STA REPORT

LIVERPOOL REVIEWS AND IMPLEMENTATION GROUP (LRiG)

Nivolumab for previously treated locally advanced or metastatic non-squamous non-small cell lung cancer [ID900]

Confidential until published

This report was commissioned by the NIHR HTA Programme as project number 14/206/12

07 March 2016

CONTAINS ACADEMIC IN CONFIDENCE AND COMMERCIAL IN CONFIDENCE DATA

UNIVERSITY OF LIVERPOOL

LIVERPOOL REVIEWS AND IMPLEMENTATION GROUP

A MEMBER OF THE RUSSELL GROUP

Title: Nivolumab for previously treated locally advanced or metastatic non-squamous non-small cell lung cancer [ID900]

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

Authors: Rumona Dickson, Director (Clinical Effectiveness), LR*i*G, University of Liverpool

Juliet Hounsome, Research Fellow, (Clinical Effectiveness), LR*i*G, University of Liverpool

Angela Stainthorpe, Research Fellow (Health Economic Modelling), LR*i*G, University of Liverpool

Ahmed Abdulla, Research Fellow (Health economics modelling), LR*i*G, University of Liverpool

Adrian Bagust, Professor (Health Economic Modelling), LR*i*G, University of Liverpool

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

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

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

Alastair Greystoke, Honorary Consultant in Medical Oncology, Newcastle Upon Tyne NHS Foundation Trust

Correspondence to: Rumona Dickson, Liverpool Reviews and Implementation Group, University of Liverpool, Room 2.12, Whelan Building, The Quadrangle, Brownlow Hill, Liverpool L69 3GB

Date completed: 07 March, 2016

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 14/206/12

Declared competing interests of the authors: None

Acknowledgements: The authors would like to thank Nigel Fleeman and his LR*i*G team for allowing us to use their previous report related to the use of nivolumab. Also thanks go to Dr

Paula Scullin (Consultant Oncologist, Belfast HSC Trust) for providing feedback on a draft version of the report and Gareth Jones for administrative assistance.

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, Hounsome J, Stainthorpe A, Abdulla A, Bagust A, Richardson M, Boland A, Kotas, E, Greystoke A. Nivolumab for previously treated locally advanced or metastatic non-squamous non-small cell lung cancer [ID900]: A Single Technology Appraisal. LR*i*G, University of Liverpool, 2016

Contributions of authors:

Dickson R	Project lead, drafted clinical results section and supervised the final report
Hounsome J	Critical appraisal of the clinical evidence, production of initial report and incorporation of peer reviewer comments
Stainthorpe A	Critique of the company economic model and proposal of alternative interpretations of the economic evidence
Abdulla A	Summary and critical appraisal of economic evidence
Bagust A	Checking and validation of the economic model and critique
Richardson M	Critical appraisal of the statistical evidence
Boland A	Critical appraisal of the clinical and economic evidence
Kotas E	Critical appraisal of the database searching
Greystoke A	Clinical advice and critical appraisal of the clinical sections of the company submission

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

Table of contents

List of tables

List of figures

List of Boxes

LIST OF ABBREVIATIONS

AE	adverse event
ALK	anaplastic lymphoma kinase
ASBI	Average Symptom Burden Index
BMS	Bristol-Myers Squibb
BSC	best supportive care
CI	confidence interval
CR	Complete response
CS	company's submission
CSR	clinical study report
DMC	Data Monitoring Committee
DoR	duration of response
ECOG	Eastern Cooperative Oncology Group
EGFR	epidermal growth factor receptor
EQ-5D	EuroQol-5 dimensions (questionnaire)
EQ-VAS	EuorQol – visual analogue scale
ERG	Evidence Review Group
HR	hazard ratio
HRQoL	health related quality of life
ICER	incremental cost effectiveness ratio
irAE	immune related adverse events
ITC	indirect treatment comparison
ITT	intention-to-treat
K-M	Kaplan-Meier
LCSS	Lung Cancer Symptom Scale
NLCA	National Lung Cancer Audit
NICE	National Institute for Health and Care Excellence
NSCLC	non-small cell lung cancer
ORR	overall response rate
OS	overall survival
PD	progressive disease
PD-1	programmed death-1
PD-L1	programmed death-ligand 1
PD-L2	programmed death-ligand 2
PFS	progression-free survival
PPS	post-progression survival
PR	partial response
PS	performance status
PSA	probabilistic sensitivity analysis
QALY	quality adjusted life year
Q2W	every 2 weeks
Q3W	every 3 weeks
RCT	randomised controlled trial
RECIST	Response Evaluation in Solid Tumours
RMST	restricted mean survival time
STA	single technology appraisal
RWD	real world data
TSAP	trial statistical analysis plan
TTD	time to treatment discontinuation
TTR	time to response

1 SUMMARY

The remit of the Evidence Review Group (ERG) is to comment on the clinical and cost effectiveness evidence submitted to the National Institute for Health and Care Excellence (NICE) as part of the single technology appraisal (STA) process. Clinical and economic evidence has been submitted to NICE by Bristol-Myers Squibb Pharmaceuticals Ltd in support of the use of nivolumab (Opdivo®) for patients who have received prior chemotherapy for locally advanced or metastatic non-squamous, non-small cell lung cancer (NSCLC). A European licence for nivolumab in this specific patient population has not been received but the company expects a decision to be made in the first quarter of 2016.

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

The company submission (CS) indicates a slight change in wording in the included population – changed from people with previously treated disease to those who have received prior chemotherapy (thus excluding epidermal growth factor receptor (EGFR) positive patients who have previously had a TKI). Although the company submission (CS) acknowledges the validity of all of the comparators in the scope they limit their analysis to available data which therefore provides comparison of nivolumab with docetaxel, nintedanib+docetaxel and best supportive care (BSC).

1.2 Summary of submitted clinical effectiveness evidence

Clinical evidence includes direct evidence of nivolumab compared with docetaxel from CheckMate 057. The trial was stopped early due to the pre-specified stopping rules related the superiority of nivolumab in relation to overall survival (OS). An indirect treatment comparison (ITC) comparing nivolumab with nintedanib+docetaxel as well as best supportive care (BSC) is provided. The company admits that the analysis of the original trial data and the ITC are limited by the fact that the proportional hazards assumption has been violated and therefore **none** of the hazard ratios (HRs) can be considered a reliable estimate of treatment effect.

CheckMate 057 provides evidence of median overall survival (OS) benefit of nivolumab over docetaxel at both 12 and 18 months (12.2 versus 9.4 and 39 versus 23 months respectively). Due to issues of pseudo progression (tumours that initially increase as a result of the treatment action before shrinking/stabilising) with nivolumab, the results for progression free survival (PFS) are less clear. Patients receiving nivolumab show less benefit at 12 month (4.2 versus 2.3 months). However, 12 month data show a reversal with PFS rates for

nivolumab versus docetaxel at 18.5 versus 8.1%. Subgroup analysis by EGFR status ('all comer'[1] population versus EGFR mutation-negative/unknown) show similar results.

The adverse event (AE) data presented indicate that nivolumab, although having a slightly different AE profile to standard cytotoxic chemotherapy, has fewer Grade 3-4 AEs than docetaxel. Data from additional non-randomised studies and studies of the use of nivolumab in patients with a variety of other cancers are provided to support this assertion. The CS makes the case that the uniqueness of the AE profile can be managed by established guidelines and that overall treatment with nivolumab is better tolerated than treatment with docetaxel alone and by association is also superior to nintedanib+docetaxel.

The ITC provides evidence using restricted mean survival time (RMST) analysis demonstrating no benefit of nivolumab versus nintedanib+docetaxel in relation to OS, PFS, overall response rate (ORR) or AEs in either the 'all comer' or EGFR mutation negative/unknown population. The comparison with BSC provides somewhat mixed results demonstrating the possible lack of homogeneity of the studies used in the comparison. No data are available for the EGFR positive population of patients.

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

The primary data provided in the CS comes from CheckMate 057 and an ITC that is limited by a lack of data to allow for comparison with all of relative comparators listed in the scope. The comparison of nivolumab is therefore limited to data related to docetaxel, nintedanib+docetaxel and BSC.

CheckMate 057 is a well conducted trial however the use of HRs in the analysis of the data cannot be considered a reliable estimate of treatment effectiveness as the CS points out that the proportional hazards assumption is violated for both OS and PFS. This limitation is also true of the ITC where only RMST analysis should be considered. The ITC is also limited by the fact differences in the patient populations included in the analysis (e.g. inclusion of patients with squamous disease, Asian population, length of follow-up etc.) The comparison with BSC provides mixed results demonstrating the effectiveness of nivolumab versus BSC in the all-comers group but not the EGFR mutation-negative/unknown patients supporting concerns that there were differences in the patient populations in the trials used in the ITC.

The CS infers that the AE experienced by patients receiving nivolumab will be fewer than those experienced by patients receiving nintedanib+docetaxel. The ERG is of the opinion that although the comparative data are limited that patients receiving docetaxel do have

[1] all comers- the term used in the CS to denote the entire population of CheckMate 057

higher rates of Grade3-4 AEs and it would be expected this would be at least the same when docetaxel was given in combination with nintedanib.

The CS makes a claim that OS in the patients receiving docetaxel in CheckMate 057 is longer then would be expected. Examination of data from other similar trials does not substantiate this claim. The CS also makes a claim that the pseudo progression seen in patients receiving nivolumab would have an effect on OS. The ERG is not convinced that the data presented support this claim.

Subgroup analyses suggest that nivolumab is statistically significantly more effective in patients with higher PD-L1 expression levels than those with lower PD-L1 expression levels. The report is however somewhat inconsistent with regards to whether all patients should therefore be tested for PD-L1.

1.4 Summary of submitted cost effectiveness evidence

The company developed a de novo cohort-based partitioned survival model in Microsoft Excel to compare the cost effectiveness of nivolumab 3mg/kg given every 2 weeks with docetaxel 75mg/m^2 given every 3 weeks as the base case comparator. The model comprised three health states: pre-progression, post-progression and death. All patients entered the model in the pre-progression state. Variants of this model structure have been used in the modelling of treatment for patients with cancer in a number of previous NICE STAs. The model time horizon was set to 20 years with a 1-week cycle length. As recommended by NICE, a discount rate of 3.5% has been used for both costs and outcomes; outcomes are measured in quality adjusted life years (QALYs). The model perspective was that of the UK NHS. Survival estimates were based on data collected from CheckMate 057 and published sources. Utility values were calculated from data collected during CheckMate 057. Resource use and costs were estimated based on information from CheckMate 057, published sources and advice from clinical and economic experts. The company also compared nivolumab versus nintedanib+docetaxel via an ITC. The company did not estimate the cost effectiveness of nivolumab versus BSC.

In the CS, the base case comparison describing nivolumab vs. docetaxel resulted in an incremental cost effectiveness ratio (ICER) per QALY gained of £103,589, with nivolumab being more expensive (+£75,452) and more effective (+1.15 life years and +0.73 QALYs) than docetaxel. The company carried out a range of deterministic sensitivity analyses. The most influential parameters were discount rate and average body weight. Other influential parameters include body surface area, utility weights, administration cost of nivolumab and progression-free state costs. The probabilistic sensitivity analysis (PSA) results show that

the probabilistic ICER of £99,291 per QALY gained has a 0% chance of being cost effective at a threshold of £30,000 per QALY gained and a 0.1% probability of being cost effective at a threshold of £50,000 per QALY gained.

The ICER per QALY gained for nivolumab versus nintedanib+docetaxel was £126,861; nivolumab had higher lifetime costs (+£62,598) and was more effective (+0.80 life years and +0.49 QALYs) than the combination therapy. The probabilistic ICER per QALY gained for nivolumab versus nintedanib+docetaxel was £111,934. Scenario analyses were undertaken by the company using different survival modelling approaches for OS and time to treatment discontinuation (TTD) and alternative treatment durations.

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

The company's decision model is structured conventionally. The economic model relies on patient level data from CheckMate 057. Projection of survival data was required to enable a lifetime equivalent evaluation. Limited data from CheckMate 057 and published sources were used to identify suitable parametric models for survival extrapolation. The ERG has identified the following main areas of concern: (i) manner in which OS, PFS and post-progression survival (PPS) have been projected, (ii) use of time to treatment discontinuation (TTD) data instead of PFS in all parts of the company model, (iii) indirect treatment comparison of nivolumab with nintedanib+docetaxel, (iv) choice of utility values used in the model, (v) nivolumab dosing calculations and (vi) treatment administration costs.

The ERG considers the company's methods to project OS and PFS to be flawed for both the intervention and the comparators. Concerns relating to the modelling of each health state are compounded by the ERG's identification of subgroups of patients within the patients treated with nivolumab. The interdependence of OS, PFS and all cause population mortality in the model also results in questionable projections for nivolumab OS and PFS. The ERG also identified problems with the company's use of TTD data as a proxy for PFS. The projection of PFS/TTD is implausibly long with an unlikely proportion of patients remaining alive at 20 years, progression-free and continuing to receive treatment. Additionally, the ERG considers the use of TTD instead of PFS data to estimate QALYs to be inappropriate. In relation to the ITC of nivolumab versus nintedanib+docetaxel, the ERG considers that piecewise PH assumptions do not hold for OS and PFS in the LUME-Lung 1 trial, thus invalidating any potential inferences made by the company. The ERG is concerned with the possible over-estimation of utility values collected as part of CheckMate 057. Throughout the duration of the trial, the number of respondents steadily declined and it is likely that participants that continued to respond to the EQ-5D questionnaires were exhibiting self-selecting behaviour and are unlikely to match the characteristics of the initial trial population.

In addition, the ERG has identified two key issues with the company model related to costs. In the model there is an over-estimation of the average cost per dose of nivolumab due to a body weight calculation error and treatment administration costs are calculated according to the number of patients in treatment mid-cycle rather than at the start of the cycle.

1.6 Summary of exploratory and sensitivity analyses undertaken by the ERG

The ERG implemented various modifications to the company's model which yielded a mixture of effects. The individual execution of model amendments resulted in both increases and decreases in the size of the estimated ICER per QALY gained for nivolumab versus the comparator treatments. The combined impact of ERG recommended model revisions resulted in an estimated ICER per QALY gained of £165,234 for nivolumab versus docetaxel and £293,232 per QALY gained for nivolumab versus nintedanib+docetaxel.

The ERG considers that the company base case results substantially underestimate the size of the ICER per QALY gained for nivolumab versus docetaxel and nivolumab versus nintedanib+docetaxel for a previously treated non-squamous NSCLC patient population.

1.7 Summary of company's case for end of life criteria being met

The company makes the following case for nivolumab versus docetaxel to be considered under NICE's end of life criteria:

- patients with advanced or metastatic squamous NSCLC have a life expectancy of less than 24 months
- data from CheckMate 057 demonstrate that nivolumab extends life by more than 3 months compared with docetaxel
- the patient population eligible for nivolumab treatment in England is expected to be small (n=3570).

The company does not make the case for nivolumab versus nintedanib+docetaxel to be considered under NICE's end of life criteria.

1.8 ERG commentary on end of life criteria

The ERG agrees that patients with advanced NSCLC have a short life expectancy of less than 24 months and that the total number of patients who would be eligible for the treatment is small. It also considers that nivolumab offers an extension to life of more than 3 months in comparison with docetaxel; the ERG estimates a mean gain of 5.8 months for nivolumab versus docetaxel. The ERG estimates a mean extension to life of 3.1 months in comparison with nintedanib+docetaxel.

1.9 ERG commentary on the robustness of submitted evidence

1.9.1 Strengths

Clinical evidence

- Checkmate 057 is a good quality trial providing direct evidence of effectiveness of nivolumab versus docetaxel in relation to OS and demonstrating an acceptable AE profile.

Cost effectiveness evidence

- The company provided a detailed submission that fulfilled the requirements of NICE's scope for the base case analysis. The ERG's requests for further clinical information were met to a good standard
- Variants of this model structure have been used in the modelling of similar treatments in a number of previous NICE STAs
- The decision model submitted by the company is generally implemented to a good standard.

1.9.2 Weaknesses and areas of uncertainty

Clinical evidence

- The validity of all assessed outcomes is limited by the fact that the proportional hazards assumption has been violated
- The comparison with all comparators in the original scope is limited by the available direct and indirect evidence.

Cost effectiveness evidence

Issues common to the modelling of nivolumab, docetaxel and nintedanib+docetaxel

- QALY calculations in the company model are linked to the time patients spend on treatment and not to their health state, which is incorrect
- The utility data used by the company lack credibility
- The model calculates treatment administration costs mid-treatment cycle when they should be applied at the start of the cycle, when treatment is received.

Issues specific to the modelling of nivolumab

- The method employed by the company to project nivolumab OS results in the model does not adequately represent the existing trial evidence from CheckMate 057
- The company's PFS model projects a small minority of patients treated with nivolumab to remain progression free throughout the lifetime of the model and to constitute 85% of those patients still alive after 20 years. It also predicts that any patient treated with nivolumab who is still in PFS by 18.4 years is cured of the disease and will never progress. The ERG considers both these outcomes to be implausible
- The company model creates an interdependence between OS and PFS projections that results in some values from the parametric OS model for nivolumab being replaced by PFS values to ensure that PFS is never greater than OS. This indicates

that at least one of the parametric models (PFS or OS) used for nivolumab is inappropriate

- In the company model, one-third of the survival gain (nivolumab versus docetaxel) occurs post-progression, but this does not take into account the subgroup of nivolumab patients treated beyond progression who continue to accrue extra survival benefit, whether due to extra treatment or other factors. ERG analysis suggests that post-progression survival constitutes 52% of survival gain when 25% of patients are treated beyond progression
- The nivolumab dosing calculations undertaken by the company are inaccurate

Issues specific to the modelling of nintedanib+docetaxel

- The proportional hazards assumptions required to validate the company's indirect method of comparing nivolumab with nintedanib+docetaxel do not hold

2 BACKGROUND

2.1 Critique of company's description of underlying health problem.

Key points from the description of the underlying health problem (lung cancer, and in particular non-squamous non-small cell lung cancer [NSCLC]) presented in the company's submission (CS) are summarised in Box 1.

Box 1 Company's overview of the underlying health problem

Lung cancer

- Lung cancer is the second most common cancer in the UK and has the highest mortality of any cancer
- Most lung cancers are diagnosed at an advanced stage, when the cancer has spread to lymph nodes and other organs in the chest (locally advanced disease and unresectable locally advanced disease; stages IIIA and IIIB) or to other parts of the body (metastatic disease; stage IV)
- In 2011, lung cancer was the underlying cause of 30,148 deaths in England and Wales
- The median survival for all lung cancer patients in England and Wales was 7.6 months
- Although lung cancer typically affects older patients (median age of diagnosis in England and Wales is 74 years), in 2013 more than one-third of patients diagnosed with lung cancer were aged between 50 and 70 years

Non-small cell lung cancer (NSCLC)

- Approximately 84% of lung cancer cases in England and Wales fall within the NSCLC category
- In 2013, there were 27,300 patients with NSCLC in England; 19,138 patients (70%) had stage IIIB or IV lung cancer
- Median survival for all stage III patients with NSCLC was 9.6 months
- Median survival for stage IV patients with NSCLC was only 3.3 months
- Data from the UK suggest the 1-year relative survival rate (by stage at diagnosis) is 71%, 48%, 35%, and 14% for stage I, II, III, and IV disease, respectively
- In addition to high mortality, a large proportion of patients experience increasingly severe morbidity as they progress from localised to metastatic disease
- Approximately 90% of patients with advanced NSCLC experience two or more disease-related symptoms, such as cough, dyspnoea, pain, anorexia, or fatigue
- These symptoms, in turn, can cause psychological distress and may have a negative impact on a patient's health-related quality of life (HRQoL)

Non-squamous NSCLC

- NSCLC can be further divided into squamous or non-squamous NSCLC, based on the cell type responsible for the tumour
- Approximately 64% of patients within England and Wales had non-squamous NSCLC in 2013
- EGFR or ALK mutations are predominantly present in non-squamous NSCLC and if present lead to the following of a slightly different care pathway.

ALK=anaplastic lymphoma kinase; EGFR=epidermal growth factor receptor; NSCLC=non-small cell lung cancer
Source: CS, Sections 3.1 and 3.3

The Evidence Review Group (ERG) considers that, in general, these key points appropriately summarise the key points related to this health problem. The ERG notes that the prevalence of epidermal growth factor receptor (EGFR) mutation is 15% in patients in Spain,[1] 10% in patients in the USA and up to 35% of patients in Asia,[2] while the prevalence of anaplastic lymphoma kinase (ALK) mutations is 3-7% in patients with NSCLC.[3] Up-to-date data for patients in the UK are currently not available.

2.2 Overview of current service provision

The ERG has summarised (as bulleted items) the key points from the company's description of current treatment options for patients with non-squamous NSCLC in Box 2. The ERG considers that these points provide an accurate overview of current service provision.

Box 2 Current treatment options for patients with stage IIIB and IV non-squamous NSCLC

Current treatment options

- The aims of therapy are to prolong survival and improve HRQoL
- Treatment of patients with non-squamous NSCLC depends on a patient's ECOG PS, comorbidities, histology, presence of mutations and personal choice
- Patients are typically treated with platinum-based doublet chemotherapy at first-line
- At second-line patients can be treated with docetaxel chemotherapy or nintedanib in combination with docetaxel
- Third-line treatment can include erlotinib (if not received previously in patients with EGFR-unknown status) and docetaxel.

EGFR-positive tumours

- At first-line NICE recommends the use of the EGFR inhibitors erlotinib, afatinib and gefitinib
- At second-line patients may receive platinum-based chemotherapy (in combination with gemcitabine, vinorelbine, pemetrexed or a taxane). If no previous EGFR-TKI therapy has been used afatinib or erlotinib may be given. In patients for whom platinum-based chemotherapy is inappropriate, patients may receive single-agent gemcitabine or vinorelbine
- Third-line treatment can include nintedanib in combination with docetaxel. Following the use of an EGFR-TKI and one other therapy, docetaxel monotherapy and BSC may be used, although these are not recommended by NICE in the third-line setting.

ALK-positive tumours

- As with ALK-negative patients, those with ALK-positive tumours may receive platinum-based chemotherapy (in combination with gemcitabine, vinorelbine, pemetrexed or a taxane) at first-line
- Crizotinib is currently available as a second-line treatment in ALK-positive patients through the Cancer Drugs Fund
- Ceritinib (current NICE appraisal suspended) received FDA and conditional EMA approval for NSCLC treated with or intolerant to crizotinib.

Issues relating to current clinical practice

- Due to their age and/or co-morbidities, many patients in the UK are unlikely to receive systemic treatment
- First-line therapy in this patient population is a platinum-based combination therapy, which is associated with high toxicity and may not be suitable for many patients
- Only 23% of patients with non-squamous NSCLC are treated with first-line therapy
- The mortality rate in these patients is high and the OS rate is low following first-line therapy
- Long-term survival, with a concomitant good HRQoL, is not currently deemed achievable with current treatments in this patient population
- BSC, such as analgesics, antiemetics, and palliative interventions, are a part of the care package offered to all patients with non-squamous NSCLC, regardless of eligibility for systemic anti-cancer therapies and line of treatment
- NICE[4] recommends five different tests for detecting EGFR status in NSCLC, test accuracy is dependent on the quality of the tissue samples available
- Turnaround times for EGFR mutation testing are from 3 to 7 days
- While sequential testing of EFGR and ALK is more cost-effective, parallel testing allows for more rapid turnaround of results

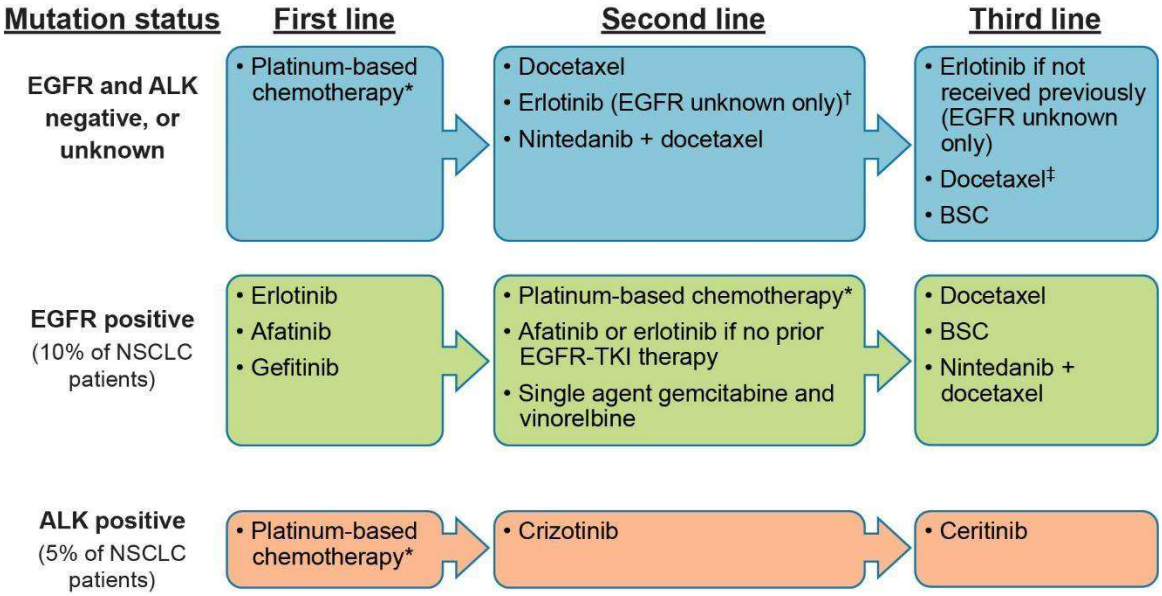
HRQoL=health related quality of life; NSCLC=non-small cell lung cancer; ECOG PS=Eastern Cooperative Oncology Group performance status; EGFR=epidermal growth factor receptor; TKI=tyrosine kinase inhibitors; NICE= National Institute for Health and Care Excellence; ALK=anaplastic lymphoma kinase; BSC=best supportive care; FDA=Food and Drug Administration; EMA=European Medicines Agency; OS=overall survival

Source: CS, Sections 3.2 and 3.5

Nivolumab has received marketing authorisation for use in the NSCLC squamous population but has not received marketing authorisation for use in the non-squamous NSCLC patient population. The company expects a decision to be made in the first quarter of 2016.

Nivolumab is a human, monoclonal immunoglobulin G4 antibody that acts as a programmed death-1 (PD-1) inhibitor; nivolumab blocks the interaction of PD-1 with programmed death-ligands 1 and 2 (PD-L1 and PD-L2).[5,6] A typical immune response to foreign antigens or cells in the body is the activation of T-cells that can destroy these antigens or cells; the PD-1 receptor is a negative regulator of T-cell activity. Engagement of PD-1 with its ligands (PD-L1 and PD-L2) results in the inhibition of T-cell activation and T-cell death. PD-1 has also been shown to control the inhibition of T-cell response in human malignancies.[7-9] Hence, nivolumab stimulates the patient's own immune system to directly fight cancer cells, resulting in destruction of the tumour. Nivolumab's mechanism of action differs from that of conventional cytotoxic anti-cancer therapies which generally destroy all rapidly dividing and fast growing cell types. The cytotoxic mode of action means that non-cancerous cells, such as hair follicles and gut mucosa, are often targeted alongside cancer cells, resulting in undesirable side effects such as hair loss and diarrhoea.

The CS provides an overview of the current treatment pathway for patients with non-squamous NSCLC and is summarised in Figure 1).



Abbreviations: ALK=anaplastic lymphoma kinase; BSC=best supportive care; EGFR=Epidermal Growth Factor Receptor; NSCLC=non-small cell lung cancer; TKI= tyrosine kinase inhibitor; UK=United Kingdom

* Platinum-based chemotherapy + gemcitabine, vinorelbine, pemetrexed or a taxane

† Until recently, erlotinib was recommended second-line in patients with EGFR mutation-negative/unknown status; however, recent NICE guidance recommends erlotinib only in patients with EGFR unknown mutation status, which is a very small subgroup of patients

‡ As erlotinib is no longer recommended in second-line for patients with EGFR mutation-negative status, docetaxel (as monotherapy or in combination with nintedanib) is the only second-line option and, as a result, will no longer be used in third-line

Source: CS, Figure 1

Figure 1 Overview of treatments in the UK for this appraisal

In the CS (Figure 8), the company proposes nivolumab as a second- or even third-line treatment option for patients with non-squamous NSCLC. The CS estimates the potential number of patients eligible for nivolumab as a second-line treatment to be 1413 (Table 1). Clinical advice given to the ERG indicates that this may be an underestimate since clinicians would consider using nivolumab to treat patients whose condition means that they would be unlikely to be able to tolerate the side effects of docetaxel or nintedanib+docetaxel.

Table 1 Estimate of those eligible for nivolumab for non-squamous NSCLC in England

Population	Proportion of patients	Number of patients	Reference
Total NSCLC	N/A	27,300	Health and Social Care Information Centre 2014b[10]
Patients with stage IIIb/IV NSCLC	N/A	19,138	Health and Social Care Information Centre 2014b[10]
Non-squamous NSCLC	64.35%	12,315	Powell 2013[11]
Second-line setting	11.5%	1413	NICE,[12] Sculier and Moro-Sibilot (2009)[13]

NICE=National Institute for Health and Care Excellence; NSCLC=non-small cell lung cancer
Source: CS, Table 121

3 CRITIQUE OF COMPANY'S DEFINITION OF DECISION PROBLEM

Table 2 summarises the decision problem described by the company in the CS in relation to the final scope issued by NICE with reasons for any differences.

3.1 Population

The CS limits the population to those patients who had previously received chemotherapy rather than those who had received any prior treatment. This therefore excludes those patients with EGFR positive non-squamous NSCLC who have *only* received an EGFR-TKI. However, the ERG concurs that this is a very small group and that the population included is appropriate.

3.2 Intervention

The intervention (nivolumab) described in the CS matches the intervention described in the final scope issued by NICE. Nivolumab (brand name Opdivo®) is administered via intravenous infusion at 3mg/kg over 60 minutes every 2 weeks.

At the time of writing, nivolumab is still awaiting marketing authorisation from the European Medicines Agency (EMA) for use in patients with non-squamous NSCLC.

3.3 Comparators

For the clinical effectiveness systematic review all comparators outlined in the NICE scope are included. However the CS notes that for the non-squamous population "The comparators listed in the final scope are representative of the standard treatments used in the NHS. However, not all are relevant comparators to nivolumab." The CS considers that the relevant comparators are docetaxel, nintedanib+docetaxel and best supportive care (BSC). Their rationale for this decision are outlined in sections 1.4 (Figure 2) and 3.2 of the CS. The basis for this decision is that there are no or limited data available to compare nivolumab to the other current standard treatments.

The ERG agrees that the available data allow only for a comparison of nivolumab with docetaxel, nintedanib+docetaxel and BSC. Currently docetaxel is the standard of care. However, with the recent approval of nintedanib+docetaxel it is expected that this will replace docetaxel monotherapy and become the standard care for patients fit enough to tolerate the treatment.

3.4 Outcomes

The outcomes in the CS match the outcomes described in the final scope. The measurements of outcomes are appropriate.

3.5 Economic analysis

As specified in the final NICE scope, the cost effectiveness of treatments was expressed in terms of the incremental cost per quality adjusted life year (QALY) gained through modelling that extended over a 20-year time horizon (equivalent to a lifetime horizon) and costs were considered from an NHS perspective.

3.6 Subgroups

The NICE scope specifies that if the evidence allows, consideration should be given to subgroups based on biological markers. The company carried out a range of subgroup analyses (including analyses by PD-L1 status) to assess clinical effectiveness.

Subgroup analyses for EGFR mutation-negative/unknown and PD-L1 to assess cost-effectiveness were also conducted and reported in CS appendices.

3.7 Other considerations

The CS does not identify any equality issues.

Table 2 NICE scope and company's decision problem

	Final scope issued by NICE	Decision problem in the company's submission	Rationale for difference
Population	People with previously treated locally advanced or metastatic non-squamous non-small cell lung cancer	Adults with locally advanced or metastatic non-squamous NSCLC after prior chemotherapy	In line with expected marketing authorisation
Intervention	Nivolumab	As per scope	
Comparator(s)	**Non-squamous EGFR-TK mutation negative or unknown tumours:** • After one prior therapy: - Docetaxel monotherapy - Erlotinib (subject to ongoing NICE appraisal) - Nintedanib in combination with docetaxel - Crizotinib (only for patients with ALK positive mutation status) - Ceritinib (only for patients with ALK positive mutation status; subject to ongoing NICE appraisal) - Best supportive care • After two prior therapies: - Docetaxel monotherapy - Erlotinib (if not received previously; subject to ongoing NICE appraisal) - Best supportive care **Non-squamous EGFR-TK mutation positive tumours:** • After one prior therapy: - Platinum therapy (in combination with gemcitabine, vinorelbine, pemetrexed or a taxane) - Single agent gemcitabine and vinorelbine (for people for whom platinum therapy is not appropriate) - Afatinib, erlotinib or gefitinib (if no previous EGFR-TKI therapy received due to delayed confirmation of mutation status; erlotinib and gefitinib subject to ongoing NICE appraisal) • After two prior therapies (an EGFR-TKI and one other therapy): - Docetaxel monotherapy - Erlotinib - Nintedanib in combination with docetaxel - Best supportive care	Base case economic analysis in a previously treated setting is nivolumab versus: • Docetaxel monotherapy • Nintedanib in combination with docetaxel	**EGFR negative/unknown** *Erlotinib* - no data from trial available *ALK mutation positive* - too few patients in trial to allow for subgroup analysis *Ceritinib* - at the time of CS not recommended by NICE – currently the appraisal has been suspended[14] *BSC* - lack of data available for comparison[15] **EGFR positive** *Platinum based therapy* - patients in trial had already received this therapy so this is not a valid comparator *Gemcitabine or vinorelbine* - no available data *Erlotinib, afatinib* - limited data *Gefitinib* - not recommended for second-line

	Final scope issued by NICE	Decision problem in the company's submission	Rationale for difference
Outcomes	The outcome measures to be considered include: • PFS • OS • ORR • AEs • HRQoL	As per scope	
Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per QALY The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared Costs will be considered from an NHS and Personal Social Services perspective The availability of any patient access schemes for the comparator technologies should be taken into account	As per scope	
Other considerations	If the evidence allows, consideration will be given to subgroups based on biological markers. If appropriate, the appraisal should include consideration of the costs and implications of additional testing for biological markers, but will not make recommendations	As per scope	

ALK=anaplastic lymphoma kinase; BSC=best supportive care; EGFR=epidermal growth factor receptor; OS=overall survival; PFS=progression-free survival; ORR=objective response rate; AE=adverse event; HRQoL=health related quality of life; QALY=quality adjusted life year
Source: CS, adapted from Table 1

4 CLINICAL EFFECTIVENESS

4.1 Critique of the methods of review(s)

4.1.1 Searches

The CS adequately describes the search strategies used to identify relevant studies relating to the use of nivolumab for the treatment of patients with previously treated locally advanced or metastatic non-squamous NSCLC. The search strategies were updated versions of the searches run in a 2013 NICE multiple technology appraisal by the Liverpool Reviews and Implementation Group.[16] The company conducted a systematic search for randomised controlled trial (RCT) evidence, the same search strategy was employed for the indirect treatment comparisons (ITC). Separate searches were conducted for the retrieval of cost effectiveness studies. The date of the searches and the full date span are included in the CS.

Clinical effectiveness

Full details of the strategies used to locate clinical evidence were reported in Section 4.1.1 and Appendix 2 of the CS. The search terms were relevant and included MeSH and free text as well as an RCT filter. No animal or language filters were used. The company searched the following databases: Medline, Medline in Process, Embase and The Cochrane Library (CENTRAL only). The company reported results from hand searches of three conference sites: American Society of Clinical Oncology (ASCO), European Society for Medical Oncology (ESMO) and World conference on Lung Cancer. Four clinical trial registries were searched: clinicaltrials.gov, International Clinical Trials Registry Platform (ICTRP) and Australian and New Zealand Clinical Trials Registry. The CS did not include details of the search terms used to search these additional resources; therefore the ERG was unable to comment on the search terms used.

Summary of searching

In summary, the ERG concludes that the company's searches were carried out to an adequate standard and accurately reflected the population and indication described in the final scope issued by NICE. The ERG is confident that no relevant references were missed.

4.1.2 Eligibility criteria

All citations were assessed for potential inclusion through two stages. Detailed eligibility criteria are presented in the CS (Table 7). The ERG considers these criteria to be essentially consistent with the NICE scope in relation to population and outcomes. The ERG notes that the comparators of two single chemotherapy agents (i.e. gemcitabine and vinorelbine) are

not included in the company's eligibility criteria, in line with the company decision problem but differing from the NICE decision problem.

The review included RCTs and studies published as full texts in English. The ERG notes that although the company's search aimed to identify RCTs which included patients with squamous and non-squamous histology, ultimately, studies were only included in the review if they either included only patients with non-squamous NSCLC or if the study included a relevant subgroup analysis describing patients with non-squamous NSCLC. The ERG concurs that these criteria were appropriate.

4.1.3 Risk of bias

A descriptive critical appraisal of all of the trials included in the systematic review was conducted by the company using multiple criteria (e.g. Centre for Reviews and Dissemination's guidance,[17] Jadad score,[18] and the Cochrane Collaboration risk of bias tool[19]). The results of the quality assessment for all of the included studies are presented in Table 10 and Appendix 3 of the CS. The ERG notes that whilst there are some minor errors in referencing, the quality assessment for the five studies used in subsequent analyses is presented and are accurate. The company also assessed the methodological quality of the company sponsored non-randomised studies that were provided as supportive evidence using the Down and Black's checklist for non-randomised studies.[20]

4.2 Critique of trials of the technology of interest

4.2.1 Identified studies in the systematic review

Thirty three RCTs were included in the company's review but only one trial (CheckMate 057)[21] assessed the clinical effectiveness of nivolumab (versus docetaxel). The trial characteristics and findings of CheckMate 057 were appropriately presented narratively in the CS. Characteristics of the other 32 RCTs included in the systematic review were reported in tables in the CS (CS, Appendices 7.14 to 7.16). The supporting evidence from the two non-randomised studies (CheckMate 153[22] and CheckMate 003[23]) were presented narratively in Section 4.11 and in Appendices 17 and 18 of the CS.

To compare nivolumab with the comparators of nintedanib+docetaxel and BSC the company conducted ITCs using evidence derived from CheckMate 057, LUME-Lung 1;[24] ISEL;[25] ISTANA;[26] and V-15-32.[27] trials. The ERG's critique of the company's ITCs is presented in Section 4.3. The ERG is not aware of any additional studies that should have been included.

4.2.2 Methodological approach for the synthesis and analysis of trials included in the systematic review

Since CheckMate 057 was the only study to provide direct evidence for nivolumab the company conducted ITCs to compare nivolumab to the other comparators (nintedanib+docetaxel, BSC). This is described in Section 4.3 of the ERG report.

4.2.3 Characteristics of the studies included in the systematic review

CheckMate 057 is a phase III open-label RCT of nivolumab versus docetaxel in patients with locally advanced or metastatic non-squamous NSCLC after failure of at least one prior platinum doublet-based chemotherapy. The key characteristics of CheckMate 057 are summarised in Table 3.

Table 3 Summary of methodology of CheckMate 057

	CheckMate 057
Location	106 sites in 22 countries worldwide Argentina, Australia, Austria, Brazil, Canada, Chile, Czech Republic, France, Germany, Hong Kong, Hungary, Italy, Mexico, Norway, Peru, Poland, Romania, Russian Federation, Singapore, Spain, Switzerland and the United States
Study design (including method of randomisation)	Global, phase III, randomised, open-label study Patients were randomised via interactive voice response system in a ratio of 1:1 Randomisation was stratified according to prior treatment with maintenance therapy vs. no maintenance and second-line therapy vs. third-line therapy
Study drugs	Nivolumab at 3 mg/kg by intravenous infusion every 2 weeks (n=292) Docetaxel at 75 mg/m^2 by intravenous infusion every 3 weeks (n=290)
Overview of patient population	Adult (≥ 18 years) patients with metastatic or recurrent non-squamous NSCLC after failure of prior platinum doublet-based chemotherapy
Primary outcomes	OS
Secondary outcomes	• Investigator-assessed ORR • Duration of response • Time to response • Investigator-assessed PFS • HRQoL • Safety and tolerability • Immunogenicity of nivolumab (exploratory outcome)
Duration of follow-up	The enrolment period was from November 2012 until December 2013. The last patient was randomised on 31 December 2013, and the last patient last visit occurred on 5 February 2015, providing a minimum follow-up of 13.2 months

HRQoL=health-related quality of life; NSCLC=non-small cell lung cancer; ORR=objective response rate; OS=overall survival; PFS=progression-free survival
Source: CS, adapted from Table 10

Due to differences in adverse events (AE), dosing and drug action CheckMate 057 was an open-label study.

Eligibility criteria for entry into CheckMate 057 are provided in the Appendices to the ERG report, Section 10.1. Clinical advice to the ERG is that the eligibility criteria for the trial are reasonable, although the prohibition of oral steroids may become problematic when

implementing the treatment in clinical practice. Patients may be on chronic low dose steroids for cancer related symptoms and/or short courses of steroids for exacerbations of chronic obstructive pulmonary disease (a frequent co-morbidity). Therefore, treatment of these patients may subsequently be delayed. This is the case for all immune-oncology drugs as the use of steroids is directly antagonistic to the mechanism of action. The company provided detailed information on permitted concomitant medications for CheckMate 057 (CS, Table 10).

Baseline characteristics of the CheckMate 057 patient population are provided in Table 4. The median age of patients in CheckMate 057 was 61 years in the nivolumab arm and 64 years in the docetaxel arm. There was a greater percentage of males in the docetaxel arm than in the nivolumab arm (58% versus 52%); this slight imbalance may favour nivolumab as the clinical advice received by the ERG suggests that male patients have poorer outcomes. However, there was also a 4% difference in the proportion of patients with Eastern Cooperative Oncology Group (ECOG) performance status (PS) score 0 in the docetaxel arm in comparison to patients in the nivolumab arm (33% versus 29%). As ECOG PS 0 patients would be expected to do better than ECOG PS 1 patients, this slight imbalance may favour docetaxel. Overall, the ERG does not consider that these differences are likely to lead to major bias and/or favour one treatment over another.

Table 4 Baseline characteristics of patients in CheckMate 057

Baseline characteristic		**CheckMate 057**	
		Nivolumab (n=292)	**Docetaxel (n=290)**
Age (years)	Median (range) <65, n (%) 65-74, n (%) ≥75, n (%)	61 (37-84) 184 (63) 88 (30) 20 (7)	64 (21-85) 155 (53) 112 (39) 23 (8)
Sex, n (%)	Male	151 (52)	168 (58)
Race, n (%)	White	267 (91)	266 (92)
Patients with quantifiable PD-L1 status at baseline, n (%)		231 (79.1%)	224 (77.2%)
PD-L1 expression level* n (%)	<1% ≥1% <5% ≥5% <10% ≥10 Not quantifiable at baseline	108 (46.8) 123 (53.2) 136 (58.9) 95 (41.1) 145 (62.8) 86 (37.2) 61 (20.9)	101 (45.1) 123 (54.9) 138 (61.6) 86 (38.4) 145 (64.7) 79 (35.3) 66 (22.8)
Smoking status, n (%)	Current/former Never smoked Unknown	231 (79) 58 (20) 3 (1)	227 (78) 60 (21) 3 (1)

Baseline characteristic		CheckMate 057	
		Nivolumab (n=292)	Docetaxel (n=290)
ECOG PS, n (%)	0	84 (29)	95 (33)
	1	208 (71)	193 (67)
	Not reported	0	1 (<1)
Disease stage, n (%)	IIIb	20 (7)	24 (8)
	IV	272 (93)	266 (92)
CNS metastases, n (%)	Yes	34 (12)	34 (12)
Median time from initial diagnosis,	Years (range)	0.8 (0.2-8.4)	0.8 (0.0-8.5)
No. of prior systemic cancer therapies received, n (%)	1	256 (88)	259 (89)
	2	35 (12)	31 (11)
	Other	1 (<1)	0
Prior radiotherapy, n (%)	Yes	139 (48)	138 (48)
Type of prior systemic cancer therapy n (%)	Prior platinum-based therapy	292 (100)	290 (100)
	Prior ALK inhibitor	1 (0.3)	2 (0.7)
	Prior EGFR-TKI	29 (9.9)	24 (8.3)
	Other – chemotherapy	292 (100)	290 (100)
	Other – experimental drugs	23 (7.9)	18 (6.2)
Time from completion of most recent prior systemic therapy regimen to randomisation, n (%)	<3 months	181 (62)	183 (63.1)
	3-6 months	59 (20.2)	56 (19.3)
	>6 months	52 (17.8)	51 (17.6)
Best response to most recent prior regimen, n (%)	CR or PR	73 (25)	68 (23.4)
	SD	103 (35.3)	96 (33.1)
	PD	111 (38.0)	116 (40.0)
	Unknown/Not reported	5 (1.7)	10 (3.4)

ALK=anaplastic lymphoma kinase; CNS=central nervous system; CR=complete response; ECOG PS=Eastern Cooperative Oncology Group performance status; EGFR=epidermal growth factor receptor; PD=progressive disease; PD-L1=programmed cell death-ligand 1; PR=partial response; SD=stable disease; TKI=tyrosine kinase inhibitor
*Percent membranous staining in ≥100 tumour cells.
Source: CS, adapted from Table 14

Overall, aside from the caveat that, in general, patients who participate in RCTs tend to be slightly younger and fitter than patients seen in clinical practice, the ERG considers that the characteristics of the patient population in CheckMate 057 are likely to be similar to the characteristics of patients treated in routine clinical practice in England.

