STA REPORT

LIVERPOOL REVIEWS AND IMPLEMENTATION GROUP (LRiG)

Osimertinib for locally advanced or metastatic EGFR and T790M mutation-positive non-small cell lung cancer [ID874]

Confidential until published

This report was commissioned by the NIHR HTA Programme as project number 15/121/09

26th April 2016

CONTAINS ACADEMIC IN CONFIDENCE AND COMMERCIAL IN CONFIDENCE DATA

UNIVERSITY OF LIVERPOOL

LIVERPOOL REVIEWS AND IMPLEMENTATION GROUP

A MEMBER OF THE RUSSELL GROUP

Title: Osimertinib for locally advanced or metastatic EGFR and T790M mutation-positive non-small cell lung cancer [ID874]

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

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

James Mahon, Director, Coldingham Analytical Services, Berwickshire

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

Sophie Beale, Research Associate (Decision Analysis), LR*i*G, University of Liverpool

Ashma Krishan, Research Associate (Medical Statistician), LR*i*G, University of Liverpool

Rabeea'h W Aslam, Senior Research Fellow (Clinical Effectiveness), LR*i*G, University of Liverpool

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

Lindsay Banks, Medicines Information Pharmacist, North West Medicines Information Centre, Pharmacy Practice Unit, Liverpool

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

Correspondence to: Janette Greenhalgh, Liverpool Reviews and Implementation Group, University of Liverpool, 2nd Floor, Whelan Building, The Quadrangle, Brownlow Hill, Liverpool L69 3GB

Date completed: 26 April, 2016

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 15/121/09

Declared competing interests of the authors: None

Acknowledgements: The authors would like to thank Dr Nick Brown, Consultant Medical Oncologist, Calderdale and Huddersfield NHS Foundation Trust, who provided feedback on the final version of the report. Dr Brown has received consultancy fees and reimbursement for attending a symposium from Boehringer Ingelheim Ltd and hospitality from Eli Lilly Ltd. We also thank Dr Catrin Tudur Smith, Reader in Medical Statistics, University of Liverpool, for her statistical advice.

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: Greenhalgh J, Mahon J, Boland A, Beale S, Krishan A, Aslam RW, Kotas E, Banks L and Green J. Osimertinib for treating locally advanced or metastatic EGFR and T790M mutation-positive non-small cell lung cancer [ID874]: A Single Technology Appraisal. LR*i*G, University of Liverpool, 2016

Contributions of authors:

Greenhalgh J	Project lead, critique of the clinical evidence, drafted the clinical results section and supervised the final report
Mahon J	Checking and validation of the economic model and critique
Boland A	Critical appraisal of the clinical and economic evidence
Beale S	Critical appraisal of the clinical and economic evidence
Krishan A	Critical appraisal of the statistical evidence
Aslam RW	Project administration, critique of the clinical evidence, drafted the clinical results section
Kotas E	Cross-check of the company submission search strategy
Banks L	Critical appraisal of the company submission
Green J	Clinical advice and critical appraisal of the clinical evidence

All authors read and commented on draft versions of the ERG report.

Table of contents

List of tables

List of figures

LIST OF ABBREVIATIONS

AE	Adverse event
AIC	Akaike Information Criterion
AURA	Clinical programme of trials assessing the clinical effectiveness of osimertinib
BIC	Bayesian Information Criterion
BICR	blinded independent central review
BOR	best overall response
BSA	body surface area
BSC	best supportive care
CI	confidence interval
CR	complete response
CS	company submission
CSR	clinical study report
ctDNA	circulating tumour DNA
DCR	disease control rate
DoR	duration of response
EAMS	Early Access to Medicines Scheme
ECOG	Eastern Cooperative Oncology Group
EGFR (-TKI)	epidermal growth factor receptor (tyrosine kinase inhibitor)
EGFRm+	epidermal growth factor receptor mutation-positive
EMA	European Medicines Agency
EORTC	European Organisation for the Treatment of Cancer
EPAR	European Public Assessment Report
EQ-5D	European Quality of Life - 5 Dimensions Questionnaire
ERG	Evidence Review Group
FACT-L	Functional Assessment of Cancer Therapy – Lung
FAS	full analysis set
GLOBOCAN	Global Burden of Cancer Study
HR	hazard ratio
HRQoL	health related quality of life
ICER	incremental cost effectiveness ratio
IMPRESS	Iressa Mutation-Positive Multicentre Treatment Beyond Progression Study
IPD	individual patient data
K-M	Kaplan-Meier
MHRA	Medicine and Healthcare Regulatory Agency
NSCLC	non-small cell lung cancer
ORR	objective response rate
OS	overall survival
PD	progressed disease
PDC	platinum doublet chemotherapy
PF	progression free
PFS	progression-free survival
PH	proportional hazards
PPS	post-progression survival
PR	partial response
PS	performance score
PSA	probabilistic sensitivity analysis
PSS	Personal and Social Services
PTDS	post-treatment discontinuation survival
QALY	quality adjusted life year
RCT	randomised controlled trial
SAE	serious adverse event
SAP	statistical analysis plan
SD	stable disease
SmPC	summary of product characteristics
T790M	A secondary mutation of the EGFR
TTD	time to treatment discontinuation

1 SUMMARY

The remit of the Evidence Review Group (ERG) is to comment on the clinical and cost effectiveness evidence submitted to the National Institute for Health and Care Excellence (NICE) as part of the single technology appraisal (STA) process. Clinical and economic evidence has been submitted to NICE by AstraZeneca UK Ltd to support the use of osimertinib (Tagrisso®) for locally advanced or metastatic epidermal growth factor (EGFR) T790M mutation-positive non-small cell lung cancer (NSCLC).

Osimertinib is licensed in Europe for the treatment of adults with locally advanced or metastatic EGFR T790M NSCLC. The European Medicines Agency (EMA) granted osimertinib a conditional marketing authorisation on 3rd February 2016. The marketing authorisation is conditional on AstraZeneca providing the final results of an ongoing phase III randomised controlled trial (RCT), AURA3, to the EMA by 30th June 2017.

The clinical evidence presented in the company submission (CS) comes from three main sources: AURAext and AURA2 (both single-arm studies) and the IMPRESS trial (a double-blind placebo-controlled RCT). Data from the two AURA studies provide evidence for the clinical effectiveness of osimertinib. The IMPRESS trial was designed to compare the efficacy of gefitinib+pemetrexed+cisplatin versus placebo+pemetrexed+cisplatin; in this trial, the placebo+pemetrexed+cisplatin combination is labelled as platinum doublet chemotherapy (PDC). The outcomes from a small subgroup of patients (n=max 61) recruited to the control arm of the IMPRESS trial, who were identified retrospectively as having the EGFR T790M mutation, are compared to the outcomes calculated from the pooled AURA dataset.

1.1 Critique of the decision problem in the company submission

Intervention

The intervention specified in the final scope issued by NICE and discussed in the CS is osimertinib. Osimertinib is administered at a dose of 80mg once daily. It is available as 40mg or 80mg film-coated tablets.

Line of treatment is not specified in the conditional EMA licence or in the final scope issued by NICE. The EMA's decision to grant a licence for all treatment lines was based on a biological assumption of effectiveness as there are no data to support the use of osimertinib in treatment-naïve patients. The company expects that, in NHS clinical practice, osimertinib will mainly be used as a second-line treatment after failure of a first-line EGFR tyrosine kinase inhibitor (EGFR-TKI). The company estimates that, if recommended by NICE,

approximately 300 patients per year in England will be eligible for treatment with osimertinib. This estimate includes patients at all lines of treatment.

It is stipulated in the summary of product characteristics (SmPC) for osimertinib that treatment should only be initiated after the patient's EGFR T790M mutation status is positively confirmed using a validated test method. EGFR testing after first-line treatment to establish the presence or absence of the EGFR T790M mutation is feasible in the NHS as the infrastructure is in place. However testing at this point in the treatment pathway is not currently standard practice in the NHS.

Population

The population described in the final scope issued by NICE is people with locally advanced or metastatic EGFR T790M mutation-positive NSCLC. This is the same as the population described in the conditional licence for osimertinib issued by the EMA.

The clinical evidence describing osimertinib submitted by the company is derived from two single-arm studies (AURAext and AURA2). These studies were designed to assess the clinical effectiveness of osimertinib in patients with locally advanced or metastatic EGFR T790M mutation-positive NSCLC who had received treatment with an EGFR-TKI prior to recruitment. Patients in the AURAext and AURA2 studies had received between 1 and 14 prior anti-cancer treatments, including an EGFR-TKI.

Comparators

There are 13 comparators listed in the final scope issued by NICE; these vary by line of treatment. The company presents comprehensive clinical effectiveness data for the unadjusted and adjusted comparison of second or further line treatment with osimertinib versus second-line PDC (specifically, placebo+pemetrexed+cisplatin).

The data used to inform the PDC comparison were obtained from the subgroup of patients (n=61) included in the control arm of the IMPRESS trial whose tumours were identified retrospectively as having the EGFR T790M mutation. Comparisons between outcomes for the selected patients from the IMPRESS trial and patients in the pooled AURA dataset were made using two approaches, a simple unadjusted comparison and an adjusted comparison. The adjusted comparison involved adjustments to control for differences in baseline characteristics between the populations in the two datasets.

The company has assumed that treatment with pemetrexed+cisplatin can be used to represent all PDC treatments, i.e. vinorelbine, gemcitabine, docetaxel or paclitaxel in combination with cisplatin or carboplatin. The ERG is aware that, for the specified

population, pemetrexed+cisplatin is the most commonly used PDC in the NHS. The company also assumes that the efficacy of docetaxel monotherapy in patients untested for EGFR T790M mutations can be used to represent the efficacy data associated with any single-agent chemotherapy for the second-, third-line and further treatment of patients with tumours exhibiting EGFR T790M mutations.

Outcomes

Clinical evidence is presented in the CS for all five outcomes specified in the final scope issued by NICE: overall survival (OS), progression-free survival (PFS), objective response rate (ORR), adverse events (AEs) and health-related quality of life (HRQoL). The currently available OS data from the AURAext and AURA2 studies (pooled) and the subgroup of patients with T790M mutations from the control arm of the IMPRESS trial are very immature (12.7% and approximately 33% respectively).

Other considerations

No subgroups were specified in the final scope issued by NICE in addition to the distinct populations specified in the comparator section. The company has submitted a Patient Access Scheme (PAS) proposal to the Department of Health.

Equality and End of Life considerations

The company has not identified any equality issues. However, the company has presented a case for osimertinib to be assessed against the NICE End of Life criteria.

1.2 Summary of clinical effectiveness evidence submitted by the company

Direct evidence

The company conducted a broad literature search and did not identify any RCTs other than the ongoing phase III AURA3 trial. The results of the AURA3 trial are expected to be available in 2017.

The company has presented results from two single-arm studies, the phase I/II AURAext study and the phase II AURA2 study. The company combines the data from these two studies in a pooled analysis. Results from the pooled AURA dataset (n=411) demonstrate that treatment with osimertinib yields an ORR of 66.1% (95% CI: 61.2 to 70.7), a finding that is consistent across all subgroups tested. Median PFS is 9.7 months (95% CI: 8.3 to not calculable). Results for OS were not available due to the immaturity of the data. The most commonly reported AEs (all grades) were diarrhoea (42.3%) and rashes and acne (41.4%). Grade 3 or 4 AEs included respiratory disorders (13%), infections (6%), investigations

(5.8%) and blood disorders (5%, AURA2). The HRQoL data collected during the AURAext and AURA2 studies suggest that osimertinib has a significant, measurable and relevant impact on patients' HRQoL and symptoms.

Unadjusted and adjusted comparisons evidence

Two comparisons, (unadjusted and adjusted), were carried out to compare the effectiveness of osimertinib with that of PDC.

The unadjusted comparison of osimertinib with PDC yields a statistically significantly higher ORR for osimertinib (66.1% versus 39.3%). The comparison of PFS also demonstrates a statistically significant difference of 4.4 months (9.7 months [95% CI 8.9 to not calculable] versus 5.3 months) in favour of osimertinib. Median OS was not reached for osimertinib and was 15.7 months for PDC.

Results from the adjusted comparison are consistent with those from the unadjusted comparison. The ORR results indicate a statistically significant improvement in favour of osimertinib compared to PDC (64.6% and 34.8% respectively; OR=4.76; 95% CI: 2.21 to 10.26; p<0.001). The disease control rate (DCR) results also indicate a statistically significant improvement in favour of osimertinib compared to PDC (92.1% and 76.1% respectively; OR=4.39; 95% CI: 1.71 to 11.28; p=0.002). The PFS results indicate a statistically significant difference in favour of osimertinib compared to PDC (HR=0.280; 95% CI: 0.185 to 0.422; p<0.0001). Median PFS is 9.7 months for the osimertinib cohort compared to 5.2 months for the PDC cohort. Analysis of OS indicated an overall hazard ratio of 1.022 (95% CI 0.387 to 2.696) for osimertinib versus PDC. The data indicated that treatment with osimertinib is better tolerated than treatment with PDC.

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

The ERG is satisfied with the company's search strategy and stated inclusion and exclusion criteria. The ERG is confident that searching was carried out to an acceptable standard and is not aware of any additional studies that should have been included in the company's systematic review.

1.3.1 Direct evidence

The ERG considers that the AURAext and AURA2 studies were designed and conducted to a good standard. In particular, the use of blinded independent central review (BICR) in the assessment of the radiological results lends robustness to the PFS outcomes. Furthermore, the use of a single treatment arm design means that the OS data from the AURAext study

and AURA2 study cannot be contaminated by treatment crossover. However, data from single-arm studies are difficult to interpret due to the lack of a comparator arm and may be subject to unplanned (and unrecognised) bias and confounding. The interpretation of the results of the AURAext and AURA2 studies is also hampered by the very immature survival data.

The extent to which the submitted evidence reflects outcomes that would be seen in NHS clinical practice is limited by lack of confidence in the magnitude of the outcomes from the AURAext and AURA2 studies. Patients included in these studies are younger and fitter than EGFRm+ patients who would be eligible for treatment with osimertinib in the NHS. Very few EGFRm+ patients in the NHS receive more than one or two treatments after an EGFR-TKI; this is in contrast to patients in the AURAext and AURA2 studies who received up to 14 treatments.

The AURAext study was open to recruitment at two centres in the UK; however, it is not clear how many patients were recruited from the UK. The AURA2 study was not open to recruitment from UK centres.

1.3.2 Unadjusted and adjusted comparisons

The ERG commends the company's efforts to carry out an adjusted comparison. However, the robustness of the outcomes is limited by the small number of patients represented in the PDC cohort. In addition, the company was unable to consider some of the important demographic and disease characteristics (e.g., number of previous EGFR-TKI treatments) required by the matching process.

In particular, the OS results from the adjusted comparison should be interpreted with caution as only 11.5% of the osimertinib data and 29.4% of the PDC data were mature at the time that the analysis was carried out. The ERG and the company agree that the OS data are too immature to allow any meaningful interpretation of results.

1.4 Summary of submitted cost effectiveness evidence

The company developed a de novo cohort-based partitioned survival model in Microsoft Excel to compare the cost effectiveness of osimertinib with PDC in patients with locally advanced or metastatic EGFR T790M mutation-positive NSCLC who have progressed on or after EGFR-TKI therapy (i.e. ≥second-line therapy). The model comprised three health states: progression-free (PF), progressed-disease (PD) and death. All patients entered the model in the PF state. Variants of this model structure have been used in previous NICE STAs. The model time horizon was set to 15 years with a 1-week cycle length. As

recommended by NICE, a discount rate of 3.5% was used for both costs and outcomes; outcomes were measured in quality adjusted life years (QALYs). The model perspective was that of the UK NHS. Survival estimates were based on data collected from the AURAext and AURA2 studies (for osimertinib) and from patients in the control arm of the IMPRESS trial with EGFR T970M positive mutations (for PDC) and other published sources. Utility values were calculated from data collected during the AURA2 study and the IMPRESS trial. Resource use and costs were estimated based on information from the AURAext and AURA2 studies and the IMPRESS trial, published sources and advice from clinical and economic experts. The company also compared osimertinib versus PDC in a second-line population only, versus docetaxel in a second-line population only and versus docetaxel in a ≥third-line population only.

The base case comparison of osimertinib versus PDC resulted in an incremental cost effectiveness ratio (ICER) per QALY gained of [redacted] with osimertinib being more expensive [redacted] and more effective [redacted] life years and [redacted] QALYs). The company carried out a range of deterministic sensitivity analyses. The most influential parameters were utility values, particularly for the PD state, and choice of discount rate. The probabilistic sensitivity analysis (PSA) results showed that the probabilistic ICER of [redacted] per QALY gained had a ≤5% chance of being cost effective at a threshold of £30,000 per QALY gained and a 35% probability of being cost effective at a threshold of £50,000 per QALY gained. The ICER per QALY gained for osimertinib versus PDC in a second-line only population was [redacted]; versus docetaxel in a second-line only population was [redacted] and versus docetaxel in a ≥third-line only population was [redacted]. The company performed scenario analyses using different survival modelling approaches, health state utility values and resource use and costs. Only the choice of survival modelling approach had a significant influence on the size of the ICER per QALY gained; using a Gompertz distribution for PFS and OS yielded an ICER of [redacted] per QALY gained.

1.5 Summary of the ERG's critique of cost effectiveness evidence

The ERG considers that there are several fundamental issues that cast doubt on the cost effectiveness results produced by the company model.

Over 90% of the QALY benefit from osimertinib estimated from the model arises when OS trial data are no longer available. The available OS data for osimertinib and PDC are not statistically significantly different and are very immature, especially so for osimertinib. The only statistically significant evidence on effectiveness incorporated in the model is an improvement in PFS with osimertinib, the extent of which is uncertain due to the single-arm nature of the AURA studies.

The ERG considers that lack of statistical significance in OS between osimertinib and PDC during the period for which data are available means that there is no basis to project differential OS. Even if there were a statistically significant difference between osimertinib and PDC for the period that data were available, as the OS data are so immature, any projection could only be speculative with the degree of uncertainty in the projection being impossible to quantify.

The populations within the AURA studies and the IMPRESS trial appear to be fitter than the EGFRm+ patients who would be expected to be seen in routine NHS practice. This casts doubt on the appropriateness of using the AURA and IMPRESS OS datasets to represent the UK EGFRm+ population even if it was fully mature.

The ERG therefore considers the OS projections employed by the company to be based on opinion rather than to be supported by evidence. To support this view, the ERG cites the wide variation in ICERs that the company shows (CS, p234) could be produced depending on the selection of different statistically plausible, if not necessarily clinically plausible, projections of OS.

The ERG considers that all of the ICERs estimated using the company OS projections – including the ERG model amendments - should therefore be treated as 'what if?' scenarios as they are not underpinned by statistically significant clinical effectiveness evidence.

Even if the company's OS projection was accurate, the company has underestimated the acquisition costs of osimertinib and failed to take into account any administration cost of osimertinib as an oral chemotherapy. Using time to treatment discontinuation data (TTD) from the AURA studies and the IMPRESS trial and a cautious estimate of the NHS Reference Cost for oral chemotherapy administration results in substantial increases in the size of the ICER per QALY gained from the company base case.

The ERG also considers that the utilities applied in the company model appear to be implausible as they are higher in the PF state (0.815) than the general population norm for patients of the same age at the start of the model (0.80). Whilst no utility values are available specifically for the population described in the CS, the ERG considers that there are alternative utility values that, whilst they are by no means perfect, may be closer to the actual values of the target population compared to the utility values used in the model.

The ERG did not identify any statistically significant difference in PFS and OS by line of treatment for osimertinib and did not consider the evidence on single-agent chemotherapy to

be convincing. As such, the ERG does not consider the results of the company's subgroup analyses to be informative.

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

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

- the available clinical effectiveness data from the IMPRESS trial suggest that patients previously treated with an EGFR-TKI have a median OS of less than 24 months
- the results of the company's economic modelling suggest a mean OS gain of over 12 months with osimertinib compared with PDC
- the number of patients eligible for treatment with osimertinib is 300 per year.

1.7 ERG commentary on End of Life criteria

The ERG agrees with the company that, in England, approximately 300 patients each year will be eligible for treatment with osimertinib. The ERG considers that patient life expectancy in the second-line setting life is less than 24 months. The ERG's view is that the OS benefit from treatment with osimertinib cannot be established with any confidence until more mature OS data are available.

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

1.8.1 Strengths

Clinical evidence

- The company provided a detailed submission that fulfilled as many requirements of the final scope issued by NICE as is currently possible given the available clinical effectiveness data
- The AURAext and AURA2 studies were of good methodological quality and included a BICR of radiological outcomes
- The company made use of the IPD available from the IMPRESS trial
- The ERG's requests for further clinical information were fulfilled promptly and to a good standard

Cost effectiveness evidence

- The economic model was well constructed, easy to navigate and there were no flaws in the algorithms
- The company has undertaken a large number of subgroup and scenario analyses to explore the impact of the uncertainty in the OS data
- The company went to great lengths to compare the cost effectiveness of osimertinib to chemotherapy even when no head-to-head trial data were available.

1.8.2 Weaknesses and areas of uncertainty

Clinical evidence

- There is no RCT evidence available to support the use of osimertinib for locally advanced or metastatic EGFR T790M mutation-positive NSCLC for any line of treatment
- The clinical evidence supporting treatment with osimertinib in the CS is derived from two single-arm studies
- The pooled OS data derived from the AURAext and AURA2 studies are very immature (12.7% mature)
- The company was unable to compare osimertinib with 11 of the comparators listed in the final scope issued by NICE due to a lack of relevant clinical effectiveness evidence
- The results of the company's unadjusted and adjusted analyses should be treated with caution due to the limited and immature survival data available
- The company's use of references in the CS was confusing and often inaccurate.

Cost effectiveness evidence

- There is no clinical or statistically significant basis to support any difference in OS between osimertinib and PDC. As such, there is no basis to project a difference in OS in the company model
- The use of PFS data, rather than TTD data, underestimates the cost of osimertinib treatment and overestimates the cost of PDC treatment
- The utility values used in the company model are high. There are alternative utility values that the ERG considers to be more plausible than those used by the company
- Treatment with osimertinib statistically significantly improves PFS compared to treatment with PDC. The ERG considers that the company base case should comprise a PFS gain for osimertinib and no OS gain. The ERG considers that hypothetical OS gains should be employed only in the company's scenario analyses.

1.9 Summary of exploratory and sensitivity analyses undertaken by the ERG

The ERG suggested five amendments that could be made to the company model. Two amendments suggested changes to the costs in the company model, two offered alternative utility values and a final amendment removed the OS gain for osimertinib over PDC with only a gain in PFS for osimertinib compared to PDC remaining.

The ERG also noted minor errors related to AE costs, discounting and the PDC costs per dose. However, as the impact of correcting these minor errors would only have a small impact on the size of the ICERs, the ERG did not include these minor errors when compiling the list of suggested model amendments. Similarly, the testing costs for the EGFR T790M mutation were estimated to have only a small impact on the size of the ICERs whether they were included or not in the model.

Application of the ERG changes to costs and the ERG's alternative utility values resulted in ICERs for osimertinib compared to PDC of [redacted] per QALY gained or [redacted] per QALY gained depending on the source of the alternative utility values used in the model. In addition, when only the improvement in PFS with osimertinib is included (i.e.., there is no OS gain as only PFS is statistically significantly improved for osimertinib versus PDC) then the ERG estimates the ICER for osimertinib compared to PDC to be [redacted] per QALY gained or [redacted]x per QALY gained, again depending on the source of the alternative utility values used in the model.

All of the ERG's revised ICERs are based on list prices.

2 BACKGROUND

2.1 Critique of the company description of underlying health problem

Section 3.1 of the company submission[1] (CS) includes an overview of locally advanced or metastatic EGFR T790M mutation-positive non-small cell lung cancer (NSCLC). Section 3.2 of the CS includes a description of the effects of the disease on patients, carers and society. Information about the life expectancy of this population in England is presented in Section 3.4 of the CS. Key points from these sections are included as bulleted items in Box 1 and Box 2. The Evidence Review Group (ERG) considers that these points appropriately summarise the underlying health problems.

Box 1 Company overview of locally advanced or metastatic EGFR T790M mutation-positive NSCLC

Lung cancer types, subtypes and incidence rates

- Lung cancer is the most common cancer worldwide and the leading cause of cancer death worldwide with an estimated annual death toll of 1.59 million people. The majority of lung cancer cases are diagnosed when patients have either locally advanced or distant metastatic disease that is not amenable to curative surgery

- Non-small cell lung cancer (NSCLC) accounts for approximately 85% of all lung cancers. Within the UK, approximately 38,000[2] people are diagnosed with lung cancer every year of which NSCLC accounts for 88%[2]

- Advanced NSCLC (aNSCLC) is further divided into subtypes depending on the molecular profile and predominant oncogenic driver of the tumour. One of these is aNSCLC with an epidermal growth factor receptor sensitising mutation (EGFRm+). The prevalence of EGFR mutations in NSCLC varies according to the different histological subtypes and patient ethnicity. In a Caucasian aNSCLC population, EGFR mutations account for approximately 10%[3] of all cases. Clinical guidelines recommend routine testing for EGFR mutations before selecting a first-line therapy for aNSCLC

- Most advanced EGFRm+ tumours initially respond to tyrosine kinase inhibitors (EGFR-TKIs), but subsequently develop resistance to therapy on average 10–14 months after commencing treatment. This can be either due to secondary mutations or via activation of bypass signalling pathways (c-Met amplification). EGFR T790M mutations account for 50% to 60%[4-6] of all cases of acquired resistance. The EGFR T790M mutation is rarely detected (approximately 1%[3] of patients) in EGFR-TKI naive tumours (also known as de novo or primary mutations).

Source: CS, Section 3.1

Box 2 Company's overview of effects of the disease on patients, carers and society

Prognosis

- The 5-year survival rate for patients in the UK who are diagnosed with locally advanced (stage IIIB) NSCLC is very low at 7% to 9% and an even worse prognosis (5 year survival equal to 1%) is associated with the presence of distant metastases (stage IV)
- Treatment with EGFR-TKIs has resulted in an improved life expectancy for patients with EGFRm+ disease of approximately 20 to 24 months[7-12] from the point of initial diagnosis
- At disease progression and for patients who develop EGFR-TKI resistance and who are treated with platinum doublet chemotherapy, median overall survival (OS) is approximately 17 months according to the results of the IMPRESS[13] trial
- The prognostic role of the EGFR T790M mutation is not fully understood. In the dataset from the IMPRESS[13] trial, median progression-free survival (PFS) was consistent (5.3 months and 5.4 months) for EGFR T790M mutation-positive and EGFR T790M mutation-negative patients respectively. The OS Kaplan-Meier (K-M) plots between the EGFR T790M mutation-positive and negative control group (treated with platinum-doublet chemotherapy) showed a degree of separation from 12 months onwards.

Effects of disease on carers and society

- Lung cancer is the most common cause of death in the UK, accounting for more than one in five cancer deaths. In 2012, 35,751 deaths from lung cancer across the UK were recorded
- In the UK, costs associated with lung cancer exceed the cost of all other cancer types
- For the year 2013 to 2014, hospital admissions in England associated with lung cancer (ICD10 C34) reached 88,350 and accounted for 108,216 completed consultant episodes and 282,717 bed days
- Most NSCLC patients experience multiple symptoms; the majority of metastatic patients endure three or more, with cough, pain and dyspnoea being the most common
- NSCLC symptoms directly affect physical functioning and well-being. This has a direct impact on patients' health related quality of life, which is significantly reduced amongst patients with early stage disease
- Chemotherapy is associated with acute, potentially life-threatening side-effects and serious longer-term toxicities**.** Chemotherapy-treated patients require frequent clinic visits
- NSCLC can cause a burden for people who provide informal care due to its direct psychological impact
- As the disease progresses, informal care givers may also experience an economic burden due to taking time off work whilst caring.

Source: CS, Section 3.2 and Section 3.4

Clinical advice to the ERG is that patients with EGFR mutation-positive (EGFRm+) NSCLC have a better prognosis than patients in an unselected advanced NSCLC population as they are younger and have fewer co-morbidities.

The ERG agrees with the company (CS, p51) that the advent of treatment with EGFR-TKIs has led to increased life expectancy for patients with EGFRm+ disease. The ERG notes that in the pivotal trials[7-12,14-16] exploring the efficacy of EGFR-TKIs, overall survival (OS) results of up to 34 months[17] have been reported.

The ERG notes that the OS outcomes from the IMPRESS[13] trial (reported in

Box 2) are preliminary outcomes as the data are immature. In addition, the data are derived from patients in the control arm of the IMPRESS trial who were retrospectively identified as having EGFR T790M mutation-positive NSCLC.

The ERG understands that the prevalence of EGFRm+ disease in the NSCLC population varies according to histological subtype and ethnicity. The prevalence in a Caucasian population is approximately 10%[3] but may be greater in other ethnicities.[17] A small proportion of patients (approximately 1%[3]) have EGFR T790M mutation-positive disease at first diagnosis. Of the EGFRm+ patients whose first-line treatment is an EGFR-TKI, between 50% and 60% are found to have developed the EGFR T790M mutation at disease progression.[4-6]

2.2 Critique of the company overview of current service provision

An overview of current service provision is presented in Section 3.3 of the CS. The company discusses the appropriate published NICE guidance[18-20] and international treatment guidelines in Section 3.5 of the CS.

It is correctly reported in Section 3.5 of the CS that there is no published NICE guidance or international guidelines that are tailored specifically to the treatment of patients with EGFR T790M mutation-positive advanced or metastatic NSCLC.

In the UK NHS, patients who have EGFRm+ disease at diagnosis are treated with an EGFR-TKI, either gefitinib, erlotinib or afatinib (TA192,[18] TA258[19] and TA310[20] respectively). Clinical advice to the ERG is that, for the (very few) patients who are identified as having primary EGFR T790M mutation-positive disease at diagnosis (approximately 1%[3]), treatment starts with an EGFR-TKI followed by an early switch to platinum doublet chemotherapy (PDC), usually pemetrexed+cisplatin, at the clinician's discretion. Not all centres routinely test for the EGFR T790M mutation.

