NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Premeeting briefing

Nintedanib for previously treated locally advanced, metastatic or locally recurrent non-small cell lung cancer

This premeeting briefing presents:

- the key evidence and views submitted by the company, the consultees and their nominated clinical specialists and patient experts and
- the Evidence Review Group (ERG) report.

It highlights key issues for discussion at the first Appraisal Committee meeting and should be read with the full supporting documents for this appraisal.

Please note that this document includes information from the ERG before the company has checked the ERG report for factual inaccuracies.

Key issues for consideration

Clinical effectiveness

Comparators

Should erlotinib be considered a comparator?

- The comparators listed in the final scope issued by NICE were docetaxel alone and erlotinib alone. Both the company and the ERG do not consider erlotinib to be an appropriate comparator because:
  - NICE recommends erlotinib for first-line treatment of people with locally advanced or metastatic non-small-cell lung cancer (NSCLC) if they test positive for the epidermal growth factor receptor tyrosine kinase (EGFR-TK) mutation

- Approximately [redacted] [AiC] of patients in the LUME-Lung 1 trial ([redacted]) [AiC] are likely to have the EGFR-TK mutation
- The ERG noted that erlotinib is often used in clinical practice in England as a first-line treatment for patients with EGFR-positive disease, and it is unlikely that a patient would receive erlotinib again as a second-line treatment.
- The ERG and company both consider that patients fit enough (that is, those patients with Eastern Cooperative Oncology Group [ECOG] performance status [PS] 0 to 1) to receive nintedanib plus docetaxel would also be considered fit enough to receive docetaxel alone rather than erlotinib.

Trial design

- The LUME-Lung 1 trial included patients with all histologic types of non-small cell lung cancer. In the trial during randomisation, stratification factors included histology defined by squamous or non-squamous, but adenocarcinoma (which is non-squamous) was not included as a specific stratification factor. Does the LUME-Lung 1 study provide robust data for the subgroup of patients with adenocarcinoma (and on which the marketing authorisation is expected to be based)?

Generalisability

- Are the results from the LUME-Lung 1 trial generalisable to patients seen in clinical practice in England?
  - The trial included 27 countries including the UK.
  - The LUME-Lung 1 trial excluded patients with clinically significant pleural effusion or evidence of cavitary or necrotic tumours, significant cardiovascular disease, and anticoagulation (except low dose heparin) or antiplatelet therapy (except for chronic low-dose therapy with aspirin ≤325mg/day). This means that the patient population could have a better prognosis than patients in England and that the modelled absolute gains in survival could be longer.
  - In the LUME-Lung 1 trial, as first-line treatment only 18.8% of people received pemetrexed and the majority received platinum-based drugs. The ERG considered that most patients in England would have pemetrexed as first-line

treatment. The company did not include first-line treatment (other than bevacizumab) as a subgroup.

- The trial also excluded people who received docetaxel first-line. The ERG considered that docetaxel is rarely used first- line in England.

Proportional hazards assumption in LUME-Lung 1 trial

Are the hazard ratios for nintedanib plus docetaxel compared with docetaxel credible?

- The company reported progression-free and overall survival as hazard ratios derived from a Cox proportional hazards model. Curves that cross or converge violate the proportional hazards assumptions. The ERG commented that the trial data does not support the proportional hazards assumption (see appendix 7, page 131 of ERG report).

Mixed-treatment comparison

Are the results from the mixed treatment comparison comparing nintedanib plus docetaxel with erlotinib valid?

- The company performed a mixed treatment comparison to compare nintedanib plus docetaxel with erlotinib. The ERG raised a number of concerns:
  - The ERG considers that the proportional hazards assumption is not supported by the LUME-Lung 1 trial data for progression-free or overall survival. As the LUME-Lung 1 trial is the only trial providing evidence for nintedanib plus docetaxel, any comparison with this trial, will mean that any estimation of the relative effectiveness of nintedanib plus docetaxel compared with erlotinib (that is a calculated hazard ratio) will lack credibility and invalidate the comparison.
  - The trials within the mixed treatment comparison enrolled patients with varying baseline characteristics and were heterogeneous. Trials varied by age, EGFR mutation status, ECOG score, sex and whether patients had smoked.
  - The company assumed that docetaxel and pemetrexed were equally effective in the mixed-treatment comparison. The ERG was not aware of any evidence that supports this assumption specifically in an adenocarcinoma population.

