Erlotinib monotherapy for the maintenance treatment of non-small cell lung cancer after previous platinum-containing therapy

ERG Report

Contains commercially in confidence data

Produced by:

Liverpool Reviews and Implementation Group (LRiG)
University of Liverpool
Room 2.12
Whelan Building
The Quadrangle
Brownlow Hill
Liverpool
L69 3GB

Tel: +44 (0) 151 794 5682/5541/5067
Fax: +44 (0) 151 794 5821
Email: LRiG@liv.ac.uk

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 08/200/01

Date completed: March 23, 2010

Authors: Adrian Bagust, Professor of Modelling in Health, Liverpool Reviews and Implementation Group, University of Liverpool

Angela Boland, Associate Director, Liverpool Reviews and Implementation Group, University of Liverpool

Michaela Blundell, Research Fellow (Medical Statistician) Liverpool Reviews and Implementation Group, University of Liverpool

Tamara Brown, Research Fellow (Clinical Effectiveness) Liverpool Reviews and Implementation Group, University of Liverpool

Helen Davis, Assistant Director, North West Medicines Information Centre, Pharmacy Practice Unit, Liverpool

Rumona Dickson, Director, Liverpool Reviews and Implementation Group, University of Liverpool

Yenal Dundar, Research Fellow (Clinical Effectiveness), Liverpool Reviews and Implementation Group, University of Liverpool

Juliet Hockenhull, Research Fellow (Clinical Effectiveness) Liverpool Reviews and Implementation Group, University of Liverpool

Carlos Martin Saborido, Research Fellow (Health Economics), Liverpool Reviews and Implementation Group, University of Liverpool

James Oyee, Research Fellow (Medical Statistician), Liverpool Reviews and Implementation Group, University of Liverpool

Vidhya Sagar Ramani, Consultant Clinical Oncologist, Clatterbridge Centre for Oncology, Bebington, Wirral, Merseyside

Correspondence: Ms. Rumona Dickson, Director (LRiG), Liverpool Reviews and Implementation Group, University of Liverpool, Room 2.12, Whelan Building, The Quadrangle, Brownlow Hill, Liverpool L69 3GB

Declared competing interests of the authors: None

Acknowledgements: Thanks to Professor Mike Pearson (University of Liverpool) for his comments on the MS, provision of clinical input and feedback on the final report.

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

This report should be referenced as follows:

Dickson R, Bagust A, Boland A, Blundell M, Brown T, Davis H, Dundar Y, Hockenhull J, Martin Saborido C, Oyee J, and Ramani, VS. Erlotinib monotherapy for the maintenance treatment of non-small cell lung cancer (NSCLC): A Single Technology Appraisal. LRiG, The University of Liverpool, 2009

Contributions of authors:

Rumona Dickson	Served as project lead, drafted the clinical results section and supervised the production of the final report
Adrian Bagust	Critical appraisal of the economic model
Angela Boland	Critical appraisal of the economic evidence
Michaela Blundell	Critical appraisal of clinical statistical approach
Tamara Brown	Critical appraisal of the clinical evidence
Helen Davis	Critical appraisal of the manufacturer's submission
Yenal Dundar	Cross checking of manufacturer's search strategies
Juliet Hockenhull	Summary and critical appraisal of the clinical evidence
Carlos Martin Saborido	Critical appraisal of economic model and compilation of economic critique
James Oyee	Critical appraisal of clinical statistical approach
Vidhya Sagar Ramini	Critical appraisal of the clinical sections of the manufacturer's submission

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

Table of Contents

Abbreviations

AE(s)	Adverse event(s)
BNF	British National Formulary
BSA	Body surface area
BSC	Best supportive care
CEAC	Cost-effectiveness acceptability curve
CHMP	Committee for Medicinal Products for Human Use
CI	Confidence interval
CSR	Clinical study report
ECOG	Eastern Cooperative Oncology Group
EMA	European Medicines Agency
EQ-5D	EuroQol 5D (a standardised instrument used as a measure of health outcome)
ERG	Evidence Review Group
EGFR	Epidermal growth factor receptor
FACT-L	Functional Assessment of Cancer Therapy-Lung
FACT-G	Functional Assessment of Cancer Therapy-General
FAS	Full analysis set
FDA	Food and Drug Administration
HR	Hazard ratio
HRQoL	Health-related quality of life
ICER	Incremental cost-effectiveness ratio
IHC	Immunohistochemistry
IPD	Individual patient data
ITT	Intention to treat
LUCADA	Lung Cancer Data
LYG	Life year gained
MS	Manufacturer submission
NICE	National Institute for Health and Clinical Excellence
NSCLC	Non-small cell lung cancer
NS	Non-squamous
OS	Overall survival
PFS	Progression-free survival
PPS	Post-progression survival
PS	Performance status
PSA	Probabilistic sensitivity analysis
PSS	Personal Social Services
QALY	Quality adjusted life year
QoL	Quality of life
RCT	Randomised controlled trial
SA	Sensitivity analysis
SD	Stable disease
SPC	Summary of Product Characteristics
STA	Single Technology Appraisal
vs	versus
VAS	Visual analogue scores
WTP	Willingness to pay

Licence announcement – March 19, 2010

This document was essentially complete when notification was received on March 19, 2010 that the Committee for Medicinal Products Use (CHMP) announced a positive opinion relevant to this appraisal:[1]

> *'Tarceva is indicated as monotherapy for maintenance treatment in patients with locally advanced or metastatic non-small cell lung cancer with stable disease after 4 cycles of standard platinum-based first-line chemotherapy.'*

We have not substantively changed the contents of the report as it includes analysis related to this specific patient group. We have however altered the discussion and research recommendations related to this announcement.

1 SUMMARY

1.1 Scope of the submission

The remit of the Evidence Review Group (ERG) is to comment on the clinical and cost-effectiveness evidence submitted to the National Institute for Health and Clinical Excellence (NICE) as part of the single technology appraisal (STA) process. Clinical and economic evidence has been submitted to NICE from Roche Ltd in support of the use of erlotinib (Tarceva®) as a treatment for patients with non-small cell lung cancer (NSCLC). The manufacturer submission (MS) describes the use of erlotinib as a maintenance therapy for patients whose disease has not progressed following the completion of four cycles of first-line platinum-based chemotherapy.

This report has been compiled prior to the announcement of two important decisions that may impact on the conclusions of this report. The first is related to the extension of the European licence for this technology. In December 2009 the Food and Drug Administration (FDA) Advisory Committee failed to approve the use of erlotinib as a maintenance therapy for patients with NSCLC. This decision is currently under review.[2] The second is the outcome of the NICE[3] appraisal of the use of pemetrexed as a maintenance treatment for patients with NSCLC. As such the ERG report includes an element of speculation.

1.2 Summary of submitted clinical-effectiveness evidence

The systematic review of the literature conducted by the manufacturer appropriately identified and described one relevant randomised controlled trial (RCT) known as the SATURN[4] trial; the clinical evidence presented in the MS is primarily derived from this double-blind, placebo-controlled RCT. The trial compared the use of erlotinib plus best supportive care (BSC) as maintenance therapy vs placebo plus BSC as maintenance therapy in patients with NSCLC who had received four cycles of platinum based-chemotherapy as a first-line treatment and whose disease had not progressed.

The MS presents the clinical outcomes of a range of patient populations including: patients in the full analysis set (FAS), also known as the intention to treat (ITT) group (n=889), and *post hoc* analyses of patients with stable disease (SD) (n=487) and patients with non-squamous histology (n=529). Using results from the unstratified analyses, in these three groups the trial demonstrated greater median progression free survival (PFS) for patients in the erlotinib arm compared with patients in the placebo arm (ITT: 12.3 weeks vs 11.1 weeks; HR 0.71; 95% CI 0.62, 0.82; p<0.0001), (SD: HR 0.68; 95% CI 0.56, 0.83; p<0.0001) and (non-squamous: HR 0.68; 95% CI 0.56, 0.82). Median overall survival (OS) was also greater for the erlotinib treated patients (ITT: 12 months vs 11 months; HR 0.81; 95% CI 0.70, 0.95), (SD: HR 0.72; 95% CI 0.59, 0.89) and (non-squamous: 13.7 months vs 10.5 months;

HR 0.79; 95% CI 0.64, 0.96; p=0.019). Using the Functional Assessment of Chronic Illness Therapy-Lung (FACT-L) questionnaire the manufacturer shows that there were no differences between the groups with respect to time to symptom progression (HR 0.91; p=0.379); time to deterioration in trial outcome index (TOI) (HR 1.06; p=0.538); or time to deterioration in quality of life (QoL) (HR 0.96; p=0.653).

As pemetrexed is also under consideration by NICE as maintenance therapy for patients with NSCLC, the manufacturer appropriately carried out an indirect comparison of pemetrexed vs erlotinib using data from the JMEN[5] trial. The indirect comparison shows that pemetrexed vs erlotinib in patients with non-squamous histology yields a statistically significant PFS benefit for patients on pemetrexed compared with erlotinib [redacted] there is no statistically significant benefit shown for pemetrexed compared with erlotinib in patients with non-squamous histology in terms of OS [redacted]

1.3 Summary of submitted cost-effectiveness evidence

The manufacturer did not identify any published cost-effectiveness analyses of erlotinib for the maintenance treatment of patients with NSCLC. The manufacturer therefore developed a *de novo* economic model to support their economic evaluation. In the MS, the economic evaluation was tailored to consider three specific patient populations: intention to treat (ITT), stable disease (SD) and non-squamous populations. The economic evaluations for the ITT and SD patient populations compare erlotinib plus BSC vs BSC; the economic evaluation for the non-squamous population compares pemetrexed with erlotinib.

For direct comparisons of erlotinib plus BSC vs BSC, the clinical data used in the MS economic model were primarily generated from the SATURN[4] trial. For the indirect comparison of pemetrexed vs erlotinib, data were derived from the JMEN[5] trial of pemetrexed maintenance treatment (pemetrexed plus BSC versus placebo plus BSC) in patients with NSCLC. The manufacturer's economic evaluation adopts a lifetime horizon (five years) for the consideration of costs and benefits and the perspective is that of the UK NHS and Personal Social Services (PSS). The manufacturer estimates incremental cost-effectiveness ratios (ICERs) and presents incremental cost per quality adjusted life year gained (QALYs) and incremental cost per life year gained for all comparisons. The manufacturer conducted both sensitivity analysis (SA) and probabilistic sensitivity analysis (PSA).

The ICERs estimated by the manufacturer comparing erlotinib plus BSC vs BSC are as follows: £55,219 per QALY gained (ITT population) and £47,743 per QALY gained (SD population). The ICER estimated by the manufacturer comparing pemetrexed vs erlotinib is [redacted] per QALY gained (non-squamous disease).

The manufacturer has presented a case for erlotinib to be considered under the NICE end of life criteria[6] as a maintenance treatment for patients with NSCLC and (i) stable disease that has not progressed after first-line chemotherapy and (ii) non-squamous histology who have not progressed after first-line chemotherapy.

1.4 Commentary on the robustness of submitted evidence

1.4.1 Strengths

The manufacturer provides evidence from a well-designed trial (SATURN[4]) of the clinical benefit of erlotinib plus BSC vs placebo plus BSC. The trial recruited a substantial number of patients in a difficult disease area. It is noteworthy that investigators' assessments of PFS outcomes were independently verified. The manufacturer also appropriately carries out an indirect comparison of pemetrexed vs erlotinib in patients with non-squamous histology.

Substantial additional data and analyses provided by the manufacturer to the ERG via the clarification process were also used in the compilation of the ERG report.

1.4.2 Weaknesses

Clinical

The ERG notes that there is only one relevant RCT (SATURN[4]). Despite being generally well-designed, the trial exhibits several weaknesses. Of note is that, despite efforts to ensure blinding of patients and investigators, as patients in the erlotinib arm were significantly more likely to develop a rash and suffer from diarrhoea than patients in the placebo group, the extent to which patients and investigators were truly blind to treatment allocation throughout the trial is uncertain.

The randomisation technique used in the trial included stratification by six different factors. Neither histology (e.g. non-squamous NSCLC) nor response to treatment (e.g. stable disease) was employed as a stratification factor. However, the manufacturer was particularly interested in these two subgroups of patients and conducted several *post-hoc* analyses using outcomes from the SATURN[4] trial. The ERG considers that the results of the *post-hoc* analyses should be considered with caution as the trial was not designed to perform this type of analysis and did not adjust for multiple testing.

Economics

The main weakness of the economic evidence presented is directly related to the type of model built by the manufacturer. The submitted models are structured around two health states (before and after disease progression), and are presented in the form of a Markov structure. However, it is important to

recognise that the models are not in fact Markov models since they are not governed by transition probabilities which impose restrictions on the way patients move from state to state. The models developed by the manufacturer allow negative post-progression survival (PPS) values to appear in later cycles. The ERG is therefore concerned about the reliability of the results generated by the submitted models.

The ERG has identified at least eight key areas where corrections and/or adjustments to the economic models are required related to: time horizon, discounting logic, cost of erlotinib, cost of second-line chemotherapy, unit costs, utility values, PFS model and OS model. Taken together, these corrections and/or adjustments have increased the size of the ICER for the ITT and SD populations (erlotinib vs placebo) and reduced the size of the ICER for the non-squamous population (pemetrexed vs erlotinib).

1.4.3 Areas of uncertainty

The generalisability of the SATURN[4] trial to patients in England and Wales is uncertain for a number of reasons:

- Only seven patients were recruited from the UK, 75% of patients in the trial were recruited from outside of Western Europe
- The trial did not include patients who had received pemetrexed as a first-line treatment (according to the MS pemetrexed is becoming the dominant first-line treatment for patients with non-squamous NSCLC); hence the response of patients to erlotinib after treatment with pemetrexed is unknown
- Paclitaxel appears to be used as a first-line treatment for a greater proportion of patients in the trial than might otherwise be the case in clinical practice in England and Wales.[7] The impact of this when generalising to patients in England and Wales is unknown
- A number of patients in the trial received second-line therapies that are not available to patients in clinical practice in England and Wales; this may affect the magnitude of the OS benefit observed in the trial.

The role of erlotinib in the treatment pathway is also uncertain and partially dependent on the result of the current NICE appraisal[3] of pemetrexed as a maintenance therapy. If both erlotinib and pemetrexed are approved as maintenance therapies by NICE then, depending on the eligible populations defined in the guidance issued by NICE, erlotinib and pemetrexed may or may not be recommended as treatment alternatives for the same group of patients; a decision by NICE regarding the outcome of the pemetrexed appraisal[3] is expected the week commencing March 22th 2010.

1.5 End of life

The ERG is of the opinion that the manufacturer has met the clinical end of life criteria[6] set out by NICE for consideration of erlotinib as a maintenance therapy for patients with stable disease who have not progressed after first-line chemotherapy.

However, the ERG is of the opinion that, based on the OS benefit demonstrated in the SATURN[4] trial, the manufacturer has not met the clincial end of life criteria[6] set out by NICE for consideration of erlotinib as a maintenance therapy for patients with non-squamous histology who have not progressed after first-line chemotherapy; the manufacturer did not demonstrate an end of life extension of greater than three months for this group of patients.

1.6 Key issues

Of primary importance to this appraisal is that the EMA has not yet approved erlotinib for use as maintenance therapy. It is unclear from the MS whether the manufacturer has applied for a licence for (i) all patients with NSCLC who have had four cycles of first-line chemotherapy and whose disease has not progressed or (ii) a subgroup of these patients (e.g. patients with stable disease or patients with non-squamous histology). If, for example, the licence is granted only for treatment of patients with non-squamous disease then a confirmed histological diagnosis will be required before patients can be offered treatment. Whilst histological testing is routinely carried out in many centres in England and Wales, this will not be available to all patients. The outcome of the EMA's deliberations on erlotinib as maintenance therapy is currently unknown.

Another key issue is that none of the patients in the SATURN[4] trial received pemetrexed as a first-line treatment. In the UK, pemetrexed is recommended as a first-line treatment by NICE[8] for patients with "other than squamous cell carcinoma". Currently, there is no clinical evidence to support the use of erlotinib as a maintenance therapy in patients who received pemetrexed as a first-line therapy.

The ERG believes that the economic model(s) submitted by the manufacturer is flawed and is therefore concerned about the validity of the results presented in the MS. The ERG offers a solution to the structural problems identified in the manufacturer's model; however, the ERG's revised ICERs are estimated to be approximately £63,000 per QALY gained (erlotinib vs placebo) for the ITT population, £60,000 per QALY gained for the SD population (erlotinib vs placebo) and [redacted] per QALY gained for patients with non-squamous histology (erlotinib vs pemetrexed).

2 BACKGROUND

2.1 Critique of manufacturer's description of underlying health problem

The MS provides an overview of the clinical problem, including epidemiology, first-line therapy provision and its limitations and offers an introduction to the use of maintenance treatment in patients with NSCLC who have not progressed after treatment with first-line chemotherapy.

The MS states that approximately 33,500 new cases of lung cancer occur each year in England and Wales.[9] Figures from the NHS Information Centre from the National Lung Cancer Audit for 2007[10] indicate that approximately 14% of reported cases were small cell cancer or mesothelioma. Of the remaining 29,000+ other reported cases only 10,500 (approximately 55%) were histologically confirmed as NSCLC.Currently histological confirmation of lung cancer is made only in 72% of patients in the UK with variation across centres (range 25%-85%).[11] In addition, a significant proportion of these cancers are typed as NSCLC without specifically being sub-classified (e.g. non-squamous, adenocarcinoma). Recent NICE guidance[8] for the first-line treatment of NSCLC now calls for histological testing and therefore histological testing rates are expected to increase.

Diagnosis of the disease tends to be late with at least 40% of patients being confirmed as having stage IIIB or IV disease.[10] Survival rates are therefore poor with survival of less than 30% of patients at one year and less than 10% at five years.[12]

2.2 Overview of current service provision

The MS provides a summary of current service provision for patients with newly diagnosed lung cancer with data provided from a variety of published sources and clinical experts. Data from national audits indicated that targets for patient treatment have not been reached with only 48% of eligible patients receiving first-line chemotherapy.[11] There were internal inconsistencies in the MS related to the proportion and number of patients who might be eligible for maintenance treatment after first-line chemotherapy. In their clarification response, the manufacturer appropriately points out the lack of definitive data related to the second-line treatment of NSCLC in England and Wales. Given the market research conducted by the manufacturer and the ERG's consultation with those responsible for the National Lung Cancer Audit, it seems reasonable to accept that 28% of patients with NSCLC are currently eligible for second-line treatment as described in the MS. An edited version of the proposed future treatment pathway with approximated patient numbers described in the MS (pg 22) is shown in Figure 2-1.