The company presents a treatment algorithm outlining the existing treatment pathway for patients with EGFRm+ NSCLC and the anticipated NHS treatment pathway for patients with EGFR T790M mutation-positive disease (see Figure 1). The ERG considers that the algorithm presented by the company reflects current clinical practice and would capture the treatment pathway in the event that osimertinib were recommended by NICE for use in the NHS. The ERG agrees with the company that pemetrexed with cisplatin or carboplatin is usually offered following disease progression on an EGFR-TKI. The ERG notes that nintedanib+docetaxel is now recommended by NICE for use after failure of first-line chemotherapy.

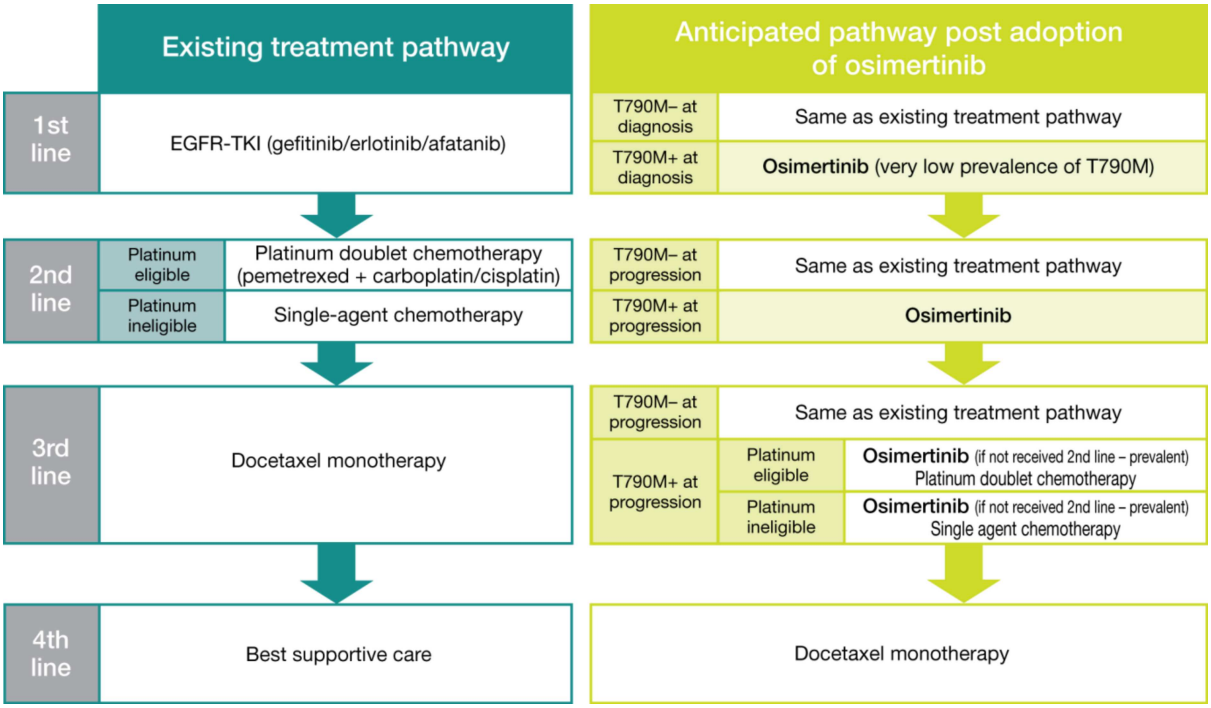


Figure 1 NHS treatment algorithm presented by the company

Source: CS, Figure 3.2

The ERG notes that the European Medicines Agency (EMA) marketing authorisation[21] is for the treatment of adult patients with locally advanced or metastatic EGFR T790M mutation positive NSCLC. The marketing authorisation does not specify a particular line of treatment; however, the company expects osimertinib will mainly be used as a second-line treatment, after failure of an EGFR-TKI (CS, p53). The company discusses the use of osimertinib as first- and third-line and further treatments in the NHS and states that:

- the number of patients treated at first-line is likely to be small given the low prevalence of the EGFR T790M mutation at diagnosis
- an existing third-line patient population is a 'one-off' group.

2.2.1 Testing for the EGFR T790M mutation in the NHS

The ERG is aware that, in the NHS, mutation testing is currently carried out via tissue biopsy. Patients are routinely tested for the presence of the EGFR mutation at diagnosis; very few tissue biopsies are carried out after treatment has commenced. Testing for the EGFR T790M mutation is not routinely carried out in the NHS either at diagnosis or after treatment failure with a first-line EGFR-TKI. The company acknowledges (CS, p44) that tissue biopsy at disease progression after treatment with an EGFR-TKI is not routinely carried out in the NHS and its introduction will therefore necessitate a change in service provision.

Company's anticipated testing protocol in the NHS

The company observes (CS, p44) that blood plasma testing (ctDNA) is becoming available to cancer patients in the NHS; however, ctDNA testing carries a high false negative rate. In line with the Summary of Patient Characteristics (SmPC),[21] the company states that all patients with a negative result for the EGFR T790M mutation following a ctDNA test should be retested using a tissue biopsy.

The company points out (CS, p44) that ctDNA testing mitigates the complications associated with the acquisition of lung tissue samples (for example, pneumothorax, infection and bleeds) and may be a preferred option for patients with later stage disease and poor performance status (PS).

The company states (CS, p45) that although the tissue testing pathway is well established within the NHS (particularly in the first-line setting), ctDNA testing is a less expensive alternative and offers more rapid results. The company reports that feasibility studies into the pathway for ctDNA processing within the NHS are expected to begin in the second quarter of 2016.

The CS (p45) includes an algorithm to illustrate the optimal testing pathway for EGFR T790M mutation status in the NHS (Figure 2) as perceived by the company. Clinical advice to the ERG is that tissue biopsy testing is available at approximately 85% of NHS treatment centres and is conducted mainly at diagnosis. Not all centres test for the presence of the EGFR T790M mutation at diagnosis and very few centres currently re-biopsy patients after treatment is initiated. This means that the use of osimertinib in the NHS will require a change in practice to facilitate EGFR T790M mutation testing following disease progression on a first-line EGFR-TKI. Given the low estimated incidence and prevalence rate of primary EGFR T790M mutation-positive disease, it is unlikely that patients will ever be routinely tested for the EGFR T790M mutation at diagnosis. Ideally, until a ctDNA test is available to the NHS, all patients would be offered a tissue biopsy at relapse after a first-line EGFR-TKI to determine their EGFR T790M mutation status and inform second-line treatment decisions; however, clinical advice to the ERG is that there are concerns about patients' willingness to tolerate the biopsy procedure.

The ERG notes that the SmPC[21] for osimertinib recommends that patients with a negative ctDNA test result should have the result confirmed by a tissue biopsy test. This means that in the absence of a ctDNA test that is 100% sensitive, the 40% to 50% of patients who do not have the EGFR T790M mutation will be offered a tissue biopsy test to confirm the negative ctDNA test result.

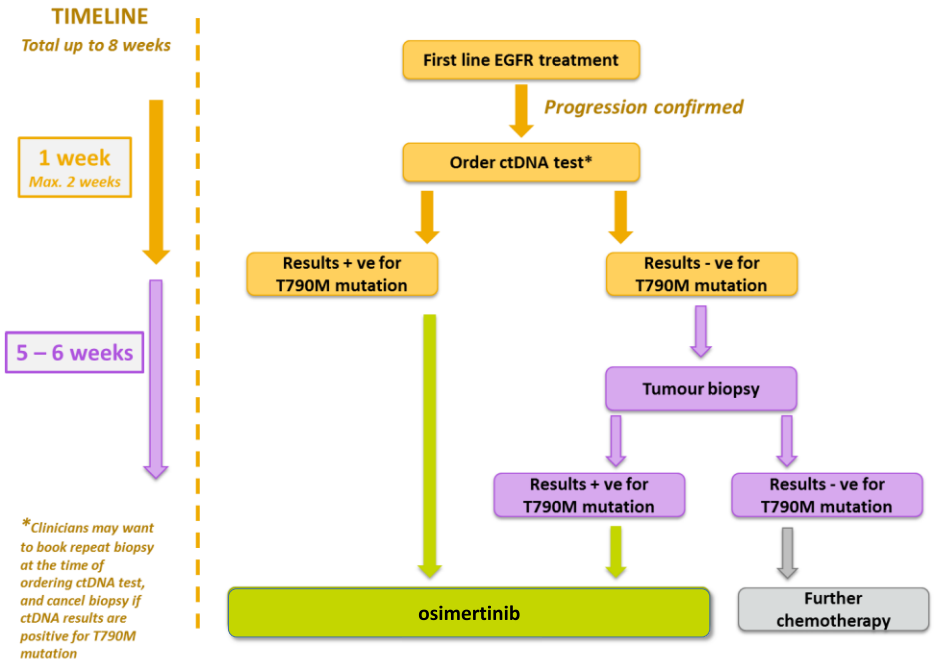


Figure 2 Company's anticipated optimal NHS testing pathway for EGFR T790M mutation status

Source: CS, Figure 2.1

2.3 Innovation

The company states (CS, p46) that osimertinib was included in the Medicine and Healthcare Regulatory Agency (MHRA) Early Access to Medicines Scheme (EAMS).[22] The purpose of the EAMS is to give patients with life-threatening or debilitating conditions access to medicines that do not yet have a marketing authorisation when there is a clear unmet clinical need.

The company provides text from the MHRA assessment report[22] (CS, p46) to support the claim that osimertinib is an innovative treatment:

> *'EGFR T790M mutation-positive lung cancer is a life threatening disease. Patients with this condition have very limited treatment options, reduced life expectancy and there is an urgent need for more therapies. In clinical studies, osimertinib was able to slow or shrink the cancer in these patients. Other currently available treatments have limited activity. The MHRA has considered the benefits of osimertinib in this difficult to treat condition and concluded that the benefits are greater than the risks'.*

The ERG agrees with the MHRA that treatment with osimertinib appears promising but cautions that the available data are derived from two single-arm studies (AURAext[23] and AURA2[24]) and the survival and safety data are currently very immature. The final outcome

data from the ongoing AURA3[1] phase III randomised controlled trial (RCT) will be more robust.

The ERG agrees with the company (CS, p58) and the MHRA that patients with EGFR T790M mutation NSCLC are a group of patients with no specific treatments available to them. The ERG considers that osimertinib appears to be better tolerated than treatment with pemetrexed+cisplatin and that, in terms of drug administration, patients generally prefer the oral method of administration (osimertinib) to intravenous infusion (PDC).

The company reports (CS, p29) that osimertinib is the first drug to be approved under the EMA's EU PRIME scheme.[25] The purpose of the EU PRIME scheme is to provide support for the development of medicines that target an unmet medical need. When the application for marketing authorisation is submitted to the EMA, medicines awarded EU PRIME status are eligible for accelerated assessment.

2.4 Number of patients eligible for treatment with osimertinib

The company estimates (CS, p245) that approximately 300 patients every year are likely to be eligible for treatment with osimertinib (Table 1). The ERG considers the company's estimate to be reasonable but notes that the figure of 10% used to estimate the proportion of patients with EGFRm+ tumours relates to Caucasian patients and that this figure is higher in other ethnic groups.[17]

Table 1 Company estimate of the number of patients in England eligible for treatment with osimertinib

Parameter	Estimated proportion of patients	Number of patients (incident)	Number of patients (prevalent)
Lung cancer diagnosis[2]		31,393	25,276
Confirmed NSCLC [2]	59%	18,447	14,853
Patients with stage III/IV disease[2]	77%	14,204	11,437
Patients tested for EGFR mutation status[2]	87%	12,372	9,961
Patients with EGFR mutation*	10%	1,237	996
EGFRm+ patients receiving 1st-line anti-cancer treatment [2]	58%	713	574
EGFRm+ patients who progress on 1st-line anti-cancer treatment and receive active treatment*	Incident 65% Prevalent 50%	463	287
Patients with T790M mutation – eligible population[6,26]	60%	278	172

Source: CS, Table 6.1
*AstraZeneca internal research

3 CRITIQUE OF COMPANY'S DEFINITION OF DECISION PROBLEM

A summary of the decision problem described by the company in the CS in relation to the final scope issued by NICE[27] is presented in Table 2. Each parameter is discussed in more detail in the text following the table.

Table 2 Summary of parameter details included in the final scope issued by NICE and the company's decision problem

NICE scope **Parameter and specification**	**Decision problem addressed in the company's submission**
Population People with locally advanced or metastatic EGFR T790M mutation-positive NSCLC	As per final scope, except that no cost effectiveness results are presented for the treatment-naïve population
Intervention Osimertinib	As per final scope
Comparator(s) For people who have not received previous treatment: afatanib, erlotinib and gefitinib	Very limited evidence describing the clinical effectiveness of erlotinib and afatanib is presented in the CS (p166-171) for previously untreated patients
For people who have received previous treatment with an EGFR-TKI: PDC (including pemetrexed+carboplatin or pemetrexed+cisplatin)	As per final scope The base case cost effectiveness analysis compares osimertinib with PDC in ≥second-line patients A subgroup analysis is provided to compare the cost effectiveness of osimertinib with PDC in second-line patients
For people who have received previous treatment with an EGFR-TKI and in whom PDC is not appropriate: single-agent chemotherapy including gemcitabine, paclitaxel, vinorelbine and docetaxel	As per final scope Clinical effectiveness data are limited. A subgroup analysis is provided to compare the cost effectiveness of osimertinib with docetaxel in second-line patients
For people who have received previous treatment with an EGFR-TKI and chemotherapy: docetaxel (+/- nintedanib), nivolumab (subject to ongoing NICE appraisal), ramucirumab (subject to ongoing NICE appraisal), single agent chemotherapy including gemcitabine, paclitaxel, vinorelbine (for those for whom treatment with docetaxel is not appropriate) and BSC	Clinical effectiveness data are limited and only relate to treatment with single-agent chemotherapy A subgroup analysis is provided to compare the cost effectiveness of osimertinib with single-agent chemotherapy at ≥third-line
Outcomes The outcome measures to be considered include: PFS, OS, ORR, AEs and HRQoL	As per final scope
Economic analysis The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per QALY 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 an NHS and Personal Social Services perspective	As per final scope

NICE scope Parameter and specification	Decision problem addressed in the company's submission
The availability of any patient access schemes for the comparator technologies should be taken into account The use of osimertinib is conditional on the presence of the T790M mutation in the EGFR gene. The economic modelling should include the costs associated with testing for EGFR T790M mutations in people with NSCLC who would not otherwise have been tested. A sensitivity analysis should be provided without the cost of additional testing	
Subgroups to be considered None (other than the subgroup populations mentioned in population)	As per final scope
Other considerations None	As per final scope

AE=adverse event; EGFR=epidermal growth factor receptor; HRQoL=health related quality of life; NICE=National Institute for Health and Care Excellence; NSCLC=non-small cell lung cancer; ORR=objective response rate; OS=overall survival; PDC=platinum doublet chemotherapy; PFS=progression-free survival; QALY=quality adjusted life year; TKI=tyrosine kinase inhibitor; BSC=best supportive care.
Source: CS, adapted from Table 1.1

3.1 Osimertinib clinical evidence

There is no direct clinical evidence comparing osimertinib with any of the comparators listed in the final scope issued by NICE. To compare osimertinib versus PDC (base case comparator) the company had to (i) pool very immature clinical data from two single-arm studies (ii) retrospectively identify patients in the control arm of the IMPRESS trial who tested positive for the EGFR T790M mutation and (iii) carry out an unadjusted and an adjusted treatment comparison. Consequently, the ERG is concerned that the clinical evidence presented by the company to support the use of osimertinib in patients who have received previous treatment with an EGFR-TKI is not robust. Furthermore, as the OS data from the pooled AURA[1] dataset were only 12.7% mature at the time of writing the CS, there are no reliable long-term safety outcome data available.

3.2 Population

The population described in the final scope issued by NICE is people with locally advanced or metastatic EGFR T790M mutation-positive NSCLC. This is the same as the population described in the conditional licence for osimertinib issued by the EMA.[3]

Treatment line is not specified in either the final scope issued by NICE or in the conditional EMA licence.[28] The company expects osimertinib to be used as a second-line treatment following relapse whilst receiving first-line treatment with an EGFR-TKI. The clinical evidence submitted by the company is derived from two single-arm studies (AURAext and AURA2) that were designed to assess the clinical effectiveness of osimertinib in patients with locally advanced or metastatic EGFR T790M mutation-positive NSCLC who had received treatment with an EGFR-TKI prior to recruitment.

The company explains (CS, p166 to p167) that clinical effectiveness data to support the use of osimertinib as a first-line treatment are limited to the experience of five patients who were treated with osimertinib as part of the AURA phase I extension study. The company states that the EMA's decision to grant a licence for the use of osimertinib in all treatment lines was based on a biological assumption of its effectiveness as a first-line treatment as no clinical studies have been conducted in treatment-naïve patients. The ERG accepts the company's explanation and notes that the European Public Assessment Report (EPAR[28]) issued by the EMA confirms the company's explanation.

For the specified patient population, clinical advice to the ERG is that patients with EGFRm+ disease who are treated in the NHS are typically aged between 65 years and 70 years and the majority are of ECOG PS 1 or 2. The ERG notes that patients in the studies discussed in the CS, AURAext and AURA2, are younger (median 62.2 years) and fitter (ECOG PS 0 or 1) than EGFRm+ patients treated in the NHS. Similarly, patients in the IMPRESS trial are also younger (mean age of 58.1 years) and fitter (ECOG PS 0 or 1) than EGFRm+ patients treated in the NHS.

The ERG notes that 12.4% of patients recruited to the AURAext and AURA2 studies had received more than five lines of prior treatment. Clinical advice to the ERG is that the majority of patients treated in the NHS are not well enough to tolerate more than one or two chemotherapy treatments after a first-line EGFR-TKI.

3.3 Intervention

The intervention specified in the final scope issued by NICE is osimertinib. Osimertinib is a small molecule irreversible inhibitor that targets the sensitising and EGFR T790M mutant forms of the EGFR-TKI.[27] It has a conditional licence in Europe for the treatment of adults with locally advanced or metastatic EGFR T790M mutation-positive NSCLC.[3] The licence is conditional on the company providing the EMA with the results of a phase III RCT (AURA3) by July 2017 (CS, p40). The company expects that, in NHS clinical practice, osimertinib will mainly be used as a second-line treatment, after failure with an EGFR-TKI.

Osimertinib is available as a film coated tablet (40mg or 80mg). The daily dose is 80mg until disease progression or unacceptable toxicity.

3.3.1 Testing for the EGFR T790M mutation

In the SmPC[21] for osimertinib it is stipulated that treatment should only be initiated after the patient's EGFR T790M mutation status is positively confirmed by a clinical laboratory test using a validated test method. It is further cautioned in Section 4.4 of the SmPC[21] that

patients' whose plasma-based ctDNA test results in a negative outcome should receive a follow-up tissue biopsy test as approximately 20% of plasma tests are believed to give false negative results (CS, p210). The company discusses the issues relevant to testing for the EGFR T790M mutation within the NHS in the CS (p43 to p45).

The company states that the pathway for acquiring, handling and testing tissue samples and reporting results is already well established in the NHS as up to 90% of treatment-naïve patients with NSCLC are routinely tested for the EGFR mutation. The company has confirmed with the UK National Quality Assessment Services (NEQAS) that the majority of laboratories (88%) currently validated for testing for EGFR mutations are able to test for the EGFR T790M mutation using existing test platforms. The company is confident that EGFR T790M mutations are already being identified and states that there are, therefore, no additional cost implications in terms of equipment, reagent or manpower. The company recognises, however, that, if osimertinib is recommended for use in the NHS, the volume of tests required to identify the presence of the EGFR T790M mutation is likely to increase, based on the increasing use of biopsy at relapse.

3.4 Comparators

There are a number of comparators listed in the final scope issued by NICE and these vary by line of treatment.

The company has provided clinical effectiveness evidence for the comparison of osimertinib with PDC for EGFR T790M patients who have been previously treated with an EGFR-TKI. The evidence for the effectiveness of PDC was obtained following a retrospective analysis of tumour samples from patients recruited to the control/PDC arm of the IMPRESS trial. Although 132 patients were recruited to this arm of the IMPRESS trial, a retrospective analysis showed that only 61 patients had tumours expressing the EGFR T790M mutation. The company used data from this subgroup of patients from the IMPRESS trial and from the pooled AURA dataset in an adjusted treatment comparison to allow the clinical efficacy of treatment with osimertinib to be compared with PDC.

The ERG agrees with the company that:

- there are no clinical effectiveness data that directly compare osimertinib with any of the other 11 comparators specified in the final scope issued by NICE
- there are no clinical effectiveness data available, either for an EGFR T790M mutation-positive population, or for an EGFRm+ population, that would allow a treatment comparison to be carried out to inform a robust comparison of osimertinib with any of the other comparators specified in the final scope issued by NICE. The only clinical evidence available allows comparison of osimertinib with PDC.

However, the ERG notes that the company uses subgroup analyses to consider additional comparators in the cost effectiveness section of the CS. Based on the information presented in the CS, the ERG assumes that relevant survival data are used directly in these analyses. For second-line patients only, osimertinib is compared with PDC and also with docetaxel. For ≥third-line patients, osimertinib is compared with single-agent chemotherapy (docetaxel). The ERG considers that all of the economic subgroup analyses rely on limited clinical evidence.

The company makes strong assumptions regarding the choice of comparators used in the economic analyses. First, the company assumes that the clinical evidence describing 'placebo+pemetrexed+cisplatin' from the IMPRESS trial can be used to represent PDC. The ERG notes that, in clinical practice, PDC may also comprise other treatments e.g. vinorelbine, gemcitabine, docetaxel or paclitaxel in combination with cisplatin or carboplatin. However, the ERG also acknowledges that, for the specified patient population, pemetrexed+cisplatin is the most commonly used PDC in the UK NHS. The company also assumes that docetaxel monotherapy efficacy data from patients (untested for EGFR T790M mutation) that are described in the studies by Park[29] and by Schuler[30] can be used to represent the efficacy data associated with any single-agent chemotherapy for the treatment of second-line and third-line or further patients, respectively, with EGFR T790M mutations.

3.5 Outcomes

Clinical evidence is reported in the CS for all of the outcomes specified in the final scope issued by NICE, overall survival (OS), progression-free survival (PFS), response rate (reported as objective response rate [ORR], disease control rate [DCR] and duration of response [DoR]), adverse events (AEs) of treatment and health related quality of life (HRQoL). The ERG notes that the OS data that are currently available from the pooled AURA dataset and the IMPRESS trial are still very immature (12.7% and approximately 33% respectively).

3.6 Economic analysis

As specified in the final scope issued by NICE, the cost effectiveness of treatments was expressed in terms of the incremental cost per quality adjusted life year (QALY) gained. Outcomes were assessed over a 15-year time period (equivalent to a lifetime horizon) and costs were considered from an NHS perspective. The company's economic model includes the costs associated with four possible testing strategies to identify patients with EGFR T790M mutation-positive tumours.

3.7 Subgroups

No subgroups were specified in the final scope issued by NICE in addition to the distinct patient populations specified in the comparator section.

3.8 Other considerations

The company did not identify any equality issues. The ERG is aware that the company has submitted a Patient Access Scheme (PAS) proposal to the Department of Health. The list prices of osimertinib, cisplatin, pemetrexed and docetaxel are used in all of the cost effectiveness analyses presented in the CS.

4 CLINICAL EFFECTIVENESS

This section provides a structured summary and critique of the clinical effectiveness evidence submitted by the company in support of the use of osimertinib for the treatment of patients with locally advanced or metastatic EGFR T790M mutation-positive NSCLC.

4.1.1 Systematic review methods

The company conducted a systematic review to identify studies of relevance to the appraisal under discussion. A summary of the systematic review methods employed by the company, with accompanying ERG comments, is presented in Table 3. Full details can be found in the CS (p60 to p72).

Overall, the ERG is satisfied that the company's systematic review methods were of an adequate standard, and were relevant to the final scope issued by NICE and to the company's decision problem. The ERG notes that, in the systematic review, the company has restricted the patient population to those who have failed treatment with an EGFR-TKI (i.e. the previously treated population). The ERG agrees with company that this is appropriate as evidence suggests that only 1% of patients are likely to have tumours with the EGFR T790M mutation at diagnosis.[3]

Table 3 Summary and ERG comment on the systematic review methods used by the company

Review method	ERG comment
Searching	
• RCT and non-RCT data searches • Databases searched included Medline, Medline in Process, Embase and CENTRAL (search strategies are described in Appendix CS, 1.1) from inception to 4th January 2016 • Grey literature was searched for clinical studies and conference abstracts	• The company states that, due to lack of data specific to the EGFR T790M mutation-positive population, a broad search was carried out to include advanced or metastatic NSCLC and EGFRm+ patients regardless of EGFR T790M status. The ERG considers it appropriate to widen the search criteria • As expected, due to the recent drug name change, the drug terms used by the company do not include the term 'osimertinib' but do include 'AZD9291' • The company limited the patient population to those who were pre-treated with at least one EGFR-TKI. The final scope issued by NICE and the licensed indication do not specify a particular line of treatment. The company states (CS, p59) that there are very limited data on the use of osimertinib in a treatment-naïve NSCLC population. The ERG agrees with the company • The ERG was able to replicate the searches • The company searched the appropriate conference abstracts • The ERG verified the data in the PRISMA flowchart presented in the CS via the clarification process • The ERG is confident that no relevant studies were missed
Eligibility criteria	
• Two independent assessors assessed study eligibility	• Use of two independent assessor improves the quality of reviews • Only articles published with full-text in the English language were considered • The patient population is defined in the inclusion criteria as those with advanced or metastatic NSCLC with acquired EGFR/or T790M mutation and at least one prior EGFR-TKI therapy. The patient population in the final scope issued by NICE and in the licence is not restricted to a particular line of therapy. The ERG accepts that there is no clinical evidence relevant to treatment in treatment-naïve patients
Data extraction	
• Two independent assessors extracted data • A pre-defined extraction form was used	• The company has not reported the method used to extract study data. Quality assurance regarding data extraction is therefore uncertain
Quality assessment and risk of bias	
• Descriptive critical appraisal of all included RCTs and non-RCTs was undertaken using the NICE recommended method[31]	• Unclear if two independent assessors were employed • The Downs and Black[32] appraisal tool was applied to the non-RCTs. The included RCT was appraised using a hybrid of criteria derived from the Jadad[33] scale and the recommendations of Centre for Reviews and Dissemination at the University of York.[34] The ERG considers the company quality assessment strategy is appropriate.

EGFR=epidermal growth factor receptor; ERG=Evidence Review Group; NSCLC=non-small cell lung cancer; PRISMA=Preferred Reporting Items for Systematic Reviews and Meta-analyses; RCT=randomised controlled trial; TKI=tyrosine kinase inhibitor;
Source: CS Table 4.21

4.1.2 Evidence synthesis

No RCTs comparing osimertinib with any treatment in patients with EGFR T790M mutation-positive NSCLC were identified.

The company presents direct evidence for the clinical efficacy of osimertinib from two single-arm studies, the AURAext study and the AURA2 study. The CS includes a narrative description of the AURAext and AURA2 studies and results from the analysis of data pooled from these two studies (i.e., pooled AURA dataset).

4.2 Critique, analysis and interpretation of trials of the technology

4.2.1 Identified studies

Key studies

The company presents evidence for the clinical effectiveness of the intervention from the AURAext (phase I/II) and the AURA2 (phase II) studies. Both are single-arm studies involving patients with EGFR T790M mutation-positive NSCLC treated with osimertinib. The company identified an ongoing phase III RCT (AURA3) comparing osimertinib with PDC; this study is not due to report until 2017. A dosing study, AURA1, has also been undertaken. However, the company considers that results from this study are not relevant to the current appraisal.

In the absence of any trial evidence comparing the clinical effectiveness of osimertinib with any of the comparators specified in the final scope issued by NICE, the company has employed evidence from the subgroup of patients in the control arm of the IMPRESS trial whose tumours were (retrospectively) identified as having the EGFR T790M mutation. The IMPRESS trial is a phase III RCT in which patients with EGFRm+ NSCLC who had progressed on treatment with an EGFR-TKI were randomised to receive either gefitinib+cisplatin+pemetrexed (intervention) or placebo+cisplatin+pemetrexed (control).

Other studies

In total, the company's search identified ten studies, the AURAext[23] and AURA2[24] study (combined as AURA pooled dataset[1]), the IMPRESS trial,[13] and eight other studies[29,35-41] that could potentially be used to inform an indirect comparison of the clinical effectiveness of osimertinib with a comparator specified in the final scope issued by NICE. The eight studies[29,35-41] assessed the use of PDC (n=3) or single-agent chemotherapy (n=5) in patients who had previously received first-line EGFR-TKI treatment. The majority of recruited patients had confirmed EGFRm+ NSCLC; however, none of the studies included testing for the presence of the EGFR T790M mutation.

Six[29,37-41] of the eight studies were retrospective observational studies rejected by the company for the following reasons: i) small patient numbers ii) the patients observed in the studies differed from the patients included in the pooled dataset of the AURAext and AURA2 studies iii) the definitions of key endpoints in the studies were considered to be inconsistent with the definitions used in the prospective studies[23,24,35,42] (CS, p67). The two other studies, both RCTs,[29,30] were considered by the company to be inappropriate due to the small number of patients recruited to the comparator arm in each trial. The ERG agrees with the company's assessment.

The company presents a summary of details about the included and excluded studies in Table 4.4 of the CS (p68). The ERG is not aware of any other studies relevant to the decision problem.

4.2.2 Methodological approach for the synthesis and analysis of data from key studies

The company employed two approaches to compare the effectiveness of osimertinib (using pooled data from the AURAext and AURA2 studies) with PDC (using data from the IMPRESS trial): an unadjusted comparison and a comparison that included adjustments for differences in patient baseline characteristics. The company explains that data from the AURAext and AURA2 studies were pooled to increase the precision of the estimate of the primary endpoint.