4.2.4 Statistical approach adopted for the conduct and analysis of studies included in the systematic review

Information relevant to the statistical approach taken by the company has been taken from the clinical study report (CSR),[28] the trial statistical analysis plan (TSAP),[29] the trial protocol,[30] and from the CS.

Trial population

For the analysis of all efficacy outcomes, the intention-to-treat (ITT) population was used. Safety outcomes were analysed using the safety population, consisting of all patients who received study medication.

Outline of analyses

The company states that an interim OS analysis was scheduled to take place when at least 380 deaths had been reported. As a consequence of this interim review (18 March 2015 data-cut), the independent data monitoring committee (DMC) declared that the trial had reached its primary endpoint, and recommended that the trial be stopped (April 2015). The trial protocol[30] was consequently modified to provide a mechanism for eligible subjects who were originally randomised to the docetaxel treatment group to receive subsequent nivolumab therapy as part of a nivolumab extension phase. However the ERG notes that this affected a very small number of patients. The results from this interim analysis are based on a minimum follow-up of 13.2 months; in the ERG report this analysis is referred to as the "12-month interim analysis" for consistency with how the term is used in the CS.

The company also provides updated results with additional follow-up, on the basis of data from a 2 July 2015 data-cut. The results from this analysis are based on a minimum follow-up of 17.1 months; in the ERG report this analysis is referred to as the "18-month updated analysis" for consistency with how the term is used in the CS. The ERG notes that although the company states that updated results are available for OS only, PFS results at the 18-month updated analysis are presented in the CS.

The ERG was initially concerned that CheckMate 057 had been stopped early for benefit as previous technology appraisals have highlighted that early closure of cancer trials can lead to exaggerated treatment effects that are not borne out in the longer term.[31-34] However, considering the 18-month updated analysis results, the ERG is of the view that stopping the trial early does not appear to have biased the efficacy results in any way since the OS data are now mature and consistent with the findings from the 12-month interim analysis.

Efficacy outcomes

The definitions, and methods of analysis, for the primary and key secondary efficacy outcomes from CheckMate 057 are listed in Table 5. The ERG is satisfied that all of the outcomes were pre-specified in the TSAP[29] and that all outcomes were fully reported in the CSR.[28]

Table 5 Analysis strategy for key efficacy endpoints

Endpoint	Definition	Statistical method
Primary outcome		
OS	Time between the date of randomisation and the date of death	OS was analysed with the use of a two-sided log-rank test stratified according to prior maintenance treatment and line of therapy. HR and CI were estimated with the use of a stratified Cox PH model. Survival curves and rates were estimated with the use of the K-M method
Secondary outcomes		
Investigator-assessed ORR*	The number of patients whose BOR was either a confirmed complete or partial response, as determined by the investigator, divided by the number of randomised patients	ORR was computed in each treatment group along with the exact 95% CI using Clopper-Pearson method. An estimate of the difference in ORRs and corresponding 95% CI was calculated using Cochran-Mantel-Haenszel methodology and adjusted by the same stratification factors as in primary analysis of OS
Investigator-assessed PFS*	Time from randomisation to the date of the first documented tumour progression as determined by the investigator using RECIST 1.1 criteria or death due to any cause	PFS was analysed with the use of a two-sided log-rank test stratified according to prior maintenance treatment and line of therapy. HR and CI were estimated with the use of a stratified Cox PH model. Survival curves and rates were estimated with the use of the K-M method

BOR=best confirmed objective response; CI=confidence interval; CR=complete response; HR=hazard ratio; KM=Kaplan-Meier; ORR=objective response rate; OS=overall survival; PFS=progression-free survival; PH=proportional hazards; PR=partial response; RECIST= Response Evaluation Criteria in Solid Tumors; CI=confidence interval
*Non-conventional benefit (i.e. a reduction in the size or number [or both] of target lesions with simultaneous appearance of new lesions or initial progression followed by either tumour reduction or no further progression for at least two tumour assessments) in patients treated beyond initial progression was not included in response-based analyses (ORR or PFS)
Source: CS, adapted from Table 10 and Table 11

Censoring methods

For OS, subjects without documentation of death were censored on the last date that the subject was known to be alive.

For PFS, subjects who did not progress or die were censored on the date of their last evaluable tumour assessment. Subjects that did not have any on study tumour assessments and did not die were censored on the date they were randomised. Subjects who started any subsequent anti-cancer therapy without a prior reported progression were censored at the last evaluable tumour assessment prior to initiation of the subsequent anti-cancer therapy.

Proportional hazard ratios

The analyses carried out by the company to generate PFS and OS hazard ratios (HRs) were conducted using Cox proportional hazards (PH) modelling. The validity of this method relies on the hazards of the two comparative drugs being proportional. To investigate the assumption of PH, the company inspected log-log plots (log cumulative hazard versus log time); if the curves for each treatment arm were approximately parallel, it was assumed that PH was valid. The company also performed statistical tests, namely the Global Test for PH assumption,[35] and a supremum test for PH assumption.

The results of the testing carried out by the company (see Appendix 10.2) indicate that the assumption of PH is violated for both OS and PFS data for CheckMate 057. Consequently, it

is inappropriate to summarise these data by using HRs and 95% confidence intervals (CI) estimated by a Cox PH model. For this reason, the ERG considers that HRs ought to be interpreted with caution. The ERG would have preferred for the company to provide a rationale for using this approach and an explanation as to why alternative approaches were not considered.

ERG assessment of statistical approach

A summary of the checks made by the ERG regarding the statistical approach adopted by the company to analyse data from CheckMate 057 is provided in Table 6.

Table 6 ERG assessment of statistical approach used to analyse CheckMate 057 data

Component	Statistical approach	ERG comments
Sample size calculation	Provided in the CS (pg 65-66)	The ERG considers that the methods used to calculate the sample size are correct
Protocol amendments	Provided in the CSR[28] (Section 4.5)	The ERG notes that all protocol amendments were carried out prior to the interim analysis, with the exception of the modification to the protocol after stopping the trial, when patients in the docetaxel arm were allowed to switch to receive nivolumab. All other amendments were not driven by the results of the trial, and are therefore not of concern
Missing data approach	Provided in the CS (pg 65-66)	The ERG is satisfied that the company took a suitable approach to handling missing data
Pre-specified subgroup analyses	Efficacy (OS, PFS, ORR) based on pre-study PD-L1 expression level Pre-specified expression level cut-off values of 1%, 5% and 10% were used Efficacy (OS, ORR and PFS) based on: • Age • Sex • Race • Region • Baseline ECOG PS • Smoking status • Presence of CNS metastases • Prior neoadjuvant vs. adjuvant treatment • Prior use of maintenance therapy • Line of therapy • EGFR mutation status • ALK translocation status • KRAS mutation status • MET receptor status • Cell type • Time from diagnosis to randomisation • Time from completion of most recent regimen to randomisation	The ERG is satisfied that the results of all subgroup analyses are provided in the CSR[28]
Adverse events	Safety was assessed through summaries of deaths, AEs, serious AEs, AEs leading to discontinuation of study drug, AEs leading to dose delay, Select AEs and specific clinical laboratory assessments	The ERG is satisfied that the results of all the AE data analyses are provided in the CSR[28]
Health related quality of life	Disease-related symptom improvement rate by week 12 as measured by the LCSS Overall health status using the EQ-5D Index and Visual Analogue Scale (exploratory outcome)	The ERG is satisfied that the methodology used to analyse HRQoL data is appropriate

AE=adverse event; CNS=central nervous system; CS=company submission; CSR=clinical study report; ECOG=Eastern Cooperative Oncology Group; EQ-5D=EuroQol-5 Dimensions; ERG=Evidence Review Group; HRQoL=health related quality of life; MET=mesenchymal epithelial transition; ORR=objective response rate; OS=overall survival; PD-L1=programmed cell death 1 ligand; PFS=progression-free survival; PS=performance status
Source: CS, CSR[28] and ERG comment

4.2.5 Assessment of risk of bias of included studies

The ERG is generally satisfied with the assessments of risk of bias that are presented in the CS (see Table 7). CheckMate 057 was not a double-blind trial but the ERG concurs that blinding patients and health professionals would have been difficult for a number of reasons i.e. different dosing regimes, different side effect profiles.

Table 7 Quality assessment of CheckMate 057

Was randomisation carried out appropriately?	Yes
Was the concealment of treatment allocation adequate?	Yes
Were the groups similar at the outset of the study in terms of prognostic factors?	Yes
Were the care providers, participants and outcome assessors blind to treatment allocation?	No
Were there any unexpected imbalances in drop-outs between groups?	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes
How closely do the RCT(s) reflect routine clinical practice?	Patients included in CheckMate 057 are thought to reflect patients seen in UK clinical practice • Comparator in the study is docetaxel, which represents standard of care in previously treated patients in the UK. • First-line treatment in the UK is a platinum-based chemotherapy; patients who had received a platinum-based chemotherapy were included in the study. • Doses for both nivolumab and docetaxel used in the study are reflective of UK clinical practice. • Baseline characteristics are similar to patients seen in UK clinical practice (e.g. ex-smokers).

RCT=randomised controlled trial
Source: CS, adapted from Table 15

4.2.6 Results from the studies included in the systematic review

Overall survival

The results for OS are provided in Table 8. Nivolumab was found to significantly improve survival in comparison to docetaxel (HR=0.73, 95% CI: 0.59 to 0.89; p=0.002) at the 12-month interim analysis. Median OS was 2.8 months longer for patients in the nivolumab arm than for patients in the docetaxel arm. OS rates were also higher for nivolumab patients than docetaxel patients (50.5% versus 30.9%). This treatment benefit with nivolumab was shown

to be consistent over time, as results from the updated analysis suggest that OS rates at 18 months are still higher in the nivolumab arm than in the docetaxel arm (39% versus 23%).

Table 8 CheckMate 057 OS results

	CheckMate 057	
	Nivolumab (n=292)	**Docetaxel (n=290)**
12-month interim analysis		
Events, n (%)	190 (65.1)	223 (76.9)
Stratified log-rank test p value		0.002
HR for death (95% CI) at 12 months		0.73 (0.59 to 0.89)
Median OS, months (95% CI)	12.2 (9.7 to 15.0)	9.4 (8.1 to 10.7)
OS rate at 12 months (95% CI)	50.5 (44.6 to 56.1)	39.0 (33.3 to 44.6)
18-month updated analysis		
OS rate at 18 months (95% CI)	39 (34 to 45)	23 (19 to 28)

CI=confidence interval; HR=hazard ratio; OS=overall survival
Source: CS, Table 16

The company also provides Kaplan-Meier (K-M) curves to demonstrate OS, as shown in Figure 2. For the first 7 months, patients in the docetaxel arm are less likely to have an OS event than patients in the nivolumab arm. At 7 months, the K-M curves begin to separate, and nivolumab appears to improve OS in comparison to docetaxel for the remainder of the follow-up period. The company states that pseudo-progression (tumours that initially increase as a result of the treatment action before shrinking/stabilising) may be responsible for the 7-month delay in OS benefit for patients treated with nivolumab.

As the assumption of PH for the two treatments is violated, HRs should be interpreted with caution.

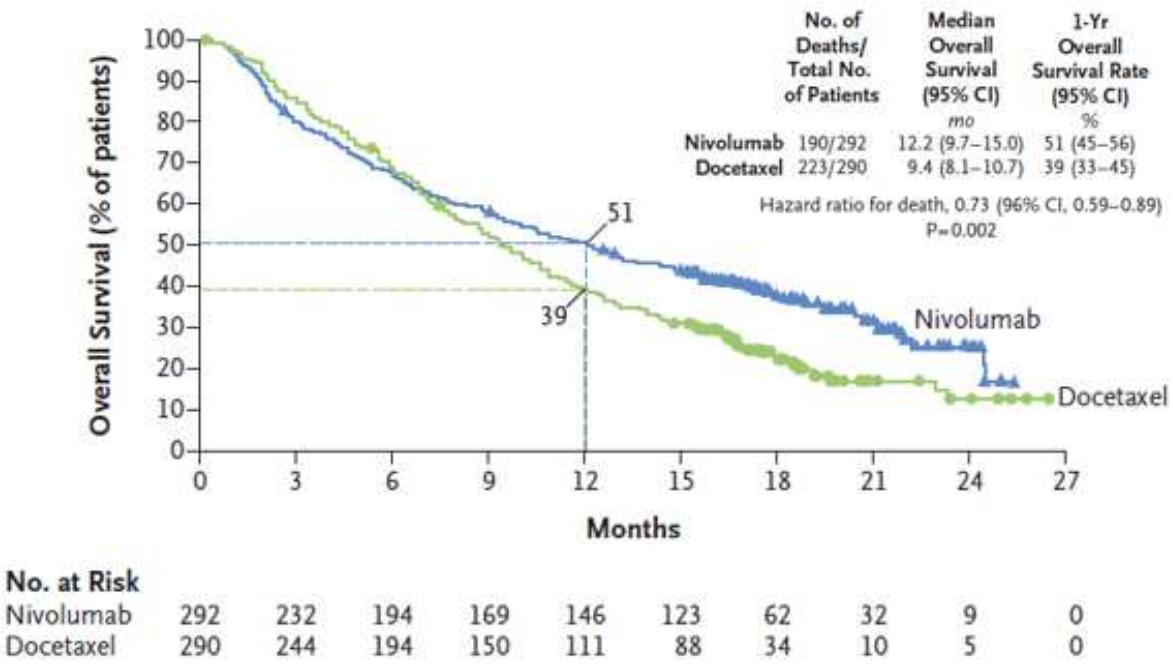


Figure 2 CheckMate 057 K-M overall survival plot – all randomised patients

Abbreviations: CI=confidence interval
Note: The analysis included all the patients who underwent randomisation. Symbols indicate censored observations, and horizontal lines the rates of OS at 1 year.
Source: CS, Figure 11

K-M curves with the additional follow-up from the 18-month updated analysis are also provided in the CS (Figure 12). The HR at the time of the updated analysis was consistent with the HR reported at the time of the interim analysis, suggesting that nivolumab statistically significantly improves OS in comparison to docetaxel (HR=0.72, 95% CI: 0.60 to 0.88; p=0.0009). Once again, the ERG is of the opinion that this HR should be interpreted with caution. 18-month OS rates were also higher for nivolumab patients than for docetaxel patients (39% versus 23%).

Progression-free survival

The results for PFS are provided in Table 9. There was no statistically significant difference between nivolumab and docetaxel in terms of PFS at the time of the 12-month interim analysis (HR=0.92, 95% CI: 0.77 to 1.11; p=0.3932). Median PFS was 1.9 months longer for patients in the docetaxel arm than for patients in the nivolumab arm (4.2 months versus 2.3 months). However, 12-month PFS rates were higher for nivolumab patients than for patients receiving docetaxel (18.5% versus 8.1%).

Table 9 CheckMate 057 PFS results

PFS at 12-month interim analysis	CheckMate 057	
	Nivolumab (n=292)	Docetaxel (n=290)
Events, n (%)	234 (80.1)	245 (84.5)
Stratified log-rank test p value		0.3932
HR for progression or death (95% CI) at 12 months		0.92 (0.77 to 1.11)
Median, months (95% CI)	2.3 (2.2 to 3.3)	4.2 (3.5 to 4.9)
PFS rate at 12 months (95% CI)	18.5 (14.1 to 23.4)	8.1 (5.1 to 12.0)

CI=confidence interval; HR=hazard ratio; PFS=progression-free survival
Source: CS, Table 17

The nivolumab and docetaxel K-M curves provided by the company, as shown in Figure 3, explain why different measures of effect for PFS favour different treatments. It is clear that the K-M curves for the two treatments show markedly different profiles. For the first 7 months, patients in the docetaxel arm are less likely to have a PFS event than those in the nivolumab arm, resulting in median PFS values which favour docetaxel. At 7 months, the K-M curves begin to separate, and nivolumab appears to improve PFS in comparison to docetaxel for the remainder of the follow-up period. Hence, the PFS rate at 12 months favours nivolumab over docetaxel. The company states that pseudo-progression may be responsible for the 7-month delay in PFS benefit for patients treated with nivolumab, although a number of theories exist for this delay, and the exact underlying mechanism is unclear.

As the assumption of PH for the two treatments is violated the HR should be interpreted with caution.

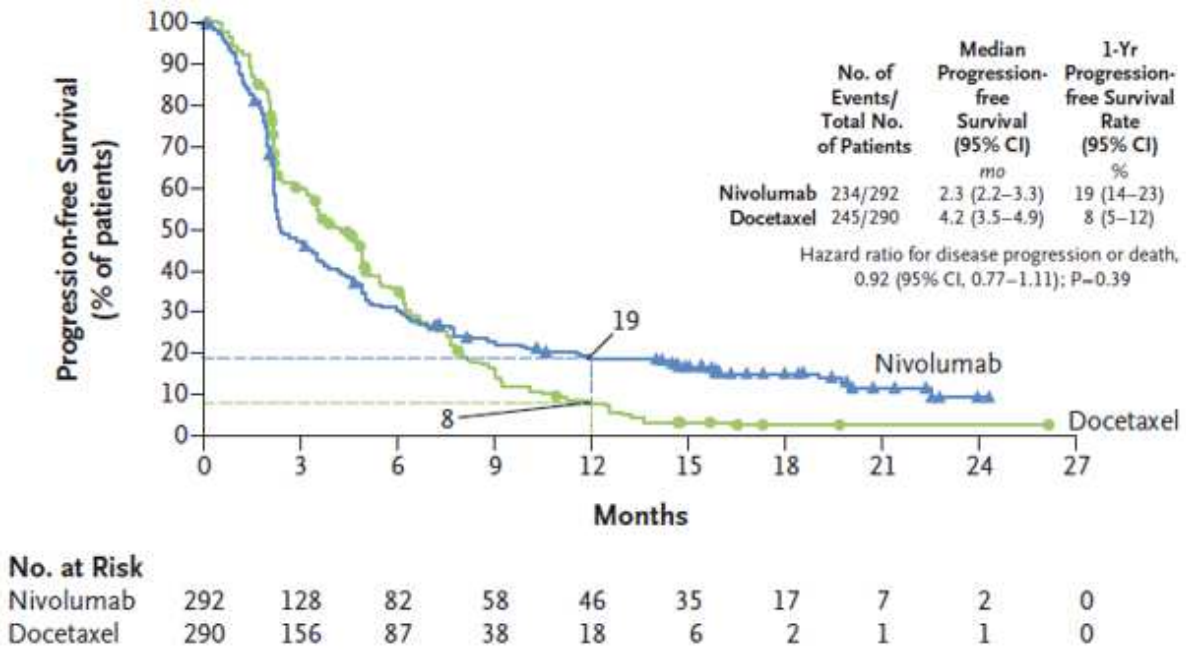


Figure 3 CheckMate 057 K-M PFS plot – all randomised patients in the study

CI=confidence interval
Source: CS, Figure 13

At the 18-month updated analysis, the HR for PFS with nivolumab versus docetaxel was 0.91 (95% CI: 0.76 to 1.09). Once again, the ERG does not believe that the use of HR is an appropriate way to summarise the PFS data. The ERG also notes that the company states that only OS results from the 18-month updated analysis are available, but then proceeds to present PFS results from this same time point.

Response

The ORR results are provided in Table 10. Nivolumab was found to statistically significantly improve ORR in comparison to docetaxel (OR 1.7, 95% CI: 1.1 to 1.6; p=0.02). Four patients in the nivolumab group (1.4%) achieved a complete response (CR) compared with one patient (0.3%) in the docetaxel group. Median time to treatment response (TTR) was slightly shorter in the nivolumab arm than in the docetaxel arm (2.1 versus 2.6 months), and median duration of response (DoR) was found to be much longer in the nivolumab arm than in the docetaxel arm (17.2 versus 5.6 months). These findings are also demonstrated by the characteristics of responses provided by the company in Figure 14 of the CS. Patients achieving a response in either arm usually responded early on in the follow-up period, and often by the time of the first scan.

Table 10 CheckMate 057 summary of response analyses

	CheckMate 057	
	Nivolumab (n=292)	**Docetaxel (n=290)**
ORR		
n, responders	56	36
% of patients (95% CI)	19 (15 to 24)	12 (9 to 17)
Odds ratio estimate (95% CI)	1.7 (1.1 to 2.6)	
P value	0.02	
TTR		
Median, months	2.1	2.6
Min-Max (months)	1.2-8.6	1.4-6.3
DOR		
N, responders	56	36
Median, months (95% CI)	17.2	5.6
Min-Max (months)	1.8-22.6+	1.2+-15.2+

CI=confidence interval; DOR=duration of response; ORR=objective response rate; TTR=time to treatment response
The + symbol indicates a censored value. The value of 1.2 was censored because the patient discontinued treatment without disease progression, and the other values were censored because the response was ongoing at the time of the analysis.
Source: CS, Table 18

Treatment beyond progression

The CheckMate 057 protocol outlines how subjects treated with nivolumab were permitted to continue treatment beyond initial Response Evaluation Criteria in Solid Tumours (RECIST 1.1) defined progressive disease (PD), as long as they met specific criteria (trial protocol,[30] Section 4.3.4). For the nivolumab treatment group, 71 patients received treatment beyond progression, 16 of whom demonstrated a non-conventional pattern of benefit. Non-conventional benefit was experienced by patients who had not experienced a best objective response of partial response (PR) or CR prior to initial progression and met at least one of the following criteria:

- Criterion 1: Appearance of a new lesion followed by decrease from baseline of ≥ 10% in the sum of the target lesions (12 patients).
- Criterion 2: Initial increase from nadir ≥ 20% in the sum of the target lesions followed by reduction from baseline of ≥ 30% (no patients).
- Criterion 3: Initial increase from nadir ≥ 20% in the sum of the target lesions or appearance of new lesion followed by at least two tumour assessments showing no further progression defined as a 10% additional increase in sum of target lesions and new lesions (7 patients).

Furthermore, 14 patients had extended nivolumab treatment (defined as >3 doses received after initial progression) and extended OS (defined as more than the median OS of 12.2 months in the nivolumab group) after initial progression but did not meet the criteria for non-conventional benefit as defined above.

Subgroup analyses

The company conducted subgroup analyses for OS using a range of pre-specified characteristics. The company presents results for subgroup analyses performed at the time of the 12-month interim analysis (CS, Figure 16). The results of the subgroup analyses using data from the 18-month updated analysis are provided in Appendix 10.3.

OS benefit for nivolumab was observed across most pre-specified subgroups, except for the following: third-line therapy, 'Rest of the World' region, never smokers and EGFR mutation-positive status. The company observes that CIs in these subgroups were wide due to small subgroup sizes, that the study was not powered to identify significant differences in these subgroups, and that the "Rest of the World" subgroup may also have been confounded by smoking status.

The ERG agrees with the company's interpretation of the results of the subgroup analyses, although to inform further investigation, the ERG requested the corresponding p values for the tests for interaction for the performed subgroup analyses. Statistically significant

subgroup differences were observed for line of therapy (p=0.0431), region (p=0.0006), and smoking status (p=0.0446), suggesting that treatment effect is statistically significantly greater for second-line patients than third-line patients, US/Canada and Europe patients than "Rest of the World" patients, and smokers than never-smokers.

In terms of PFS, HRs favoured nivolumab in comparison to docetaxel for all pre-specified subgroups, except for third-line therapy, Europe and 'Rest of the World' region, females, never smokers, kirsten rat sarcoma 2 viral oncogene homolog (KRAS) mutation not detected, EGFR mutation-positive status, ALK translocation status not reported, and prior adjuvant treatment (CSR,[28] Section 7.4.1).

Efficacy results by EGFR mutation status

The efficacy of nivolumab was analysed according to EGFR mutation status. Many patients had unknown EGFR mutation status as the test was not mandatory as per the study protocol;[30] EGFR mutation status was either reported by the investigator or collected from case-report forms if the patients had been tested for EGFR mutation status as part of routine care prior to study entry.

The results for OS (Table 11) suggest that the nivolumab benefit observed for the overall trial population in comparison to docetaxel is also seen within the EGFR mutation-negative/unknown population (HR=0.7, 95% CI: 0.6 to 0.8). However, results for the EGFR mutation-positive patients suggest that there is no statistically significant difference between nivolumab and docetaxel within this population, with the HR actually favouring docetaxel (HR=1.18, 95% CI: 0.69 to 2.00).

As the ERG has previously mentioned, the presented HRs are not an appropriate way to summarise treatment effect. In the EGFR mutation-negative unknown population, median OS was 3.5 months longer for patients in the nivolumab arm than for patients in the docetaxel arm (12.8 months versus 9.3 months). However, in the EGFR mutation-positive patients, median OS was 2.3 months longer for docetaxel patients than for nivolumab patients (11.5 months versus 9.2 months).

The ERG notes that the p value for interaction provided for the EGFR mutation status subgroup analysis at the 18-month data-cut was not statistically significant (p=0.4689).

Table 11 CheckMate 057 treatment effect on OS by EGFR mutation status

EGFR mutation status	Nivolumab		Docetaxel		HR for nivolumab vs. docetaxel (95% CI)
	Event no. (pt no.)	Median OS (95% CI)	Event no. (pt no.)	Median OS (95% CI)	
12-month interim analysis					
Positive (n=82)	31 (44)	9.2 (5.2 to 13.1)	25 (38)	11.5 (5.8 to 17.8)	1.2 (0.7 to 2.0)
Negative (n=340)	104 (168)	13.6 (10.4 to 18.4)	133 (172)	9.3 (7.7 to 10.7)	NR
Unknown (n=160)	55 (80)	11.3 (7.7 to 15.7)	65 (80)	9.3 (7.2 to 12.0)	NR
Negative/unknown combined (n=500)	159 (248)	12.8 (10.0 to 15.7)	198 (252)	9.30 (8.0 to 10.6)	0.7 (0.6 to 0.9)
18-month updated analysis					
Positive (n=82)	33 (44)	9.31 (5.2 to 15.7)	27 (38)	11.53 (5.8 to 17.0)	1.1 (0.7 to 1.9)
Negative/unknown combined (n=500)	173 (248)	12.8 (10.0 to 15.7)	209 (252)	9.3 (8.0 to 10.6)	0.7 (0.6 to 0.8)

CI=confidence interval; HR=hazard ratio; OS=overall survival; no=number; pt=patient
Source: CS, Table 20 and text in Section 4.8.2 and clarification response-question A3

The company also provided a K-M plot (CS, Figure 17) to demonstrate OS in the EGFR mutation-negative/unknown patients; the K-M curves for nivolumab and docetaxel suggest a similar pattern of survival to the overall patient population. For the first 7 months, patients in the docetaxel arm are less likely to have an OS event than those in the nivolumab arm. At 7 months, the K-M curves begin to separate, and nivolumab appears to improve OS in comparison to docetaxel for the remainder of the follow-up period.

In the CS the company reports that the 1-year OS rate for docetaxel is higher than expected in CheckMate 057 and suggests that this underestimates the true survival benefit of nivolumab (CS, Section 4.71). For comparison, the company cites OS results from the docetaxel arm of CheckMate 017. However, the ERG is unclear why this study was chosen for comparison as this trial only included squamous patients. A more appropriate study for comparison may be the TAILOR[36] study or one of the other three studies[24,26,27] with a docetaxel arm that were included in the ITC. The median OS and 1-year survival results in these studies are reported in Table 12 and demonstrate similar rates to those found in CheckMate 057.

Table 12 Comparison of OS in docetaxel arms of comparator studies

Study	Median OS (95% CI) mths	1 year OS (95% CI)
CheckMate 057	9.4 (8.1 to 10.7)	39.0% (33 to 45)
CheckMate 017[37]	6.0 (5.1 to 7.3)	24% (17 to 31)
Tailor[36]	8.2 (5.8 to 10.9)	39.6% (36.1 to 43.4)
Lume-Lung 1 adenocarcinoma[24]	10.3	44.7% (38.9 to 49.8)
ISTANA[26]	12.2 (NR)	NR
V-15-32[27]	14 (11.7 to 16.5)	53.7% (NR)

OS=overall survival; NR=not reported

The HRs for PFS (Table 13) suggest that there were no statistically significant differences between nivolumab and docetaxel for either the EGFR mutation-positive or EGFR mutation-negative/unknown patients. Once again, these HRs need to be interpreted with caution. For EGFR mutation-positive patients, median PFS was 2.7 months longer for patients in the docetaxel arm than for patients in the nivolumab arm (4.8 months versus 2.1 months). In the EGFR mutation-negative/unknown patients, median PFS was 1.6 months longer for docetaxel patients than for nivolumab patients (4.2 months versus 2.6 months). The company also provided PFS results for the EGFR mutation negative/unknown patients from the 18-month updated analysis; the reported HRs concurred with the HRs reported at the 12-month interim analysis.

Table 13 CheckMate 057 treatment effect on PFS by EGFR mutation status

EGFR mutation status	N	Nivolumab (n=292)		Docetaxel (n=290)		HR for nivolumab vs. docetaxel (95% CI)
		Events (patients)	Median PFS (95% CI)	Events (patients)	Median PFS (95% CI)	
12-month interim analysis						
Positive	82	39 (44)	2.1 (1.6 to 3.3)	29 (38)	4.8 (2.1 to 6.9)	1.5 (0.9 to 2.47
Negative	340	131 (168)	3.1 (2.2 to 4.2)	144 (172)	3.9 (3.5 to 4.9)	-
Unknown	160	64 (80)	2.3 (2.1 to 5.0)	72 (80)	4.7 (2.2 to 5.5)	-
Negative/unknown combined	500	195 (248)	2.6 (2.2 to 3.7)	216 (252)	4.2 (3.5 to 4.9)	0.8 (0.7 to 1.0)
18-month updated analysis						
Negative/unknown combined	500	198 (248)	2.6 (2.2 to 3.7)	218 (252)	4.2 (3.5 to 4.9)	0.82 (0.7 to 1.0)

CI=confidence interval; PFS=progression-free survival; no=number; pt=patient
Source: CS, Table 21 and text in section 4.8.2

The company also provided a K-M plot (CS, Figure 18) to demonstrate PFS in the EGFR mutation-negative/unknown patients; the K-M curves for nivolumab and docetaxel suggest a similar pattern of PFS to the overall patient population.

4.2.7 Efficacy results by PD-L1 expression level

The efficacy of nivolumab was also analysed according to PD-L1 status in terms of both OS and PFS; 78% (455/582) of randomised patients had an evaluable PD-L1 status (231/292 in the nivolumab arm and 224/290 in the docetaxel arm). The company used three different categorisations to investigate the impact of PD-L1 status on treatment efficacy (<1% versus ≥1%, <5% versus ≥ 5%, and <10% versus ≥ 10%).

Firstly, the company used HRs to summarise treatment effect for each of the subgroups (CS, Figure 19). HRs are not an appropriate way to measure treatment effectiveness for nivolumab in comparison to docetaxel, and therefore the ERG does not believe that the results presented in Figure 19 of the CS are valid.

Instead, the ERG considered the median OS results in order to evaluate how PD-L1 status might influence the efficacy of nivolumab. The company provided K-M graphs and median OS for each subgroup for each categorisation of PD-L1 status, with additional follow-up from the 18-month updated analysis (CS, Figure 20 and Figure 21). Median OS was 10.5, 9.8, and 9.9 months for nivolumab patients compared to 10.1, 10.1, and 10.3 for docetaxel subjects in PD-L1 negative subgroups defined by the <1%, <5%, and <10% expression levels, respectively.

ORR results by PD-L1 expression status are shown in Table 14. Higher ORRs were observed in nivolumab patients versus docetaxel patients for high expressors at each of the pre-specified PD-L1 expression levels (31% versus 12% by the ≥1% expression level, 36% versus 13% by the ≥5% expression level and 37% versus 13% by the ≥10% expression level). In low expressors, ORRs were lower in nivolumab patients in comparison to docetaxel patients, (9% versus 15% by the <1% expression level, 10% versus 14% by the <5% expression level and 11% versus 14% by the <10% expression level).

Table 14 Outcomes nivolumab vs. docetaxel by baseline PD-L1 expression level

	Baseline PD-L1 expression level						**Not quantifiable†**
	1%*		**5%***		**10%***		
	<1%	**≥1%**	**<5%**	**≥5%**	**<10%**	**≥10%**	
Nivolumab							
n (%)	108 (47)	123 (53)	136 (59)	95 (41)	145 (63)	86 (37)	61 (21)
ORR‡ n (%) [95% CI]	10 (9) [5 to 16]	38 (31) [23 to 40]	14 (10) [6 to 17]	34 (36) [26 to 46]	16 (11) [6 to 17]	32 (37) [27 to 48]	8 (13) [6 to 24]
Median DOR, months (95% CI) N	18.3 (4.2 to NE) 10	16.0 (8.4 to NE) 38	18.3 (5.5 to NE) 14	16.0 (8.4 to NE) 34	18.3 (7.5 to NE) 16	16.0 (6.9 to NE) 32	7.3 (2.2 to NE) 8
Docetaxel							
n* (%)	101 (45)	123 (55)	138 (62)	86 (38)	145 (65)	79 (35)	66 (23)
ORR‡ n (%) [95% CI]	15 (15) [9 to 23]	15 (12) [7 to 19]	19 (14) [9 to 21]	11 (13) [7 to 22]	20 (14) [9 to 21]	10 (13) [6 to 22]	6 (9) [3 to 19]
Median DOR, months (95% CI) n	5.6 (4.2 to 9.9) 15	5.6 (3.0 to 5.7) 15	5.6 (4.2 to 7.1) 19	5.6 (3.0 to 7.0) 11	5.6 (4.2 to 7.1) 20	5.6 (1.6 to 6.2) 10	6.6 (2.8 to 14.2) 6
OR‖ (95% CI)	0.6 (0.2 to 1.5)	3.2 (1.6 to 6.7)	0.7 (0.3 to 1.6)	3.8 (1.7 to 9.0)	0.8 (0.4 to 1.7)	4.1 (1.8 to10.1)	1.5 (0.4 to 5.6)

CI=confidence interval; DOR=duration of response; NE=not evaluable; OR=odds ratio; ORR=objective response rate; PD-L1=programmed death-ligand 1
* Number and percent of evaluable patients with membranous staining at the respective expression level in ≥ 100 tumour cells.
† Number and percent of randomised patients with PD-L1 expression not quantifiable.
‡ Confirmed complete and partial responses per RECIST v1.1 criteria, as assessed by the investigator. CI based on the Clopper-Pearson method.
‖ Ratio of nivolumab over docetaxel
Source: C
S, Table 22

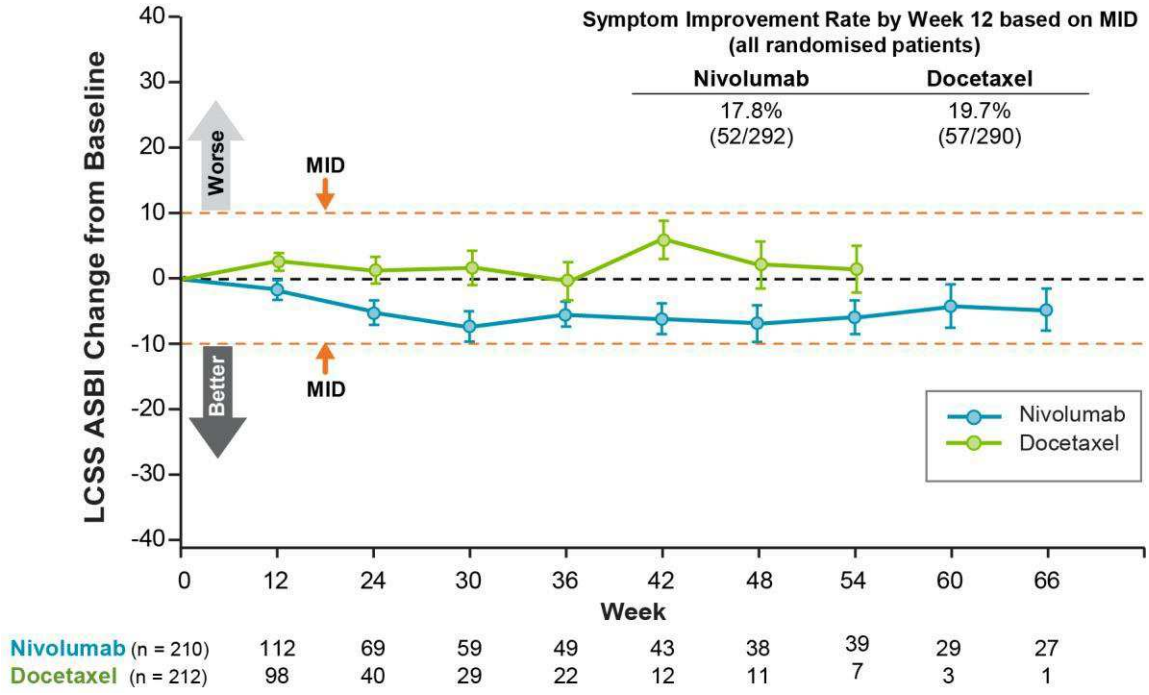
Both high expressors and low expressors experienced a benefit in terms of DOR, which was greater in nivolumab patients than in docetaxel patients for all pre-specified PD-L1 expression levels.

4.2.8 Health related quality of life

In CheckMate 057 the effect of nivolumab treatment on patients' HRQoL was measured according to the Lung Cancer Symptom Scale (LCSS) Average Symptom Burden Index (ASBI) score (which is the mean score computed from the six symptom-specific questions of the LCSS), the EuroQol 5-Dimensions utility index (EQ-5D) and the EuroQol Visual Analogue Scale (EQ-VAS) at each assessment point. As described in Table 55 of the CS and on pg 258 of the CSR,[28] assessments for both EQ-5D/EQ-VAS and LCSS were performed at every other cycle in the nivolumab arm (i.e. at 4 weeks, 8 weeks, 12 weeks, etc) or at every cycle in the docetaxel arm (i.e. 3 weeks, 6 weeks, 12 weeks, etc) for the first 24 weeks on study, then every 6 weeks thereafter in both arms for the remainder of the study. The differences in time of initial measurement reflect the fact that nivolumab is administered every 2 weeks while docetaxel is administered every 3 weeks. The scores were also assessed at 100 days following the last dose administered to patients and every 3 months for the first 12 months, and every 6 months after. Disease-related symptom improvement rate is defined as a decrease of 10 points or more from baseline in average symptom burden by week 12.

LCSS questionnaire completion rates were ≥75% at baseline and ≥65% through to week 66. From week 66 compliance rates were lower but remained at ≥45%, though by week 42 for the docetaxel group the number of available patients had fallen to below 20.

Results of the LCSS ASBI are shown in Figure 15 of the CS. Mean (standard deviation) baseline scores for LCSS ASBI were 26.2 (16.2) for nivolumab patients and 24.4 (15.5) for docetaxel patients. By week 12, the rate of disease-related symptom improvement was comparable between the groups i.e. 17.8% for nivolumab patients and 19.7% for docetaxel patients. The ERG concurs that this is correct as any difference would not be clinically significant as neither curve crosses the minimally important difference i.e. a change of ≥10 points (See Figure 4).



Source Horn, Brahmer, Reck, Borghaei, Spigel, Steins, *et al.* [38]

Abbreviations: ASBI = Average Symptom Burden Index; LCSS = Lung Cancer Symptom Scale; MID = Minimally Important Difference

Higher scores indicate greater symptom burden. Mean (standard deviation) scores at baseline were 24.8 (15.9) for nivolumab and 24.4 (15.8) for docetaxel. Only time points that had patient-reported outcome data available for ≥ 5 patients in either treatment arm are plotted on the graph. MID consists of a change of ≥ 10 points.

Figure 4 CheckMate 057: change in LCSS ASBI (on treatment)

The company also states (CS, p80) that, at the two follow-up visits after treatment discontinuation, the average symptom burden for both groups indicated a worsening of symptoms relative to baseline (range 3.6-6.3). However, no further details are reported in the CS, so the ERG cannot provide further comment.

Completion rates for the EQ-VAS were high and were greater than 50% for most of the on-treatment assessments and were similar across the two arms of the tria. 'The average EQ-VAS increased over time for both treatment groups (although the increase began later for docetaxel patients), indicating better overall health status for patients remaining on treatment. The average EQ-VAS score exceeded the average baseline score by more than the 7-point MID from week 16 through week 72 in the nivolumab group and from week 36 through week 48 for the docetaxel group. For both treatment groups, the EQ-VAS assessments in the follow-up visits following discontinuation returned to values in the region of the baseline scores (range: 60.6-66.4)'.

4.2.9 Adverse events

Comparative safety data from CheckMate 057 demonstrated that nivolumab had a more favourable safety profile than docetaxel (Table 15). There was one drug related death in

each arm of the trial. Any Grade treatment related AEs were lower in the nivolumab arm (69% versus 88%) as were Grade 3-4 AE (10% versus 54%). The overall number of AEs is similar in each group.

Table 15 Summary of safety profiles in CheckMate 057

Type of AE	Patients with each type of AE (%)	
	Nivolumab (n=287)	Docetaxel (n=268)
All cause and any Grade AE	280 (98)	265 (99)
All cause Grade 3 to 4 AE	132 (46)	180 (67)
Treatment related AE- all	199 (69)	236 (88)
All cause AE leading to discontinuation	48 (17)	58 (22)
All cause Grade 3 to 4 AE leading to discontinuation†	38 (13)	34 (13)
All cause Select AEs Grade 3 to 4		
Endocrine	0 (0)	0 (0)
Gastrointestinal	2 (1)	3 (1.1)
Hepatic	3 (1.0)	2 (0.7)
Pulmonary	4 (1.4)	1 (0.4)
Renal	0 (0)	0 (0)
Skin	2 (0.7)	0 (0)

AE=adverse event; SAE=serious adverse event
Source: CS, Table 32

The AE profile of nivolumab is different to that of standard chemotherapy because of the action of the drug and therefore the company describes a set of 'Select AEs'. The company claims that these Select AEs are manageable and patients may be successfully treated using the recommended treatment algorithm guidelines provided in the Summary of Product Characteristics.[39] The Select AEs are defined as immune-related adverse events (irAEs) that may require more frequent monitoring and treatment with immune modulating medications. A more detailed list of AEs that occurred in ≥5% of patients is presented in Table 16. As can be seen the most important differences are the increased rates of AE related to neutropenia and febrile neutropenia in the docetaxel patients.

The company provides additional clinical data from two non-randomised studies[22,23] of nivolumab (in patients with a variety of different types of cancer including NSCLC, melanoma and renal cancer); the safety profile of nivolumab in these two studies is similar to the safety profile of nivolumab that was seen in CheckMate 057.