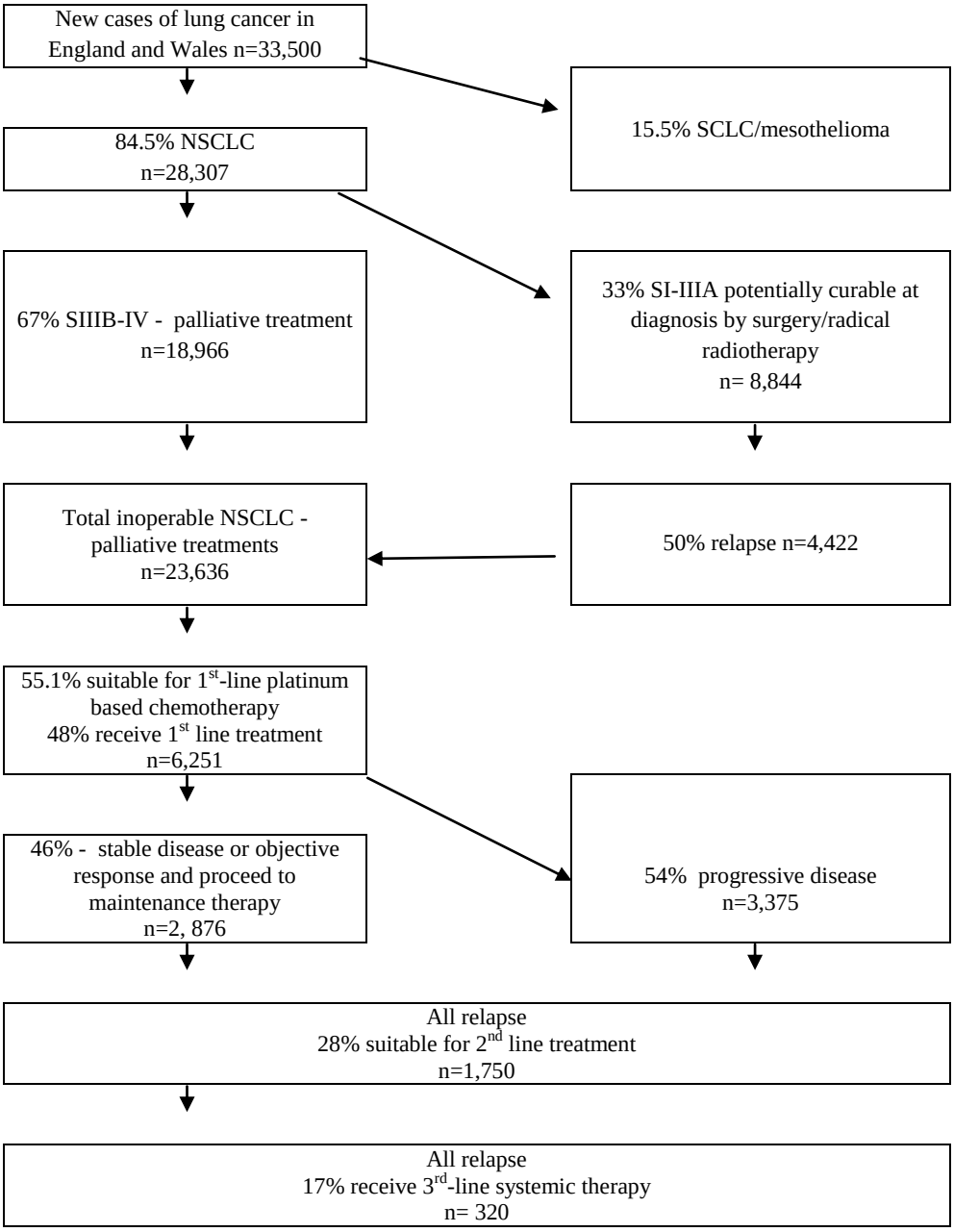

Figure 2-1 Proposed future NSCLC treatment pathway including maintenance therapy

The MS outlines clearly the historical context and limitations of current first-line chemotherapy options for patients with NSCLC. The history and rationale for the development of maintenance therapy in this patient population as well as the future position of maintenance treatment in the care pathway are described by the manufacturer. The MS makes the points that the use of gemcitabine, docetaxel or vinorelbine for maintenance therapy is not justified because these drugs have not 'shown sufficient activity to offset their toxicity'.[13-16]

The MS appropriately discusses the use of pemetrexed as a maintenance therapy for patients with NSCLC who have non-squamous histology. At the time of writing, the outcome of the NICE appraisal regarding the use of pemetrexed as a maintenance treatment for patients with NSCLC is not completed. Published timelines for this appraisal indicate that the Final Appraisal Document will be sent to consultees in the week commencing March 22nd, 2010.

The MS states that since the release of NICE guidance[8] approving the use of pemetrexed as a first-line treatment for patients with NSCLC who have non-squamous histology, the number of patients receiving pemetrexed in England and Wales is increasing. The MS considers that this will limit the use of pemetrexed as a maintenance therapy (if approved by NICE), thus making erlotinib, should it be approved by NICE, the only currently available maintenance alternative. Recent evidence provided to the Appraisal Committee during the appraisal of gefitinib as a first-line treatment for patients with NSCLC supports the position that the use of pemetrexed in first-line therapy for patients with non-squamous histology is increasing.[17] However, this situation may change with reports (Personal communication; Dr Scott; UCLH NHS Foundation Trust; January 2010) that the cost of gemcitabine has decreased by as much as 85% and so the use of pemetrexed may decrease. In addition, it is worth noting that the evidence in the submission does not include any patients who received pemetrexed as first-line therapy.

3 CRITIQUE OF DECISION PROBLEM

The final scope issued by NICE and the manufacturer's definition of the decision problem are described in the MS (pg 11) and an edited version of the summary is presented in Table 3-1.[12] The manufacturer's additions to the scope are appropriate.

Table 3-1 Overview of decision problem

	Final scope issued by NICE	Decision problem addressed in the MS
Population	People with advanced or metastatic (stage IIIB and IV) NSCLC whose disease has not progressed following treatment with platinum-based first-line chemotherapy.	As scope
Intervention	Erlotinib monotherapy (150mg/day in oral tablet form) Dose reduction cases of adverse events (100mg or 50mg/day).	As scope
Comparator(s)	Best supportive care, which may include palliative radiotherapy care, corticosteroids (without maintenance therapy) and watchful waiting alone. Additionally, for people with non-squamous NSCLC: pemetrexed monotherapy may be included as a comparator, dependent on the outcome of the ongoing STA: pemetrexed for maintenance treatment of NSCLC.	As scope The pemetrexed non-squamous analysis is provided and is considered only of relevance upon a positive recommendation being published for pemetrexed.
Outcomes	The outcome measures to be considered include: -overall survival -progression free survival -tumour response rate -adverse effects of treatment -health-related quality of life.	As scope These outcomes are covered in the submission.
Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. The time horizon for the economic evaluation should reflect the life expectancy of patients with NSCLC. Costs will be considered from an NHS and Personal Social Services perspective.	As scope
Subgroups to be considered	If the evidence allows, subgroups will be considered. These may include subgroups defined by: performance status, histology (squamous/ non-squamous), smoking status, EGFR mutational status, and response to first line treatment. Guidance will only be issued in accordance with the Marketing Authorisation.	We will consider subgroups as appropriate and as far as evidence allows us.
Special considerations	None noted	As scope

NSCLC= non-small cell lung cancer; MS= manufacturer submission

3.1 Licence indications

At the time of writing, erlotinib was being considered by the EMA as a maintenance therapy for patients with NSCLC who have not progressed after first-line chemotherapy. The MS does not provide any details of the licence application submitted to the EMA. However the manufacturer has modelled the cost effectiveness of erlotinib based on three distinct patient populations: ITT, SD and non-squamous. This suggests to the ERG that the manufacturer anticipates that the Marketing Authorisation for erlotinib as a maintenance therapy may be focussed on a specific patient population. The ERG was made aware of their decision when this report was essentially complete.

Erlotinib is also currently being considered by the FDA.[18] The FDA initially rejected the licence extension of the use of erlotinib as a first-line maintenance treatment in patients with locally advanced or metastatic NSCLC. However, the FDA has since extended the review period by an additional 90 days, to April 18 2010, following the manufacturer's submission of further data in support of the application.[2]

3.2 Population

The decision problem indicates that the relevant patient population is made up of patients with advanced or metastatic (stage IIIB and IV) NSCLC whose disease has not progressed following treatment with platinum-based first-line chemotherapy. This is consistent with the patient population described in the primary trial cited by the manufacturer and used as the main source of clinical evidence in the MS. In the MS, patients who had not progressed demonstrated complete response (CR), partial response (PR) or stable disease (SD) (see Table 4-3 later in this document for definitions of each of these states).

3.3 Intervention

Erlotinib is administered as a 150 mg tablet once per day. In the event that patients experience adverse reactions, most commonly rash (in 50% of patients) and diarrhoea (in 20% of patients), the dose is titrated down until symptoms are managed with the lower dose and other symptom specific treatments.

3.4 Comparators

In the UK, current management of patients with NSCLC following first-line chemotherapy is BSC and monitoring. This continues until disease progression when, depending on the performance status (PS) of the patient, second-line chemotherapy may be provided.

In the clinical and economic sections of the MS, the manufacturer considers three distinct patient groups: ITT, SD and non-squamous populations. For the ITT and SD populations, the manufacturer appropriately uses data from the SATURN[4] trial as it compares erlotinib vs placebo. However, for the non-squamous population, as NICE is currently appraising pemetrexed as a maintenance therapy for this group of patients, the manufacturer uses data from the JMEN[5] trial (pemetrexed vs placebo) and performs an indirect treatment comparison. Given the ongoing NICE appraisal of pemetrexed this is appropriate.[3]

3.5 Outcomes

The decision problem described by the manufacturer in the MS states that PFS, OS, tumour response rates, health related quality of life(HRQoL) and adverse events (AE) will be considered; these are all included in the MS and match the final decision problem issued by NICE.

3.6 Time frame

At the time of writing, in the main RCT described by the manufacturer to support its clinical argument, 97% of erlotinib patients had progressed and the maximum duration of censored OS was 34 months.
The economic model uses a five year time frame which is taken to be equivalent to a life-time horizon.

3.7 Other relevant factors

In the final scope issued by NICE, it is stated that consideration should be given to subgroups. These may include subgroups defined by: PS, histology (squamous vs non-squamous), smoking status, epidermal growth factor receptor (EGFR) mutation status and response to first-line treatment. In the clinical section of the MS, the manufacturer considers each of these subgroups in turn; the manufacturer also considers the benefit of erlotinib in patients with stable disease.

4 CLINICAL EFFECTIVENESS

4.1 Critique of manufacturer's approach

Table 4-1 provides an outline of the key background/clinical information and its location within the MS. Its purpose is to signpost the reader to the main areas of background/clinical information within the MS.

Table 4-1 Key clinical information in the MS

Key information	Pages in the MS
Description of technology	8-11
Statement of decision problem	11-12
Context	19-27
Equity and equality	28
Literature search:	
Search strategies	29-30
Study selection	30-35
Clinical effectiveness evidence:	
Trial information	36-70
Results: main and subgroups	71-82
Results: HRQoL analysis	83
Results: Safety	103-108
Indirect/mixed treatment comparisons	86-103

4.1.1 Critique of search

The results of two clinical reviews are reported in the MS; one to inform the review of clinical effectiveness and the second to identify studies for inclusion in the indirect analysis. The first of these is described here while the results of the second are included in the summary and critique of the indirect comparison exercise discussed in section 4.6.

In the MS, the information describing the systematic reviews of the literature undertaken by the manufacturer reveals that review activities were carried out by an un-named author. There is no indication that review decisions (e.g. application of inclusion criteria)or data extraction tables were cross checked by a second reviewer.

The MS provides a clear description of the searches carried out to identify primary relevant research. The ERG re-ran the manufacturer's searches and is confident that no relevant studies have been missed.

4.1.2 Critique of inclusion/exclusion criteria

The explicit inclusion/exclusion criteria used in the review are not described in the MS. The stated inclusion criteria are broad:

> '*that the trial should be informative with regard to the patient population covered by the anticipated Marketing Authorisation for erlotinib*.' (MS, pg 31)

It may be argued that the use of such broad criteria is appropriate to identify all relevant studies for inclusion in the review. However, more specific inclusion criteria are required in order to transparently describe the process used to identify studies for inclusion in the review.

4.1.3 Included and excluded studies

The search conducted by the manufacturer identified three studies for potential inclusion in the review: SATURN,[4] ATLAS[19] and D0410.[20] References provided by the manufacturer related to the first two studies include numerous conference abstracts, some of which contain interim data analysis; the ERG notes that no published papers were identified. For the purposes of this ERG report we use the Clinical Study Report (CSR)[4] to reference the SATURN study and the citations provided in the MS for the ATLAS[19] trial and the D0410 both published as stracts.[20] An appropriate QUORUM flow diagram describing the review process is provided by the manufacturer. No studies were identified that provided a direct comparison of erlotinib with any other active treatment in the maintenance setting for patients with NSCLC.

Although three studies were identified by the searches, D0410[20] was excluded from the review by the manufacturer. The study assessed the clinical effectiveness of erlotinib as a maintenance therapy; however, the study population had received first-line chemoradiation. The manufacturer states that this paper was excluded from the review as the clinical effectiveness of erlotinib after chemoradiation is not relevant to the decision problem.

The manufacturer includes the ATLAS[19] trial in the review, critically appraises the study and discusses its results. However, the manufacturer does not use the results of the ATLAS[19] trial to support the use of erlotinib as a maintenance therapy as the study "has limited direct relevance to the clinical situation in the UK" (MS, pg 83). The ATLAS[19] trial included patients who had received bevacizumab as a first-line therapy (not used in the UK) and, after completion of successful chemotherapy, bevacizumab was

compared with bevacizumab plus erlotinib in the maintenance setting. The ERG agrees with the manufacturer that only the SATURN4 trial is relevant to the decision problem.

4.2 Description of the included study

The SATURN[4] trial is:

> *A multi-centre, double-blind randomized, Phase III study to evaluate the efficacy of Tarceva® or placebo following 4 cycles of platinum-based chemotherapy in patients with histologically documented, advanced or recurrent (Stage IIIB and not amenable for combined modality treatment) or metastatic (Stage IV) non-small cell lung cancer (NSCLC) who have not experienced disease progression or unacceptable toxicity during chemotherapy*. (MS, page 24)

The study characteristics of the SATURN[4] trial are presented in Table 4-2. The ERG highlights that the full clinical paper describing this trial is not yet published. Full details of the inclusion/exclusion criteria used in the SATURN[4] trial are presented in the MS (pg 43-47).

Table 4-3 provides definitions of the key outcome measures and terms used within the trial. Assessments related to progression (RECIST criteria), ECOG performance status (PS) and QoL were measured at baseline and then every six weeks as per the trial protocol.

Table 4-2 SATURN study characteristics

Study	Trial design and patients	Intervention	Comparator	Inclusion criteria	Exclusion criteria	Outcomes
SATURN[4]	RCT Phase III 26 countries 70% of participants from Eastern Europe and Southeast Asia 7 UK patients	Erlotinib (n=451) 150mg/day (oral) Continued until disease progression Dose reduced in the event of toxicity	Placebo (n=438)	Adult patients with NSCLC following on from completion of successful platinum doublet chemotherapy Presence of measurable disease according to RECIST criteria ECOG PS = 0-1 Numerous other clinical markers	Prior treatment with EGFR inhibitors (e.g. gefitinib, cetuximab) Other previous disease treatment regimens or signs of metastasis	Primary: PFS: Assessed at 6 weeks and then every 6 weeks to week 48 then every 12 weeks. Determined by the investigator and a central reviewer and used CT, spiral CT or MRI Secondary: OS (all cause mortality); TTP; RR; quality of life

OS=Overall survival; PFS=Progression free survival; RCT= randomised controlled trial; RECIST= response evaluation criteria in solid tumours; PS= performance status; CT=computed tomography; MRI= magnetic resonance imaging; ECOG= Eastern Clinical Oncology Group; EGFR= epidermal growth factor receptor; IHC=immunohistochemistry; TTP= time to disease progression; RR= response rate

Table 4-3 Definitions used in the SATURN[4] trial

Terms used	Definition
Complete response (CR)	Disappearance of all target lesions
EGFR status	A positive EGFR expression status is defined as having at least 10% of total tumour cells stained for EGFR. All evaluations will be performed following an existing optimized protocol (CSR, pg 1762).
Overall survival	Determined from the date of randomisation to the date of death irrespective of the cause of death.
Partial response (PR)	At least a 30% decrease in the sum of the LD of target lesions, taking as reference the baseline sum LD
Progression free survival	Measured using objective progression (RECIST) plus clinical progression (based on relevant clinical findings – if any) according to the investigator generated dataset
Progressive disease (PD)	At least a 20% increase in the sum of the LD of target lesions, taking as reference the smallest sum LD recorded since the treatment started or the appearance of one or more new lesions
Quality of life	This was measured using the standardised Functional Assessment of Cancer Therapy-Lung (FACT-L)
Response rate	For patients with PR or SD at randomisation a best response of CR or PR required follow-up measurements meeting the criteria for CR or PR at 2 consecutive visits at least 4 weeks apart at any time post baseline. For best response SD follow-up measurements must have met the SD criteria at least 6 weeks post baseline (CSR, page 1776).
Stable disease (SD)	Neither sufficient shrinkage to qualify for PR nor sufficient increase to qualify for PD, taking as reference the smallest sum LD since the treatment started
Time to progression (TTP)	TTP is defined as the time from randomisation to the first date disease progression is recorded using RECIST criteria

LD=lesion dimension; CSR= clinical study report

4.3 Critique of the manufacturer's approach to validity in the SATURN trial

A single trial (SATURN[4]) makes up the basis of the clinical and cost-effectiveness evidence in the MS. This section outlines the strengths and weaknesses of the SATURN[4] trial. Data in this section are taken from the MS as well as from data subsequently provided as a part of the STA clarification process (including the clinical study report) from the manufacturer.

The MS provides a critical appraisal of the SATURN4 trial and the manufacturer comments on the items listed in the CONSORT checklist;[21] a summary of this critique and the ERG's comments are included in Appendix 1. Overall, the study was of good quality with concealment of allocation, random allocation, appropriate sample size calculations and adequate patient follow-up.