4.2.3 Statistical approach adopted for the conduct and analysis of data from included studies

A full description and critique of the AURAext and AURA2 studies and the IMPRESS trial is presented in this section of the ERG report. Information relevant to the statistical approach taken by the company to analyse data from these sources has been taken directly from the clinical study reports (CSR),[23,24,42] the statistical analysis plans (TSAP),[43-45] the protocols and from the CS.

Trial populations

For the AURAext and AURA2 studies, all efficacy outcomes other than PFS, the sensitivity analysis of ORR and best objective response (BOR) by blinded independent central review (BICR), and investigator RECIST outcomes were analysed using the 'evaluable for response' analysis population. This specific population comprises patients who received at least one dose of osimertinib and had measureable disease at baseline according to an independent review of imaging data. PFS, the sensitivity analyses of ORR and BOR by BICR, and investigator RECIST outcomes were analysed using the full analysis set (FAS) population; this population comprises all patients who received at least one dose of osimertinib.

For the analysis of all efficacy outcomes in the IMPRESS trial, the FAS population was used. The FAS population follows the intention-to-treat (ITT) principle so all patients were analysed according to the treatment arm to which they were initially randomised, regardless of which treatment they actually received. Safety outcomes were analysed using the safety analysis set, consisting of all patients who received at least one dose of study medication.

Outline of analyses

Patient recruitment to the AURAext study started in May 2014 and finished in October 2014. The data cut-off for the data presented in the CS was 1st May 2015.

The AURA2 study started recruiting patients in June 2014 and the last patient was recruited in October 2014. The data cut-off for the data presented in the CS was also 1st May 2015. For the AURA2 study no formal interim analyses were planned. However, the investigators analysed the data at approximately 3 months and 8 months after the last patient was recruited. The results presented in the CS are from the 8-month data cut-off. The final database lock will be at the end of the study, at 12-24 months after the last patient was recruited.

Patient recruitment to the IMPRESS trial started in March 2012 and the last patient was recruited in December 2013. The primary data cut-off for this trial was 5th May 2014. Two data cut-offs were planned, the primary data cut-off for the primary PFS analysis and the final data cut-off for the final OS analysis. The primary PFS analysis was conducted on a total of 205 progressions (77.4% maturity). At the time of the primary PFS analysis, the OS data were also analysed (87 patient deaths had occurred, 33% maturity). [redacted]

Study outcomes

The definitions and methods of analysis for the primary and secondary efficacy outcomes from the AURAext and AURA2 studies and the IMPRESS trial are listed in

Table 4. The ERG is satisfied that all of the outcomes were pre-specified in the TSAP and that all of the outcomes were fully reported in the CSR.

Table 4 Analysis strategy for key efficacy endpoints

Endpoint	Definition	Statistical method
AURAext and AURA2: primary outcome		
ORR	The percentage of patients with at least one visit response of CR or PR that was confirmed at least 4 weeks later according to RECIST 1.1 by BICR	The analysis of ORR was presented together with 95% exact (Clopper-Pearson) CI by study and overall for the pooled AURA dataset. Overall ORR based on the pooled data was calculated as the number (%) of patients with BOR of confirmed CR or PR from both studies. A similar analysis of ORR was also presented by treatment cohort (2nd- versus ≥3rd-line) and overall. The ORR in each treatment cohort based on the pooled data was calculated as the number (%) of patients with BOR of confirmed CR or PR from each treatment cohort across two studies
AURAext and AURA2: secondary outcomes		
DoR	The time from the date of first documented response, (that is	DoR (months) in responding patients based on the BICR was summarised using the median and 95% CI. The median was

	subsequently confirmed) until the date of documented progression or death in the absence of disease progression	calculated using the K-M method. The number and percentage of responding patients remaining in response at >3; >6; >9; >12 months was summarised. Analyses were presented by study and for the pooled AURA dataset. A K-M plot was presented for the overall pooled population
DCR	The percentage of patients who had a BOR of CR or PR or SD for at least 6 weeks (allowing for a 1-week visit window)	DCR presented together with 95% exact (Clopper-Pearson) CIs
Tumour shrinkage	Tumour size is the sum of the longest diameters of the TLs. The best percentage change in tumour size from baseline was determined for each patient, ie, maximum reduction from baseline or the minimum increase from baseline in the absence of a reduction from baseline based on all post-baseline assessments prior to progression or the start of subsequent therapy	To assess the depth of tumour shrinkage, the proportion of patients who achieved >30%, >50% and >75% reduction in TL tumour size was summarised descriptively. The percentage change in TL tumour size from baseline was summarised using descriptive statistics and presented for each visit
PFS	The time from date of first dose until the date of objective disease progression as defined by RECIST or death (by any cause in the absence of progression) regardless of whether the patient withdrew from osimertinib therapy or received another anticancer therapy prior to progression	PFS was displayed in a K-M plot for the pooled population. The total number of events, median PFS (calculated from the K-M plot, with 95% CIs), and the percentage PFS at 3, 6, 12 and 18 months was summarised by study and overall for the pooled AURA dataset. Similar analyses of PFS were presented by treatment cohort and overall. A K-M plot was presented for each treatment cohort
OS	The time from the date of first dose until death due to any cause	Any patient not known to have died at the time of analysis was censored based on the last recorded date on which the patient was known to be alive
IMPRESS: primary outcome		
PFS	The time from randomisation until objective disease progression as detailed in RECIST or death (by any cause in the absence of progression) regardless of subsequent treatment	The primary analysis compared PFS between treatment groups using a Cox PH model that included terms for treatment and age (<65, ≥65 years), and prior response to gefitinib (SD versus PR and CR combined). The HR (gefitinib: placebo) was estimated together with its 95% CI and p-value
IMPRESS: secondary outcomes		
OS	The time from the date of randomisation until death due to any cause. Any patient not known to be dead at the time of analysis was censored based on the last recorded date on which the patient was known to be alive	The analysis of OS compared the OS between treatment groups using a PH model adjusted for adjusted for age (<65 years, ≥65 years) and prior response to gefitinib. The HR (gefitinib: placebo) was estimated together with its 95% CI and p-value. A K-M plot of OS was presented and the median survival time from the K-M curve was presented. The median OS, 9, 12, and 18-month rates were also presented
ORR	The number (%) of patients with at least 1 visit response of CR or PR. Data obtained up until progression or last evaluable assessment, in the absence of progression were included in the assessment of ORR. This was irrespective of whether or not patients discontinued treatment or received a subsequent therapy prior to progression	The response rate was calculated for each randomised treatment based on the percentage of patients who had a BOR (based on RECIST) of CR or PR. Objective tumour response was compared between the randomised treatment groups using a logistic regression model. The model allowed for the effect of randomised treatment and the same covariates as used in the analysis of PFS. The odds ratio for treatment (gefitinib: placebo) was estimated from the model as was the 95% CI and p-value. The p-value was based on twice the change in log-likelihood resulting from the addition of a treatment factor to a model containing the covariates detailed above
DCR	The percentage of patients who achieved disease control at 6 weeks following randomisation. Disease control was defined as a best objective response of CR, PR or SD	The DCR was analysed using the same methodology as ORR

	≥6 weeks. If a patient experienced a CR/PR very shortly after starting treatment but then progressed or became NE by 6 weeks, then they were not included as having disease control at 6 weeks	
Symptoms and HRQoL	Data on symptoms and HRQoL were assessed using the FACT-L questionnaire. FACT-L has been validated with respect to its psychometric properties and sensitivity to clinical changes	The change from baseline was summarised for each of the FACT-L total score, TOI and LCS by randomized treatment, for each week that HRQoL was assessed and where there were 20 or more patients with available data across treatment groups. The mean change from baseline and 95% CI at each of these weeks were also plotted for each treatment group separately. The number and percentage of patients with each of the best overall responses were presented for each treatment group. The HRQoL improvement rates (for FACT-L total score, TOI and LCS) were summarised descriptively by treatment groups and analysed using the same methodology as ORR. The improvement rate was calculated for each randomised treatment group. The time to worsening data were analysed using a PH model including terms for treatment received and the covariates as defined for PFS. The HR and 95% CI and p-value were presented. The KM curves for time to worsening were also plotted. The median was presented, in addition to the number of patients who worsened by 3 and 6 weeks
AURA2 study only: EQ-5D-5L and EQ-VAS	To assess utilities to support health technology assessment and health economic modelling in patients	Simple summaries of the data were provided and included the frequency of response to each of the 5 questions by protocol-led visit. A summary at each protocolled visit of the expected number of questionnaires and the actual number of questionnaires received was also presented. This included the number of questionnaires received as a percentage of the expected number at each protocol-led visit. In addition, EQ-5D-5L scores and individual questions from the EQ-5D-5L questionnaire were summarised at each scheduled time point and by treatment group using descriptive statistics

BICR=blinded independent central review; BOR=best objective response; CI=confidence interval; CR=complete response; DCR=disease control rate; DoR= duration of response; EQ-5D=euro quality of life – 5 dimensions; HR=hazard ratio; KM=Kaplan-Meier; NE=not evaluable; ORR=overall response rate; OS=overall survival; PFS=progression-free survival; PH=proportional hazards PR=partial response; RECIST= Response Evaluation Criteria in Solid Tumours; SD=stable disease; TL=target lesion
Source: CS (adapted from sections 4.11.1.2 and 4.11.2.2) and protocol (Section 5.7.4.2)

Censoring methods

The censoring methods employed in the AURAext study, the AURA2 study and the IMPRESS trial are shown in Table 5.

Table 5 Censoring methods

	OS	PFS
AURAext and AURA2 studies	Any patient without documentation of death at the time of analysis was censored based on the last recorded date on which the patient was known to be alive	Patients who had not progressed or died at the time of analysis were censored at the time of the latest date of assessment from their last evaluable RECIST assessment. However, if the patient had progressed or died after three or more missed RECIST assessments visits, the patient was censored at the time of the latest evaluable RECIST assessment prior to the missed visits. If the patient had no evaluable visits or did not have baseline data they were censored at zero days unless they died within two visits of baseline
IMPRESS trial	Any patient without documentation of death at the time of analysis was censored based on the last recorded date on which the patient was known to be alive	Patients, who had not progressed or died at the time of the statistical analysis, were censored on the date of their last target lesion/non-target lesion (TL/NTL) assessment from their last evaluable RECIST assessment. If a patient had progressed or died after ≥2 missed visits, the patient was censored at the time of the latest evaluable RECIST assessment. If the patient had no evaluable visits or did not have baseline data, the patient was censored at zero days.

OS=overall survival; PFS=progression-free survival; RECIST=Response Evaluation Criteria In Solid Tumours
Source: CS, p99 and p132

Subgroup analyses

Subgroup analyses were conducted, using the pooled AURA dataset to explore ORR and duration of response (DoR) by BICR across the key subgroups listed in Table 6. The analysis of ORR together with 95% exact (Clopper-Pearson) confidence intervals (CIs) was presented by treatment cohort and overall for each category of subgroup. Forest plots of ORR by BICR were constructed for each treatment cohort and for the overall study population. DoR by BICR was summarised using the median and 95% CI by both treatment cohort and the overall population for each of the subgroup categories.

The IMPRESS trial subgroup analyses were conducted to explore the consistency of treatment effect of PFS across the key subgroups listed in Table 6. A Cox Proportional Hazards (PH) model was used to investigate the treatment effect in each of the subgroups. The hazard ratios (HRs) and 95% CIs were summarised and presented in a forest plot, along with the overall primary analysis results.

Health related quality of life

Three patient reported outcome questionnaires were used in the AURAext and AURA2 studies to collect data on the impact of osimertinib on patients' disease-related symptoms and HRQoL:

- The European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire – Core 30 items (EORTC QLQ-C30[46])
- The Organisation for Research and Treatment of Cancer Quality of Life Questionnaire – Lung Cancer 13 items (EORTC QLQ-LC13[47])
- The EuroQoL–5 dimensions–5 levels (EQ-5D-5L[48]) questionnaire (AURA2 study only).

These HRQoL outcomes were exploratory endpoints, so only summary data were collected and no statistical testing was carried out.

In the IMPRESS trial, data on symptoms and HRQoL were collected using the Functional Assessment of Cancer Therapy–Lung (FACT-L[49]) questionnaire, the EQ-5D-3L questionnaire and the EQ visual analogue scale (EQ-VAS[48]). The company presents the results from the FACT-L questionnaire in various different forms, including the results from change from baseline, the improvement rates for the FACT-L questionnaire data, the frequency of responses received for each question and the number of EQ-5D VAS completed.

The ERG considers that the approaches taken by the company to explore the available HRQoL data are acceptable.

Proportional hazards

The analyses carried out by the company to generate PFS and OS HRs to demonstrate the relative effectiveness of the two treatments assessed by the IMPRESS trial were conducted using Cox PH modelling. The validity of this method relies on the hazards of the two comparative drugs being proportional.

To investigate the assumption of PH, the company inspected Log-log plots (log cumulative hazard versus log time); if the curves for each treatment arm were approximately parallel, the assumption of PH was valid. If inspection of the Log-log plot raised any concerns, a time dependent covariate was fitted to the model to assess the extent to which any deviance from being parallel represents random variation.

As part of the clarification process, the ERG asked the company to provide details of the analyses that had been undertaken to determine whether the assumption of PHs holds for

the PFS data. The company explained that the assumption of PH was tested by examining plots of the log(event) versus log(time) time data from both arms of the IMPRESS trial. Visual examination of the graph suggested that the resultant lines were not parallel, which raised concerns about the validity of the PH assumption. To assess the extent to which the deviance from being parallel was due to random variation, a time dependent covariate was fitted to the Cox PH model. The p-value for the test of non-proportional hazards was [redacted] (provided in the company clarification response to question A6), suggesting that there is insignificant evidence of non-proportionality.

No details are provided in the CS to suggest that any testing has been carried out to test whether the assumption of PHs holds for the OS data. However, examination of the Kaplan-Meier (K-M) data presented in the CS (Figure 4.19, p154) suggests that the PH assumption may hold when considering the OS data from all patients without EGFR T790M mutation-positive disease (irrespective of treatment) and the patients in the control arm with EGFR T790M mutation-positive disease.

Pooling data from the AURAext and AURA2 studies

The company pooled the data from the AURAext and AURA2 studies to produce a single dataset that was subsequently used to calculate summary efficacy and safety end-points. The company states that using pooled data increases the precision of outcome estimates.

Pooling the data from the two AURA studies was carried out by merging the individual patient data (IPD). Study identifiers were anonymised and variables which were calculated/computed were derived identically across the studies. The company considers that the two AURA studies are largely comparable, having similar inclusion/exclusion criteria as well as similar baseline demographic and disease characteristics.

During the clarification process, the ERG asked the company to undertake analyses to show whether the efficacy results from the pooled AURA dataset differ from those generated using the AURAext and AURA2 study data independently. The company repeated the analyses for the two primary efficacy endpoints (PFS and ORR) comparing the data for the AURAext and AURA2 studies separately with PDC. The PFS and ORR analyses for the individual AURAext and AURA2 studies were conducted versus PDC using the same methodology as for the pooled analysis and the results indicate that the pooled results are consistent with the results achieved when the AURAext and AURA2 studies are compared separately versus PDC.

The ERG considers that it is reasonable to pool data from the AURAext and AURA2 studies, and the approach taken to do so was appropriate.

ERG assessment of statistical approach

A summary of the checks made by the ERG regarding the statistical approach adopted by the company to analyse data from the AURAext study, the AURA2 study and the IMPRESS trial is provided in Table 6.

Table 6 ERG assessment of statistical approaches used to analyse data from the AURAext and AURA2 studies and the IMPRESS trial

Component	AURAext and AURA2 studies		IMPRESS trial	
	Statistical approach	ERG comments	Statistical approach	ERG comments
Sample size calculation	Provided in the CS (p96-97)	The ERG considers that the methods used to calculate the sample size are correct	Provided in the CS (p135)	The ERG considers that the methods used to calculate the sample size are correct
Protocol amendments	Provided in the CSR (Section 5.8.1)	The ERG notes that the changes detailed in the protocol amendments are unlikely to have been driven by the results of the trial and are, therefore, not a cause for concern. All protocol amendments were carried out prior to the analyses being conducted	Provided in the CSR (Section 5.8.1)	The ERG notes that the changes detailed in the protocol amendments are unlikely to have been driven by the results of the trial and are therefore not a cause for concern. All protocol amendments were carried out prior to any analyses being conducted
Missing data approach	No details provided	It is not possible for the ERG to review these results as they have not been provided	No details provided	It is not possible for the ERG to review these results as they have not been provided
Subgroup analyses	For ORR and DoR • Patients who received EGFR-TKI or those whose treatment prior to study start was not an EGFR-TKI • Ethnicity (Asian or Non-Asian) • Gender (Male or Female) • Age at screening (<65 or ≥65) • Mutation status prior to start of study (Exon 19 deletion or L858R or Other) • Duration of most recent prior EGFR-TKI (<6 months or ≥6 months) • Smoking history (never or ever) • Brain metastases at entry (yes or no) • Patients with EGFR T790M	It is not possible for the ERG to review these results as the pooled results used to perform the subgroup analyses are not provided in the CSR	For PFS: • Region (Asia or European Union) • Time from progression to randomisation (≤2 weeks or >2 weeks) • Smoking history (never or current/former) • Prior response to gefitinib (SD or PR and CR combined) • Exon 19 deletion (present or absent/unknown) • L858R mutation (present or absent/unknown) • Age (<65 years or ≥65 years) • Gender (male or female) • Disease stage at diagnosis	The ERG is satisfied that the results of all subgroup analyses are provided in the CSR

	mutation positive or patients that are EGFR T790M negative • Region (North America or Asia or Europe or rest of world)		(1=locally advanced or 0=metastatic, 'other') • Time to progression for initial gefitinib (≤10 months or >10 months) • Site of disease at baseline (brain/CNS or non-brain/CNS) • WHO performance status (0=normal activity or 1=restricted activity)	
Adverse events	Safety was assessed through summaries of most common AEs, SAEs and patients who had at least one adverse event	It is not possible for the ERG to review these results as results generated from the pooled dataset are not provided in the CSRs	Safety was assessed through summaries of most common AEs, SAEs, AEs leading to treatment discontinuation, AEs of CTCAE grade 3 or higher, dose interruptions and dose reductions	It is not possible for the ERG to review these results as they are not provided in the CSR
Health-related quality of life	• EORTC-QLQ-C30 • EORTC QLQ-LC13 • EQ-5D-5L questionnaire	The ERG is satisfied that the methodology used to analyse HRQoL data is appropriate	• FACT-L • EQ-5D questionnaire	The ERG is satisfied that the methodology used to analyse HRQoL data is appropriate

AE=Adverse Event; CR=complete response; CS=company submission; CSR=clinical study report; DoR=duration of response; EGFR-TKI=epidermal growth factor receptor tyrosine kinase inhibitor; EORTC-QLQ-C30=European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire version 3; EQ-5D=EuroQol-5 Dimensions; ERG=Evidence Review Group; FACT-L=Functional Assessment of Cancer Therapy—Lung; HRQoL=health- related quality of life; ORR=overall response rate; OS=overall survival; PFS=progression-free survival; PR=partial response; SAE=Serious Adverse Event; SD=stable disease; WHO=World Health Organisation
Source: CS, CSRs and ERG comment

4.2.4 Key study characteristics of the included studies

The company reports (CS, p87) that the AURAext study is a phase II extension study and is part of an overarching phase I/II, open label, dose-escalation, expansion and extension cohort study programme, known as AURA. The company describes the AURA2 study as a phase II, open-label study.

There are no published papers that describe the AURAext and AURA2 studies. However, as well as information presented in the CS, the company has provided the CSRs for both studies.

The key characteristics of the AURAext and AURA2 studies are listed in Table 7. The studies are being conducted internationally and include a combined total of 411 patients with EGFR T790M mutation-positive NSCLC who have progressed following previous treatment. All patients are treated with 80mg of osimertinib once daily. Previous treatments include, but are not limited to, an EGFR-TKI. The ERG notes that the AURAext study inclusion criteria stipulate previous treatment with an EGFR-TKI and other anti-cancer treatments, whilst the AURA2 study criteria stipulate previous treatment with an EGFR-TKI and PDC (other previous lines of treatment are also permitted). The ERG notes that the AURAext study was open to recruitment at two centres in the UK; however, it is not clear how many patients have been recruited from the UK. The AURA2 study was not open to recruitment from UK centres.

Table 7 Key characteristics of the AURAext and AURA2 studies

	AURAext (osimertinib 80mg)	AURA2 (osimertinib 80mg)
Location	International (including 2 UK centres)	International (no UK centres)
Design	Phase II extension, open label, single arm	Phase II, open label, single arm
Population	N=201 patients with T790M mutation-positive EGFR NSCLC Two patient cohorts: 1. Patients with disease progression following first-line therapy with an EGFR-TKI 2. Patients with disease progression following treatment with an EGFR-TKI and other anti-cancer treatments	N=210 patients with T790M mutation-positive EGFR NSCLC Two patient cohorts: 1. Patients with disease progression following first-line therapy with an EGFR-TKI 2. Patients with disease progression following treatment with an EGFR-TKI and a platinum-based doublet (possibly other lines of treatment also)
Intervention	Osimertinib (80mg) until disease progression or cessation of clinical benefit	Osimertinib (80mg) until disease progression or cessation of clinical benefit
Primary outcome	ORR	ORR
Secondary outcomes	Duration of response, disease control rate, tumour shrinkage, PFS, OS, safety, HRQoL	Duration of response, disease control rate, tumour shrinkage, PFS,OS, safety, HRQoL
Duration of study	The first patient started treatment on 14th May 2014 and the last patient started treatment on 21st October 2014. The data cut-off for the present appraisal was 1st May 2015	The first patient started treatment on 13th June 2014 and the last patient started treatment on 27th October 2014. The data cut-off for the present appraisal was 1st May 2015

EGFR=epidermal growth factor receptor; HRQoL=health related quality of life; ORR=overall response rate; OS=overall survival; PFS=progression-free survival; TKI-tyrosine kinase inhibitor
Source: CS, p87 to p98

The key baseline characteristics of patients included in the AURAext and AURA2 studies are listed in Table 8. The ERG notes that the baseline characteristics of the patients in the two studies are similar.

Table 8 Key characteristics of patients in the AURAext and AURA2 studies

	AURAext (osimertinib 80mg) N=201	AURA2 (osimertinib 80mg) N=210	Total N=411
Mean age (sd)	61.4 (10.58)	62.9 (10.91)	62.2 (10.76)
Age group n (%)			
<50 years	30 (14.9)	20 (9.5)	50 (12.2)
≥50 to <65 years	86 (42.8)	88 (41.9)	174 (42.3)
≥65 to <75 years	64 (31.8)	69 (32.9)	133 (32.4)
≥75 years	21 (10.4)	33 (15.7)	54 (13.1)
WHO PS n (%)			
0	68 (33.8)	84 (40)	152 (37)
1	132 (65.7)	126 (60)	258 (62.8)
2	1 (0.5)	0	1 (0.2)
Female n (%)	133 (66.2)	146 (69.5)	279 (67.9)
Race n (%)			
White	76 (38.2)	72 (34.3)	148 (36.2)
Asian	114 (57.3)	132 (62.9)	246 (60.1)
Other/not reported	9 (4.5)	6 (2.9)	15 (3.7)
Number of prior anti-cancer treatments n (%)			
1	61 (30.3)	69 (32.9)	130 (31.6)
2	49 (24.4)	45 (21.4)	94 (22.9)
3	33 (16.4)	38 (18.1)	71 (17.3)
4	22 (10.9)	22 (10.5)	44 (10.7)
5	14 (7.0)	7 (3.3)	21 (5.1)
>5	22 (10.9)	29 (13.8)	51 (12.4)
Mean (sd)	2.8 (1.92)	3.0 (2.43)	2.9 (2.20)
Min and max	1 and 11	1 and 14	1 and 14
Number of prior EGFR-TKI treatments n (%)			
1	111 (55.2)	131 (62.4)	242 (58.9)
2	47 (23.4)	42 (20)	89 (21.7)
3	33 (16.4)	18 (8.6)	51 (12.4)
4	7 (3.5)	9 (4.3)	16 (3.9)
5	2 (1.0)	4 (1.9)	6 (1.5)
>5	1 (0.5)	6 (2.9)	7 (1.7)
Mean (sd)	1.7 (0.98)	1.8 (1.34)	1.7 (1.18)
Min and max	1 and 6	1 and 9	1 and 9

	AURAext (osimertinib 80mg) N=201	AURA2 (osimertinib 80mg) N=210	Total N=411
Histology type n (%)			
Adenocarcinoma: NOS	171 (85.1)	170 (81)	341 (83)
Adenocarcinoma: acinar	11 (5.5)	10 (4.8)	21 (5.1)
Adenocarcinoma: papillary	10 (5.0)	17 (8.1)	27 (6.6)
Adenocarcinoma: bronchiolo-alveolar	3 (1.5)	1 (0.5)	4 (1.0)
Adenocarcinoma: solid with mucous formation	0	2 (1.0)	2 (0.5)
Adenosquamous carcinoma	1 (0.5)	1 (0.5)	2 (0.5)
Squamous cell carcinoma	0	2 (1.0)	2 (0.5)
Other	5 (2.5)	7 (3.3)	12 (2.9)
EGFR mutation type n (%)			
T790M	197 (98)	208 (99)	405 (98.5)
Exon 19 deletion	142 (70.6)	137 (65.2)	279 (67.9)
L858R	51 (25.4)	67 (31.9)	118 (28.7)
G719X	4 (2.0)	4 (1.9)	8 (1.9)
S768I	3 (1.5)	3 (1.4)	6 (1.5)
Exon 20 insertion	2 (1.0)	1 (0.5)	3 (0.7)
EGFR T790M only	5 (2.5)	1 (0.5)	6 (1.5)
Overall disease classification n (%)			
Metastatic	197 (98)	198 (94.3)	395 (96.1)
Locally advanced	4 (2.0)	12 (5.7)	16 (3.9)
Brain metastases n (%)	74 (36.8)	87 (41.4)	161 (39.2)
Visceral metastases n (%)	173 (86.1)	168 (80)	341 (83)
Baseline sum of target lesions mean (sd)	60.7 (37.08)	59.7 (40.57)	60.2 (38.82)
Baseline sum of target lesions tumour size			
<40mm	65 (32.3)	68 (32.4)	133 (32.4)
40 to 79mm	86 (42.8)	90 (42.9)	176 (42.8)
80 to 119mm	31 (15.4)	26 (12.4)	57 (13.9)
≥120mm	17 (8.5)	15 (7.1)	32 (7.8)

EGFR=epidermal growth factor receptor; PS=performance status; sd=standard deviation; TKI=tyrosine kinase inhibitor; WHO=World Health Organisation
Source: CS, Table 4.17 to Table 4.20

Figures in Table 8 show that 32% of patients in the AURAext and AURA2 studies received osimertinib as a second-line treatment after an EGFR-TKI. This means that the majority of patients (68%) received osimertinib as a third-line (or greater) treatment.

The ERG notes that, in comparison to patients recruited to the AURAext and AURA2 studies, patients with EGFRm+ NSCLC currently treated in the NHS are older and are less fit. Clinical advice to the ERG is that, typically, patients in this population treated in the NHS are aged between 65 and 70 years and the majority have an ECOG PS of 1 or 2. Patients recruited to the two AURA studies have a mean age of 62 years and an ECOG PS of 0 or 1.

Clinical advice to the ERG is that the ethnic case-mix of patients treated in the NHS is the reverse of that in the AURAext and AURA2 studies. In these two studies, 36.2% of patients were described as white and 60% were described as Asian.

In addition, figures in Table 8 show that, despite patients in the AURAext and AURA2 studies having experienced a substantial number of anti-cancer treatments prior to study entry, they were considered fit enough for treatment with osimertinib (ECOG PS 0 or 1).

The ERG further notes that there are patients included in the AURAext and AURA2 studies who have received multiple EGFR-TKI treatments (up to six in the AURAext study and up to nine in the AURA2 study). Clinical advice to the ERG is that in NHS clinical practice, patients are typically treated with only one EGFR-TKI (although a second may be offered in the case of toxicity, or if the patient is not considered to be fit enough to receive chemotherapy).

The vast majority (96%) of patients included in the AURAext and AURA2 studies have tumours that are of adenocarcinoma histology and have metastatic disease (96%). This disease profile is consistent with EGFRm+ patients treated in the NHS.

4.2.5 Assessment of risk of bias for the AURA studies

The company conducted a risk of bias assessment for the AURAext and AURA2 studies using the Downs and Black checklist[32] (Table 9). The Downs and Black checklist[32] is listed in the NICE methods guide[50] as being appropriate for use when assessing cohort studies. The results of the company assessment of the AURA studies, with accompanying ERG comments, are shown in Table 9.

In general, the ERG agrees with the company assessment, but differs in responses to Q24 and Q25 (allocation concealment and adjustment for confounders). The company omitted to include Q27 of the checklist and the ERG has added it to Table 9. The ERG considers that the AURAext and AURA2 studies were designed, conducted and reported to a good standard. The ERG notes that the blinded independent review of the radiological outcomes in the AURAext and AURA2 studies lends weight to the efficacy results. However, the ERG highlights that the AURAext and AURA2 studies are non-randomised, single-arm studies without a control group. As a consequence, the results of the studies cannot be considered as reliable or robust as the results of a RCT (outcomes could, for example, be the result of chance, patient characteristics or the Hawthorne effect). In addition, the OS data available from the AURAext and AURA2 studies are very immature (pooled OS dataset is 12.7% mature).