Cost effectiveness

The ERG identifies that the main driver of cost-effectiveness is overall survival, but that the half-cycle correction, the calculation of costs in the stable disease state, and limit docetaxel cycles to 4 (as in UK practice) also change the cost-effectiveness estimates.

Extrapolation of trial results-key driver of cost effectiveness

Are the modelled survival extrapolations valid?

- The company used Kaplan-Meier survival curves for progression-free and overall survival from LUME-Lung 1 to extrapolate over a lifetime horizon by choosing parametric curves, and then replacing the Kaplan-Meier curves. In sensitivity analyses, the company added the chosen curve onto the end of the Kaplan Meier curves. The company took 2 approaches:
    - Joint models including data from both treatment groups with a term for treatment and using the same distributions for each arm.
    - Separate models using statistical models fitted to each randomised treatment arm separately and which may use different distributions for each arm.
- The company validated the extrapolation taking advice from a group of UK clinicians and against data from the British National Lung Cancer Audit (LUCADA) Data set and data from the American Surveillance, Epidemiology and End (SEER). The ERG could not assess whether this approach was valid because the company did not provide references to clarify which data it used (including a date of extracting the data, selection criteria and duration of follow-up).
- The ERG stated that because the evidence rests on only one trial, the company should make better use of existing data rather than replace it with a parametric function. The ERG notes that because of a survival effect, the data from a few people whom live the longest exert undue influence on the chosen extrapolation curve. When the ERG takes its approach of maximising the trial data, the ICERs rise. Which approach to survival modelling is better?

Half-cycle correction

The ERG commented that the company costed both nintedanib plus docetaxel and docetaxel alone using average number of patients receiving treatment across each cycle. The ERG commented that adjusting mid-cycle is not accurate for docetaxel or for nintedanib because patients receive treatment on the first day of a 3-week cycle. Is it appropriate to include a mid-cycle correction?

Calculation of costs in the stable disease state

The cost of care for patients who had finished active treatment, but remained in a stable condition differed between the model and the figures supplied to the company by clinicians. In the model the company assumed that these patients accrued the costs of receiving palliative nursing care every week and a bone scan every 3 weeks. In the ERG's opinion, this reflected an error which significantly reduces the care costs of patients in a stable condition after second-line treatment. Are the costs in the company's model underestimated?

Docetaxel cycles

The ERG noted that the company's' model followed the protocol used in the LUME-Lung 1 trial which allowed patients to have unlimited docetaxel treatment. The ERG explained that in the UK patients are restricted to 4 cycles of docetaxel because of unacceptable adverse events. Although the company's model allowed the number of cycles to be restricted, the ERG found an error which limited the cycles to 5 rather than to 4. How many docetaxel cycles should be included in the model?

Plausible ICER

- The base-case ICER from the company for the comparison of nintedanib plus docetaxel with docetaxel alone was £50,776 per QALY gained and the base-case ICER provided by the ERG in its exploratory analyses was £85,292 per QALY gained. What is the most plausible ICER for this comparison?
- The ICER from the company for the comparison of nintedanib plus docetaxel compared with erlotinib ranged from £27,008 per QALY gained (assuming a 0% discount on the list price of erlotinib) to £36,318 per QALY gained (assuming a 50% discount on the list price of erlotinib). The ICERs provided by the ERG from

its exploratory analyses for this comparison ranged from £28,307 per QALY gained (assuming a 0% discount on the list price of erlotinib) to £38,375 per QALY gained (assuming a 50% discount on the list price of erlotinib). Given the concerns raised by the company and ERG regarding the appropriateness of erlotinib as a comparator and the ERG's concerns about the robustness of the results from the mixed-treatment comparisons, are the ICERs presented for this comparison a suitable basis for decision making?