The ERG noted two amendments were made to the SATURN[4] trial protocol on 23rd July 2007 (seven items) then on 29th August 2008 (three items). There are two important aspects related to the first of these which was made just a week before the planned interim analysis. The first is related to the number of chemotherapy cycles for patients who were receiving cisplatin and vinorelbine. It was changed from four-weekly cycles to three-weekly cycles based on a study by Gridelli et al.[22] The second was related to tumour response evaluation assessment schedules; these were also changed from two weeks to four weeks prior to chemotherapy and from one week to two weeks prior to starting erlotinib/placebo on the grounds that this would avoid unnecessary restriction for patient enrolment. Any amendments made to a trial protocol raises concerns regarding the robustness of the results of the trial as the modifications may directly impact on outcomes especially for those who were recruited before the protocol modifications came into effect. It is not possible to assess the impact of the two amendments on the results of the SATURN[4] trial.

The SATURN[4] trial was conducted in 110 centres across 26 countries; large numbers were recruited and randomisation was applied centrally via an Interactive Voice Response System rather than within centre or country. However, ensuring uniformity of general clinical practice across and within so many centres is problematic; the results of a trial can only be generalisable if the trial protocol is specific and executed efficiently by investigators. The manner in which the protocol is implemented should be clear to all investigators to ensure that the same systems and procedures are in place across all centres (i.e. to minimise potential clustering) and reduce protocol violations. From data described in the CSR, the ERG notes that 36 serious protocol violations occurred in the SATURN[4] trial, the most common being

described as (i) missing post-baseline tumour assessments and (ii) Progressed Disease at baseline, this therefore calls into question the uniformity of the application of protocol guidance in the trial.

The SATURN[4] trial was a double-blind study in which patients, investigators and the sponsor were all unaware of treatment assignments; blinding was achieved by the use of a matched placebo. Despite the double-blind nature of the study, it is difficult to judge to what extent blinding was successfully maintained throughout the trial period due to major concerns related to the side effects of erlotinib since 48.5% of patients in the erlotinib arm experienced a rash compared to only 4.9% of patients in the placebo arm. Blinding is especially important in trials with a primary outcome of PFS as PFS relies on investigator assessment and therefore could be subject to potential assessment bias.[23] However, as the robustness of the PFS assessment by the investigator was corroborated by the results of the independent central review of radiological and clinical data, this should not be an issue in the SATURN[4] trial.

Of the 889 patients in the trial, seven patients were recruited from four sites in the UK. In total, only 24% of patients were recruited from Western Europe and thus the generalisability of the results to patients in England and Wales might be limited. Such contextual diversity and small numbers may undermine some of the benefits of randomisation and may also cast doubt on the applicability of the results to any one country.

Eligible patients were randomised to receive either erlotinib or placebo using a minimisation allocation technique, an adaptive randomisation method using six factors; EGFR protein expression by Immunohistochemistry (IHC), stage of disease at start of chemotherapy, ECOG PS, chemotherapy regimen, smoking status and region. Justification for only one of these factors, EGFR status, is provided in FDA documentation.[24] There is no rationale for the selection of these factors in the MS. Minimisation sequentially assigns patients to treatment by attempting to minimise the total imbalance between treatment groups over important prognostic factors. One main disadvantage of minimisation approach is that it is predictable. This means allocation of patients can sometimes be guessed by investigators which can lead to selection bias. However, this is unlikely to be a problem in the SATURN[4] trial given the number of factors employed in she stratification process. It is worth noting that unlike stratified randomisation which considers combinations of levels of prognostic factors as mutually exclusive groups and balances within each stratum separately, minimisation considers important prognostic factors together. Minimisation aims for marginal balance among factors rather than within-strata balance. In the SATURN[4] trial the balance within all of the stratified groups appears to have been achieved.

In the SATURN[4] trial patient groups were similar at baseline (Table 4-4). However, the ERG notes that the data demonstrate the expected gender distribution (more males than females) and the usual RCT study population (i.e. slightly younger and healthier patient population than might be seen in clinical practice).

Table 4-4 Patient characteristics

	Erlotinib n=438	**Placebo n=451**
Female	117 (27%)	113 (25%)
Male	321 (73%)	338 (75%)
Age (mean years)	59.8	59.7
White	370(84%)	376(83%)
Other	68(16%)	76 (17%)
ECOG PS 0	135 (31%)	145 (32%)
ECOG PS 1	303 (69%)	306 (68%)
Smoking status		
Current	239 (55%)	254 (56%)
Never smoked	77 (18%)	75 (17%)
Stopped >1year ago	122 (28%)	122 (27%)
EGFR IHC status		
Positive	308 (70%)	313 (69%)
Negative	62 (14%)	59 (13%)
Indeterminate	16 (4%)	24 (5%)
Missing	52 (12%)	55 (12%)
NSCLC stage IV	322 (74%)	342 (76%)
NSCLC unresectable IIIB	116 (26%)	109 (24%)
Prior chemotherapy		
Carboplatin+docetaxel	12 (3%)	12 (3%)
Gemcitabine+cisplatin	118 (27%)	117 (26%)
Other	307 (70%)	321 (71%)
Africa	8 (2%)	8 (2%)
Eastern Europe	207 (47%)	219 (49%)
North America	22 (5%)	20 (4%)
South East Asia	89 (20%)	94 (21%)
Western Europe	112 (26%)	110 (24%)
Tumour Histology		
Squamous cell	166 (38%)	194 (43%)
BAC	3 (<1%)	6 (1%)
Adenocarcinoma	202 (46%)	192 (43%)
Large cell	21 (5%)	24 (5%)
Other	46 (11%)	35 (8%)

IHC= immunohistochemistry; BAC= bronchioloalveolar carcinoma; NSCLC= non-small cell lung cancer; ECOG PS= European Clinical Oncology Group

First-line treatment was one of the stratification factors using the minimisation technique. As noted earlier, no justification is provided for the choice of this factor or the grouping of interventions selected. The stratification considered only three types of first-line treatment (gemcitabine+cisplatin, carboplatin+docetaxel, and others); from the data presented by the manufacturer at least seven different first- line treatments were administered to patients. A summary of the first-line treatments received by

patients in the SATURN[4] trial is presented in Table 4-5; the table shows that treatments are similar across both groups.

In the UK NICE currently recommends the following first-line chemotherapy treatments for patients with NSCLC: combination chemotherapy with a platinum-based drug (cisplatin or carboplatin) plus a second cytotoxic drug (gemcitabine, paclitaxel, docetaxel, vinorelbine and most recently pemetrexed for non-squamous cell carcinoma).[25] Patients in the trial received these recommended treatments except for pemetrexed. This is unfortunate as noted in the MS by the manufacturer:

> *"...[in the UK]pemetrexed is now viewed as the non-platinum drug of choice for use in combination with platinum for the first-line chemotherapy of predominantly non-squamous NSCLC – a view recently endorsed by NICE in TA 124- non-squamous patients finishing first-line platinum doublet treatment are unlikely to be pemetrexed naïve in the future"*. (MS, pg 112)

The ERG also notes that the use of paclitaxel in the SATURN[4] trial is also higher that would be expected in the UK.

Table 4-5 First-line treatment as presented in the CSR

(CSR, pg497-8)	**Erlotinib (n=433)**		**Placebo (n=445)***	
	N	**%**	**N**	**%**
Cisplatin plus paclitaxel	52	12	55	12
Cisplatin plus gemcitabine	118	27	117	26
Cisplatin plus docetaxel	16	4	24	5
Cisplatin plus vinorelbine	33	8	32	7
Carboplatin plus paclitaxel	82	19	83	19
Carboplatin plus gemcitabine	122	28	130	29
Carboplatin plus docetaxel	13	3	13	3
Total	436	101	454	101

*numbers slightly different to those in other provided tables

Use of PFS as the primary outcome measure in RCTs, especially RCTs designed to assess maintenance-targeted therapies, is of concern and is open for debate. Consistency of assessment is critical as any difference in tumour assessment could induce bias in the PFS result. In the SATURN[4] trial tumour assessments were performed using more than one approach. These included computed tomography scan,

spiral computed tomography or magnetic resonance imaging; for each patient the same method was to be used throughout the study. It is documented in the MS that if more than one method of measurement was used, the data from the most accurate method according to RECIST criteria should be recorded. Despite measures undertaken by the investigators to maintain uniform tumour assessment, the ERG believes the reliability of PFS results is of concern given the large number of centres and different methods of tumour assessments.

The data presented in the MS did not provide a clear picture of the treatments that the patients received after disease progression in the SATURN[4] trial. Examination of the CSR[4] provided some of this information. Unfortunately, there was a lack of consistency between the information provided in the MS and the CSR and the ERG was unable to reconcile these differences. In the MS (page 49) the total number and proportion of patients receiving at least one post-study treatment in the placebo and erlotinib groups were 325(72%) and 309(71%) respectively. As can be seen in Table 4-6, which includes data taken from the CSR, the reported numbers are much lower (290[64%] and 241[55%], respectively).

The manufacturer acknowledges that the administration of post-progression treatments and the provision of erlotinib to placebo patients in the SATURN[4] trial would be likely to attenuate any impact of study treatment on the secondary endpoint of OS but does not offer further consideration of this impact. Thus the interpretation of OS and other outcomes, including HRQoL, that are evaluated in the post-progression phase of the SATURN[4] trial is of concern.

The type and distribution of post-progression therapies administered to patients in the SATURN[4] trial are not sufficiently similar to those administered to patients with NSCLC in the UK after failure of first-line chemotherapy to dismiss concerns about the generalisability of the trial to patients in the UK. Only docetaxel and erlotinib are currently approved by NICE for the second-line treatment of patients with NSCLC. It is clear from the data provided by the manufacturer that a wide range of treatments, including experimental treatments, and subsequent lines of therapy were received by patients in both arms of the SATURN[4] trial. As administration of post-progression therapies directly influences estimates of OS, the ERG considers that the OS benefit demonstrated in the SATURN[4] trial may not be replicable in an NHS setting; it is not possible to speculate whether OS estimates would improve or worsen in a UK setting.

Table 4-6 Post study treatments

From CSR (page 659-661)	Erlotinib n=438 (%)	Placebo n=451 (%)
Total patients with at least one treatment	241(55)	290(64)
Erlotinib	9(2)	48(11)
Docetaxel	102(23)	110(24)
Paclitaxel	14(3)	14(3)
Pemetrexed	61(14)	64(14)
Gemcitabine	19(3)	31(7)
Vinorelbine	23(5)	34(8)
Carboplatin/Doublet*	10(2)	14(3)
Cisplatin/Doublet**	8(2)	12(3)
Gemcitabine/Vinorelbine	3(1)	7(2)
Carboplatin	18(4)	35(8)
Cisplatin	18(4)	18(4)
Surgical and medical procedures	84(19)	101(22)
Gefitinib	8(2)	21(5)

*Docetaxel, paclitaxel, gemcitabine, pemetrexed, vinorelbine and etoposide.
** Docetaxel, paclitaxel, gemcitabine and vinorelbine

4.3.1 Outcomes and results of the SATURN trial

The primary outcome reported in the SATURN[4] trial is PFS. The MS also appropriately provides analyses of secondary outcomes of OS, tumour response rates and HRQoL as stated in the final scope issued by NICE. Results of the full analysis set (FAS) outcomes of the SATURN[4] trial are presented in Table 4-7. Subgroup analysis (using stratification factors) and *post hoc* analyses conducted by the manufacturer are presented in Table 4-8.

As can be seen from Table 4-7 using the FAS, the median PFS is 12.3 weeks for the erlotinib group and 11.1 weeks for the placebo group. Although this difference of 1.2 weeks between the groups is reported as being statistically significantly different with a HR of 0.71 and a 95% CI of 0.62, 0.82, it represents a small clinical difference. A similar situation is reported for OS using the FAS where the median OS is 12 months for the erlotinib group and 11 months for the placebo group with a HR of 0.81 and a 95% CI of 0.7, 0.95.

In Table 4-7 the non-stratified analysis reported in the MS shows statistically significant results in the erlotinib group for PFS and OS. However, when the results of the Log rank stratified analysis were received with the clarification response, the ERG noted that the statistically significant OS benefit was no longer apparent. In the clarification response, the manufacturer explained that a multiple Cox regression analysis was also carried out and that erlotinib generated a statistically significant OS benefit compared with placebo using this method.

Table 4-7 SATURN study outcomes (FAS)

			Non-stratified analysis		Stratified analyses	
Endpoint	**Erlotinib (n=438)**	**Placebo (n=451)**	**HR (95% CI)**	**P value (Log rank)**	**HR (95% CI)**	**P value**
Primary						
PFS (Weeks – median) (95% CI)	12.3 (12.0, 13.3)	11.1 (8.1, 11.7)	0.71 (0.62, 0.82)	<0.0001	0.70 (0.59, 0.84)	Log rank p<0.0001
Secondary						
OS (Months – median) (95% CI)	12.0 (10.6, 13.9)	11.0 (9.9, 12.1)	0.81 (0.7, 0.95)	0.008	0.85 (Log rank) CI not provided	Log rank p=0.0839
					0.82 CI not provided	Multiple Cox Regression* p=0.0103
Estimated 1 year survival rate n/N; %	204/438; 47%	196/451; 43%				
Complete tumour response n(%) (95% CI)	4(0.9%) (0.3, 2.3)	3(0.7%) 0.1, 2.0				
Partial tumour response n(%) (95% CI)	48 (11.0%) (8.2, 14.3)	21(4.7%) (2.9, 7.1)				
Stable disease n(%) (95% CI)	212 (48.6%) (43.8, 53.4)	202(45.4%) (40.7, 50.1)				
Progressive disease n(%) (95% CI)	155(35.6%) (31.1, 40.2)	212 (47.6%) (42.9, 52.4)				
Quality of life						
Time to symptom progression Time to deterioration in TOI Time to deterioration in quality of life	No difference between treatment arms		0.91 (0.74,1.12) 1.06 (0.87,1.31) 0.96 (0.79,1.16)	p=0.38 p=0.54 p=0.65		

*additional factors are ECOG performance status and smoking status only
CI= confidence interval: FAS-Full analysis set; OS= overall survival; PFS= progression free survival; HR= hazard ratio; TOI= trial outcome index

4.3.2 Subgroup analysis

As previously stated, in the SATURN[4] trial randomisation was undertaken using minimisation technique to ensure a balance between the treatment arms for the following factors: EGFR status (positive IHC vs negative IHC); stage of disease (stage IIIB vs stage IV); ECOG PS (0 vs 1); first-line chemotherapy (gemcitabine+cisplatin vs carboplatin+docetaxel vs other); smoking status (current vs former vs never) and geographical region (North America, South America, Western Europe, Eastern Europe, South East Asia and Africa). The manufacturer presents analyses for PFS and OS for each of the stratification factors. Also of interest to the ERG is the *post hoc* subgroup analysis provided for PFS and OS by histology (non-squamous) and response to treatment (stable disease).

Progression free survival and overall survival

Table 4-8 Non-stratified subgroup analysis: PFS and OS

Factor	HR for PFS (95%CI)	HR for OS (95%CI)
EGFR IHC		
Positive	0.69 (0.58, 0.82)	0.77 (0.64, 0.93)
Negative	0.77 (0.51, 1.14)	0.91 (0.59, 1.38)
Indeterminate	0.78 (0.55, 1.11)	0.96 (0.67, 1.38)
Stage of disease		
IIIB	0.83 (0.82, 1.10)	0.81 (0.59, 1.11)
IV	0.68 (0.58, 0.81)	0.81 (0.68, 0.97)
ECOG status		
0	0.59 (0.45, 0.77)	0.86 (0.65, 1.13)
1	0.77 (0.65, 0.92)	0.78 (0.65, 0.94)
First-line treatment		
Gemcitabine+cisplatin	0.79 (0.59, 1.04)	0.95 (0.71, 1.29)
Other	0.69 (0.58, 0.81)	0.77 (0.64, 0.92)
Smoking status		
Never	0.56 (0.38, 0.81)	0.69 (0.45, 1.05)
Current smoker	0.80 (0.67, 0.97)	0.88 (0.72, 1.08)
Past smoker	0.66 (0.50, 0.88)	0.75 (0.56, 1.00)
Region		
Eastern Europe	0.77 (0.63, 0.95)	0.88 (0.71, 1.10)
Western Europe	0.75 (0.56, 0.99)	0.80 (0.59, 1.09)
North America	0.79 (0.41, 1.52)	0.80 (0.40, 1.60)
South East Asia	0.56 (0.40, 0.78)	0.71 (0.50, 1.02)
Histology		
Squamous cell carcinoma	0.76 (0.60, 0.95)	0.86 (0.68, 1.10)
Non-squamous cell carcinoma	0.68 (0.56, 0.82)	0.79 (0.64, 0.96)
Response to first-line treatment		
Good (CR/PR) reduction	0.74 (0.60, 0.92)	0.94 (0.74, 1.20)
Less than adequate (SD) reduction	0.68 (0.56, 0.83)	0.72 (0.59, 0.89)

PFS= progression free survival; OS= overall survival; CI= confidence interval; IHC= immunohistochemistry; NSCLC= non-small cell lung cancer; ECOG= Eastern Clinical Oncology Group; CR= complete response; PR= partial response; SD= stable disease

EGFR status

The emphasis on the assessment of EGFR status comes from negotiations with the FDA and their request for a post approval commitment study to evaluate the relationship between EGFR protein expression and clinical outcome.[24]

In the SATURN[4] trial, tumour biomarker analyses were performed in the following order: EGFR IHC, EGFR FISH, Kras mutations and EGFR mutations. Approximately 70% of patients in the SATURN[4] trial were EGFR IHC positive; however only 5.5% of patients were identified as being EGFR mutation positive. There was a high rate of indeterminate (8%) and missing (42%) samples when the EGFR mutation test was used.

As shown in the MS (Figure 9 of the MS, pg 75) patients with EGFR IHC positive status who are treated with erlotinib have a highly statistically significant PFS benefit compared to patients on placebo, a result that is not shown for EGFR IHC negative patients. Results of the same analysis related to OS show a similar result. Results of analysis related to EGFR FISH, EGFR mutation status and KRAS mutation status do not demonstrate this difference. These results are consistent with the current debates[26] related to the predictive capacity of EGFR IHC status and the MS concludes therefore that mandatory testing for mutation status is not required prior to treatment.

Disease stage

Results related to disease stage demonstrate that, compared to placebo, patients with stage IV disease are statistically likely to benefit more from erlotinib in terms of PFS and OS than patients with stage IIIB.

ECOG status

Results of the analysis by ECOG status at baseline vary. In terms of PFS, all patients taking erlotinib are statistically significantly likely to benefit from treatment compared with placebo. In terms of OS, only patients with PS equal to 1 show a statistically significant benefit compared with placebo.

First-line treatment

Although the trial population was stratified by first-line treatment, the stratification considered only three groups; gemcitabine +cisplatin (n=235), carboplatin+ docetaxel (n=26) and others (n=639). The numbers of patients in each group is approximate as there is a numerical discrepancy in the tables presented in the MS. In addition, since only 26 patients received carboplatin+docetaxel, they were combined with the "other" group in the subgroup analysis. Therefore the resultant analysis compares

approximately 235 patients with the remaining patients (n>650). The results of the analysis should therefore be interpreted with caution due to the artificial nature of the groupings.