Table 9 Results of company quality assessment of the AURAext and AURA2 studies with ERG comments

Downs and Black checklist item	AURAext	AURA2	ERG comment
Reporting			
Q1: Aim clearly described	Yes	Yes	Agree
Q2: Outcomes clearly described	Yes	Yes	Agree
Q3: Patients characteristics clearly described	Yes	Yes	Agree
Q4: Interventions clearly described	Yes	Yes	Agree
Q5: Principal confounders clearly described	Yes	Yes	Agree
Q6: Main findings clearly described	Yes	Yes	Agree
Q7: Random variability for the main outcome provided	Yes	Yes	Agree
Q8: Adverse events reported	Yes	Yes	Agree
Q9: Lost to follow up reported	Yes	Yes	Agree
Q10: Actual p-value reported	No	No	Agree
External validity and bias			
Q11: Sample asked to participate representative of the population	Yes	Yes	Partially agree
Q12: Sample agreed to participate representative of the population	Yes	Yes	Agree
Q13: Staff participating representative of the patient's environment	Yes	Yes	Agree
Q14: Attempt to blind participants	No	No	Agree
Q15: Attempt to blind assessors	Yes	Yes	Agree
Q16: Data dredging results stated clearly	Yes	Yes	Agree
Q17: Analysis adjusted for length of follow up	Yes	Yes	Agree
Q18: Appropriate statistics	Yes	Yes	Agree
Q19: Reliable compliance	Yes	Yes	Agree
Q20: Accurate outcome measures	Yes	Yes	Agree
Statistical bias and power			
Q21: Same population	Yes	Yes	Agree
Q22: Participants recruited at the same time	Yes	Yes	Agree
Q23: Randomised?	No	No	Agree
Q24: Adequate allocation concealment?	UTD	UTD	No. The studies were not randomised
Q25: Adequate adjustment for confounders?	UTD	UTD	Disagree. Subgroup analyses were conducted to assess key factors that may impact on outcomes
Q26: Loss of follow up reported?	Yes	Yes	Agree
Q27: Did the study have sufficient power to detect a clinically important event? (score between 1 and 5)	Not addressed in CS	Not addressed in CS	5. Sample sizes were calculated and presented in the CS

UTD=unable to determine; ERG=Evidence Review Group
Source: CS, Table 4.21

4.3 Results from the AURAext and AURA2 studies

Results reported in the CS for both the AURAext and AURA2 studies use data from the 1st May 2015 data-cut. The company states (CS, p118) that median follow-up for PFS by BICR is 6.9 months in the AURAext study and 6.7 months in the AURA2 study. Median follow-up for OS is 8.3 months in the AURAext study and 7 months in the AURA2 study (CS, p119). The ERG notes that 83 patients continued osimertinib treatment for at least 7 days after progression, the median duration of treatment with osimertinib treatment after progression was 1.6 months (range 0.4 to 8.4).

The company presents individual study data from the AURAext and AURA2 studies and also the results from the analyses of pooled AURAext and AURA2 study data. The focus of the CS is on results from the pooled AURA dataset. The company claims (CS, p112) that the AURAext and AURA2 studies are comparable in terms of patient populations, design and outcome measures. The ERG agrees with the company that the AURA studies are comparable and that it is reasonable to combine the data from the two studies.

4.3.1 Objective response rate (primary outcome)

The BICR assessment of the pooled 'evaluable for response' dataset (Table 10) yielded an ORR of 66.1% (95% CI: 61.2 to 70.7). The results of two sensitivity analyses (64.2% and 70.6 %) were similar to the BICR result.

Table 10 Summary of overall response rate

Analysis set Study	N	Number of patients with confirmed response	ORR (%)	95% CI
BICR assessment of 'evaluable for response' analysis set				
AURAext	199	122	61.3	54.2 to 68.1
AURA2	199	141	70.9	64.0 to 77.1
Total	398	263	66.1	61.2 to 70.7
BICR assessment of FAS (sensitivity analysis)				
AURAext	201	122	60.7	53.6 to 67.5
AURA2	210	142	67.6	60.8 to 73.9
Total	411	264	64.2	59.4 to 68.9
Investigator assessment of FAS (sensitivity analysis)				
AURAext	201	142	70.6	63.8 to 76.8
AURA2	210	148	70.5	63.8 to 76.6
Total	411	290	70.6	65.9 to 74.9

BICR=blinded independent central review; CI=confidence interval; FAS=full analysis set: ORR=overall response rate
Source: CS, Table 4.22

Subgroup analyses

The ORR data were analysed across patient subgroups. The subgroups included line of treatment, Asian or non-Asian, male or female, <65 years or ≥65 years, EGFR mutation

status (M+ or M-), duration of prior EGFR-TKI treatment (<6 months or ≥6 months), brain metastases (yes or no), smoking status (ever or never), last treatment prior to enrolment (<30 days or ≥30 days or not EGFR), EGFR T790M detected in plasma sample (positive or negative), and region of origin (North America or Asia or Europe and Rest of World).

Figure 4.8 in the CS (p114) illustrates that ORRs by BICR range from 58.9% (patients with L858R mutations, n=112) to 72.2% (patients for whom the EGFR T790M mutation was not detected via a plasma sample, n=162). The company highlights (CS, p113) that the ORR for second-line patients (66.9%) is very similar to the ORR for ≥third-line patients (65.7%).

During the clarification process, the ERG requested the p-values for the tests for interaction for the performed subgroup analyses. Statistically significant subgroup differences were observed for ethnicity [redacted]), mutation status prior to start of study [redacted]), and EGFR T790M status in baseline plasma sample [redacted]). These results (from company clarification response to question A8) suggest that the treatment effect is statistically significantly greater for Asian patients than for non-Asian patients, for patients with Exon 19 deletion mutation present than for patients with L858R mutation present, and for patients with an EGFR T790M mutation that is detected in blood plasma than for patients in whom the mutation is not detected in blood plasma.

The company reports other measures of patient response to treatment:

- **Best objective response by BICR in the 'evaluable for response' population'**: In the pooled population, two patients (0.5%) had a complete response (CR) to treatment whilst 261 patients (65.6%) had a partial response (PR). The findings were similar for PR in the AURAext and AURA2 studies (61.3% and 69.8% respectively)
- **Duration of response by BICR**: Median DoR had not been reached (maturity of 22.8%). A K-M analysis using pooled data from patients who had responded to treatment estimated that 94.9% of patients would achieve a response lasting at least 3 months and that 55.3% of patients would achieve a response lasting at least 9 months. Duration of response by investigator assessment (maturity of 27.6%) was 8.5 months (95% CI: 8.5 to not calculable)
- **Disease control rate by BICR in the FAS:** In the pooled population, the DCR was 91%. Similar DCRs were recorded for patients in the AURAext and AURA2 studies (90.5% and 91.5% respectively)
- **Tumour shrinkage by BICR in the 'evaluable for response' population**: Mean percentage change from baseline in target lesion size in the pooled dataset was 45.01% (standard deviation [SD] 28.01). Similar tumour shrinkage rates were reported in both the AURAext and AURA2 studies (-41.09% [SD 24.71] and -48.94% [SD 30.54] respectively). The company reports that evidence of tumour shrinkage was generally noted at the first follow up scan at 6 weeks.

4.3.2 Progression-free survival and overall survival

The company reports that at the time of the data-cut, all patients had been followed up for at least 6 months. Median PFS, calculated using the pooled AURA dataset, was 9.7 months (95% CI: 8.3 to not calculable) (Table 11). Median PFS was not calculable from the AURAext study data and, using AURA2 study data, was 8.6 months (95% CI: 8.3 to 9.7).

The company reports that the OS data are immature (12.7% for the pooled AURA dataset) and median OS has not yet been reached in either the AURAext study or in the AURA2 study.

Table 11 Summary of progression-free survival and overall survival (FAS)

	AURAext (osimertinib 80mg) (n=201)	AURA2 (osimertinib 80mg) (n=210)	Total (osimertinib 80mg) (n=411)
Progression-free survival by BICR			
Total number of events	80	79	159
Median PFS months (95% CI)	NC (8.1 to NC)	8.6 (8.3 to 9.7)	9.7 (8.3 to NC)
Median follow-up (months)	6.9	6.7	6.8
% Progression-free at 3 months (95% CI)	81.5 (75.3 to 86.2)	84.9 (79.2 to 89.1)	83.2 (79.2 to 86.5)
% Progression-free at 6 months (95% CI)	72.0 (65.1 to 77.8)	69.7 (62.8 to 75.7)	70.9 (66.1 to 75.1)
% Progression-free at 9 months (95% CI)	54.6 (46.4 to 62.1)	47.7 (36.2 to 58.4)	51.9 (45.3 to 58.1)
Overall survival			
Total number of deaths	28	24	52
Median OS	NC	NC	NC
Survival at 3 months % (95% CI)	96.5 (92.80 to 98.32)	97.1 (93.72 to 98.70)	96.8 (94.59 to 98.14)
Survival at 6 months % (95% CI)	93.0 (88.41 to 95.77)	91.7.0 (86.97 to 94.76)	92.3 (89.27 to 94.54)
Survival at 9 months % (95% CI)	84.0 (77.49 to 88.74)	87.1 (80.83 to 91.49)	85.3 (80.85 to 88.71)
Patients in survival follow-up n (%)	168 (83.6)	181 (86.2)	349 (84.9)
Median follow-up (months)	8.3	7.0	7.4

CI=confidence interval; NC=not calculable; OS=overall survival; PFS=progression-free survival; BICR=blinded independent committee review; FAS=full analysis set
Source: CS, Table 4.27 and Table 4.28

4.4 Health related quality of life

The HRQoL data presented in the CS (p120 to p124) were collected using the European Organisation for Research and Treatment of Cancer Quality of Life questionnaires, the EORTC-QLQ-30[46] and the lung cancer specific questionnaire EORTC-QLQ-L13.[47]

Assessment points were at baseline and at each clinic visit up to week 42. During the first 6 months, the company reports a >90% completion rate in the AURAext study and a >70% completion rate in the AURA2 study.

The methods of collecting data to complete the two EORTC questionnaires differed between the AURAext and AURA2 studies in that paper-based and electronic hand-held devices were used respectively, and the company provides this as the reason why they did not pool the collected HRQoL data (CS, p120).

The EORTC-QLQ-C30[46] questionnaire comprises a measure of global health status, five functional dimensions (physical, social, role, emotional, and cognitive functioning) and nine symptom domains (fatigue, nausea and vomiting, pain, dyspnoea, insomnia, appetite loss, constipation, diarrhoea and financial difficulties). The EORTC-QLQ-LC13[47] questionnaire is a measure of symptoms that are specific to lung cancer and includes 13 items. The company highlights (CS, p120) that no major differences were noted in patient reported outcomes between second-line patients and ≥third-line patients.

As part of the AURA2 study, data were also collected using the EQ-5D-5L questionnaire and the EQ-VAS.[50] The ERG notes that the AURA2 population baseline index score (a population with advanced or metastatic NSCLC) is higher than the UK population norm for the 55-64 years of age group. However, this difference in index score is difficult to interpret as a) there is no UK value set for the EQ-5D-5L tool and b) the UK population norms were estimated using the EQ-5D-3L tool.

AURAext

Responses to the EORTC-QLQ-C30 indicated:

- a consistent positive responses on symptomatic domains and quality of life domains up to week 42
- a clinically significant improvement in overall global health status from week 12 to week 30 (44% and 48% of patients)
- a clinically meaningful increase in diarrhoea reported at week 6 (37% of patients) and week 30 (26% of patients)
- that, at 6 months, 62% of patients had not reported any deterioration in dyspnoea, cough or pain.

Responses to the EORTC-QLQ-LC13 indicated:

- a clinically meaningful improvement from baseline, starting at week 6 and at each time point, was reported for dyspnoea (35% to 45% of patients), cough (31% to 39% of patients), chest pain (28% to 33% of patients), pain in arm or shoulder (17% to 28% of patients) and pain in other parts of the body (36% to 39% of patients)

- a clinically meaningful worsening in sore mouth was reported, starting at week 12 and for the follow-up time points, by 19% to 27% of patients. The remaining patients reported stability or improvement in sore mouth.

AURA2

Responses to the EORTC-QLQ-C30 indicated there is a:

- consistent positive response on symptomatic domains and quality of life domains up to week 30
- clinically meaningful improvement in social functioning for the first 24 weeks (38% to 44% of patients)
- clinically meaningful improvement in appetite loss at weeks 6, 18 and 24 (29% to 34% of patients)
- clinically meaningful improvement in insomnia at weeks 12 and 18 (34% to 35% of patients)
- clinically meaningful improvement in fatigue at week 18 (49.6% of patients)
- clinically meaningful worsening in diarrhoea reported at weeks 6 (30% of patients).

Responses to the EORTC-QLQ-LC13 indicated that there is a:

- clinically meaningful improvement from baseline, starting from week 2/4 to week 36, was reported for dyspnoea (25% to 39% of patients), cough (31% to 40% of patients) and chest pain (20% to 31% of patients). From week 6 to week 36, 23% to 29% of patients reported improvements in arm or shoulder pain, but this improvement was not clinically meaningful.

Responses to the EQ-5D-5L and VAS indicated that:

- from week 12 onwards, patients treated with osimertinib experienced a clinically significant improvement from baseline. The minimal important difference (MID) for cancer was ≥7.5 on VAS and 0.1 on the Health Utilities Index[51] [HUI] score).

The ERG cautions that as the HRQoL data are reported separately for each study, the results are based on relatively small numbers of respondents. At baseline, the number of respondents who completed the EQ-5D-5L questionnaire and the EQ-VAS was 175; by week 36, only 30 patients completed the questionnaires. Further details relating to HRQoL measures are reported in the CS (Tables 4.29 and 4.30).

4.4.1 Key study characteristics of the IMPRESS trial

The IMPRESS trial is a double-blind, placebo-controlled, multi-centre RCT. Patients (n=265) with EGFRm+ NSCLC who had progressed on treatment with gefitinib were randomised in a 1:1 ratio to receive either gefitinib+pemetrexed+cisplatin or placebo+pemetrexed+cisplatin (PDC).

As noted previously in this report, the efficacy of osimertinib was compared with efficacy data from the subset of patients in the control arm (PDC) of the IMPRESS trial who were

identified (retrospectively) as having EGFR T790M mutation-positive disease. The company had access to tumour samples from 98% of patients in the IMPRESS trial and was, therefore, able to undertake this retrospective identification. The company reports (CS, p147) that 54% of patients from the IMPRESS trial tested positive for the EGFR T790M mutation and that this prevalence rate is consistent with reported prevalence rates in other studies. The ERG agrees that the estimated prevalence of EGFR T790M mutation-positive NSCLC is likely to be approximately 50% to 60% in patients who have progressed on or after treatment with a first-line EGFR-TKI.[4,5,52] In the control (PDC) arm of the IMPRESS trial, 61 patients (46.2%) were identified as having EGFR T790M mutation-positive disease.

The early overall population results from the IMPRESS trial are available in a published peer-reviewed paper[13] and, in addition to information presented in the CS, the company has provided the CSR for the IMPRESS trial. The company expects that the final OS results from the IMPRESS trial will be published in [redacted] (company clarification response). The key characteristics of the IMPRESS trial are provided in Table 12.

Table 12 Key characteristics of the IMPRESS trial

Characteristic	IMPRESS trial
Location	Europe and Asia-Pacific region (no UK centres)
Design	Phase III, double- blind, placebo-controlled RCT, 61 centres in 11 countries
Population	265 patients with locally advanced or metastatic EGFRm+ NSCLC. Patients had progressed after first-line treatment with gefitinib 4 months minimum duration of treatment with gefitinib with a response lasting at least 4 months or stable disease for at least 6 months
Intervention	Gefitinib (250mg daily), cisplatin (75mg per m^2) and pemetrexed (500mg per m^2) for up to six cycles. After six cycles, patients continued on gefitinib until progression
Comparator	Placebo (once daily), cisplatin+pemetrexed (as per intervention). After six cycles patients continued on placebo until progression
Primary outcome	PFS (investigator-assessed)
Secondary outcomes	OS, ORR, disease control rate, HRQoL, safety
Duration of study	The first patient was randomised on 29th March 2012 and the last patient was randomised on 20th December 2013. The data cut-off for the present appraisal was 5th May 2014.

EGFRm+=epidermal growth factor receptor mutation-positive; HRQoL=health related quality of life; ORR=overall response rate; OS=overall survival; PFS=progression-free survival
Source: CS, p128 to p130

The key baseline characteristics of patients recruited to the IMPRESS trial, including the subgroup of patients with EGFR T790M mutation-positive disease who were randomised to the control (PDC) arm of the trial, are shown in Table 13.

Overall, the patient characteristics are well balanced between the two treatment arms. However, the ERG notes that the patients in the intervention arm are older than patients in the control arm (59.3 years versus 57 years). The subgroup of patients (n=61) in the

IMPRESS trial who were randomised to the control arm and (later) identified as having EGFR T790M mutation-positive disease are younger than the overall trial population (55.8 years versus 58.1 years). In all other respects, however, the patients with EGFR T790M mutation-positive disease appear to have similar characteristics to the whole IMPRESS trial population.

Clinical advice to the ERG is that patients in the population of interest treated in clinical practice in the NHS are older and less fit than the patients recruited to the IMPRESS trial. As noted earlier in Section 4.2.4 of this report, clinical advice to the ERG is that patients in the population of interest who are treated in the NHS are typically aged between 65 years and 70 years and the majority have an ECOG PS of 1 or 2. The overall patient population in the IMPRESS trial has a mean age of 58.1 years and an ECOG PS of 0 or 1. Clinical advice to the ERG is that the case-mix of Asian and white patients treated in the NHS is the reverse of the case-mix reported in the IMPRESS trial. In the IMPRESS trial 21% of patients are described as white and 77.7% are described as Asian.

Table 13 Key baseline characteristics of patients participating in the IMPRESS trial

	Gefitinib+cisplatin +pemetrexed	PDC	PDC EGFR T790M mutation-positive patients
Number of patients	133	132	61
Mean age (sd)	59.3 (10.63)	57 (11.25)	55.8 (10.20)
Age group n (%)			
<65 years	90 (67.7)	98 (74.2)	51 (83.6)
>65 years	43 (32.3)	34 (25.8)	10 (16.4)
WHO PS n (%)			
0	55 (41.4)	53 (40.2)	22 (36.1)
1	78 (58.6)	79 (59.8)	36 (63.9)
Disease stage at baseline n (%)			
Metastatic disease	124 (93.2)	119 (90.2)	58 (95.1)
Locally advanced disease	7 (5.7)	7 (9.3)	3 (4.9)
Female n (%)	87 (65.4)	84 (63.6)	38 (62.3)
Race n (%)			
White	29 (21.8)	29 (22)	12 (19.7)
Asian	104 (78.2)	102 (77.3)	48 (78.7)
Black or African American	0	1 (0.8)	1 (1.6)
Time to progression for initial gefitinib treatment n (%)			
≤10 months	52 (39.1)	58 (43.9)	NR
>10 months	81 (60.9)	74 (56.1)	NR
Never smoked n (%)	88 (66.2)	91 (68.9)	NR
Adenocarcinoma histology n (%)	126 (94.8)	131 (99.2)	59 (96.7)
Brain metastases at baseline n (%)	44 (31)	31 (23.5)	NR
Exon 19 deletion n (%)	85 (63.9)	86 (65.2)	NR
L858R n (%)	40 (30.1)	42 (31.8)	NR

PDC=platinum doublet chemotherapy, specifically placebo+pemetrexed+cisplatin; NR=not reported; PS=performance status; sd=standard deviation; TKI=tyrosine kinase inhibitor; WHO=World Health Organisation; NOS=not otherwise stated
Source: CS, Table 4.34 and Table 4.35

4.4.2 Assessment of risk of bias for the IMPRESS trial

The company conducted a risk of bias assessment for the IMPRESS trial using the minimum criteria recommended in the NICE methods guide.[50] The company has rated the overall quality of the trial using the Jadad[33] score (maximum score of 5) and has rated the allocation concealment aspect of the IMPRESS trial using a grading system where A means adequate and D means no allocation concealment was attempted.

Overall, the ERG agrees with the company assessment of risk of bias (Table 14) and considers the IMPRESS trial to be of good quality.

Table 14 Company assessment of risk of bias for the IMPRESS trial with ERG comments

Assessment criteria	Company assessment	ERG comment
JADAD score	4	The ERG considers that the IMPRESS trial warrants the maximum score of 5
Allocation concealment	A	Agree. Central block randomisation using interactive voice response system would prevent knowledge of treatment allocation
Was randomisation carried out appropriately?	Low risk. Patients were assigned to treatment arms via central block randomisation in a 1:1 ratio using interactive web response system or interactive voice response system during the first visit (initial screening)	Agree
Were the groups similar at the outset of the study in terms of prognostic factors?	Low risk. The baseline characteristics between the two treatment arms were well balanced	Agree. However, patients randomised to the control arm were slightly younger than those randomised to the intervention arm
Were the care providers, participants and outcome assessors blind to treatment allocation?	Low risk. This was a double-blind study. All study investigators and participants were masked to treatment allocation. To ensure masking of study investigators and participants, all gefitinib and placebo packaging was identical. Apart from safety reasons, nobody was allowed access to the randomisation scheme or study results until completion of the randomised treatment period to minimise any potential bias in data handling and to safeguard the integrity of the masking of study investigators	Agree
Were there any unexpected imbalances in dropouts between groups?	Low risk. Study withdrawals were adequately reported and incorporated in the patient flow diagram	Agree
Is there any evidence to suggest that the authors measured more outcomes than they reported?	Low risk. The authors measured all outcomes as reported in the protocol (NCT01544179)	Agree
Did the analysis include an ITT analysis? Was this appropriate and were appropriate methods used to account for missing data?	Low risk. The safety and efficacy analysis was performed using mITT and ITT populations respectively	Agree

ITT=intention to treat; mITT=modified intention to treat;
Source: CS, Table 4.36

4.4.3 Results from the IMPRESS trial

The company presents the results from the IMPRESS trial based on the data that were collected up until 5th May 2015 (CS, p148 to p156). Median follow-up for PFS was 11.2 months. Both results for the overall trial population and those for the subgroup of patients in the control arm with EGFR T790M mutation-positive disease are presented in this section.

4.4.4 Progression-free survival (investigator-assessed) and overall survival

Progression-free survival (primary outcome)

At the time of the analysis, the PFS data in the FAS population were 77.4% mature. Figures in Table 15 show no statistically significant differences between the PFS results for the intervention and control arms of the trial (PFS=5.4 months in both arms). The PFS result for the control arm EGFR T790M mutation-positive patient subgroup (PFS=5.3 months) was similar to the PFS results for the overall trial population.

The company presents a K-M plot of the PFS data (CS, Figure 4.16) for the overall trial population and states: i) that the curves for the treatment arms do not cross and ii) the treatment effects of the intervention and the control therapies were consistent over time. The ERG agrees with the company that the PH assumption appears valid.

The results of the company subgroup analyses for the overall population of the IMPRESS trial are presented in Figure 4.17 of the CS. Subgroup analyses were conducted for: age (<65 years or ≥65 years), male or female, region of origin (Asia or Europe), previous response to gefitinib (CR+PR or stable disease), EGFR mutation subtype (Exon 19 deletion or L838R deletion), smoking status (present or former or never), disease state at diagnosis (metastatic or non-metastatic), time from progression to randomisation (>2 weeks or ≤2 weeks), time to progression for initial gefitinib therapy ≤10 months or >10 months), brain metastases at baseline (yes or no) and WHO PS (0 or 1). Significant interactions were noted for:

- Asia (HR=0.80) versus Europe (HR=0.95)
- Never smokers (HR=0.70) versus current or former smokers (HR=1.16)
- Exon 19 deletion present (HR=0.76) versus Exon 19 absent/unknown Exon 19 (HR=0.97)
- WHO PS 0 (HR=0.68) versus WHO PS 1 (HR=0.95).

During the clarification process, the ERG requested the corresponding p-values for the tests for interaction for these analyses. However, the company only sent results for the tests of interaction for the control arm of the IMPRESS trial. This means that the ERG is unable to

assess which of the subgroup analyses are significant and therefore cannot comment on the company's conclusions.

The company also presents a PFS K-M data plot by treatment arm and biomarker status (CS, Figure 4.18). The company states that a (non-significant) treatment effect of gefitinib was recorded in patients without EGFR T790M mutation-positive disease (HR=0.67; 95% CI: 0.43 to 1.03, p=0.0745). In contrast, no treatment effect of gefitinib was recorded for patients with EGFR T790M mutation-positive disease (HR=0.97; 95% CI: 0.67 to 1.42, p=0.8829). The company interprets this finding as providing support to the biological hypothesis that, in the absence of the EGFR T790M mutation, the tumour may still respond to treatment with an EGFR-TKI.

Overall survival

At the time of the analysis, FAS population OS data from the IMPRESS trial were immature. The company reports that follow-up for survival is ongoing and that more mature data will be available (following a [redacted]

The analysis of OS data (Table 15) from the overall IMPRESS trial population demonstrates a statistically significant treatment effect for patients in the control arm (HR=1.62; 95% CI: 1.05 to 2.52, p=0.029). Median OS is 17.2 months (95% CI 15.6 to NR) in the control arm compared with 14.8 months (95% CI: 10.4 to 19.0) in the intervention arm. OS for the subgroup of patients in the control arm with EGFR T790M mutation-positive disease was 15.7 months, which is lower than the OS for the whole control arm population (17.2 months) but higher than that for the intervention arm population (14.8 months). The ERG cautions that the OS results are based on immature data.

The company reports that 45.9% of patients in the intervention arm and 54.5% of patients in the control arm had received further anti-cancer treatment after discontinuation of their study treatment (CS, p153).

Table 15 Progression-free survival and overall survival from the IMPRESS trial

	Gefitinib+cisplatin +pemetrexed	PDC	PDC EGFR T790M mutation-positive patients
Number of patients	133	132	61
Progression-free survival (investigator-assessed)			
Total number of events	98	107	51
Median PFS months (95% CI)	5.4 (4.5 to 5.7)	5.4 (4.5 to 5.5)	5.3
Median follow-up (months)	11.2	11.2	NR
% Progression-free at 4 months (95% CI)	73.5 (64.8 to 80.4)	67.8 (59 to 75.2)	NR
% Progression-free at 6 months (95% CI)	40.8 (31.5 to 49.8)	36 (27.4 to 44.6)	NR
% Progression-free at 8 months (95% CI)	28.2 (19.8 to 37.1)	17.3 (10.9 to 25.1)	NR
Overall survival			
Total number of deaths	50	37	20
Median OS months (95% CI)	14.8 (10.4 to 19.0)	17.2 (15.6 to NR)	15.7
Hazard ratio (95% CI)	1.62 (1.05 to 2.52), p-value=0.029		NR

CI=confidence interval; NR=not reached; OS=overall survival; PFS=progression-free survival: PDC=platinum doublet chemotherapy, specifically placebo+pemetrexed+cisplatin
Source: CS: Table 4.37, Table 4.38 and p153

Objective response rate

The ORR reported in the CS is based on the percentage of patients with a BOR of CR or PR (according to RECIST criteria) using investigator-assessed data. The proportion of patients with a response is similar in the intervention (31.6%) and control (34.1%) arms of the trial for the overall populations (OR=0.92; 95% CI: 0.55 to 1.55, p-value=0.760). Similarly, just over a third (39.3%) of patients in the control arm with EGFR T790M mutation-positive disease were assessed as responding.

Health related quality of life

HRQol in the IMPRESS trial was measured using the Functional Assessment of Cancer Therapy – Lung (FACT-L[49]) questionnaire and the EQ-VAS. The company states (CS, p155) that the results from the FACT-L questionnaire: i) are not relevant to the decision problem and ii) are not used in the comparison with the pooled AURA dataset. The company has, therefore, not included a detailed discussion of the FACT-L data collected during the IMPRESS trial. The ERG agrees with the company's decision.

The EQ-VAS scores for a subset of the FAS population are summarised in Table 4.40 of the CS (p156). The company states that the patient response rate for completion and evaluability rate of the questionnaires was similar in the intervention and control arms of the

trial. The ERG highlights that the EQ-VAS scores are higher than those collected as part of the AURA2 study.

4.5 Adverse events from the AURAext study, the AURA2 study and the IMPRESS trial

Adverse event data from the AURAext study, the AURA2 study, the pooled AURA dataset and the IMPRESS trial are reported in the CS (p157 to p161). A comparison of the pooled AURA dataset and in the IMPRESS trial AE data is presented and discussed as part of the adjusted comparison (CS, p78 and p79).

4.5.1 Pooled AURA dataset

The company reports (CS, p157) that median treatment duration in the AURAext study was 8.2 months and 7.4 months in the AURA2 study.

The ERG notes (Table 16) that, in the pooled AURA dataset, the majority (97.6%) of patients treated with osimertinib experienced an AE and almost one third (29.4%) of patients experienced an AE of grade 3 or higher. One fifth (20.2%) of patients had a serious adverse event (SAE) and one fifth of patients (19.7%) had their dose of osimertinib reduced due to AEs. Seventeen patients (4.1%) discontinued their treatment due to AEs and nine patients (2.2%) died as a result of an AE. It is reported in the CSRs for the AURAext and AURA2 studies that grade 3 or above AEs included respiratory disorders (13%), infections (6%), investigations (5.8%) and blood disorders (5% in AURA2). It is also reported that more SAEs were experienced by patients receiving osimertinib as a third-line (or greater) treatment than were experienced by patients receiving osimertinib as a second-line treatment.

The company reports that the most commonly reported AEs were consistent across the studies and are consistent with the AEs known to be associated with EGFR-TKI treatment.

Table 16 Categories of AEs from the pooled AURA dataset, FAS

Category of adverse event	AURAext N=201, n (%)	AURA2 N=210, n (%)	Total N=411, n (%)
Patients with any AE	198 (98.5)	203 (96.7)	401 (97.6)
AE ≥grade 3	60 (29.9)	61 (29.0)	121 (29.4)
SAEs	41 (20.4)	42 (20.0)	83 (20.2)
Fatal SAEs	4 (2)	5 (2.4)	9 (2.2)
AEs leading to discontinuation	9 (4.5)	8 (3.8)	17 (4.1)
AEs leading to dose modification	40 (20.4)	41 (19.5)	81 (19.7)

AE=adverse event; SAE=serious adverse event, FAS=full analysis dataset
Source: CS, Table 4.41

The information in Table 17 shows that, in the pooled AURA dataset, diarrhoea and rashes and acne were the most frequently reported AEs at any grade (42.3% and 41.4%

respectively). The company states (CS, p158) that the incidences of decreased appetite, fatigue and nausea were 'mostly mild in nature and non-serious.' Clinical advice to the ERG is that, in NHS clinical practice, incidences of diarrhoea and fatigue can be difficult to manage, particularly in an elderly population.