End-of-life

- The company stated that nintedanib meets NICE's requirements for end-of-life criteria.
    - The company stated that patients with advanced non-small cell lung cancer have a short life expectancy of less than 24 months. Using the extrapolated results from LUME-Lung 1 and the model outputs, the company considers the median overall survival of patients treated with docetaxel alone to be 10.23 months and the mean to be 15.96 months.
    - Both the company and ERG estimate the total eligible population in England for nintedanib plus docetaxel to be under 800.
    - The company states that:
        - nintedanib plus docetaxel will extend life by a mean of 3.96 months when compared with docetaxel alone. The company also suggests that the extension to overall survival when comparing with erlotinib will be a mean of 5.16 months.
    - The ERG calculated that the **mean** extension to overall survival would be 3.05 months for the base-case scenario when comparing nintedanib plus docetaxel to docetaxel alone. The ERG says that this may overestimate mean survival gain because of the trial population, on which the estimate is based, is likely to live longer than patients in clinical practice in England at a similar stage of disease.
        - The ERG was only able to carry out a partial comparison of nintedanib plus docetaxel with erlotinib and was unable to calculate a mean overall survival.

1 Remit and decision problems

1.1 The remit from the Department of Health for this appraisal was to appraise the clinical and cost effectiveness of nintedanib within its licensed indication for previously treated locally advanced or metastatic non-small cell lung cancer (NSCLC). The Department of Health revised the remit in November 2014 to reflect the positive Committee for Medicinal Products for Human Use (CHMP) opinion related to nintedanib in combination with docetaxel. The updated remit was to appraise the clinical and cost effectiveness of nintedanib within its licensed indication for previously treated locally advanced, metastatic or locally recurrent non-small cell lung cancer of adenocarcinoma tumour histology.

Table 1. Decision problem

	Final scope issued by NICE	Decision problem addressed in the submission	Comments from the company	Comments from the ERG
Pop.	Adults with locally advanced or metastatic non-small cell lung cancer that has progressed following prior chemotherapy	Patients with locally advanced, metastatic or recurrent non-small cell lung cancer of adenocarcinoma tumour histology after first-line chemotherapy		The ERG noted that the population addressed in the company submission differed to the population specified in the scope. The decision problem addressed by the company is in line with the CHMP positive opinion and the anticipated full marketing authorisation for nintedanib.
Int.	Nintedanib in combination with docetaxel	As in final scope		As per final scope
Com.	docetaxel monotherapy erlotinib	Primary analysis: docetaxel monotherapy Secondary	The company states that erlotinib is not a relevant comparator to nintedanib plus	The ERG agreed with the company that docetaxel monotherapy should be the