Smoking status and region

Current, past and never smokers appear to derive a statistically significant benefit from erlotinib in terms of PFS compared with placebo; this difference between groups is not evident in the OS data. Similar results are shown in the analyses by region with the exception that compared with placebo, the small North American cohort does not derive a statistically significant benefit from erlotinib in terms of PFS or OS.

4.3.3 Post hoc analyses

Non-squamous and squamous histology

The ERG notes that disease histology was not a stratification factor and all comparisons are the result of *a post hoc* analysis. Patients with non-squamous histology derive a statistically significant PFS and OS benefit from erlotinib compared with placebo. In addition, the MS cites a conference presentation from Cappuzzo et al[27] who report that erlotinib leads to a 3.2 month increase in median OS from 10.5 months to 13.7 months in the entire non-squamous population compared with placebo.

The ERG notes that although the evidence shows a statistically significant PFS benefit for erlotinib compared with placebo regardless of histology, in the case of OS such a benefit is seen only in patients with non-squamous histology. The ERG carried out projection modelling for OS as described in later in section 5.5.4. For patients with non-squamous histology, the ERG's revised OS benefit is estimated to be less than three months (2.7 months); this estimate is lower than the 3.2 months described in the MS.

Stable disease

The manufacturer presents an analysis reporting that in patients with stable disease the use of erlotinib demonstrates a consistent statistically significant benefit in terms of PFS and OS; responding (CR+PR) patients show a statistically significantly benefit in terms of PFS only. The manufacturer does not present median OS in months for stable disease patients; the manufacturer simply states that erlotinib leads to an increase in 3.3 months for patients with stable disease compared with placebo. The ERG carried out projection modelling for OS as described in later in section 5.5.4. The ERG agrees with the manufacturer that for patients with SD there is an expected life extension of at least three months. The ERG estimates a survival advantage for erlotinib patients of approximately 4.2 months which is higher than the estimate of 3.3 months described in the MS.

4.4 Health related quality of life

Health-related quality of life was assessed in the SATURN[4] trial using the FACT-L version 4 questionnaire, which consists of 27 general health questions (FACT-G) and nine lung cancer questions (FACT-L subscale); the FACT-L was administered to study patients during each six week assessment. This tool has been validated[28] and it was found that clinically relevant changes in scores were estimated as two to three points for the FACT-L subscale and five to six points for the Trial Outcome Index (TOI) of the FACT-L; the TOI is defined as the sum of the scores of physical well being, functional wellbeing and lung cancer scores of the FACT-L instrument.

As shown in Table 4-7 the MS reports no statistically significant differences between the treatment arms in relation to time to symptom progression, time to deterioration in TOI or time to deterioration in QoL. The MS reports a *post hoc* analysis of time to pain and time to analgesics; both times were longer in patients receiving erlotinib (HR 0.61, p=0.08 and HR 0.66, p=0.12, respectively).

4.5 Safety/adverse events

In the SATURN[4] trial 78.8% of patients in the erlotinib arm experienced an adverse event compared with 54.2% of placebo patients. The reported rates of grade 1 or 2 AEs of any kind is 54% in the erlotinib arm compared with 42% in the placebo arm of the trial. As would be expected the proportion of grade 3 or 4 events is higher in the erlotinib arm (24.7% vs 12.1%). The MS reports that the majority of patients required neither dose modification nor terminated treatment due to these AEs. A total of 11% of patients required dose modification of erlotinib to manage their AEs. The most frequently reported events in erlotinib patients were rash (reported in 50% of patients) and diarrhoea (reported in 20% of patients).

4.5.1 Description and critique of the statistical approach used in the trial

The SATURN[4] trial was designed with PFS as a primary outcome in two separate patient populations. The first population is described as a full analysis set (FAS/ITT) and is defined as patients with locally advanced or metastatic NSCLC who have not progressed after four cycles of first-line treatment with platinum-based chemotherapy. The second patient population is a subgroup of the FAS who have positive EGFR protein expression as assessed by IHC.

An interim analysis was undertaken in the SATURN[4] trial; neither methods nor results were clearly described in the MS. In response to the ERG's clarification letter, the manufacturer provided the CSR from the SATURN[4] trial as well as a large amount of additional data. In the SATURN [4] trial, one

interim analysis of PFS and safety was planned after approximately 365 events (disease progression or death) had occurred (50% of events), and this was expected to be on 30th July 2007. The interim analysis was based on an acceptable approach using a Lan-DeMets alpha spending function with an O'Brien-Fleming boundary to maintain an overall alpha of 0.05. The MS appropriately adjusted for the two co-primary analyses using pre-specified alpha values of 0.03 and 0.02 for FAS and EGFR IHC positive patients groups, respectively.

The ERG was concerned with the fact that, in the MS, the manufacturer did not adequately describe the protocol violations that occurred in the SATURN[4] trial. However, the manufacturer provided further details of protocol violations in the clarification response. The additional data provided by the manufacturer show similar rates of protocol violations for each of the trial arms (erlotinib: 19/438; placebo: 17/451) and provided information regarding how these were managed in the analysis. The major protocol violations are listed in Appendix 2.

The manufacturer presented results for several exploratory analyses on the FAS and also for different subgroups. There are often problems with this type of analysis; in particular, the ERG is concerned with possible loss of statistical power and the consequences of not adjusting for multiple testing. Statistical tests on subgroups only have power to detect substantially larger effects on the outcome of interest. In subgroup analysis, where a group of factors may influence the outcome, the risk of false-positive results is high. One way to partly overcome these problems is to use an interaction term in a regression model. This approach is not clearly explained in the MS but, based on the evidence presented in the CSR; the ERG concludes that this appears to have been investigated by the manufacturer. The ERG considers that it may have been more appropriate to consider the separate subgroups within the same statistical model as opposed to subdividing the population for most of the subgroup analyses presented. However, there is no evidence to suggest that the approach employed by the manufacturer has introduced any significant bias into the results.

4.6 Indirect comparison

The MS reports the results of a systematic review of the literature designed to identify pemetrexed maintenance studies that might be used to provide data for an indirect comparison of pemetrexed with erlotinib (MS, pg 86-88). The search strategy identified only one relevant study for use in the indirect comparison.[5] An appropriate QUOROM flow diagram of the pemetrexed study selection process is presented by the manufacturer. The ERG was involved in the recent appraisal of pemetrexed as maintenance therapy and is confident that no relevant studies have been missed.

The only relevant study identified for use in the indirect comparison was the JMEN[5] trial; this RCT compared pemetrexed plus BSC vs placebo plus BSC in patients (n=663) who had received four cycles of platinum-based chemotherapy and whose disease had not progressed. The MS points out what the manufacturer believes to be important differences between the patient populations in the JMEN[5] trial and the patients in the SATURN[4] trial. These differences are summarised in Table 4-9.

Table 4-9 Differences in characteristics of patient populations (JMEN and SATURN)

	JMEN[5] trial	SATURN[4] trial
Never smokers	30%	18%
Patients with squamous histology	28%	43%
Asian patients	31%	21%
Patients receiving post-progression treatments	55%	72%

In principle, differences in the baseline characteristics of patients from different trials may influence the size of the health effect estimated in an indirect analysis. However, the ERG believes any perceived differences in the patient populations of the JMEN[5] and SATURN[4] trials are not considered to be important as the available trial evidence shows that patients in the JMEN[5] and SATURN[4] trials respond to treatment with placebo in the same way.

In contrast, the ERG considers that the generalisability of the JMEN[5] and SATURN[4] trials is important and merits discussion. The generalisability of the SATURN[4] trial is discussed in section 4.3. The ERG is of the opinion[29] that the generalisability of the JMEN[5] trial to UK clinical practice is uncertain: there were no UK trial centres, there was a high proportion of Asian patients who are known to respond better to lung cancer treatments than other ethnicities; patients received a high rate of post-progression treatments that are not commonly administered in the UK; none of the patients received vinorelbine or pemetrexed as a first line therapy; and patients in the JMEN trial received unlimited cycles of pemetrexed as maintenance, this is unlikely to occur in the UK.

The JMEN[5] and SATURN[4] trials share a common comparator (no treatment until disease progression in patients achieving at least SD after first-line chemotherapy) so it is possible to perform a simple indirect comparison of hazard ratios for PFS and OS for the interventions vs the common control. The manufacturer appropriately recognised that this approach has limitations because of differences in the baseline characteristics of the patients in the two studies. The manufacturer submitted a description of the indirect comparison analysis and it is clear that the analyses were performed for PFS and OS in the ITT and non-squamous patient populations. However, the results of the indirect comparison were not presented in the main submission and the approach taken by the manufacturer was not clearly described; the ERG requested and received additional information regarding the methods used in the indirect comparison.

Based on the information described in the MS and the manufacturer's clarification response, it appears that the manufacturer employed an adjusted indirect comparison approach based on the methodology proposed by Bucher.[30] This approach is widely used in absence of head-to-head evidence and the method maintains the randomisation from each trial and compares the summary estimates of pooled interventions with CIs.

Despite the appropriateness of the method, the ERG is concerned with the approach used by the manufacturer to estimate the variance of log hazard ratios for the indirect estimates between pemetrexed and erlotinib. The manufacturer used $\left[\frac{\ln(upper\ limit)-ln(HR)}{(1.96)}\right]^2$ to estimate the variance for log hazard ratio for each trial. This approach ignores information from the lower confidence interval and therefore underestimates the variance of the indirect estimate. A précised approach would be to use the method proposed by Parmar et al.[31] This method utilises information from both the upper and lower confidence intervals for the log hazard ratio; the formula for log hazard variance is $\left[\frac{\ln(upper\ limit)-ln(lower\ limit)}{(2*Z\ score\ of\ the\ upper\ limit)}\right]^2$. As the ERG's revised indirect estimates are very similar to those presented by the manufacturer, the manufacturer's results are presented in Table 4-10.

Table 4-10: Direct and indirect evidence (PFS and OS) in the SATURN and JMEN trials

		Direct evidence vs placebo		Indirect evidence: Manufacturer estimate	
Patient population	**Drug**	**HR (PFS) (95% CI)**	**HR (OS) (95% CI)**	**HR (PFS) (95% CI)**	**HR (OS) (95% CI)**
ITT	Erl	0.71 (0.61, 0.84)	0.81 (0.70, 0.95)		
	Pem	0.50 (0.42, 0.61)	0.79 (0.65, 0.95)	[redacted]	[redacted]
Non-squamous	Erl	0.68 (0.56, 0.82)	0.79 (0.64, 0.96)		
	Pem	0.44 (0.36, 0.55)	0.70 (0.56, 0.88)	[redacted]	[redacted]
Squamous	Erl	0.76 (0.60, 0.95)	0.86 (0.68, 1.10)		
	Pem	0.69 (0.41, 0.98)	1.07 (0.77, 1.50)	[redacted]	[redacted]

Erl= erlotinib; pem = pemetrexed; ITT= intention to treat; HR= hazard ratio; CI= confidence interval; PFS= progression free survival; OS= overall survival; ERG= evidence review group

In summary, the clinical data from the two trials show that both erlotinib and pemetrexed improve PFS and OS when used as maintenance treatment in an unselected patient population and both provide benefit to patients with non-squamous histology. The indirect comparison shows that pemetrexed vs erlotinib in patients with non-squamous histology yields a statistically significant PFS benefit for patients on pemetrexed; however, this difference was not statistically significant for patients with non-squamous histology in terms of OS. Given the perceived differences in the patient populations of the two studies and the ERG's view that the generalisability of both trials to patients in the UK is uncertain, the ERG agrees with the manufacturer that these results should be interpreted with caution when considering maintenance treatment for patients in the UK.

4.7 Summary of clinical evidence

4.7.1 Clinical results

Direct comparison: erlotinib vs placebo

- The main source of clinical evidence described in the MS is from the SATURN[4] trial
- SATURN[4] trial includes patients with NSCLC who have not progressed after first-line chemotherapy
- Compared with placebo, erlotinib demonstrated a statistically significant increase in PFS in the ITT (29%), SD (32%) and non-squamous populations (32%)
- Compared with placebo, erlotinib demonstrated a statistically significant increase in OS in the ITT (19%), SD (28%) and non-squamous populations (21%)
- One year survival rates are also statistically significantly increased for patients in the erlotinib arm and there was no evidence of statistically significant differences between erlotinib patients and placebo patients with respect to HRQoL outcomes
- In the SATURN[4] trial 78.8% of patients in the erlotinib arm experienced an adverse event compared with 54.2% of placebo patients; 50% and 20% patients in the erlotinib arm suffered from a rash and diarrhoea, respectively, compared with 6% and 4% of patients in the placebo arm
- Approximately 70% of patients in the SATURN[4] trial were EGFR IHC positive; however only 5.5% of patients were identified as being EGFR mutation positive (high rate of indeterminate (8%) and missing (42%) samples).

Indirect comparison: pemetrexed vs erlotinib

- The manufacturer appropriately uses clinical data from the JMEN[5] trial to indirectly compare pemetrexed with erlotinib in patients with non-squamous histology
- The MS sets out differences between patients in the SATURN[4] and JMEN[5] trials including: higher proportion of never smokers, patients with squamous histology and Asian patients in the JMEN[5] trial compared to the SATURN[4] trial
- Results of the indirect comparison show that, in patients with non-squamous histology, pemetrexed yields a statistically significant PFS benefit compared with erlotinib but pemetrexed does not yield a statistically significant OS benefit compared with erlotinib.

4.7.2 Clinical issues

Direct comparison: erlotinib vs placebo

- The ERG considers the SATURN4 trial to be generally well-designed but notes the following weaknesses: rationale for the six stratification factors used as part of the randomisation process is inadequate; blinding was difficult to maintain as patients given erlotinib are more likely to suffer from rashes and diarrhoea; the manufacturer appears to be focussed on patients with non-squamous histology or stable disease yet only post-hoc analyses have been conducted
- The ERG considers that the generalisability of the results of the SATURN4 trial to patients in England and Wales is uncertain due to the following:
    - Only seven patients were recruited from the UK, 75% of patients in the trial were from outside of Western Europe, first-line chemotherapy treatments are not consistent (in terms of proportions of patients receiving various treatments) with current UK practice, the range and frequency of post-progression therapies administered to patients in the trial do not appear to be similar to those available in the UK

- In the SATURN4 trial none of the patients received pemetrexed as a first-line therapy. This means that the trial does not provide clinical evidence to support the use of erlotinib as a maintenance therapy after first-line treatment with pemetrexed. Pemetrexed is recommended by NICE for use in patients in the UK with "other than squamous cell carcinoma"
- The role of erlotinib in the treatment pathway will depend on the outcome of the current NICE appraisal of pemetrexed as a maintenance therapy and whether pemetrexed becomes a dominant first-line therapy in the UK for patients with non-squamous NSCLC
- Due to the small number of samples available for EGFR mutation status testing, the relationship between a positive EGFR IHC result and EGFR mutation positive status remains uncertain.

Indirect comparison: pemetrexed vs erlotinib

- The ERG agrees with the manufacturer that there are differences in the characteristics of the patients in the trials used in the indirect comparison exercise; however, the ERG has compared the placebo arms of the SATURN[4] and JMEN[5] trials for both PFS and OS and found that they are very closely matched, suggesting that any potential differences in baseline patient characteristics are not important
- As the ERG is concerned with the generalisability of the results of the SATURN[4] and JMEN[5] trials to patients in the UK, the ERG agrees that the results of the indirect comparison of pemetrexed vs erlotinib must be interpreted with caution.

5 ECONOMIC EVALUATION

5.1 Introduction

This section provides a structured critique of the economic evidence submitted by the manufacturer of erlotinib. The two key components of the economic evidence presented in the MS are (i) a systematic review of the relevant literature and (ii) a report of the manufacturer's *de novo economic* evaluation. See Table 5-1 for a summary of key information points. The manufacturer also provided an electronic version of the EXCEL based economic model.

Table 5-1 Key information in the MS

Key information	**Pages (MS)**	**Key tables/figures (MS)**
Details of the systematic review of the economic literature	113	-
Technology, patients, comparator, perspective, time horizon	114-118	Table 17
Framework for model-based evaluation	118-131	Figures 23-26; Tables 19-26
Clinical evidence used in economic evaluation	131-134	Table 27-30
Measurement and valuation of health benefits	134-137	Tables 31-32
Resource identification, measurement and valuation	137-146	Tables 33-43
Methods of sensitivity analysis and validity assessment	146-152	Tables 44-47
Results – base case analysis	152-159	Tables 48-54; Figure 30
Results – subgroup analysis	152-159	-
Results – sensitivity analysis	159-167	Tables 55-57; Figures 31-38
Results – end of life criteria	178-179	-
Assessment of factors relevant to the NHS and other parties	170-178	Tables 60-69

5.2 Overview of manufacturer's cost-effectiveness review

The MS provides a brief description of the review of published cost-effectiveness evidence undertaken by the manufacturer. The databases searched and the search terms used appear to be reasonable; however, only exclusion criteria were explicitly stated. The search by the manufacturer did not identify any relevant studies for inclusion in the review. Although there is no mention of searching in-house databases for relevant studies, the ERG is confident that no relevant published studies are available for inclusion in the review.

5.3 Overview of manufacturer's economic evaluation

The manufacturer undertook a *de novo* economic evaluation of erlotinib plus BSC compared with BSC alone for the maintenance treatment of NSCLC in patients who have not progressed after four cycles of treatment with platinum based first-line chemotherapy. The manufacturer presents three separate economic evaluations, one for each population being evaluated (ITT, SD and non-squamous). Each economic evaluation follows the same structure unless otherwise stated.

5.3.1 Description of the manufacturer's economic model

The economic model is designed to estimate the lifetime direct NHS costs and total QALYs for each population being evaluated. In the model, patients are assumed to be in one of three possible discrete health states at any given time: "progression free survival" (PFS), "progressed" or "death". The structure of the economic model is shown in Figure 5-1. The model takes an Area-Under-the-Curve (AUC) approach constructed using EXCEL™ with a cycle length of one month. Patients enter the model in the PFS health state; at the end of the cycle the patient can remain in PFS or move to the "progressed" health state or die.