Table 17 AEs from the pooled AURA dataset (≥10% of patients), FAS

	AURAext N=201, n (%)		AURA2 N=210, n (%)		Total N=411, n (%)	
Adverse event	Any grade	≥grade 3	Any grade	≥grade 3	Any grade	≥grade 3
Diarrhoea	93 (46.3)	2 (1.0)	81 (38.6)	2 (1.0)	174 (42.3)	4 (1.0)
Rashes+acnes	81 (40.3)	1 (0.5)	87 (41.4)	1 (0.5)	170 (41.4)	2 (0.5)
Dry skin	43 (21.4)	0	52 (24.8)	0	95 (23.1)	0
Paronychia	40 (19.9)	0	32 (15.2)	0	72 (17.5)	0
Nausea	35 (17.4)	2 (1.0)	34 (16.2)	0	69 (16.8)	2 (0.5)
Decreased appetite	36 (17.9)	2 (1.0)	29 (13.8)	1 (0.5)	65 (15.8)	3 (0.7)
Constipation	30 (14.9)	1 (0.5)	32 (15.2)	1 (0.5)	62 (15.1)	1 (0.2)
Cough	32 (15.9)	0	25 (11.9)	1 (0.5)	57 (13.9)	1 (0.2)
Fatigue	25 (12.4)	2 (1.0)	32 (15.2)	0	57 (13.9)	2 (0.5)
Pruritus	25 (12.4)	0	32 (15.2)	0	57 (13.9)	0
Back pain	27 (13.4)	1 (0.5)	25 (11.9)	2 (1.0)	52 (12.7)	3 (0.7)
Stomatitis	27 (13.4)	0	22 (10.5)	0	49 (11.9)	0
Platelet count decreased	27 (13.4)	1 (0.5)	20 (9.5)	1 (0.5)	47 (11.4)	2 (0.5)
Headache	22 (10.9)	0	20 (9.5)	1 (0.5)	42 (10.2)	1 (0.2)

FAS=full analysis dataset
Source: CS, Table 4.42

4.5.2 IMPRESS trial

The company reports (CSR, p108) that the mean total treatment duration in the intervention and control arms of the IMPRESS trial was 165 days and 155 days respectively. The data in Table 18 show that, across all categories of AEs, patients in each of the trial arms experienced similar rates of AEs.

Table 18 Categories of common AEs (≥10%) from the IMPRESS trial, safety analysis set

Category of AE	Gefitinib+cisplatin +pemetrexed (N=133)	PDC (N=132)
Patients with any AE	126 (95.5)	130 (98.5)
AE ≥grade 3	59 (44.7)	55 (41.7)
Fatal AEs	5 (3.8)	8 (6.1)
AEs leading to discontinuation	10 (7.6)	13 (9.8)
AEs leading to dose modification	6 (4.5)	7 (5.3)
SAEs	NR (28)	NR (21.2)
SAE leading to discontinuation	(3)	(8)

AE=adverse event; PDC=platinum doublet chemotherapy, specifically placebo+cisplatin+pemetrexed; SAE=serious adverse event
Source: CS, p159 and published paper (Soria 2015)

The data in Table 19 list the most commonly reported AEs experienced by patients in the IMPRESS trial. The three most frequent AEs in the intervention and control arms of the trial were nausea (64.4% and 61.4%), decreased appetite (49.2% and 34.1%) and vomiting (41.7% and 33.3%).

The company reports (CS, p161) that decreased neutrophil count, anaemia, neutropenia and decreased white blood cell count were the most commonly experienced (≥5%) AEs of grade 3 or higher.

Table 19 AEs from the IMPRESS trial (≥10% of patients), safety analysis set

Adverse event	Gefitinib+cisplatin +pemetrexed (N=132), n (%)	PDC (N=132), n (%)
Any AE	126 (95.5)	130 (98.5)
Nausea	85 (64.4)	81 (61.4)
Decreased appetite	65 (49.2)	45 (34.1)
Vomiting	55 (41.7)	44 (33.3)
Anaemia	42 (31.8)	33 (25.0)
Constipation	34 (25.8)	35 (26.5)
Diarrhoea	44 (33.3)	19 (14.4)
Neutropenia	29 (22.0)	28 (21.2)
Fatigue	28 (21.2)	23 (17.4)
Leucopenia	27 (20.5)	22 (16.7)
Asthenia	15 (11.4)	30 (22.7)
Neutrophil count decreased	16 (12.1)	22 (16.7)
Pyrexia	22 (16.7)	14 (10.6)
Cough	18 (13.6)	15 (11.4)
White blood cell count decreased	17 (12.9)	13 (9.8)
Headache	10 (7.6)	19 (14.4)
Dyspnoea	16 (12.1)	10 (7.6)
Back pain	11 (8.3)	14 (10.6)
Rash	14 (10.6)	11 (8.3)
Stomatitis	14 (10.6)	5 (3.8)

AE=adverse event; PDC=platinum doublet chemotherapy, specifically placebo+cisplatin+pemetrexed
Source: CS, Table 4.43

4.6 Critique of trials included in the unadjusted and adjusted comparisons

4.6.1 Methodological approach to the unadjusted and adjusted comparisons

The company employed two methods to compare the clinical effectiveness results from the pooled AURA dataset with those from the subgroup of patients in the control arm of the IMPRESS trial with EGFR T790M mutation-positive disease. The two methods were a simple unadjusted comparison and an adjusted comparison. The adjusted comparison involved adjustments to control for differences in baseline characteristics between the populations in the two datasets.

Unadjusted comparison

The unadjusted comparison simply involved comparing key efficacy outcomes (ORR, PFS and OS) from the two datasets.

Adjusted comparison

The methods employed by the company to carry out the adjustments and estimate outcomes are summarised in Box 3.

Box 3 Summary of adjustment approach and outcome calculation methods

Patients from the AURAext and AURA2 studies were matched with patients from the control arm of the IMPRESS trial with EGFR T790M mutation-positive disease. Matching was based on baseline demographic and disease characteristics. Patients for whom there was no match were dropped from the analysis at this point.

Prior to estimating efficacy outcomes, differences between baseline demographic and disease characteristics were accounted for using a three-step process:

a. selection of baseline variables that were statistically significantly different between groups (based on a p-value of <0.2)
b. generation of a propensity score to represent aggregated differences in variables selected and trimming of the data set by removal of patients for which there was no similar propensity score in the alternative group and
c. incorporation of propensity score as covariate in analysis of treatment effect of osimertinib for each endpoint to adjust for remaining differences between the two groups.

The estimated propensity score was defined as the conditional probability that the distribution of observed baseline covariates will be similar between treated and untreated patients, i.e. patients are equally likely to be treated with osimertinib or PDC in the absence of baseline differences. It acts as a proxy for randomisation.

The baseline demographic and disease characteristics that were used in the regression

model to estimate the propensity scores and the final trimmed dataset were age, ethnicity, baseline target lesion size and smoking history. The final datasets included in the adjusted analysis comprised 287 patients from the AURA studies and 51 patients from the control arm of the IMPRESS trial. The trimmed dataset is referred to as the T790M+ adjusted dataset.

For the matched populations, the treatment effect of osimertinib versus PDC was assessed for key efficacy and safety endpoints as follows:

- PFS: Cox PH model with treatment as a factor and propensity score as a covariate
- OS: based on independent assessment review and performed at the time of the PFS analysis using a Cox PH model
- ORR and DCR: carried out using logistic regression with treatment as a factor and propensity score as a covariate.

PH=proportional hazards; ORR=objective response rate; PDC=platinum doublet chemotherapy; PFS=progression-free survival; OS=overall survival; DCR=disease control rate
Source: CS, Section 4.10.3

ERG critique of the company's adjusted comparison

The ERG appreciates the lengths taken by the company to facilitate a comparison of the effectiveness of osimertinib with PDC, but considers that only a well-controlled, head-to-head RCT can avoid unobserved confounding.

The ERG notes that the adjustments that have been made only relate to age, ethnicity, baseline target lesion size and smoking history, all of which (except for age) are well balanced between the two datasets. No adjustments have been made to account for differences in either line of treatment (including number of previous EGFR-TKIs) or brain metastases; clinical advice to the ERG suggests that these may be important prognostic factors.

In addition, when testing for statistically significant differences in baseline variables, the company used a p-value of <0.2 instead of a conventional significance level of 0.05; the rationale behind this choice is not explained. Furthermore, the company has not provided details of all of the baseline summary variables that were tested for possible inclusion; the ERG has, therefore, been unable to check whether there are any further uncontrolled differences between the trials.

2. Over and above concerns about the immaturity of the OS data, the ERG notes that of the datasets was substantially reduced following adjustments (

Table 21). The pooled AURA dataset was reduced from n=411 to n=287 for the PFS and OS analyses, and to n=277 for the ORR and DCR analyses. The T790M+ adjusted dataset was reduced from n=61 to n=51 for the PFS and OS analyses, and to n=46 for the ORR and DCR analyses.

Characteristics of studies included in the clinical efficacy comparisons

The company provides baseline characteristics related to patients in the pooled AURA dataset (n=411) and to the subgroup of patients in the control arm of the IMPRESS trial with EGFR T790M mutation-positive disease (n=61). The baseline characteristics of the patients in the individual datasets used in the adjusted comparisons are unknown.

The baseline characteristics of the patients included in the pooled AURA dataset and in the subgroup of patients in the control arm of the IMPRESS trial with EGFR T790M mutation-positive disease are shown in Table 4.7 of the CS. The eligibility criteria used to recruit patients to the two AURA studies and to the IMPRESS trial differ slightly. The pooled AURA dataset includes both patients receiving second-line and patients receiving subsequent lines of treatment, whilst the IMPRESS trial only recruited patients who had received one prior EGFR-TKI therapy. Furthermore, whilst patients in the AURAext and AURA2 studies were not required to have had a prior treatment response to an EGFR-TKI, patients in the IMPRESS trial had to have had a prior objective clinical benefit (as measured by CR or PR) and a minimum duration on first-line gefitinib treatment of 4 months. Another key difference in eligibility criteria is that the two populations used different methods to identify EGFR T790M mutation status. The Roche Cobas method was used in the AURAext and AURA2 studies and the BEAMing digital PCR method was used (retrospectively) to identify patients with T790M mutation-positive disease recruited to the IMPRESS trial.

Despite the differences in inclusion criteria, overall, the baseline patient demographic characteristics of patients included in the pooled AURA dataset and the subgroup of patients in the control arm of the IMPRESS trial with EGFR T790M mutation-positive disease are well balanced. The key differences between datasets are in terms of age and presence of brain metastases. The subgroup of patients in the control arm of the IMPRESS trial with EGFR T790M mutation-positive disease is slightly younger than patients in the pooled AURA dataset, with a mean age of 55.8 years compared to 62.2 years. Only 16.4% of patients in the subgroup of patients in the control arm of the IMPRESS trial with EGFR T790M mutation positive disease were ≥65 years, whereas in the pooled AURA dataset population 45.5% of patients were ≥65 years old. Furthermore, compared with patients in the pooled AURA dataset, fewer patients in the control arm of the IMPRESS trial who had EGFR T790M mutation-positive disease had brain metastases at baseline (40.4% versus 34.4%

respectively). The company considers that both the age and the brain metastases imbalances may have a prognostic effect favouring the subgroup of patients in the control arm of the IMPRESS trial with EGFR T790M mutation-positive disease.

Key differences between the pooled AURA dataset population and the subgroup of patients in the control arm of the IMPRESS trial with EGFR T790M mutation-positive disease are summarised in Table 20.

Table 20 Key baseline differences between the AURA studies and the IMPRESS subgroup

Demographic characteristics		Pooled AURA dataset	EGFR T790M mutation-positive population
Number		411	61
Indication		≥Second-line	Second-line
Treatment		Osimertinib 80mg	Placebo+PDC
Age (years)	Mean (SD)	62.2 (10.76)	55.8 (10.20)
	Median (min-max)	63 (35-89)	55 (38-79)
	% ≥65 years	187 (45.5%)	10 (16.4%)
Brain metastatic at baseline		166 (40.4%)	21 (34.4%)

EGFR=epidermal growth factor receptor; SD=standard deviation; PDC=platinum doublet chemotherapy
Source: CS, Table 4.7

4.6.2 Assessment of risk of bias of the trials included in the unadjusted and adjusted comparisons

The company conducted an assessment of the risk of bias for the AURAext and AURA2 studies, which is discussed in Section 4.2.5, and for the IMPRESS trial, which is discussed in Section 4.4.2, of this report. The ERG considers that the assessments of risk of bias conducted by the company were appropriate. The ERG considers that the AURA studies were designed and conducted to a good standard (with the caveat that both were single-arm studies) and that the IMPRESS trial was of very good quality.

4.6.3 Results from the unadjusted and adjusted comparisons

Unadjusted comparison of study results

The results of the unadjusted comparison are provided in

Table 21.

The ORR observed within the pooled AURA dataset (ORR=66.1%) was significantly higher than the ORR observed in the IMPRESS EGFR T790M mutation-positive control group (ORR=39.3%). Furthermore, patients in the pooled AURA dataset had a median PFS of 9.7 months compared to 5.3 months for the IMPRESS EGFR T790M mutation-positive control group, indicating a statistically significant difference of 4.4 months. However, the results need to be interpreted with caution as, at the time of analysis, the 95% CIs for PFS were either not calculable or not reached as the data were very immature with only 12.7% of patients having had OS events in the pooled AURA dataset and only 32.8% of patients in the EGFR T790M mutation-positive control group of the IMPRESS trial having had OS events. The median OS was not reached for patients in the pooled AURA dataset. In the control arm of the IMPRESS trial, median OS was 15.7 months for patients with EGFR T790M mutation-positive disease; the company did not report 95% CIs.

Adjusted comparison of study results

The results from the adjusted comparison for ORR, PFS and OS are provided in

Table 21.

The ORR results indicate a statistically significant improvement in favour of osimertinib compared to PDC (64.6% and 34.8% respectively, OR=4.76; 95% CI: 2.21 to 10.26; p<0.001). Similarly, the DCR results indicate a statistically significant improvement in favour of osimertinib compared to PDC (92.1% and 76.1% respectively, OR=4.39; 95% CI: 1.71 to 11.28; p=0.002). The PFS results indicate a statistically significant difference in favour of osimertinib compared to PDC (HR=0.280; 95% CI: 0.185 to 0.422; p<0.0001). Median PFS is 9.7 months for the osimertinib cohort compared to 5.2 months in the matched PDC cohort. Due to the very small number of patients experiencing events (osimertinib, n=33; PDC, n=15) median OS could not be calculated (HR=1.022; 95% CI 0.387 to 2.696; p=0.9654).

The ERG investigated whether the PH assumption employed by the company to calculate PFS and OS HRs hold by digitising the data presented in Figure 4.2 (PFS) and Figure 4.3 (OS) of the CS and then plotting the cumulative hazard associated with osimertinib treatment versus the cumulative hazard associated with PDC treatment (H-H plot). The PFS H-H plot suggests that the PH assumption does not hold for PFS and, therefore, the PFS HR result must be interpreted with caution. Interpretation of the OS H-H plot is less clear and the issue is complicated by the lack of data. However, based on the data available, it would not be unreasonable to assume that hazards are broadly proportional.

The ERG notes that key efficacy results from the adjusted and unadjusted analyses are very similar (

Table 21).

Table 21 Comparison of key efficacy outcomes (unadjusted and adjusted)

Study / Outcome		Unadjusted data sets			Adjusted data sets	
		Pooled AURA dataset	PDC (IMPRESS trial)		Pooled AURA dataset	IMPRESS trial T790M+adjusted dataset
			Whole population	EGFR T790M+ subgroup		
Number		411	132	61	PFS/OS: 287 ORR: 277	PFS/OS: 51 ORR: 46
ORR	Total responses, n (%)	264 (66.1)	45 (34.1)	24 (39.3)	179 (64.6)	16 (34.8)
PFS	Total events, n (%)	159 (38.9)	107 (81.1)	51 (83.6)	106 (36.9)	42 (82.4)
	Median, months (95% CI)	9.7 (8.9 to NC)	5.4 (4.6 to 5.5)	5.3 (NR)	9.7 (8.3 to NC)	5.3 (4.0 to 6.1)
OS	Total events, n (%)	52 (12.7)	37 (28)	20 (32.8)	33 (11.5)	15 (29.4)
	Median, months (95% CI)	NR	17.2 (15.6 to NC)	15.7 (NR)	NC (NC to NC)	21.7 (12.55 to NC)

CI=confidence interval; NC=not calculable; NR=not reported; ORR=overall response rate; OS=overall survival; PDC=platinum doublet chemotherapy, specifically placebo+pemetrexed+cisplatin; PFS=progression-free survival
Source: CS, Table 4.8, Table 4.11, Table 4.12 and Table 4.14

Conclusions of the clinical effectiveness section

The phase I/II study evidence presented in the CS in support of the clinical effectiveness of osimertinib for treating EGFR T790M mutation-positive NSCLC suggests that osimertinib may be a promising treatment for this population.

Direct evidence - key issues and uncertainties

The AURAext and AURA2 studies were designed as single-arm studies. This raises challenges in interpreting study results. The lack of results from a comparator arm means that how much of the reported effects of osimertinib are the result of treatment, the natural course of the disease or a placebo effect is unclear. The lack of a comparator arm also means that no direct comparison of the clinical effectiveness of osimertinib with any of the comparators listed in the final scope issued by NICE is available.

A particular difficulty in this appraisal is the lack of mature survival data. The OS in the pooled AURA dataset has reached only 12.7% maturity; this clearly precludes any reliable assessment of the OS benefit of treatment with osimertinib.

The ERG questions the generalisability of the results from the AURAext and AURA2 studies to the population of interest treated in the NHS. The patients recruited to the studies were younger and fitter (ECOG PS 0 or 1) than similar patients seen in NHS clinical practice. The

majority (two-thirds) of recruited patients received osimertinib as a third- fourth- or fifth-line treatment following a first-line EGFR-TKI and a first-line chemotherapy. The ERG is aware that very few patients seen in clinical practice in the NHS are well enough to tolerate more than one or two chemotherapy treatments. Patients from only two UK centres contributed to the data in the pooled AURA dataset.

The company has pooled IPD data from the AURAext and AURA2 datasets and generated efficacy results from this dataset. The company explains that the rationale behind this approach was to improve the precision of outcomes. The ERG considers that, as these two studies are very similar in terms of recruitment criteria and patient baseline characteristics, it was reasonable to pool the data. Furthermore, results generated independently using data from the two studies are similar to results generated from the pooled dataset.

Unadjusted and adjusted comparisons

The company should be commended for the effort that they have taken to formulate a comparator dataset. The comparator dataset comprises patients recruited to the control (PDC) arm of the IMPRESS trial who were (retrospectively) identified as having EGFR T790M mutation-positive disease. The ERG, however, has concerns that data from single-arm, non-controlled studies (AURAext and AURA2) are compared with data from a retrospectively identified subgroup participating in a good quality placebo-controlled, double-blind RCT (IMPRESS). Furthermore, this IMPRESS subgroup only includes 61 patients and OS data are only 32.8% mature.

The ERG commends the company for attempting to control for differences in baseline dataset differences by carrying out an adjusted comparison. However, [redacted]

The results from the unadjusted comparison indicate that osimertinib is more clinically effective, as measured by PFS and ORR than treatment with PDC (median OS has not yet been reached). The safety data suggest that treatment with osimertinib is more tolerable than treatment with PDC.

Other key issues and uncertainties

The evidence presented in the CS compares the clinical effectiveness of osimertinib with PDC. No evidence is available to compare osimertinib with any of the other 11 comparators specified in the final scope issued by NICE.

The mutation testing protocol required for the use of osimertinib is not in place in the NHS. T790M mutation testing after first-line treatment to establish the presence or absence of the EGFR T790M mutation is feasible as the infrastructure is in place; however, EGFR T790M mutation testing after first-line treatment is not standard practice in the NHS.

5 COST EFFECTIVENESS

This section provides a structured critique of the economic evidence submitted by the company to support the cost effectiveness of osimertinib. The company model focuses on patients with locally advanced or metastatic EGFR T790M mutation-positive NSCLC who have progressed on or after EGFR-TKI therapy.

The two key components of the economic evidence presented in the CS are (i) a systematic review of the relevant literature and (ii) a report of the company's de novo economic evaluation. The company also provided an electronic version of their economic model that was developed in Microsoft Excel.

5.1 Objective of the company systematic review

The company's search was conducted to identify evidence to support the development of the company's cost effectiveness model. The company carried out a single review to identify studies that included descriptions of economic evaluations as well as information on resource use and costs. Initially, the review focussed on identifying evidence relating to patients with EGFR-TKI mutation-positive disease and/or T790M mutations with acquired resistance to an EGFR-TKI. However, due to a lack of available evidence, the remit of the review was broadened to include patients:

a) harbouring EGFR and/or T790M mutations following any prior therapy
b) with unknown EGFR and T790M mutation status following treatment failure with an EGFR-TKI.

Details of the search strategies employed by the company are provided in Appendix A2.1 of the CS. The data sources for the economic systematic review are outlined in Table 22. The searches were conducted in January 2016.

5.1.1 Eligibility criteria used in study selection

Table 22 Data sources for the economic systematic review

Search strategy component	Sources	Date limits
Electronic database searches Key biomedical electronic literature databases recommended by HTA agencies	MEDLINE® MEDLINE® In-process Excerpta Medical Database (Embase®) Cochrane® Central Register of Controlled Trials (CENTRAL) Cochrane National Health Service Economic Evaluation Database (NHS EED) EconLit®	01 Jan 2004 to 21 Jan 2016
Conference proceedings	HTA International	2012–2016
	International Society for Pharmacoeconomics and Outcomes Research (ISPOR)	
	European Society for Medical Oncology (ESMO)	
	American Society of Clinical Oncology (ASCO)	

Source: CS, Table 5.1

5.1.2 Inclusion criteria

The inclusion/exclusion criteria used by the company to facilitate study selection are presented in Table 23. The company used different study designs to identify economic evaluations (S1) and quality of life studies (S2).

Table 23 Inclusion/exclusion criteria for the economic review

	Economic evaluations	Rationale
Patient population (P)	Age: adults aged ≥18 years Gender: any Race: any Disease: patients with advanced/metastatic NSCLC who are EGFR and/or T790M mutant and who have failed at least one EGFR-TKI ±other anticancer regimens	The patient population of interest to the review comprised adult patients with advanced/metastatic NSCLC of any race and gender because NSCLC can occur at any age but is most common in adults aged between 40 years and 70 years. Therefore, studies focusing solely on children and adolescents were not included in this review
Intervention (I)	Osimertinib	This is the intervention of interest within the decision problem
Comparator (C)	Any pharmacological intervention Placebo Best supportive care	The searches for economic review were not restricted to any interventions in order to collate all available published economic evidence in patients with advanced/metastatic NSCLC harbouring EGFR/T790M mutations following prior therapy
Outcome (O)	Studies were not be excluded based on the reported outcomes	The aim of the review was to identify relevant economic evaluations that also reported costs

	Economic evaluations	Rationale
Study design 1 (S1)	All economic evaluation studies based on models Cost-effectiveness analysis Cost-utility analysis Cost-minimisation analysis Cost-benefit analyses Budget impact models Resource use studies Cost/economic burden of illness	The aim of the review was to identify relevant economic evaluations that also reported costs
Study design 2 (S2)	Randomised controlled trials Database studies Prospective observational studies Retrospective observational studies	The aim of the review was to identify relevant studies that reported quality of life data
Line of therapy	Second- or further-line of therapy	This is the relevant line of treatment
Search timeframe	2004 to 2016	This period was deemed relevant to reflect models that are representative of the current NSCLC landscape
Language	Only studies with the full-text published in English language were included	It is expected that the majority of evidence in this disease area will be available in the English language
Exclusion criteria	Reviews, letter to the editors, and editorials Case studies/case series Case reports Cross-sectional studies	The design of such studies was not relevant to the decision problem These are generally smaller studies with higher risk of bias, hence excluded
	Studies investigating the role of radiotherapy, chemo-radiotherapy, hormonal therapy, or surgery only were excluded Studies investigating the role of maintenance/consolidation therapy after surgery were also excluded Adjuvant or neo-adjuvant therapy were excluded No subgroup analysis	Only pharmacological interventions (chemotherapies and targeted therapies) were considered as relevant comparators for osimertinib Studies that included children and adults and did not provide subgroup analysis for the adult populations Studies which enrol a mixed population of stage I, II, IIIa, and stage IIIb/IV NSCLC and did not provide subgroup analysis for the disease stage IIIb/IV

Source: CS, Table 5.2

5.1.3 Included and excluded studies

The company identified five studies[51-55] for inclusion in a qualitative synthesis. None of the studies included osimertinib as an intervention or as a comparator. None of the studies included patients with locally advanced or metastatic EGFR T790M mutation-positive NSCLC who had progressed on or after EGFR-TKI therapy.

The company reported that it did find one conference abstract describing a study that included the primary population. However, no reference to this conference proceeding was listed in the CS.

5.1.4 Findings from cost effectiveness review

The company provided a summary of the five included studies.[51-55] However, the company did not comment on the results from any of these studies.

5.2 ERG critique of the company's literature review

The ERG is satisfied with the company's search strategies and is confident that there are no studies that fully meet the company's inclusion criteria. The databases searched and search terms used appear to be reasonable. The ERG considers the wider search for published economic literature (e.g. inclusion of a broad population of patients) to be appropriate when taking into account the shortage of relevant clinical and economic data for the specific patient population of interest to this appraisal.

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

5.4 ERG's summary of company's submitted economic evaluation

The company base case cost effectiveness analysis compares osimertinib with platinum doublet chemotherapy (PDC – specifically, cisplatin+pemetrexed), and adopts a lifetime horizon of 15 years. In the CS (p181), the company states that the economic evaluation is carried out from the perspective of the NHS and Personal Social Services (PSS) and includes the resource use and costs associated with treatment acquisition, treatment administration, disease management, AEs and EGFR T790M mutation testing. In the model, the cycle length is 1-week to facilitate comparison with most chemotherapy regimens and a half-cycle correction is employed. In line with the current NICE Reference Case,[50] costs and quality adjusted life years (QALYs) are discounted at an annual rate of 3.5%.

The company also carried out scenario analyses and subgroup analyses to explore the cost effectiveness of osimertinib versus PDC, and versus single-agent chemotherapy, in different patient populations.

5.4.1 Model structure

The company developed a de novo cohort based survival model that comprised three health states: progression-free (PF), progressed disease (PD) and death. The partitioned survival model is similar to that of other treatments for advanced cancers that have been submitted to NICE as part of the STA process. In the model, OS = PF + PD. The structure of the company model is shown in Figure 3.

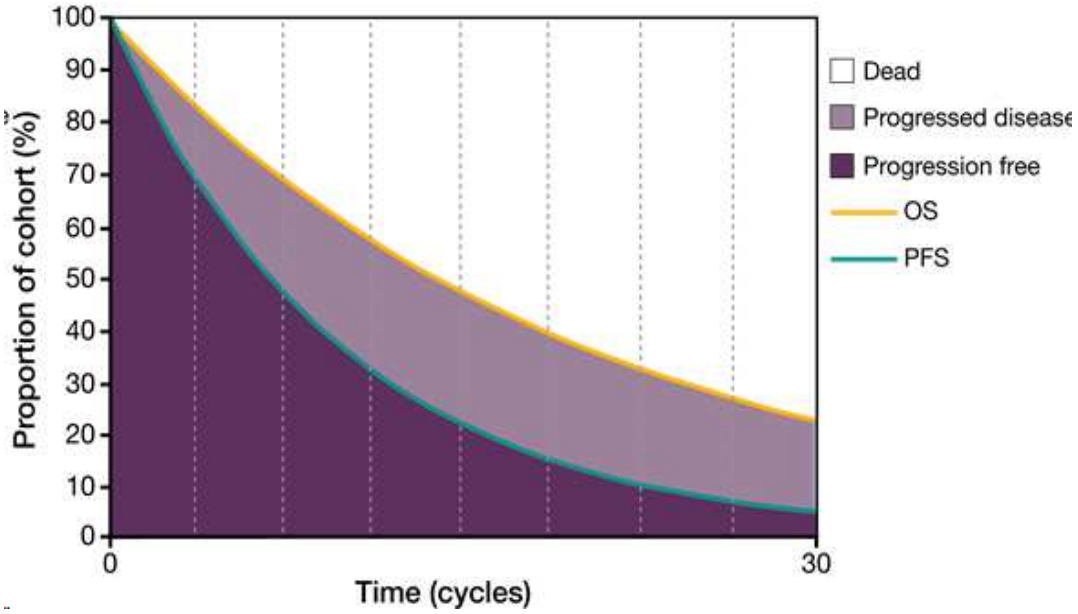


Figure 3 Partitioned survival analysis model structure

OS=overall survival; PFS= progression-free survival
Source: CS, Figure 5.2

As described in Section 5.5.2 of the CS, the company assumes that the three health states represent the key sequence of events that patients may experience over the course of their treatment, with the additional assumption that these events are progressive, mutually exclusive and irreversible.

5.4.2 Population

The economic evaluation considers patients with locally advanced or metastatic EGFR T790M mutation-positive NSCLC who have progressed on or after EGFR-TKI therapy, i.e. the model is relevant to patients requiring second-line or further-line treatment. The age of patients starting treatment in the model is 62.17 years; this is based on the average age of patients in the AURAext and AURA 2 studies. The body surface area (BSA) of patients in the model is assumed to be 1.68m^2.

5.4.3 Interventions and comparators

Osimertinib is implemented in the model in line with the anticipated licensed dose, i.e. one 80mg tablet to be taken once per day. The base case comparator is PDC (i.e., pemetrexed+cisplatin); this treatment is administered every 3 weeks by intravenous infusion at a dose of 500mg/m^2 over 10 minutes for pemetrexed, and at a dose of 75mg/m^2 over 2 hours for cisplatin. Pemetrexed+cisplatin is the current standard of care in the NHS for patients with EGFR mutation-positive disease who have failed first-line treatment with an EGFR-TKI.