Table 16 Summary of treatment related AE in ≥5% of treated patients in CheckMate 057

	Nivolumab n=287		Docetaxel n=268	
	Any Grade	Grade 3–4	Any Grade	Grade 3–4
Event	Number of patients with an event (%)			
Any event	199 (69)	30 (10)	236 (88)	144 (54)
Fatigue	46 (16)	3 (1)	78 (29)	13 (5)
Nausea	34 (12)	2 (1)	70 (26)	2 (1)
Decreased appetite	30 (10)	0	42 (16)	3 (1)
Asthenia	29 (10)	1 (<1)	47 (18)	6 (2)
Rash	27 (9)	1 (<1)	8 (3)	0
Pruritus	24 (8)	0	4 (1)	0
Diarrhoea	22 (8)	2 (1)	62 (23)	3 (1)
Hypothyroidism	19 (7)	0	0	0
Arthralgia	16 (6)	0	16 (6)	0
Vomiting	15 (5)	0	20 (8)	0
Constipation	13 (5)	0	21 (8)	2 (1)
Peripheral oedema	8 (3)	0	28 (10)	1 (<1)
Pyrexia	8 (3)	0	17 (6)	0
Myalgia	7 (2)	1 (<1)	30 (11)	0
Anaemia	6 (2)	1 (<1)	53 (20)	7 (3)
Dysgeusia	5 (2)	0	25 (9)	0
Paraesthesia	5 (2)	0	20 (7)	0
Pain	4 (1)	0	14 (5)	0
Peripheral neuropathy	3 (1)	0	25 (9)	3 (1)
Stomatitis	3 (1)	0	23 (9)	2 (1)
Mucosal inflammation	2 (1)	0	20 (7)	5 (2)
Lacrimation increased	1 (<1)	0	14 (5)	0
Alopecia	1 (<1)	0	67 (25)	0
Neutrophil count decreased	1 (<1)	1 (<1)	19 (7)	16 (6)
Neutropenia	1 (<1)	0	83 (31)	73 (27)
Febrile neutropenia	0	0	27 (10)	26 (10)
Leukopenia	0	0	27 (10)	22 (8)
White blood cell count decreased	0	0	22 (8)	12 (4)

Note: A patient may be recorded as having more than one adverse event within a category
Source: adapted from CS, Table 34

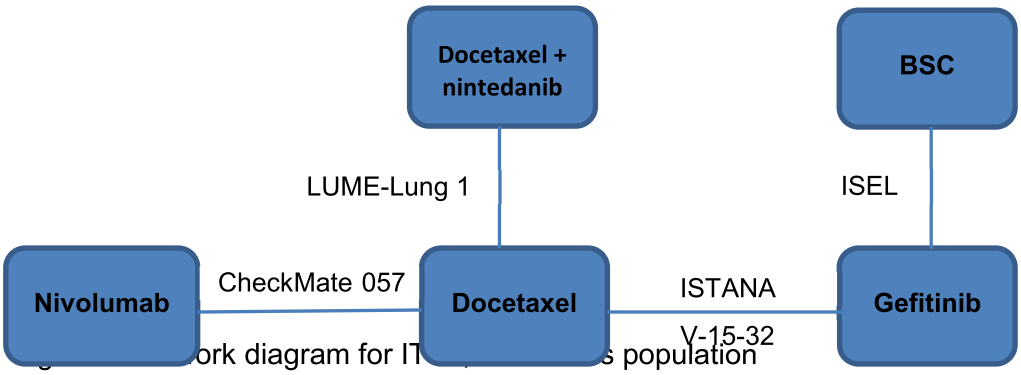
4.3 Critique of trials identified and included in the indirect comparisons

4.3.1 Studies identified for inclusion in the indirect comparisons

Using broad criteria, 33 trials were eligible for inclusion in the company's original systematic review of clinical effectiveness data; these trials included patients as described in the company's decision problem, i.e. previously treated patients with locally advanced or metastatic non-squamous NSCLC. This population is henceforth referred to as the all-comers population. However, as indicated in the NICE scope, the company also considered the subpopulations of EGFR mutation-negative/unknown and EGFR mutation-positive

patients. Thirty studies reported data for the EGFR mutation-negative/unknown population, and ten studies reported data for the EGFR mutation-positive population (see Appendix 7.1.4 Table 11).

For the all-comers population, i.e. any EGFR status, only five studies (Figure 5) contained data and formed a network which enabled ITCs between nivolumab and the relevant comparators to be carried out: CheckMate 057,[21] ISEL,[25] ISTANA;[26] LUME-Lung 1;[24] V-15-32.[27]



ork diagram for IT s population

BSC=best supportive care
References: ISEL,[25] ISTANA,[26] LUME-Lung 1,[24] V-15-32[27]
Source: CS, adapted from Appendix 7.17 Figure 6

For the EGFR mutation-negative/unknown population, only four studies (Figure 6) contained data and formed a network which enabled ITCs between nivolumab and the relevant comparators to be carried out: CheckMate 057;[21] LUME-Lung 1;[24] ISEL;[25] and ISTANA.[26]

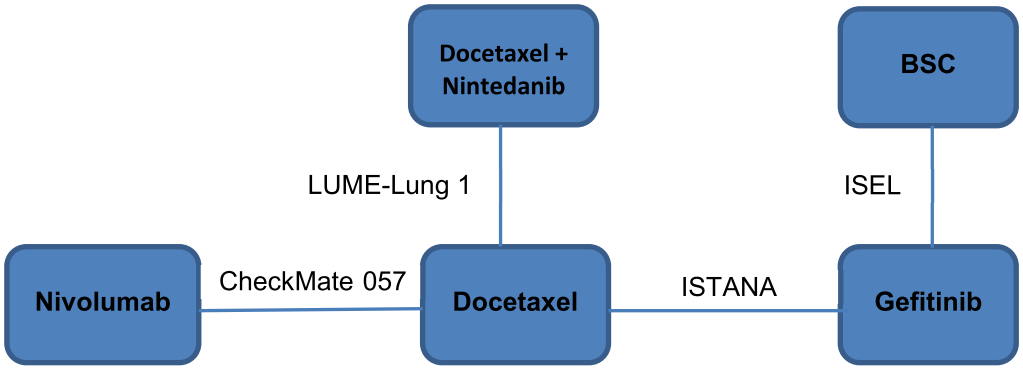


Figure 6 Network diagram for ITCs, EGFR mutation-negative/unknown population

BSC=best supportive care
References: ISEL,[25] ISTANA,[26] LUME-Lung 1,[24]
Source: CS, adapted from Appendix 7.17 Figure 4

For the EGFR mutation-positive population, no studies formed a network enabling ITCs between nivolumab and the relevant comparators to be carried out. No ITCs were conducted for this patient population.

The ERG did not identify any additional studies that met the company's eligibility criteria for inclusion in either the all-comers or EGFR mutation-negative/unknown networks. The key characteristics of the RCTs used to inform the ITCs are summarised in Table 17.

Table 17 Summary of RCTs reporting data for previously treated non-squamous NSCLC population and included in analyses

Study	Design	Location	Intervention/ comparators (n)	Duration	Patient population
Studies connected in BOTH networks (EGFR mutation-negative/unknown AND all-comers NSQ NSCLC)					
LUME-Lung 1[24]	Randomised, multicentre international, double-blind, placebo controlled, phase III	27 countries (211 centres)	Docetaxel (659) Docetaxel+nintedanib (655)	31·7 months	ECOG PS 0-1 At least one measurable target lesion One previous first-line chemotherapy regimen
ISTANA[26]	Randomised, multicentre, open-label, active-controlled, phase III	Korea (6 centres)	Docetaxel (79) Gefitinib (82)	NR	Age ≥18 years WHO PS 0-2 Histologically or cytologically confirmed NSCLC with stage IIIB or IV disease One previous platinum-based chemotherapy regimen Progressive or recurrent disease following previous chemotherapy
ISEL[25]	Randomised, multicentre international, double-blind, placebo controlled, phase III	28 countries (210 centres)	BSC (563) Gefitinib+BSC (1129)	7.2 months	Age ≥18 years WHO PS 0-3 Histologically or cytologically proven, locally advanced or metastatic NSCLC At least one previous platinum-based chemotherapy regimen
CheckMate 057[28]	Randomised, multicentre international, open-label, active-controlled, phase III study	22 countries (106 sites)	Nivolumab (292) Docetaxel (290)	30 months	Age ≥18 years Stage IIIB/Stage IV or recurrent or progressive non-squamous NSCLC ECOG PS 0-1 Failed at least one prior platinum-based doublet chemotherapy regimen
Study connected ONLY in network of all-comers NSQ NSCLC					
V-15-32[27]	Randomised, multicentre, open label, active-controlled, phase III study	Japan (50 sites)	Docetaxel (244) Gefitinib (245)	21 months	Age ≥20 years Histologically or cytologically confirmed NSCLC (stages IIIB to IV) Failure of prior treatment with one or two chemotherapy regimens (≥ 1 platinum-based regimen) Life expectancy of 3 months or greater WHO PS 0-2

ECOG=Eastern Cooperative Oncology Group; EGFR=epidermal growth factor receptor; NSCLC=non-small cell lung cancer; NR=not reported; NSQ=non-squamous; PS=performance status; WHO=World Health Organization
Source: CS, Table 23

4.3.2 Methodological approach to the indirect comparisons

The company performed the ITCs using the Bucher method,[40] as described in Appendix 7.2 of the CS. The Bucher method can be used to obtain indirect estimates of treatment effect when there are no closed loops in the network of evidence. As is evident from Figure 5 and Figure 6, there were no closed loops in the either network of evidence and hence the ERG is satisfied that the modelling approach chosen by the company was appropriate.

ITCs often incorporate reported HRs from the included studies, and the outputs from these ITCs are estimated HRs for pairs of treatments which take into consideration both direct and indirect evidence. As there are no closed loops in the networks of evidence provided by the company, the HRs generated by the ITCs are based on indirect evidence only.

However, the company highlights that the validity of using HRs within ITCs relies on the assumption that the hazards of drugs within each study and across different studies included in the network of evidence are proportional. To test the assumption of PH, the company regenerated individual patient data from published K-M curves using methodology proposed by Guyot and colleagues.[41] The company inspected log-log plots (log cumulative hazard versus log time) and, if the curves for each treatment arm were approximately parallel, it was assumed that the assumption of PH was valid. The company also performed statistical tests, namely the Global Test for PH assumption,[35] and a supremum test for PH assumption.

The company states that if the PH assumption was violated for data taken from the included trials they would use differences in restricted mean survival time (RMST) instead of HRs in the ITCs as RMST is a measure of treatment effect which does not rely on the assumption of PH. This is based on work by Royston[42] in the analysis of OS in RCTs. This is the first time that the ERG have seen this measure of treatment effect used in an ITC. The ERG can think of no reason why it would not be appropriate to use this measure of treatment effect in an ITC and their approach using WinBUGs to fit the model sounds reasonable, although we do have very limited information in the CS about the actual analysis.

RMST is defined as the area under the survival curve up to the time t*, where t* is the follow-up period of clinical interest. From a clinical perspective, this measure can be interpreted as the 'life expectancy' between randomisation (t=0) and a particular time horizon (t=t*). The company chose t* to be the minimum follow-up time of all the trials included in the analysis. The company explains that the benefit of implementing this methodology is that the approach does not make any assumptions about the distribution of the data. The company used R to calculate RMST for each treatment arm, and the difference between the RMST in

the two arms, along with CIs, for each study included in the network of evidence. After calculating the RMST for each study, the company used WinBUGS to perform the ITCs, consequently estimating the RMST differences for treatments for which there is no direct evidence.

The results of the company assessments of PH for the included studies were not provided in the CS, and, instead of using the results of these assessments to determine whether to perform ITCs using HRs or RMST differences, the company performed each ITC using both HRs and RMST differences.

For the results of an ITC that uses HRs as data inputs to be credible, the assumption of PH must hold both within and across trials. As previously discussed in Section 4.2.4 of this ERG report, the results of the testing carried out by the company (see Appendix 10.2) indicate that the assumption of PH is violated for both OS and PFS data for CheckMate 057. As data from CheckMate 057 are used in every ITC, the mathematics used to calculate estimated HRs for indirect estimates of effect are fundamentally compromised. Consequently, the ERG is of the opinion that none of the ITC results that were generated using HRs (as opposed to RMST differences) are valid.

The company also provided the results of testing the PH assumption for the other trials included in the ITCs. However, as PH was violated for CheckMate 057, this rendered every ITC result for OS and PFS (which was generated using HRs) meaningless, and so there is no reason to consider PH for the other trials included in the ITCs. The ERG does not therefore report the HRs for any of these analyses and urges that any HRs reported in the CS are interpreted with extreme caution.

4.3.3 Characteristics of studies included in the indirect comparisons

The patient populations of the five trials[21,24-27] included in the networks of evidence differed due to differences in eligibility criteria; CheckMate 057 recruited previously treated patients with only non-squamous advanced and/or metastatic NSCLC, whereas the other four studies[24-27] included patients with both squamous and non-squamous NSCLC. Furthermore, CheckMate 057[21], ISEL,[25] ISTANA,[26] and V-15-32[27] recruited patients who had failed a platinum-based chemotherapy and had PS 0-1, PS 0-2, PS 0-3 and PS 0-2, respectively; however, LUME-Lung 1[24] included patients who had failed one line of chemotherapy and had a PS 0-1. In addition V-15-32[27] recruited patients from Japan who may have significantly different responses.

Baseline characteristics of the patients recruited to RCTs that were included in the ITCs are reported in Table 18 and Table 19.

Table 18 Baseline characteristics of studies for previously treated non-squamous NSCLC population

Study	Treatment arm	N	Smokers, n (%)				PS (ECOG*/WHO†), n (%)					
			Current	Former	Never	Current or former	PS 0	PS 0-1	PS 1	PS 2	PS 3	PS 2-3
Studies connected in BOTH networks (EGFR mutation-negative/unknown AND all-comers NSQ NSCLC)												
LUME-Lung 1[24] Adenocarcinoma population	Docetaxel	336	59 (17.6)	162 (48.2)	115 (34.2)	221 (65.8)	99 (29.5)	-	237 (70.5)			
	Docetaxel+ nintedanib	322	56 (17.4)	151 (46.9)	115 (35.7)	207 (64.3)	96 (29.8)	-	226 (70.2)			
ISTANA[26]	Docetaxel	79	0 (0.0)	43 (54.4)	36 (45.6)	43 (54.4)	3† (3.8)	-	71† (89.9)	5† (6.3)		
	Gefitinib	82	1 (1.2)	51 (62.2)	30 (36.6)	52 (63.4)	2† (2.4)	-	74† (90.2)	6† (7.3)		
ISEL[25]	BSC	563	97 (17)	340 (60)	125 (22)	437 (77)	70† (12)	-	318† (56)	145† (26)	29† (5)	
	Gefitinib+BSC	1129	201 (18)	678 (60)	250 (22)	879 (78)	140† (12)	-	598† (53)	332† (29)	55† (5)	
CheckMate 057[28]	Nivolumab	292	-	-	58 (19.9)	231 (79.1)	84 (28.8)	-	208 (71.2)		0 (0)	
	Docetaxel	290	-	-	60 (20.7)	227 (78.3)	95 (32.8)	-	193 (66.6)		1 (0.3)	
Study connected ONLY in network of all-comers NSQ NSCLC												
V-15-32[27]	Docetaxel	244	-	-	87 (35.7)	157 (64.3)		93* (38.1)	141* (57.8)	10* (4.1)		
	Gefitinib	245	-	-	71 (29)	174 (71)		85* (34.7)	149* (60.8)	11* (4.5)		

BSC=best supportive care; ECOG=Eastern Cooperative Oncology Group; EGFR=epidermal growth factor receptor; NSCLC=non-small cell lung cancer; NSQ=non-squamous; PS=performance status; WHO=World Health Organization

* PS rated on the ECOG scale

† PS rated on the WHO scale

Source: CS, adapted from Table 24; data for the adenocarcinoma population from LUME-Lung 1[24] were taken from data published as part of the previous NICE appraisal (TA347)[43]

Table 19 Summary of baseline characteristics of studies for previously treated non-squamous NSCLC population and included in analysis

Study	Treatment arm	N	Disease stage (%)			EGFR mutation status	Histology	Median age (range), years	Male (%)
			Stage III	Stage IV	Stage III/ IV				
Studies connected in BOTH networks (EGFR mutation-negative/unknown AND all-comers NSQ NSCLC)									
LUME-Lung 1[24] Adenocarcinoma population	Docetaxel	336	NR	NR	NR	EGFR mutation-negative/unknown: 100%	NSQ: 100% SQ: 0%	Mean (SD): 58.6 (9.5)	61.9
	Docetaxel + nintedanib	322	NR	NR	NR		NSQ: 100% SQ: 0%	Mean (SD): 58.5 (10.1)	63
ISTANA[26]	Docetaxel	79	-	-	100	Unclear	NSQ: 86.3% SQ: 13.7%	58 (20-73)	57
	Gefitinib	82	-	-	100		NSQ: 79.3% SQ: 20.7%	57 (21-74)	67.1
ISEL[25]	BSC	563	39	50	-	EGFR mutation-negative/unknown: 87.9% EGFR mutation-positive: 12.1%	NSQ: 67% SQ: 33%	61 (31–87)	67
	Gefitinib + BSC	1129	44	47	-		NSQ: 65% SQ: 35%	62 (28–90)	67
CheckMate 057[28]	Nivolumab	292	6.8	93.2	-	EGFR mutation testing: N=212 EGFR mutation-positive: 20.7% EGFR mutation-negative/unknown: 79.3%	NSQ: 100%	61 (37-84)	51.7
	Docetaxel	290	8.3	91.7	-	EGFR mutation testing: N=210 EGFR mutation-positive: 18.1% EGFR mutation-negative/unknown: 79.9%	NSQ: 100%	64 (21-85)	57.9
Study connected ONLY in network of all-comers NSQ NSCLC									
V-15-32[27]	Docetaxel	244	20.5	61.5	-	EGFR mutation testing: N=54 EGFR mutation-positive: 54.4%	NSQ: 83.2% SQ: 16.8%	NR	61.9
	Gefitinib	245	19.2	64.9	-		NSQ: 84.9% SQ: 15.1%	NR	61.6

BSC=best supportive care; EGFR=epidermal growth factor receptor; NR=not reported; NSQ=non-squamous; SQ=squamous
Source: CS, adapted from Table 25; data for the adenocarcinoma population from LUME-Lung 1[24] were taken from data published as part of the previous NICE appraisal (TA347)[43]

Baseline characteristics for trials included in all-comers network

The ERG considered baseline characteristics for the included trials in order to determine whether performing an ITC for the all-comers population was appropriate.

The ERG notes that the company presented baseline characteristics for the whole trial population of LUME-Lung 1[24] in Table 24 of the CS, despite the fact that only data for the adenocarcinoma population were used in the ITC. In order to allow comparisons between the LUME-Lung 1[24] adenocarcinoma population and the other trials included in the ITC, the ERG presents baseline characteristics for the adenocarcinoma population for LUME-Lung 1[24] (Table 18). These data were published as part of a previous NICE appraisal (TA347).[43]

Due to differences in eligibility criteria, the proportions of patients with different types of ECOG PS varied considerably between the included trials. The trials also differed with regards to the smoking status of the patient populations.

Although only a small number of patients in V-15-32[27] had EGFR mutation status tested (54 patients out of 489), the results of this testing indicated that a large proportion of patients were EGFR mutation-positive (54.4%). The other trials all had large (≥75%) percentages of EGFR mutation-negative/unknown patients.

Perhaps the most important source of variability within the populations of the trials included in the all-comers network is the inclusion of both squamous and non-squamous patients in ISEL,[25] ISTANA,[26] and V-15-32.[27] The company states that subgroup data were provided for the non-squamous patients for the trials that included both squamous and non-squamous patients, but the data inputs provided by the company that were used in the ITCs were for the whole trial populations. The ERG notes that the trials which recruited both squamous and non-squamous patients recruited a majority of non-squamous patients, although the ERG is disappointed that the company did not list the variability in the proportion of non-squamous patients as a limitation of the ITCs, and the company did not consider whether it was appropriate to compare these trial populations to the trial population of CheckMate 057 which consisted of solely non-squamous patients.

The ERG notes that heterogeneity within the trials would be more likely to affect the ITC of nivolumab versus BSC than the ITC of nivolumab versus nintedanib+docetaxel. The main heterogeneity in the network is observed between the trials used to link BSC to nivolumab (i.e. CheckMate 057,[7] ISEL,[25] ISTANA,[26] and V-15-32[27]). The ERG is of the opinion that this heterogeneity ought to be considered when interpreting results presented for nivolumab versus BSC, although it is very difficult to assess how the overall treatment effect estimate

would be affected by these differences. The ERG advises that the results of the ITC for nivolumab versus BSC ought to be interpreted with caution.

Only two trials contribute evidence to the ITC of nivolumab versus nintedanib+docetaxel (CheckMate 057[7] and LUME-Lung 1[24]), and the ERG is of the opinion that baseline characteristics are fairly similar across these trials. CheckMate 057 included fewer male patients (51.7% in the nivolumab arm, 57.9% in the docetaxel arm) than LUME-Lung 1[24] (61.9% in the docetaxel arm, 63% in the docetaxel plus nintedanib arm). As male patients are expected to do slightly worse than females, this heterogeneity could favour nivolumab, although the ERG notes that the LUME-Lung 1[24] population consisted of solely EGFR mutation-negative/unknown patients, who are likely to have better treatment outcomes than a population consisting of both EGFR mutation-positive and -negative/unknown patients (CheckMate 057). Overall, the ERG does not consider differences between these two trials to be concerning. However, it is important to note that the ERG did not have access to data summarising the disease stage of patients in the adenocarcinoma population of the LUME-Lung 1 trial,[24] and so it is not possible to compare the patients of the adenocarcinoma subpopulation of LUME-Lung 1[24] and the patient population of CheckMate 057 in this respect.

Baseline characteristics for trials in EGFR mutation-negative/unknown network

No baseline characteristics were provided for the subpopulation of EGFR mutation-negative/unknown patients for any of the included trials. Following the company's response to the clarification letter, it became clear to the ERG that, with the exception of CheckMate 057, the company used data for the whole trial populations in the ITCs for EGFR mutation-negative/unknown patients. The company justified this approach by stating that in all of these trials, either the whole population was EGFR negative or the proportion of EGFR negative patients was ≥80% of the whole population.

The company used data for the EGFR mutation-negative/unknown subgroup of patients from CheckMate 057. However, there are no baseline characteristics presented for the subgroup of EGFR mutation-negative/unknown patients from CheckMate 057, and consequently it is not possible to assess whether the patient populations of the trials included in the EGFR mutation-negative/unknown network are comparable.

4.3.4 Assessment of risk of bias of the trials included in the indirect comparisons

The company conducted an assessment of the risk of bias of the studies included in the ITCs and the results presented in the CS and are shown in Appendix 10.4 of the ERG report and discussed below.

The CS indicates that it is unclear whether randomisation was carried out appropriately in ISTANA,[26] ISEL[25] and V-15-32[27] so it is not possible to assess whether there was a risk of bias. All studies were at low risk of bias for differences between the groups on prognostic factors at the outset of the studies.

Whilst LUME-Lung 1[24] and ISEL[25] were considered to be at low risk of bias, CheckMate 057,[28] ISTANA[26] and V-15-32[27] trials were considered to be at a high risk of bias for blinding due to being open-label trials. However, the ERG notes that due to the difference in the treatment schedules and AE profiles it would be challenging to compare any combination of these treatments with each other in a blinded manner.

Only ISTANA[26] did not report unexpected imbalances in drop-outs between groups and was therefore assessed to at high risk of bias for this item. There was no evidence to conclude whether all outcomes assessed were reported in the ISEL[25] trial so it was not possible to assess the risk of bias for this item. All trials were assessed as having low risk of bias for analysis using an ITT.

4.3.5 Results from the indirect comparisons

Individual study results

The results of the individual studies included in the ITCs were provided by the company (CS, Table 27). The ERG has updated Table 27 of the CS to correct minor errors which were identified as part of the clarification process; this table is provided in Appendix 10.5 of this ERG report.

Indirect treatment comparison results

The results of the ITCs carried out by the company are provided in Table 20.

As PH was violated for CheckMate 057, this rendered every ITC result for OS and PFS (which was generated using HRs) meaningless, and so there is no reason to consider PH for the other trials included in the ITCs. The ERG does not therefore report the HRs for any of these analyses and urges that any HRs reported in the CS are interpreted with extreme caution.

The results shown in Table 20 are slightly different to those shown in Table 28 of the CS, as the ERG has incorporated the updated ITC results, which use the updated data inputs. The ERG is satisfied that the company's approach was appropriate, and the results of the ITC are very similar when using either set of data inputs.

For OS in the all-comers patient population, no statistically significant differences between nivolumab and nintedanib+docetaxel were identified. Similarly, no statistically significant differences were found in terms of PFS or ORR. For AEs, no treatment benefit was observed for nivolumab in comparison to nintedanib+docetaxel when considering the outcome "any adverse event"; however, nivolumab was demonstrated to statistically significantly reduce the risk of Grade 3-4 AEs ([redacted]) (Table 20).

Considering the comparison of nivolumab and BSC in the all-comers patient population, nivolumab was shown to statistically significantly improve OS (RMST difference (95%CI); [redacted]) (Table 20).

For OS, PFS and ORR in the EGFR mutation-negative/unknown patient population, there were no statistically significant differences between nivolumab and nintedanib+docetaxel.

The comparison of nivolumab and BSC in the EGFR mutation-negative/unknown patient population failed to demonstrate any statistically significant differences in terms of OS (Table 20).

Table 20 Results of the indirect treatment comparison

Outcome	Nivolumab vs nintedanib+docetaxel	Nivolumab vs. BSC
Patient population: All-comers NSQ NSCLC		
OS RMST difference (95% CI); p value	[redacted]	[redacted]
PFS RMST difference (95% CI); p value	[redacted]	[redacted]
ORR RR (95% CI); p value	[redacted]	[redacted]
Any adverse event RR (95% CI); p value	[redacted]	[redacted]
Any Grade 3/4 adverse event RR (95% CI); p value	[redacted]	[redacted]
Patient population: EGFR mutation-negative/unknown NSQ NSCLC		
OS RMST difference (95% CI); p value	[redacted]	[redacted]
PFS RMST difference (95% CI); p value	[redacted]	
ORR RR (95% CI); p value	[redacted]	[redacted]

BSC=best supportive care; CI=confidence interval; EGFR=epidermal growth factor receptor; NSCLC=non-small cell lung cancer; NSQ=non-squamous; ORR=objective response rate; OS=overall survival; PFS=progression-free survival; RMST=restricted mean survival time; RR=relative risk
Source: CS, adapted from Table 28

Additional analysis requested by the ERG

The ERG noted that the ITC for nivolumab versus nintedanib+docetaxel used data from the LUME-Lung 1[24] second-line patient population, and CheckMate 057 second and third-line patient population. The ERG therefore requested that the company perform an indirect comparison of nivolumab versus nintedanib+docetaxel using data only from the second-line population from CheckMate 057. The results of this analysis are provided in Appendix 10.6. The ERG is satisfied that the results of the ITC for nivolumab+docetaxel are robust; results generated by the additional analysis requested by the ERG are in accordance with those from the original analysis.

ERG critique of the company's results from the ITCs

The ERG's main concern when considering the results of the ITCs is that none of the results of the ITCs which were performed using HRs (as opposed to RMST differences) are valid.

The ERG has not previously seen RMST differences used as the measure of effect in an ITC. The ERG can think of no reason why it would not be appropriate to use this measure of treatment effect in an ITC, although the ERG would be more confident in the validity of the approach of the company had provided detailed justification for the use of RMST in ITCs, including references to published articles describing the methodology used, and an explanation of strengths and limitations of the approach. The ERG notes that these analyses are sometimes based on reasonably short follow-up periods, as shown in Table 21.

For the comparison of nivolumab versus BSC (both all-comers and EGFR mutation-negative/unknown networks), OS data are truncated for inclusion in the ITC at 13 months, as this is the minimum follow-up time of ISEL.[25] Each of the other studies had a minimum follow-up time of at least 25 months for this outcome. Similarly, for the comparison of nivolumab versus nintedanib+docetaxel (both all-comers and EGFR mutation-negative/unknown networks), PFS data are truncated for inclusion in the ITC at 12 months, as this is the minimum follow-up time of LUME-Lung 1.[24] Data from CheckMate 057 also contributed to this ITC and this trial had a minimum follow-up time of 25 months. The ERG is of the opinion that the company applied an appropriate method to overcome the problem of non-PH, but the RMST method is limited in that long-term data cannot be incorporated into the ITCs even if available.

Table 21 Minimum follow-up for each ITC

	All-comers network		EGFR mutation-negative/unknown network	
	OS	**PFS**	**OS**	**PFS**
Nivolumab versus nintedanib+docetaxel	25 months	12 months	28 months	12 months
Nivolumab versus BSC	13 months	N/A	13 months	N/A

BSC=best supportive care; EGFR= epidermal growth factor receptor; N/A=not applicable; OS=overall survival; PFS=progression-free survival

The ERG's interpretation of the ITC results which were calculated using RMST differences (or risk ratios in the case of ORR and AE outcomes) is as follows:

- In the all-comers patient population, no statistically significant differences between nivolumab and nintedanib+docetaxel were found for OS, PFS, ORR or "any AE". Nivolumab was found to statistically significantly reduce the risk of Grade 3/4 AEs in comparison to nintedanib+docetaxel
- In the all-comers patient population, nivolumab was shown to statistically significantly improve OS in comparison to BSC.
- In the EGFR mutation-negative/unknown patient population, no statistically significant differences between nivolumab and nintedanib+docetaxel were found for OS, PFS, or ORR.
- In the EGFR mutation-negative/unknown patient population, no statistically significant differences were observed in terms of OS between nivolumab and BSC.

The ERG noted that variability within the trials would be more likely to affect the indirect comparison of nivolumab versus BSC than the indirect comparison of nivolumab versus nintedanib+docetaxel (see Section 4.3.3). However, it is very difficult to assess how the overall treatment effect estimates would be affected by heterogeneity. The ERG notes that a statistically significant treatment benefit was observed for nivolumab in comparison to BSC in terms of OS for the all-comers network, but not for the EGFR mutation-negative/unknown network. This result is somewhat surprising in that EGFR mutation-negative/unknown patients might be expected to respond better to nivolumab than patients in the all-comers population. This unexpected result could be a consequence of heterogeneity within one or both networks of evidence, which decreases the validity of indirect estimates of treatment effect.

Finally, the ERG notes that it is not possible to assess whether the patient populations of the trials included in the EGFR mutation-negative/unknown network are comparable, since no baseline characteristics for the EGFR mutation-negative/unknown subgroup of CheckMate

057 were presented. Consequently, it is possible that there may be important differences between this subpopulation and the trial populations of other trials included in the EGFR mutation-negative/unknown network. This issue introduces uncertainty as to the validity of estimates of treatment effect generated for this network.

4.4 Summary and critique of supportive evidence from non-randomised studies

In addition to the Phase III RCT (CheckMate 057), evidence from two non-RCTs was also submitted by the company: a Phase IIIb/IV, open-label study (CheckMate 153[22]) and a single-arm Phase I dose-escalation study (CheckMate 003[23]). The characteristics and findings relating to these trials are reported in the CS in Section 4.11.

4.5 Conclusions of the clinical effectiveness section

The primary data provided in the CS are derived from CheckMate 057 and an ITC that is limited by the use of HRs and a lack of data to allow for comparison with all of relative comparators listed in the scope. Comparison is therefore limited to nivolumab with docetaxel, nintedanib+docetaxel and BSC.

CheckMate 057 is a well conducted trial however the use of HRs in the analysis of the data cannot be considered a reliable estimate of treatment effectiveness as the CS points out that the proportional hazards assumption is violated for both OS and PFS. This limitation is also true of the ITC where only RMST analysis should be considered. The ITC is also limited by the differences in the patient populations of patients included in the analysis (e.g. inclusion of patients with squamous disease, Asian population, length of follow-up etc.) Since it is expected that nintedanib+docetaxel will replace the use of docetaxel alone in the treatment of these patents then consideration of the comparison of nivolumab to this combination treatment is important. The results of the ITC show no difference in terms of OS, PFS, ORR or AE in the 'all comer' or EGFR mutation-negative/unknown populations. The comparison with BSC provides mixed results demonstrating the effectiveness of nivolumab versus BSC in the all-comers group but not in the EGFR mutation-negative/unknown patients, raising concerns that there were differences in the patient populations in the trials used in the ITC.

The CS indicates that the AE experienced by patients receiving nivolumab will be fewer than those experienced by patients receiving nintedanib+docetaxel. The ERG is of the opinion that although the comparative data are limited that patients receiving docetaxel do have a higher rate of Grade 3-4 AEs and it would be expected this would be at least the same when docetaxel was given in combination with nintedanib.

The company makes a claim that OS in the patients receiving docetaxel in CheckMate 057 is longer then would be expected. Examination of data from other similar trials does not substantiate this claim. The CS also makes a claim that the pseudo progression seen in patients receiving nivolumab would have an effect on OS. The ERG is not convinced that the data presented support this claim.

Subgroup analyses suggest that nivolumab is statistically significantly more effective in patients with higher PD-L1 expression levels than those with lower PD-L1 expression levels. The report is however somewhat inconsistent with regards to where all patients should therefore be tested for PD-L1.

5 COST EFFECTIVENESS

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

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

5.1.1 Objective of cost effectiveness review

The company performed a search to identify economic evaluations, resource use and costs and utility values. The full search strategies are documented in Appendix 11 of the CS and are outlined in Section 5.1.1. The search was performed in February 2015 using the following databases: MEDLINE; MEDLINE in Process; EMBASE; EconLIT, NHSEED, CENTRAL and HTAD. The reported population terms and drug names in the database strategies were considered to be comprehensive by the ERG. Both an economics search filter and a health related quality of life (HRQoL) search filter were added to the search, and consequently the results were not limited to cost effectiveness alone. The searches carried out in EconLit included only NSCLC search terms; the ERG deems this approach to be appropriate due to the small numbers of studies retrieved from these databases. It is not documented whether any further hand searches were carried out as part of the cost effectiveness searches. The same search strategy was used for the measurement and valuation of health effects searches.

5.1.2 Eligibility criteria used in study selection

The inclusion/exclusion criteria that were used to facilitate study selection are presented in Table 22. The ERG is satisfied that these criteria are relevant to the decision problem.

Table 22 Inclusion/exclusion criteria for the company's cost effectiveness review

	Inclusion criteria	Exclusion criteria
Population	Adults diagnosed with locally advanced or metastatic NSCLC previously treated with at least one previous line of chemotherapy	Patients aged <18 years Patients with stage I-IIIa disease Chemotherapy treatment-naïve patients
Intervention	Nivolumab	Studies investigating first-line treatment for NSCLC Studies assessing nivolumab as an adjuvant or neoadjuvant therapy Studies evaluating nivolumab in combination with radiotherapy Studies comparing different doses of the same intervention (i.e. dose-ranging studies), two formulations of the same intervention and intervention with two different routes of administration
Comparator	Any pharmacological intervention Placebo BSC Afatinib Docetaxel Erlotinib Gefitinib Nintedanib (in combination with docetaxel) Pemetrexed monotherapy Ceritinib Crizotinib Platinum-based chemotherapy in combination with gemcitabine, vinorelbine, pemetrexed, or a taxane	Non-pharmacological interventions, other than BSC
Outcome	All reported outcomes	-
Study design 1 (S1)*	All economic evaluation studies based on: Cost-effectiveness analysis Cost-utility analysis Cost-minimisation analysis Budget-impact models	Studies reporting only cost and resource use data where no formal economic analysis has been undertaken
Study design 2 (S2)*	Randomised controlled trials Database studies Prospective observational studies Retrospective observational studies	Animal/in vitro studies Single-arm studies Studies with no subgroup data for disease and adult population Reviews, letters to the editors and editorials Conference abstracts prior to 2012
Line of therapy	Second- or further-line of therapy	First-line of therapy
Search timeframe	2000 to 2015 (last 15 years)	Prior to 2000
Language	Only studies with the full-text published in English included	Studies with the full-text published in languages other than English

BSC=best supportive care; NSCLC=non-small cell lung cancer
*Within the single systematic review, two sets of study design criteria (S1 and S2) were used to identify relevant economic evaluations and relevant studies reporting data on quality of life in second-line or later-line patients with NSCLC
Source: CS, Table 42

5.1.3 Included and excluded studies

None of the studies identified by the company's search evaluated the cost effectiveness of treatments in a non-squamous only population and, furthermore, no studies considered treatment with nivolumab. The company identified four relevant appraisals (Crizotinib [TA296[44]], Erlotinib [TA162[45]], Erlotinib and gefitinib [Review of TA162 and TA175)[46]] and Nintedanib [TA379[43]]) and these were used to inform the development of the company economic model (Table 44 of the CS). Two relevant UK-based cost effectiveness studies[47,48] were also identified by the search. Both studies included patients with NSCLC who had been previously treated (CS, Table 43); one study[47] compared docetaxel with BSC and the other study[48] compared erlotinib with docetaxel. Holmes[47] reported an incremental cost per life year gained (LYG) for docetaxel versus BSC of £13,863. Lewis[48] found erlotinib to be dominant when compared with docetaxel. The models described in these two studies[47,48] and the four relevant models[43-46] submitted previously as part of technology appraisals, all used a three-state partitioned survival model representing progression-free (PF) disease, progressive disease (PD) and death.

5.1.4 Findings from the cost effectiveness review

The company did not identify any studies that evaluated the cost effectiveness of nivolumab compared to any comparator in a non-squamous patient population. Summary details relating to the two UK-based published cost effectiveness studies[47,48] and four published NICE technology appraisals[43-46] considered to be relevant to the company's review question are reported in the CS (Tables 43-44).

5.2 ERG critique of the company's literature review

The ERG is satisfied with the company's search strategy and stated inclusion/exclusion criteria and is confident that the company did not miss any relevant published papers at the time of submission. The ERG considers the wider search for published economic literature (e.g. inclusion of squamous patient population) to be appropriate when taking into account the shortage of relevant clinical and economic data for patient populations with advanced or metastatic non-squamous NSCLC.

5.3 ERG summary and critique of the economic evaluation submitted by the company

5.3.1 Model structure

The company developed a cohort-based partitioned survival model comprising three mutually exclusive health states: progression-free (PF), progressed disease (PD) and death. The de novo economic model was developed in Microsoft Excel and the structure is consistent with previous economic evaluations submitted to NICE as part of appraisals of treatments for patients with advanced NSCLC (including nivolumab [after chemotherapy] for patients with squamous NSCLC[49] and other metastatic cancers (e.g. Nintedanib TA347,[43] Erlotinib TA258[50] and Bevacizumab TA212[51]). A schematic of the company's model is shown in Figure 7.

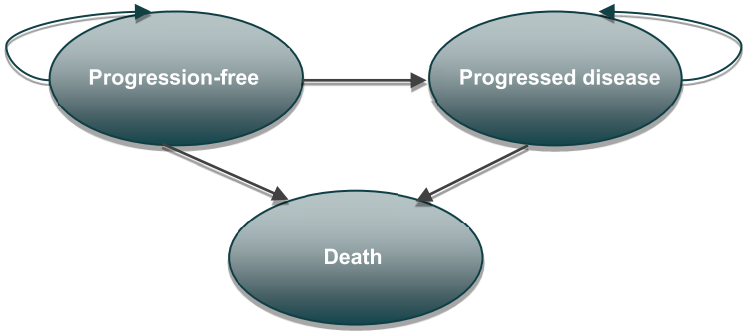


Figure 7 Schematic of company's model

Source: CS, Figure 26

The base case evaluates the cost effectiveness of nivolumab compared with (i) docetaxel and (ii) nintedanib+docetaxel. These two comparators represent the current standard of care in the second-line setting in the UK NHS. Patients who have failed platinum-based chemotherapy enter the model in the PF health state. Patients who remain in PF are treated with nivolumab, docetaxel or nintedanib+docetaxel. At the end of each cycle a patient can remain in the same health state or transition to PD or death.

A restriction in the model is that patients cannot transition to an improved health state. Costs and health-related utilities are allocated to each health state and multiplied by state occupancy to calculate the weighted costs and quality adjusted life years (QALYs) per cycle. In the model, cycle length is 1 week to accommodate the different dosing regimens of nivolumab (every 2 weeks) and docetaxel (every 3 weeks). All patients are treated until treatment discontinuation, which may be beyond progression, and this is consistent with the CheckMate 057 protocol.[52]

In the company model the PF health state occupancy is modelled using time to treatment discontinuation (TTD), rather than PFS. This means that costs and utilities are based on actual treatment duration. The ERG does not recommend the use of TTD data for the estimation of treatment benefit/utilities.

5.3.2 Population

The economic evaluation considers previously treated adult patients with advanced or metastatic non-squamous NSCLC, which is consistent with the decision problem, expected marketing authorisation and population included in CheckMate 057. This patient population is a subgroup of the population described in the final NICE scope.

5.3.3 Interventions and comparators

In the model, nivolumab treatment is implemented in line with the anticipated licensed dose, i.e. 3mg/kg over 60 minutes as an intravenous infusion every 2 weeks.

The base case comparator in the economic analysis is docetaxel, administered at a dose of 75mg/m^2 every 3 weeks via intravenous infusion. The company also compares nivolumab with nintedanib+docetaxel, nintedanib is taken as two tablets per day on a 21-day cycle and docetaxel is administered at a dose of 75mg/m^2 every 3 weeks via intravenous infusion. Due to docetaxel being the current standard of care in previously treated patients with non-squamous NSCLC in the UK,[21] the company assumes that it is the treatment most likely to be displaced by the introduction of nivolumab.

Subsequent treatments

In the model it is assumed that nivolumab and docetaxel are second-line treatments and that patients can only receive one further line of therapy following progression (third-line therapy). However, the company did not provide details about the duration of subsequent treatments used in CheckMate 057. The duration of third-line therapy was derived from real world data (RWD) as reported in an observational study (CA209-116[53]) in which treatment patterns, outcomes and healthcare resource use in patients with advanced NSCLC in Europe were investigated. In the model, the time until treatment discontinuation in patients in a third-line setting is reported to be [redacted] days.

5.3.4 Perspective, time horizon and discounting

The company states that the economic evaluation is undertaken from the perspective of the NHS and Personal Social Services. The time horizon is set at 20 years, in line with a

previous NICE STA[49] in this disease area (Table 45 of the CS) and taking into account the typical age of patients at diagnosis. Both costs and outcomes are discounted at a rate of 3.5% per annum; a half-cycle correction is implemented in the model.

5.3.5 Treatment effectiveness and extrapolation: nivolumab versus docetaxel

The company economic model relies on patient level data from CheckMate 057. The follow-up period in this trial was shorter than the required length of the economic analysis (a lifetime equivalent) and extrapolation of the OS and TTD data from the trial was necessary to enable the partitioned survival method to be used. Extrapolation involved the identification of suitable parametric survival models for OS and TTD data.

Overall survival

Log-cumulative hazards, log-cumulative odds and standardised normal curve plots were generated to determine whether patient level data from CheckMate 057 indicated proportional hazards. The analyses that were carried out by the company confirmed that the assumption of proportional hazards did not hold for OS. Therefore, both independent survival models and single survival models adjusted for shape and scale were then assessed. The Akaike Information Criterion and Bayesian Information Criterion goodness-of-fit values for the selected parametric distributions and long-term clinical plausability of the extrapolated model were used to establish the best fitting survival model. As stated by the company (CS, p149), the long-term clinical plausibility of the extrapolated model was based on validation against available nivolumab clinical study data with longer follow-up than CheckMate 057: CheckMate 003,[54] the National Lung Cancer Audit (NLCA) registry[10] (UK) and input from UK clinicians. NLCA data[10] were available for up to 5 years.

According to the company, in terms of statistical fit, the three best-fitting parametric survival models for nivolumab are the 2-knot spline, log-normal and generalised gamma distributions. Correspondingly, the three best-fitting parametric survival models are the gamma, generalised gamma and 1-knot spline distributions for docetaxel. Based on all of the evidence (statistical and visual fit, validation against CheckMate 003[54] and NLCA data[10] clinical input and NICE DSU guidance[55]), the generalised gamma model was used in the company base case for OS extrapolation of nivolumab and docetaxel.