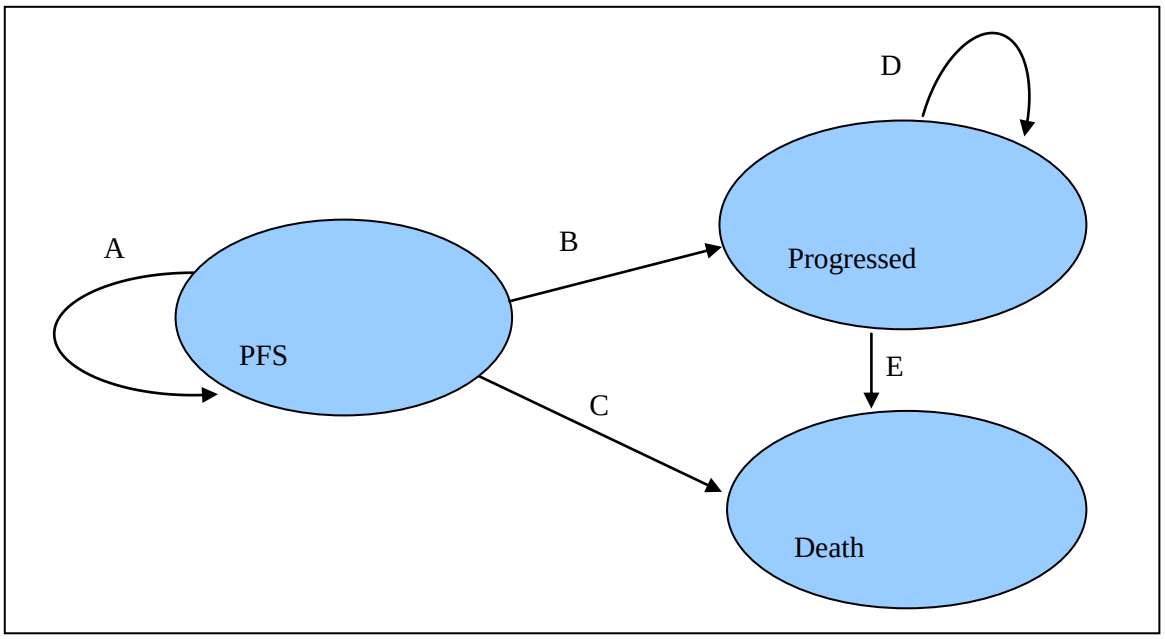

Figure 5-1 Structure of the manufacturer's model

The health states in the model reflect those health states measured within the SATURN[4] trial (erlotinib plus BSC as first-line maintenance vs placebo plus BSC) and the JMEN[5] trial (pemetrexed as first-line maintenance vs "watch and wait" treatment); both of these trials are discussed extensively in the clinical effectiveness section of this report.

The "progressed" health state represents the period of time between first treatment relapse and death; this health state therefore includes the possible sequence of remission and relapse of further treatments. Death is an absorbing health state within the model; this health state is not a true health state as there are no explicit transitions to this health state.

5.3.2 Parameters and values

The base case model parameters and vales are presented in Table 5-2.

Table 5-2 Parameters and values used by the manufacturer in the economic model

Model variable	Value	Source
Costs		
Monthly PFS health state supportive care	£361.44	TA162,[32] 2008 inflated by PSSRU 2009
Monthly progressed state supportive care	£1,088.93	TA162,[32] 2008 inflated by PSSRU 2009
Drug costs†		
Monthly cost of erlotinib	£1,415.30	BNF 58[33] less 14.5% existing erlotinib PAS (TA162)
Monthly cost of pemetrexed	£2,188.03	BNF 58[33]
Monthly cost of concomitant medication for pemetrexed	£19.84	BNF 58, [33] 1LM pemetrexed submission, 2009
Post-progression drug cost		
ITT population	£325 (erlotinib) £440 (placebo)	BNF58[33]
SD population	£322 (erlotinib) £483 (placebo)	BNF58[33]
Non-squamous population	£226 (erlotinib) £413 (placebo)	BNF58[33]
Mean dose of erlotinib		
Mean dose for the ITT population	140.70mg	SATURN[4] trial
Mean dose for the SD population	141.16mg	SATURN[4] trial
Mean dose for the NS population	138.14mg	SATURN[4] trial
Drug administration costs†		
Erlotinib. Monthly pharmacy preparation cost	£13.50	CPORT/2009/ NCAT, [34] 2009/ erlotinib and pemetrexed SmPCs
Pemetrexed IV and its concomitant medication. Pharmacy preparation cost (per cycle)	£37.35	CPORT, 2009/ NCAT,[34]2009/ erlotinib[35] and pemetrexed [36]SmPCs
Pemetrexed delivery cost (per cycle)	£212	Ref cost 07-08[37] Deliver simple parenteral chemotherapy at first attendance Code SB12Z
Adverse events costs		
Rash	£129.00	TA162[32], PSSRU, 2009,
Diarrhoea	£261.55	TA162[32], PSSRU, 2009
Neutropenia	£330.93	1LM pemetrexed submission to NICE, 2009[7]
Anaemia	£615.04	1LM pemetrexed submission to NICE, 2009[7]
Fatigue	£38.90	1LM pemetrexed submission to NICE, 2009[7]
HR pemetrexed vs erlotinib		
PFS	[redacted]	SATURN[4] trial, JMEN[5] trial
OS	[redacted]	SATURN[4] trial, JMEN[5] trial
Utilities		
PFS health state	0.695: ITT population 0.685 SD/NS populations	SATURN[4] trial, Kind et al,[38] 2005, Kind et al,[39] 2009
Progressed health state	0.47	Naffees et al[40], 2008
Discount rates		
Costs	3.5%	NICE Guide to Methods[41]
QALYs	3.5%	NICE Guide to Methods[41]

†Costs are provided by month, the costs provided in this table have been adjusted to account for 30.4375 days per month; PFS= progression free survival; OS= overall survival; QALYs= quality adjusted life years; ITT= intention to treat; SD= stable disease; NS= non-squamous; vs= versus; BNF= British National Formulary; PSSRU= Personal and Social Services Resource Use; HR= hazard ratio; TA= technology appraisal; 1LM= first line maintenance; PAS = payment access scheme

5.3.3 Treatment effectiveness within the MS

The clinical data used in the manufacturer's economic evaluations are mostly taken directly from the SATURN[4] trial, which is described in section four of this ERG report.

For the purposes of survival analysis, patient level data on PFS and OS were obtained from the SATURN[4] study for the ITT, SD and non-squamous populations. Since the vast majority of erlotinib patients had progressed at the point of follow up, a mature data set of PFS outcomes from the SATURN[4] study was available. The manufacturer considered it reasonable to use the SATURN[4] Kaplan-Meier curves directly within the model to estimate mean PFS in the ITT and SD populations.

In order to perform the indirect comparison [redacted].

To estimate OS a variety of parametric functions was evaluated. Log Logistic, Gamma and Log Logistic functions were judged by the manufacturer to be the best fit for the ITT, SD and non-squamous populations respectively (Table 5-3).

Table 5-3 Summary of chosen curves for PFS and OS

	Intention to treat (ITT) population	Stable disease (SD) population	Non-squamous (NS) population
Progression free survival	SATURN Kaplan-Meier	SATURN Kaplan-Meier	[redacted]
Overall survival	Log Logistic	Gamma	[redacted]

5.3.4 Population

Three separate economic evaluations reflecting three separate patient populations have been performed by the manufacturer (Table 5-4).

Table 5-4 Populations included in the economic evaluation

Erlotinib population	Comparator	Comments	SATURN BSC arm (%)	SATURN Erlotinib arm (%)	JMEN Pemetrexed arm (%)
ITT	Best supportive care	ITT population in the SATURN trial	451 (50.7%)	438 (49.3%)	-
Stable disease	Best supportive care	Main base case – this population derived most OS benefit in SATURN	252 (51.7%)	235 (48.2%)	-
Non-squamous	Pemetrexed	Non-squamous population from SATURN and JMEN trials	-	257 (57%)	325 (74%)

ITT= intention to treat; BSC= best supportive care; SD= stable disease; NS= non-squamous; OS= overall survival

5.3.5 Comparator technology

For the ITT and SD patient populations, the comparator technology in the economic evaluation is BSC, this is appropriate. For the non-squamous population, there is the possibility that pemetrexed may be approved by NICE therefore pemetrexed is the appropriate comparator in the economic evaluation.

5.3.6 Health related quality of life

The manufacturer reports that, in line with the NICE Guide to Methods,[41] the measurement of changes in HRQoL in PFS has been reported directly from patients in the SATURN[4] trial using the FACT-L. However, in order to comply fully with the NICE Guide to Methods,[41] two data transformations were performed by the manufacturer. Firstly, the FACT-L scores were transformed into EQ-5D (visual analogue score) using the methods outlined in Kind.[38] Since visual analogue scores are not recommended by the NICE Guide to Methods,[41] a further mapping from EQ-5D (visual analogue score) to EQ-5D (time-trade off) was applied using the methods outlined in Kind.[39] The manufacturer gives a description of the steps taken to perform both transformations.

Since no utility data for the progressed health state was collected in the SATURN[4] trial, the manufacturer carried out a literature review and used utility data from the 2008 quality of life publication by Nafees.[40] The Nafees[40] paper is a Lilly sponsored UK study. Health states in the Nafees[40] study were developed by five oncologists and five oncologist specialist nurses and the questions, using standard gamble techniques, were asked of 100 members of the general public. Table 5-2 shows the utilities used in the manufacturer's economic evaluations.

5.3.7 Resources and costs

The MS states that NHS resources were estimated in order to capture all relevant costs associated with the treatment of patients with NSCLC. Resource use and unit cost details are presented in Table 5-5 and total average per patient costs for the ITT, SD and non-squamous populations are presented in Table 5-6.

Table 5-5 Resource use and unit cost data sources used in the MS

Resource	Utilisation rate data source	Unit cost data source
Drug		
Erlotinib (150mg once daily)	SATURN trial	BNF 58[33] (including 14.5% reduction)
Pemetrexed 500mg/m^2/BSA=1.8m^2	JMEN trial	BNF 58[33]
Pemetrexed concomitant medication	JMEN trial	BNF 58[33]
Drug administration and preparation		
Erlotinib preparation	NCAT pathway	CPORT[34]
Pemetrexed administration	JMEN trial	NHS reference costs (2007-08) [37]
Pemetrexed preparation	NCAT pathway	CPORT[34]
Best supportive care (monthly)		
Progression free health state	Expert panel meeting (2006)	TA162[42] (2006) inflated to PSSRU 2009
Progressed health state	Expert panel meeting (2006)	TA162[42] (2006) inflated to PSSRU 2009
Adverse event		
Erlotinib: rash	SATURN trial	TA162[42] (2006) inflated to PSSRU 2009
Erlotinib: diarrhoea	SATURN trial	TA162[42] (2006) inflated to PSSRU 2009
Pemetrexed: neutropenia	JMEN trial	Pemetrexed maintenance MS 2009[7]
Pemetrexed: anaemia	JMEN trial	Pemetrexed maintenance MS 2009[7]
Post progression drug treatment		
Erlotinib patients	SATURN trial	BNF 58[33] or other published sources
Pemetrexed patients	Patients are assumed to have the same post progression drug options as placebo patients in SATURN	BNF 58[33] or other published sources

MS= manufacturer's submission; BNF= British National Formulary; BSA= body surface area; CPORT= Chemotherapy Online Planning Resource Tool; TA= Technology Appraisal; PSSRU= Personal Social Services Resource Use; TA= technology appraisal NCAT= National Action Cancer Team

Table 5-6 Total average per patient costs by population

	ITT		SD		Non-squamous	
Cost component (£)	**Erlotinib**	**Placebo**	**Erlotinib**	**Placebo**	**Erlotinib**	**Pemetrexed**
Mean cost of PFS	£8,543	£1,373	£8,466	£1,348	**£8,721**	**£23,724**
Costs of drug	£6,430	£0	£6,396	£0	**£6,617**	**£17,853**
Administration/pharmacy	£65	£0	£65	£0	**£69**	**£2,924**
Cost of supportive care in PFS	£2,036	£1,373	£1,995	£1,348	**£2,021**	**£2,923**
Cost of adverse events	£12	£0	£11	£0	**£15**	**£24.64**
Mean cost of progression	£16,569	£18,034	£15,662	£15,034	**£16,748**	**£16,840**
Mean total cost	£25,112	£19,407	£24,129	£16,382	**£25,470**	**£40,564**

ITT= intention to treat; SD= stable disease; NS= non-squamous; PFS= progression free survival

5.3.8 Perspective, time horizon and discounting

The economic evaluation was undertaken from the perspective of the NHS and Personal Social Services. The time horizon set was a lifetime horizon. In the base case analysis, the time horizon was assumed to be five years, after which most patients had died; the maximum evaluation time horizon in the model was 15 years. Both costs and benefits were discounted at 3.5% per annum.

5.3.9 Model validation

The MS states that the internal validation and debugging of the model was performed by Outcomes International. The following validation procedures were performed:

- Check of completeness of reported results (health outcomes, economic outcomes) as compared to other published economic evaluations targeting the same indication

- Execution of selected extreme tests to check the plausibility of model outcomes. Extreme testing was applied to the following parameters: treatment efficacy, adverse event costs, cost of study drugs and administration, discount rates, and health utilities.

5.3.10 Results included in manufacturer's submission

The manufacturer presents ICERs for 'erlotinib vs placebo' and 'pemetrexed vs erlotinib'. This implies that placebo (no maintenance treatment) is the current standard of NHS care - i.e. that pemetrexed has not been recommended by NICE. However, in the second comparison it implies that erlotinib is the current standard against which pemetrexed is to be compared. In the current, where a decision on pemetrexed is expected to be issued by NICE prior to consideration of the case for erlotinib, this is illogical and it is more appropriate to present results for "erlotinib vs pemetrexed". It happens that, since erlotinib is both less effective and less expensive than pemetrexed, the cost per QALY gained is the same for either comparison (since Q / C = -Q / -C). In the ERG's summary of corrections and amendments made to the submitted models (section 6.1), the ERG considers that relabelling the ICERs as "erlotinib vs pemetrexed" is more relevant to the decision problem being considered.

Base case results

Results for both the incremental cost per QALY gained as well as the cost per life year gained are presented in the MS for the ITT, SD and non-squamous populations (Table 5-7).

Table 5-7 Incremental cost per QALY gained and incremental cost per life year gained ratios

Cost-utility results (ITT)	Erlotinib	Placebo	Incremental	Cost per life year gained	Cost per QALY gained
Mean life years (yrs)	1.446	1.299	0.147	**£38,896**	**£55,219**
Mean QALYs	0.788	0.685	0.103		
Mean total cost	£25,112	£19,407	£5,706		
Cost-utility results (SD)	**Erlotinib**	**Placebo**	**Incremental**	**Cost per life year gained**	**Cost per QALY gained**
Mean life years (yrs)	1.385	1.108	0.277	**£27,968**	**£47,743**
Mean QALYs	0.750	0.587	0.162		
Mean total cost	£24,129	£16,382	£7,747		
Cost-utility results (NS)	**Pemetrexed**	**Erlotinib**	**Incremental**	**Cost per life year gained**	**Cost per QALY gained**
Mean life years (yrs)	NA	NA	NA	[redacted]	[redacted]
Mean QALYs	[redacted]	[redacted]	[redacted]		
Mean total cost	[redacted]	[redacted]	[redacted]		

ITT= intention to treat; SD= stable disease; NS= non-squamous; QALY= quality adjusted life year gained; NA= not available

5.3.11 Sensitivity analysis

One-way sensitivity analysis

The manufacturer has undertaken extensive one-way SA. The incremental cost-effectiveness results for the one-way SAs for the comparison of erlotinib vs placebo in the ITT and SD populations and pemetrexed vs erlotinib in the non-squamous population are reproduced here from the MS (Table 5-8). The associated tornado diagrams ranking the changes made, in terms of the impact on the ICER, are presented in the MS (MS, pg 161- 183). The MS states that the following parameters have the largest impact on the ICERs: changes in utility values, selection of parametric curves, cost of BSC disease progression and changes to the hazard ratio when comparing pemetrexed vs erlotinib in the non-squamous population.

Table 5-8 One-way sensitivity analysis

Sensitivity analyses	ITT (erlotinib vs placebo) ICERs	SD (erlotinib vs placebo) ICERs	[redacted]
Base case	£55,219	£47,743	[redacted]
PFS utility ↓ by 20%	£69,517	£54,624	[redacted]
PFS utility ↑by 20%	£45,799	£42,402	[redacted]
OS utility ↓ by 20%	£54,907	£51,560	[redacted]
OS utility ↑ by 20%	£55,534	£44,452	[redacted]
Pharmacy preparation ↓to minimum values	£54,797	£47,477	[redacted]
Pharmacy preparation ↑to maximum values	£55,282	£47,783	[redacted]
Cost of BSC PFS health state ↓ by 50%	£52,010	£45,749	[redacted]
Cost of BSC PFS health state ↑by 50%	£58,428	£49,737	[redacted]
Cost of BSC progressed OS health state ↓ by 50%	£55,611	£42,597	[redacted]
Cost of BSC progressed OS health state ↑by 50%	£54,827	£52,889	[redacted]
Cost of post-progression drug treatment ↓ by 50%	£61,914	£50,953	[redacted]
Cost of post-progression drug treatment ↑by 50%	£48,524	£44,533	[redacted]
Cost of treating AE ↓ by 50%	£55,161	£47,709	[redacted]
Cost of treating AE ↑ by 50%	£55,277	£47,776	[redacted]
Treatment dose ↓by 10%	£48,996	£43,801	[redacted]
Treatment dose ↑ by 10%	£61,441	£51,685	[redacted]
Pemetrexed drug administration cost lower quartile	N/A	N/A	[redacted]
Pemetrexed drug administration cost upper quartile	N/A	N/A	[redacted]
PFS HR (lower confidence limit)	N/A	N/A	[redacted]
PFS HR (upper confidence limit)	N/A	N/A	[redacted]
OS HR (lower confidence limit)	N/A	N/A	[redacted]
OS HR (upper confidence limit)	N/A	N/A	[redacted]
4 years time horizon	£57,083	£48,696	[redacted]
6 years time horizon	£54,038	£47,250	[redacted]
Gamma function for both PFS and OS	£68,185	£51,853	[redacted]
Log Logistic function for both PFS and OS	£61,853	£50,473	[redacted]
Log Normal function for both PFS and OS	£62,521	£50,129	[redacted]
Gompertz function for both PFS and OS	£55,435	£49,874	[redacted]
Weibull function for both PFS and OS	£55,583	£50,000	[redacted]
Exponential function for both PFS and OS	£52,855	£47,411	[redacted]

ITT= intention to treat; SD= stable disease; NS= non-squamous; ICER= incremental cost effectiveness ratio; BSC = best supportive care; OS= overall survival; PFS= progression free survival; AE= adverse event; N/A = not applicable

Probabilistic sensitivity analysis

The manufacturer also undertook PSA. In the MS, probabilistic cost-effectiveness results, including scatter plots, are presented for erlotinib vs placebo in the ITT (PSA= £55,464 per QALY gained) and SD (PSA= £45,270 per QALY gained) populations. Cost-effectiveness acceptability curves are presented for erlotinib vs placebo in the ITT and SD populations and for pemetrexed vs erlotinib in the non-squamous population. The manufacturer concludes the following:

1. The probability of erlotinib being cost effective compared to placebo at a threshold of £30,000 and £50,000 per QALY is 30% and 46% respectively for the ITT population
2. The probability of erlotinib being cost effective compared to placebo at a threshold of £50,000 per QALY is 55% for the SD population
3. There is a high degree of certainty that pemetrexed is not cost effective compared to erlotinib in the non-squamous population.

