor for the manuscript. This summary has not been externally peer reviewed by PharmacoEconomics.

## Compliance with Ethical Standards

**Funding** This project was funded by the National Institute for Health Research (NIHR) Health Technology Assessment Programme. Please visit the HTA programme website for further project information (<https://www.nihr.ac.uk/funding-and-support/funding-for-research-studies/funding-programmes/health-technologyassessment>).

**Conflicts of interest** Willem Witlox, Antoinette van Asselt, Robert Wolff, Nigel Armstrong, Gill Worthy, Annette Chalker, Titas Buksnys, Lisa Stirk, Jos Kleijnen, Manuela Joore and Sabine Grimm have no conflicts of interest to declare.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

1. National Institute for Health and Care Excellence. Durvalumab for treating locally advanced unresectable non-small-cell lung cancer after platinum-based chemoradiation [TA578]. 2019. <https://www.nice.org.uk/guidance/ta578/evidence>. Accessed 22 Aug 2019.
2. 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):1919–29.
3. AstraZeneca. Durvalumab for treatment of locally advanced, unresectable, stage III non-small cell lung cancer in adults whose tumours express PD-L1 on  $\geq 1\%$  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. pp. 356.
4. 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).
5. Reck M, Rodriguez-Abreu D, Robinson AG, Hui R, Czoszi T, Fulop A, et al. Pembrolizumab versus chemotherapy for PD-L1-positive non-small-cell lung cancer. *N Engl J Med*. 2016;375(19):1823–33.
6. National Institute for Health and Care Excellence (NICE). Position statement on use of the EQ-5D-5L valuation set. London: NICE; 2017. [https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisal-guidance/eq5d5\\_L\\_nice\\_position\\_statement.pdf](https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisal-guidance/eq5d5_L_nice_position_statement.pdf). Accessed 25 Oct 2018.
7. National Health Service (NHS). Reference costs 2016/17: National schedule of reference costs 1. 2017. <https://improvement.nhs.uk/resources/reference-costs/>. Accessed 6 Jul 2018.
8. Curtis L, Burns A. Unit costs of health and social care 2017. PSSRU; 2017. <https://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2017/>. Accessed 10 Jan 2018.
9. Haymarket Media Group Ltd. MIMS UK; 2018. Available at: <https://www.mims.co.uk/>. Accessed 3 Jan 2018.
10. Department of Health and Social Care. eMIT National Database, 2017, 2018 (updated: 5 January 2018). <https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-marketing-information-emit>. Accessed 5 Jan 2018.
11. 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):iv1–21.
12. National Institute for Health and Care Excellence. Pembrolizumab for untreated PD-L1-positive metastatic non-small-cell lung cancer: NICE technology appraisal guidance 531. London: NICE; 2018. <https://www.nice.org.uk/guidance/ta531>. Accessed 7 Aug 2018.
13. 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. <http://scharr.dept.shef.ac.uk/nicedsu/wp-content/uploads/sites/7/2017/06/Partitioned-Survival-Analysis-final-report.pdf>. Accessed 20 Sept 2018.
14. 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). <http://scharr.dept.shef.ac.uk/nicedsu/wp-content/uploads/sites/7/2016/03/NICE-DSU-TSD-Survival-analysis.updated-March-2013.v2.pdf>. Accessed 24 Oct 2018.
15. National Institute for Health and Care Excellence. Guide to the processes of technology appraisal. 2018. <https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisals/technology-appraisal-processes-guide-apr-2018.pdf>. Accessed 3 Apr 2018.