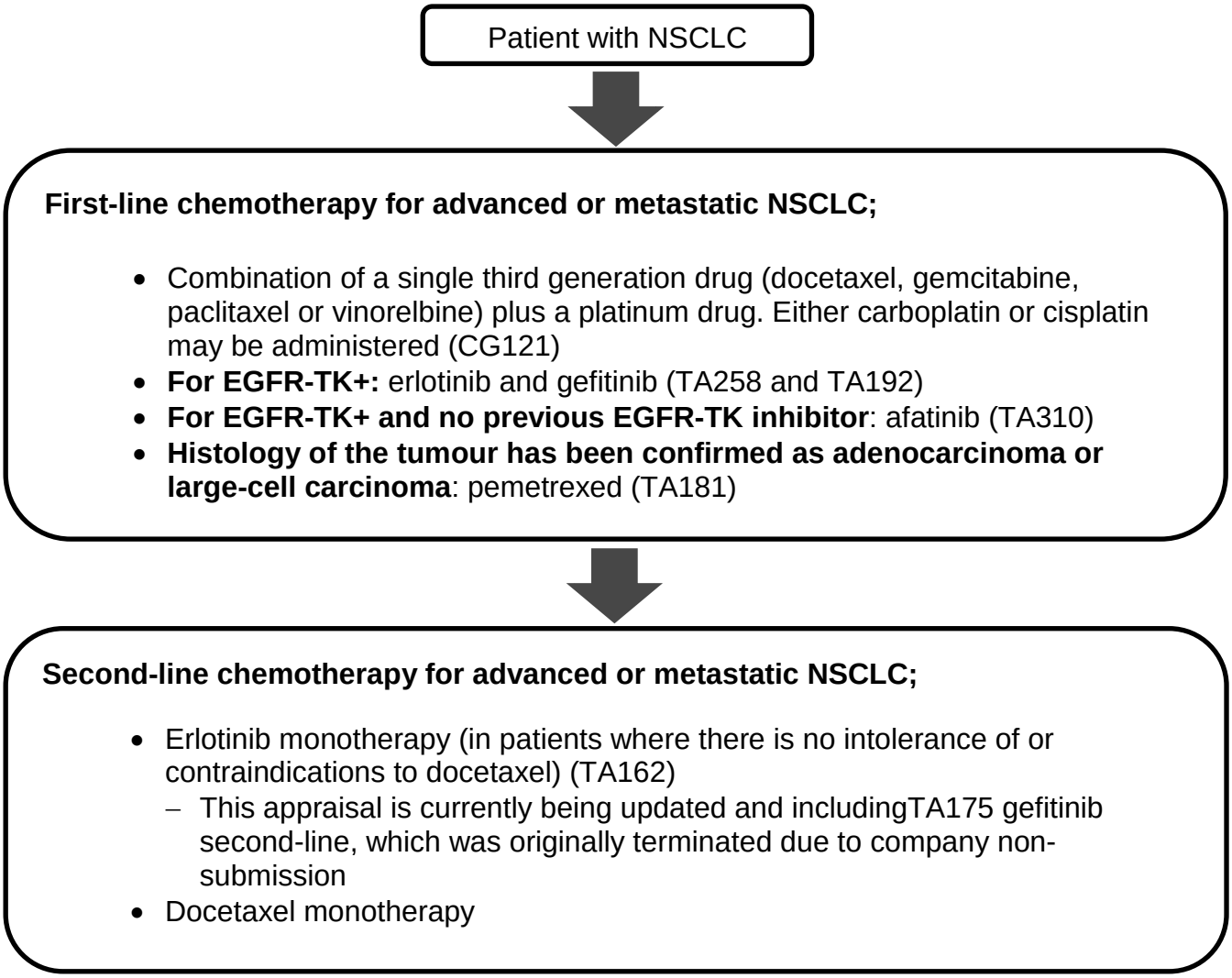
		analysis: erlotinib monotherapy As in final scope	docetaxel and this is only considered a comparator by the company for secondary analyses. The company considered that patients fit enough for treatment with docetaxel would receive docetaxel rather than erlotinib.	comparator for the primary analysis and erlotinib the comparator in the secondary analysis. The ERG considered that the characteristics of patients who are considered suitable for second-line erlotinib treatment are different from those who are considered suitable for docetaxel treatment. It also stated that erlotinib is likely to be preferred when patients have a poorer European Cooperative Oncology Group score and/or have EGFR-positive tumours, docetaxel is the most appropriate comparator to nintedanib plus docetaxel in the second-line setting.
Out.	Overall survival Progression free survival Response rates Adverse effects of treatment Health-related quality of life	As in final scope		The ERG considered that the company had included all the outcomes in the final scope

2 The technology and the treatment pathway

2.1 Nintedanib is a small molecule triple angiokinase inhibitor targeting three receptor classes that have a key role in angiogenesis and tumour growth:

vascular endothelial growth factor receptors (VEGFR), fibroblast growth factor receptors (FGFR) and platelet-derived growth factor receptors (PDGFR) α and β. VEGF and its receptor VEGFR-2 are crucial for the formation of new tumour vessels. Nintedanib interferes with steps in the angiogenesis signalling cascade impacting tumour growth and spread.