5.4 Assessment of the manufacturer's economic model

Table 5-9 shows closely the manufacturer's submitted economic evaluations accord with the requirements for a base case analysis set out in the NICE reference case checklist.[41] Table 5-10 summarises the ERG's appraisal of the economic evaluations conducted by the manufacturer using the Drummond 10-point checklist.[43]

The ERG's main criticism of the submitted economic models is that, although the models are structured around two health states in the form of a Markov structure, they are not Markov models; this means that there is no guarantee that post-progression survival (PPS) estimates will not take negative values, thus concerns are raised about the reliability of the results generated. Scrutiny of the submitted model/economic evaluation by the ERG has highlighted several weaknesses including: the manufacturer has not presented full results for the PSA as only pairwise comparisons were described in the MS; for the SD and non-squamous populations, the discount rate was applied on a daily basis after the first year, NICE prefers the use of an annual discount rate after the first year; the costs of erlotinib and pemetrexed were miscalculated; finally the ERGs is not confident that the utility values used in the model are the most appropriate.

Table 5-9 NICE reference case checklist[41]

Attribute	Reference case	Does the *de novo* economic evaluation match the reference case?
Comparator(s)		BSC in the SD and ITT populations Pemetrexed in the non-squamous populations
Perspective costs	NHS and Personal Social Services	The economic evaluation is carried out from the perspective of the NHS, no PSS costs are described in the MS
Perspective benefits	All health effects on individuals	Health effects to the individual are captured via QALYs
Form of economic evaluation	Cost-effectiveness analysis	Cost-effectiveness analysis
Time horizon	Sufficient to capture differences in costs and outcomes	Time horizon adopted is five years - this is appropriate
Synthesis of evidence on outcomes	Systematic review	No systematic review was undertaken. All survival data are derived from the SATURN and JMEN trials - this is appropriate
Outcome measure	Quality adjusted life years	QALYs used - this is appropriate
Health states for QALY	Described using a standardised and validated instrument	Utility (PFS) – original data collected in SATURN trial and transformed twice before being considered for use. Utility (progressed) taken from Nafees study – this study was not designed to capture the QoL of patients on maintenance therapy
Benefit valuation	Time-trade off or standard gamble	The Nafees QoL study used standard gamble techniques - this is appropriate
Source of preference data for valuation of changes in HRQoL	Representative sample of the public	The main QoL study (Nafees) was based on responses from 100 members of the general public. It is not clear how representative this sample is of the general public in the UK
Discount rate	An annual rate of 3.5% on both costs and health effects	Benefits and costs have been discounted using a rate of 3.5% - this rate has been applied incorrectly
Equity	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	All QALYs estimated by the model have the same weight
Sensitivity analysis	Probabilistic sensitivity analysis	The manufacturer only carried out PSA on pairwise comparisons and did not present a CEAF for the three treatment strategies (non-squamous population)

PSS= Personal Social Services; MS= manufacturer submission; RCT= randomised controlled trial; QALYs= quality adjusted life years; PSA= probabilistic sensitivity analysis; ERG= Evidence Review Group; HRQoL= health related quality of life; CEAF= cost-effectiveness acceptability frontier

Table 5-10 Critical appraisal checklist[43]

Item	Critical appraisal	ERG comment
Was a well-defined question posed in answerable form?	Yes	The manufacturer answered the decision problem set by NICE
Was a comprehensive description of the competing alternatives given?	Yes	The manufacturer's description of the comparator(s) was adequate
Was the effectiveness of the programme or services established?	Partially	The effectiveness of erlotinib maintenance therapy is established using data from the SATURN trial for specific patient populations
Were all the important and relevant costs and consequences for each alternative identified?	Yes	Key costs and outcomes were identified
Were costs and consequences measured accurately in appropriate physical units?	Partially	The BSA approach used to estimate pemetrexed costs was flawed. The ERG revised the costs of erlotinib in the economic model to take account of drug wastage. The ERG offered the manufacturer favourable estimates of PFS and OS for use in the model
Were the cost and consequences valued credibly?	Partially	The MS used 2007/08 NHS Ref Costs – the ERG used the recently available 2008/09 NHS Ref costs
Were costs and consequences adjusted for differential timing?	Partially	Costs and benefits were discounted, but the method of discounting was not applied correctly
Was an incremental analysis of costs and consequences of alternatives performed?	Yes	ICERs (incremental cost per QALY gained and incremental cost per life year gained) were presented for three different patient populations
Was allowance made for uncertainty in the estimates of costs and consequences?	Partially	The manufacturer only carried out PSA on pairwise comparisons and did not present a CEAF for the three treatment strategies (non-squamous population)
Did the presentation and discussion of study results include all issues of concern to users?	Yes	The results are presented and discussed in detail and an end of life treatment case has been proposed by the manufacturer

ERG= Evidence Review Group; QALY= quality adjusted life year; SA= sensitivity analysis; PSA= probabilistic sensitivity analysis; ICER= incremental cost-effectiveness ratio; BSA= body surface area; CEAF= cost-effectiveness acceptability frontier

5.5 Detailed critique of manufacturer's economic model

The manufacturer has submitted three distinct models, implemented as a series of Excel worksheets. The layout of the model(s) is generally clear and the tables are adequately labelled. Unless otherwise stated, the critique presented in this section applies to all submitted models. In section 5.4.1 to 5.4.6 major issues that are apparent from examination of the models are discussed. In section 5.4.7 to 5.4.9, less important issues apparent from examination of the models are discussed.

5.5.1 Model design

The submitted models are structured around two health states (before and after disease progression), and are presented in the form of a Markov structure. However, it is important to recognise that the models are not in fact Markov models since they are not governed by transition probabilities which impose restrictions on the way patients move from state to state. Instead, the modellers have chosen to employ parametric projection models of PFS and OS as the basic building blocks of the analysis, deriving the number of patients in PPS at any time as the difference between OS and PFS. It should be noted that in general there is no guarantee that the resulting PPS estimates will not take negative values, especially in later time periods. It is only possible for negative values to be definitively ruled out if the same projective function is used in both OS and PFS, if the estimates are jointly estimated from the trial data, and if a function is selected with the proportional hazards characteristic. This is not the case with any of the preferred manufacturer's base case models and in all three models negative PPS values appear in later cycles, raising concerns about the reliability of any results generated by the submitted models. The remedies to this problem are:

- to redesign the models as genuine Markov models, deriving time-dependent transition probabilities for disease progression and for death from both states; or

- to calibrate parametric models for PFS and PPS, and then validate the projected OS (= PFS + PPS) against the trial OS data.

5.5.2 Costs of erlotinib as maintenance therapy

The costs of erlotinib therapy (and associated dispensing costs) are calculated each month on the basis of the average number of patients remaining progression-free during the month. This is incorrect since drugs are administered at the beginning of each month to all eligible patients regardless of whether or not their disease progresses during the period. In addition, the costs for erlotinib therapy in the submitted models take no account of drug wastage. Erlotinib is dispensed in packs of 30 doses to be self-administered in the patient's home. Any part-used packs at the time when a patient

discontinues treatment for any reason are discarded and will not be dispensed to another patient. When wastage is estimated, based on the timing of disease progression seen within the SATURN[4] trial, the mean cost of treatment with erlotinib increases by 13.6%. In the stable disease population this increases the ICER for erlotinib vs placebo by £12,269 (+ 25%) per patient and in the non-squamous population the ICER decreases by £23,380 (-13%).

The manufacturer has introduced an optional feature which aims to reduce the cost of erlotinib by relating it to the average number of patients still progression free during each cycle. This is both incompatible with the way the drug is dispensed in packs, but also amounts to applying the mid-cycle correction twice over, when it is not applicable at all for this cost item. This feature is superseded by the ERG revised cost estimate described earlier which encompasses pack wastage.

The manufacturer has applied a cost reduction based on mean drug exposure calculations expressed in terms of the mean daily dose (in mg) per patient. These calculations cover both dose reductions (seen in 9.6% of patients) and dose interruptions (in 15.6% of patients). However, 82% of patients received the full daily dose throughout the trial. This method of cost adjustment does not reflect the impact on the NHS of such events. Dose reductions require the prescription of an additional pack of lower dose tablets, leading to the discarding of unused tablets of one sort or the other (depending on whether the change is temporary or permanent). In addition, the lower dosage tablets are priced at a higher per tablet price than the normal 150mg pack, therefore incurring higher NHS costs. In principle dose interruptions may lead to fewer packs of erlotinib being dispensed to some (but not all) such patients, but without more detailed information from the SATURN[4] trial it is not possible to quantify this effect, and how it compares with the additional cost of dose reductions. On balance the ERG considers that it is not appropriate to alter the estimated cost to the NHS of erlotinib for dose reductions or interruptions.

5.5.3 Costs of pemetrexed as maintenance therapy

Chemotherapy drug acquisition costs for pemetrexed maintenance therapy are presented within the submitted non-squamous model as a drug cost per month including mandatory concomitant medications. The authors have adopted an overall mean body surface area (BSA) value[44] of 1.8 m^2 but have not recognised that separate calculations are necessary for male and female UK patients nor have they incorporated the effect of the distribution of BSA values within the population on drug wastage. In addition the gender balance must also be aligned with that reported in the SATURN[4]trial population.

Taking account of these factors, the ERG has re-estimated the cost of pemetrexed drugs on the same basis as was used in the ERG's recent appraisal[29] of pemetrexed as maintenance therapy in patients with non-squamous advanced NSCLC, including concomitant medications. The net cost per patient month is estimated as £2,170.53, compared to the manufacturer's estimate of £2,207.87. The impact of this adjustment on the cost-effectiveness results for pemetrexed is quite modest. The mean incremental cost per patient is reduced by about £300, with reductions in ICERs of 1.5 – 2.0%.

5.5.4 Survival modelling and projection of OS and PFS

The results obtained with the submitted model depend upon projective modelling of OS and PFS beyond the trial period to estimate the lifetime gain in patient outcomes expected to arise from erlotinib maintenance therapy, compared with either normal monitoring of the patient's condition (equivalent to placebo in the SATURN[4] trial), or to pemetrexed maintenance therapy. The model authors have provided six parametric models for OS and PFS calibrated from the SATURN[4] trial data as a basis for projecting benefits for up to 15 years (base case 5 years).

The estimation of treatment benefits in the three submitted models raises some complex issues; quite different results may be obtained depending upon the assumptions and preferences of the modeller. In particular, it is important to consider which measures of patient benefit should be calculated, and which methods should be used in their estimation.

Choice of outcomes: The manufacturer has designed the economic model with PFS and OS as the primary measures of patient experience. However, the model structure implies that PFS and the time from disease progression to death (PPS) are the essential components which give rise to OS. In the submitted model, PPS is not estimated directly, but only as the difference between OS and PFS. It can be argued that since PPS data are available from the SATURN trial for as many patients as for OS, it would be logical to model and project PPS directly and then define OS as the sum of PFS and PPS. A related issue is the influence of subsequent courses of second-line chemotherapy on PPS;

since some patients die at progression, and others are not able to receive second-line chemotherapy, the therapeutic casemix of patients may differ between trial arms, and the patient experience with or without subsequent chemotherapy can also alter the overall mean PPS (and hence OS). In previous NSCLC appraisals this issue has been found to be important, and the ERG have therefore investigated its potential impact on estimated outcomes in this case, using analyses from the SATURN[4] data provided by the manufacturer.

Projection methods: It is common practice to attempt to fit standard parametric functions to empirical survival data, and routines are available within commercial statistical software for this purpose. However, little consideration is given to the assumptions implicit in such analyses, and whether they are appropriate to generating models with the express purpose of projecting beyond the available data. Curve-fitting, and the statistical tests used to assess relative 'goodness of fit' between candidate formulations are essentially descriptive, in that they relate solely to the extent of correspondence between the available data and the calibrated standard function. It cannot be presumed that such a mechanistic process will necessarily yield clinically or physiologically credible results when projected into the future. In addition, the normal practice of fitting curves to the whole of the available data may not be sensible, especially for analysis of clinical trials where 'in process' alterations in key aspects of patient care are an essential part of trial design. This inevitably conflicts with the standard mathematical functions which are founded on smooth continuous functions, without abrupt alterations in trends. In the real world several aspects of RCTs may contribute to difficulties achieving a good 'model fit' arising from the inappropriateness of the approach taken including the following:

a) Trial inclusion/exclusion criteria frequently include direct or indirect stipulations which minimise or remove altogether the likelihood of specific events occurring in the first few weeks of the trial.

b) The action of a prescribed drug takes time to achieve its full effect, partly due to the pharmacokinetic/dynamic profile of the drug, and partly due to the time required for the active agent to achieve its full effect at the target site(s). Conversely, when the period of active treatment comes to an end its effects may dissipate gradually over several weeks. This is also relevant where no 'washout' period is allowed between prior courses of treatment (first-line chemotherapy in this case) and commencing the trial interventions.

c) Additional confounding is potentially introduced by the availability of subsequent courses of active chemotherapy which may further complicate the dynamic nature of the event hazard rate following disease progression.

d) There is also the possibility that the patient population is essentially heterogeneous in relation to the event risk of interest, leading to progressive survivor bias as members of one subgroup suffer death at a faster rate than other patients.

As a consequence of these influences, it is not surprising that fitting a standard parametric survival function to the full clinical trial dataset rarely produces a satisfactory correspondence to the calculated survival trajectory. Moreover, since the reliability of fit at later periods is increasingly sensitive to diminishing patient numbers, calibrating a parametric function from the full patient data may be a particularly unsatisfactory basis for projecting events beyond the trial data collection period.

An alternative approach is to examine the cumulative hazard plot from a Kaplan-Meier survival analysis on the trial data with a view to identifying indications of the long-term trends which may persist after any short-term transient effects have dissipated.[45] This is often more likely to reveal the underlying temporal dynamic which should inform projection. The ERG has followed this approach in preparing their own estimates of PFS, PPS and OS to compare with these presented by the manufacturer, since it is clear that the findings of the economic assessment in this case could be sensitive to projection methods.

In addition, the ERG has considered the importance of (c) above by carrying out separate projections of PPS for patients who did and did not receive second-line chemotherapy following disease progression.

One further point is worthy of note. It is common practice for statisticians to fit a parametric function simultaneously to both arms of a clinical trial, distinguishing between the arms by allowing only one of the functions parameters to differ by type of treatment. This is generally applied to all the results to be expressed, or later modified on the assumption of proportionality of hazards. However, this is a strong assumption to make, particularly where there is good reason to expect the modes of action of treatment in the arms of the trial to be qualitatively different (as in the case of a comparison between a disease-modifying agent and a placebo/no treatment). The ERG has not made this assumption in arriving at their own projection estimates.

Progression-free survival estimation: Table 5-11 details mean PFS estimates for the Kaplan-Meier analysis (area under the Kaplan-Meier plot), the six parametric models fitted by the manufacturer, and the ERG's own analysis. In the ITT and stable disease populations, the manufacturer did not consider any of the parametric models sufficiently accurate to employ in the economic analysis, and therefore chose to use the AUC value instead. The ERG models are based on the observation that in each case a steady linear hazard trend was observed after the first 12 months. Therefore the AUC data were

used for the first year, followed by projection using an exponential curve calibrated by linear regression on the cumulative hazard beyond 12 months. In each case the ERG's projections are similar to the AUC data and the manufacturer's preferred option.

Table 5-11 Estimated mean PFS using various models (months)

ITT population	At 5 years			Lifetime		
Method / model	Erlotinib	Placebo	Difference	Erlotinib	Placebo	Difference
AUC	**5.690**	**3.814**	**1.876**	**5.690**	**3.814**	**1.876**
ERG model	**5.710**	**3.897**	**1.813**	**5.721**	**3.904**	**1.817**
Weibull	5.547	3.805	1.742	5.547	3.805	1.742
Exponential	5.556	3.815	1.741	5.556	3.815	1.741
LogLogistic	4.883	3.737	1.146	5.003	3.807	1.197
LogNormal	5.049	3.824	1.224	5.059	3.827	1.232
Gompertz	5.556	3.815	1.741	5.556	3.815	1.741
Gamma	5.255	4.378	0.877	5.473	4.529	0.944
SD population	At 5 years			Lifetime		
Method / model	Erlotinib	Placebo	Difference	Erlotinib	Placebo	Difference
AUC	**5.564**	**3.744**	**1.821**	**5.564**	**3.744**	**1.821**
ERG model	**5.566**	**3.739**	**1.826**	**5.571**	**3.742**	**1.828**
Weibull	5.470	3.759	1.711	5.470	3.759	1.711
Exponential	5.494	3.766	1.727	5.494	3.766	1.727
LogLogistic	5.046	3.718	1.328	5.163	3.780	1.383
LogNormal	5.159	3.732	1.427	5.168	3.735	1.434
Gompertz	5.494	3.766	1.727	5.494	3.766	1.727
Gamma	5.425	4.094	1.331	5.576	4.174	1.402
NS population	At 5 years			Lifetime		
Method / model	Erlotinib	Placebo	Difference	Erlotinib	Placebo	Difference
AUC	5.718	3.599	2.120	5.718	3.599	2.120
ERG model	**5.675**	**3.453**	**2.222**	**5.683**	**3.453**	**2.231**
Weibull	5.622	3.537	2.085	5.622	3.537	2.085
Exponential	5.623	3.558	2.065	5.623	3.558	2.065
LogLogistic	4.754	3.725	1.029	4.877	3.800	1.077
LogNormal	4.957	3.756	1.201	4.968	3.760	1.208
Gompertz	**5.623**	**3.558**	**2.065**	**5.623**	**3.558**	**2.065**
Gamma	5.103	4.466	0.636	5.344	4.653	0.691

AUC = area under the Kaplan-Meier survival plot; SD= stable disease; NS= non-squamous; ERG= evidence review group. Models shown in bold type are preferred options for either the manufacturer or the ERG

Post-progression survival estimation: In response to a request from the ERG, the manufacturer provided information relating to the survival of patients following disease progression, stratified by whether or not patients received subsequent second-line chemotherapy. These data have been analysed to identify appropriate survival models to project the remaining lifetime of patients following disease progression. In each case the area under the Kaplan-Meier curve was used for the first 12 months, with projection applied only to the later period. For the ITT populations and stable disease populations Weibull models were found to provide the best correspondence to the trial data, whereas in the non-squamous population an exponential model with a transient effect was employed to reflect early risk suppression evident from the data. Separate model parameters have been estimated for each combination of patient population, trial arm and use of second-line chemotherapy, with the exception of patients receiving subsequent chemotherapy in the non-squamous population

where no meaningful difference was observed by trial arm so that a joint parameter was estimated. These projections were used to produce estimated long-term post-progression survival estimates, weighted by the relative proportions of patients who did and did not receive second-line chemotherapy in each population/treatment arm combination (Table 5-12). This suggests that some patients experience extended benefit from erlotinib maintenance therapy after disease progression in the ITT and stable disease populations, but not in the non-squamous population.