Time to treatment discontinuation

The choice of a parametric survival model for TTD was informed by assessment of whether the assumption of proportional hazards was violated. This was performed by visual inspection of the log-cumulative hazards, log-cumulative odds, and standardised normal curve plots. The results of the Grambsch and Therneau's correlation test between Schoenfeld residuals and log of time indicated that the null hypothesis for proportional hazards should be rejected ($p<0.05$).

In terms of statistical fit, the two best-fitting parametric survival models for nivolumab are the 1-knot spline odds and generalised gamma. Correspondingly, the two best-fitting parametric survival models are the generalised gamma and gamma distribution for docetaxel. The company determined that the generalised gamma model should be used as the base case for TTD extrapolation for nivolumab and docetaxel (Table 23) on the grounds that the generalised gamma distribution provided a good fit for both treatment arms in terms of goodness-of-fit statistics and internal validation against long-term nivolumab clinical study data (CheckMate 003[56]). The company also noted that using the generalised gamma model maintained consistency between the functional forms used for OS and TTD extrapolation.

Table 23 Summary of survival distributions for TTD and OS used in the base case

Survival models explored	Best-fitting parametric curve
Time to treatment discontinuation	
Base case	Docetaxel: Generalised gamma Nivolumab: Generalised gamma
Scenario analysis	Docetaxel: Gamma Nivolumab: 1-knot spline odds
Overall survival	
Base case	Docetaxel: Generalised gamma Nivolumab: Generalised gamma
Scenario analysis	Docetaxel: Gamma Nivolumab: 2-knot spline hazards

Source: CS, Table 52

Nivolumab versus nintedanib+docetaxel comparison

K-M graphs from the LUME-Lung 1[24] trial (nintedanib+docetaxel versus docetaxel+placebo) were digitised by the company to estimate proxy patient-level data. Specifically, data for the adenocarcinoma population were used in the analysis. According to the company, it is indicated that there is no difference in OS between nintedanib+docetaxel versus docetaxel+placebo up to 6 months, thus the HR is assumed to be 1 to this time point. An estimated HR of 0.75 was assumed for ≥6 months. For PFS, a HR of 1 is assumed up to 2

months and a HR of 0.98 is assumed for ≥ 2 months. Table 24 summarises the output of this analysis.

Table 24 Summary of OS and PFS hazard ratios for nintedanib+docetaxel

Efficacy	HR	95% CI
OS after 6 months	0.75	(0.60 to 0.93)
PFS after 2 months	0.98	(0.73 to 1.33)

CI=confidence interval; HR=hazard ratio; OS=overall survival; PFS=progression-free survival
Source: CS, Table 74

In the economic model, the HRs applied to the docetaxel data to estimate the nintedanib+docetaxel patient flows, according to the relevant time points, are outlined in Table 24.

5.3.6 Health related quality of life

Systematic searches to identify HRQoL studies were performed as part of the company's systematic literature review. However, none of the identified studies evaluated nivolumab and none were performed in a UK-based population. The utility values incorporated into the model are those derived from CheckMate 057. Utility data were collected in the trial using the EuroQol-5D preference-based health state utility questionnaire (EQ-5D Utility Index). The schedule of EQ-5D assessments is outlined in Table 25.

Table 25 Schedule of EQ-5D assessments in CheckMate 057

EQ-5D assessment schedule	On-study assessment			Follow-up assessment (visit 1 and 2)	Survival assessment (beyond 100 days from the last study dose)
	Every 4 weeks for the first 6 months	Every 3 weeks for the first 6 months	Every 6 Weeks thereafter	Within the first 100 days from the last dose of study	Every 3 months for the first 12 months, and every 6 months thereafter
Nivolumab	✓		✓	✓	✓
Docetaxel		✓	✓	✓	✓

EQ-5D=EuroQol 5-Dimension
Source: CS, Table 55

In all, 82.2% of nivolumab patients and 76.6% of docetaxel patients completed the EQ-5D assessment at baseline. For baseline and at least one post-baseline visit, completion rates fell to 70.5% and 69.7% for nivolumab and docetaxel, respectively. No adjustments were made for missing data when analysing the EQ-5D data. Data from screening visits (up to 28 days before) were used in place of missing baseline data.

The mean utility values derived from patients with advanced NSCLC based on the analysis of CheckMate 057 (using a UK scoring algorithm[57]) are 0.688 (PD) and 0.739 (PF) with an overall utility of 0.728 across all disease states. The HRQoL of patients with advanced NSCLC is lower than the mean utility value of 0.86 that is derived from a representative sample of adults drawn from a national Health Survey of England in 2008.[58] The utility values used in the economic model are summarised in Table 26.

Table 26 UK-specific mean EQ-5D values by health state

Tumour response category (N=number of assessments)	UK (Mean)	Standard deviation	95% CI
Overall (N=1132)	0.728	NA	NA
PD (N=219)	0.688	0.298	0.665 to 0.712
PF (including SD/PR/CR) (N=913)	0.739	0.233	0.729 to 0.748

CI=confidence interval; CR=complete response; NA=not available; PD=progressive disease; PF=progression-free; PR=partial response; SD=stable disease; UK=United Kingdom
Sources: CS, Table 56

Adverse events

The economic model incorporates the quality of life impact of treatment related AEs of Grade 3 or higher severity which occurred in ≥2% of patients in CheckMate 057. The disutility per episode for the included AEs is shown in Table 27. The expected disutility per patient was calculated according to the incidence rates of the included AEs from CheckMate 057 according to treatment arm, this was applied separately in the first cycle only (i.e. without discounting) as a single disutility quantum. Disutility values could not be identified for all AEs; therefore, in the base case, where information was not available, a disutility of 0 was assumed. In addition to the AE disutility applied in the first cycle, the company applied the disutility of each AE separately.

Table 27 Disutilities of AEs

Adverse event	Disutility	Reference
Fatigue	-0.07346	Nafees 2008[59]
Asthenia	-0.07346	Nafees 2008[59]
Pain	0	Assumption
Dyspnoea	-0.050	Doyle 2008[60]
Pleural effusion	0	Assumption
Hyperglycemia	0	Assumption
Pneumonia	-0.008	Marti 2013[61]
Neutrophil count decreased	0	Assumption
White blood cell count decreased	-0.05	NICE 2015[43]
Anaemia	-0.07346	Nafees 2008[59]
Neutropenia	-0.08973	Nafees 2008[59]
Febrile neutropenia	-0.09002	Nafees 2008[59]
Leukopenia	-0.08973	Nafees 2008[59]
Diarrhoea	-0.0468	Nafees 2008[59]
Increased ALT	-0.05	NICE 2015[43]
Increased AST	0	Assumption
Hyponatraemia	0	Assumption

ALT=alanine Aminotransferase; AST=aspartate Transaminase
Source: CS, Table 57

The AE data in the economic model for nintedanib+docetaxel were taken directly from the LUME-lung 1[24] trial. Table 28 presents the proportion of patients experiencing Grade 3-4 AEs (≥2% incidence).

5.3.7 Resources and costs

Intervention costs

The drug acquisition costs by pack/vial size and the acquisition costs of each treatment cycle for the treatments are presented in Table 29 and Table 30 respectively.

Table 28 AEs included in the economic model based on LUME-Lung 1 (Grade 3 and 4 severity)

Type of AE	Rate for nintedanib+docetaxel
Fatigue	5.5%
Asthenia	2.0%
Pain	0.0%
Dyspnoea	4.9%
Pleural effusion	1.0%
Hyperglycemia	1.1%
Pneumonia	2.6%
Neutrophil count decreased	32.0%
White blood cell count decreased	16.4%
Anaemia	1.1%
Neutropenia	12.1%
Febrile neutropenia	7.0%
Leukopenia	2.9%
Diarrhoea	6.5%
Increased ALT	7.8%
Increased AST	3.4%
Hyponatraemia	2.1%

AE=adverse event; ALT=alanine aminotransferase; AST=aspartate transaminase
Source: CS, Table 75, Reck 2014[24]

Table 29 Drug acquisition costs (initial treatments)

Drug	Tablet dose/vial concentration	Pack size/vial volume	Cost per vial/pack	Source
Nivolumab	10mg/ml	4ml	£439.00 (£10.98/mg)	UK list price (CS Table 60)
		10ml	£1,097.00 (£10.98/mg)	
Docetaxel	10mg/ml	2ml	£7.45 (£0.37/mg)	eMit[62]
		8ml	£25.73 (£0.32/mg)	
		16ml	£35.35 (£0.22/mg)	
Dexamethasone	-	100 tablets	£5.16 cost per 21-day cycle	eMit[62]
Nintedanib	150mg	60 tablets	£2,151.10	PharmaTimes[63]

BNF=British National Formulary; NICE=National Institute for Health and Care Excellence; UK=United Kingdom
Source: CS, Table 60

Table 30 Drug acquisition cost per dose (initial treatments)

Drug	Total dose per administration	No. of vials per packs	Method of administration	Total drug cost per dose	Frequency of administration
Nivolumab	3mg/kg	1.19 × 10-ml vial*+ 1.84 × 4-ml vial	IV; no vial sharing (i.e. round up to nearest full vials)	£2,538.25	Every 2 weeks
Docetaxel	75mg/m^2	1.79 × 2-ml + 0.65 × 8 ml + 0.35 × 16-ml* + dexamethasone	IV; no vial sharing (i.e. round up to nearest full vials)	£47.59	Every 3 weeks
Nintedanib + docetaxel	-	2 tablets per day × 21 days plus the cost per dose of docetaxel	Oral	£1,553.29	2 tablets a day for 21 days =1 dose

IV=Intravenous
*The 4-ml vial (nivolumab) and 16-ml vial (docetaxel) are used in the base case because these are the smallest and cheapest vial sizes, respectively
Source: CS, Table 61

Subsequent treatments

The model includes costs of subsequent treatments for patients with PD based on the distribution of subsequent therapies observed in CheckMate 057. Table 31 presents drug acquisition costs per dose for these subsequent treatments. The treatment duration for all subsequent therapies is [redacted], based on RWD collected in an observational study (CA209-116[53]) in which the treatment patterns, resource use and outcomes of patients with advanced NSCLC in Europe were explored. The company made an assumption that the pooled RWD collected from European countries were applicable to clinical practice in the UK.

Table 31 Drug acquisition cost per dose (subsequent treatments)

Drug	Total dose required per administration	No. of vials / packs	Method of administration	Total drug cost per dose	Frequency of administration
Pemetrexed	500mg/m^2	2.15 × 500-ml vials	IV; no vial sharing (i.e. round up to nearest full vials)	£1,723.02	Every 3 weeks
Carboplatin	400mg/m^2	0.80 × £3.43 + 1.22 × £7.69 + 1.15 × £20.17	IV; no vial sharing (i.e. round up to nearest full vials)	£35.42	Every 4 weeks
Gemcitabine	1,000mg/m^2	1.44 × 200mg + 5.63 1,000mg	IV; no vial sharing (i.e. round up to nearest full vials)	£56.20	Every 4 weeks (once per week for 3 weeks, followed by 1 week off treatment)
Docetaxel	75mg/m^2	1.79 × 2-ml + 0.65 × 8-ml+ 0.35 × 16-ml + dexamethasone	IV; no vial sharing (i.e. round up to nearest full vials)	£47.59	Every 3 weeks
Erlotinib	150mg	1/30 pack (30 × 150mg)	Oral; vial sharing is N/A	£54.38	Daily

IV=Intravenous; N/A=not applicable
Source: CS, Table 63

Treatment administration costs

The costs of treatment administration for nivolumab and docetaxel, as applied in the model, are shown in Table 32.

Table 32 Cost per administration

Treatment	Type of administration		Currency code	Cost per administration*	Source
Nivolumab	Deliver simple parenteral chemotherapy at first attendance	Outpatient setting	SB12Z	£167.34	NHS reference costs 2013-2014[64]
Docetaxel	Deliver simple parenteral chemotherapy at first attendance	Outpatient setting	SB12Z	£167.34	NHS reference costs 2013-2014[64]

NHS=National Health Service
*All administration costs are assumed to be for first attendances in a cycle due to the length of time between administrations (for nivolumab and docetaxel, it is every 2 weeks and 3 weeks, respectively). All costs are inflated to June 2015 values.
Source: CS, Table 64

Health care costs

Patient monitoring, disease management and terminal care costs are provided in Table 33. A one-off, end of life/terminal care cost is applied to patients who enter the death state. These weighted costs reflect treatment received in various care settings.

Table 33 Health care costs

Type of cost	Health state	Cost*
Monitoring cost – nivolumab, docetaxel or nintedanib+docetaxel	Progression-free	£151.89 per 4 weeks
Disease management	Progression-free	£313.55 per 4 weeks
Disease management	Progressed disease	£766.62 per 4 weeks
Terminal care	Death	£3,628.70 (one off)

*2015 National Reference Costs for unit costs were unavailable at the time of submission, the company inflated costs to June 2015 values
Source: Adapted from CS, Tables 65-69

Adverse event costs

The base case analysis includes all Grade 3-5 AEs (regardless of causality) with ≥2% incidence in the nivolumab or docetaxel arms of CheckMate 057. AE costs and management costs per episode were sourced from NHS Reference Costs[65] guided by the currency codes used in recent NICE submissions in NSCLC and melanoma.[9,43,46] A summary of costs is presented in Table 34.

Table 34 Cost of adverse events

AEs from CheckMate 057	Cost per episode*	Mean number of episodes per AE treatment course	Source
Fatigue	£3,015.13	1	NHS Reference Costs 2013-2014[64]
Asthenia	£3,015.13	1	NHS Reference Costs 2013-2014[64]
Pain	£122.00	1	NHS Reference Costs 2013-2014[64]
Dyspnoea	£0.00	1	Assumption based on ipilimumab NICE STA submission for melanoma[9]
Pleural effusion	£553.00	1	NHS Reference Costs 2013-2014[64]
Hyperglycemia	£652.00	1	NHS Reference Costs 2013-2014[64]
Pneumonia	£1,822.85	1	NHS Reference Costs 2013-2014[64]
Neutrophil count decreased	£0.00	1	Assumption
White blood cell count decreased	£423.00	1	NICE 2015[43]
Anaemia	£978.00	1	NICE 2015[43]
Neutropenia	£354.72	1	NHS Reference Costs 2013-2014[64]
Febrile neutropenia	£5,489.94	1	Erlotinib and gefitinib (post-chemotherapy) MTA (rev TA162, TA175) [ID620]
Leukopenia	£354.72	1	Assumed to be same as neutropenia based on medical opinion
Diarrhoea	£1,796.00	1	NICE 2015[43]
Increased ALT	£587.00	1	NICE 2015[43]
Increased AST	£336.00	1	NICE 2015[43]
Hyponatraemia	£652.00	1	Assumed to be same as hyperglycaemia based on medical opinion

AE=adverse event; ALT=alanine aminotransferase; AST=aspartate transaminase; ID=in development; MTA=multiple technology appraisal; NHS=National Health Service; TA=technology appraisal
*All adverse event costs originating from pre-2015 sources are inflated to June 2015 values
Source: CS, Table 71

5.3.8 Cost effectiveness results

Base case results

The base case analysis was based on the generalised gamma curves for all extrapolations (OS and TTD). Life years were undiscounted. In comparison to docetaxel, nivolumab generated 0.73 incremental QALYs and 1.15 incremental life years, and the nivolumab-treated cohort incurred an increase in total costs. The incremental cost effectiveness ratio (ICER) was £103,589 per QALY gained. Total costs, LYG, QALYs, and incremental cost per QALY gained for nivolumab versus docetaxel are presented in Table 35.

In comparison to nintedanib+docetaxel, nivolumab generated 0.49 incremental QALYs and 0.80 incremental life years with a higher total cost. The ICER was £126,861 per QALY gained. Expected QALYs for nivolumab, docetaxel and nintedanib+docetaxel disaggregated

by health state are shown in Table 36 and Table 37. Predicted (per patient) resource use costs included in the company model are presented in Table 38 and Table 39.

Table 35 Base case results

Treatment	Total cost (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	Incremental cost per QALY (£)
Nivolumab	93,306	2.24	1.42				
Docetaxel	17,854	1.09	0.70	75,452	1.15	0.73	103,589
Nintedanib +docetaxel	30,708	1.44	0.93	62,598	0.80	0.49	126,861

LYG=life year gained; QALY=quality adjusted life year
Source: CS, Table 76

Table 36 Summary of QALY gain per patient by health state - nivolumab vs. docetaxel

Health state	QALY intervention (nivolumab)	QALY comparator (docetaxel)	Incremental QALYs	% absolute incremental QALYs
PF	[redacted]	[redacted]	[redacted]	[redacted]
PD	[redacted]	[redacted]	[redacted]	[redacted]
AE disutility	[redacted]	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]	[redacted]

AE=adverse event; PD=progressed disease; PF=progression-free; QALY=quality adjusted life year
Note: No utility is assigned to the death state
Source: CS, Table 79

Table 37 Summary of QALY gain per patient by health state - nivolumab vs. nintedanib+docetaxel

Health state	QALY intervention (nivolumab)	QALY comparator (nintedanib+docetaxel)	Incremental QALYs	% absolute incremental QALYs
PF	[redacted]	[redacted]	[redacted]	[redacted]
PD	[redacted]	[redacted]	[redacted]	[redacted]
AE disutility	[redacted]	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]	[redacted]

AE=adverse event; PD=progressed disease; PF=progression-free; QALY=quality adjusted life year
Note: No utility is assigned to the death state
Source: CS, Table 80

Table 38 Discounted cost per patient (disaggregated) - nivolumab vs. docetaxel

Health state	Cost intervention (nivolumab)	Cost comparator (docetaxel)	Incremental costs	% absolute incremental costs
Disease management cost: PF	[redacted]	[redacted]	[redacted]	[redacted]
Disease management cost: PD*	[redacted]	[redacted]	[redacted]	[redacted]
Drug acquisition cost	[redacted]	[redacted]	[redacted]	[redacted]
Administration cost	[redacted]	[redacted]	[redacted]	[redacted]
Monitoring cost	[redacted]	[redacted]	[redacted]	[redacted]
Subsequent treatment	[redacted]	[redacted]	[redacted]	[redacted]
AEs	[redacted]	[redacted]	[redacted]	[redacted]
Total treatment cost	[redacted]	[redacted]	[redacted]	[redacted]

AE=adverse event; PD=progressed disease; PF=progression-free
*PD includes the costs of managing patients who have progressed and end-of-life and terminal care. No costs are assigned to the death state
Source: CS, Table 80

Table 39 Discounted cost per patient (disaggregated) - nivolumab vs. nintedanib+docetaxel

Health state	Cost intervention (nivolumab)	Cost comparator (nintedanib+docetaxel)	Incremental costs	% absolute incremental costs
Disease management cost: PF	[redacted]	[redacted]	[redacted]	[redacted]
Disease management cost: PD*	[redacted]	[redacted]	[redacted]	[redacted]
Drug acquisition cost	[redacted]	[redacted]	[redacted]	[redacted]
Administration cost	[redacted]	[redacted]	[redacted]	[redacted]
Monitoring cost	[redacted]	[redacted]	[redacted]	[redacted]
Subsequent treatment	[redacted]	[redacted]	[redacted]	[redacted]
AEs	[redacted]	[redacted]	[redacted]	[redacted]
Total treatment cost	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: AE=adverse event; PD=progressed disease; PF=progression-free
*PD includes the costs of managing patients who have progressed and end-of-life and terminal care. No costs are assigned to the death state
Source: CS, Table 81

5.3.9 Sensitivity analyses

Deterministic sensitivity analysis

One-way sensitivity analyses of nivolumab versus docetaxel and nivolumab versus nintedanib+docetaxel were undertaken by varying cost, utility and OS base case parameter values by their confidence intervals or +/-20%, based on data availability (Table 40 and Table 41).

Table 40 Results of deterministic analysis vs. docetaxel

Parameter	Analysis	Incremental costs (£)	Incremental QALYs	Incremental cost per QALY (£)
Base case analysis		75,452	0.73	103,589
Discount rate: costs	Lower	85,895	0.73	117,928
	Higher	69,973	0.73	96,068
Discount rate: outcomes	Lower	75,452	0.88	85,753
	Higher	75,452	0.65	116,472
Average body weight	Lower	63,650	0.73	87,386
	Higher	87,528	0.73	120,169
BSA	Lower	75,500	0.73	103,655
	Higher	75,360	0.73	103,463
				Costs
Cost: PF state	Lower	74,908	0.73	102,843
	Higher	75,995	0.73	104,335
Cost: PD state	Lower	74,911	0.73	102,848
	Higher	75,992	0.73	104,331
Terminal cost	Lower	75,481	0.73	103,630
	Higher	75,422	0.73	103,549
Administration cost: nivolumab	Lower	74,567	0.73	102,375
	Higher	76,336	0.73	104,804
Administration cost: docetaxel	Lower	75,639	0.73	103,847
	Higher	75,264	0.73	103,331
Monitoring cost: nivolumab	Lower	75,054	0.73	103,043
	Higher	75,849	0.73	104,135
Monitoring cost: docetaxel	Lower	75,572	0.73	103,755
	Higher	75,331	0.73	103,424
				Outcomes
Utility weight, PFS	Lower	75,452	0.72	104,546
	Higher	75,452	0.73	102,743
Utility weight, PD	Lower	75,452	0.72	104,484
	Higher	75,452	0.73	102,672

BSA=body surface area; PD=progressed disease; PF=progression-free; PFS=progression-free survival; QALY =quality adjusted life year
Source: CS, Table 104

Table 41 Results of deterministic analysis versus nintedanib+docetaxel

Parameter	Analysis	Incremental costs (£)	Incremental QALYs	Incremental cost per QALY (£)
Base case analysis		62,598	0.49	126,861
Discount rate: costs	Lower	72,736	0.49	147,407
	Higher	57,308	0.49	116,141
Discount rate: outcomes	Lower	62,598	0.63	99,830
	Higher	62,598	0.42	147,671
Average body weight	Lower	50,796	0.49	102,943
	Higher	74,674	0.49	151,335
BSA	Lower	62,646	0.49	126,959
	Higher	62,506	0.49	126,676
Costs				
Cost: PF state	Lower	62,058	0.49	125,767
	Higher	63,138	0.49	127,956
Cost: PD state	Lower	62,686	0.49	127,040
	Higher	62,510	0.49	126,683
Terminal cost	Lower	62,663	0.49	126,993
	Higher	62,533	0.49	126,730
Administration cost: nivolumab	Lower	61,713	0.49	125,069
	Higher	63,482	0.49	128,654
Administration cost: docetaxel	Lower	62,809	0.49	127,289
	Higher	62,387	0.49	126,433
Monitoring cost: nivolumab	Lower	62,200	0.49	126,055
	Higher	62,996	0.49	127,668
Monitoring cost: docetaxel	Lower	62,734	0.49	127,137
	Higher	62,462	0.49	126,585
Outcomes				
Utility weight, PFS	Lower	62,598	0.49	128,588
	Higher	62,598	0.50	125,347
Utility weight, PD	Lower	62,598	0.49	126,601
	Higher	62,598	0.49	127,134
Survival outcomes				
HR on PFS: nintedanib+docetaxel	Lower	60,246	0.49	123,209
	Higher	64,293	0.50	129,442
HR on OS: nintedanib +docetaxel	Lower	59,328	0.28	214,630
	Higher	65,217	0.66	98,353

BSA=body surface area; PD=progressed disease; PF=progression-free; PFS=progression-free survival; QALY=quality adjusted life year
Source: CS, Table 105

Scenario analyses

The scenario analyses involved varying the survival modelling approaches applied to OS and TTD data and duration of treatment. With regards to the alternative treatment duration scenarios, treatment stopping rules were implemented by terminating all treatment-related costs (i.e. acquisition, administration and monitoring) at either 1 or 2 years. The treatment stopping rule was applied in order to represent patients who experienced a durable response i.e. maintenance of clinical benefit after treatment discontinuation prior to progression. The influence of each change on the size of ICER per QALY gained is presented in Table 42.

Table 42 Scenario analyses results

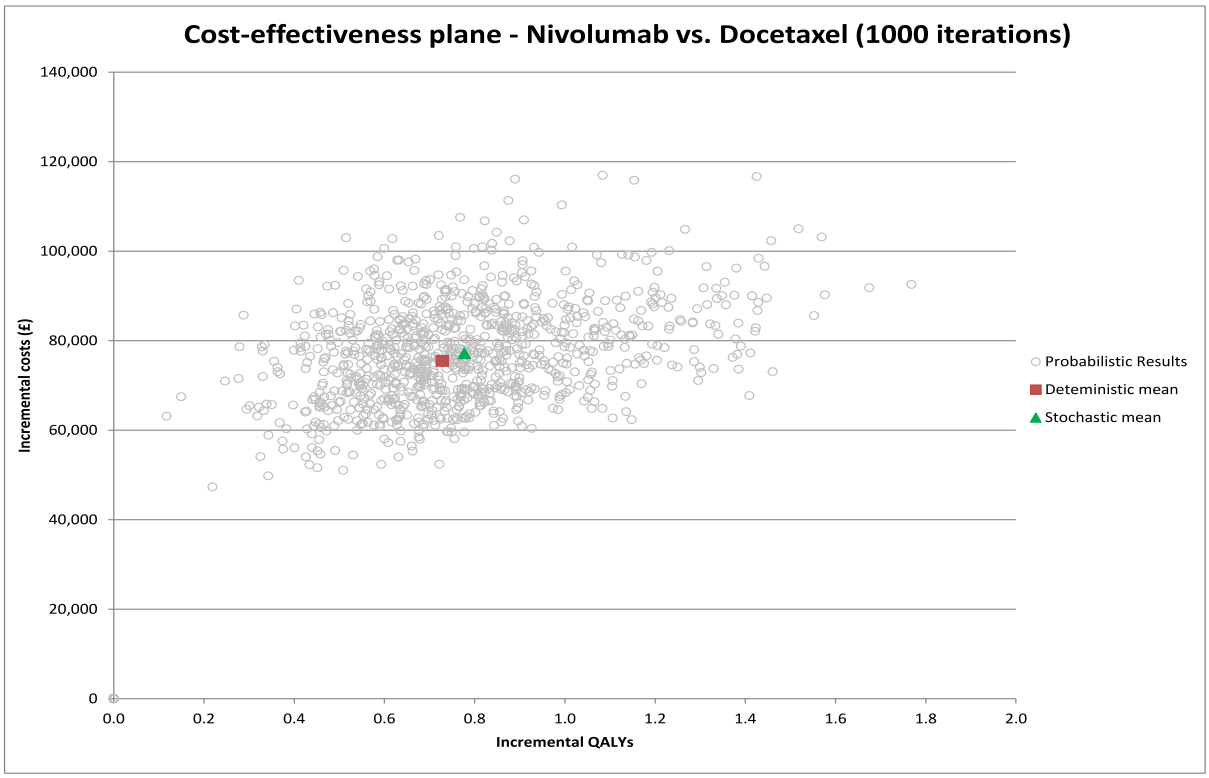
Description	ICER per QALY gained
Base case – nivolumab vs. docetaxel	£103,589
Base case – nivolumab vs. nintedanib+docetaxel	£126,861
Survival analysis - OS	
2-knot spline hazards model for nivolumab and gamma distribution for docetaxel – nivolumab vs. docetaxel	£144,594
2-knot spline hazards model for nivolumab and gamma distribution for docetaxel– nivolumab vs. nintedanib+docetaxel	£195,348
Survival analysis - TTD	
1-knot spline hazards model for nivolumab and gamma distribution for docetaxel – nivolumab vs. docetaxel	£120,773
1-knot spline hazards model for nivolumab and gamma distribution for docetaxel – nivolumab vs. nintedanib+docetaxel	£149,112
Duration of treatment	
1-year treatment stopping rule for nivolumab - nivolumab vs. docetaxel	£46,860
1-year treatment stopping rule for nivolumab - nivolumab vs. nintedanib+docetaxel	£43,122
2-year treatment stopping rule for nivolumab - nivolumab vs. docetaxel	£60,955
2-year treatment stopping rule for nivolumab - nivolumab vs. nintedanib+docetaxel	£63,928

ICER=incremental cost effectiveness ratio; OS=overall survival; QALY=quality adjusted life year; TTD=time to treatment discontinuation
Source: CS, adapted from Tables 108, 111, 114, 117

Probabilistic sensitivity analysis

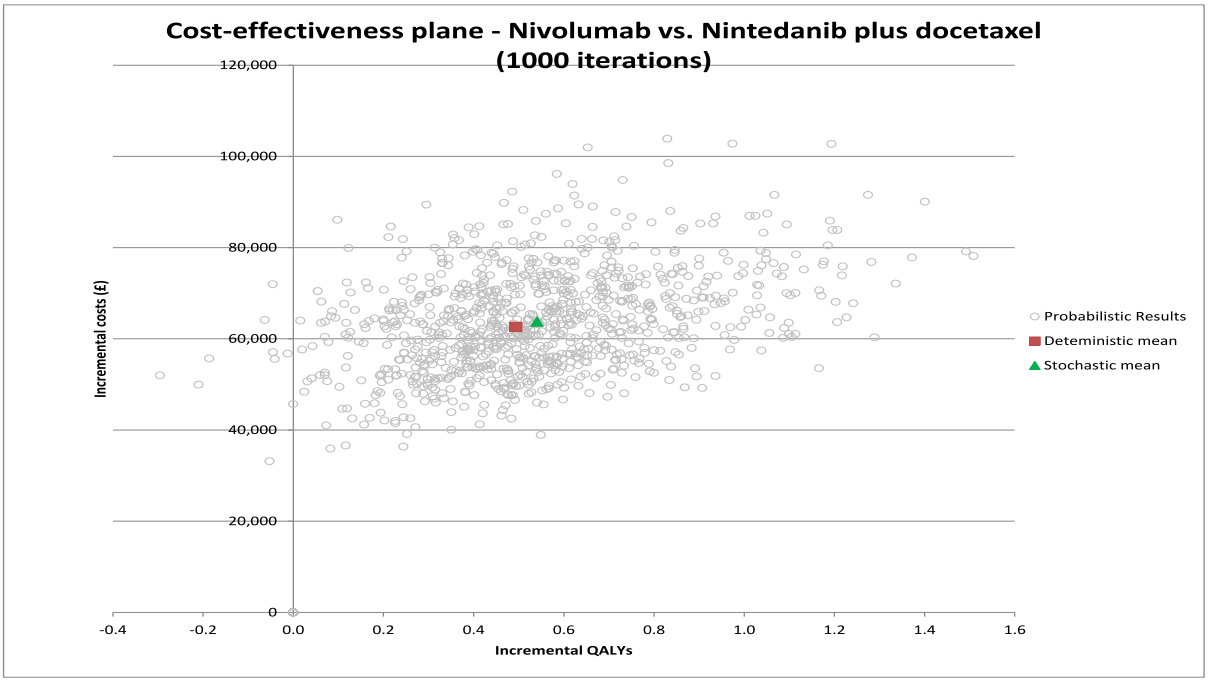
The company undertook probabilistic sensitivity analysis (PSA) to derive the mean ICER per QALY gained for nivolumab versus docetaxel and nivolumab versus nintedanib+docetaxel. The PSA was run for 1000 iterations. The probabilistic ICER versus docetaxel is £99,291 per QALY gained compared with £103,589 per QALY gained in the deterministic analysis. The probabilistic ICER versus nintedanib+docetaxel is £111,934 per QALY gained compared with £126,861 per QALY gained in the deterministic analysis. For these comparisons, the

cost effectiveness planes are shown in Figure 8 and Figure 9 and the cost effectiveness acceptability curve for both comparators is shown in Figure 10.



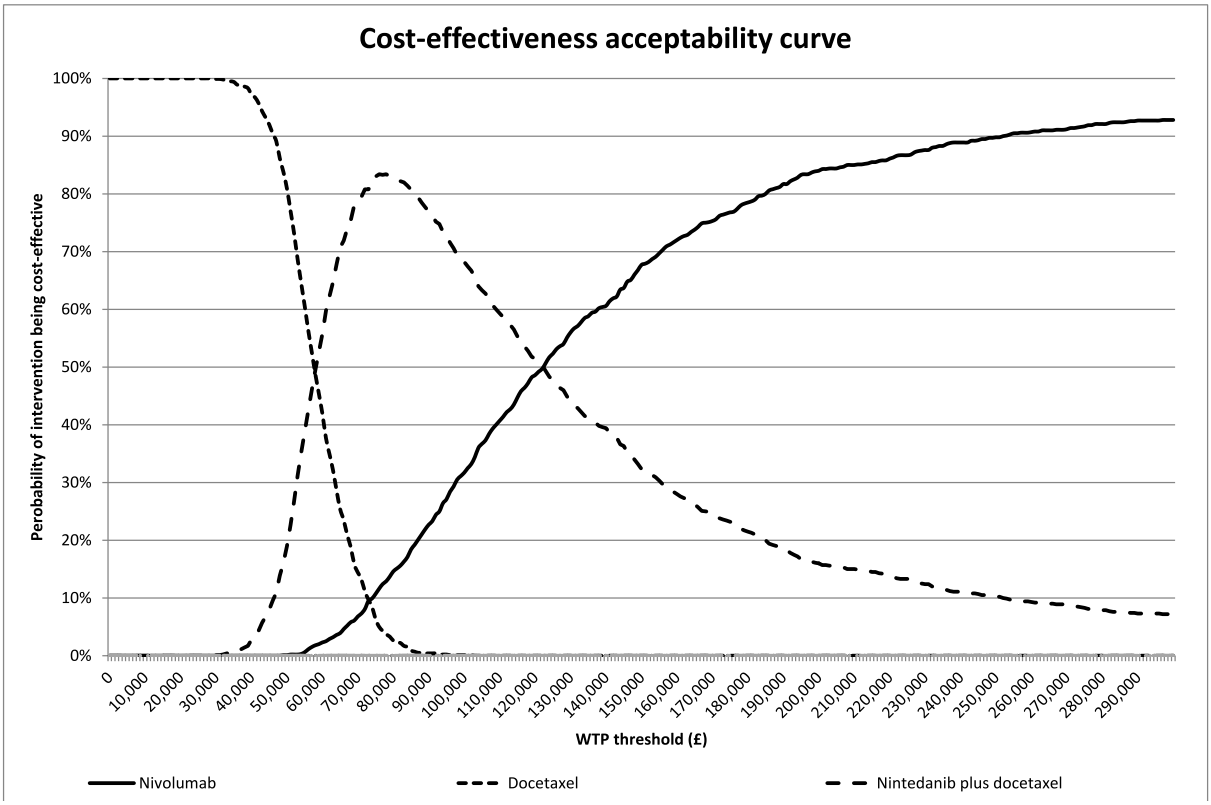
QALY=quality adjusted life year
Source: CS, Figure 43

Figure 8 Scatter plot-cost effectiveness of nivolumab vs. docetaxel (1,000 iterations)



QALY=quality adjusted life year
Source: CS, Figure 44

Figure 9 Scatter plot for cost effectiveness of nivolumab vs. nintedanib+docetaxel (1,000 iterations)



WTP=willingness to pay
Source: CS, Figure 45

Figure 10 Cost effectiveness acceptability curve of nivolumab vs. docetaxel and nintedanib+docetaxel

5.3.10 Model validation and face validity check

The company states that their survival models were validated against data from CheckMate 057,[28] CheckMate 003[54] and the NLCA dataset.[10] In addition, during model development, external clinical and health economic experts attended four workshops and provided advice during ad hoc consultations.

5.4 Exploratory and sensitivity analyses undertaken by the ERG

5.4.1 NICE reference case and Drummond critical appraisal

A summary of the checklists for the reference case and the Drummond critical appraisal are presented in Table 43 and Table 44.

Table 43 NICE Reference case checklist completed by ERG

Attribute	Reference case	Does the de novo economic evaluation match the reference case?
Decision problem	The scope developed by NICE	Yes
Comparator(s)	As listed in the scope developed by NICE	Partial. BSC was not subject to a full economic evaluation
Perspective costs	NHS and PSS	Partial. The model only includes NHS costs. Personal Social Service costs have not been considered
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	Patient related direct health effects are considered. No impact on carers has been considered in the model
Form of economic evaluation	Cost utility analysis with fully incremental analysis	Yes
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	Yes – 20 year time horizon
Synthesis of evidence on outcomes	Based on systematic review	Yes – data primarily taken from CheckMate 057
Outcome measure	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of HRQoL in adults	Yes – health effects are expressed in QALYs and the EQ-5D instrument has been used to collect HRQoL data
Source of data for measurement of HRQoL	Reported directly by patients and/or carers	Yes
Source of preference data for valuation of changes in HRQoL	Representative sample of the UK population	Yes
Discount rate	The same annual rate for both costs and effects (currently 3.5%)	Benefits and costs have been discounted at the 3.5% rate
Equity	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	Yes
Evidence on resource use and costs	Costs should relate to NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS	Partial. NHS costs, valued at relevant prices, have been used. PSS costs are not included in the model

EQ-5D=EuroQol 5-dimension; HRQoL=health related quality of life; NHS=National Health Service; NICE=National Institute for Health and Care Excellence; PSS=Personal Social Services; QALY=quality adjusted life year

Table 44 Critical appraisal checklist for the economic analysis completed by the ERG

Question	Critical appraisal	ERG comment
Was a well-defined question posed in answerable form?	Yes	-
Was a comprehensive description of the competing alternatives given?	Yes	-
Was the effectiveness of the programme or services established?	Partially	Limited data available from Checkmate 057
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?	Yes	-
Were the cost and consequences valued credibly?	Partially	The ERG considers that the company's OS and PFS/TTD projections lack clinical credibility and overestimate the effectiveness of nivolumab
Were costs and consequences adjusted for differential timing?	Yes	Discount rate of 3.5% per annum
Was an incremental analysis of costs and consequences of alternatives performed?	Yes	ICER calculated correctly
Was allowance made for uncertainty in the estimates of costs and consequences?	Yes	Deterministic, scenario and probabilistic sensitivity analyses were undertaken
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 company

ICER=incremental cost effectiveness ratio; PFS=progression-free survival; TTD= time to treatment discontinuation

5.5 The company model

The company's Microsoft Excel spreadsheet model is constructed according to conventional practice and is generally implemented correctly. The company is to be commended for taking previous ERG comments into account regarding the implementation of certain features of the model.

5.5.1 Health-state modelling: key issues

The elements of the company's cost effectiveness evidence that cause most concern to the ERG relate to the modelling of patients in two health states: PFS and post-progression survival (PPS). In the model, PPS is based on the modelling of PFS and OS and it is assumed that PPS = OS-PFS. The ERG considers the modelling of PFS, PPS and OS to be flawed for both the intervention and the comparators. The issues become more problematic when nivolumab is compared with nintedanib+docetaxel rather than when compared with

docetaxel monotherapy. The ERG's proposed amendments to the company model have a substantial impact on the size of each of the estimated ICERs per QALY gained.

The specific issues of concern identified by the ERG relate to the modelling of each health state and are compounded by the results of the ERG's examination of subgroups of patients within the group of patients treated with nivolumab; these subgroups have not been discussed in the CS. The ERG identified two distinct patient subgroups according to whether patients received treatment with nivolumab post-progression. Information describing the baseline characteristics of the patient subgroups was not available to the ERG at the time of analysis. Therefore, the ERG cannot ascertain whether there are fundamental differences between the groups other than that some patients received treatment with nivolumab after progression (PPTx) and other patients did not receive nivolumab after progression (no-PPTx).

The specific survival modelling issues identified by the ERG are as follows:

- the interdependence in the model between OS, PFS and all-cause mortality rates results in implausible projections for nivolumab PFS in particular, but also casts doubt on the reliability of the model used to estimate nivolumab OS
- survival gain is predominantly accrued in the PFS state for nivolumab in the company model, whereas the trial evidence suggests that nivolumab has a substantial post-progression benefit over docetaxel. This is particularly true for the PPTx patient subgroup
- the gamma parametric model chosen to model OS for nivolumab is not a good fit to the K-M data from CheckMate 057. The CheckMate 003[23] data used to validate the nivolumab OS model are inappropriate as the survival profiles are different
- TTD data have been used instead of PFS data in all parts of the company model. There are two key issues. First, the projection of TTD data as a proxy for PFS data is implausibly long and results in 85% of patients being still alive at 20 years, remaining progression-free and being on treatment. Second, the ERG considers that TTD data should only be used for estimating costs and not for estimating QALYs accruing in the different health states
-

- the piecewise proportional hazards assumption is not supported for OS or PFS in the LUME-Lung 1[24] trial, which invalidates the company's indirect method of comparing nivolumab with nintedanib+docetaxel.

The ERG deals with several of these issues using its preferred method of modelling survival i.e. using as much direct trial (K-M) data as possible, and only projecting future survival for the subgroup of patients remaining at risk towards the end of the reported trial data. This method ensures that as much of the available trial evidence as possible is used and limits uncertainty to the projection period only.

Other issues identified by the ERG relate to the use of utility values from CheckMate 057, an error in the calculation of nivolumab dosing and the timing of treatment administration costs. The ERG has also performed a sensitivity analysis based on the company's 1- and 2-year stopping rule scenarios.

5.5.2 Nivolumab treatment subgroups

The protocol[52] for the pivotal CheckMate 057 trial notes that there is accumulating evidence that a minority of patients treated with immunotherapy may derive clinical benefit despite exhibiting initial evidence of progressed disease. As a result, patients in the trial were permitted to continue on study treatment beyond progression (as defined by RECIST 1.1) as long as:

- they continued to derive clinical benefit from nivolumab (as assessed by investigator) and did not have rapidly progressing disease;
- they tolerated the drug;
- they had stable PS;
- the intervention to prevent serious complications of disease progression would not be delayed and the subject had provided written consent.

Further progression was defined as an additional 10% increase in tumour volume from the time of the initial progression, at which point treatment was discontinued permanently.

During the clarification process, the ERG requested details of the number of patients treated beyond progression. The ERG also requested survival data split by whether nivolumab patients received treatment-post progression. The ERG did not request clarification data on the baseline characteristics of the different nivolumab treatment subgroups, so is unable to provide the results of further analyses based on more detailed patient information.

As stated in the clarification response, during CheckMate 057, 25% of all nivolumab patients (n=72) had received treatment beyond progression by the 18-month data lock, and 22.5% (n=16) of these patients met the criteria for 'non-conventional benefit'. Non-conventional benefiters are subjects whose best confirmed objective response was not PR/CR and who met at least one of the following:

- appearance of a new lesion followed by decrease from baseline of at least 10% in sum of target lesions; or
- initial increase from nadir >=20% in sum of target lesions followed by reduction from baseline of at least 30%; or
- initial increase from nadir >=20% in sum of target lesions or appearance of new lesion followed by at least 2 tumour assessments showing no further progression defined as 10% additional increase in sum of target lesions and new lesions.

Based on the proportion of nivolumab patients who had received treatment beyond progression at the time of the 18-month data lock, the ERG has assumed that 25% of patients treated with nivolumab are permitted to continue treatment beyond progression. This assumption affects the analysis of PPS and OS, but not PFS.