Figure 1. Treatment pathway for patient with non-small cell lung cancer



2.2 NICE technology appraisal guidance 162 (Erlotinib for the treatment of non-small-cell lung cancer) recommends erlotinib within its licensed indication, as an alternative to docetaxel as a second-line treatment option for patients with non-small-cell lung cancer (NSCLC) only on the basis that it is provided by the manufacturer at an overall treatment cost

(including administration, adverse events and monitoring costs) equal to that of docetaxel. Erlotinib is not recommended for the second-line treatment of locally advanced or metastatic NSCLC in patients for whom docetaxel is unsuitable (that is, who are intolerant to or have contraindications to docetaxel) or for third-line treatment after docetaxel therapy.

2.3 NICE technology appraisal guidance 162 is currently being updated. Publication of the reviewed guidance is expected in December 2014.

Table 2. Technology

	Nintedanib plus docetaxel	Docetaxel (generic)	Erlotinib
Marketing authorisation	Received in December 2014. [Nintedanib] in combination with docetaxel for the treatment of adult patients with locally advanced, metastatic or locally recurrent non-small cell lung cancer (NSCLC) of adenocarcinoma tumour histology after first-line chemotherapy'	Indicated for the treatment of patients with locally advanced or metastatic non-small cell lung cancer after failure of prior chemotherapy	Indicated for the treatment of patients with locally advanced or metastatic NSCLC after failure of at least one prior chemotherapy regimen
Administration method	Oral	Intravenous	Oral
Cost	200 mg twice daily. Dose adjustments to 150 mg or 100 mg twice daily are permitted in patients who experience adverse events. £2,151 [taken from company submission]	Administered on day 1 of a 21 day cycle at a dose of 75 mg/m^2. If required doses can be reduced to 60 mg/m^2. Docetaxel 10 mg/mL 2-mL vial = £138.33 8-mL vial = £454.53 16-mL vial = £1069.50; Docetaxel 20 mg/mL 1-mL vial = £160.00 4-mL vial = £530.00 7-mL vial = £900.00 [BNF 68, September 2014]	150 mg daily. Erlotinib 25 mg 30-tab pack = £378.33; Erlotinib 100 mg, 30-tab pack = £1324.14; Erlotinib 150 mg, 30-tab pack = £1631.53 Erlotinib has a confidential patient access scheme in place.

See summary of product characteristics for details on adverse reactions and contraindications.

3 Comments from consultees

3.1 A patient and carer organisation suggested that patients with locally advanced or metastatic non-small cell lung cancer have few second-line treatments, with only docetaxel and erlotinib being available and there is no cure. Patients consider symptoms such as breathlessness difficult to manage and the patient and carer organisation suggests that a treatment, such as nintedanib, with anti-tumour activity may provide the best option for symptom relief. In the anecdotal patient experiences reported to the organisation, patients report side effects associated with docetaxel as an issue.

3.2 The patient and carer organisation explained that improving quality of life and even small extensions in duration of life are of considerable significance to patients and their families. It also suggested that the availability of new treatments offers hope to patients. For this organisation, prognosis is very poor so even relatively small benefit in extension to life can be disproportionately large for these patients.

3.3 The patient and carer organisation suggested that "inverse weighting for duration of life" should be considered when appraising nintedanib. When considering the cost of treatment, the patient and carer organisation does not consider it appropriate to give the same weighting to the final 6 months of life as to all other 6 months of life. The patient and carer organisation also suggested that the addition of nintedanib to docetaxel seemed to be well tolerated. Nintedanib being available as an oral treatment was also seen positively by the organisation.