Table 5-12 Estimated mean PPS (months)

ITT population	**Lifetime**		
Sub-group	Erlotinib	Placebo	Difference
Received 2nd line chemotherapy	15.755	13.952	1.803
No 2nd line chemotherapy	6.017	5.936	0.081
Weighted mean	13.350	12.008	1.342
SD population	Lifetime		
Sub-group	Erlotinib	Placebo	Difference
Received 2nd line chemotherapy	13.948	11.060	2.888
No 2nd line chemotherapy	5.570	4.465	1.105
Weighted mean	11.656	9.321	2.335
NS population	Lifetime		
Sub-group	Erlotinib	Placebo	Difference
Received 2nd line chemotherapy	14.997	14.997	0.000
No 2nd line chemotherapy	4.259	3.310	0.949
Weighted mean	12.908	12.437	0.471

ITT= intention to treat; NS=non-squamous

Overall survival estimation: The ERG carried out projection modelling for OS directly from the trial data, and also indirectly, combining the PFS and PPS estimates shown above. In principle the ERG prefers the latter method, since it avoids the confounding of effects arising in two distinct phases of treatment, and also allows case-mix differences related to post-progression therapies to be incorporated. This option suggests rather greater patient benefits from use of erlotinib compared to the manufacturer's preferred models, except in the case of the non-squamous population (Table 5-13).

Table 5-13 Estimated mean OS using various models (months)

ITT population	At 5 years			Lifetime		
Method / model	Erlotinib	Placebo	Difference	Erlotinib	Placebo	Difference
AUC	15.944	13.852	2.092	15.944	13.852	2.092
ERG OS model	17.816	14.603	3.213	18.526	14.742	3.783
ERG PFS+PPS	**18.556**	**15.674**	**2.882**	**19.071**	**15.912**	**3.159**
Weibull	16.946	14.385	2.561	17.046	14.412	2.634
Exponential	18.175	15.223	2.952	18.976	15.550	3.426
LogLogistic	**17.993**	**16.120**	**1.873**	**20.771**	**18.324**	**2.448**
LogNormal	18.206	16.226	1.980	20.247	17.681	2.566
Gompertz	16.873	14.326	2.546	16.910	14.336	2.574
Gamma	17.698	15.473	2.225	18.667	16.050	2.617

SD population	At 5 years			Lifetime		
Method / model	Erlotinib	Placebo	Difference	Erlotinib	Placebo	Difference
AUC	14.176	12.582	1.595	14.176	12.582	1.595
ERG OS model	16.521	12.948	3.573	17.117	12.932	4.186
ERG PFS+PPS	**17.178**	**13.063**	**4.115**	**17.227**	**13.064**	**4.163**
Weibull	16.430	12.799	3.632	16.499	12.806	3.693
Exponential	17.695	13.373	4.322	18.399	13.533	4.865
LogLogistic	17.457	14.393	3.064	19.855	15.985	3.870
LogNormal	17.614	14.206	3.409	19.218	15.032	4.186
Gompertz	17.252	14.106	3.146	17.302	14.115	3.186
Gamma	**17.163**	**13.631**	**3.532**	**17.948**	**13.943**	**4.005**

NS population	At 5 years			Lifetime		
Method / model	Erlotinib	Placebo	Difference	Erlotinib	Placebo	Difference
AUC	15.243	13.814	1.429	15.243	13.814	1.429
ERG OS model	17.865	13.944	3.921	18.406	14.011	4.395
ERG PFS+PPS	**18.391**	**15.696**	**2.695**	**18.591**	**15.890**	**2.701**
Weibull	17.751	14.665	3.086	17.937	14.711	3.226
Exponential	18.913	15.428	3.485	19.883	15.780	4.103
LogLogistic	**19.061**	**16.485**	**2.575**	**22.700**	**19.188**	**3.512**
LogNormal	19.334	16.528	2.806	22.216	18.387	3.829
Gompertz	17.690	14.618	3.071	17.778	14.640	3.138
Gamma	18.768	15.811	2.957	20.289	16.634	3.655

AUC = area under the Kaplan_Meier survival plot. Models shown in bold type are preferred options for either the manufacturer or the ERG; NS=non-squamous

Survival estimates for pemetrexed maintenance therapy: The manufacturer of erlotinib estimates survival times for pemetrexed therapy in the case of the non-squamous population by deriving HRs from an indirect comparison of the SATURN[4] and JMEN[5] clinical trials. The ERG has compared the placebo arms of these two arms for both PFS and OS, and found that they are very closely matched; the Kaplan-Meier curves follow very similar trajectories with frequent crossovers. It seems reasonable therefore to conclude that there is little benefit in computing hazard ratios. Instead we have adopted the survival functions previously developed by the ERG for the recent appraisa[29]l of pemetrexed maintenance therapy (based on patient data from the JMEN[5] trial) for direct comparison with the erlotinib estimates shown above, thus avoiding the necessity of assumptions of proportionality of hazards.

5.5.5 Costs of second-line chemotherapy

The manufacturer's models include estimates of the cost of subsequent courses of chemotherapy through an elaborate analysis of recorded treatments in the SATURN[4] trial. The total costs are then converted to average monthly costs which are added to the state-based monthly cost of BSC. This approach is problematic for a number of reasons:

- a wide range of licensed and experimental treatments are included which are not recommended for use in UK, where only erlotinib and docetaxel are approved by NICE;

- erlotinib is included in the list of drugs used in both arms of the trial, whereas use of erlotinib for maintenance therapy would normally preclude its further use in second-line treatment;

- the conversion of costs to a monthly average and then reapplying them to survival estimates to recover cost totals risks generating unintended distortions and bias in the resulting incremental costs.

Since the agreement between NICE and the manufacturer for the supply of erlotinib in second-line use was based on net equivalent costs for erlotinib and docetaxel, a simpler approach is to use a fixed cost per course of treatment multiplied by the proportion of progressed patients receiving second-line chemotherapy in the SATURN[4] trial, spread pro-rata over the post-progression survivors in each cycle.

5.5.6 Utility values

The SATURN[4] trial did not collect data using a generic health-utility instrument, using only a disease specific quality of life instrument (FACT-L). The manufacturer was only able to identify a mapping algorithm to associate QoL data with EuroQol visual analogue scores (VAS), which are a measure of patient general well-being and not designed to act as a health-utility measure. A further mapping algorithm was then employed to convert VAS scores to EQ-5D scores in order to supply values for use in the submitted models. Since each step in this process involves substantial uncertainty in addition to that involved in the original data collection, the ERG considers that the credibility of the specific values obtained cannot be deemed very high. As an alternative, the ERG has employed a source[40] used in previous NSCLC appraisals, which then provides consistency with the appraisal carried out which considered pemetrexed for maintenance therapy in the non squamous population. The ERG has applied the same method used previously to incorporate the incidence of serious adverse events into the basic utility estimates for the PFS state. These are 0.6732 for erlotinib, 0.6568 for pemetrexed and 0.6628 for placebo/BSC; the Nafees[40] value for PPS is 0.53. Applying these to the manufacturer's models has the effect of generally increasing the QALYs obtained in each arm of the comparison and also the incremental QALYs gained for the two maintenance therapies compared to no treatment.

5.5.7 Discounting method

Costs and outcome are discounted in two of the submitted models (stable disease and non-squamous) on a continuous daily basis after the first year from the time of randomisation, but not in the ITT model where discounting is applied annually. It is conventional in the UK to discount annually (i.e. no discounting in the first year, followed by use of a single discount factor for each successive twelve month period) to match the annual publication of price base information (e.g. NHS Reference Costs). Amending the method of discounting in this way leads to minor alterations increasing both incremental costs and QALYs, so that the ICER in reduced by less than £200 per QALY gained in the stable disease population.

5.5.8 Time horizon of evaluation

In principle, cost-effectiveness should be assessed over the full remaining lifetime of the patient population. The manufacturer has chosen to present results for a period of five years. However, the model is designed to allow extended horizons of up to 15 years. Altering this parameter to its maximum increases both incremental costs and outcomes by small amounts, leading to a minor (about 2.5%) reduction in cost-effectiveness ratios.

5.5.9 Updating unit costs

NHS reference costs[46] for 2008/9 have recently been released allowing some unit costs in the submitted models to be updated. Those costs principally affected are those for the monthly cost of BSC in both health states and the administration cost of pemetrexed therapy. The impact of these changes is relatively small and varies between models and depending on the methods of survival projection selected.

5.6 Summary of ERG model critique

The ERG identified a potentially serious problem with the design of the submitted models, in that the use of independent projective survival functions for PFS and OS allows negative post-progression patient numbers to be generated, compromising both lifetime cost and outcome estimates. It has not been possible to carry out a thorough investigation of the extent of the errors introduced by this problem, but all economic results produced by the original manufacturer's model must be treated with caution. An alternative method of estimating PFS and PPS has been used by the ERG which overcomes this problem and should therefore be considered more reliable.

The derivation of patient costs used in the model was found to be inadequate in three respects:

- the mean cost to the NHS of treating patients with erlotinib was found to have been seriously under-estimated
- the mean cost of providing second-line chemotherapy was under-estimated
- other unit costs could be revised using more up-to-date NHS data.

In each case the effect of modifying the model is to increase the incremental cost of erlotinib pus BSC compared to BSC alone, so that the ICER is increased substantially.

Several other modelling and data problems were identified (time horizon, discounting, utility values and projective modelling of PFS and OS), all of which increased the incremental patient benefit more than any changes in incremental cost, so that in each case the ICER was reduced in favour of erlotinib.

However, the issues related to costs have a dominant effect, so that the ICER in the SD population increases to about £60,000 per QALY gained (see Table 6-2).

6 ADDITIONAL WORK UNDERTAKEN BY THE ERG

Table 6-1 to Table 6-3 provide a detailed set of results for the base case analyses obtained from the manufacturer's models, indicating the individual effects of each of the amendments/corrections implemented by the ERG, together with final revised base case results combining all the changes. For each population, the original ICER has increased substantially for estimates of the cost effectiveness of erlotinib vs placebo, but reduced for estimates of the cost effectiveness of erlotinib vs pemetrexed.

In all cases the revised results obtained by the ERG indicate high cost-effectiveness ratios of around £60,000 per QALY gained or greater. The major contributions to these changes were the inclusion of wastage in the acquisition cost of erlotinib, and the cost of second-line chemotherapy.

It has not been possible to carry out PSA of these results because the complexity of deriving all the necessary measures of uncertainty and reprogramming the models was not possible within the time available. (It should be noted that the manufacturer only carried out PSA on pairwise comparisons, and did not present a cost-effectiveness acceptability frontier for the three treatment strategies for the non-squamous population).

Table 6-1Effect of corrections/amendments made by the ERG to submitted ITT model for the base case analysis

ITT population

Erlotinib vs placebo

	Costs per patient			QALYs per patient				Difference from initial base case		
	Erlotinib	Placebo	Increment	Erlotinib	Placebo	Increment	ICER	Inc cost	Inc QALY	ICER
Manufacturer base case	**£25,112**	**£19,407**	**£5,706**	**0.7881**	**0.6848**	**0.1033**	**£55,219**	-	-	-
Extended time horizon	£28,063	£21,939	£6,124	0.8704	0.7501	0.1203	£50,899	£419	0.0170	-£4,320
Discounting logic corrected	£25,267	£19,543	£5,723	0.7928	0.6886	0.1042	£54,904	£18	0.0009	-£315
Cost of erlotinib corrected	£27,231	£19,407	£7,825	0.7881	0.6848	0.1033	£75,727	£2,119	0.0000	£20,508
Cost of 2nd line CTX corrected	£26,336	£19,271	£7,066	0.7881	0.6848	0.1033	£68,382	£1,360	0.0000	£13,163
Unit costs updated	£26,991	£21,204	£5,788	0.7881	0.6848	0.1033	£56,013	£82	0.0000	£794
Revised utility values	£25,112	£19,407	£5,706	0.8336	0.7307	0.1029	£55,431	£0	-0.0004	£212
ERG PFS model	£24,946	£19,248	£5,698	0.7914	0.6873	0.1040	£54,768	-£7	0.0007	-£451
ERG OS model	£25,151	£17,647	£7,504	0.7892	0.6395	0.1498	£50,103	£1,799	0.0465	-£5,116
Revised base case	**£31,035**	**£19,753**	**£11,282**	**0.8677**	**0.6898**	**0.1778**	**£63,440**	**£5,577**	**0.0745**	**£8,221**

Table 6-2 Effect of corrections/amendments made by the ERG to submitted SD model for the base case analysis

Stable disease population

Erlotinib vs placebo

	Costs per patient			QALYs per patient				Difference from initial base case		
	Erlotinib	Placebo	Increment	Erlotinib	Placebo	Increment	ICER	Inc cost	Inc QALY	ICER
Manufacturer base case	**£24,129**	**£16,382**	**£7,747**	**0.7497**	**0.5875**	**0.1623**	**£47,743**	-	-	-
Extended time horizon	£25,001	£16,771	£8,230	0.7739	0.5972	0.1768	£46,557	£483	0.0145	-£1,186
Discounting logic corrected	£24,266	£16,476	£7,790	0.7538	0.5900	0.1638	£47,559	£43	0.0015	-£184
Cost of erlotinib corrected	£26,119	£16,382	£9,738	0.7497	0.5875	0.1623	£60,012	£1,991	0.0000	£12,269
Cost of 2nd line CTX corrected	£25,431	£16,659	£8,772	0.7497	0.5875	0.1623	£54,061	£1,025	0.0000	£6,318
Unit costs updated	£25,918	£17,872	£8,046	0.7497	0.5875	0.1623	£49,584	£299	0.0000	£1,842
Revised utility values	£24,129	£16,382	£7,747	0.7998	0.6284	0.1714	£45,197	£0	0.0091	-£2,545
ERG PFS model	£23,954	£16,460	£7,493	0.7505	0.5863	0.1642	£45,649	-£253	0.0019	-£2,094
ERG OS model	£23,803	£15,672	£8,132	0.7407	0.5698	0.1709	£47,574	£385	0.0087	-£169
Revised base case	**£29,344**	**£17,745**	**£11,599**	**0.8075**	**0.6120**	**0.1955**	**£59,336**	**£3,852**	**0.0332**	**£11,593**

Table 6-3 Effect of corrections/amendments made by the ERG to the submitted non-squamous model for the base case analysis

[redacted]

6.4

7 END OF LIFE CRITERIA

7.1 Introduction

This section provides an overview and critique of the manufacturer's case for erlotinib as an end of life maintenance treatment for patients with NSCLC who have not progressed after first-line chemotherapy. The NICE end of life treatment criteria[6] has three key points:

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

7.2 Application of end of life treatment criteria

The NICE end of life criteria[6] is discussed below for the case of erlotinib as a maintenance therapy for patients with NSCLC. The manufacturer makes the case for erlotinib as a maintenance therapy for (i) patients with SD after first-line chemotherapy and (ii) patients with non-squamous disease who have not progressed after first-line chemotherapy.

7.2.1 Patient life expectancy of less than 24 months

The manufacturer makes the case that the OS of untreated patients with NSCLC is 11 months. The ERG agrees with the manufacturer that the life expectancy of patients with stage IIIB/IV NSCLC is likely to be less than 24 months for all patient populations.

7.2.2 Life extension of at least three months

The ERG carried out projection modelling for OS as described in section 5.5.4. The ERG agrees with the manufacturer that for patients with SD there is an expected life extension of at least three months. The ERG estimates a survival advantage for erlotinib patients of approximately 4.2 months which is higher than the estimate of 3.3 months described in the MS.

For patients with non-squamous histology, the ERG's revised OS benefit is estimated to be less than three months (2.7 months); this estimate is lower than the 3.2 months described in the MS.

For the FAS/ ITT population, the ERG estimates a survival advantage of just over 3.1 months for patients treated with erlotinib vs placebo; this estimate is much higher than the one month survival advantage that is described in the MS. However, the ERG highlights that the FAS/ITT group includes patients with squamous disease for whom treatment with erlotinib does not appear to offer a statistically significant OS advantage.

7.2.3 Licensed for a small population

As part of the end of life proposal the manufacturer presents data on the total number of patients eligible for treatment with erlotinib; this total is made up of patients receiving erlotinib as a maintenance treatment (the numbers vary according to the licence awarded) as well as patients receiving erlotinib as a second-line treatment. In the MS the total number of eligible patients varies depending on whether or not erlotinib is licensed as a maintenance treatment for patients with stable disease who have not progressed after first-line chemotherapy (n=2965) or as a maintenance treatment for all patients who have not progressed after first-line chemotherapy (3,888). The ERG notes that the size of the market share for the manufacturer will be affected by the outcome of the NICE appraisal for pemetrexed as a maintenance therapy; if pemetrexed is approved then the erlotinib market share will probably be smaller than if pemetrexed is not approved.

From the sources of information identified by the manufacturer as being relevant, the ERG was not able to verify or refute the size of the total patient population as described in the MS. However, given the nature of the disease and the limited availability of treatments, the ERG is of the opinion that, irrespective of the wording of the EMA licence, erlotinib will be licensed for what is considered by NICE to be a small patient population.

7.2.4 End of life criteria: ERG conclusion

The ERG is of the opinion that the manufacturer has met the criteria[6] set out by NICE for consideration of erlotinib as a maintenance therapy for patients with stable disease who have not progressed after first-line chemotherapy. However, the ERG is of the opinion that, based on the OS benefit demonstrated in the SATURN[4] trial, the manufacturer has not met the (life extension by three months) criteria[6] set out by NICE for consideration of erlotinib as a maintenance therapy for patients with non-squamous histology who have not progressed after first-line chemotherapy.

8 DISCUSSION

NOTE: This section was revised on March 22, 2010 to take into consideration the notification[1] of the approval of erlotinib maintenance therapy by the EMA for the SD patient population.