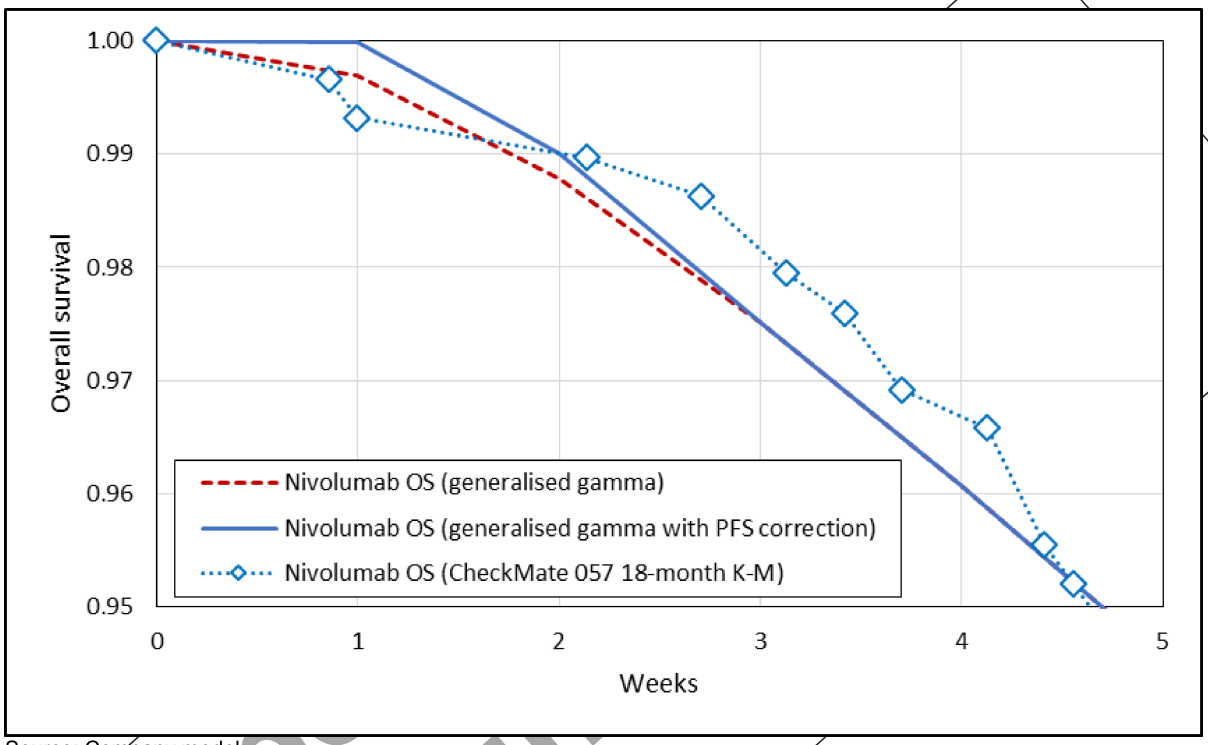
It is not valid to split the PFS K-M data according to whether nivolumab patients were treated beyond progression, as progression is a prerequisite of belonging to the PPTx group. Membership of the post-progression treatment subgroups is decided once a patient has left the progression-free state, so it is not appropriate to reverse-assign them to either the PPTx or no-PPTx group whilst in PFS. If other characteristics were identified that could predict whether a patient would receive treatment beyond progression whilst the patient was still progression-free, then the split would be valid. The ERG has not had access to detailed patient data that might be able to identify such patients before they progress.

5.5.3 Interdependence of health-state models

The company model is built with two 'check and substitute' mechanisms that link PFS, OS and all-cause mortality rates in a way that is not credible and undermines the projection of PFS and OS, particularly for patients receiving nivolumab.

First, OS is linked with PFS to ensure that PFS is never greater than OS for any treatment in the company model. Should the modelled curves result in a greater value for PFS than OS in any given week, the PFS value is used instead of the OS value. It can be seen in Figure 11 that the distribution used to model nivolumab OS produces lower values than the distribution used to model nivolumab PFS in the first 2 weeks. The company model then substitutes PFS

values for OS in the first 2 weeks so that there are not more people in PFS than people who are alive. The choice of parametric distribution for either PFS or OS (or both) is therefore inappropriate, as their combination produces implausible values and cannot be used without adjustment. Figure 11 also emphasises the uncertainty in the fit of the generalised gamma curve to the K-M data during this early period.

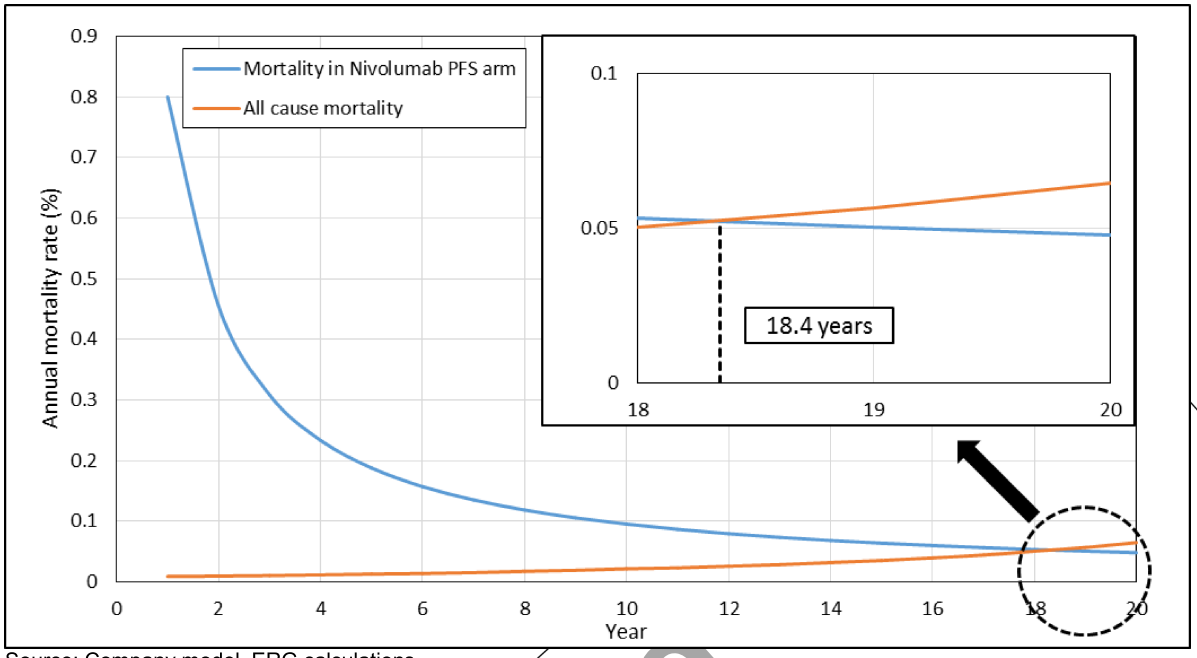


Source: Company model

Figure 11 Nivolumab OS with modelled generalised gamma curve, PFS correction and CheckMate 057 18-month K-M data

Second, the nivolumab arm is subject to a 'check and substitute' mechanism to ensure that disease specific mortality rates do not fall below all-cause mortality rates. Projections for both PFS and OS are compared to age- and sex- adjusted all-cause mortality rates and, should the latter be greater than the modelled rates, a substitution is made.

Figure 12 shows that the gamma model used by the company to model PFS for patients treated with nivolumab projects annual mortality rates that fall below all-cause mortality rates 18.4 years after patients begin treatment. Hence, the model forecasts that any patient who remains in PFS for 18.4 years will never progress and is essentially cured of the disease. This is a very strong assumption for the company to make without providing supporting clinical evidence.



Source: Company model, ERG calculations

Figure 12 Relationship between nivolumab PFS and all-cause annual mortality rates in the company model

5.5.4 Post-progression survival: nivolumab and docetaxel

When compared with docetaxel, the company model estimates that patients treated with nivolumab accrue 31% of mean survival gain during PPS; this 31% gain equates to a survival gain of 4.3 months.

On inspection of the cohort trace for nivolumab (Figure 13), it is clear that the proportion of survival gain attributable to PPS is influenced considerably by the implausibly long PFS tail in the nivolumab arm. In the company model, PFS is modelled with a tail so long that 85% of the nivolumab patients who are still alive at 20 years are in PFS and are still receiving treatment. In comparison, almost all of the patients treated with docetaxel (>99.9%) are estimated to have left the progression-free state by 1.8 years when only 17% of these patients are still alive.

Source: CS, Figure 40

Figure 13 Cohort trace for nivolumab up to 20 years (company model base case analysis)

The ERG considers the PPS estimates from the model to be unreliable as a consequence of a flawed approach to modelling PFS. The ERG therefore requested PPS K-M data from CheckMate 057 to perform an independent analysis.

Examination of the PPS data at the time of the 18-month data lock shows that there is little difference in survival rates for all patients in the nivolumab and docetaxel arms immediately after progression, the curves then separate around 5 months and then converge again at around 20 months (Figure 14). This implies that, compared to docetaxel, nivolumab has only a small incremental effect on PPS. However, the amalgamated all-patient nivolumab PPS data conceal substantial differences between the PPTx and the no-PPTx subgroups. PPS for patients treated with nivolumab until disease progression have PPS indistinguishable from patients treated with docetaxel (log-rank test, p=0.84), whereas patients treated with nivolumab beyond progression have a much better chance of survival post-progression than other patients treated with nivolumab or patients treated with docetaxel (Figure 15).

Source: Clarification response-question B1c

Figure 14 Nivolumab (all patients) vs. docetaxel PPS K-M (18-month data cut)

PPTx=Received treatment post progression; no PPTx=Did not receive treatment post progression
Source: Clarification response-question B1g

Figure 15 Nivolumab (PPTx patients), nivolumab (no PPTx patients) and docetaxel PPS K-M data (18-month data cut)

The effect of these differences in PPS between the PPTx and no-PPTx nivolumab subgroups versus docetaxel depends on the proportion of patients receiving each treatment who die in PFS and on the proportion of patients in each of the nivolumab subgroups.

5.5.5 Overall survival: nivolumab versus docetaxel

The company employed an independent-curve approach to modelling OS for nivolumab and docetaxel, as it noted that proportional hazards models and single survival models were not appropriate due to the crossing of the nivolumab and docetaxel OS curves of CheckMate 057 at around 7 months. The company explored the use of a number of OS models and concluded that separate generalised gamma models had the most appropriate fit to both the nivolumab and docetaxel arms of the trial.

The ERG has identified three flaws in the company approach to modelling OS for patients in the nivolumab arm:

- the chosen gamma curve systematically underestimates most of the K-M data and so represents a poor fit to the nivolumab data from CheckMate 057
- the K-M data from CheckMate 003[23] Phase 1b clinical trial that were used to validate the projection exhibit a different survival profile to the data from CheckMate 057
- the modelled OS curve in the company model does not relate appropriately to the modelled PFS curve, as noted in Section 5.5.3 of this ERG report.

Inspection of the company model OS curve against the 18-month K-M data for nivolumab shows that the fitted distribution systematically underestimates the trial data from 7 months to 20 months (Figure 16). This means that the fitted curve has not adequately incorporated all of the evidence on survival from CheckMate 057 for nivolumab patients who live beyond 7 months and relies too heavily on the pattern of survival during the first 7 months from randomisation. It is desirable to use all of the available clinical data when projecting survival. However, it is not always possible to fit a single parametric curve to the K-M data from time 0 without systematically misrepresenting that data to some extent. The principle objective of fitting a curve to K-M data is to be able to project a trend beyond the limits of available evidence, so it is preferable to closely model trends that are established later in the data and trends that might reasonably be expected to continue in the long-term rather than to seek a

parametric distribution that fits well to earlier K-M data but does not adequately capture the later evidence.

Source: Clarification response-question B1a, company model

Figure 16 Nivolumab OS: 18-month K-M data vs. company model

At the time of writing the CS, the company had only 12 months of data with which to project 20 years of survival. However, the company was able to provide the ERG with 18 months of data during the clarification process. The company attempted to mitigate the uncertainty inherent in extrapolating immature data by comparing potential OS models to other clinical studies and to RWD, namely the single-arm, Phase 1b CheckMate 003[23] trial and the UK's NLCA database.[10]

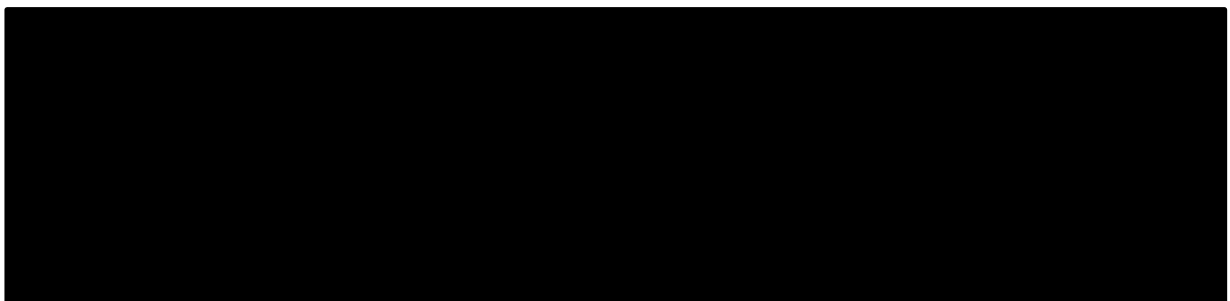
Figure 17 compares the K-M OS data from the CheckMate 057[28] and CheckMate 003[23] trials. It is clear from this plot that the survival profiles differ markedly between the two trials from around 7 months. The ERG therefore considers CheckMate 003[23] trial data to be unsuitable for validating projections based on data from CheckMate 057.

Source: Clarification response-question B1a, BMS 2015h[54]

Figure 17 Nivolumab OS K-M data from CheckMate 057 and CheckMate 003

It is good practice wherever possible to use the same functional form to model survival in both the intervention and comparator arms. So, although the company's generalised gamma model appears to be a better fit to the docetaxel 18-month OS data from CheckMate 057 than it is to the nivolumab K-M OS data, the ERG re-analysed the K-M OS data for both arms of CheckMate 057 to investigate alternative methods of extrapolating survival.

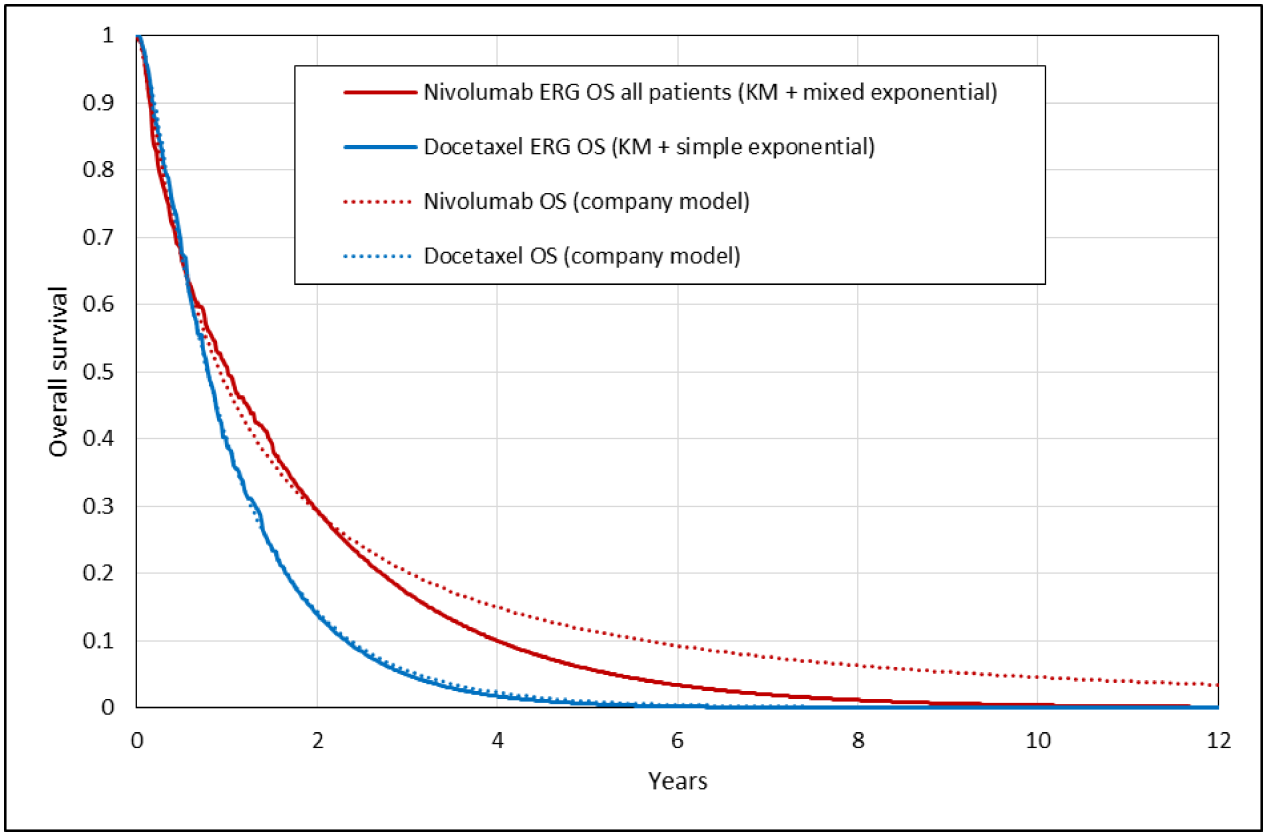
The ERG examined the cumulative hazard plot of the 18-month OS K-M data from CheckMate 057 (Figure 18). It is clear that both of the nivolumab subgroups (PPTx and no-PPTx) and the survival of patients in the docetaxel arm can be satisfactorily modelled using simple exponential distributions; from around 8 months for the nivolumab PPTx patients and docetaxel patients, and from 12 months for the no-PPTx patients. Long-term hazards in the nivolumab subgroups are very similar and much of the difference in survival occurs before 10 months.



Source: Adapted from clarification response-questions B1a & B1e

Figure 18 Cumulative hazard plot of OS K-M data from CheckMate 057 (18-month data cut)

As exponential models fit well to OS K-M data for both nivolumab subgroups, the ERG built a mixed exponential model based on 25% of patients receiving nivolumab beyond progression to project OS for the full nivolumab cohort. The ERG then appended the mixed exponential model to the K-M data for the full nivolumab cohort and modelled the docetaxel arm using K-M data followed by a simple exponential projection (Figure 19).



Source: Company model, clarification response-question B1a

Figure 19 Nivolumab and docetaxel OS: ERG model and company model

When the ERG compared its alternative projections for nivolumab OS to the three best-fitting survival models from the CS, it noted that its exponential extrapolation produced similar values to the company's 2-knot spline model (Figure 20). This is not unexpected as the spline model is a piecewise model that projects exponentially beyond the limit of the K-M data. The company reported that the 2-knot spline model had the best statistical fit of all of the models that were explored, but rejected this model on the basis that it projected lower OS rates at 2, 3 and 4 years than the K-M data from CheckMate 003[23] trial.

Source: Company model, clarification response-question B1a

Figure 20 Nivolumab OS ERG mixed exponential extrapolation and company 2-knot spline model

The company model projects an implausibly long PFS tail for patients receiving nivolumab, with 85% of the patients who are still alive at 20 years still being in PFS. It can be seen from Figure 21 that the ERG's amended OS projection dips below the company's PFS projections at approximately 5 years. The model's 'check and substitute' mechanism discussed in Section 5.5.3 is activated at this point, which artificially increases the ERG's OS projections to ensure that OS is not below PFS. It also means that implementing the ERG's alternative OS projections in isolation in the company model results in there being no patients in the PD state from 5 years onwards, so all patients who had progressed before 5 years have died and no further patients progress. These are purely functions of the interaction between the ERG's and company's projections, and the ERG does not consider them to be plausible clinical scenarios.

Source: Company model, ERG calculations

Figure 21 Effect of company's nivolumab PFS model on ERG nivolumab OS model

Applying the ERG's exponential projections to both the nivolumab and docetaxel arms from 7 months to 20 years reduces OS gain in the company model by 58% to 5.8 months. The estimated ICER per QALY gained is increased by £40,395 to £143,984.

5.5.6 Progression-free survival and time to treatment discontinuation: nivolumab versus docetaxel

The company does not use PFS data from CheckMate 057 in its model. The company uses TTD data from CheckMate 057 as a proxy for PFS data in order to be able to capture the extra treatment received by nivolumab patients who were treated beyond progression. However, the base case model that the company uses to project TTD data for both nivolumab and docetaxel is a poor fit to the available K-M data for both TTD and PFS and is not an appropriate surrogate for either.

The company concluded that a generalised gamma model was the most appropriate fit to the TTD K-M data in both the nivolumab and docetaxel arms. Figure 22 compares the generalised gamma curves used in the company model with the TTD K-M data from the 18-

month data lock provided by the company during the clarification process. It is clear that the company model is a poor fit to the TTD data and substantially overestimates time on treatment in the early part of the model for patients in both trial arms.

Source: Company model, clarification response-question B1d

Figure 22 TTD K-M data and company model for nivolumab and docetaxel

The CS did not include any PFS projections based on PFS K-M data from CheckMate 057 as the number of patients in PFS was estimated from projections of TTD K-M data. When the company's TTD models are compared against the PFS K-M data for both nivolumab and docetaxel, it is clear that the gamma curves are again inappropriate. Figure 23 shows that the TTD model used to estimate PFS for nivolumab overestimates almost all of the data and captures only a few of the final points. Conversely, the company TTD model serves to underestimate PFS for docetaxel.

Source: Company model, clarification response-question B1b

Figure 23 PFS K-M data and company model for nivolumab and docetaxel

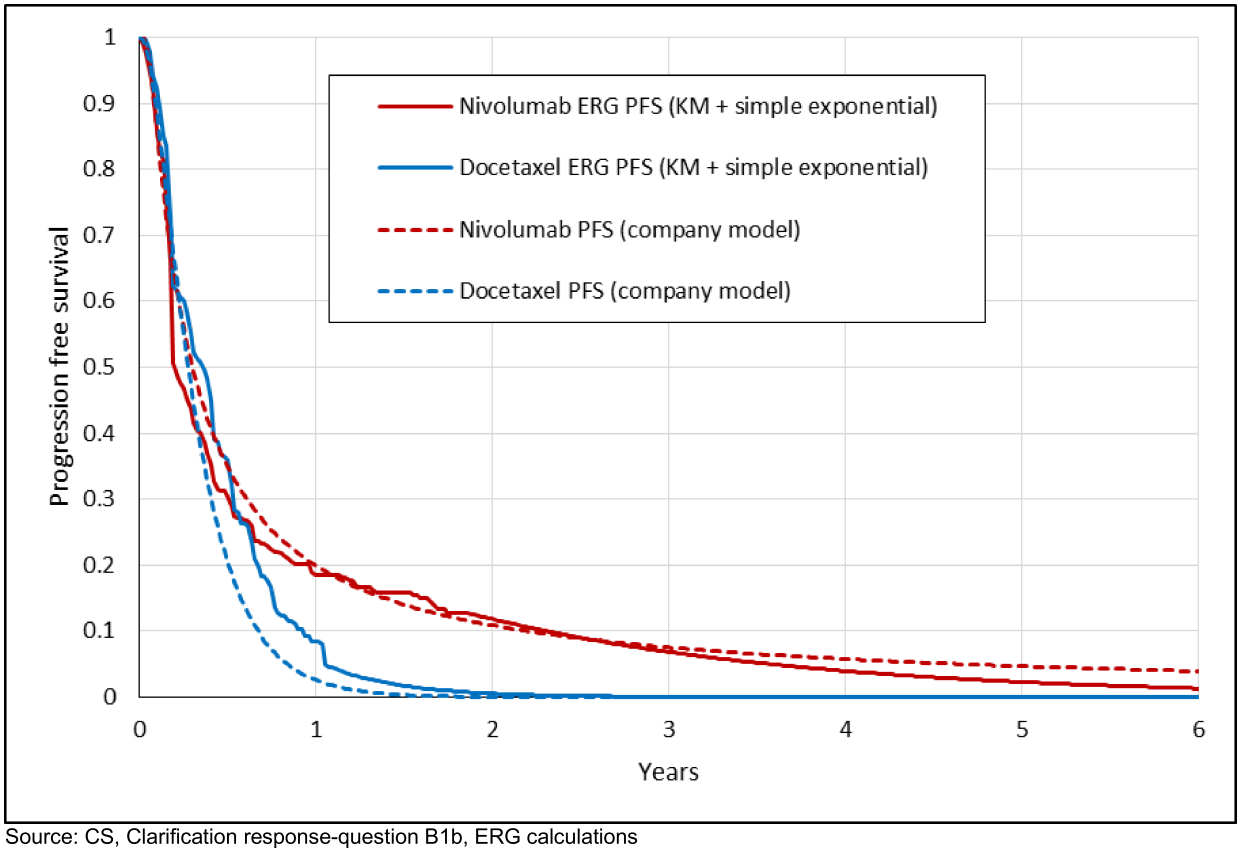
Progression-free survival projections

A cumulative hazard plot from 3 months indicates that nivolumab PFS hazards are constant from around 8 months; an exponential curve is therefore an appropriate method of projecting PFS for nivolumab in the long-term (Figure 24). Docetaxel PFS also exhibits constant hazards from around 8 months onwards, allowing exponential projections to be fitted to the end of the trial data.

Source: Adapted from clarification response-question B1b

Figure 24 Cumulative hazard plot of nivolumab and docetaxel PFS KM data (CheckMate 057)

Figure 25 compares the ERG's preferred PFS models for nivolumab and docetaxel with the company's generalised gamma model. The ERG's models decrease PFS gain in the company model by 57.9% to 4 months, as nivolumab PFS is reduced by shortening the long tail and docetaxel PFS is increased primarily in the first year.



Source: CS, Clarification response-question B1b, ERG calculations

Figure 25 ERG and company PFS projections for nivolumab and docetaxel

Using the ERG's PFS projections instead of the company TTD projections in the company model reduces the size of the ICER per QALY gained by £22,649 to £80,940 due to decreases in treatment acquisition, treatment administration and treatment monitoring costs.

Time to treatment discontinuation projections

Inspection of Figure 26 shows nivolumab TTD and PFS data to be very similar, particularly after 3 months; however, the ERG considers it appropriate to capture the differences between the PFS and TTD K-M data because of their relevance to treatment cost (TTD) and QALY (PFS) calculations.

Source: Clarification response-question B1b and B1d, ERG calculations

Figure 26 Nivolumab PFS and TTD K-M data and long-term trends

All patients in the docetaxel arm of CheckMate 057 had finished treatment by the time of the 18-month data cut (Figure 27). This means that no projections were necessary to estimate TTD for docetaxel, as the area under the K-M curve provides the best estimate of mean treatment duration for all of these patients.

Source: Clarification response-question B1d, ERG calculations

Figure 27 ERG TTD for nivolumab and docetaxel

The ERG's TTD estimates decrease the projected time that patients spend receiving nivolumab from 14.5 months to 10.3 months and the time that patients spend receiving docetaxel from 5 months to 4.2 months.

Using the ERG's TTD projections instead of the company TTD projections in the company model reduces the size of the ICER per QALY gained by £22,077 to £81,513 due to proportionately greater decreases in treatment acquisition, treatment administration and treatment monitoring costs for nivolumab.

5.5.7 Overall survival: nivolumab versus nintedanib+docetaxel

There is no direct clinical evidence to compare nivolumab with nintedanib in combination with docetaxel (nintedanib+docetaxel) for patients with progressed non-squamous lung cancer. The company notes that it was not possible to carry out a conventional ITC for the comparison of nivolumab versus nintedanib+docetaxel as the standard proportional hazards assumption was shown in TA347[43] not to hold for OS in the LUME-Lung 1[24] trial for the adenocarcinoma population.

The company analysed K-M OS data for the adenocarcinoma subgroup digitised from the published LUME-Lung 1 trial[24] and concluded that the OS K-M data for nintedanib+docetaxel

versus docetaxel+placebo had a 2-part proportional-hazard profile; i.e. there was no difference between the two treatments up to 6 months (HR=1) and this was followed by a separation of the curves showing benefit for nintedanib+docetaxel (HR=0.75).

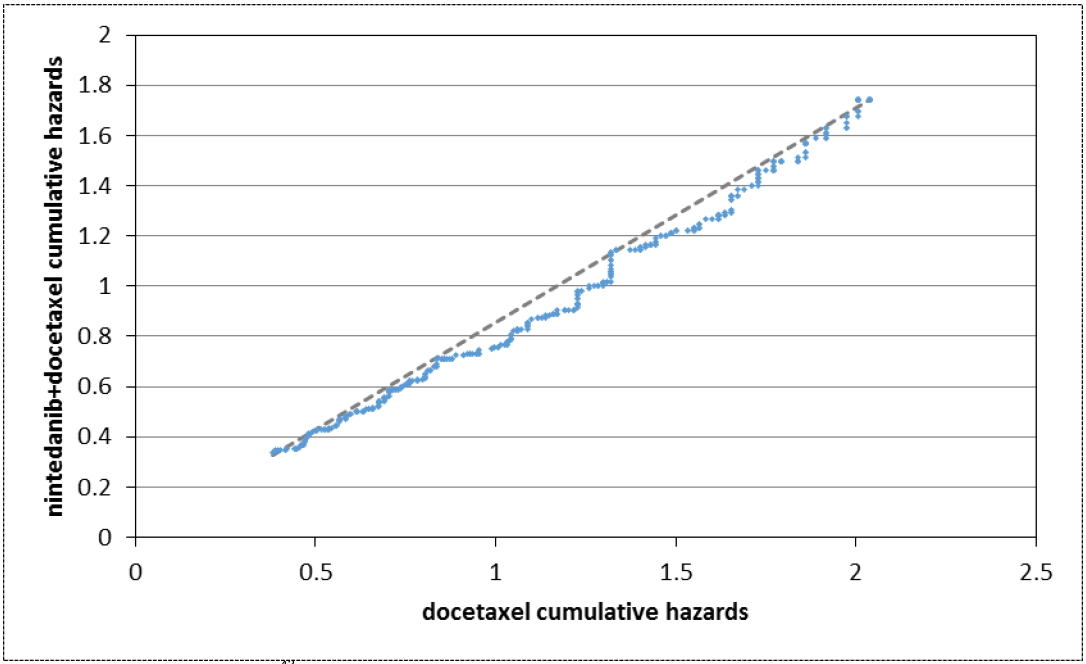
This approach relies on the proportional hazards assumption holding between 0 and 6 months, and again from 6 months onwards. In the CS there was no evidence presented to suggest that these proportional hazards assumptions had been tested by the company in any way beyond a visual inspection of the OS curves. The ERG independently digitised published data from TA347,[43] which contains more mature OS data from the LUME-Lung 1[24] trial than have been included in the CS, in order to estimate the proxy patient-level data required to test the proportional hazards assumptions.

Figure 28 shows an H-H plot of the docetaxel versus nintedanib+docetaxel arms of the LUME-Lung 1[24] trial from 6 months to 30 months. Points are dotted randomly around a line joining the first and last points in an H-H plot if hazards are proportional in the two arms. However, Figure 28 shows that the data points curve systematically below the line. This indicates that the docetaxel hazard is diverging from the nintedanib+docetaxel hazard, so the hazard ratio is increasing. The proportional hazards assumption is therefore not valid (Lee & Pirie, p value=0.0235). The ERG thus investigated alternative ways to compare nivolumab with nintedanib+docetaxel.

The ERG is of the opinion that baseline characteristics are fairly similar between the docetaxel arms of the CheckMate 057[28] and LUME-Lung 1[24] (adenocarcinoma population) trials. This means that, if there is sufficient evidence to suggest that the (comparator) docetaxel arms of the CheckMate 057[28] and LUME-Lung 1[24] trials are equivalent, the intervention arms of both trials (nivolumab and nintedanib+docetaxel) may be compared without adjustment. However, it is important to note that the ERG did not have access to data summarising the disease stage of patients in the adenocarcinoma population of the LUME-Lung 1 trial,[24] d so it is not possible to compare the two trials in this respect.

The ERG compared digitised K-M OS data for the adenocarcinoma population in the docetaxel+placebo arm of the LUME-Lung 1[24] trial with OS data from the docetaxel arm of the CheckMate 057 trial to investigate whether the OS outcomes for patients receiving docetaxel were significantly different in the two trials. The ERG concluded, by visual inspection of Figure 29 and by statistical test, that the docetaxel-treated populations from the two trials could be treated as equal. The ERG thus considers it is credible to compare

unadjusted nintedanib+docetaxel OS data from the LUME-Lung 1[24] trial with nivolumab OS data from the CheckMate 057 trial.



Source: Adapted from TA347[43]

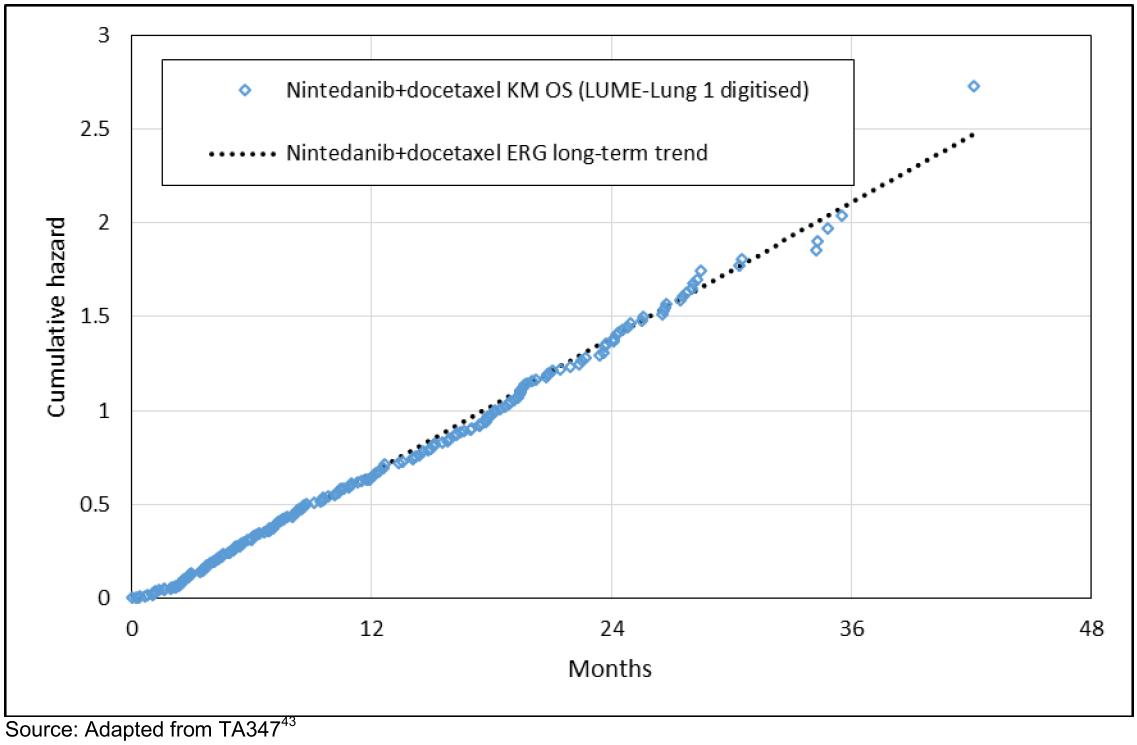
Figure 28 H-H plot of nintedanib+docetaxel OS vs. docetaxel OS (6 to 30) months

[redacted]

Figure 29 OS in docetaxel arms of CheckMate 057 and LUME-Lung 1 trials

Inspection of the cumulative hazard plot of the nintedanib+docetaxel OS K-M data from the LUME-Lung 1[24] trial (Figure 30) shows that a simple linear trend is established by 300 days

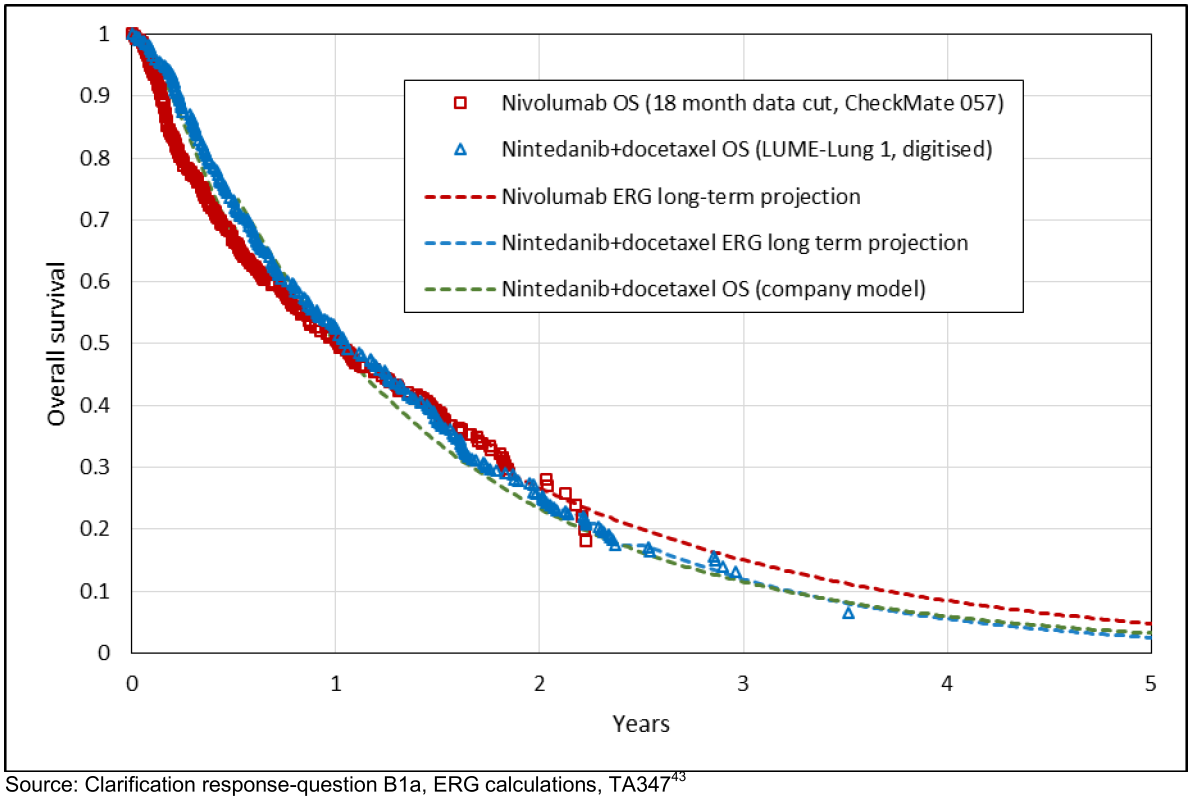
and continues indefinitely. This indicates that OS can be estimated by use of a simple exponential projective model beyond the limits of the K-M data (i.e. there is a constant hazard irrespective of time).



Source: Adapted from TA347[43]

Figure 30 Nintedanib OS from LUME-Lung 1 trial

A comparison of the nintedanib+docetaxel K-M OS data from LUME-Lung 1[24] trial and the K-M data for nivolumab from CheckMate 057 reveals very little difference in OS between the two treatments. When appropriate exponential long-term projections (see Section 5.5.5 for an explanation of the nivolumab OS projection) are applied to both arms (Figure 31), survival gain for nivolumab versus nintedanib+docetaxel is reduced in the company model by 70% to 3.1 months due to the reduction in nivolumab OS. The ERG and company projections for nintedanib+docetaxel OS result in very similar values. The estimated size of the ICER per QALY gained increases by £121,977 to £248,838 due to a substantial decrease in the incremental life years gained.



Source: Clarification response-question B1a, ERG calculations, TA347[43]

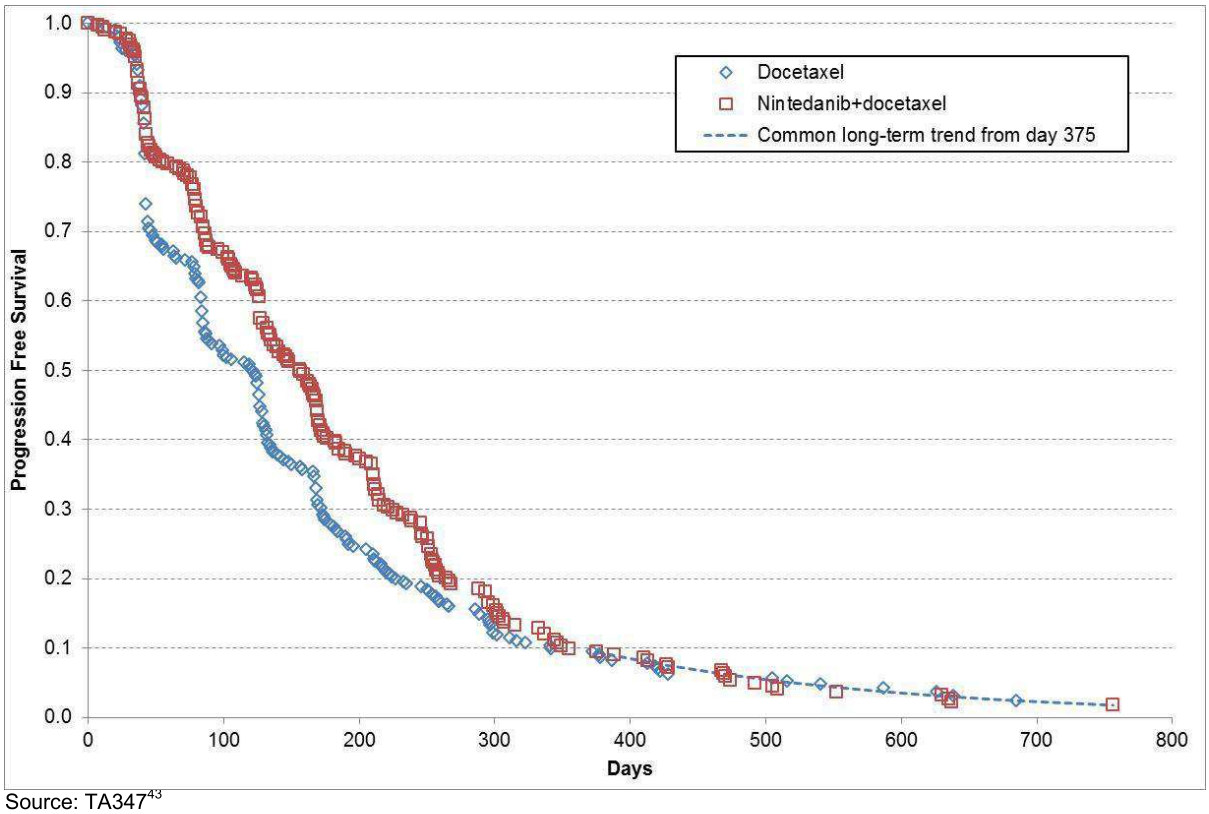
Figure 31 Nivolumab and nintedanib+docetaxel OS K-M data plus ERG projections

5.5.8 Progression-free survival: nivolumab versus nintedanib+docetaxel

When modelling PFS for patients treated with nintedanib+docetaxel, the company analysed digitised K-M data for the adenocarcinoma subgroup from the LUME-Lung 1[24] trial to compare outcomes in the nintedanib+docetaxel and docetaxel+placebo arms. The company concluded that PFS in the LUME-Lung 1[24] trial could also be described as a two-part proportional-hazard profile: equal in both arms to 2 months, then a HR of 0.98 from 2 months onwards. The company applied this two-part profile to the docetaxel TTD data from CheckMate 057 in order to model PFS for nintedanib+docetaxel. This approach again relies on the proportional hazards assumption holding independently in the two stages identified by the company.

The ERG report for TA347[43] contains more mature PFS data from the LUME-Lung 1[24] trial than have been included in the CS. Figure 32 shows that the early delay in progression for some patients receiving nintedanib+docetaxel (where the PFS curves begin to separate at around 6 weeks or 42 days) progressively dissipates over the course of a few months. The curves then converge at around 1 year, when the PFS experiences of patients in both arms of the LUME-Lung 1[24] trial are indistinguishable. Since the two arms are clearly separated in

the early part of the plot yet identical by 1 year, the HRs cannot be constant over time and the proportional hazards assumption - even from 2 months onwards - is invalidated.