In a scenario analysis, osimertinib is compared with up to six cycles of docetaxel monotherapy; this treatment is administered every 3 weeks by intravenous infusion at a dose of 75mg/m^2 over 60 minutes. Up to four cycles of docetaxel is the current standard of care in the NHS for patients with non-squamous lung cancer who do not have EGFR mutation-positive disease and who have failed first-line treatment with pemetrexed+cisplatin.

5.4.4 Perspective, time horizon and discounting

The company states that the economic evaluation is undertaken from the perspective of the NHS/PSS. In the model the maximum lifetime is set at 15 years, this is slightly shorter than other similar models[56] that have been recently submitted to the NICE STA process. Both costs and benefits are discounted at a rate of 3.5% per annum.

5.4.5 Treatment effectiveness and extrapolation in the base case

Disease progression and overall survival model inputs (osimertinib versus PDC)

Data from the AURAext (n=201) and AURA2 (n=210) studies were pooled and used to demonstrate the PFS and OS associated with treatment with osimertinib. At the time of data cut-off, the PFS data were 39% mature and the OS data were 12.7% mature. PDC clinical data were derived from patients (n=60) with T970M mutation-positive disease in the control arm of the IMPRESS trial. The company recognised that, due to the immaturity of the data, it was difficult to test proportional hazards assumptions. Therefore, the analysis used independent survival models for osimertinib and comparator treatments.

The company followed standard guidance for fitting and selecting survival functions. A full step-wise description of the statistical analysis undertaken by the company, which was based on NICE DSU guidance[57] is presented in the CS (Section 5.3.5 to Section 5.3.9, Figure 5.3). The company investigated the use of a range of parametric models: Gompertz, Generalised Gamma, Log-normal, Log-logistic, Exponential and Weibull. In accordance with the DSU guidance[57], the company selected the same parametric models for both treatment arms.

Based on visual inspection only, the Gompertz, Weibull and Generalised Gamma distributions appeared to provide the most adequate estimates of PFS for PDC. For osimertinib, the Weibull model appeared to provide the best fit to the pooled AURA dataset.

Based on statistical goodness-of-fit tests (Akaike Information Criterion [AIC] and Bayesian Information Criterion [BIC]), the Weibull distribution had the best fit for osimertinib and the second best fit for PDC, whilst the Gompertz distribution had the best fit for PDC and the

second best fit for osimertinib. The company concluded that the goodness-of-fit for OS was not conclusive and differed between studies.

In summary, the Gompertz distribution was selected for PFS as it had the best visual fit for both osimertinib and PDC. The Weibull distribution was selected for OS as it appeared to produce the most reasonable fit to the non-parametric OS data that are currently available from the AURAext and AURA2 studies and from the IMPRESS trial. OS and PFS survival curves used in the base case are shown in Figure 4.

The median PFS and OS for osimertinib and PDC are shown in Figure 5. The data show that, compared with PDC, treatment with osimertinib results in an incremental PFS gain of 4.8 months and an incremental OS gain of 10.6 months.

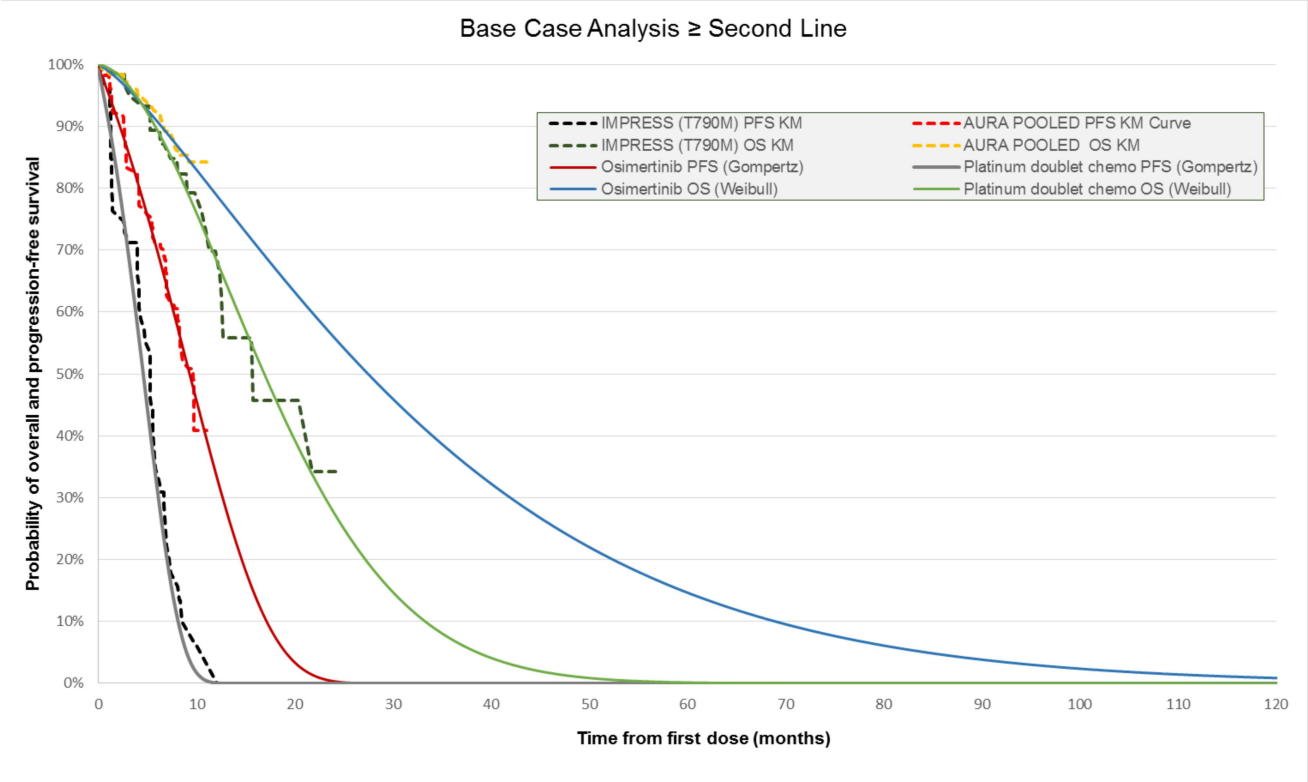


Figure 4 Overall survival and progression-free survival curves used in the base case analysis

OS=overall survival; PFS=progression-free survival; KM=Kaplan-Meier
Source: CS, Figure 5.7

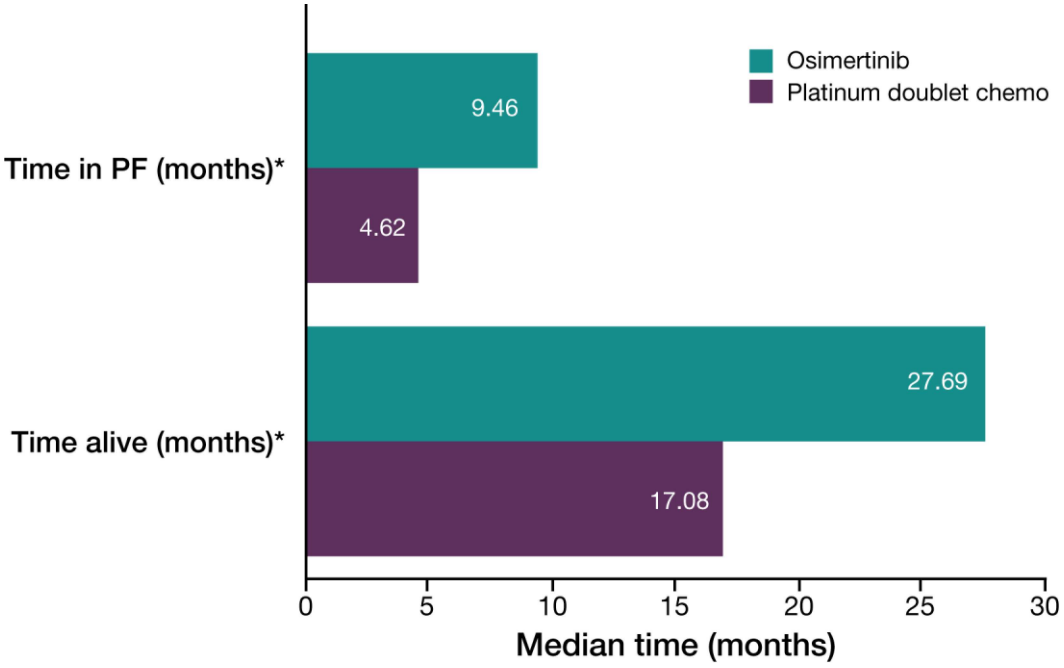


Figure 5 Median duration of the parametric distributions used in the base case analysis

PF=progression-free; *undiscounted results
Source: CS, Figure 5.8

Data from single-agent chemotherapy studies – second-line only subgroup

There are no published survival data available to demonstrate the effect of docetaxel monotherapy on patients with locally advanced or metastatic EGFR T790M mutation-positive NSCLC who have progressed on or after EGFR-TKI therapy. The company, therefore, has assumed that docetaxel has the same efficacy as pemetrexed in the second-line only setting. Data from the study by Park[29] were used in the company model; this study included 37 patients in the pemetrexed arm whose EGFR T790M mutation status is unknown.

Data from single-agent chemotherapy studies – ≥third-line only subgroup

The company used data from the Schuler[30] study in their model; in this study there were 68 patients included in the analysis. The company assumed that all single-agent chemotherapies had the same efficacy in the ≥third-line setting. The study by Shuler[30] was not specific to patients with EGFR T790M mutation-positive disease and documentation of EGFR mutation status was not mandatory prior to study entry, instead a clinically enriched EGFR inclusion criterion was used.

For simplicity, the parametric distributions selected to model PFS and OS for these subgroup analyses were the same as the distributions used in the base case analysis.

5.4.6 Health related quality of life

The original systematic review carried out by the company did not identify any HRQoL or utility studies that were relevant to the decision problem described in the final scope issued by NICE.

HRQoL data were collected using the EQ-5D-5L during the AURA2 study. As an EQ-5D-5L tariff has not been formally published or recommended by NICE, the EQ-5D-5L crosswalk index values[58] for the UK were applied. The utility values collected during the AURA2 study are shown in Table 24.

Table 24 Average EQ-5D-5L utility values for progression-free and progressed disease states from AURA2 study

Health state	N	Mean utility	Standard deviation
Base case analysis (≥second-line population)			
Progression-free	158	0.815	0.183
Post-progression	39	0.678	0.314
Second-line only population			
Progression-free	50	0.853	0.139
Post-progression	11	0.726	0.319
≥Third-line population			
Progression-free	108	0.798	0.198
Post-progression	28	0.659	0.316

Source: CS, Table 5.15

HRQoL data were collected using the EQ-5D-3L during the IMPRESS trial. The utility values collected from patients in the control arm of the IMPRESS trial are shown in Table 25.

Table 25 Average EQ-5D-3L index value from IMPRESS (control arm)

Health state	N	Mean utility	Standard deviation
Progression-free	117	0.779	0.210
Post-progression	88	0.679	0.271

Source: CS, Table 5.16

Treatment specific utility values were not used in the company base case analysis for the PF and PD health states. Instead, the company used values from the AURA2 study only. The company is confident that this is the most appropriate approach to adopt for the base case analysis. However, to test this assumption, the company applied treatment specific utility values in scenario analyses.

5.4.7 Adverse events

Utility decrements, due to grade 3 or grade 4 AEs were included in the company base case analysis. The company assumed that the disutility associated with AEs lasted for period of 4 weeks. To estimate utility decrements, the company mainly used previously published values from a study by Nafees.[59] The AE disutilities used in the company model are shown in Table 26.

Table 26 Disutilities associated with adverse events

Adverse event	Disutility	Source
Diarrhoea	0.047	Nafees 2008[59]
Rash (grouped term)	0.032	Nafees 2008[59]
Nausea	0.048	Nafees 2008[59]
Fatigue/asthenia	0.073	Nafees 2008[59]
Vomiting	0.048	Nafees 2008[59]
Febrile neutropenia	0.090	Nafees 2008[59]
Neutropenia/Leucopenia/ neutrophil count decreased	0.090	Nafees 2008[59]
Febrile neutropenia	0.090	Nafees 2008[59]
Anaemia	0.073	Assumed to be same as fatigue/asthenia
Platelet count decreased	0.05	Assumption based on previous STA[56]
Oedema peripheral	0.05	Assumption
Constipation	0.05	Assumption
Cough	0.05	Assumption
Stomatitis	0.05	Assumption
Headache	0.05	Assumption
Back pain	0.05	Assumption

Source: CS, Table 5.18

In the model, AEs were entered as one-off events. The costs of the AEs used in the model are listed, with sources, in Table 27.

Table 27 Cost of adverse events

Adverse event	Cost	Source/comment
Diarrhoea	£431.54	NHS Reference Costs 2014–15[60] FZ36G-FZ36Q Gastrointestinal Infections with Multiple Interventions – Non-elective short stay (Weighted Average) [NHS 2015]
Rash (grouped term)	£435.92	NHS Reference Costs 2014–15[60] JD07A-JD07K Skin Disorders with Interventions – Non-elective short stay (Weighted Average) [NHS 2015]
Nausea/vomiting	£449.94	NHS Reference Costs 2014–15[60] FZ91A-FZ91M Non-Malignant Gastrointestinal Tract Disorders with Multiple Interventions – Non-elective short stay (Weighted Average) [NHS 2015]
Decreased appetite	£83.00	NHS Reference Costs 2014–15[60] Assumed one outpatient dietician visit [NHS 2015]
Platelet count decreased	£502.63	NHS Reference Costs 2014–15 [60] SA12G-SA12K Thrombocytopenia – Non-elective short stay (Weighted Average) [NHS 2015]
Neutropenia/ leucopenia/ neutrophil count decreased	£478.31	NHS Reference Costs 2014–15[60] SA35A-SA35E Agranulocytosis – Non-elective short stay (Weighted Average) [NHS 2015]
Fatigue/asthenia/anaemia	£610.63	NHS Reference Costs 2014–15[60] SA01G-SA01K Acquired Pure Red Cell Aplasia or Other Aplastic Anaemia – non-elective short stay (Weighted Average) [NHS 2015]
Oedema peripheral	£365.66	NHS Reference Costs 2014–15[60] WH10A-WH10B Unspecified Oedema – Non-elective short stay (Weighted Average) [NHS 2015]
Constipation	£0.00	Assumed to be zero cost
Cough	£0.00	Assumed to be zero cost
Stomatitis	£0.00	Assumed to be zero cost – as per ipilimumab TA319[61] NICE submission [NICE 2014]
Headache	£0.00	Assumed to be zero cost – as per ipilimumab TA319[61] NICE submission [NICE 2014]
Febrile neutropenia	£2,426.86	NHS Reference Costs 2014–15[60] SA35A-SA35E Agranulocytosis – Non-elective long stay (Weighted Average) [NHS 2015]
Back pain	£421.67	NHS Reference Costs 2014–-15[60] HC32H-HC32K: Low Back Pain Without Interventions – Non-elective short stay (Weighted Average) [NHS 2015]

Source: CS, Table 5.30

5.4.8 Resources and costs

EGFR T790M mutation testing

It is assumed in the model that only identified EGFR T790M mutation-positive patients are treated with osimertinib. In the CS, the company describes four possible testing strategies:

1. tissue biopsy
2. ctDNA (plasma) test followed by tissue biopsy in patients identified as being EGFR T790M negative by ctDNA (plasma) test
3. ctDNA (plasma) test alone
4. tissue biopsy followed by ctDNA (plasma) test.

EGFR T970M incidence and test sensitivity and specificity of all four tests are presented in the CS (Table 5.19). In addition, the EGFR T790M diagnostic strategy outputs are also reported in the CS (Table 5.20). However, only testing strategies 1 and 2 are considered in the base case analysis where the company assumes that 20% of patients undergo tissue biopsy alone and 80% undergo ctDNA (plasma) test followed by tissue biopsy. The same diagnostic strategies, incidence and test performance estimates apply in all scenario analyses.

The cost of EGFR T790M mutation testing includes the acquisition cost of the test plus other costs that are incurred during the visit to undertake the test (Table 28).

Table 28 EGFR T790M test costs

Resource	Tissue biopsy	ctDNA	Source/Comment
Test cost	£147	£147	Tissue biopsy: based on cost of cobas EGFR Test[62] ctDNA: assumed to be same as tissue biopsy
Sample procedure	£578	£325	Tissue biopsy: £578 NHS Ref Costs: DZ70Z Endobronchial Ultrasound Examination of Mediastinum[60] ctDNA: assumption
Total cost	**£725**	**£472**	**Total costs applied in the model**

Source: CS, Table 5.21

Drug acquisition costs of initial treatment

The treatment dosing, administration and drug acquisition costs used in the model are shown in Table 29. At the discretion of the investigators, patients could continue to receive osimertinib beyond disease progression. The dosages for pemetrexed+cisplatin and docetaxel monotherapy are based on average patient characteristics, in terms of body weight, BSA and glomerular filtration rate, of patients included in the AURAext and AURA2

studies. The base case analysis uses patient characteristics from all of the patients in the AURAext and AURA2 studies, whilst data from the second-line only and ≥third-line subgroups are used in the subgroup analyses. The patient characteristics used to inform drug acquisition costs are presented in the CS (Table 5.22). Treatment dosing, administration and acquisition costs are presented in Table 29.

Table 29 Treatment dosing, administration and drug acquisition costs used in the model

		Osimertinib	PDC		Docetaxel monotherapy
			Pemetrexed	Cisplatin	
Label information	Admin method	Oral	IV	IV	IV
	Dose per admin	78.9mg	500mg/m^2	75mg/m^2	75mg/m^2
	Frequency	Once daily	Once every third week		
	Duration	TDP	TDP or maximum 6 doses		
Package information	Formulation	80mg	100mg	1mg	20mg/ml
	Pack size	30	1	10	7
	Price	£4722	£160.00	£3.24	£20.95
Dosing used in model	Required dose	80mg	840mg/m^2	126mg/m^2	126mg/m^2
	Vials/caps per admin (with waste)	1.00	9.00	13.00	1.00
	Vials/caps per admin (without waste)	1.00	8.40	12.59	0.90

PDC=platinum doublet chemotherapy; TDP=treatment until disease progression; admin=administration
Source: CS, Table 5.23 and BNF 2015; eMIT (accessed January 2016)

Drug administration costs

For osimertinib (oral medication), administration costs were assumed to be £0. The drug administration costs for all intravenous therapies comprise the costs of chemotherapy infusion and premedication with dexamethasone. In the model, administration costs are applied to all patients on treatment and are shown in Table 30.

Table 30 Unit costs, resource use and total administration costs used in the model

Treatment	Cost item	Unit cost	Sum	Source
Platinum doublet chemotherapy	Chemotherapy IV infusion – First attendance	£239.12	£245.16	NHS Ref Costs 2015; DH 2011[60,63]
	Dexamethasone (8mg/day for 3 days)	£6.04		
	Chemotherapy IV infusion – Subsequent attendances	£326.46	£332.50	
	Dexamethasone (8mg/day for 3 days)	£6.04		
Docetaxel monotherapy	Chemotherapy IV infusion – First attendance	£239.12	£251.19	
	Dexamethasone (16mg/day for 3 days)	£12.07		
	Chemotherapy IV infusion – Subsequent attendances	£326.46	£338.53	
	Dexamethasone (16mg/day for 3 days)	£12.07		

Source: CS, Table 5.24

Drug costs

A summary of the drug monitoring costs used in the company model is shown in Table 31. Costs were taken from NHS Reference Costs (2014-15).[60] Frequencies were based on data submitted to NICE as part of the nintedanib STA submission[56] and were applied to all treatments.

Table 31 Unit costs, resource use and total weekly monitoring costs used in the model

Treatment	Cost item	Numbers per week	Unit cost	Sum
Osimertinib	–	–	–	£0.00
Platinum doublet chemotherapy	Liver function test	0.153	£7.00	£4.61
	Renal function test	0.153	£10.00	
	Complete blood count	0.667	£3.00	
Docetaxel	Complete blood count	0.667	£3.00	£2.00

Source: CS, Table 5.25

Subsequent treatment costs

The company assumes that patients who progress whilst on treatment with second-line osimertinib will subsequently be treated with PDC and then with single-agent pemetrexed or docetaxel. The company assumes that patients who progress on second-line PDC will subsequently be treated with single-agent chemotherapy and then with best supportive care. In the model the distribution of patients across subsequent treatments for each second-line treatment is based on UK clinical expert opinion. The duration of all subsequent treatments is assumed to be the same as the modelled duration of the second-line treatment. The cost of subsequent treatment is applied as a one-off cost for all patients entering the PD state (Table 32).

Table 32 Distribution and costs of subsequent treatment

To ↓ / From →	Osimertinib	PDC	Docetaxel
Base case analysis (≥2nd-line) and 2nd-line only subgroup			
Platinum doublet chemotherapy	80%	0%	0%
Docetaxel monotherapy	50%	50%	15%
Best supportive care	70%	50%	85%
Total	200%*	100%	100%
≥3rd-line subgroup			
Platinum doublet chemotherapy	0%	N/A	0%
Docetaxel monotherapy	50%	N/A	15%
Best supportive care	50%	N/A	85%
Total	100%	N/A	100%
Total cost per patient on subsequent treatment (≥second-line)	£7,304	£609	£183

PDC=platinum doublet chemotherapy
*Note that total proportion for osimertinib is 200% in 2L or ≥2L setting to reflect that patients will have two subsequent treatments following progression on osimertinib treatment
Source: CS, Table 5.26

Health state unit costs and resource use

In the CS (Tables 5.27 and 5.28), the company gives a detailed summary of the costs and resource use associated with disease management in the PF and PD health states. In addition, the company provides a breakdown of a one-off 'end of life/terminal care' cost that is applied in the model (CS, Table 5.29). In summary, the total weekly cost of disease management in the PF and PD health states is £77.42 and £139.52 respectively. The overall weighted 'end of life/terminal care' cost is £3,905.26.

5.4.9 Cost effectiveness results

Total costs, life years gained (LYG), QALYs and incremental costs per QALY gained for the cost effectiveness comparison of treatment with osimertinib versus PDC are shown in Table 33. In the base case, osimertinib generates more benefits than PDC [redacted]LYG and +[redacted] QALYs) at an increased cost of [redacted]. The company base case incremental cost effectiveness ratio (ICER) for osimertinib versus PDC is [redacted] per QALY gained.

Table 33 Base case results

Technologies	Total costs	Total LYG	Total QALYs	Δ Costs	Δ LYG	Δ QALYs	ICER per QALY gained
Osimertinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PDC	[redacted]	[redacted]	[redacted]				

PDC=platinum doublet chemotherapy; Δ=change; LYG=life years gained; QALY=quality adjusted life year; ICER=incremental cost effectiveness ratio
Source: CS, Table 5.32

5.4.10 Sensitivity analyses

Deterministic sensitivity analyses

The company undertook one-way sensitivity analyses by varying key model parameters by +/- 20% around the mean values applied in the base case. The results of these analyses are shown in Table 34 and Figure 6.

The results show that the ICER per QALY gained is most sensitive to the utility values used in the model, particularly for the PD state. In addition, the model is sensitive to the choice of discount rate.

Table 34 Results of deterministic sensitivity analysis – osimertinib vs PDC

Parameter		Parameter values			Lower value (ICER)	Upper value (ICER)
		Lower value	Base case	Upper value		
Body surface area (m²)		1.34	1.68	2.02	[redacted]	[redacted]
Discount rate	Costs	0.0%	3.5%	[redacted]	[redacted]	[redacted]
	Outcomes	0.0%	3.5%	[redacted]	[redacted]	[redacted]
Disease management	PF	£62	£77	[redacted]	[redacted]	[redacted]
	PD	£112	£140	[redacted]	[redacted]	[redacted]
	TC	£3,124	£3,905	[redacted]	[redacted]	[redacted]
Drug acquisition cost: PDC		£369	£461	£554	[redacted]	[redacted]
Drug acquisition cost: Docetaxel		£5	£6	£8	[redacted]	[redacted]
Testing cost	ctDNA	£378	£472	[redacted]	[redacted]	[redacted]
	Biopsy	£752	£940	[redacted]	[redacted]	[redacted]
Health state utility	Osimertinib: PF	0.652	0.815	[redacted]	[redacted]	[redacted]
	Osimertinib: PD	0.542	0.678	[redacted]	[redacted]	[redacted]
	PDC: PF	0.652	0.815	[redacted]	[redacted]	[redacted]
	PDC: PD	0.542	0.678	[redacted]	[redacted]	[redacted]

PF=progression-free; PD=progressed disease; TC=terminal care; PDC=platinum doublet chemotherapy; ICER=incremental cost effectiveness ratio
Source: CS, Table 5.39

Figure 6 Tornado diagram – osimertinib versus PDC

Source: CS, Figure 5.11

Probabilistic sensitivity analysis

The company undertook a probabilistic sensitivity analysis (PSA) to derive the mean ICER per QALY gained for the comparison of osimertinib versus PDC. The PSA was run for 10,000 iterations. Results from the PSA are shown in Table 35. The probabilistic ICER per

QALY gained for osimertinib versus PDC is [redacted], which is comparable to the deterministic ICER per QALY gained of [redacted].

Table 35 Average results based on the probabilistic sensitivity analysis (10,000 iterations)

Treatment	Total costs	QALYs	Δ costs	Δ QALYs	ICER per QALY gained
Osimertinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PDC	[redacted]	[redacted]			

Δ= change; PDC=platinum doublet chemotherapy; QALY=quality adjusted life year; ICER=incremental cost effectiveness ratio
Source: CS, Table 5.37

The cost effectiveness plane and cost effectiveness acceptability curve for the comparison of osimertinib versus PDC are shown in Figure 7 and Figure 8.

At a cost effectiveness threshold of £50,000 per QALY gained osimertinib has a 35% probability of being cost effective compared with PDC. At a cost effectiveness threshold of £30,000 per QALY gained osimertinib has a <2% probability of being cost effective compared with PDC.

Figure 7 Cost effectiveness plane for osimertinib versus PDC

Source: CS, Figure 5.9

Figure 8 Cost effectiveness acceptability curve for osimertinib versus PDC

Source: CS, Figure 5.10

5.4.11 Scenario analyses

The company undertook several scenario analyses to explore alternative approaches to: survival modelling, as well as different values for health state utilities, resource use and costs. Full descriptions of the scenario analyses undertaken by the company are described in Section 5.8.3 of the CS. The results of the scenario analyses for osimertinib versus PDC are shown in Table 36.

Table 36 Results of scenario analyses for osimertinib versus PDC

Scenario	Total costs (£)		Total QALYs		Incremental		ICER per QALY gained
	Osimertinib	PDC	Osimertinib	PDC	Costs	QALYs	
Base case	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Survival modeling							
IMPRESS ITT population PFS/OS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PFS & OS distribution – Log logistic (both arms)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PFS & OS distribution – Log normal (both arms)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PFS & OS distribution – Weibull (both arms)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PFS & OS distribution – Generalised Gamma (both arms)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PFS & OS distribution – Gompertz (both arms)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PFS & OS distribution – exponential (both arms)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Health state utility values							
Treatment-specific utilities (osimertinib/AURA; PDC/IMPRESS)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PD utility decrement (Nafees[59]): -0.1798 (both arms)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Resource use and costs							
Exclude T790M test costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment after RECIST progression - osimertinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Assume pemetrexed generic costs (75% discount)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

PD=progressed disease; PFS=progression-free survival; OS=overall survival; PDC=platinum doublet chemotherapy; ICER=incremental cost effectiveness ratio; QALY=quality adjusted life year gained; ITT=intention to treat
Source: CS, Table 5.40

5.4.12 Subgroup analyses

To explore the cost effectiveness of osimertinib compared with other the comparators listed in the final scope issued by NICE, the company performed three main separate subgroup analyses:

- osimertinib versus PDC in a second-line population only
- osimertinib versus docetaxel monotherapy in a second-line population only
- osimertinib versus single agent chemotherapy in a ≥third-line population only.

In each of the subgroup analyses, the following parameters were dependent on line of treatment: patient demographics (CS, Table 5.22), survival data (CS, Section 5.3), safety (CS, Table 5.14) and subsequent treatments (CS, Table 5.26).

Osimertinib versus PDC in a second-line population only

Using data specific to the second-line only population of the AURAext and AURA2 studies (for osimertinib) and data from the IMPRESS trial (for PDC), the ICER per QALY gained for osimertinib versus PDC is [redacted] as shown in Table 37. When health state utility values from the AURA2 study were applied to both treatment arms for the second-line only population, the ICER per QALY gain decreased slightly to [redacted].

Table 37 Subgroup analysis – osimertinib versus PDC (second-line only population)

Treatment	Total cost	Total QALYs	Δ costs	Δ QALYs	ICER per QALY gained
Osimertinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PDC	[redacted]	[redacted]			

Δ=change; PDC=platinum doublet chemotherapy; QALY=quality adjusted life years; ICER=incremental cost effectiveness ratio
Source: CS, Table 5.42

Osimertinib versus docetaxel monotherapy in a second-line population only

Using data specific to the second-line only population of the AURAext and AURA2 studies (for osimertinib) and data from the study by Park[29] (for single-agent docetaxel), the ICER per QALY gained for osimertinib versus docetaxel is [redacted] as shown in Table 38. When health state utility values from the AURA2 study were applied to both treatment arms for the second-line only population, the ICER per QALY gain decreased slightly to [redacted].