8.1 Summary of clinical and cost-effectiveness issues

The manufacturer presents a case for the use of erlotinib as a maintenance treatment for patients with NSCLC whose disease has not progressed following four cycles of first-line platinum-based chemotherapy. The systematic review carried out by the manufacturer identified a single relevant RCT (SATURN[4]) which compared the use of erlotinib plus BSC vs placebo plus BSC. The SATURN[4] trial was a large, double-blind, multi-centred, placebo controlled trial (n=889). The ERG has highlighted a number of concerns related to the conduct of the trial (including reliance on *post hoc* analyses, choice of stratification factors and security of blinding) and noted the limited generalisability of the results to the UK population (large number of patients from Eastern Europe and Asia and limited comparability of first- and second-line treatments).

The EMA licence[1] stipulates that erlotinib is approved for "...patients with stable disease after four cycles of standard platinum-based first-line chemotherapy", as per the first-line chemotherapies administered in the SATURN[4] trial. In the UK, pemetrexed has recently been approved as a first-line treatment for patients with non-squamous NSCLC, however the ERG notes that none of the patients in the SATURN[4] trial received pemetrexed as a first-line therapy.

The manufacturer also used clinical data from the JMEN[5] trial to indirectly compare erlotinib with pemetrexed as a maintenance treatment for patients with non-squamous NSCLC. Again, the ERG is concerned with the limited generalisability of the JMEN[5] results to a UK population (no trial centres in the UK, one third of patients were of Asian origin, not all second-line chemotherapies administered in the UK, none of the patients received pemetrexed or vinorelbine).

The MS presents data indicating a statistically significant improvement in PFS and OS in the ITT, SD and non-squamous patient populations (erlotinib plus BSC vs placebo plus BSC). Limited clinical data were included in the MS related to the SD population and although additional data were provided via the clarification process it is important to note that the analysis of this dataset was *post hoc*. In the SD population the HR for PFS (95%CI) is 0.68 (0.56, 0.83) and for OS is 0.72 (0.59, 0.89). Only the baseline demographics of this group of patients are known; histology or stage of disease is not reported in the MS.

The ERG notes that BSC may not be the only maintenance treatment option for those patients with non-squamous disease; the treatment alternatives for patients with non-squamous disease are dependent on the result of this current appraisal and the outcome of the ongoing NICE appraisal[3] related to the use of pemetrexed as maintenance treatment for patients with NSCLC who have not progressed following four cycles of platinum-based chemotherapy. The outcome of the appraisal of pemetrexed as maintenance therapy is not yet in the public domain and the ERG cannot speculate on the outcome or its implications for the treatment of patients with NSCLC.

The ERG's view is that the manufacturer has met the end of life criteria[6] set out by NICE for consideration of erlotinib as maintenance therapy for patients with SD who have not progressed after first-line platinum-based chemotherapy.

The manufacturer has not been able to clarify the effect of EGFR status on a patient's response to treatment with erlotinib. The trial used a range of different EGFR tests, however high rates of indeterminate and missing samples means that there is not enough evidence to allow meaningful interpretation of the data collected. The MS concludes that there is no need to test the EGFR status of patients before administering erlotinib; the ERG is unaware of any data to contradict this.

The ERG believes that the economic model(s) submitted by the manufacturer is flawed and is therefore concerned about the validity of the results presented in the MS. The ERG offers a solution to the structural problems identified in the manufacturer's model; however, the ERG's revised ICERs are estimated to be approximately £63,000 per QALY gained (erlotinib vs placebo) for the ITT population, £60,000 per QALY gained for the SD population (erlotinib vs placebo) and £96,000 per QALY gained for patients with non-squamous histology (erlotinib vs pemetrexed).

8.2 Implications for research

The EMA licence[1] states that erlotinib is approved for those patients with SD after four cycles of standard platinum-based first-line chemotherapy. Limiting the patient population in this way has an impact on the type of research that needs to be conducted. Within the SD patient population, the effect of histology on OS outcomes will need to be investigated.

In addition, as that none of the patients in the SATURN[4] trial received pemetrexed as first-line therapy, the clinical effectiveness of erlotinib as maintenance therapy in this population of patients is not known and needs to be investigated.

The limited number of UK patients in the SATURN[4] trial means that efforts should be made to encourage UK patients to participate in future lung cancer trials in order to improve the generalisability of trial results to UK patients.

Given that maintenance therapy is a new addition to the treatment pathway there is a need to bring together all of the available clinical- and cost-effectiveness data and also to encourage, where appropriate, head to head comparisons of therapies.

The ongoing debate about the importance of EGFR status continues and further investigation in this area is required; including specification of the most appropriate tests to be used to ascertain EGFR status in the UK population.

9 REFERENCES

1. European Medicines Agency. Summary of opinion - Tarceva (erlotinib). 2010; Available from: http://www.ema.europa.eu/pdfs/human/opinion/TarcevaII_8608010en.pdf.
2. Biospace. US FDA panel rejects the use of Tarceva. 2009; Available from: <http://www.biospace.com/news_story.aspx?NewsEntityId=165836&Source=TopBreaking>.
3. National Institute for Health and Clinical Excellence. Lung cancer (non-small-cell) - pemetrexed (maintenance): final scope. 2009; Available from: http://guidance.nice.org.uk/TA/WaveCRS2/47.
4. Roche Ltd. Erlotinib - Clinical research report - No. 10314602009.
5. Ciuleanu T, Brodowicz T, Kim J, Krzakowski M, Wu Y, Bover I, et al. Maintenance pemetrexed plus best supportive care versus placebo plus best supportive care for non-small-cell lung cancer: a randomised, double-blind, phase 3 study. Lancet. 2009; 374(1432-1440).
6. National Institute for Health and Clinical Excellence. Supplementary Advice to the Appraisal Committees - End of Life Treatments. 2009; Available from: http://www.nice.org.uk/aboutnice/howwework/devnicetech/endoflifetreatments.jsp?domedia=1&mid=88ACDAE5-19B9-E0B5-D422589714A8EC6D.
7. Eli Lilly. Single Technology Appraisal - Pemetrexed in the maintenance treatment of non-small cell lung cancer. Manufacturer submission. 2009 [January]; Available from: <http://www.nice.org.uk/guidance/index.jsp?action=download&o=46610>.
8. National Institute for Health and Clinical Excellence. Pemetrexed for the first-line treatment of non-small-cell lung cancer. 2008 [cited 2010 January]; Available from: http://www.nice.org.uk/nicemedia/pdf/TA181Guidance.pdf.
9. Cancer Research UK. CancerStats: Lung Cancer Incidence - UK. 2006 [January]; Available from: <http://info.cancerresearchuk.org/cancerstats/types/lung/incidence/>.
10. NHS Information Centre. National Lung Cancer 2007 Audit. 2007; Available from: <http://www.ic.nhs.uk/webfiles/Services/NCASP/audits%20and%20reports/7089_Lung_Cancer_V5.pdf.
11. NHS Information Centre. National Lung Cancer 2008 Audit 2009 [cited 2010 January]; Available from: <http://www.ic.nhs.uk/webfiles/Services/NCASP/audits%20and%20reports/IC23090809_NHS_IC_Lung_Cancer_AUDIT_2009_Interactive_for_web.pdf
12. Roche Ltd. Erlotinib for the 1st line maintenance treatment of non-small cell lung cancer - submission to NICE Single Technology Appraisal Process 2010.
13. Brodowicz T, Krazkowski M, Zwitter M, Tzekova V, Ramlau R, for Central European Cooperative Oncology Group (CECOG), et al. Cisplatin and gemcitabine first-line chemotherapy followed by maintenance gemcitabine or best supportive care in advanced non-small cell lung cancer: A phase III trial. Lung Cancer. 2006; 52:155-63.
14. Fidias P, Dakhil S, Lyss A, Loesch D, Waterhouse D, Bromund J, et al. Phase III study of immediate compared with delayed docetaxel after front-line therapy with gemcitabine plus carboplatin in advanced non-small-cell lung cancer. Journal of Clinical Oncology. 2009; 27(4):591-98.
15. Westeel V, Quoix E, Moro-Sibilot D, Mercier M, Breton J, for the French Thoracic Oncology Collaborative Group (GCOT), et al. Randomized study of maintenance vinorelbine in responders with advanced non-small-cell lung cancer. Journal of the National Cancer Institute. 2005; 97:499-506.

16. Belani C, Brodowica T, Ciuleanu J, Kim M, Krazkowski E, Laack Y, et al. Maintenance pemetrexed (Pem) plus best supportive care (BSC) versus placebo (Plac) plus BSC: A randomized phase III study in advanced non-small cell lung cancer (NSCLC). Journal of Clinical Oncology. 2009; 27:18s.
17. National Institute for Health and Clinical Excellence. Lung cancer (non-small-cell, first line) - gefitinib: appraisal consultation document. NICE; 2010 [February]; Available from: http://www.nice.org.uk/guidance/index.jsp?action=article&o=47235.
18. Food and Drug Administration. Questions for Oncologic Drugs Advisory Committee Meeting - Tarceva - December 2009. 2010 [January]; Available from: http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/OncologicDrugsAdvisoryCommittee/UCM195716.pdf.
19. Miller V, O'Connor P, Soh C, Kabbinavar F, for the ATLAS Investigators. A randomized, double-blind, placebo-controlled, phase IIIb trial (ATLAS) comparing bevacizumab (B) therapy with or without erlotinib (E) after completion of chemotherapy with B for first-line treatment of locally advanced, recurrent, or metastatic non-small cell lung cancer (NSCLC). Journal of Clinical Oncology. 2009; 27:18S.
20. Rigas J, Carey M, Rubin M, Dragnev K, Aitken C, Hassan G, et al. Efficacy of maintenance erlotinib versus placebo in patients with unresectable stage III non-small cell lung cancer (NSCLC) following concurrent chemoradiation Journal of Thoracic Oncology. 2009; 4(9):A 3.6.
21. Moher D, Schulz K, Altman D, for the CONSORT Group. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. JAMA. 2001; 285(15):1987-91.
22. Gridelli C, Gallo C, Di Maio M, Barletta E, Illiano A, Maione P, et al. A randomised clinical trial of two docetaxel regimens (weekly vs 3 week) in the second-line treatment of non-small-cell lung cancer. The DISTAL 01 study. British Journal of Cancer. 2004; 91(12):1996-2004.
23. U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER). Guidance for industry, clinical trial endpoints for the approval of cancer drugs and biologics. 2005.
24. Roche Ltd. Supplemental NDA: first-line maintenance therapy in patients with locally advanced or metastic NSCLC-Briefing document December, 2009. 2009; Available from: http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/OncologicDrugsAdvisoryCommittee/UCM193921.pdf.
25. National Institute for Health and Clinical Excellence. Lung cancer: The diagnosis and treatment of lung cancer. National cost-impact report. 2005 [cited 2010 January]; Available from: <http://www.nice.org.uk/nicemedia/pdf/cg024costings.pdf>.
26. Pinter F, Papay J, Almasi A, Sapi Z, Szabo E, Kanya M, et al. Epidermial growth factor receptor (EGFR) high gene copy number and activating mutations in lung adenocarcinomas are not consistently accompanied by positivity for EGFR protein in standard immunohistochemistry. Journal of Molecular Diagnostics. 2008; 10(2).
27. Cappuzzo F, Coudert R, Wierzbicki R, Park K, Custers F, on behalf of the SATURN Investigators, et al. Overall survival analyses from the SATURN phase III placebo-controlled study of erlotinib as first-line maintenance therapy in advanced non-small-cell lung cancer. European Journal of Cancer Supplement. 2009; 7(3):12.
28. Cella D, Eton D, Fairclough D, Bonomi P, Heyes A, Silberman C, et al. What is a clinically meaningful change on the Functional Assessment of Cancer Therapy - Lung (FACT-L): Results from the Eastern Cooperative Oncology Group (ECOG) Study 5592. Journal of Clinical Epidemiology. 2002; 55:285-95.

29. ERG. ERG report: Pemetrexed for the maintenance treatment of locally advanced or metastatic non-small cell lung cancer2009 [cited 2010 20th March]: Available from: http://www.nice.org.uk/nicemedia/pdf/LungCancerPemetrexedERGReport.pdf.
30. Bucher H, Guyatt G, Griffith L, Walter S. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. Journal of Clinical Epidemiology. 1997; 50:683-91.
31. Parmar M, Torri V, Stewart L. Extracting summary statistics to perform meta-analyses of the published literature for survival endpoints. Statistics in Medicine. 1998; 17:2815-34.
32. National Institute for Health and Clinical Excellence. Lung cancer (non-small cell) - erlotinib, single technology appraisal. 2006 [07/2009]; Available from: http://guidance.nice.org.uk/TA162.
33. British National Formulary. BNF 58 2009 [cited 2010 March]. Available from: http://bnf.org/bnf/.
34. Chemotherapy planning online resource tool: Available from: http://cport.co.uk/Home.aspx?ReturnUrl=%2fDefault.aspx.
35. Roche Ltd. Tarceva (erlotinib) Summary of Product Characteristics (SmPC). 2009; Available from: http://www.medicines.org.uk/searchresult.aspx?search=tarceva.
36. Eli Lilly. Alimta (pemetrexed) Summary of Product Characteristics (SmPC). 2009 [updated September 2009.]; Available from: http://www.medicines.org.uk.
37. Department of Health. NHS reference costs 2007-2008. Available from: http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_098945.
38. Kind P, Macran S. Eliciting social preference weights for functional assessment of cancer therapy-lung health states. Pharmacoeconomics. 2005; 23(11):1143-53.
39. Kind P. Calibrating FACT-L on an EQ-5D utility metric: refining the use of a VAS estimation procedure. York: University of York 2009.
40. Nafees B, Stafford M, Gavriel S, Bhalla S, Watkins J. Health state utilities for non small cell lung cancer. Health and Quality of Life Outcomes. 2008; 6(84).
41. National Institute for Health and Clinical Excellence. Guide to the methods of technology appraisal. London: National Institute for Health and Clinical Excellence; 2008; Available from: <http://www.nice.org.uk/media/B52/A7/TAMethodsGuideUpdatedJune2008.pdf>
42. National Institute for Health and Clinical Excellence. Tarceva (erlotinib) for the treatment of non-small cell lung cancer -TA162. 2006 [cited 2010 January]; Available from: <http://www.nice.org.uk/guidance/index.jsp?action=download&o=37396>.
43. Drummond M, Jefferson T. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working. British Medical Journal. 1996; 313(7052):275-83.
44. Sacco JJ, 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(1).
45. Cox DR, Oakes D. Analysis of Survival Data: Chapman and Hall/CRC; 1984.
46. Department of Health. NHS Reference Costs 2008-20092010 [cited 2010 March]: Available from: http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_111591.

10 APPENDIX

Appendix 1 The manufacturer's approach to validity assessment and the ERG's critique

NICE evaluative criteria	SATURN[4]	ERG comments
How was allocation concealed?	Central randomisation and drug pack number allocation was performed using an Interactive Voice Response System (IVRS)	Adequate
What randomisation technique was used?	Randomisation sequence generated centrally and used to programme the IVRS which was accessed by site investigators Randomization was stratified using an adaptive method (minimization [3]) that ensured a balance between the treatment arms for the following factors: • EGFR protein expression by IHC (EGFR positive versus EGFR negative versus EGFR undetermined); • Stage of disease at start of chemotherapy (IIIb versus IV); • ECOG PS (0 versus 1); • Chemotherapy regimen (gemcitabine plus cisplatin versus carboplatin plus docetaxel versus other); • Smoking status (current smoker [includes patients who had stopped smoking within a year] versus former smoker versus never smoked); • Region (North America, South America, Western Europe, Eastern Europe, South East Asia and Africa).(Page 49 CSR)	Adequate
Was a justification of the sample size provided?	CSR states it was powered to perform two primary analyses, to compare PFS between the two treatment arms in all patients and in patients who had EGFR IHC positive tumours. To detect a 25% improvement in median PFS (HR=0.8) with 80% power at a two-sided 3% significance level, then 731 events (progression or death) were required. If 50% of patients were expected to be EGFR positive then this would lead to 215 randomised patients per arm for testing the treatment difference in PFS for EGFR + patients. For the secondary endpoint of overall survival in order to detect a 25% improvement in median survival with erlotinib (HR=0.8) with 80% power at a two-sided 5% significance then 641 events are required.	Appropriate

NICE evaluative criteria	SATURN[4]	ERG comments
Was follow-up adequate?	18 month recruitment with 6 months follow-up for PFS and 15 months follow-up for OS	Apprpriate
Were the individuals undertaking the outcomes assessment aware of allocation?	Study was double blind with sponsor, investigators and patients unaware of the treatment assignment of each patient. Pharmacokinetic data collected and released to analyst but not released with individual patient identification until database was closed. Un-blinding at progression only in specific circumstances Roche personal remained blinded when preparing statistical outputs	Rash is an adverse event in 50% of patients and diarrhoea in 20% of patients. Therefore efforts to blind sponsors, investigators and patients was not possible.
Was the design parallel-group or crossover? Indicate for each crossover trial whether a carry-over effect is likely	This was not a cross over trial. However, following progression patients would have been provided second-line therapy at the discretion of the investigator.	
Was the RCT conducted in the UK (or were one or more centres of the multinational RCT located in the UK)? If not, where was the RCT conducted, and is clinical practice likely to differ from UK practice?	Only one UK site Multi-centre trial with majority of patients recruited from Eastern Europe and Southeast Asia Fist-line treatments similar to UK but proportions different Second –line treatments more varied than used in the UK	Not comparable to UK population
How do the patients included in the RCT compare with patients who are likely to receive the intervention in the UK? Consider factors known to affect outcomes in the main indication, such as demographics, epidemiology, disease severity, setting.	Patient population younger and fitter than those in UK practice.	This is consistent with other similar trials
For pharmaceuticals, what dosage regimens were used in the RCT? Are they within those detailed in the Summary of Product Characteristics?	Yes	Appropriate
Were the study groups comparable?	Yes	Appropriate
Were the statistical analyses used appropriate?	Generally correct except for *post hoc* analysis	Implications noted
Was an intention-to-treat analysis undertaken?	Yes	Appropriate
Were there any confounding factors that may attenuate the interpretation of the results of the RCT(s)?	No	

Appendix 2: Types of major protocol violation in the SATURN[4] trial:

- Failure to receive at least 1 dose of study medication
- Incorrect study medication given versus randomised treatment arm (crossover from erlotinib to placebo or vice versa)
- Failure to receive four cycles of platinum-based chemotherapy
- Failure to undergo at least one post-baseline efficacy assessment (unless patient died before first post-baseline tumour assessment)
- Receipt of previous anti-cancer treatments specifically listed in the exclusion criteria of the protocol
- Absence of measurable disease at screening
- Absence of CR, PR or SD at baseline
- Absence of histologically documented stage IIIb or IV NSCLC
- Presence of malignancy other than carcinoma in situ of the cervix or basal or squamous cell skin cancer
- Lung tumour resection following response to chemotherapy before baseline.