Source: TA347[43]

Figure 32 PFS for the nintedanib+docetaxel and docetaxel arms (adenocarcinoma only) from LUME-Lung 1

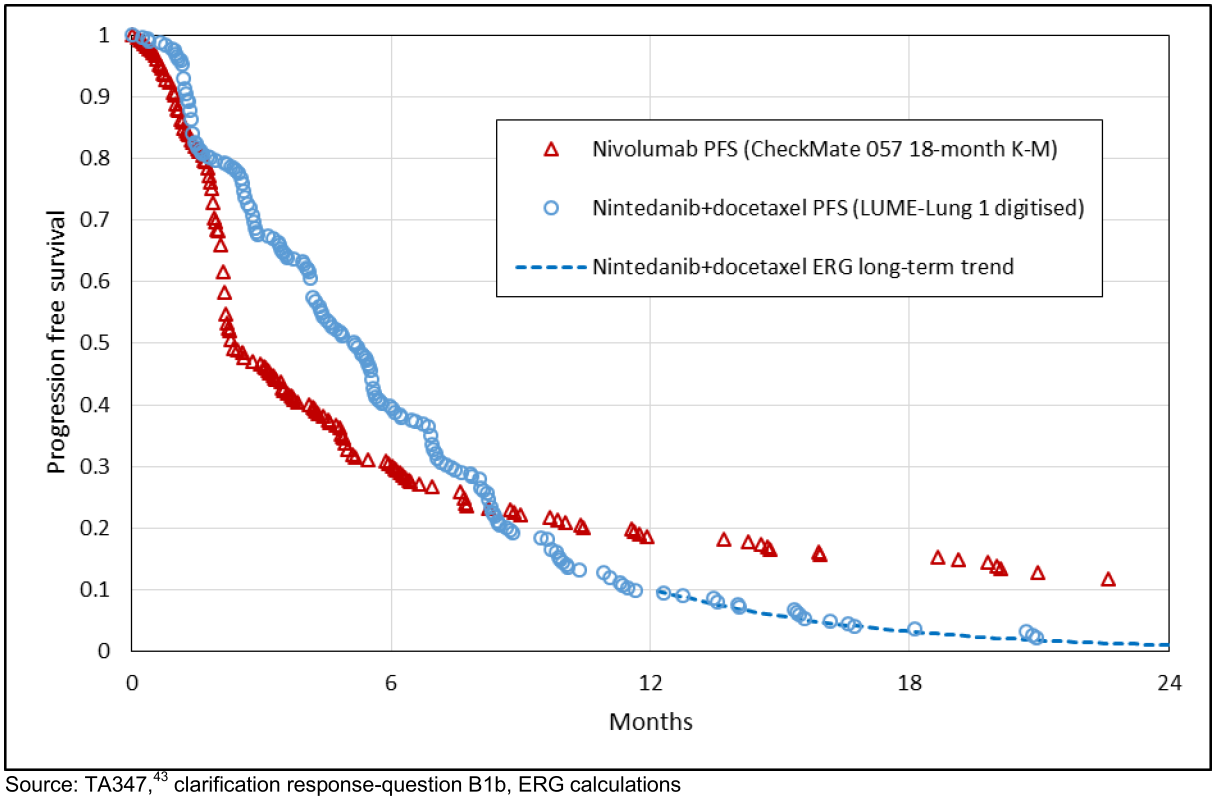
The ERG examined PFS data from the LUME-Lung 1[24] and CheckMate 057[28] trials to investigate whether PFS outcomes in the docetaxel arms were demonstrably different between the studies. Visual inspection of Figure 33 suggests that there is very little to separate the unadjusted K-M PFS data from the docetaxel arms of the two trials and this is confirmed by statistical testing. Since the PFS outcomes in the comparator arms from the two trials may be treated as equivalent, the ERG deems it is credible to compare the unadjusted nintedanib+docetaxel K-M PFS data from the LUME-Lung 1[24] trial with nivolumab K-M PFS data from CheckMate 057.

Figure 33 Docetaxel PFS in CheckMate 057 and LUME-Lung 1

The ERG used unadjusted nintedanib+docetaxel K-M PFS data from LUME-Lung 1 [TA347][43] to 375 days. Since the nintedanib+docetaxel and docetaxel arms of the LUME-Lung 1 trial[24] converge and the docetaxel arms of the CheckMate 057[28] and LUME-Lung 1[24] trials can be considered equivalent, the ERG applied the long-term trend from the docetaxel arm of CheckMate 057 to project PFS for nintedanib+docetaxel (Figure 34).

Using the ERG's preferred PFS models for nivolumab and nintedanib+docetaxel reduces nivolumab PFS gain over nintedanib+docetaxel in the company model by 42.3% to 5.8 months.

Amending the model to use the ERG's PFS estimates for nivolumab and for nintedanib+docetaxel reduces the ICER per QALY gained by £39,660 to £87,202 due to a proportionately greater reduction in treatment costs for nivolumab. Amending the model to use the ERG's TTD estimates for nivolumab, and the ERG's PFS projection for nintedanib+docetaxel reduces the size of the ICER per QALY gained versus nintedanib+docetaxel by £38,714 to £88,147, again due to a proportionately greater reduction in treatment costs for nivolumab.



Source: TA347,[43] clarification response-question B1b, ERG calculations

Figure 34 Nivolumab vs. nintedanib+docetaxel PFS

5.5.9 Calculating costs and QALYs using PFS and TTD

The company used TTD to represent time spent on treatment and time spent in PFS for all treatments compared in the company model. The ERG considers this approach to be flawed, as it can misrepresent the incremental costs and QALYs of treatments and their comparators.

Some costs that patients accrue are linked to the treatment they receive (i.e. treatment acquisition, treatment administration and treatment monitoring) and some are connected to their health state (i.e. disease-management costs). Some QALYs that patients accrue are linked to the treatment they receive (i.e. treatment-related AE disutilities) and some are connected to their health state (i.e. health state utilities). Therefore the ERG considers it preferable to use PFS and PPS when possible to estimate the costs and QALYs associated with health states and TTD to estimate costs and QALYs associated with treatment and has investigated the effect of using this approach on cost effectiveness estimates.

TTD data are also available for nintedanib+docetaxel from the LUME-Lung 1 trial [TA347];[43] however, it was not possible to use them in the company model. The company economic model does not have the facility to accommodate separate calculations for TTD for the two

elements of the nintedanib+docetaxel combination therapy, so the ERG used PFS as a proxy to model treatment-related costs for nintedanib+docetaxel. The projected costs of the nintedanib+docetaxel comparator in the ERG's revised ICER per QALY gained should thus be treated with caution.

Using the ERG's preferred models for PFS and TTD for nivolumab and docetaxel in the relevant parts of the economic model decreases the size of the ICER per QALY gained versus docetaxel by £19,996 to £83,594. Using the ERG's preferred models for PFS and TTD for nivolumab and for PFS for nintedanib+docetaxel reduces the size of the ICER per QALY gained versus nintedanib+docetaxel by £39,491 to £87,371.

5.5.10 Nivolumab dosing calculations

The company has used a log-normal distribution of body weight, weighted by sex, to calculate the average number of doses of nivolumab received by patients. Whilst the method is robust, the company has made a mistake when implementing the method in the model, which leads to a small overestimate in the average cost per dose of nivolumab.

The ERG has corrected the body weight calculation error, which decreases the average cost per full dose of nivolumab by £50.84 to £2,487.41. This amendment results in a £1,855 reduction in the size of the ICER per QALY gained versus docetaxel to £101,734 and a £2,738 reduction in the size of the ICER per QALY gained versus nintedanib+docetaxel to £124,123.

5.5.11 Treatment administration costs

The company model correctly estimates costs for the intervention and comparators based on the number of patients in treatment at the start of any given cycle. However, administration costs are calculated based on the number of patients in treatment mid-cycle. Given that treatment is acquired and administered at the same time, administration costs should also be calculated at the start of the cycle. The ERG has amended this oversight in the company model. Calculating administration costs at the beginning of the cycle reduces the size of the ICER per QALY gained versus docetaxel by £1,187 to £102,403 and increases the estimated ICER per QALY gained versus nintedanib+docetaxel by £26 to £126,887.

5.5.12 Health state utilities

The ERG identified several limitations with regard to completion rates and the health state utility estimates associated with the EQ-5D data collected during CheckMate 057. 73.9% of randomised patients completed the EQ-5D assessment at baseline. Despite 81.6% of participants being alive at week 12, the respective completion rates at this time point were 40.8% and 38.9% for nivolumab and docetaxel patients. The corresponding completion rates for nivolumab and docetaxel arms at 24 weeks were 26.9% and 14.8% while 69.4% of patients in CheckMate 057 were still alive.

It is likely that patients' decisions to continue completing the EQ-5D questionnaire throughout the trial period have been subject to various influences. Even if the possible self-selecting behaviour and response bias attributable to patients that completed the EQ-5D questionnaire were to be ignored, the ERG remains concerned that the characteristics of patients who completed the EQ-5D instrument are unlikely to match the characteristics of the initial trial population. Improvements in observed mean utility scores over time were observed (ERG report, Appendix 11.3). However, the ERG considers that the substantial differences deemed attributable to nivolumab treatment compared with docetaxel treatment cannot be considered reliable. The ERG considers it is likely that patients who continue to complete HRQoL assessments are those with better health status and higher ECOG PS scores than non-respondents. An important implication of this finding is that self-selection is likely to cause health state utility values to be overestimated. This phenomenon was previously observed in the NICE appraisal for nivolumab and squamous NSCLC patients.[49]

Health state utility values from CheckMate 057 indicate that patients in the PF health state have a mean utility score of 0.739 compared with patients in PD who have a mean utility score of 0.688. The ERG analysis of EQ-5D data by region provided utility estimates of 0.735 and 0.654 for the corresponding PF and PD states in European patients. Testing the effect of EQ-5D values obtained exclusively from European patients, as carried out by the ERG (Appendix 10.7), results in a slight increase in the overall ICERs per QALY gained when comparing nivolumab with both comparators.

The effects of using alternative utility values from (i) the study by Nafees et al[59] and (ii) a combination of EQ-5D values from CheckMate 057 with a Dutch lung cancer study by van den Hout et al[66] were investigated by the ERG.

Nafees et al[59] obtained utility values for health states describing second-line treatments for NSCLC from UK participants using a Standard Gamble (SG) approach. Values from the Nafees et al[59] study were previously used for patients treated with second-line chemotherapy in a systematic review and economic evaluation of first-line chemotherapy for NSCLC.[67] Substituting 0.65 for the PF state and 0.43 for the PD state into the company model reduced the incremental QALYs gained per patient for nivolumab versus docetaxel by 18%, and increased the size of the estimated ICER per QALY gained by £22,347 to £125,936. For nivolumab versus nintedanib+docetaxel, incremental QALYs gained per patient are reduced by 8% and the estimated ICER per QALY gained increases by £13,537 to £140,399.

In Holland, van den Hout et al[66] studied alternative palliative radiotherapy delivery models for patients with NSCLC. EQ-5D utility values were obtained using a UK valuation set over 52 weeks. The ERG has estimated patient utility in the PPS state to be 0.545 using stable data from this Dutch trial for both treatment arms. The ERG has also calculated an additional disutility associated with the terminal care phase. The company model structure does not capture terminal disutility therefore the ERG applied an adjustment to the stable PPS utility value to spread terminal disutility over the mean duration of PPS from CheckMate 057. This adjustment resulted in a utility value for PD of 0.476. Taking into account the increasing EQ-5D utility estimates over time in CheckMate 057, the ERG has selected early EQ-5D assessment results where participant responses were most stable i.e. limited analysis of overall EQ-5D values for the PF state during the first 12 weeks after randomisation for European patients alone. This method generated a utility value for the PF health state of 0.713. The implementation of these alternative utility values in the company model reduced the incremental QALYs gained per patient for nivolumab versus docetaxel by 10.2% and increased the estimated ICER per QALY gained by £11,853 to £115,443. For nivolumab versus nintedanib+docetaxel, the incremental QALYs gained per patient decreased by 1.4% and the estimated ICER per QALY gained increased by £2,055 to £128,916.

The utility values calculated from van den Hout et al[66] and CheckMate 057 are the ERG's preferred values, as they take into account terminal disutility, which is not otherwise accounted for in the company model. The alternative utility values considered by the ERG are outlined in Table 45.

Table 45 UK-specific mean EQ-5D values by source

Source	PF Mean	PD Mean
CheckMate 057 – all patients	0.739	0.688
CheckMate 057 – European patients	0.735	0.654
Nafees study[59]	0.65	0.43
CheckMate 057 & van den Hout study[66]	0.713	0.476

ICER=incremental cost effectiveness ratio; PD=progressed disease; PF=progression free
Source: Clarification response – Tables 4-7, Nafees 2009,[59] van den Hout 2006[66]

5.5.13 Adverse event utility decrements

In the company model, 17 AEs are selected to represent the effects of AEs on health-related utility. Disutility estimates associated with the selected AEs were derived from the following sources: the Nafees study[59] for fatigue, asthenia, anaemia, neutropenia, febrile neutropenia, leukopenia and diarrhoea, a study by Marti et al[68] for pneumonia, a study by Doyle, Lloyd and Walker[60] for dyspnoea and a previous Technology Appraisal[43] for white blood cell count and increased ALT. Utility values were unavailable for pain, pleural effusion, decreased neutrophil count, increased aspartate transaminase and hyponatraemia and the company therefore assumed that a disutility of 0 was associated with these AEs. The study by Marti et al[68] included a standard gamble exercise involving South and Central American parents of hospitalised children aged 3 to 36 months, and considered the disutility of a 7 day stay in hospital followed by recovery to full health. The relevance of utility values estimated in this study to elderly patients with metastatic lung cancer undergoing second-line chemotherapy is therefore questionable. The Doyle, Lloyd and Walker study[60] was less sophisticated than the study carried out by Nafees et al[59] and included only three symptoms and omitted PD. The ERG considers that the company's approach i.e. combining a single estimated parameter value from the Doyle, Lloyd and Walker model[60] with parameters from the study by Nafees et al[59] and from a previous appraisal,[43] is inconsistent and is therefore inappropriate.

The ERG is concerned that the estimated disutility effect of AEs in the model is necessarily understated to an unknown extent. The company applied utility decrements to AEs by multiplying the Grade 3 or 4 AE incidence rates (≥2%) of selected AEs from CheckMate 057 with the corresponding disutility values and summing them to a single disutility quantum for each treatment. Disutilities associated with AEs are applied only once during the first cycle of the company model. This technique assumes that patients suffering an AE only suffer a

single episode due to the use of incidence rates per person rather than use of event rates. Additionally, the ERG considers the assumption that all AEs and any consequent effects on patient health last no longer than 1 week in the model to lack validity. Moreover, the disutility of each AE was applied separately in the model thus introducing a potential for interaction or double counting when multiple AEs occur concurrently. The ERG is not able to assess the potential size of these problems due to lack of data, but this is not expected to have a substantial effect on the model results.

5.5.14 1- and 2- year stopping rules

The company has conducted sensitivity analyses in which medical dose caps are applied to treatment with nivolumab. These caps halt all treatment-related costs (i.e. acquisition, administration and monitoring) at either 1 or 2 years, but assume clinical efficacy across treatments and the costs and disutlilties associated with AEs remain equal to those experienced with uncapped treatments.

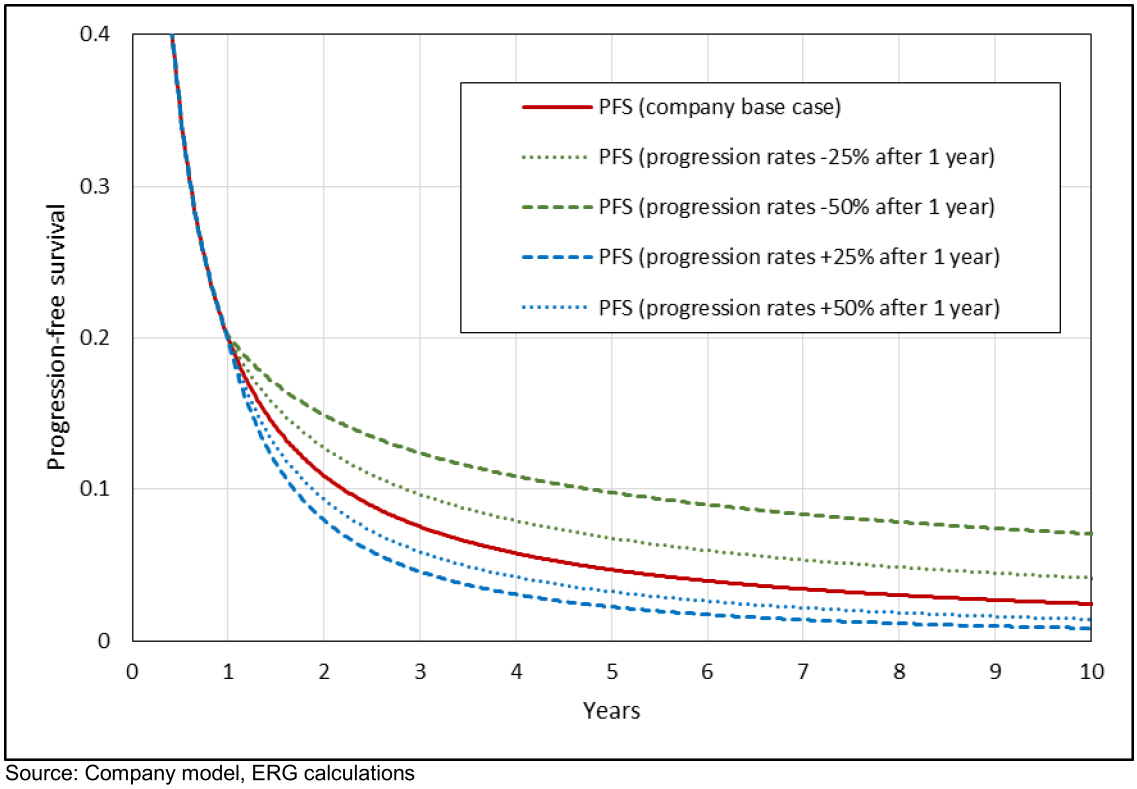
Evidence is currently lacking for the duration of benefit beyond an imposed treatment cap. The company references two trials whose results suggest that a durable benefit might be feasible for some patients who stop treatment with nivolumab before disease progression: the Phase Ib CheckMate 003[23] trial and the Phase IIIb/IV CheckMate 153[22] safety study.

CheckMate 003[23] implemented a 96-week (1.8 year) stopping rule and results from the September 2014 data analysis indicate that some patients who responded to treatment and were still in PFS at 96 weeks continued to respond to treatment. Data on post-treatment benefit in this study is limited to those patients who experienced either a CR or a PR. According to the CS (Figure 48), 13 patients with non-squamous NSCLC experienced a response in CheckMate 003[23] and one patient was still receiving treatment at the 96 week cut off. Three non-squamous patients (23%) had a response that lasted beyond 96 weeks.

Results from the CheckMate 003[23] trial indicate possible ongoing benefit for some patients who stop treatment before disease progression. However, there are insufficient data available from this trial to enable robust modelling of PFS or OS if treatment with nivolumab were to be capped at 1 or 2 years.

The CheckMate 153[22] trial randomised patients who were still in PFS at 1 year to either stop nivolumab or to continue nivolumab treatment until progression. Detailed survival data from CheckMate 153[22] trial were not available to the ERG at the time of writing the ERG report.

There is a lack of data describing the clinical effect of stopping treatment with nivolumab before progression. The ERG investigated the sensitivity of the estimated ICERs per QALY gained to 25% and 50% increases and decreases in nivolumab progression rates and mortality rates (which affect PFS and OS) after treatment is halted at 1 or 2 years. An example of the effect of varying progression rates on PFS is shown in Figure 35.



Source: Company model, ERG calculations

Figure 35 Nivolumab PFS with varying mortality rates after 1 year

The ERG explored scenarios where progression and mortality rates for nivolumab either increased or decreased after treatment was capped at 1 or 2 years. The ERG scenarios were included to investigate the potential effect on cost effectiveness of changes in clinical efficacy due to capped treatment with nivolumab. The ERG scenarios did not investigate the potential effect of changes to costs and disutilities associated with AEs as a result of capping treatment.

The company base case ICER per QALY gained with a 1-year medical dose cap is £46,860 for nivolumab versus docetaxel. This ICER per QALY gained ranges from £26,521 to £85,844 when varying progression and mortality rates were applied. The company base case ICER per QALY gained with a 2-year medical dose cap is £60,955 for nivolumab

versus docetaxel. This ICER per QALY gained ranges from £39,690 to £135,323 when varying progression and mortality rates were applied.

The company base case ICER per QALY gained with a 1-year medical dose cap is £43,122 for nivolumab versus nintedanib+docetaxel. This ICER per QALY gained ranges from £21,942 to £140,494 when varying progression and mortality rates were applied. The company base case ICER per QALY gained with a 2-year medical dose cap is £63,928 for nivolumab versus nintedanib+docetaxel. This ICER per QALY gained ranges from £36,366 to £198,460 when varying progression and mortality rates were applied.

The full results of the ERG's analysis of the sensitivity of the 1- and 2-year medical dose caps to changes in the progression and mortality rates for nivolumab patients are given in Appendix 10.8

5.6 Conclusions of the cost effectiveness section

The various changes implemented by the ERG for the comparison of nivolumab versus docetaxel and for nivolumab versus nintedanib+docetaxel yield a mixture of effects. When implemented individually, these revisions both increase and decrease the size of the ICERs per QALY gained. However, the combined effect of all of the ERG's preferred changes yields a revised base case ICER of £165,234 per QALY gained for nivolumab versus docetaxel and of £293,232 for nivolumab versus nintedanib+docetaxel.

The ERG considers that the company's base case results substantially underestimate the size of the most probable ICER per QALY gained for both nivolumab versus docetaxel and nivolumab versus nintedanib+docetaxel in previously treated patients with non-squamous NSCLC.

6 IMPACT ON THE ICER OF ADDITIONAL ERG ANALYSES

The ERG has made the following changes to the submitted company model to address the points raised in Section 5:

- use of ERG preferred OS estimates (R1)
- use of ERG preferred PFS estimates (R2)
- use of ERG preferred treatment duration estimates (based on TTD) for nivolumab and docetaxel (R3)
- application of ERG preferred PFS and TTD estimates to relevant cost and QALY categories for nivolumab and docetaxel (R4)

- ERG TTD for nivolumab and ERG PFS for nintedanib+docetaxel (R5)
- ERG PFS for nivolumab disease costs and QALYs, ERG TTD for nivolumab treatment costs and AEs; ERG PFS for nintedanib+docetaxel disease costs and QALYs (R6)
- nivolumab dosing calculations (R7)
- treatment administration costs at the start of each cycle (R8)
- use of preferred health state utility values (R9)
- use of health state utility values from study by Nafees[59] (R10).

Details of all Microsoft Excel revisions made by the ERG to the company's model are presented in Appendix 10.8 of this report.

6.1.1 Summary of ERG revisions to company model

The cost effectiveness results obtained by applying each of the ERG's model revisions for nivolumab versus docetaxel are shown in Table 46 and Table 47.

Revisions R2 and R3 (shaded rows) are superseded by R4 for nivolumab versus docetaxel, and revisions R2 and R5 (shaded rows) are superseded by R6 for nivolumab versus nintedanib+docetaxel.

The ERG's revised base case analysis (Scenario B) yields an ICER per QALY gained of £165,234 for nivolumab versus docetaxel, which is £61,644 per QALY gained higher than the company's original ICER. The ERG's revised base case for the comparison of nivolumab and docetaxel generates both incremental costs (-£22,109) and benefits (-0.41 QALYs) that are lower than those generated by the company.

The ERG's revised base case analysis (Scenario C) yields an ICER per QALY gained of £293,232 for nivolumab versus nintedanib+docetaxel, which is £166,370 per QALY gained higher than the company's original ICER. The ERG's revised base case for the comparison of nivolumab and docetaxel generates both incremental costs (-£27,482) and benefits (-0.37 QALYs) that are lower than those generated by the company.

Table 46 Cost effectiveness (nivolumab vs. docetaxel): ERG revisions to company base case

Model scenario ERG revision	Nivolumab			Docetaxel			Incremental			ICER	ICER
	Cost £	QALYs	Life years	Cost £	QALYs	Life years	Cost £	QALYs	Life years	£/QALY[+]	Change
A. Company base case	***93,306***	***1.424***	***2.243***	***17,854***	***0.696***	***1.095***	***+75,452***	***+0.728***	***+1.149***	***103,589***	**-**
R1) ERG OS	89,873	1.184	1.806	17,666	0.683	1.072	+72,207	+0.501	+0.734	143,984	**+40,395**
R2) ERG PFS	76,044	1.410	2.243	18,715	0.702	1.095	+57,328	+0.708	+1.149	80,940	**-22,649**
R3) ERG TTD	76,568	1.411	2.243	17,991	0.693	1.095	+58,577	0.719	+1.149	81,513	**-22,077**
R4) ERG PFS for disease costs and QALYs, ERG TTD for treatment costs and AEs	76,123	1.410	2.243	16,915	0.702	1.095	+59208	+0.708	+1.149	83,594	-19996
R7) Nivolumab dosing calculations	91,955	1.424	2.243	17,854	0.696	1.095	+74,100	+0.728	+1.149	101,734	**-1,855**
R8) Treatment administration costed at start of cycle	93,347	1.424	2.243	18,759	0.696	1.095	+74,587	+0.728	+1.149	102,403	**-1,187**
R9) ERG utility values (Van den Hout[66] + CheckMate 057)	93,306	1.186	2.243	17,854	0.532	1.095	+75,452	+0.654	+1.149	115,443	**+11,853**
R10) Utility values from study by Nafees[59]	93,306	1.076	2.243	17,854	0.477	1.095	+75,452	+0.599	+1.149	125,936	**+22,347**
B. ERG revised base case A+R1, R4, R7:R9	***70,124***	***0.870***	***1.806***	***16,781***	***0.547***	***1.072***	***+53,343***	***+0.323***	***+0.734***	***165,234***	***+61,644***

Costs and QALYs discounted; life years undiscounted
ERG=Evidence Review Group; HR=hazard ratio; OS=overall survival; PFS=progression-free survival; QALYs=quality adjusted life years; TTD=time to treatment discontinuation
R2 and R3 (shaded) are superseded by R4
* Interpret with caution due to interdependence of nivolumab ERG OS and company PFS projection in the company model
[+] Rounding errors account for difference between ICERs calculated using the incremental cost and QALY values given in the table and ICERs in this column

Table 47 Cost effectiveness results (nivolumab vs. nintedanib+docetaxel): ERG revisions to company base case

Model scenario ERG revision	Nivolumab			Nintedanib+docetaxel			Incremental			ICER	ICER
	Cost £	**QALYs**	**Life years**	**Cost £**	**QALYs**	**Life years**	**Cost £**	**QALYs**	**Life years**	**£/QALY[+]**	**Change**
A. Company base case	***93,306***	***1.424***	***2.243***	***30,708***	***0.931***	***1.440***	***+62,598***	***+0.493***	***+0.803***	***£126,861***	**-**
R1) ERG OS	89,873	1.184	1.806	30,709	0.946	1.457	+59,164	+0.238	+0.349	248,838	**+121,977**
R2) ERG PFS	76,044	1.410	2.243	34,974	0.939	1.440	+41,069	+0.471	+0.803	87,202	**-39,660**
R5) ERG TTD for nivolumab treatment costs and AEs, ERG PFS for nintedanib+docetaxel disease costs and QALYs	76,568	1.411	2.243	34,974	0.939	1.440	+41,593	+0.472	+0.803	88,147	**- 38,714**
R6) ERG PFS for nivolumab disease costs and QALYs, ERG TTD for nivolumab treatment costs and AEs; ERG PFS for nintedanib+docetaxel disease costs and QALYs	76,123	1.410	2.243	34,974	0.939	1.440	+41149	+0.471	+0.803	87,371	-39491
R7) Nivolumab dosing calculations	91,955	1.424	2.243	30,708	0.931	1.440	+61,247	+0.493	+0.803	124,123	**-2,738**
R8) Treatment administration costed at start of cycle	93,347	1.424	2.243	30,736	0.931	1.440	+62,611	+0.493	+0.803	126,887	**+26**
R9) ERG utility values (Van den Hout[66] + CheckMate 057)	93,306	1.186	2.243	30,708	0.700	1.440	+62,598	+0.486	+0.803	128,916	**+2,055**
R10) Utility values from Nafees[59]	93,306	1.076	2.243	30,708	0.630	1.440	+62,598	+0.446	+0.803	140,399	**+13,537**
C. ERG revised base case A+R1, R6:R9	***70,124***	***0.870***	***1.806***	***35,007***	***0.750***	***1.457***	***+35,116***	***+0.120***	***+0.349***	***£293,232***	***+166,370***

Costs and QALYs discounted; life years undiscounted
ERG=Evidence Review Group; HR=hazard ratio; OS=overall survival; PFS=progression-free survival; QALYs=quality adjusted life years
R2 and R5 (shaded) are superseded by R6
* Interpret with caution due to interdependence of nivolumab ERG OS and company PFS projection in the company model
[+] Rounding errors account for difference between ICERs calculated using the incremental cost and QALY values given in the table and ICERs in this column

7 END OF LIFE

The company makes a case that nivolumab fulfils the criteria set by NICE for end of life treatment. Namely:

- the life expectancy of the patient population was short (median life expectancy <24 months)
- the number of patients who would be eligible for the treatment is small
- the increase in mean OS is >3 months.

The details of the company's case for nivolumab meeting the NICE end of life criteria are outlined in Table 48.

Table 48 Company end of life criteria

Criterion	Data available
The treatment is indicated for patients with a short life expectancy, normally less than 24 months	Patients with advanced or metastatic NSCLC have a short life expectancy of less than 24 months.[10]
There is sufficient evidence to indicate that the treatment offers an extension to life, normally of at least an additional 3 months, compared with current NHS treatment	The proportional hazards assumption used in the published OS analysis has been shown not to hold (CS, Section 4.10). Therefore, mean OS from the cost effectiveness model may be considered more appropriate. This estimates mean OS, over the model time horizon of 20 years, to be 26.8 months for nivolumab compared with 13.09 months for docetaxel. This means that nivolumab is anticipated to extend life by greater than 3 months compared with docetaxel.
The treatment is licensed or otherwise indicated for small patient populations	The non-squamous NSCLC patient population potentially eligible for nivolumab treatment is expected to be very small (estimated 1413 patients). Nivolumab is also indicated for the treatment of advanced (unresectable or metastatic) squamous NSCLC in adults. The expected number of eligible patients for which nivolumab is being appraised in that submission is 853. Nivolumab is also indicated for the treatment of advanced (unresectable or metastatic) melanoma in adults. The expected number of eligible patients for which nivolumab is being appraised in that submission is 1304.

Source: CS, Table 40

Abbreviations: NHS = National Health Service; NSCLC = Non-Small Cell Lung Cancer

The ERG agrees that patients with advanced NSCLC have a short life expectancy of less than 24 months and that the total number of patients who would be eligible for the treatment is small. It also considers that nivolumab offers an extension to life of more than 3 months in comparison with docetaxel, but of just over 3 months in comparison with nintedanib+docetaxel; the ERG estimates a mean gain of 5.8 months for nivolumab versus docetaxel and a mean gain of 3.1 months for nivolumab versus nintedanib+docetaxel.

The ERG also agrees that nivolumab is licensed for a small patient population. Overall, the ERG considers that the combination of nivolumab meets NICE's end of life criteria.

8 OVERALL CONCLUSIONS

8.1 Strengths of the clinical and cost effectiveness evidence

- Checkmate 057 is a good quality trial providing direct evidence of effectiveness of nivolumab versus docetaxel in relation to OS and demonstrating an acceptable AE profile.

- The company provided a detailed submission that fulfilled the requirements of NICE's scope for the base case analysis. The ERG's requests for further clinical information were met to a good standard

- Variants of this model structure have been used in the modelling of similar treatments in a number of previous NICE STAs

- The decision model submitted by the company is generally implemented to a good standard.

8.2 Weaknesses and areas of uncertainty

- The validity of all assessed outcomes is limited by the fact that the proportional hazards assumption has been violated

- The comparison with all comparators in the original scope is limited by the available direct and indirect evidence.

Issues common to the modelling of nivolumab, docetaxel and nintedanib+docetaxel

- QALY calculations in the company model are linked to the time patients spend on treatment and not to their health state, which is incorrect

- The utility data used by the company lack credibility

- The model calculates treatment administration costs mid-treatment cycle when they should be applied at the start of the cycle, when treatment is received.

Issues specific to the modelling of nivolumab

- The method employed by the company to project nivolumab OS results in the model does not adequately represent the existing trial evidence from CheckMate 057

- The company's PFS model projects a small minority of patients treated with nivolumab to remain progression free throughout the lifetime of the model and to constitute 85% of those patients still alive after 20 years. It also predicts that any patient treated with nivolumab who is still in PFS by 18.4 years is cured of the disease and will never progress. The ERG considers both these outcomes to be implausible

- The company model creates an interdependence between OS and PFS projections that results in some values from the parametric OS model for nivolumab being replaced by PFS values to ensure that PFS is never greater than OS. This indicates that at least one of the parametric models (PFS or OS) used for nivolumab is inappropriate

- In the company model, one-third of the survival gain (nivolumab versus docetaxel) occurs post-progression, but this does not take into account the subgroup of nivolumab patients treated beyond progression who continue to accrue extra survival benefit, whether

due to extra treatment or other factors. ERG analysis suggests that post-progression survival constitutes 52% of survival gain when 25% of patients are treated beyond progression

- The nivolumab dosing calculations undertaken by the company are inaccurate.

Issues specific to the modelling of nintedanib+docetaxel

- The proportional hazards assumptions required to validate the company's indirect method of comparing nivolumab with nintedanib+docetaxel do not hold.

9 REFERENCES

1. Rossell R, Moran T, Queralt C, Porta R, Cardenal F, Camps C. Screening for epidermal growth factor receptor mutations in lung cancer. N Engl J Med. 2009; 361:958-67.
2. My Cancer Genome: Genetically Informed Cancer Medicine. EGFR in lung cancer. 2015 [cited 2016 February]; Available from: http://www.mycancergenome.org/content/disease/lung-cancer/egfr/313/.
3. My Cancer Genome: Genetically Informed Cancer Medicine. ALK in lung cancer. 2015 [cited 2016 February]; Available from: http://www.mycancergenome.org/content/disease/lung-cancer/alk/66/.
4. NICE. EGFR-TK mutation testing in adults with locally advanced or metastatic non-small-cell lung cancer: NICE diagnostics guidance 9. 2013; Available from: https://www.nice.org.uk/guidance/dg9.
5. Chen DS, Irving BA, Hodi FS. Molecular pathways: Next-generation immunotherapy--inhibiting programmed death-ligand 1 and programmed death-1. Clin Cancer Res. 2012; 18:6580-7.
6. Wang C, Thudium KB, Han M, Wang XT, Huang H, Feingersh D*, et al.* In vitro characterization of the anti-PD-1 antibody nivolumab, BMS-936558, and in vivo toxicology in non-human primates. Cancer Immunol Res. 2014; 2:846-56.
7. Brahmer JR, Drake CG, Wollner I, Powderly JD, Picus J, Sharfman WH*, et al.* Phase I study of single-agent anti-programmed death-1 (MDX-1106) in refractory solid tumors: safety, clinical activity, pharmacodynamics, and immunologic correlates. J Clin Oncol. 2010; 28:3167-75.
8. Freeman GJ, Long AJ, Iwai Y, Bourque K, Chernova T, Nishimura H*, et al.* Engagement of the PD-1 immunoinhibitory receptor by a novel B7 family member leads to negative regulation of lymphocyte activation. J Exp Med. 2000; 192:1027-34.
9. NICE. Ipilimumab for previously untreated unresectable malignant melanoma (TA319). 2014; Available from: https://www.nice.org.uk/guidance/ta319 (Accessed 15/05/2015).
10. Health and Social Care Information Centre. National Lung Cancer Audit Report: Report for the audit period 2013. 2014; Available from: http://www.hqip.org.uk/assets/NCAPOP-Library/NCAPOP-2014-15/HSCICNLCA-2014finalinteractivereport.pdf (Accessed 08/01/2015).
11. Powell HA, Tata LJ, Baldwin DR, Stanley RA, Khakwani A, Hubbard RB. Early mortality after surgical resection for lung cancer: An analysis of the English National Lung cancer audit. Thorax. 2013; 68:826-34.
12. NICE. Gefitinib for first-line treatment of locally advanced or metastatic non-small-cell lung cancer [TA192]; Costing template. NICE; 2010 [cited 2016 February]; Available from: https://www.nice.org.uk/guidance/ta192/resources.
13. Sculier JP, Moro-Sibilot D. First- and second-line therapy for advanced nonsmall cell lung cancer. Eur Respir J. 2009; 33:915-30.
14. National institute for Health and Care Excellence (NICE). Lung cancer (non-small-cell, anaplastic lymphoma kinase positive, previously treated) - ceritinib [ID729]. 2016 [updated 01/02/16]; Available from: http://www.nice.org.uk/guidance/indevelopment/gid-tag478 [Accessed February 2016].
15. Shepherd FA, Dancey J, Ramiau R, Mattson K, Gralla R, O'Rourke M*, et al.* A randomized trial of docetaxel versus best supportive care in patients with non-small-cell lung cancer previously treated with platinun-based chemotherapy. J Clin Oncol. 2000; 18:2095-103.

16. Greenhalgh J, Bagust A, Boland A, Dwan K, Beale S, Hockenhull J*, et al.* Erlotinib and gefitinib for treating non-small-cell lung cancer that has progressed following prior chemotherapy (review of NICE technology appraisals 162 and 175) (Protocol). Health Technol Assess Rep. 2013; 19:1-134.
17. Centre for Reviews and Dissemination (CRD). CRD's guidance for undertaking reviews in healthcare: Systematic Reviews (3rd Edition). York: CRD, University of York 2008.
18. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJM, Gavaghan DJ*, et al.* Assessing the quality of reports of randomized clinical trials: Is blinding necessary? Control Clin Trials. 1996; 17:1-12.
19. Cochrane Collaboration. The Cochrane Collaboration's tool for assessing risk of bias. 2011 [cited 2015]; Available from: http://ohg.cochrane.org/sites/ohg.cochrane.org/files/uploads/Risk%20of%20bias%20assessment%20tool.pdf.
20. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. J Epidemiol Community Health. 1998; 52:377-84.
21. Borghaei H, Paz-Ares L, Horn L, Spigel DR, Steins M, Ready N*, et al.* Nivolumab versus docetaxel in advanced nonsquamous non-small-cell lung cancer. N Engl J Med. 2015; 373:1627-39.
22. Hussein M, McCleod M, Chandler J, Blumenschein G, Schwartzberg L, Burris H*, et al.* Safety and efficacy of nivolumab in the ongoing trial of PD-L1/-patient population with metastatic non-small cell lung cancer (CheckMate 153). 16th World Conference on Lung Cancer; Denver, Colorado, USA2015.
23. Gettinger SN, Horn L, Gandhi L, Spigel DR, Antonia SJ, Rizvi NA*, et al.* Overall Survival and Long-Term Safety of Nivolumab (Anti-Programmed Death 1 Antibody, BMS-936558, ONO-4538) in Patients With Previously Treated Advanced Non-Small-Cell Lung Cancer. J Clin Oncol. 2015.
24. Reck M, Kaiser R, Mellemgaard A, Douillard J, Orlov S, Krzakowski M*, et al.* Docetaxel plus nintedanib versus docetaxel plus placebo in patients with previously treated non-small-cell lung cancer (LUME-Lung 1): A phase 3, double-blind, randomised controlled trial. Lancet Oncol. 2014; 15:143-55.
25. Thatcher N, Chang A, Parikh P, Rodrigues Pereira J, Ciuleanu T, von Pawel J*, et al.* Gefitinib plus best supportive care in previously treated patients with refractory advanced non-small-cell lung cancer: Results from a randomised, placebo-controlled, multicentre study (Iressa Survival Evaluation in Lung Cancer). Lancet. 2005; 366:1527-37.
26. Lee D, Park K, Kim J, Lee J, Shin S, Kang J*, et al.* Randomized phase III trial of gefitinib versus docetaxel in non-small cell lung cancer patients who have previously received platinum-based chemotherapy. Clin Cancer Res. 2010; 16:1307-14.
27. Maruyama R, Nishiwaki Y, Tamura T, Yamamoto N, Tsuboi M, Nakagawa K*, et al.* Phase III study, V-15-32, of gefitinib versus docetaxel in previously treated Japanese patients with non-small-cell lung cancer. J Clin Oncol. 2008; 26:4244-52.
28. Bristol-Meyers Squib. Final Clinical Study Report for Study CA209057: An open-label randomized phase iii trial of bms-936558 (nivolumab) versus docetaxel in previously treated metastatic nonsquamous non-small cell lung cancer (NSCLC). Bristol-Meyers Squib confidential support document for STA [ID900] 2015.
29. Statistical Analysis Plan for Clinical Study Report: An Open-Label Randomized Phase III Trial of Nivolumab versus Docetaxel in Previously Treated Metastatic Non-squamous Non-small cell Lung Cancer (NSCLC)(2014).
30. Bristol-Meyers Squib. Clinical Protocol CA209057: An Open-Label Randomized Phase III Trial of BMS-936558 (Nivolumab) versus Docetaxel in Previously Treated Metastatic Non-squamous Non-small cell Lung Cancer (NSCLC) Revised Protocol No.:04 Incorporates Amendment(s):08 2015.

31. Bassler D, Briel M, Montori VM, Lane M, Glasziou P, Zhou Q*, et al.* Stopping randomized trials early for benefit and estimation of treatment effects: systematic review and meta-regression analysis. JAMA. 2010; 303:1180-7.
32. Trotta F, Apolone G, Garattini S, Tafuri G. Stopping a trial early in oncology: for patients or for industry? Ann Oncol. 2008; 19:1347-53.
33. Pocock SJ. When to stop a clinical trial. BMJ. 1992; 305:235-40.
34. Ioannidis JPA, Karassa FB. The need to consider the wider agenda in systematic reviews and meta-analyses: breadth, timing, and depth of the evidence. BMJ. 2010; 341.
35. Grambsch PM, Therneau TM. Proportional hazards tests and diagnostics based on weighted residuals. Biometrika. 1994; 81:515-26.
36. Garassino MC, Martelli O, Broggini M, Farina G, Veronese S, Rulli E*, et al.* Erlotinib versus docetaxel as second-line treatment of patients with advanced non-small-cell lung cancer and wild-type EGFR tumours (TAILOR): a randomised controlled trial. Lancet Oncol. 2013; 14:981-8.
37. Brahmer JR, Reckamp KL, Baas P, Crinò L, Eberhardt W, Poddubskaya E*, et al.* Nivolumab versus docetaxel in advanced squamous-cell non-small cell lung cancer. N Engl J Med. 2015; 373.
38. Horn L, Brahmer JR, Reck M, Borghaei H, Spigel DR, Steins M*, et al.* Phase 3, randomized trial (CheckMate 057) of nivolumab vs docetaxel in advanced non-squamous non-small cell lung cancer: Subgroup analysis of patient-reported outcomes. 18 ECCO-40th ESMO European Cancer Congress; September 25029 2015; Viena2015.
39. Bristol-Meyers Squib. Summary of product characteristics: Opdivo 10mg/ML. Bristol-Meyers Squib confidential support document for STA [ID900] 2015.
40. Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. J Clin Epidemiol. 1997; 50:683-91.
41. Guyot P, Ades A, Ouwens MJ, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. BMC Med Res Methodol. 2012; 12:9.
42. Royston P, Parmar MK. Restricted mean survival time: an alternative to the hazard ratio for the design and analysis of randomized trials with a time-to-event outcome. BMC Med Res Methodol. 2013; 13:152.
43. NICE. Nintedanib for previously treated locally advanced, metastatic, or locally recurrent non‑small‑cell lung cancer (TA347). 2015; Available from: https://www.nice.org.uk/guidance/indevelopment/gid-tag449 (Accessed 09/10/2015).
44. NICE. Crizotinib for previously treated non-small-cell lung cancer associated with an anaplastic lymphoma kinase fusion gene (TA296). 2013; Available from: https://www.nice.org.uk/guidance/ta296 (Accessed 09/10/2015).
45. NICE. Erlotinib for the treatment of non-small-cell lung cancer (TA162). 2012; Available from: https://www.nice.org.uk/guidance/ta162 (Accessed 09/10/2015).
46. NICE. Erlotinib and gefitinib for treating non-small-cell lung cancer that has progressed following prior chemotherapy (Review of TA162 and TA175): appraisal consultation document. 2014; Available from: http://www.nice.org.uk/guidance/gid-tag347/resources/erlotinib-and-gefitinib-for-treating-nonsmallcell-lung-cancer-that-has-progressed-following-prior-chemotherapy-review-of-ta162-and-ta175-appraisal-consultation-document (Accessed 10/09/2015).
47. Holmes J, Dunlop D, Hemmett L, Sharplin P, Bose U. A cost-effectiveness analysis of docetaxel in the second-line treatment of non-small cell lung cancer. Pharmacoeconomics. 2004; 22:581-9.
48. Lewis G, Peake M, Aultman R, Gyldmark M, Morlotti L, Creeden J*, et al.* Cost-effectiveness of erlotinib versus docetaxel for second-line treatment of advanced non-small-cell lung cancer in the United Kingdom. J Int Med Res. 2010; 38:9-21.