Table 38 Subgroup analysis – osimertinib versus docetaxel monotherapy (second-line only population)

Treatment	Total cost	Total QALYs	Δ costs	Δ QALYs	ICER per QALY gained
Osimertinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Docetaxel	[redacted]	[redacted]			

Δ=change; QALYS=quality adjusted life years; ICER=incremental cost effectiveness ratio
Source: CS, Table 5.43

Osimertinib versus single agent chemotherapy in a ≥third-line population only

Using data specific to the ≥third-line only population of the AURAext and AURA2 studies (for osimertinib) and data from the study by Shuler[30] (for single-agent docetaxel), the ICER per QALY gained for osimertinib versus single agent chemotherapy is [redacted] as shown in Table 39. When health state utility values from the AURA2 study for the ≥third-line population were applied to both treatment arms, the ICER per QALY gain increased slightly to [redacted]

Table 39 Subgroup analysis – osimertinib versus single agent chemotherapy (≥third-line population)

Treatment	Total cost	Total QALYs	Δ costs	Δ QALYs	ICER per QALY gained
Osimertinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Single agent chemotherapy	[redacted]	[redacted]			

Δ=change; QALYS=quality adjusted life years; ICER=incremental cost effectiveness ratio
Source: CS, Table 5.4

5.4.13 Model validation and face validity check

In order to validate the de novo cost effectiveness analysis, the company carried out the following checks:

1. the predicted model outcomes for the base case analysis were compared to the observed pooled AURA dataset to confirm that the model behaved as expected and produced PFS and OS curves similar to the observed data

2. as it was not possible to verify predicted long-term outcomes due to the absence of external long-term data, the company analysed the relationship between PFS and OS data from other trials in advanced NSCLC to check if the company model estimated a ratio of OS to PFS that could be considered to be clinically plausible. The results from these analyses are available in the CS (Table 5.47).

In summary, the company is confident that the base case cost effectiveness estimates generated by the model are valid and are unlikely to be biased in favour of osimertinib.

5.5 ERG critique of company's submitted economic evaluation

5.5.1 NICE reference case checklist

Table 40 NICE Reference case checklist completed by ERG

Attribute	Reference case	Does the de novo economic evaluation match the reference case?
Defining the decision problem	The scope developed by NICE	Yes – although first-line patients were included in the scope, the ERG agrees with the company that these comprise a very small number of patients. The ERG considers that there is insufficient evidence of the clinical effectiveness of osimertinib in a first-line setting to include these patients in this economic evaluation
Comparator(s)	As listed in the scope developed by NICE	Yes
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	Partial. The model only includes NHS costs. Personal Social Service costs have not been considered
Perspective on costs	NHS and PSS	Patient related direct health effects are considered. No impact on carers has been considered in the model
Type of economic evaluation	Cost-utility analysis with fully incremental analysis	Yes
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	Yes – 15 year time horizon
Synthesis of evidence on health effects	Based on systematic review	Yes
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of HRQoL in adults	Yes
Source of data for measurement of health-related quality of life	Reported directly by patients and/or carers	Yes
Source of preference data for valuation of changes in HRQoL	Representative sample of the UK population	Partially – crosswalk values used for EQ-5D-5L
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	Yes
Evidence on resource use and costs	Costs should relate to NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS	Yes
Discounting	The same annual rate for both costs and effects (currently 3.5%)	Yes

EQ-5D-5L=EuroQol-5 dimension, 5 levels; QALY=quality adjusted life year

5.5.2 Drummond checklist

Table 41 Critical appraisal checklist for the economic analysis completed by the ERG

Question	Critical appraisal	ERG comment
Was a well-defined question posed in answerable form?	Yes	
Was a comprehensive description of the competing alternatives given?	Yes	
Was the effectiveness of the programme or services established?	Partial	The immaturity of the OS data means that currently no statistically significant evidence exists that the treatment extends OS. However, a PFS benefit is statistically significant.
Were all the important and relevant costs and consequences for each alternative identified?	Partial	The ERG considers that the company should have included more detail relating to adverse events in their model
Were costs and consequences measured accurately in appropriate physical units?	Partial	The ERG revised the following parameter estimates in the company's model: utility values, treatment costs and administration costs
Were the cost and consequences valued credibly?	Yes	
Were costs and consequences adjusted for differential timing?	Yes	However, discounting was applied weekly rather than annually
Was an incremental analysis of costs and consequences of alternatives performed?	Yes	
Was allowance made for uncertainty in the estimates of costs and consequences?	Partial	Deterministic and probabilistic sensitivity analyses were undertaken. Further scenario analyses were required to test the key assumptions in the model
Did the presentation and discussion of study results include all issues of concern to users?	No	Key scenario analyses were not undertaken – notably around time on treatment and the efficacy of treatment being limited to a change in PFS - and therefore the results were not presented

OS=overall survival; PFS=progression-free survival; ERG=Evidence Review Group

5.6 Detailed critique of company's economic model

Sections 5.6.1 to 5.6.5 of the ERG report provide details of five issues that have a major impact on the cost effectiveness results generated by the company model (i.e., estimation of OS, estimation of PFS, cost of osimertinib therapy, QoL and cost of administering osimertinib). Issues that only have a minor impact on the cost effectiveness results are described in Section 5.6.6.

The company provided the model in Microsoft Excel. The ERG considers that it was well constructed with no flaws in the algorithms and was straightforward to use.

5.6.1 Overall survival estimation

Company model results for the comparison of osimertinib versus PDC suggest that over 90% of the QALY gains associated with treatment with osimertinib are generated after 10 months, i.e. during the period when no trial data on osimertinib are available. This means that confidence in the results generated by the company model is highly dependent on the degree to which the projection of osimertinib OS employed in the company model, past the point trial data for osimertinib is available, reflects reality.

The company considered a range of distributions to model OS for both osimertinib and PDC with each distribution used to represent OS over the whole model time horizon (15 years). The company first used statistical tests (AIC and BIC) to help choose a distribution to represent OS. The ERG notes that AIC and BIC do not provide an indication that a particular distribution has an acceptable goodness of fit or, indeed, they do not identify the extent to which one distribution may fit the data better than another. However, the company's choice of distribution was also influenced by visual inspection, as well as a discussion around clinical plausibility. The ERG, therefore, considers that the company's approach to selecting a distribution to represent OS for both osimertinib and PDC was broadly acceptable given the paucity of relevant survival data available – especially for osimertinib. The ERG also notes that it would have been preferable to use all of the available clinical trial data before employing the statistical distribution, rather than using the company's choice of distribution over the whole time horizon. However, in this case, given that the trial data are only available for 10 months for osimertinib, the ERG considers that using actual OS data before employing a distribution would have made an insignificant impact on the cost effectiveness results.

Whilst the approach employed by the company to select a parametric model to represent OS is satisfactory for both osimertinib and PDC, the ERG considers there are a number of

issues that cast substantial doubt on the plausibility of the OS projections for osimertinib and, to a lesser extent but still significantly, for PDC in the company model.

Lack of statistical evidence of differential OS

Any extrapolation beyond the period for which trial data are available is implicitly based on the assumption that the OS associated with osimertinib treatment is different from the OS associated with PDC treatment, and that such a difference is shown to exist within the period during which trial data are available.

Table 4.14 in the CS reported an OS hazard rate for the adjusted analysis of [redacted]

To further explore the potential statistical difference in OS between osimertinib and PDC, the ERG requested K-M OS data from the IMPRESS trial and the two AURA studies as part of the clarification process. ERG analyses of these data show that there is no statistically significant difference between the (pooled) K-M OS data for patients with EGFR T790M mutations participating in the two AURA studies and the K-M OS data for patients with EGFR T790M mutations included in the control arm of the IMPRESS trial (Log-rank test, p=0.33). This indicates that there is no statistical evidence that the two K-M data sets were derived from patients with different OS.

With no statistically significant evidence of difference in OS between osimertinib and PDC during the period that trial data are available, the ERG considers that there is no statistical justification to support the use of different OS projections for osimertinib and PDC.

Proportion of benefit arising from extrapolation

The ERG considers that any ICER that relies on a QALY benefit that is over 90% generated by a projection is highly uncertain and that this level of uncertainty renders its use in decision-making questionable. The ERG considers that statistical and modelling techniques can only be employed to help describe this uncertainty and cannot be used to overcome it. Whilst OS data for PDC from the IMPRESS trial are more mature than the OS data for osimertinib (approximately 33% for PDC compared to 12.7% for osimertinib), the ERG considers the projections of OS for PDC to be only slightly less uncertain than for osimertinib.

The ERG applauds the company for investigating this uncertainty and for demonstrating how this uncertainty affects the size of the estimated ICERs. For example, in the CS (p234), the company provides the results of a scenario analysis where the use of different projection

methods is investigated. The cost effectiveness results vary depending on the method used. For example, using a Log normal distribution for PFS and OS generates an ICER per QALY gained of [redacted] whereas using a Gompertz distribution for PFS and OS generates an ICER per QALY gained of [redacted].

The ERG considers that all of the distributions that were used in the scenario analysis can, visually, be considered to provide a good fit to the available osimertinib OS K-M data. However, the ERG is not confident that any of the ICERs generated by the company model are sufficiently robust to inform decision-making.

Company acknowledgement of weakness of OS data

As part of the clarification process the ERG requested post-progression survival (PPS) and post-treatment discontinuation survival data (PTDS). The company did not supply these data, in part because the data were too immature. The exact response from the company was:

> *We believe that there are likely to be significant issues with the interpretation of these time-to-event outcomes due to the immaturity of currently available data and the high level of censoring of patients in the post-progression period from AURAext/2.* (Source: Company clarification response. QB1)

Patients who die before progression, or who die before treatment discontinuation, are considered to be in the PPS state and treatment is discontinued at the point of death. The ERG considers that the maturity of the PPS and PTDS data, and the level of censoring are, therefore, identical to the maturity and censoring of the OS data sets for the two AURA studies and the IMPRESS trial. This point is also highlighted in the CS where the following statement is made about the limitations of the current evidence base:

> *While confidence regarding the analyses of the primary endpoints of ORR, secondary endpoints of PFS and safety/tolerability assessments can be considered high, caution should be exercised when interpreting the results of the OS analyses. The OS data are very immature at the time of analysis (Osimertinib 11.5% maturity [adjusted analysis, n=287] and platinum doublet chemotherapy 29.4% maturity [adjusted analysis, n=51]). Consequently in the matched adjusted comparison in both groups the KM risk set beyond 12 months is very limited (n <15 patients) leading to unstable estimates beyond this time point, especially for the estimation of medians.* (Source: CS, p164)

In summary, the ERG agrees with the company that the OS, PPS and PTDS data from the pooled AURA dataset and the IMPRESS trial are immature and should be interpreted with caution.

Weak link between PFS and OS

The company presents data on time spent in the PFS and OS states, as predicted by the company model, as a means of justifying the distribution chosen to represent OS (CS, Table 5.47). The ERG acknowledges that the ratio of time in OS to PFS predicted by the company model is similar for patients treated with osimertinib (2.85) and for patients treated with PDC (2.96), and is within the range of ratios observed in other trials for patients receiving other second-line treatments. However, the reported range across studies is large (between 2.18 and 5.38 for active treatment arms and 2.24 and 7.60 in control arms), which suggests that the relationship between OS and PFS is complex and that there is no basis to assume that the same, or similar, OS/PFS ratios exist. This view is supported by the authors of a DSU[57] report that contains details of a literature review that was undertaken to examine the relationship between PFS and OS in people with advanced or metastatic cancers. The authors found that the evidence supporting a relationship between PFS and OS varies considerably by cancer type and, furthermore, is not always even consistent within one cancer type. In addition, they advise that:

> *…any cost-effectiveness analysis which makes a strong assumption regarding the relationship between PFS and OS should be treated with caution.* (Source: DSU report,[57] p39)

Generalisability of trial data to the UK population

Aside from concerns about the reliability of the OS representation of both osimertinib and PDC within the company model, the ERG considers that, even if the OS data were fully mature, they would not reflect the experience of the population described in the final scope issued by NICE. This is because the populations included in the two AURA studies and the IMPRESS trial have a baseline ECOG PS of 0 or 1 and, in the two AURA studies, have received at least one (and up to 14) previous lines of treatment. Clinical advice to the ERG is that only about 40% of NHS patients with advanced NSCLC are likely to have an ECOG PS of 0 to 1. Furthermore, very few NHS patients with metastatic NSCLC receive even three lines of treatment and, at the most, 30% of patients whose cancer progresses on or following treatment with an EGFR-TKI are well enough to receive chemotherapy.

The absence of trial data on patients with ECOG PS ≥2 and on patients who are too unwell to receive chemotherapy means that the impact of osimertinib or PDC in a UK population

who would be eligible for treatment under the final scope issued by NICE is not fully known. However, the ERG considers that the OS of a population that is less well than the populations included in the two AURA studies and in the IMPRESS trial is likely to be shorter than the OS demonstrated by the aforementioned trials.

5.6.2 Progression-free survival estimation

The PFS K-M data from the two AURA studies and the EGFR T790M mutation-positive population included in the control arm of the IMPRESS trial are statistically significantly different (Log-rank test, $p<0.001$). There is, therefore, statistical justification to assume differential PFS for osimertinib and PDC over the period for which IPD are available, and then to extrapolate the difference past that time point. This approach should only be carried out if acknowledging that the PFS data from a single arm phase II study (such as the AURA2 study) have greater potential for bias than the PFS data from a phase III RCT (such as the IMPRESS trial). The ERG considers that the projections employed by the company should be considered with a degree of caution.

As was the case when considering the projection of OS, PFS data were incorporated into the company model through the use of parametric curves estimated from the available PFS data from the two AURA studies and from the subgroup of patients with EGFR T790M mutations in the control arm of the IMPRESS trial. The ERG considers that whilst the company's approach to parametric model selection is broadly acceptable, it is preferable if IPD can be incorporated directly into a model where it is available and that a parametric curve should only be employed when those data become unavailable or unreliable. However, in this case, the trial data so closely match the curves chosen by the company that implementation of the available K-M data prior to introducing a parametric survival curve, rather than using the parametric model from time zero, does not result in a significant change to the size of the ICERs generated by the company model. In fact, using K-M data in the model followed by the extrapolations suggested by the company leads to a change in the incremental QALYs of just [redacted] and an increase in the ICER per QALY gained for the comparison of osimertinib with PDC of [redacted] (an increase of approximately [redacted]. The ERG, therefore, considers that the method employed within the company model to represent PFS is, in this instance, satisfactory.

In summary, the results of analyses carried out by the ERG demonstrate that, when comparing the survival of patients receiving osimertinib with the survival of patients receiving PDC, there is no statistically significant difference between available K-M OS data but there is a statistically significant difference between available K-M PFS data. The ERG has carried out a scenario analysis that reflects this, i.e. differential PFS between osimertinib and PDC

but OS equal to that modelled by the company for PDC (due to the slightly greater maturity of the data from EGFR T790M mutation-positive patients in the control arm of the IMPRESS trial compared to the AURA studies). This analysis generates incremental QALYs of [redacted] and an ICER of [redacted] per QALY gained.

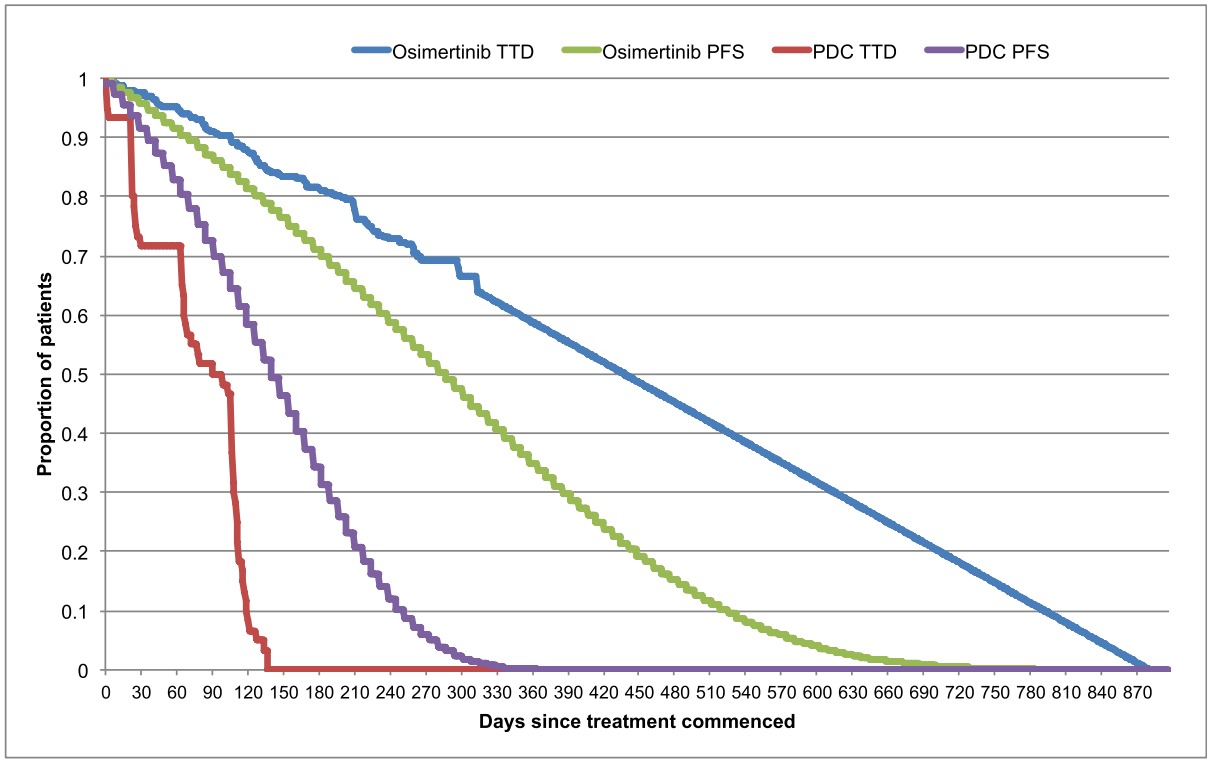
5.6.3 Cost of osimertinib treatment

In the two AURA studies, patients were permitted to continue receiving osimertinib after disease progression. As such, PFS is not a good basis for estimating treatment cost. As part of the clarification process, the ERG requested TTD data from the two AURA studies and for the EGFR T790M mutation-positive patients in the control arm of the IMPRESS trial. The pooled TTD data from the two AURA studies were 36% mature, with the last event recorded on day 313. The TTD data from EGFR T790M mutation-positive patients in the control arm the IMPRESS trial were 100% mature.

ERG analysis of the TTD data from the two AURA studies showed that they followed a simple linear decline. The ERG estimated the linear trend between days 0 and 313 and then continued the trend after day 313 to estimate TTD beyond the point that data were available. This resulted in an estimate of all patients stopping treatment with osimertinib by day 880, or around 2.5 years. If the hazard rate were to become constant at any point past day 313 then the TTD data would follow an exponential curve. If an exponential curve were used, it is likely that there would be a longer tail of patients remaining on treatment than is suggested by the linear projection of TTD with resultant higher costs for osimertinib.

The cost of PDC treatment is primarily driven by PFS status but is limited to a maximum number of four cycles of treatment in the model to match the protocol for pemetrexed-cisplatin therapy. The TTD data from the EGFR T790M mutation-positive patients in the control arm of the IMPRESS trial are complete and have been used, by the ERG, directly in the company model to generate an alternative estimate of the cost of treatment with PDC instead of using PFS as in the company base case.

A comparison of TTD and PFS data using pooled data from the two AURA studies (osimertinib) and data from the EGFR T790M mutation-positive patients included in the control arm (PDC) of the IMPRESS trial is displayed in Figure 9.



TTD=time to treatment discontinuation; PFS=progression-free survival; PDC=platinum doublet chemotherapy
Source: ERG analyses using TTD data requested via clarification process

Figure 9 TTD as estimated by the ERG and PFS as in the company model

5.6.4 Health related quality of life

The utility values used in the company's base case are taken directly from the AURA2 study. The figures used in the company model are 0.815 for the PF state and 0.678 for the post-progression state. These figures are the same, irrespective of whether patients are treated with osimertinib or PDC. The ERG considers that these values may not represent the HRQoL of the population with EGFR T790M disease treated in the NHS in a second-line setting as:

- the health states were taken from patients who were not from the UK
- the ECOG PS of patients was 0 or 1. According to clinical advice to the ERG, this would not be the case for a UK population where a number of patients with ECOG PS ≥2 would be treated.

In addition, there are several other factors that cast doubt on the validity of the utility values used in the company model:

- the HRQoL tool used in AURA2 was the EQ-5D-5L questionnaire and, as acknowledged within the CS (p205), this tool does not yet have a validated health state valuation set for the UK
- the mean utility value of people aged 55-64 in the UK is 0.80.[64] Whilst this mean utility value includes some people who are very ill, it seems implausible that a patient with advanced NSCLC will have a higher utility value (0.815) than the average person in the UK who is of a similar age.

There are no published alternative utility values that relate explicitly to the population of interest. On balance, the ERG considers that there are two studies that provide utility values that could be closer to the real utility of the target population than those in used in the company model: utility values collected during the LUME-Lung 1[65] trial and utility values reported in the Nafees study.[59]

The LUME-Lung 1[65] trial compares treatment with nintedanib+docetaxel with placebo+docetaxel in a population of previously treated patients with locally advanced or metastatic NSCLC (adenocarcinoma tumour histology). The utility values collected during the trial (and used in the STA[56]) range from 0.66 to 0.71 for patients in the PF state, whilst 0.64 is used to represent the utility of patients in the post-progression state. The population in the LUME-Lung 1[65] trial was slightly younger than the population in the AURA2 study (58 vs 62 years) and included fewer patients with brain metastases (10% vs 40%).

In the CS for the appraisal of nintedanib,[56] a range of utility values, adjusted for ECOG PS and brain metastases, are reported over a period of 30 weeks for patients in the PF state. The ERG considers that the midpoint value at 15 weeks (0.687) is a fair value to use to represent utility whilst in the PF state and has, therefore, used this value, along with the value of 0.64 to represent utility in the post-progression state.

The ERG considers that figures collected during the LUME-Lung 1[65] trial are likely to provide inaccurate utility estimates, especially for patients in the post-progression state where utility for (potentially) several years is derived at, or shortly after, the point of disease progression. In addition, the population included in the LUME-Lung 1[65] trial is, on average, likely to be healthier than patients with advanced or metastatic NSCLC being treated on or after failure of a TKI (e.g., LUME-Lung 1[65] trial patients had fewer brain metastases than patients in the AURA studies).

The ERG has also explored the impact on QALYs and ICERs of using lower utility values as provided in the Nafees[59] study. Nafees[59] estimated a utility value from a general population

assessment of NSCLC health states using the standard gamble method. The resultant figures are 0.653 for stable disease and 0.47 for progressed disease. Whilst valuations of health states from Nafees[59] are taken from the general population, the health states themselves were not taken from patients but are simple descriptors based upon breast cancer health states. A single 'stable' and 'progressed' state was also described rather than the range of health states that would be experienced by patients in PF and post-progression states. As such, the Nafees[59] values are also flawed but they do provide an alternative estimation of ICERs and it is possible that they may provide a better reflection of the experience of those with advanced or metastatic NSCLC than currently available from trial data, even if the way they were derived was not robust.

The impact on QALYs and resultant ICERs from the use of the different utility values considered by the company and by the ERG is summarised in Table 42.

Table 42 Utility values applied in company model and considered by ERG with resultant ICERs

Source	Patient population	Utility elicitation tool	PFS	PPS	QALYs		ICER per QALY gained
					Osimertinib	PDC	
AURA2 (company base case)	Second-line NSCLC after EGFR-TKI; ECOG status 0-1; T790M patients only	EQ-5D-5L	0.815	0.678	[redacted]	[redacted]	[redacted]
LUME-Lung 1[65] (ERG preferred value)	Second-line NSCLC with PDC as first-line (96% of patients); T790M status unconfirmed; ECOG status 0-1	EQ-5D-3L	0.687	0.64	[redacted]	[redacted]	[redacted]
Nafees[59]	General UK population	Bespoke standard gamble	0.653	0.47	[redacted]	[redacted]	[redacted]

PFS=progression-free survival; PPS=post-progression survival; QALYs=quality adjusted life years; ICER=incremental cost effectiveness ratio
Source: CS, Table 5.17, published studies

5.6.5 Osimertinib administration cost

The company model does not include a cost for the administration of osimertinib. Clinical advice to the ERG is that osimertinib is provided, on a monthly basis, in a nurse led clinic. The 2014-15 NHS Reference Cost[60] to deliver exclusively oral chemotherapy (SB11Z, setting: "Other") is £128. This is the lowest reference cost in the Reference Cost Schedule[60] for the delivery of oral chemotherapy. There is no simple way to apply this cost to the company model so the estimated cost of administering osimertinib has been calculated using the TTD data (provided in response to an ERG request during the clarification process). The effect of introducing a cost for administering osimertinib is to increase the total

cost per patient by [redacted] and to increase the ICER for osimertinib versus PDC to [redacted] per QALY gained.

5.6.6 Minor amendments

The ERG has identified issues relating to several of the parameter values used in the company model. However, exchanging the company values for the ERG's preferred values only has a minor impact on the size of the ICERs generated by the model.

Calculation of PDC costs per dose

The PDC treatment costs used in the company base case are based upon the age, weight and gender distribution of patients in the AURA studies.

Sacco[66] identified the characteristics of UK patients receiving palliative chemotherapy. The ERG considers that the characteristics of this patient group are more likely to reflect the characteristics of NHS patients undergoing second-line therapy than the characteristics used in the company base case. The estimated values for UK lung cancer patients in this group from the study by Sacco[66] have a body weight of 63.4kg for females and 74.7kg for males, with mean body surface area (BSA) of 1.66m^2 for females and 1.89m^2 for males. Application of these values in the company model results in an ICER for the comparison of osimertinib versus PDC of [redacted] per QALY gained, [redacted] per QALY gained.

Model structure

The company has developed a partitioned survival model and this structure has been used in previous NICE appraisals[56,67] of drugs for the treatment of advanced or metastatic cancer. This structure suffers from the limitation that it produces a counterintuitive finding i.e., the less time that patients stay in the PFS state, the more cost effective the intervention becomes. This model limitation has been discussed in previous ERG reports[68,69] as being challenging as it makes exploring the impact of assumptions around PFS on the cost effectiveness results problematic. Whilst this is not a major issue in this model, as the concerns around OS dominate the uncertainty in the ICER, the impact of choice of model structure on the ability to properly explore uncertainty in model parameters and assumptions should be fully considered in the CS.

Model time horizon

The time horizon used in the company model is 15 years. The ERG considers that this is optimistic given the population described in the final scope issued by NICE and the case, put forward by the company, that the Appraisal Committee should consider osimertinib as a life-extending, End of Life treatment.[50] However, as only [redacted] of the QALY gain for patients treated with osimertinib rather than PDC is accrued between years 10 and 15, i.e. 0.7% of the total (from baseline) QALY gain, the ERG considers that whilst a 15-year time horizon is probably optimistic, its use makes an insignificant difference to the size of the ICERs generated by the company model.

Discounting method

Discounting was applied continuously after year one to each weekly cycle. This is an incorrect application of the discount rate as costs and benefits should be summed over 12 months and then the annual discount rate should be applied. The company model cannot be easily modified to apply discounting correctly; however, the ERG does not consider that correcting this error will make any noticeable difference to the size of the ICERs generated by the company model.

Adverse events

In selecting AEs, the company has focussed only on events that are classified as being ≥grade 3. According to the Common Terminology Criteria for Adverse Events v4.0 (CTCAE),[70] a grade 3 AE is described as '*Severe or medically significant but not immediately life-threatening; hospitalisation or prolongation of hospitalisation indicated; disabling; limiting self-care ADL. Where self-care ADL (activities of daily living) refers to bathing, dressing and undressing, feeding self, using the toilet, taking medications, and not bed-ridden'*. Given this definition, some of the values chosen by the company appear to be arbitrary and implausible, both in terms of cost and disutility.

For example, there are zero costs associated with constipation, cough, stomatis and headache, and a low cost is associated with 'decreased appetite'. In addition, the disutility values associated with fatigue in particular and for AEs in general that would require hospitalisation appear to be low.

As osimertinib appears to be more tolerable than PDC, application of zero costs and implausibly low utility values for AEs, therefore, produces conservatively high ICERs when comparing the cost effectiveness of osimertinib and PDC. In addition, the AEs included in the company model result in £103 of additional cost, and 0.016 QALY loss, for PDC

compared to osimertinib. If all AEs were excluded from the model, then the ICER for osimertinib compared to PDC would rise by [redacted] to [redacted] per QALY gained.

In summary, the ERG considers that some of the costs and disutilities associated with AEs that are used in the company model may be unrealistic. However, as the magnitude of these parameter values only has a minor impact on the size of the ICER per QALY gained for the comparison of osimertinib versus PDC, and the approach taken by the company is conservative, the ERG has not reanalysed the ICERs using different AE costs and disutilities.

Testing costs

The ERG undertook a scenario analysis excluding testing costs for the EGFR T790M mutation. This reduced costs per patient by £1,351 thus lowering the ICER by [redacted] to [redacted] per QALY gained.

5.7 Subgroup analysis

The company undertook a number of subgroup analyses. These included considering different lines of treatment and use of docetaxel monotherapy rather than PDC.

As part of the clarification process the ERG requested OS, PFS and TTD data, by line of treatment, provided separately for the AURAext and AURA2 studies. Analyses carried out by the ERG found no statistical difference (using the Log rank test) for OS, PFS or TDD by line of treatment, irrespective of whether the data were considered by study (AURAext or AURA2) or if the pooled AURA dataset was used. As such, the ERG considers that the subgroup analyses by line of treatment, whilst correctly undertaken by the company, are not informative.

The ERG considers that the analysis of the cost effectiveness of osimertinib versus docetaxel, whilst worthy of pursuit by the company, is severely limited by the available clinical effectiveness data. In the analysis, the company has assumed that all single-agent chemotherapy drugs share equivalent efficacy that this is independent of EGFR T790M mutation status, to the point that the analysis against single agent chemotherapy (docetaxel in the company model) would be insufficiently robust even if AURA study OS data were mature.

6 IMPACT ON THE ICER OF ADDITIONAL CLINICAL AND ECONOMIC ANALYSES UNDERTAKEN BY THE ERG

A summary of the impact of the ERG's amendments to the company model on the cost effectiveness of osimertinib versus PDC for the treatment of patients with advanced or metastatic EGFR T790M mutation-positive disease in the second or further line setting, after failure of an EGFR-TKI, is included in ERG=Evidence Review GroupTable 43. The ERG has only implemented changes that have a major impact on the size of the ICERs and has not included changes relating to the minor issues described in Section 5.5.6. Details of all of the Microsoft Excel revisions made by the ERG to the company model are presented in Appendix 1 (and in the associated spreadsheet).