4 Clinical-effectiveness evidence

Overview of the clinical trial

4.1 The main evidence in the company's submission comes from the LUME-Lung 1 trial (n=1,314) which was a phase-III multi-centre, double-blind, randomised (1:1) controlled trial comparing nintedanib plus docetaxel with placebo plus docetaxel. The trial was carried out in 211centres in 27 countries (including the UK). All patients in the trial, in whom first-line chemotherapy had failed, had either locally advanced, metastatic (94.2% of at randomisation) or locally recurrent non-small cell lung cancer. Of the 1, 314 patients randomised, 759 patients had non-squamous cell carcinoma (of whom 658 had adenocarcinoma) and 555 had squamous cell carcinoma. The company stratified the randomisation by 4 variables: European Cooperative Oncology Group [ECOG] score (0 or 1), previous bevacizumab treatment (yes or no), presence of brain metastases (yes or no) and histology (squamous or non-squamous). Of note, randomisation did not further stratify non-squamous histology into the presence or absence of adenocarcinoma. Patients in the nintedanib group were randomised to nintedanib (200 mg) twice daily, on day 2 and 21 of a 21-day cycle, plus docetaxel (75 mg/m^2) on day 1 of the 21-day cycle. On the possibility of adverse effects, the trial design specified reducing the dose of nintedanib from 200 mg to 150 mg twice daily and then again from 150 mg to 100 mg twice daily and reducing the dose of docetaxel from 75 mg/m^2 to 60 mg/m^2. Patients in the placebo group were randomised to placebo twice daily on day 2 and 21 of a 21-day cycle, and docetaxel dosing as in the nintedanib group. In this group reducing docetaxel (from 75 mg/m^2 to 60 mg/m^2) was permitted if adverse events occurred. Patients stopped treatment when their disease progressed or if they experienced unacceptable adverse events. Patients in the nintedanib group who received at least 4 cycles of nintedanib plus docetaxel could thereafter receive nintedanib alone. The trial investigators followed patients every 6 weeks before disease progression and every 6 to

8 weeks after disease progression until the patient died or was lost to follow up.

4.2 The primary efficacy analyses were intention-to-treat. Progression-free survival, a radiologic measure, was the primary outcome in the LUME-Lung 1 trial and was defined as time from date of randomisation to date of disease progression, or to date of death, whichever occurred earlier. Progression-free survival was determined by a central independent review by radiologists using the modified Response Evaluation Criteria in Solid Tumours. The primary progression free-analysis was to be carried out when 713 patients had received a centrally assessed progression-free event (cut off November 2010) to detect a hazard ratio (HR) of 0.78 with 90% statistical power. The median follow-up in the trial was 7.1 months (interquartile range: 3.8-11.0 months) at the time of the primary progression-free survival analysis (November 2010). The trial was unblinded in July 2011, during the database lock for the primary endpoint analysis of centrally assessed progression-free survival and interim overall survival. See Figure 4 on page 81 of company's submission for timing of amendments to the trial.

4.3 The key secondary endpoint in the LUME-Lung 1 trial was overall survival. Overall survival was defined as the time from date of randomisation to date of death (irrespective of cause of death). The company calculated that 1,151 patients would need to have died to permit investigators to detect an 18% increase in median overall survival or a hazard ratio of 0.85. The statistical plan called for a single 'interim look' at the same time as the final analysis for progression-free survival which, to be considered statistically significant, had to achieve a p value of <0.00043, whereas the final analysis had to achieve a p value of <0.04984. The median follow-up was 31.7 months (interquartile range: 27.8-36.1 months) at the time of the final overall survival analysis (February 2013). See Figure 4 on page 81 of the company's submission for timing of amendments to the trial.