49. NICE. Lung cancer (non-small-cell, squamous, metastatic) - nivolumab (after chemotherapy) [ID811]. 2015; Available from: https://www.nice.org.uk/guidance/indevelopment/gid-tag506/documents (Accessed 12/10/2015).
50. NICE. Erlotinib for the first-line treatment of locally advanced or metastatic EGFR-TK mutation-positive non-small-cell lung cancer (TA258). 2012; Available from: https://www.nice.org.uk/guidance/ta258 (Accessed 09/10/2015).
51. NICE. Bevacizumab in combination with oxaliplatin and either fluorouracil plus folinic acid or capecitabine for the treatment of metastatic colorectal cancer (TA212). 2010; Available from: https://www.nice.org.uk/guidance/ta212 (Accessed 21/10/2015).
52. Bristol-Meyers Squib. Clinical Protocol CA209057: An Open-Label Randomized Phase III Trial of BMS-936558 (Nivolumab) versus Docetaxel in Previously Treated Metastatic Non-squamous Non-small cell Lung Cancer (NSCLC) Revised Protocol No.:02 Incorporates Amendment(s):06. Bristol-Meyers Squib confidential support document for STA [ID900] 2015.
53. Bristol-Myers Squibb. Nivolumab: Non-interventional Study Report for Study CA209-116; Treatment Patterns, Outcomes and Resource Use Study for Advanced Stage Non-Small Cell Lung Cancer (Squamous and Non-Squamous) in Europe (DCN: 2015-03-06). Report date 15 April 2015. Data on file. 2015b.
54. Bristol-Meyers Squib. CA209-003: Overall Survival Update on advanced non-small cell lung cancer (NSCLC) patients from study CA209-003. Bristol-Meyers Squib confidential support document for STA [ID900 2015.
55. Latimer N. NICE DSU technical support document 14: Survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data. University of Sheffield: Decison Support Unit, School of Health and Related Research 2011.
56. Bristol-Myers Squibb. Nivolumab: Final Clinical Study Report for Study CA209017; An Open-label Randomized Phase III Trial of BMS-936558 (Nivolumab) versus Docetaxel in Previously Treated Advanced or Metastatic Squamous Cell Non-small Cell Lung Cancer (NSCLC) (DCN:930086504); Report dated 26 Feburary 2015. Data on file. 2015a.
57. Dolan P. Modeling valuations for EuroQol health states. Med Care. 1997; 35:13.
58. Anokye NK TP, Green C, Pavey TG, Taylor RS. Physical Activity and Health Related Quality of Life. BMC Public Health. 2012; 12:108.
59. Nafees B SM, Gavriel S, Bhalla S, Watkins J. Health State Utilities for Non Small Cell Lung Cancer. Health Qual Life Outcomes. 2008; 6.
60. Doyle S LA, Walker M. Health state utility scores in advanced non-small cell lung cancer. Lung Cancer. 2008; 62:374-80.
61. Marti SG CL, Bardach A, Galante J, Lopez A, et al. A Cost-Effectiveness Analysis of a 10-Valent Pneumococcal Conjugate Vaccine in Children in Six Latin American Countries. Cost Eff Resour Alloc. 2013; 11:1-17.
62. Department of Health. Electronic Market Information Tool. Commercial Medicines Unit; 2014; Available from: http://cmu.dh.gov.uk/electronic-market-information-tool-emit (Accessed 20/10/2015).
63. McKee S. NICE U-turn on Boehringer's Vargatef for lung cancer. PharmaTimes Digital; 2015 [cited 2016]; Available from: http://www.pharmatimes.com/article/15-06-02/NICE_U-turn_on_Boehringer_s_Vargatef_for_lung_cancer.aspx (Accessed 07/03/2016).
64. Department of Health. Department of Health reference costs 2013-14. London: UK Department of Health 2014.
65. Health Do. National Schedule of reference costs 2013/2014. 2015; Available from: https://www.gov.uk/government/publications/nhs-reference-costs-2013-to-2014 (Accessed 09/10/2015).

66. van den Hout W. Cost–utility analysis of short- versus long-course palliative radiotherapy in patients with non-small-cell lung cancer. JNCI. 2016; 98:1786-94.

67. Brown T, Pilkington G, Bagust A, Boland A, Oyee J, Tudur-Smith C*, et al.* Clinical effectiveness and cost-effectiveness of first-line chemotherapy for adult patients with locally advanced or metastatic non-small cell lung cancer: a systematic review and economic evaluation. Health Technol Assess. 2013; 17:1-278.

68. Marti SG, Colantonio L, Bardach A, Galante J, Lopez A, Caporale J*, et al.* A cost-effectiveness analysis of a 10-valent pneumococcal conjugate vaccine in children in six Latin American countries. Cost effectiveness and resource allocation : C/E. 2013; 11:21.

10 APPENDICES

10.1 CheckMate 057 eligibility criteria

Table 49 Inclusion/exclusion criteria for entry into CheckMate 057

	CheckMate 057 eligibility criteria
Inclusion criteria	• Patients with histologically or cytologically documented locally advanced non-squamous NSCLC who presented with stage IIIb/ stage IV or recurrent or progressive disease following multimodal therapy (radiation therapy, surgical resection or definitive chemoradiation therapy for locally advanced disease) • ECOG PS ≤ 1 • Patients must have had measurable disease by computed tomography or magnetic resonance imaging per RECIST 1.1 criteria; radiographic tumour assessment was performed within 28 days of randomisation. • Target lesions may have been located in a previously irradiated field if there was documented (radiographic) disease progression in that site. • Patients who received study therapy after acceptable prior therapy as specified below: • Patients who received study therapy as second-line of treatment: • Patients must have experienced disease recurrence or progression during or after one prior platinum doublet-based chemotherapy regimen for advanced or metastatic disease. • First-line therapy was defined as therapy used to treat advanced disease. Each subsequent line of therapy was preceded by disease progression. A switch of an agent within a regimen in order to manage toxicity did not define the start of a new line of therapy. Patients must have received at least two cycles of platinum doublet-based chemotherapy before discontinuation for toxicity. Experimental therapies, when given as separate regimen, were considered as separate line of therapy. Maintenance therapy following platinum doublet-based chemotherapy was not considered as a separate regimen of therapy and could include continuation of one or more of the agents used in the first-line therapy regimen or switch to another non–cross-resistant agent. The initiation of maintenance therapy required the lack of progressive disease with front-line therapy. Treatment given for locally advanced disease was not considered as a line of therapy for advanced disease. Patients with recurrent disease > 6 months after platinum-containing adjuvant, neoadjuvant or definitive chemo-radiation therapy given for locally advanced disease, who also subsequently progressed during or after a platinum doublet-based regimen given to treat the recurrence, were eligible. • Patients who received platinum-containing adjuvant, neoadjuvant or definitive chemo-radiation therapy given for locally advanced disease and who developed recurrent (local or metastatic) disease within 6 months of completing therapy were eligible. Adjuvant or neoadjuvant platinum-based chemotherapy (after surgery and/or radiation therapy) followed by recurrent or metastatic disease within 6 months of completing therapy was considered as first-line therapy for advanced disease. • Patients who received study therapy as third-line of treatment must have experienced disease recurrence or progression during or after a separate EGFR or ALK TKI regimen in addition to one prior platinum doublet-based chemotherapy regimen (regardless of order of administration). • Patients who received an EGFR-TKI (erlotinib, gefitinib or experimental) in addition to a platinum doublet-based chemotherapy must have had a tumour with a known activating EGFR mutation. Patients with a tumour with EGFR mutation-negative/unknown status who received an EGFR-TKI after failure of a prior platinum doublet-based chemotherapy were excluded. • Patients who received an ALK inhibitor (crizotinib or experimental) in addition to a platinum doublet-based chemotherapy must have had a tumour with a known ALK translocation.

	• A formalin-fixed, paraffin-embedded tumour tissue block or unstained slides of tumour sample (archival or recent) must have been available for biomarker evaluation. Specimens must have been received by the central laboratory prior to randomisation. Biopsy should have been excisional, incisional or core needle. Fine needle aspiration was insufficient.
Exclusion criteria	• Patients with untreated CNS metastases were to be excluded. Patients were eligible if CNS metastases had been treated and the patient had neurologically returned to baseline (except for residual signs or symptoms related to the CNS treatment) for at least 2 weeks prior to enrolment. In addition, patients must have been either off corticosteroids or on a stable or decreasing dose of ≤ 10 mg daily prednisone (or equivalent). • Patients with carcinomatous meningitis. • Any serious or uncontrolled medical disorder or active infection with hepatitis or human immunodeficiency virus that may have been reactivated. • Other active malignancy requiring concurrent intervention. • Patients with previous malignancies (except non-melanoma skin cancers and the following in situ cancers: bladder, gastric, colon, cervical/dysplasia, endometrial, melanoma or breast) were excluded unless a complete remission was achieved at least 2 years prior to study entry AND no additional therapy was required or anticipated to be required during the study period. • Patients with a condition requiring systemic treatment with either corticosteroids (> 10 mg daily prednisone equivalent) or other immunosuppressive medications within 14 days of randomisation. Corticosteroids with minimal systemic absorption (inhaled or topical steroids), and adrenal replacement steroid doses > 10 mg daily prednisone equivalent, were permitted in the absence of active autoimmune disease. • Patients with active, known or suspected autoimmune disease. Patients with type I diabetes mellitus, hypothyroidism only requiring hormone replacement, skin disorders (such as vitiligo, psoriasis or alopecia) not requiring systemic treatment or conditions not expected to recur in the absence of an external trigger were permitted to enrol. • All toxicities attributed to prior anti-cancer therapy other than alopecia and fatigue must have resolved to Grade 1 (NCI CTCAE version 4) or baseline before administration of study drug. • Prior therapy with anti-tumour vaccines or other immuno-stimulatory anti-tumour agents. • Prior therapy with anti-PD-1, anti-PD-L1, anti-PD-L2, anti-CD137 or anti-CTLA-4 antibody (including ipilimumab or any other antibody or drug specifically targeting T-cell co-stimulation or checkpoint pathways). • Prior treatment with docetaxel. • Patients with interstitial lung disease that was symptomatic or may interfere with the detection or management of suspected treatment-related pulmonary toxicity. • Patients were to have recovered from the effects of major surgery or significant traumatic injury at least 14 days before the first dose of study treatment.

ECOG PS=Eastern Cooperative Oncology Group performance status; RECIST 1.1=Response Evaluation in Solid Tumours; EGFR=epidermal growth factor receptor; ALK=anaplastic lymphoma kinase; TKI=tyrosine kinase inhibitor; CNS=central nervous system; NCI CTCAE= National Cancer Institute Cancer Therapy Evaluation Program; PD-1= programmed death-1; PD-L1= programmed death-ligand 1; PD-L2= programmed death-ligand 2

10.2 PH assumption testing of CheckMate 057

- **Overall survival**

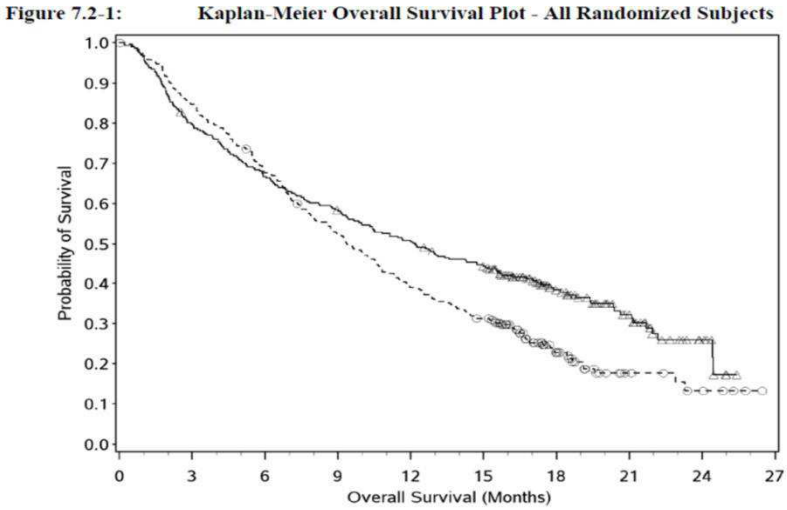


Figure 36 K-M curve of OS

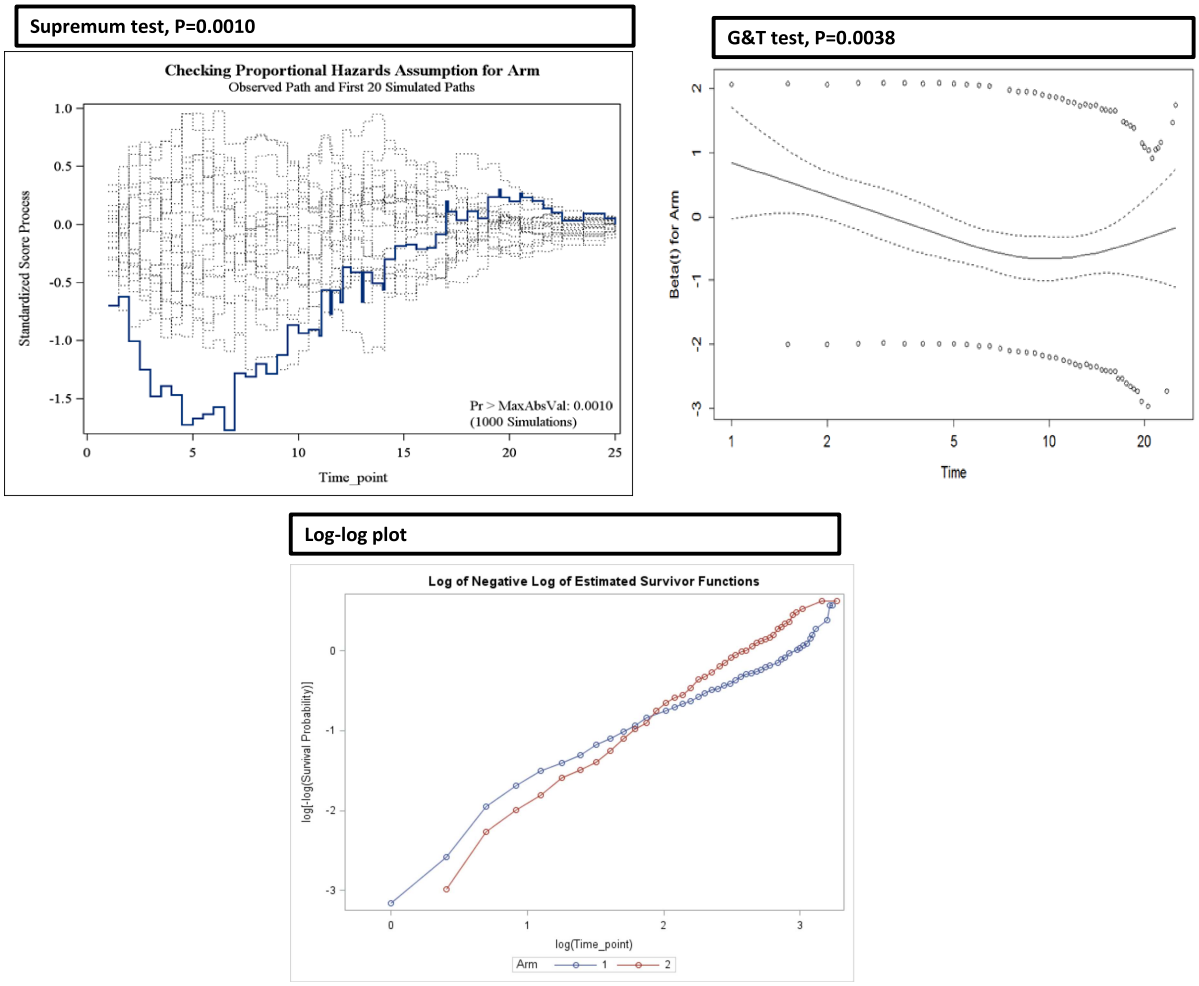


Figure 37 PH assumption tests-OS

- **PFS survival**

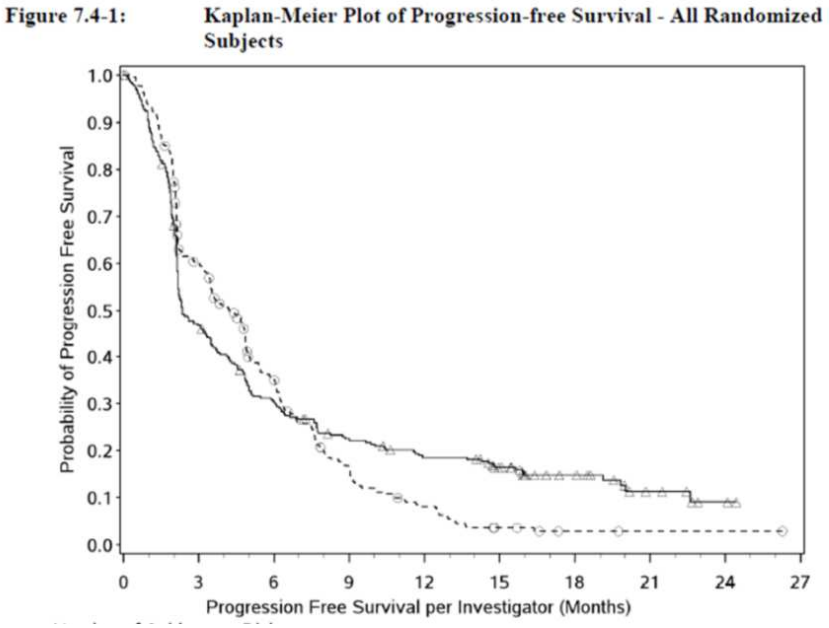


Figure 38 K-M Curve of PFS

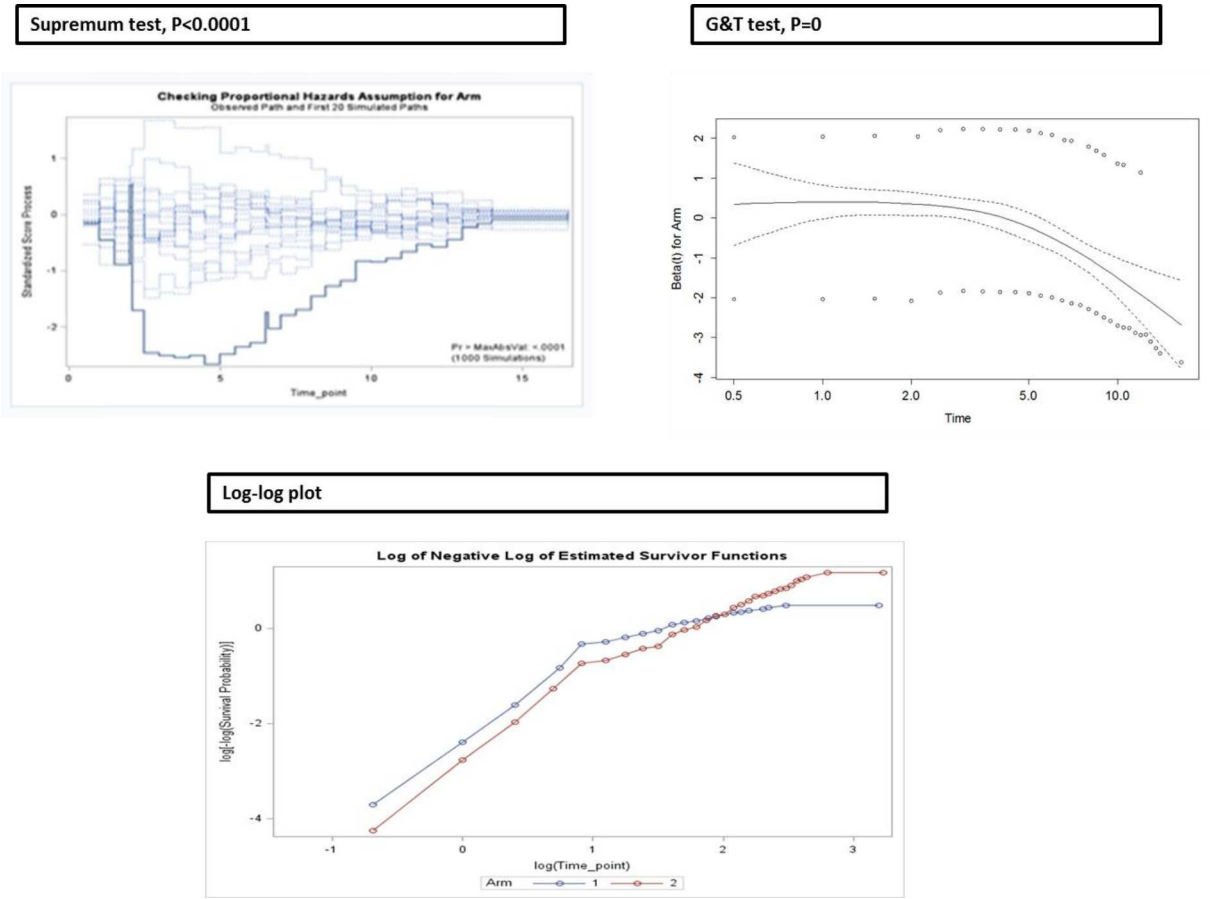


Figure 39 PH assumption tests - PFS

10.3 Subgroup analyses on 18 month data

Table 50 CheckMate 057: Treatment effect on OS in predefined subsets: 18-month data cut

Subset	N	HR, nivolumab vs. docetaxel (95% CI)	Test for interaction P value
Prior use of maintenance therapy			0.7577
Yes	233	0.78 (0.58 to 1.06)	
No	349	0.71 (0.56 to 0.91)	
Line of therapy			0.0431
Second-line	514	0.68 (0.56 to 0.83)	
Third-line/other	68	1.29 (0.74 to 2.25)	
Region			0.0006
US/Canada	215	0.54 (0.39 to 0.74)	
Europe	269	0.74 (0.57 to 0.98)	
Rest of World	98	1.54 (0.96 to 2.48)	
Age categorisation (years)			0.9960
<65	339	0.77 (0.60 to 0.99)	
≥65 and <75	200	0.68 (0.49 to 0.93)	
≥75	43	0.76 (0.37 to 1.56)	
Sex			0.3484
Male	319	0.69 (0.53 to 0.89)	
Female	263	0.82 (0.62 to 1.08)	
Race			
White	533	0.72 (0.59 to 0.88)	
Baseline ECOG PS			0.5236
0	179	0.63 (0.44 to 0.90)	
≥1	402	0.78 (0.62 to 0.97)	
Smoking status			0.0446
Yes	458	0.66 (0.54 to 0.82)	
Other	124	1.08 (0.70 to 1.65)	
EGFR mutation status			0.4689
Positive	82	1.12 (0.67 to 1.86)	
Not detected	342	0.64 (0.50 to 0.82)	
Not reported	158	0.76 (0.53 to 1.09)	
ALK translocation status			0.2970
Positive	13	0.50 (0.12 to 2.04)	
Not detected	254	0.71 (0.53 to 0.94)	
Not reported	315	0.79 (0.61 to 1.02)	
KRAS mutation status			0.9695
Positive	28	0.57 (0.32 to 1.02)	
Not detected	60	0.96 (0.65 to 1.43)	
Not reported	204	0.71 (0.56 to 0.89)	
MET receptor status			
Not reported	566	0.72 (0.59 to 0.87)	
Cell type			0.2536
Adenocarcinoma	541	0.76 (0.63 to 0.93)	
Other	41	0.51 (0.25 to 1.02)	

Subset	N	HR, nivolumab vs. docetaxel (95% CI)	Test for interaction P value
Time from diagnosis to randomisation			0.6366
<1 year	350	0.78 (0.62 to 0.99)	
Other	232	0.68 (0.50 to 0.92)	
Time from completion of most recent regimen to randomisation			0.2018
<3 months	364	0.82 (0.65 to 1.04)	
3-6 months	114	0.76 (0.49 to 1.16)	
>6 months	103	0.46 (0.28 to 0.76)	
Prior neo-adjuvant			0.8844
Yes	19	0.76 (0.26 to 2.21)	
No	563	0.74 (0.61 to 0.89)	
Prior adjuvant			0.5121
Yes	42	0.92 (0.45 to 1.87)	
No	540	0.73 (0.60 to 0.89)	
CNS metastases			0.3246
Yes	69	0.98 (0.59 to 1.65)	
No	513	0.71 (0.58 to 0.87)	

ALK=anaplastic lymphoma kinase; CI=confidence interval; CNS=central nervous system; ECOG PS=Eastern Cooperative Oncology Group performance status; EGFR=epidermal growth factor receptor; HR=hazard ratio; KRAS=kirsten rat sarcoma 2 viral oncogene homolog; MET=mesenchymal epithelial transition; OS=overall survival; US=United States
Source: Company response to the ERG clarification letter, Table 2

10.4 Quality assessment of RCTs included in the ITC

Table 51 Summary of quality assessment of RCTs included in the analysis

Study ID	Jadad[1] [8] score	AC grade	Was randomisation carried out appropriately?	Were the groups similar at the outset of the study in terms of prognostic factors?	Were the care providers, participants and outcome assessors blind to treatment allocation?	Were there any unexpected imbalances in drop-outs between groups?	Is there any evidence to suggest that the authors measured more outcomes than they reported?	Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?
LUME-Lung 1[24]	4	A	Low risk; Randomisation was carried out appropriately as treatment was assigned by IVRS or IWRS. Allocation concealment was adequate.	Low risk; Demographics and baseline characteristics were well balanced between the two treatment	Low risk; This was a double-blind study. Patients and investigators were masked to assignment, and none of the individuals directly involved in the conduct and analysis of the study had access to treatment allocation before the final database lock	Low risk; The withdrawals, completers, and the specific reasons for withdrawal were adequately reported at data cut-off	Low risk; Author has measured all the outcomes that have been reported in published protocol land in clinical study registry.	Low risk; The efficacy and safety analysis was done using ITT/mITT population
ISTANA[26]	1	B	Not clear; This was a randomised study but the method of randomisation was not reported. Allocation concealment was also not reported	Low risk; The treatment groups were well balanced for baseline characteristics, with the exception of slightly fewer females (33% versus 43%) and never-smokers (37% versus 46%) in the gefitinib treatment group than in the docetaxel group.	High risk: This was an open-label study	High Risk: Withdrawals were not reported.	Low risk; The author analysed all the primary outcomes in this final analysis as described in the protocol and the clinical study registry.	Low risk; ITT and mITT population was analysed for efficacy and safety outcomes
ISEL[25]	4	A	Not clear; This was a	Low risk; The baseline	Low risk; This was a	Low risk;	Not clear; There	Low risk; The safety

Study ID	Jadad[1] [8] score	AC grade	Was randomisation carried out appropriately?	Were the groups similar at the outset of the study in terms of prognostic factors?	Were the care providers, participants and outcome assessors blind to treatment allocation?	Were there any unexpected imbalances in drop-outs between groups?	Is there any evidence to suggest that the authors measured more outcomes than they reported?	Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?
			randomised study and the randomisation was carried out by minimisation method.	characteristics of the two treatment arms were well balanced.	double-blind study and the details of blinding were reported. Physically identical tablets and packaging, assigned by the central registration and randomisation centre were used to ensure masking of both patients and investigators.	Withdrawals and reasons for withdrawals were reported	was no evidence to conclude whether all outcomes assessed were reported or not	and efficacy analysis was done using ITT/mITT population
CheckMate 057[28]	3	A	Low risk; the patients were randomised to the two active treatments using IVRS; Allocation concealment was adequate.	Low risk; the baseline characters in the two treatment arms were well balanced.	High risk; the study was conducted as an open-label study.	Low risk; study withdrawals were adequately reported.	Low risk; the authors measured all outcomes as reported in the protocol.	Low risk; the efficacy and safety analysis were performed using ITT and mITT analysis respectively.
V-15-32[27]	2	B	Not clear; This was a randomised study but the method of randomisation was not reported	Low risk; Treatment groups were generally well balanced for baseline demographics except for some small imbalances in smoking history (7% fewer never-smokers and 10% more ex-smokers in the gefitinib arm)	High risk; This was an open-label study	Low risk; The withdrawals and the specific reasons for withdrawal were reported	Low risk; There was a published protocol that describes that the author has measured all the outcomes that have been reported	Low risk; ITT population was considered for both safety and efficacy analysis

AC=allocation concealment; ITT=Intent-to-Treat; IVRS=interactive voice response system; IWRS=interactive web-based response system; mITT=modified Intent-to-Treat
Note: The Jadad Score is used to assess quality of RCTs, allocating them a score between 0 (very poor) and 5 (rigorous) .
Source: CS, adapted from Table 25

10.5 Outcomes data used in indirect treatment comparisons

Table 52 Summary of data from studies reporting data for pre-treated non-squamous NSCLC population and included in analysis

Study ID (acronym)	Treatment (N)	ORR, n (%)	OS (HR) (95% CI)	OS (RMST; Mean (SE; 95% CI) (months)	PFS HR (95% CI)	PFS (RMST; months)	Any adverse event	Any Grade 3/4 adverse event
Studies connected in BOTH networks (EGFR mutation-negative/unknown AND all-comers NSQ NSCLC)								
LUME-Lung 1[24]	Docetaxel (659) Results presented for adenocarcinoma subgroup (n=336)	12 (3.6%)	0.83 (0.70 to 0.99)	For all-comers at tau=13 months: 9.313 (0.228; 8.866 to 9.76). For EGFR mutation-negative/unknown at tau=28 months: 13.213 (0.512; 12.211 to 14.216).	0.77 (0.62 to 0.96)	For all-comers at tau=12 months: 4.173 (0.241; 3.70 to 4.645)	314/333 (94%)	228/333 (68%)
	Docetaxel+nintedanib (655) Results presented for adenocarcinoma subgroup (n=322)	15 (4.7%)		For all-comers at tau=13 months: 9.726 (0.241; 9.253 to 10.2). For EGFR mutation-negative/unknown at tau=28 months: 14.767 (0. 565; 13.659 to 15.874).		For all-comers at tau=12 months: 4.826 (0.258; 4.32 to 5.332)	308/320 (96%)	243/320 (76%)
ISTANA[26]	Docetaxel (79)	6 (7.6%)	0.87 (0.61 to 1.24)	For all-comers at tau=13 months: 9.743 (0.495; 8.772 to 10.713)	0.634 (0.459 to 0.875)	NR	NR	NR
	Gefitinib (82)	23 (28%)		For all-comers at tau=13 months: 9.949 (0.482; 9.004 to 10.90)		NR	NR	NR
ISEL[25]	BSC (563)	NR	0.84 (0.68 to 1.03)	For all-comers at tau=13 months: 6.752 (0.314; 6.138 to 7.367)	NR	NR	NR	NR
	Gefitinib	NR		For all-comers at tau=13		NR	NR	NR

Study ID (acronym)	Treatment (N)	ORR, n (%)	OS (HR) (95% CI)	OS (RMST; Mean (SE; 95% CI) (months)	PFS HR (95% CI)	PFS (RMST; months)	Any adverse event	Any Grade 3/4 adverse event
	+BSC (1,129)			months: 7.508 (0.233; 7.052 to 7.963)				
CheckMate 057	Nivolumab (292)	All-comers NSQ NSCLC: 56 (19.2%) Pooled EGFR mutation-negative/unknown NSCLC: 51 (21%)	All-comers NSQ NSCLC: 0.73 (0.59 to 0.89) Pooled EGFR mutation-negative/unknown NSCLC: 0.69 (0.56 to 0.85)*	For all-comers at tau=13 months: 9.108 (0.273; 8.572 to 9.463) For EGFR mutation-negative/unknown at tau=28 months: 14.976 (0.665; 13.673 to 16.28)	All-comers NSQ NSCLC: 0.92 (0.77 to 1.11) Pooled EGFR mutation negative/unknown NSCLC: 0.83 (0.68 to 1.02)*	For all-comers at tau=12 months: 5.116 (0.251; 4.624 to 5.696) For EGFR mutation-negative/unknown at tau=12 months: 6.336 (0.364; 5.622 to 7.05)	280 (98%)	132 (46%)[†]
	Docetaxel (290)	All-comers NSQ NSCLC: 36 (12.4%) Pooled EGFR mutation-negative/unknown NSCLC: 30 (12%)		For all-comers at tau=13 months: 8.894 (0.251; 8.402 to 9.386) For negative/unknown at tau=28 months: 12.325 (0.599; 11.151 to 13.498)		For all-comers at tau=12 months: 5.263 (0.221; 4.831 to 5.696) For EGFR mutation-negative/unknown at tau=12 months: 5.684 (0.303; 5.09 to 6.277)	265 (99%)	180 (67%)[†]
Study connected ONLY in network of all-comers NSQ NSCLC								
V-15-32[27]	Docetaxel (244)	24 (9.8%)	1.01 (0.87 to 1.27)	For all-comers at tau=13 months: 10.323 (0.240; 9.853 to 10.793)	0.89 (0.73 to 1.09)	NR	236 (99%)	195 (82%)
	Gefitinib (245)	45 (18.4%)		For all-comers at tau=13 months: 9.432 (0.275; 8.893 to 9.971)		NR	242 (99%)	99 (41%)

CI=confidence interval; EGFR=epidermal growth factor receptor; HR=hazard ratio; NR=not reported; NSCLC=non-small cell lung cancer; NSQ=non-squamous; ORR=objective response rate, OS=overall survival; PFS=progression-free survival; RMST=restricted mean survival time; SE=standard error; tau=truncation time

*These data inputs are different to those presented in Table 27 of the CS; these are updated data inputs (from an additional analysis conducted by the company) which are more appropriate for use in the ITC
†These data were not provided in Table 27 of the CS, but the ERG observed that results for the any Grade 3/4 adverse event were reported in the ITC and so the ERG requested the data inputs as part of the clarification process
Source: CS, adapted from Table 27

10.6 Indirect comparison of nivolumab versus nintedanib+docetaxel for the second-line population only

Outcome	Nivolumab vs. nintedanib+docetaxel
Patient population: All-comers NSQ NSCLC	
OS (RMST difference (95% CI]; p value)	[redacted]
PFS (RMST difference [95% CI]; p value)	[redacted]
ORR (RR [95% CI]; p value)	[redacted]
Any adverse event (RR [95% CI]; p value)	[redacted]
Any Grade 3/4 adverse event (RR [95% CI]; p value)	[redacted]
Patient population: EGFR mutation-negative/unknown NSQ NSCLC	
OS (RMST difference [95% CI]; p value)	[redacted]
PFS (RMST difference [95% CI]; p value)	[redacted]
ORR (RR [95% CI]; p value)	[redacted]

CI=confidence interval; EGFR=epidermal growth factor receptor; HR=hazard ratio; ORR=objective response rate; OS=overall survival; PFS=progression-free survival; RMST=restricted mean survival time; RR=relative risk

10.7 Health state utility

Table 53 UK-specific mean EQ-5D values by region

Region	Overall Mean	95% CI	PF Mean	95% CI	PD Mean	95% CI	ICER
USA & Canada (n=215)	0.751	0.747, 0.755	0.758	0.753, 0.763	0.730	0.723, 0.738	£100,279
Europe (n=268)	0.716	0.716, 0.717	0.735	0.732, 0.739	0.654	0.648, 0.661	£105,307
Other (n=99)	0.713	0.712, 0.715	0.717	0.715, 0.720	0.699	0.689, 0.708	£105,278

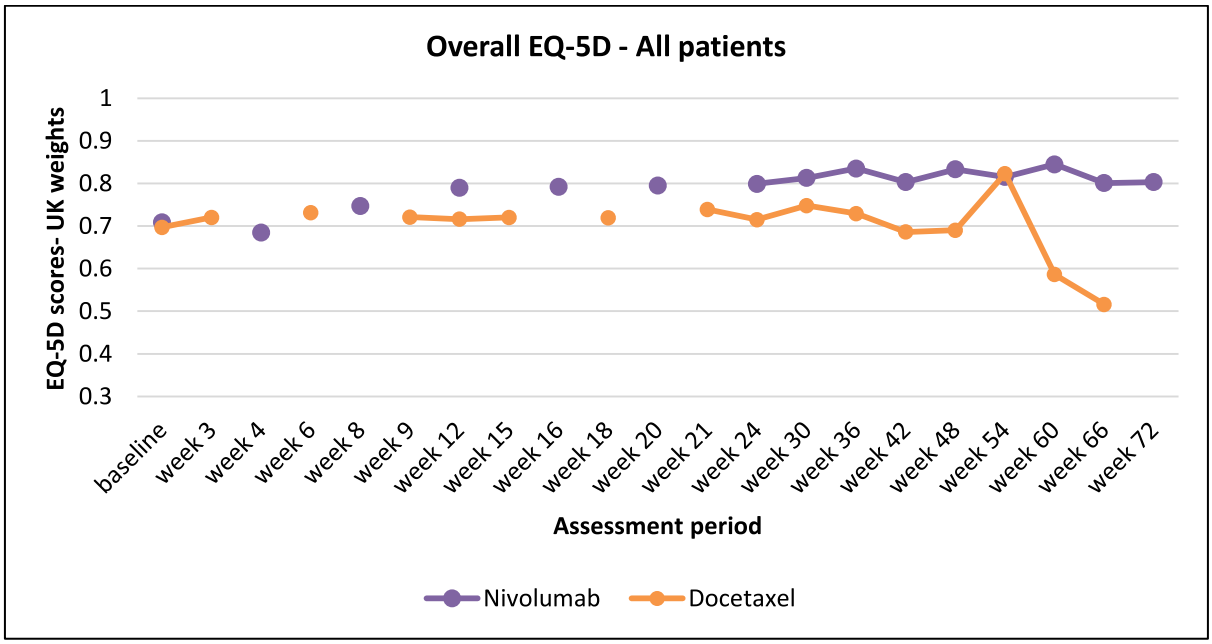
CI=confidence interval; ICER=incremental cost effectiveness ratio; PD=progressed disease; PF=progression free
Source: adapted from company's response to clarification letter, Tables 4-7

Table 54 UK-specific mean EQ-5D values by treatment arm and region

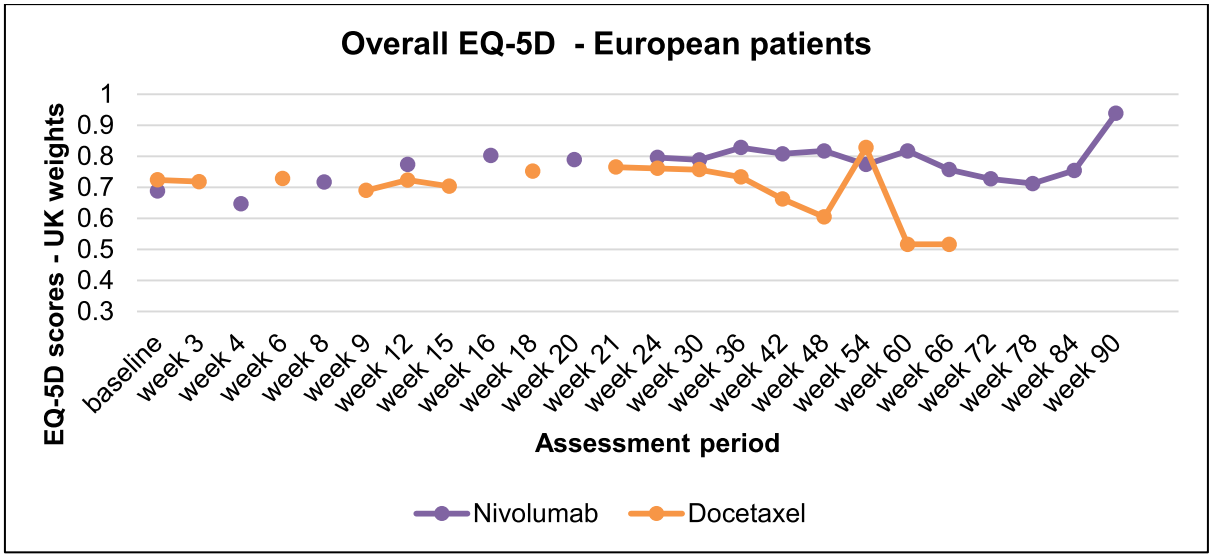
Region	Overall Mean	PF Mean	PF Mean
All			
Nivolumab	0.749	0.761	0.716
Docetaxel	0.705	0.716	0.650
USA & Canada (n=215)			
Nivolumab	0.784	0.791	0.755
Docetaxel	0.698	0.709	0.681
Europe (n=268)			
Nivolumab	0.725	0.743	0.682
Docetaxel	0.707	0.729	0.605

Other (n=99)			
Nivolumab	0.726	0.727	0.723
Docetaxel	0.703	0.711	0.644

PD= progressed disease; PF=progression-free
Source: adapted from company's response to clarification letter, Tables 4-7



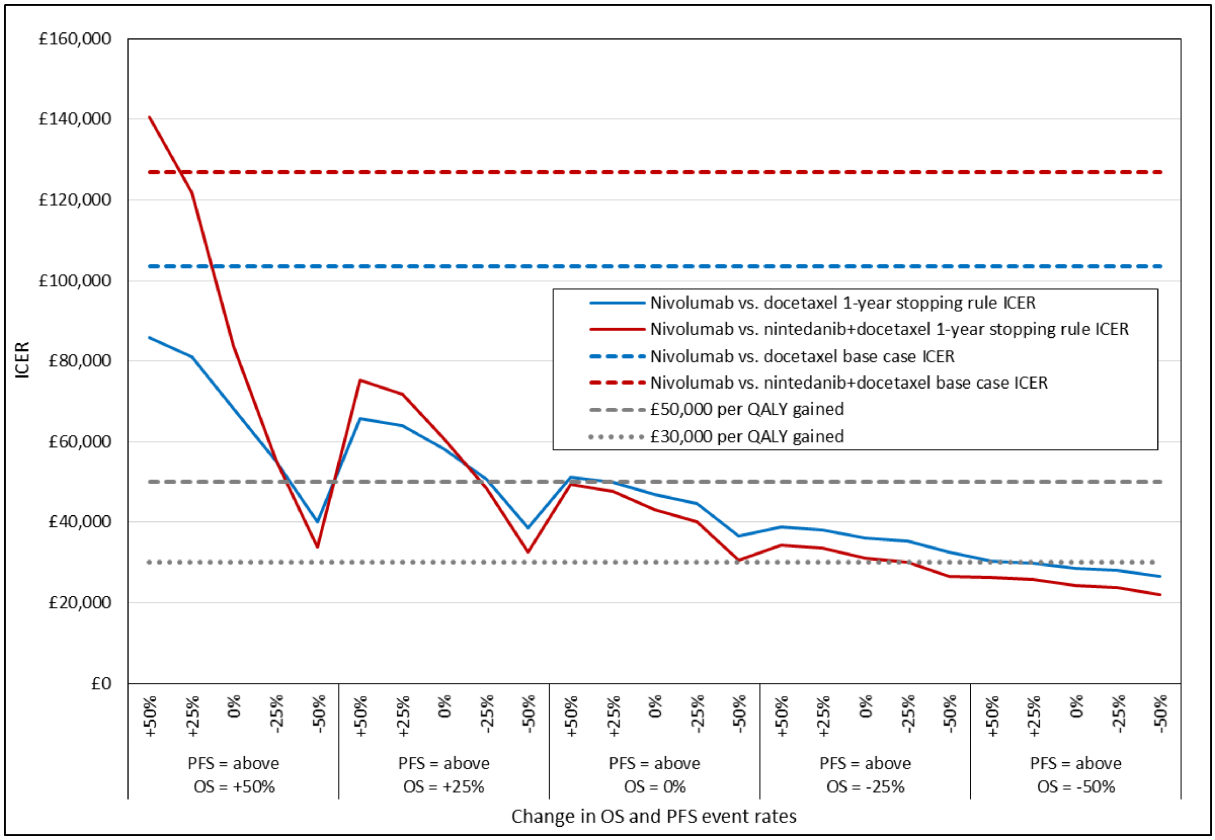
EQ-5D=EuroQol 5-dimension
Source: Company's response to clarification letter, Table 4
Figure 40 UK-specific mean EQ-5D for all patients



EQ-5D=EuroQol 5-dimension
Source: Company's response to clarification letter, Tables 5-7
Figure 41 UK-specific mean EQ-5D for European patients

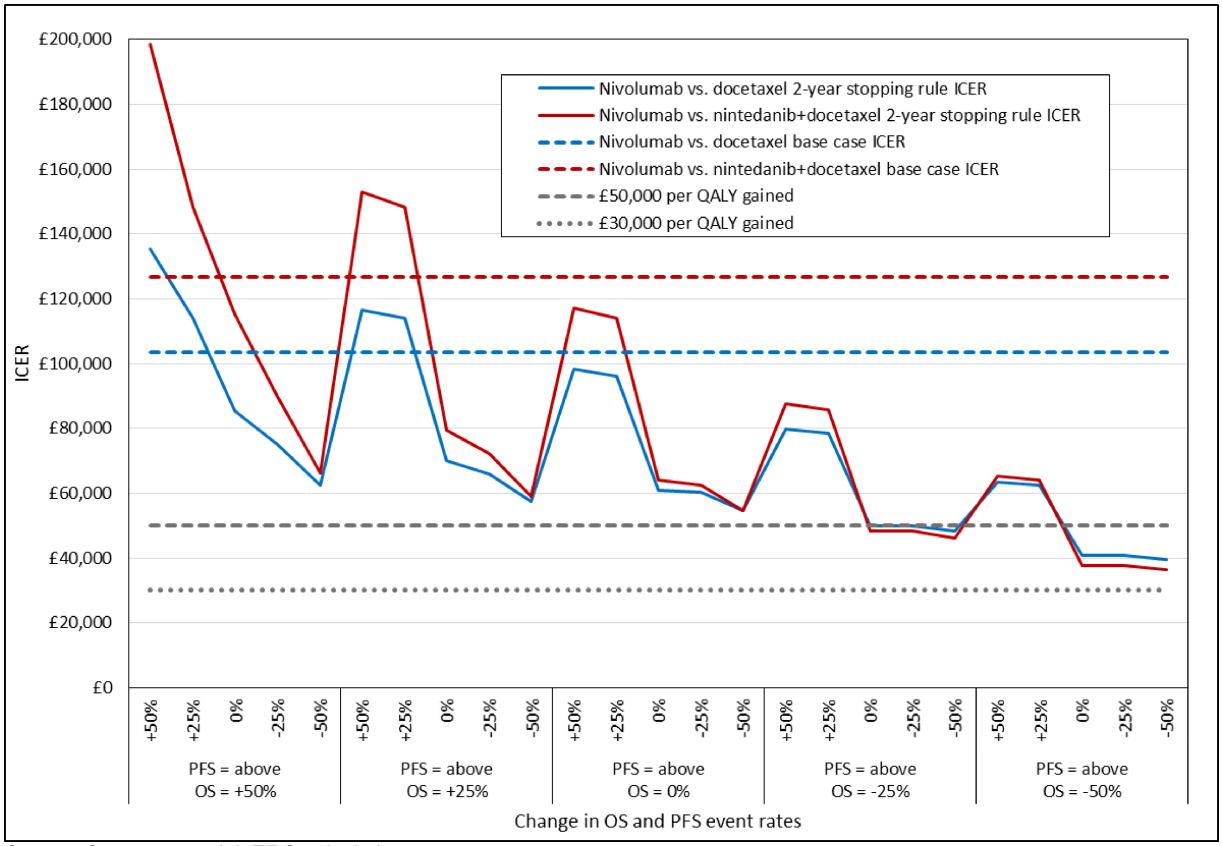
10.8 Sensitivity analysis of 1- and 2- year stopping rules

The following charts show the potential effects of varying the progression rates for PFS and mortality rates for OS by -50%, -25%, 0%, +25% and +50% after stopping treatment with nivolumab at 1 year (Figure 42) and 2 years (Figure 43).