If the company's projection of OS were replicated in the NHS, then the biggest impact on the size of the ICERs from the ERG amendments would arise from the use of more accurate costs for the acquisition and administration of osimertinib and PDC. Use of TTD data to calculate the acquisition costs of osimertinib and PDC, in combination with an administration cost for osimertinib, increases the incremental cost of osimertinib compared to PDC from [redacted] in the company base case to [redacted]. This results in an increase in the ICER from [redacted] per QALY gained (Scenario A) to [redacted] per QALY gained (Scenario B).

Changes in utility values result in smaller changes to the size of the ICERs compared to changing costs. However, these utility value changes increase the size of the ICER substantially from the company base case ICER. Applying the ERG utility amendments increases the ICER for osimertinib compared to PDC from the company base case ICER to [redacted] per QALY gained using the LUME-lung 1[65] values (R3) and [redacted] per QALY gained using the values from Nafees[59] (R4).

Applying both cost and utility changes to the company base case ICER results in ICERs of [redacted] per QALY gained using the LUME-lung 1[65] utility values (Scenario C) and [redacted] per QALY gained using the Nafees[59] values (Scenario D).

If only the improvement in PFS is modelled, with equal OS for osimertinib and PDC due to the lack of statistically significant evidence to suggest otherwise, then application of the ERG cost amendments results in an ICER of [redacted] per QALY gained (Scenario E). Application of the ERG cost and utility amendments results in an ICER of [redacted] per QALY gained using the LUME-lung 1[65] utility values (Scenario F) and [redacted] per QALY gained using the Nafees[59] values (Scenario G).

ERG=Evidence Review GroupTable 43 ERG adjustments to company base case

Model scenario and revisions	Osimertinib			PDC			Incremental			ICER	
	Cost	**QALYs**	**Life years**	**Cost**	QALYs	Life years	Cost	QALYs	Life years	**£/QALY**	**Change from base case**
A. Company's base case	[redacted]										
R1) Use of TTD data to cost drug acquisition	[redacted]										
R2) Application of administration cost for osimertinib	[redacted]										
B. Base case + (R1:R2)	[redacted]										
R3) LUME-Lung 1[65] utility	[redacted]										
C. Base case + (R1:R3)	[redacted]										
R4) Nafees[59] utility	[redacted]										
D. Base case + (R1:R2 and R4)	[redacted]										
R5) Osimertinib generates a gain in PFS but not OS compared to PDC	[redacted]										
E. Base case + (R1:R2 and R5)	[redacted]										
F. Base case + (R1:R3 and R5)	[redacted]										
G. Base case + (R1:R2, R4:R5)	[redacted]										

ICER=incremental cost effectiveness ratio; OS=overall survival; PFS=progression-free survival; PDC=platinum doublet chemotherapy; QALY=quality adjusted life year; TTD=time to treatment discontinuation

6.1 Conclusions of the cost effectiveness section

The ERG considers that there are several fundamental issues that cast doubt on the cost effectiveness results produced by the company model.

First, over 90% of the QALY benefit from osimertinib estimated from the model arises when OS trial data are no longer available. The available OS data for osimertinib and PDC are not statistically significantly different and are very immature, especially so for osimertinib. The only statistically significant evidence on effectiveness incorporated in the model is an improvement in PFS with osimertinib, the extent of which is uncertain due to the single-arm nature of the AURA studies.

The ERG considers that lack of statistical significance in OS between osimertinib and PDC during the period for which data are available means that there is no basis to project differential OS.

With such immature OS data, even if there were a statistically significant difference between osimertinib and PDC during the period that data were available, any projection could only be speculative with the degree of uncertainty in the projection impossible to quantify. For example, whether the OS curve has a multi-phase distribution, or if the phases shift post the point of data that are available, is unknown.

Even if the OS data were sufficiently mature, the populations within the AURA studies and the IMPRESS trial appear to be fitter than the patients who would be expected to be seen in routine NHS practice. This casts doubt on the appropriateness of using the AURA and IMPRESS OS datasets to represent the UK population. Even if the OS projections were accurate for the patients in the trials, they would probably overestimate survival for the average target UK patient eligible for treatment with osimertinib.

The ERG therefore considers that the OS projections employed by the company are based on opinion rather than on robust clinical effectiveness evidence. To support this view, the ERG cites the wide variation in ICERs that the company shows (CS, p234) could be produced depending on the selection of different statistically plausible, if not necessarily clinically plausible, projections of OS.

The ERG considers that all of the ICERs estimated using the company OS projections – including the ERG model amendments - should therefore be treated as ‘what if?’ scenarios as they are not underpinned by statistically significant clinical effectiveness evidence.

Second, even if the company's OS projection was accurate, the company has underestimated the acquisition costs of osimertinib and failed to take into account any administration cost of osimertinib as an oral chemotherapy – for the latter costs there are established NHS Reference Cost Schedules.[60] Using TTD data from the AURA studies and the IMPRESS trial and a cautious estimate of the NHS Reference Cost[60] for oral chemotherapy administration results in substantial increases in the size of the ICER per QALY gained from the company base case.

Third, utilities applied in the model appear to be implausibly high. Although drawn from AURA2 data, the EQ-5D-5L rather than EQ-5D-3L was used. Whilst no utility values are available specifically for the population described in the model, the ERG considered that there are alternative utility values that, whilst they are by no means perfect, could be used instead of the utility values used by the company.

The ERG did not identify a statistically significant difference in PFS and/or OS by line of treatment for osimertinib and did not consider the clinical effectiveness evidence on single-agent chemotherapy to be convincing. As such, the ERG does not consider the results of the company's subgroup analyses, by line of treatment or when osimertinib is compared to docetaxel, to be informative.

Application of the ERG changes to costs and the ERG's alternative utility values results in ICERs for osimertinib compared to PDC of [redacted] per QALY or [redacted] per QALY gained depending on the alternative utility values used in the model. If only the improvement in PFS with osimertinib is then included (PFS data being the only statistically significant effectiveness evidence included in the model) then the ERG estimates the ICER for osimertinib compared to PDC to be [redacted] per QALY gained or [redacted] per QALY gained, again depending on the alternative utility values used in the model.

7 END OF LIFE CRITERIA

The company puts forward the case (CS, Section 4.13) that osimertinib meets the NICE End of Life criteria.[50] These criteria are:

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

The company claims that osimertinib meets these criteria because:

- the available data from the IMPRESS trial (second-line treatment) suggest that patients previously treated with an EGFR-TKI have a median OS of less than 24 months
- the results from the company's economic model suggests that the mean OS associated with treatment with osimertinib is 13.5 months longer than OS associated with PDC treatment
- the number of patients eligible is 300 per year.

The ERG agrees with the company that available median OS data from the PDC arm of the IMPRESS trial indicate that patients with advanced NSCLC with T790M mutation and previously treated with an EGFR-TKI have a life expectancy of less than 24 months. The ERG also agrees that the eligible population is small. However, the ERG considers that, at this stage, there is no clinical effectiveness evidence to suggest that, when compared with PDC, treatment with osimertinib results in an OS gain of at least 3 months. Results of analyses carried out by the ERG show that when treatment with osimertinib is compared with PDC there is currently no clinical or statistically significant difference in OS. This may be, as stated by the company, due to the immaturity of the OS data, but at this stage any claim that treatment with osimertinib results in better OS than treatment with PDC is speculation.

8 OVERALL CONCLUSIONS

Clinical effectiveness data

The clinical effectiveness evidence presented by the company in support of the use of osimertinib for patients with EGFR T790M mutations who have previously failed EGFR-TKI treatment cannot be considered to be robust. There are no results available from RCTs that include osimertinib as an intervention or as a comparator, and the only on-going RCT of osimertinib (versus PDC) is not due to report until 2017. The company has submitted evidence for the clinical effectiveness of osimertinib from two ongoing phase I/II single-arm studies. Of the 411 patients recruited to these studies 31.4% had received osimertinib as second-line therapy and 68.6% as ≥third-line therapy. Data from AURAext and AURA2 are immature; the OS and PFS data from the pooled dataset are 12.7% and 38.9% mature, respectively. The immaturity of the PFS data also means that the safety profile of osimertinib should be viewed with caution. The EMA also noted the limitations of the company's data when they issued a conditional licence for osimertinib.

There are doubts about whether the clinical data used to demonstrate the relative effectiveness of osimertinib compared to treatment with PDC are robust. To create this dataset the company used data from the control (PDC) arm of the IMPRESS trial (a phase III RCT). Tumour samples from patients in this arm were tested (retrospectively) for the EGFR T790M mutation. The resultant dataset was small (n=61) and although the PFS data for this group are relatively mature (83.6%), the OS data are only 32.8% mature. The company should be commended for the effort taken to create a comparator dataset. They should also be commended for applying a methodology to adjust for different patient characteristics between trials. [redacted]

Lack of mature survival data has hindered the company's claims that treatment with osimertinib is more clinically effective, or more cost effective, than PDC in patients with T790M mutations who have failed EGFR-TKI treatment. Using the limited survival data available it is impossible for the company to put forward a robust argument in support of osimertinib using traditional methods of analysis (e.g. RCT results, indirect treatment comparisons or life-time economic evaluations). The ERG acknowledges the company's efforts to showcase the strengths of osimertinib. However, until more mature data are available the strengths and weaknesses associated with treatment with osimertinib will remain unclear.

Model OS gain

The ERG considers that the company's base case ICER for the comparison of osimertinib with PDC is implausible due to treatment with osimertinib being associated with an unsubstantiated OS gain. The immaturity of the data means that there is no clinical evidence on which to base this gain. Furthermore, analyses undertaken by the ERG demonstrate that there is no statistical evidence to support this OS gain. Therefore, any ICER that relies on this assumption is inherently flawed. When the only adjustment made to the company base case is to remove the OS gain, the ICER for the comparison of treatment with osimertinib versus PDC rises to [redacted] per QALY gained.

Utility values

The utility values used in the company's model were collected during the AURA2 study. However, the EQ-5D-5L index score at baseline is higher than that for the UK population of the same age (albeit that the latter were collected using the EQ-5D-3L tool). Furthermore, during the AURA2 study, utility was measured every 6 weeks whilst patients were receiving treatment (up to a maximum of 42 weeks) and, at all of these time points (except for week 42 when the questionnaire was only completed by four people), the index score was higher than at baseline. In addition, very little information is available from the AURA2 study about patient utility after treatment progression. Values are only available at the point of treatment discontinuation and at 28 days follow up. Values are also provided 'Post IP follow-up' but the timing of this event is not clear. The ERG, therefore, considers the utility values used in the company base case are problematic and offers two alternative sources. The ERG recognises that both alternatives also have limitations. However, they do provide alternative and less optimistic perspectives. As both sets of ERG preferred utility values are lower than those used in the company base case, their use leads to an increase in the size of the base case ICER.

EGFR T790M mutations

Currently, there is no routine EGFR T790M testing of tumours in NHS clinical practice. The organisational infrastructure may already be in place but, as for all new testing protocols, EGFR T790M testing of tumours requires careful NHS planning as well as the cooperation of NHS staff and patients.

There is a possibility of a tertiary acquired mutation identified after treatment with osimertinib for EGFR T790M mutation-positive NSCLC.[71] Osimertinib is one of a number of EGFR T790M targeted drugs (e.g., rociletinib by Clovis Oncology; BI 1482694 by Boehringer Ingelheim) and, as new clinical trial data are published and T790M testing becomes established in clinical practice, current gaps in efficacy data and safety profiles will be filled.

8.1 Implications for research

The protocols for AURAext and AURA2 studies permitted continuation of treatment after confirmed disease progression. Continuation of post-progression treatments appears to be becoming commonplace in oncology trials and the ERG suggests that it would be useful to record, and report, both PFS and TTD outcomes from RCTs and routine clinical practice.

9 REFERENCES

1. AstraZeneca UK Ltd. Osimertinib for locally advanced or metastatic EGFR and T790M mutation-positive NSCLC: Company submission to NICE 2016.
2. Royal College of Physicians. National Lung Cancer Audit Annual Report (audit period 2014). London: Royal College of Physicians 2015.
3. European Medicines Agency. Tagrisso: European Public Assessment Report (EPAR) 2015.
4. Kuiper J, Heideman D, Thunnissen E, Paul M, van Wijk A, Postmus P*, et al.* Incidence of T790M mutation in (sequential) rebiopsies in EGFR-mutated NSCLC patients. Lung Cancer. 2014; 85:19-24.
5. Oxnard GR, Arcila ME, Sima CS, Riely GJ, Chmielecki J, Kris MG*, et al.* Acquired resistance to EGFR tyrosine kinase inhibitors in EGFR-mutant lung cancer: Distinct natural history of patients with tumors harboring the T790M mutation. Clin Cancer Res. 2011; 17:1616-22.
6. Yu HA, Arcila ME, Rekhtman N, Sima CS, Zakowski MF, Pao W*, et al.* Analysis of tumor specimens at the time of acquired resistance to EGFR-TKI therapy in 155 patients with EGFR-mutant lung cancers. Clin Cancer Res. 2013; 19:2240-7.
7. Fukuoka M, Wu YL, Thongprasert S, Sunpaweravong P, Leong SS, Sriuranpong V*, et al.* Biomarker analyses and final overall survival results from a phase III, randomized, open-label, first-line study of gefitinib versus carboplatin/paclitaxel in clinically selected patients with advanced non-small-cell lung cancer in Asia (IPASS). J Clin Oncol. 2011; 29:2866-74.
8. Han JY, Park K, Kim SW, Lee DH, Kim HY, Kim HT*, et al.* First-SIGNAL: First-line single-agent iressa versus gemcitabine and cisplatin trial in never-smokers with adenocarcinoma of the lung. J Clin Oncol. 2012; 30:1122-8.
9. Maemondo M, Inoue A, Kobayashi K, Sugawara S, Oizumi S, Isobe H*, et al.* Gefitinib or chemotherapy for non-small-cell lung cancer with mutated EGFR. N Engl J Med. 2010; 362:2380-8.
10. Mitsudomi T, Morita S, Yatabe Y, Negoro S, Okamoto I, Tsurutani J*, et al.* Gefitinib versus cisplatin plus docetaxel in patients with non-small-cell lung cancer harbouring mutations of the epidermal growth factor receptor (WJTOG3405): An open label, randomised phase 3 trial. Lancet Oncol. 2010; 11:121-8.
11. Rosell R, Carcereny E, Gervais R, Vergnenegre A, Massuti B, Felip E*, et al.* Erlotinib versus standard chemotherapy as first-line treatment for European patients with advanced EGFR mutation-positive non-small-cell lung cancer (EURTAC): A multicentre, open-label, randomised phase 3 trial. Lancet Oncology. 2012; 13:239-46.
12. Zhou C, Wu YL, Chen G, Feng J, Liu XQ, Wang C*, et al.* Final overall survival results from a randomised, phase III study of erlotinib versus chemotherapy as first-line treatment of EGFR mutation-positive advanced non-small-cell lung cancer (OPTIMAL, CTONG-0802). Ann Oncol. 2015; 26:1877-83.
13. Soria JC, Wu YL, Nakagawa K, Kim SW, Yang JJ, Ahn MJ*, et al.* Gefitinib plus chemotherapy versus placebo plus chemotherapy in EGFR-mutation-positive non-small-cell lung cancer after progression on first-line gefitinib (IMPRESS): A phase 3 randomised trial. Lancet Oncol. 2015; 16:990-8.
14. Inoue A, Kobayashi K, Maemondo M, Sugawara S, Oizumi S, Isobe H*, et al.* Updated overall survival results from a randomized phase III trial comparing gefitinib with carboplatin-paclitaxel for chemo-naive non-small cell lung cancer with sensitive EGFR gene mutations (NEJ002). Ann Oncol. 2013; 24:54-9.
15. Mok TS, Wu YL, Thongprasert S, Yang CH, Chu DT, Saijo N*, et al.* Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. N Engl J Med. 2009; 361:947-57.
16. Yang JCH, Wu YL, Schuler M, Sebastian M, Popat S, Yamamoto N*, et al.* Afatinib versus cisplatin-based chemotherapy for EGFR mutation-positive lung

adenocarcinoma (LUX-Lung 3 and LUX-Lung 6): analysis of overall survival data from two randomised, phase 3 trials. Lancet Oncology. 2015; 16:141-51.
17. Shigematsu H, Lin L, Takahashi T, Nomura M, Suzuki M, Wistuba II, *et al.* Clinical and biological features associated with epidermal growth factor receptor gene mutations in lung cancers. Journal of the National Cancer Institute. 2005; 97:339-46.
18. National Institute for Health and Care Excellence. TA192: Gefitinib for the first-line treatment of locally advanced or metastatic NSCLC. NICE; 2010; Available from: https://www.nice.org.uk/guidance/ta192/resources/gefitinib-for-the-firstline-treatment-of-locally-advanced-or-metastatic-nonsmallcell-lung-cancer-82598553248965 [accessed March 2016].
19. National Institute for Health and Care Excellence. TA258: Erlotinib for the first-line treatment of locally advanced or metastatic EGFR-TK mutation-positive NSCLC. NICE; 2012; Available from: https://www.nice.org.uk/guidance/ta258/resources/erlotinib-for-the-firstline-treatment-of-locally-advanced-or-metastatic-egfrtk-mutationpositive-nonsmallcell-lung-cancer-82600498244293 [accessed March 2016].
20. National Institute for Health and Care Excellence (NICE). TA310: Afatinib for treating epidermal growth factor receptor mutation-positive locally advanced or metastatic non-small-cell lung cancer. 2014; Available from: https://www.nice.org.uk/guidance/ta310 [accessed April 2016].
21. eMC (Electronic Medicines Consortium). TAGRISSO 40mg and 80mg film-coated tablets. 2016; Available from: https://www.medicines.org.uk/emc/medicine/31496 [accessed March 2016].
22. Medicines and Healthcare Products Regulatory Agency. Osimertinib assessment report 2015.
23. AstraZeneca UK Ltd. Clinical Study Report: AURAext 2016.
24. AstraZeneca UK Ltd. Clinical Study Report: AURA2 2016.
25. European Medicines Agency. PRIME: Priority medicines. 2016; Available from: http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general_content_000660.jsp&mid=WC0b01ac058096f643.
26. Pao W, Miller VA, Politi KA, Riely GJ, Somwar R, Zakowski MF, *et al.* Acquired resistance of lung adenocarcinomas to gefitinib or erlotinib is associated with a second mutation in the EGFR kinase domain. PLoS Med. 2005; 2:e73.
27. National Institute for Health and Care Excellence. Osimertinib for locally advanced or metastatic, EGFR and T790M mutation positive non-small-cell lung cancer: Final scope. 2016.
28. European Medicines Agency. Tagrisso (osimertinib). 2016; Available from: http://www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/human/medicines/004124/human_med_001961.jsp&mid=WC0b01ac058001d124 [accessed April 2016].
29. Park S, Keam B, Kim SH, Kim KH, Kim YJ, Kim JS, *et al.* Pemetrexed singlet versus non pemetrexed-based platinum doublet as second-line chemotherapy after first-line epidermal growth factor receptor (EGFR) tyrosine kinase inhibitor failure in non-small cell lung cancer patients with EGFR mutations. Cancer Res Treat. 2015; 47:630-7.
30. Schuler M, Yang JC, Park K, Kim JH, Bennouna J, Chen YM, *et al.* Afatinib beyond progression in patients with non-small-cell lung cancer following chemotherapy, erlotinib/gefitinib and afatinib: Phase III randomized LUX-Lung 5 trial. Ann Oncol. 2016; 27:417-23.
31. National Institute for Health and Care Excellence (NICE). Single technology appraisal: User guide for company evidence submission template. 2015; Available from: https://www.nice.org.uk/article/pmg24/resources/non-guidance-single-technology-appraisal-user-guide-for-company-evidence-submission-template-pdf. [accessed April 2016].
32. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. J Epidemiol Community Health. 1998; 52:377-84.

33. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, *et al.* Assessing the quality of reports of randomized clinical trials: Is blinding necessary? Control Clin Trials. 1996; 17:1-12.
34. Centre for Reviews and Dissemination. Systematic Reviews: CRD's guidance for undertaking reviews in health care. York: University of York 2009.
35. Halmos B, Pennell NA, Fu P, Saad S, Gadgeel S, Otterson GA, *et al.* Randomized phase II trial of erlotinib beyond progression in advanced erlotinib-responsive non-small cell lung cancer. Oncologist. 2015; 20:1298-303.
36. Kasahara K, Kiura K, Nogami N, Takayama K, Takiguchi Y, Hirashima T, *et al.* Phase 2: Docetaxel with or without ramucirumab as therapy for non-small cell lung cancer (NSCLC) with Epidermal Growth Factor Receptor (EGFR) activating mutations after prior EGFR tyrosine kinase inhibitor (TKI) and platinum-based chemotherapy. European Cancer Congress; Vienna 2015.
37. Kim KS, Kim YC, Oh IJ, Na KJ, Ahn SJ, Choi YD, *et al.* The efficacy and safety of pemetrexed after the failure of gefitinib in patients with adenocarcinoma with EGFR mutation. J Thorac Oncol. 2013; 8:S1174-S.
38. Miyauchi E, Inoue A, Kobayashi K, Maemondo M, Sugawara S, Oizumi S, *et al.* Efficacy of chemotherapy after first-line gefitinib therapy in EGFR mutation-positive advanced non-small cell lung cancer-data from a randomized Phase III study comparing gefitinib with carboplatin plus paclitaxel (NEJ002). Jpn J Clin Oncol. 2015; 45:670-6.
39. Tseng JS, Yang TY, Chen KC, Hsu KH, Yu CJ, Liao WY, *et al.* Prior EGFR tyrosine-kinase inhibitor therapy did not influence the efficacy of subsequent pemetrexed plus platinum in advanced chemonaive patients with EGFR-mutant lung adenocarcinoma. Onco Targets Ther. 2014; 7:799-805.
40. Wu JY, Shih JY, Yang CH, Chen KY, Ho CC, Yu CJ, *et al.* Second-line treatments after first-line gefitinib therapy in advanced nonsmall cell lung cancer. Int J Cancer. 2010; 126:247-55.
41. Zhou CZ, Qin YY, Xie ZH, Zhang JX, Ou-Yang M, Li SY, *et al.* Efficacy of third-line pemetrexed monotherapy versus pemetrexed combination with bevacizumab in patients with advanced EGFR mutation-positive lung adenocarcinoma. Chin J Cancer Res. 2014; 26:705-10.
42. AstraZeneca Ltd. Clinical Study Report: IMPRESS. 2014.
43. AstraZeneca UK Ltd. Statistical Analysis Plan: IMPRESS. 2014.
44. AstraZeneca UK Ltd. Statistical Analysis Plan: AURAext 2015.
45. AstraZeneca UK Ltd. Statistical Analysis Plan: AURA2 2015.
46. European Organisation for Research and Treatment of Cancer Quality of Life Group. EORTC QLQ-C30 QoL questionnaire. 2015; Available from: http://groups.eortc.be/qol/eortc-qlq-c30 [accessed January 2016].
47. European Organisation for Research and Treatment of Cancer Quality of Life Group (EORTC). EORTC QLQ - LC13 1994; Available from: http://groups.eortc.be/qol/sites/default/files/img/specimen_lc13_english.pdf.
48. The EuroQoL Group. EQ-5D-3L. EuroQoL; 2016; Available from: http://www.euroqol.org/about-eq-5d/how-to-use-eq-5d.html [accessed January 2016].
49. FACIT. FACT-L. 2016; Available from: http://www.facit.org/FACITOrg/Questionnaires [accessed April 2016].
50. National Institute for Health and Care Excellence (NICE). Guide to the methods of technology appraisal 2013. 2013; Available from: http://www.nice.org.uk/article/pmg9/resources/non-guidance-guide-to-the-methods-of-technology-appraisal-2013-pdf (Accessed 01/06/2015).
51. Carlson JJ, Garrison LP, Ramsey SD, Veenstra DL. The potential clinical and economic outcomes of pharmacogenomic approaches to EGFR-Tyrosine Kinase Inhibitor therapy in non–small cell lung cancer. Value in Health. 2009; 12:20-7.
52. Chouaid C, Le Caer H, Corre R, Crequit J, Locher C, Falchero L, *et al.* Cost analysis of erlotinib versus chemotherapy for first-line treatment of non-small-cell lung cancer

in frail elderly patients participating in a prospective phase 2 study (GFPC 0505). Clin Lung Cancer. 2013; 14:103-7.
53. Chouaid C, Le Caer H, Locher C, Dujon C, Thomas P, Auliac JB*, et al.* Cost effectivenes of erlotinib versus chemotherapy for first-line treatment of non small cell lung cancer (NSCLC) in fit elderly patients participating in a prospective phase 2 study (GFPC 0504). BMC Cancer. 2012; 12:301.
54. Horgan AM, Bradbury PA, Amir E, Ng R, Douillard JY, Kim ES*, et al.* An economic analysis of the INTEREST trial, a randomized trial of docetaxel versus gefitinib as second-/third-line therapy in advanced non-small-cell lung cancer. Ann Oncol. 2011; 22:1805-11.
55. Patel K, Montouchet C, Cheynel J, Ruff L. A cost-effectiveness analysis of EGFR-TK mutation status-guided 1st- and 2nd-line treatment of stage III/IV non-small cell lung cancer in the UK. ISPOR; 8-12 November Amsterdam 2014.
56. National Institute for Health and Care Excellence. TA347: Nintedanib for previously treated locally advanced, metastatic, or locally recurrent non‐small cell lung cancer. NICE; 2015; Available from: https://www.nice.org.uk/guidance/ta347/chapter/1-guidance [accessed March 2016].
57. Decision Support Unit. A review of studies examining the relationship between progression-free survival and overall survival in advanced or metastatic cancer: School of Health and Related Research, University of Sheffield 2012.
58. EuroQoL. EQ-5D-5L Value Sets. 2013; Available from: http://www.euroqol.org/about-eq-5d/valuation-of-eq-5d/eq-5d-5l-value-sets.html [accessed April 2016].
59. 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.
60. Department of Health. NHS Reference Costs 2014 to 2015. 2015; Available from: https://www.gov.uk/government/publications/nhs-reference-costs-2014-to-2015.
61. National Institute for Health and Care Excellence (NICE). TA319: Ipilimumab for previously untreated advanced (unresectable or metastatic) melanoma. 2014; Available from: https://www.nice.org.uk/guidance/ta319 [accessed April 2016].
62. National Institute for Health and Care Excellence (NICE). DG9: EGFR-TK mutation testing in adults with locally advanced or metastatic non-small-cell lung cancer. 2013; Available from: https://www.nice.org.uk/guidance/dg9/chapter/the-diagnostic-tests.
63. Department of Health. Drugs and pharmaceutical electronic market information (eMit). 2011; Available from: https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit [accessed April 2016]
64. Kind P, Hardman G, Macran S. UK population norms for EQ-5D. Centre for Health Economics, University of York. 1999.
65. Reck M, Kaiser R, Mellemgaard A, Douillard JY, Orlov S, Krzakowski M*, et al.* Docetaxel plus nintedanib versus docetaxel plus placebo in patients with previously treated non-small-cell lung cancer (LUME-Lung 1): A phase 3, double-blind, randomised controlled trial. Lancet Oncol. 2014; 15:143-55.
66. Sacco JJ, Botten J, Macbeth F, Bagust A, Clark P. The average body surface area of adult cancer patients in the UK: a multicentre retrospective study. PLoS One. 2010; 5:e8933.
67. National Institute for Health and Care Excellence. TA366: Pembrolizumab for advanced melanoma not previously treated with ipilimumab. NICE; 2015; Available from: https://www.nice.org.uk/guidance/ta366 [accessed April 2016].
68. Fleeman N, Bagust A, Boland A, Dwan K, Richardson M, Beale S*, et al.* Nintedanib for previously treated locally advanced or metastatic non small cell lung cancer: A Single Technology Appraisal: LRiG, University of Liverpool 2014.
69. Greenhalgh J, Mahon J, Richardson M, Krishan A, Aslam RW, Beale S*, et al.* Pembrolizumab for treating advanced melanoma previously untreated with ipilimumab: A Single Technology Appraisal: LRiG, University of Liverpool 2015.

70. U.S.Department of Health and Human Services. Common Terminology Criteria for Adverse Events v4.0 (CTCAE) 2010; Available from: http://evs.nci.nih.gov/ftp1/CTCAE/CTCAE_4.03_2010-06-14_QuickReference_5x7.pdf [accessed April 2016}.
71. Yu Z, Boggon TJ, Kobayashi S, Jin C, Ma PC, Dowlati A*, et al.* Resistance to an irreversible epidermal growth factor receptor (EGFR) inhibitor in EGFR-mutant lung cancer reveals novel treatment strategies. Cancer Research. 2007; 67:10417-27.

10 APPENDICES

Appendix 1: ERG changes to submitted company model

ERG Section 6 results table revision	Associated detail	Implementation instructions
R1. TTD data for on treatment costs	OSI_TTD.xlsx. These changes generate alternative costs in the model. QALY changes resulting from these changes should be ignored.	For osimertinib ***In Workbook OSI_TTD.xls*** In Sheet 'Values' Copy range **O4:O783** ***In company model*** In Sheet 'PatFlow_B' Paste values to cells **G13:G792** For PDC ***In Workbook OSI_TTD.xls*** In Sheet 'Values' Copy range **P4:P783** ***In company model*** In Sheet 'PatFlow_B' Paste values to cells **O13:O792**
R2. Calculation of osimertinib administration cost	OSI_TTD.xlsx	Calculated by ERG. Workings can be found in OSI_TTD.xlsx in Sheet 'Values' **Column E**
R3. LUME-Lung 1 utility		In Sheet 'CountryData' Set value in cell I679 = 0.687 Set value in cell I680 = 0.640
R4. Nafees utility		In Sheet 'CountryData' Set value in cell I679 = 0.653 Set value in cell I680 = 0.470
R5. Osimertinib generates a gain in PFS but not OS		In Sheet 'PatFlow_B' Copy cells **Q13:Q792**

ERG Section 6 results table revision	Associated detail	Implementation instructions
compared to PDC		Paste values in range **I13:I792** Enter formula in cell **H13 '=1-G13-I13'** Copy formula in cell **H13** to **H14:H792**