Source: Company model, ERG calculations

Figure 42 ICERs for 1-year stopping rule with varying PFS and OS event rates



Source: Company model, ERG calculations

Figure 43 ICERs for 2-year stopping rule with varying PFS and OS event rates

10.9 ERG Revisions to company's model: Nivolumab STA

All revisions are activated by a logic switch with 0 = unchanged, 1 (or any non-zero number) = apply ERG modification.
All scenarios are activated by a logic switch with 0 = unchanged, 1 = apply first ERG scenario option, 2 (or other number >1) = apply second ERG scenario option.
Logic switches are indicated by range variables Mod_*letter* where letter = A - J.
A menu of revisions, scenarios and Mod names appears below and on the 'Results' worksheet together with summary results as used to transfer to the ERG report.

Revision	Name	Description
R1	Mod_A	ERG OS
R2	Mod_B	ERG PFS
R3	Mod_C	Nivolumab vs. docetaxel: ERG TTD
R4	Mod_D	Nivolumab vs. docetaxel: ERG PFS for disease costs and QALYS, ERG TTD for treatment costs and AEs
R5	Mod_E	Nivolumab vs. nintedanib+docetaxel: ERG TTD for nivolumab, ERG PFS for nintedanib+docetaxel
R6	Mod_F	Nivolumab vs. nintedanib+docetaxel: ERG PFS for disease costs and QALYS, ERG TTD for treatment costs and Aes for nivolumab; ERG PFS for nintedanib+docetaxel
R7	Mod_G	Nivolumab dosing calculations
R8	Mod_H	Calculate treatment administration costs at start cycle
R9	Mod_I	Use ERG utility values (Van den Hout + CheckMate 057)
R10	Mod_J	Use Nafees et al. utility values

Scenario	Name	Description
S1	Mod_Z	Sensitivity analysis PFS 1 year stopping rule: increased progression rates 0 = no change in mortality after 52 weeks, 1 = +25% mortality after 52 weeks, 2= +50% mortality after 52 weeks
S2	Mod_Y	Sensitivity analysis OS 1 year stopping rule: increased mortality rates 0 = no change in mortality after 52 weeks, 1 = +25% mortality after 52 weeks, 2= +50% mortality after 52 weeks
S3	Mod_V	Sensitivity analysis PFS 1 year stopping rule: decreased progression rates 0 = no change in mortality after 52 weeks, 1 = -25% mortality after 52 weeks, 2= -50% mortality after 52 weeks
S4	Mod_U	Sensitivity analysis OS 1 year stopping rule: decreased mortality rates 0 = no change in mortality after 52 weeks, 1 = +25% mortality after 52 weeks, 2= -50% mortality after 52 weeks
S5	Mod_X	Sensitivity analysis PFS 2 year stopping rule: increased progression rates 0 = no change in mortality after 104 weeks, 1 = +25% mortality after 104 weeks, 2= +50% mortality after 104 weeks
S6	Mod_W	Sensitivity analysis OS 2 year stopping rule: increased mortality rates 0 = no change in mortality after 104 weeks, 1 = +25% mortality after 104 weeks, 2= +50% mortality after 104 weeks
S7	Mod_T	Sensitivity analysis PFS 2 year stopping rule: decreased progression rates 0 = no change in mortality after 104 weeks, 1 = -25% mortality after 104 weeks, 2= -50% mortality after 104 weeks
S8	Mod_S	Sensitivity analysis OS 2 year stopping rule: decreased mortality rates 0 = no change in mortality after 104 weeks, 1 = -25% mortality after 104 weeks, 2= -50% mortality after 104 weeks

Instructions for modifying the company model

- Move all sheets from ID900_ERG survival estimates.xlsx into end of company model
- For each sheet given in the 'Sheet' column below:
    - copy formulae from the 'modification' column in the table below
    - paste into the cells referred to in the 'Cell' column in the table below

Sheet	Cell(s)	ERG revision/scenario number and row title	Modification name	Modification
Response and Survival	H39:H1079	R1) ERG OS	Mod_A	=(IF(INT_OS="Spline",mSpline(INT_OSsplineform,INT_OSnoknots,INT_OSsplineparams,INT_OSknots,INT_OSsplinecoef,E39:E1079),mSurvival_func(INT_OS,INT_OS_Scale,INT_OS_Shape,E39:E1079,INT_OS_Q))*IF(AND(Mod_A=0,Mod_Y=0,Mod_W=0,Mod_U=0,Mod_S=0),1,0))+('ERG OS'!C12:'ERG OS'!C1052*IF(Mod_A=1,1,0))+('ERG stopping rule'!D11:'ERG stopping rule'!D1052*IF(Mod_Y=1,1,0))+('ERG stopping rule'!E11:'ERG stopping rule'!E1052*IF(Mod_Y=2,1,0))+('ERG stopping rule'!J11:'ERG stopping rule'!J1052*IF(Mod_W=1,1,0))+('ERG stopping rule'!K11:'ERG stopping rule'!K1052*IF(Mod_W=2,1,0))+('ERG stopping rule'!P11:P1052*IF(Mod_U=1,1,0))+('ERG stopping rule'!Q11:Q1052*IF(Mod_U=2,1,0))+('ERG stopping rule'!V11:V1052*IF(Mod_S=1,1,0))+('ERG stopping rule'!W11:W1052*IF(Mod_S=2,1,0))
		S2) Sensitivity analysis OS 1 year stopping rule: increased mortality rates	Mod_Y	
		S4) Sensitivity analysis OS 2 year stopping rule: dcreased mortality rates	Mod_W	
		S6) Sensitivity analysis OS 1 year stopping rule: decreased mortality rates	Mod_U	
		S8) Sensitivity analysis OS 2 year stopping rule: decreased mortality rates	Mod_S	

Sheet	Cell(s)	ERG revision/scenario number and row title	Modification name	Modification
Response and Survival	G39:G1079	S1) Sensitivity analysis PFS 1 year stopping rule: increased progression rates S3) Sensitivity analysis PFS 2 year stopping rule: increased progression rates S5) Sensitivity analysis PFS 1 year stopping rule: decreased progression rates S7) Sensitivity analysis PFS 2 year stopping rule: decreased progression rates	Mod_Z Mod_X Mod_V Mod_T	=(IF(INT_PFS="Spline", mSpline(INT_PFSsplineform, INT_PFSnoknots, INT_PFSsplineparams, INT_PFSknots, INT_PFSsplinecoef, E39:E1079), mSurvival_func(INT_PFS, INT_PFS_Scale, INT_PFS_Shape, E39:E1079, INT_PFS_Q)))*IF(AND(Mod_Z=0,Mod_X=0,Mod_V=0,Mod_T=0),1,0)+('ERG stopping rule'!B11:'ERG stopping rule'!B1052*IF(Mod_Z=1,1,0))+('ERG stopping rule'!C11:'ERG stopping rule'!C1052*IF(Mod_Z=2,1,0))+('ERG stopping rule'!H11:'ERG stopping rule'!H1052*IF(Mod_X=1,1,0))+('ERG stopping rule'!I11:'ERG stopping rule'!I1052*IF(Mod_X=2,1,0))+('ERG stopping rule'!N11:N1052*IF(Mod_V=1,1,0))+('ERG stopping rule'!O11:O1052*IF(Mod_V=2,1,0))+('ERG stopping rule'!T11:T1052*IF(Mod_T=1,1,0))+('ERG stopping rule'!U11:U1052*IF(Mod_T=2,1,0))
	J39:J1079	R1) ERG OS	Mod_A	=(IF(TRT1_OS="Spline",mSpline(TRT1_OSsplineform,TRT1_OSnosplines,TRT1_OSsplineparams,TRT1_OSknots,TRT1_OSsplinecoef,E39:E1079),mSurvival_func(TRT1_OS,TRT1_OS_scale,TRT1_OS_shape,E39:E1079,TRT1_OS_Q))*IF(Mod_A=0,1,0))+ ('ERG OS'!D12:'ERG OS'!D1052*IF(Mod_A=1,1,0))
	W38:W64	R1) ERG OS	Mod_A	=(MAX($J38,V38)*IF(Mod_A=0,1,0))+('ERG OS'!E11*IF(Mod_A=1,1,0))
	W65:W1079	R1) ERG OS	Mod_A	=(MAX($J65^TRT2_HR_OS_user,V65)*IF(Mod_A=0,1,0))+('ERG OS'!E38*IF(Mod_A=1,1,0))
	R38:R1079	R3) Nivolumab vs. docetaxel: ERG TTD R5) Nivolumab vs. nintedanib+docetaxel: ERG TTD for nivolumab, ERG PFS for nintedanib+docetaxel	Mod_C Mod_E	=(O38*IF(AND(Mod_C=0,Mod_E=0),1,0))+('ERG TTD'!C11*IF(OR(Mod_C=1,Mod_E=1),1,0))
	T38:T1079	R3) Nivolumab vs. docetaxel: ERG TTD	Mod_C	=(I38*IF(Mod_C=0,1,0))+('ERG TTD'!O11*IF(Mod_C=1,1,0))

Sheet	Cell(s)	ERG revision/scenario number and row title	Modification name	Modification
Response and Survival	V38:V46	R2) ERG PFS R5) Nivolumab vs. nintedanib+docetaxel: ERG TTD for nivolumab, ERG PFS for nintedanib+docetaxel R6) Nivolumab vs. nintedanib+docetaxel: ERG PFS for disease costs and QALYS, ERG TTD for treatment costs and Aes for nivolumab; ERG PFS for nintedanib+docetaxel	Mod_B Mod_E Mod_F	=($I38*IF(AND(Mod_B=0,Mod_E=0,Mod_F=0),1,0))+('ERG PFS'!E11*IF(OR(Mod_B=1,Mod_E=1,Mod_F=1),1,0))
	V47:V1079	R2) ERG PFS R5) Nivolumab vs. nintedanib+docetaxel: ERG TTD for nivolumab, ERG PFS for nintedanib+docetaxel R6) Nivolumab vs. nintedanib+docetaxel: ERG PFS for disease costs and QALYS, ERG TTD for treatment costs and Aes for nivolumab; ERG PFS for nintedanib+docetaxel	Mod_B Mod_E Mod_F	=(($I47^TRT2_HR_PFS_user)*IF(AND(Mod_B=0,Mod_E=0,Mod_F=0),1,0))+('ERG PFS'!E20*IF(OR(Mod_B=1,Mod_E=1,Mod_F=1),1,0))

Sheet	Cell(s)	ERG revision/scenario number and row title	Modification name	Modification
Cost	E10	R1) ERG OS	Mod_A	=((PF_cost/4)*((('Patient flow - 1'!$P14)*IF(AND(Mod_B=0,Mod_D=0,Mod_F=0),1,0))+('ERG health states'!E11*IF(OR(Mod_B=1,Mod_D=1,Mod_F=1),1,0))))*$C10
		R2) ERG PFS	Mod_B	+ ((PD_cost/4)*((('Patient flow - 1'!$Q14)*IF(AND(Mod_B=0,Mod_D=0,Mod_F=0),1,0))+(((('ERG health states'!I11)*IF(AND(Mod_A=1,OR(Mod_B=1,Mod_D=1,Mod_F=1)),1,0)))+('ERG health
		R4) Nivolumab vs. docetaxel: ERG PFS for disease costs and QALYS, ERG TTD for treatment costs and AEs	Mod_D	states'!M11*IF(AND(Mod_A=0,OR(Mod_B=1,Mod_D=1,Mod_F=1)),1,0))))*$C10 + ((terminal_cost)*('Patient flow - 1'!J14))*C10
		R6) Nivolumab vs. nintedanib+docetaxel: ERG PFS for disease costs and QALYS, ERG TTD for treatment costs and Aes for nivolumab; ERG PFS for nintedanib+docetaxel	Mod_F	
Cost	E11:E1049	R1) ERG OS	Mod_A	=((PF_cost/4)*((('Patient flow - 1'!$P15)*IF(AND(Mod_B=0,Mod_D=0,Mod_F=0),1,0))+('ERG health states'!E12*IF(OR(Mod_B=1,Mod_D=1,Mod_F=1),1,0))))*$C11
		R2) ERG PFS	Mod_B	+ ((PD_cost/4)*((('Patient flow - 1'!$Q15)*IF(AND(Mod_B=0,Mod_D=0,Mod_F=0),1,0))+(((('ERG health states'!I12)*IF(AND(Mod_A=1,OR(Mod_B=1,Mod_D=1,Mod_F=1)),1,0)))+('ERG health
		R4) Nivolumab vs. docetaxel: ERG PFS for disease costs and QALYS, ERG TTD for treatment costs and AEs	Mod_D	states'!M12*IF(AND(Mod_A=0,OR(Mod_B=1,Mod_D=1,Mod_F=1)),1,0))))*$C11 + ((terminal_cost)*('Patient flow - 1'!J15-'Patient flow - 1'!J14))*C11+E10
		R6) Nivolumab vs. nintedanib+docetaxel: ERG PFS for disease costs and QALYS, ERG TTD for treatment costs and Aes for nivolumab; ERG PFS for nintedanib+docetaxel	Mod_F	
	N10	R1) ERG OS	Mod_A	=((PF_cost/4)*((('Patient flow - 1'!$AH14)*IF(AND(Mod_B=0,Mod_D=0),1,0))+('ERG health states'!Y11*IF(OR(Mod_B=1,Mod_D=1),1,0))))*$C10
		R2) ERG PFS	Mod_B	+ ((PD_cost/4)*((('Patient flow - 1'!$AI14)*IF(AND(Mod_B=0,Mod_D=0),1,0))+(((('ERG health states'!AC11)*IF(AND(Mod_A=1,OR(Mod_B=1,Mod_D=1)),1,0)))+('ERG health
		R4) Nivolumab vs. docetaxel: ERG PFS for disease costs and QALYS, ERG TTD for treatment costs and AEs	Mod_D	states'!AG11*IF(AND(Mod_A=0,OR(Mod_B=1,Mod_D=1)),1,0))))*$C10 +(terminal_cost)*('Patient flow - 1'!AB14)*C10

Sheet	Cell(s)	ERG revision/scenario number and row title	Modification name	Modification
Cost	N11:N1049	R1) ERG OS R2) ERG PFS R4) Nivolumab vs. docetaxel: ERG PFS for disease costs and QALYS, ERG TTD for treatment costs and AEs	Mod_A Mod_B Mod_D	=((PF_cost/4)*((('Patient flow - 1'!$AH15)*IF(AND(Mod_B=0,Mod_D=0),1,0))+('ERG health states'!Y12*IF(OR(Mod_B=1,Mod_D=1),1,0))))*$C11 + ((PD_cost/4)*((('Patient flow - 1'!$AI15)*IF(AND(Mod_B=0,Mod_D=0),1,0))+((('ERG health states'!AC12)*IF(AND(Mod_A=1,OR(Mod_B=1,Mod_D=1)),1,0)))+('ERG health states'!AG12*IF(AND(Mod_A=0,OR(Mod_B=1,Mod_D=1)),1,0))))*$C11 + ((terminal_cost)*MAX(0, ('Patient flow - 1'!AB15-'Patient flow - 1'!X15))*$C11) + N10
	F10	R2) ERG PFS R4) Nivolumab vs. docetaxel: ERG PFS for disease costs and QALYS, ERG TTD for treatment costs and AEs R6) Nivolumab vs. nintedanib+docetaxel: ERG PFS for disease costs and QALYS, ERG TTD for treatment costs and Aes for nivolumab; ERG PFS for nintedanib+docetaxel	Mod_B Mod_D Mod_F	=IF(econ_dose_cap_on, IF(B10 <= econ_dose_cap, 1, 0), 1)*IF(dose_cap_on, IF(B10 <= dose_cap, 1, 0), 1)*IF(Cap_on, IF(B10 <= trt_cap, 1,0), 1) *((INT_acq*(('Patient flow - 1'!D14*IF(AND(Mod_D=0,Mod_F=0),1,0))+('ERG health states'!C11*IF(Mod_B=1,1,0))+('ERG TTD'!C11*IF(OR(Mod_D=1,Mod_F=1),1,0))))*1 + (INT_PD_Trt*INT_PD_doses*INT_acq*('Patient flow - 1'!E14))) + (0*('Patient flow - 1'!E14))*$C10
	F11:F1049	R2) ERG PFS R4) Nivolumab vs. docetaxel: ERG PFS for disease costs and QALYS, ERG TTD for treatment costs and AEs R6) Nivolumab vs. .nintedanib+docetaxel: ERG PFS for disease costs and QALYS, ERG TTD for treatment costs and Aes for nivolumab; ERG PFS for nintedanib+docetaxel	Mod_B Mod_D Mod_F	=IF(MOD($A11, INT_periodicity) = 0, 1, 0)*IF(econ_dose_cap_on, IF(B11 <= econ_dose_cap, 1, 0), 1)*IF(dose_cap_on, IF(B11 <= dose_cap, 1, 0), 1)*IF(Cap_on, IF(B11 <= trt_cap, 1,0), 1) *((INT_acq*(('Patient flow - 1'!D15*IF(AND(Mod_B=0,Mod_D=0,Mod_F=0),1,0))+('ERG health states'!C12*IF(Mod_B=1,1,0))+('ERG TTD'!C12*IF(OR(Mod_D=1,Mod_F=1),1,0))))*$C11 + (INT_PD_Trt*INT_PD_doses*INT_acq*MAX(0,(('Response and survival'!BC38 - 'Response and survival'!BC39)*$C11))) + (0*('Patient flow - 1'!E15))*$C11) + F10

Sheet	Cell(s)	ERG revision/scenario number and row title	Modification name	Modification
Cost	O10	R2) ERG PFS R4) Nivolumab vs. docetaxel: ERG PFS for disease costs and QALYS, ERG TTD for treatment costs and AEs	Mod_B Mod_D	=(TRT1_acq*(('Patient flow - 1'!V14*IF(AND(Mod_B=0,Mod_D=0),1,0))+('ERG health states'!W11*IF(Mod_B=1,1,0))+('ERG TTD'!O11*IF(Mod_D=1,1,0)))*$C10) + (0*('Patient flow - 1'!W14)*$C10)
	O11:O1049	R2) ERG PFS R4) Nivolumab vs. docetaxel: ERG PFS for disease costs and QALYS, ERG TTD for treatment costs and AEs	Mod_B Mod_D	=IF(MOD($A11, TRT1_periodicity) = 0, 1, 0) *(TRT1_acq*((('Patient flow - 1'!V15*IF(AND(Mod_B=0,Mod_D=0),1,0))+('ERG health states'!W12*IF(Mod_B=1,1,0))+('ERG TTD'!O12*IF(Mod_D=1,1,0)))*$C11 + 0*('Patient flow - 1'!W15*$C11))) + O10
	G10	R2) ERG PFS R4) Nivolumab vs docetaxel: ERG PFS for disease costs and QALYS, ERG TTD for treatment costs and AEs R6) Nivolumab vs. nintedanib+docetaxel: ERG PFS for disease costs and QALYS, ERG TTD for treatment costs and Aes for nivolumab; ERG PFS for nintedanib+docetaxel R8) Calculate treatment administration costs at start cycle	Mod_B Mod_D Mod_F Mod_H	=IF(dose_cap_on, IF(B10<= dose_cap, 1, 0), 1)*IF(Cap_on, IF(B10 <= trt_cap, 1,0), 1)*(INT_admin*(('Patient flow - 1'!$P14*IF(AND(Mod_B=0,Mod_D=0,Mod_F=0,Mod_H=0),1,0))+('Patient flow - 1'!$D14*IF(AND(Mod_B=0,Mod_D=0,Mod_F=0,Mod_H=1),1,0))+('ERG health states'!E11*IF(AND(Mod_B=1,Mod_H=0),1,0))+('ERG health states'!C11*IF(AND(Mod_B=1,Mod_H=1),1,0))+('ERG TTD'!E11*IF(AND(Mod_H=0,OR(Mod_D=1,Mod_F=1)),1,0)) +('ERG TTD'!C11*IF(AND(Mod_H=1,OR(Mod_D=1,Mod_F=1)),1,0))*$C10+ (INT_PD_Trt*INT_PD_doses*INT_admin*('Patient flow - 1'!$Q14)*C10)+ (0*('Patient flow - 1'!$Q14))*$C10))

Sheet	Cell(s)	ERG revision/scenario number and row title	Modification name	Modification
Cost	G11:G1049	R2) ERG PFS R4) Nivolumab vs. docetaxel: ERG PFS for disease costs and QALYS, ERG TTD for treatment costs and AEs R6) Nivolumab vs. nintedanib+docetaxel: ERG PFS for disease costs and QALYS, ERG TTD for treatment costs and Aes for nivolumab; ERG PFS for nintedanib+docetaxel R8) Calculate treatment administration costs at start cycle	Mod_B Mod_D Mod_F Mod_H	=IF(MOD($A11, INT_periodicity) = 0, 1, 0)*IF(dose_cap_on, IF(B11 <= dose_cap, 1, 0), 1)*IF(Cap_on, IF(B11 <= trt_cap, 1,0), 1) *(((INT_admin*(('Patient flow - 1'!$P15 *IF(AND(Mod_B=0,Mod_D=0,Mod_F=0,Mod_H=0),1,0))+('Patient flow - 1'!$D15*IF(AND(Mod_B=0,Mod_D=0,Mod_F=0,Mod_H=1),1,0))+('ERG health states'!E12*IF(AND(Mod_B=1,Mod_H=0),1,0))+('ERG health states'!C12*IF(AND(Mod_B=1,Mod_H=1),1,0))+('ERG TTD'!E12*IF(AND(Mod_H=0,OR(Mod_D=1,Mod_F=1)),1,0)) +('ERG TTD'!C12*IF(AND(Mod_H=1,(OR(Mod_D=1,Mod_F=1))),1,0))))*$C11 + (INT_PD_Trt*INT_PD_doses*INT_admin*MAX(0,(('Response and survival'!BC38) - ('Response and survival'!BC39)))*C11)) + (0*('Patient flow - 1'!$Q15))*$C11) + G10
	P10	R2) ERG PFS R4) Nivolumab vs. docetaxel: ERG PFS for disease costs and QALYS, ERG TTD for treatment costs and AEs R8) Calculate treatment administration costs at start cycle	Mod_B Mod_D Mod_H	=(TRT1_admin*(('Patient flow - 1'!$AH14*IF(AND(Mod_B=0,Mod_D=0,Mod_H=0),1,0))+('Patient flow - 1'!V14*IF(AND(Mod_H=1,Mod_B=0,Mod_D=0),1,0))+('ERG health states'!Y11*IF(AND(Mod_B=1, Mod_H=0),1,0))+('ERG health states'!W11*IF(AND(Mod_B=1, Mod_H=1),1,0))+('ERG TTD'!Q11*IF(AND(Mod_D=1, Mod_H=0),1,0))+('ERG TTD'!O11*IF(AND(Mod_D=1, Mod_H=1),1,0)))*$C10)+ (0*('Patient flow - 1'!$AI14))*$C10
	P11:P1049	R2) ERG PFS R4) Nivolumab vs. docetaxel: ERG PFS for disease costs and QALYS, ERG TTD for treatment costs and AEs R8) Calculate treatment administration costs at start cycle	Mod_B Mod_D Mod_H	=IF(MOD($A11, TRT1_periodicity) = 0, 1, 0)*(TRT1_admin*(('Patient flow - 1'!$AH15*IF(Mod_D=0,1,0))+('Patient flow - 1'!V15*IF(AND(Mod_H=1,Mod_B=0,Mod_D=0),1,0))+('ERG health states'!Y12*IF(AND(Mod_B=1, Mod_H=0),1,0))+('ERG health states'!W12*IF(AND(Mod_B=1, Mod_H=1),1,0))+('ERG TTD'!Q12*IF(AND(Mod_D=1,Mod_H=0),1,0))+('ERG TTD'!O12*IF(AND(Mod_D=1,Mod_H=1),1,0))))*$C11+ 0*('Patient flow - 1'!W15)*$C11 + (0*('Patient flow - 1'!$AI15)*$C11) + P10

Sheet	Cell(s)	ERG revision/scenario number and row title	Modification name	Modification
Cost	Y10	R8) Calculate treatment administration costs at start cycle	Mod_H	=((TRT2_admin*(('Patient flow - 1'!$AZ14*IF(Mod_H=0,1,0))+'Patient flow - 1'!$AN14*IF(Mod_H=1,1,0)))*$C10 + (0*('Patient flow - 1'!$BA14))*$C10)
	Y11:Y1049	R8) Calculate treatment administration costs at start cycle	Mod_H	=IF(MOD($A11, TRT2_periodicity) = 0, 1, 0)*(TRT2_admin*(('Patient flow - 1'!$AZ15*IF(Mod_H=0,1,0))+('Patient flow - 1'!$AN15*IF(Mod_H=1,1,0)))*$C11 + (0*('Patient flow - 1'!$BA15))*$C11) + Y10
	H10	R2) ERG PFS R4) Nivolumab vs. docetaxel: ERG PFS for disease costs and QALYS, ERG TTD for treatment costs and AEs R6) Nivolumab vs. nintedanib+docetaxel: ERG PFS for disease costs and QALYS, ERG TTD for treatment costs and Aes for nivolumab; ERG PFS for nintedanib+docetaxel	Mod_B Mod_D Mod_F	=(((INT_mon/4)*(('Patient flow - 1'!$P14*IF(AND(Mod_B=0,Mod_D=0,Mod_F=0),1,0))+('ERG health states'!E11*IF(Mod_B=1,1,0))+('ERG TTD'!E11*IF(OR(Mod_D=1,Mod_F=1),1,0))))+(INT_PD_Trt*INT_PD_time*(INT_mon/4)*('Patient flow - 1'!$Q14)*C10))*IF(dose_cap_on, IF(B11 <= dose_cap, 1, 0), 1)*IF(Cap_on, IF(B11 <= trt_cap, 1,0), 1)*$C10+(0*('Patient flow - 1'!$Q14))*$C10
	H11:H1049	R2) ERG PFS R4) Nivolumab vs. docetaxel: ERG PFS for disease costs and QALYS, ERG TTD for treatment costs and AEs R6) Nivolumab vs. nintedanib+docetaxel: ERG PFS for disease costs and QALYS, ERG TTD for treatment costs and Aes for nivolumab; ERG PFS for nintedanib+docetaxel	Mod_B Mod_D Mod_F	=(((INT_mon/4)+ (INT_PD_Trt*INT_PD_time*(INT_mon/4)*MAX(0,(('Response and survival'!BC38) - ('Response and survival'!BC39)))*C11))*IF(dose_cap_on, IF(B11 <= dose_cap, 1, 0), 1)*IF(Cap_on, IF(B11 <= trt_cap, 1,0), 1)*(('Patient flow - 1'!$P15*IF(AND(Mod_B=0,Mod_D=0,Mod_F=0),1,0))++('ERG health states'!E12*IF(Mod_B=1,1,0))+('ERG TTD'!E12*IF(OR(Mod_D=1,Mod_F=1),1,0))))*$C11+(0*('Patient flow - 1'!$Q15))*$C11 + H10
	Q10	R2) ERG PFS R4) Nivolumab vs. docetaxel: ERG PFS for disease costs and QALYS, ERG TTD for treatment costs and AEs	Mod_B Mod_D	=((TRT1_mon/4)*(('Patient flow - 1'!$AH14*IF(AND(Mod_B=0,Mod_D=0),1,0))+('ERG health states'!Y11*IF(Mod_B=1,1,0))+('ERG TTD'!Q11*IF(Mod_D=1,1,0))))*$C10 + (0*'Patient flow - 1'!$AI14*$C10)

Sheet	Cell(s)	ERG revision/scenario number and row title	Modification name	Modification
Cost	Q11:Q1049	R4) Nivolumab vs. docetaxel: ERG PFS for disease costs and QALYS, ERG TTD for treatment costs and AEs	Mod_D	=((TRT1_mon/4)*(('Patient flow - 1'!$AH15*IF(AND(Mod_B=0,Mod_D=0),1,0))+('ERG health states'!Y12*IF(Mod_B=1,1,0))+('ERG TTD'!Q12*IF(Mod_D=1,1,0))))*$C11 + (0*('Patient flow - 1'!$AI15)*$C11) + Q10
	I10	R1) ERG OS R2) ERG PFS R4) Nivolumab vs. docetaxel: ERG PFS for disease costs and QALYS, ERG TTD for treatment costs and AEs R6) Nivolumab vs. nintedanib+docetaxel: ERG PFS for disease costs and QALYS, ERG TTD for treatment costs and Aes for nivolumab; ERG PFS for nintedanib+docetaxel	Mod_A Mod_B Mod_D Mod_F	=((INT_subtrt_cost*INT_subtrt_prop)+ (INT_subtrt_admin_cost*INT_subtrt_prop)+(INT_subtrt_mon_cost*INT_subtrt_prop))*('Patient flow - 1'!$Q14*IF(AND(Mod_B=0,Mod_D=0,Mod_F=0),1,0)+MAX('ERG TTD'!G11,'ERG health states'!G11)*IF(AND(Mod_A=1,OR(Mod_B=1,Mod_D=1,Mod_F=1)),1,0)+MAX('ERG TTD'!K11,'ERG health states'!K11)*IF(AND(Mod_A=0,OR(Mod_B=1,Mod_D=1,Mod_F=1)),1,0))*C10
	I11:I1049	R1) ERG OS R2) ERG PFS R4) Nivolumab vs. docetaxel: ERG PFS for disease costs and QALYS, ERG TTD for treatment costs and AEs R6) Nivolumab vs. nintedanib+docetaxel: ERG PFS for disease costs and QALYS, ERG TTD for treatment costs and Aes for nivolumab; ERG PFS for nintedanib+docetaxel	Mod_A Mod_B Mod_D Mod_F	=((INT_subtrt_cost*INT_subtrt_prop)+ (INT_subtrt_admin_cost*INT_subtrt_prop)+ (INT_subtrt_mon_cost*INT_subtrt_prop)) *MAX(0,((('Response and survival'!BC38) - ('Response and survival'!BC39))*IF(AND(Mod_B=0,Mod_D=0,Mod_F=0),1,0))+(IF('ERG TTD'!G12>'ERG health states'!G12,'ERG TTD'!G12-'ERG TTD'!G11,'ERG health states'!G12-'ERG health states'!G11)*IF(AND(Mod_A=1,OR(Mod_B=1,Mod_D=1,Mod_F=1)),1,0))+(IF('ERG TTD'!K12>'ERG health states'!K12,'ERG TTD'!K12-'ERG TTD'!K11,'ERG health states'!K12-'ERG health states'!K11)*IF(AND(Mod_A=0,OR(Mod_B=1,Mod_D=1,Mod_F=1)),1,0)))*C11)+I10

Sheet	Cell(s)	ERG revision/scenario number and row title	Modification name	Modification
	R10	R1) ERG OS R2) ERG PFS R4) Nivolumab vs. docetaxel: ERG PFS for disease costs and QALYS, ERG TTD for treatment costs and AEs	Mod_A Mod_B Mod_D	=((TRT1_subtrt_cost*TRT1_subtrt_prop) + (TRT1_subtrt_admin_cost*TRT1_subtrt_prop) + (TRT1_subtrt_mon_cost*TRT1_subtrt_prop)) *(('Patient flow - 1'!$AI14*IF(AND(Mod_B=0,Mod_D=0),1,0))+('ERG health states'!AA11*IF(AND(Mod_A=1,OR(Mod_B=1,Mod_D=1)),1,0))+('ERG health states'!AE11*IF(AND(Mod_A=0,OR(Mod_B=1,Mod_D=1)),1,0)))*C10
	R11:R1049	R2) ERG PFS R4) Nivolumab vs. docetaxel: ERG PFS for disease costs and QALYS, ERG TTD for treatment costs and AEs	Mod_B Mod_D	=((TRT1_subtrt_cost*TRT1_subtrt_prop) + (TRT1_subtrt_admin_cost*TRT1_subtrt_prop)+(TRT1_subtrt_mon_cost*TRT1_subtrt_prop))*MAX(0, ((('Response and survival'!BE38-'Response and survival'!BE39)*IF(AND(Mod_B=0,Mod_D=0),1,0))+(('ERG health states'!W11-'ERG health states'!W12)*IF(OR(Mod_B=1,Mod_D=1),1,0)))*C11) +R10
Nivolumab cost	I9	R7) Nivolumab dosing calculations	Mod_G	=((LN(D9)-I36^2/2)*IF(Mod_G=0,1,0))+((LN(D9)-I10^2/2)*IF(Mod_G=1,1,0))
	J9	R7) Nivolumab dosing calculations	Mod_G	=((LN(E9)-J36^2/2)*IF(Mod_G=0,1,0))+((LN(E9)-J10^2/2)*IF(Mod_G=1,1,0))
Outcomes	F12	R9) Use ERG utility values (Van den Hout + CheckMate 057) R10) Use Nafees et al. utility values	Mod_I Mod_J	=0.739*IF(AND(Mod_I=0,Mod_J=0),1,0)+0.713*IF(Mod_I=1,1,0)+0.65*IF(Mod_J=1,1,0)
	F13	R9) Use ERG utility values (Van den Hout + CheckMate 057) R10) Use Nafees et al. utility values	Mod_I Mod_J	=0.688*IF(AND(Mod_I=0,Mod_J=0),1,0)+0.476*IF(Mod_I=1,1,0)+0.43*IF(Mod_J=1,1,0)

Sheet	Cell(s)	ERG revision/scenario number and row title	Modification name	Modification
Outcome	F10	R1) ERG OS	Mod_A	=(utility_PFS*((('Patient flow - 1'!$P14*IF(AND(Mod_B=0,Mod_D=0,Mod_F=0),1,0))+('ERG health states'!E11*IF(OR(Mod_B=1,Mod_D=1,Mod_F=1),1,0)))/52)*$C10) + (utility_PD*((('Patient flow - 1'!$Q14*IF(AND(Mod_A=0,Mod_B=0,Mod_D=0,Mod_F=0),1,0))+('ERG health states'!Q11*IF(AND(Mod_A=1,Mod_B=0,Mod_D=0,Mod_F=0),1,0))+('ERG health states'!M11*IF(AND(Mod_A=0,OR(Mod_B=1,Mod_D=1,Mod_F=1)),1,0))+('ERG health states'!I11*IF(AND(Mod_A=1,OR(Mod_B=1,Mod_D=1,Mod_F=1)),1,0)))/52)*$C10)
		R2) ERG PFS	Mod_B	
		R4) Nivolumab vs docetaxel: ERG PFS for disease costs and QALYS, ERG TTD for treatment costs and AEs	Mod_D	
		R6) Nivolumab vs. nintedanib+docetaxel: ERG PFS for disease costs and QALYS, ERG TTD for treatment costs and Aes for nivolumab; ERG PFS for nintedanib+docetaxel	Mod_F	
Outcome	F11:F1049	R1) ERG OS	Mod_A	=(utility_PFS*((('Patient flow - 1'!$P15*IF(AND(Mod_B=0,Mod_D=0,Mod_F=0),1,0))+('ERG health states'!E12*IF(OR(Mod_B=1,Mod_D=1,Mod_F=1),1,0)))/52)*$C11) + (utility_PD*((('Patient flow - 1'!$Q15*IF(AND(Mod_A=0,Mod_B=0,Mod_D=0,Mod_F=0),1,0))+('ERG health states'!Q12*IF(AND(Mod_A=1,Mod_B=0,Mod_D=0,Mod_F=0),1,0))+('ERG health states'!M12*IF(AND(Mod_A=0,OR(Mod_B=1,Mod_D=1,Mod_F=1)),1,0))+('ERG health states'!I12*IF(AND(Mod_A=1,OR(Mod_B=1,Mod_D=1,Mod_F=1)),1,0)))/52)*$C11) + F10
		R2) ERG PFS	Mod_B	
		R4) Nivolumab vs. docetaxel: ERG PFS for disease costs and QALYS, ERG TTD for treatment costs and AEs	Mod_D	
		R6) Nivolumab vs. nintedanib+docetaxel: ERG PFS for disease costs and QALYS, ERG TTD for treatment costs and Aes for nivolumab; ERG PFS for nintedanib+docetaxel	Mod_F	
	L10	R1) ERG OS	Mod_A	=(utility_PFS*((('Patient flow - 1'!$AH14*IF(AND(Mod_B=0,Mod_D=0),1,0))+('ERG health states'!Y11*IF(OR(Mod_B=1,Mod_D=1),1,0)))/52)*$C10) + (utility_PD*((('Patient flow - 1'!$AI14*IF(AND(Mod_A=0,Mod_B=0,Mod_D=0),1,0))+('ERG health states'!AK11*IF(AND(Mod_A=1,Mod_B=0,Mod_D=0),1,0))+('ERG health states'!AG11*IF(AND(Mod_A=0,OR(Mod_B=1,Mod_D=1)),1,0)))+('ERG health states'!AC11*IF(AND(Mod_A=1,OR(Mod_B=1,Mod_D=1)),1,0)))/52)*$C10
		R2) ERG PFS	Mod_B	
		R4) Nivolumab vs. docetaxel: ERG PFS for disease costs and QALYS, ERG TTD for treatment costs and AEs	Mod_D	

Sheet	Cell(s)	ERG revision/scenario number and row title	Modification name	Modification
Outcome	L11:L1049	R1) ERG OS R2) ERG PFS R4) Nivolumab vs. docetaxel: ERG PFS for disease costs and QALYS, ERG TTD for treatment costs and AEs	Mod_A Mod_B Mod_D	=(utility_PFS*((('Patient flow - 1'!$AH15*IF(AND(Mod_B=0,Mod_D=0),1,0))+('ERG health states'!Y12*IF(OR(Mod_B=1,Mod_D=1),1,0)))/52)*$C11) + (utility_PD*((('Patient flow - 1'!$AI15*IF(AND(Mod_A=0,Mod_B=0,Mod_D=0),1,0))+('ERG health states'!AK12*IF(AND(Mod_A=1,Mod_B=0,Mod_D=0),1,0))+('ERG health states'!AG12*IF(AND(Mod_A=0,OR(Mod_B=1,Mod_D=1)),1,0))+('ERG health states'!AC12*IF(AND(Mod_A=1,OR(Mod_B=1,Mod_D=1)),1,0)))/52)*$C11) + L10
	R10	R1) ERG OS R2) ERG PFS R5) Nivolumab vs. nintedanib+docetaxel: ERG TTD for nivolumab, ERG PFS for nintedanib+docetaxel R6) Nivolumab vs. nintedanib+docetaxel: ERG PFS for disease costs and QALYS, ERG TTD for treatment costs and Aes for nivolumab; ERG PFS for nintedanib+docetaxel	Mod_A Mod_B Mod_E Mod_F	=(utility_PFS*((('Patient flow - 1'!$AZ14*IF(AND(Mod_B=0,Mod_E=0,Mod_F=0),1,0)))+('ERG health states'!AS11*IF(OR(Mod_B=1,Mod_E=1,Mod_F=1),1,0)))/52)*$C10+(utility_PD*((('Patient flow - 1'!$BA14*IF(AND(Mod_A=0,Mod_B=0,Mod_E=0,Mod_F=0),1,0))+('ERG health states'!BE11*IF(AND(Mod_A=1,Mod_B=0,Mod_D=0,Mod_E=0,Mod_F=0),1,0))+('ERG health states'!BA11*IF(AND(Mod_A=0,OR(Mod_B=1,Mod_E=1,Mod_F=1)),1,0))+('ERG health states'!AW11*IF(AND(Mod_A=1,OR(Mod_B=1,Mod_E=1,Mod_F=1)),1,0)))/52)*$C10)
	R11:R1049	R1) ERG OS R2) ERG PFS R5) Nivolumab vs. nintedanib+docetaxel: ERG TTD for nivolumab, ERG PFS for nintedanib+docetaxel R6) Nivolumab vs. nintedanib+docetaxel: ERG PFS for disease costs and QALYS, ERG TTD for treatment costs and Aes for nivolumab; ERG PFS for nintedanib+docetaxel	Mod_A Mod_B Mod_E Mod_F	=(utility_PFS*((('Patient flow - 1'!$AZ15*IF(AND(Mod_B=0,Mod_E=0,Mod_F=0),1,0)))+('ERG health states'!AS12*IF(OR(Mod_B=1,Mod_E=1,Mod_F=1),1,0)))/52)*$C11+(utility_PD*((('Patient flow - 1'!$BA15*IF(AND(Mod_A=0,Mod_B=0,Mod_E=0,Mod_F=0),1,0))+('ERG health states'!BE12*IF(AND(Mod_A=1,Mod_B=0,Mod_D=0,Mod_E=0,Mod_F=0),1,0))+('ERG health states'!BA12*IF(AND(Mod_A=0,OR(Mod_B=1,Mod_E=1,Mod_F=1)),1,0))+('ERG health states'!AW12*IF(AND(Mod_A=1,OR(Mod_B=1,Mod_E=1,Mod_F=1)),1,0)))/52)*$C11)+R10