4.4 The company used another trial, LUME-Lung 2, which compared nintedanib or matching placebo in combination with pemetrexed and which was stopped for futility, as 'hypothesis generating' and changed the analysis plan of LUME-Lung 1 to reflect this. From analysis of LUME-Lung 2, the company identified 'time since start of first-line therapy', and adenocarcinoma histology as potentially interactive with trial treatment in the LUME-Lung 1 trial. The statistical analysis of LUME-Lung 1 was amended before database lockdown for the final overall survival analysis. The company tested overall survival in a sequential fashion (a hierarchical overall survival statistical analysis): first patients with adenocarcinoma whose disease had progressed within 9 months of starting first-line therapy, followed by all patients with adenocarcinoma, and finally the overall trial population.

4.5 The focus of the company's submission to NICE was on patients with adenocarcinoma as this was the population that the company expected the marketing authorisation for nintedanib to specify (this was confirmed in the CHMP's positive opinion [see Table 2]). The company did present the results of the primary progression-free survival analysis for the overall trial population and overall survival for patients with adenocarcinoma whose disease had progressed with 9 months of starting treatment wherever necessary because these populations comprised the hierarchical overall statistical analysis.

4.6 The company considered the baseline characteristics of patients in LUME-Lung 1 with adenocarcinoma, such as sex, age, race, smoking status and ECOG score, to be similar between the treatment groups, and to patients diagnosed with adenocarcinoma. Overall, 62.5% of patients with adenocarcinoma were men, and the mean age was 58.5 (standard deviation 10.1) years. The majority of patients were white (76.9%), 70.4% of patients had an ECOG performance status of 1, and 7.4% of patients had brain metastases at baseline. 95.9% of patients with adenocarcinoma had received platinum-based therapy first-line and 6.8% of patients had

received bevacizumab treatment before entry into the trial. Data on EGRF mutation was not routinely collected in the LUME-Lung 1. However, during the clarification process, the company stated that it had collected these data retrospectively for a sample ([redacted]) [AiC] of patients and demonstrated that [redacted] [AiC]. For further details of the baseline characteristics of patients with adenocarcinoma, see Table 10 page 61 of the company's submission.

ERG comments

4.7 The ERG considered that the LUME-Lung 1 trial was well designed and conducted with a low risk of bias and reasonably mature data. The ERG noted that the trial included an unknown number of trial sites in the UK.

4.8 The ERG commented that the eligibility criteria for the trial meant that the trial results were not generalisable to England. The trial excluded patients with clinically significant pleural effusion, or evidence of cavitary or necrotic tumours, with significant coronary disease, or on anticoagulation (except low dose heparin) or antiplatelet therapy. The ERG considered the trial population to have a better prognosis than patients seen in clinical practice in England.

4.9 The ERG highlighted differences in the proportion of patients in the trial who had received first-line pemetrexed and third-line treatments compared with those treated in England. The ERG commented that more patients in England receive pemetrexed first-line than in LUME-Lung 1 (18.8%). The ERG also commented that patients in England are less likely to receive third-line treatment ([redacted]) [CiC] relative to the trial (55.8%).

4.10 The ERG noted that the company limited in submission to patients with adenocarcinoma, even though only approximately 50% of the patients in the LUME-lung 1 trial had adenocarcinoma which itself was neither a stratification factor at randomisation nor a pre-defined subgroup. The ERG

commented that patients with adenocarcinoma constituted the majority of patients with non-squamous cell carcinoma in the trial, and that non-squamous cell carcinoma was a stratification factor. The ERG commented that among patients with adenocarcinoma, the baseline characteristics were well balanced across the 2 groups in the trial suggesting that the analyses were acceptable.

Clinical trial results

4.11 The results for progression-free and overall survival for the adenocarcinoma population in LUME-Lung 1 are presented in Table 3 and figures 2, 3 and 4. The results for progression-free and overall survival for the overall trial population in LUME-Lung 1 are presented in Table 4.

Table 3. Progression–free and overall survival results for the adenocarcinoma population in LUME-Lung 1 (cut off November 2010 and February 2013) (see Table 18 and 19, pages 87 and 89 of the company's submission)

Outcome		**LUME-Lung 1**		
		Nintedanib plus docetaxel)	**Placebo plus docetaxel**	**Hazard ratio (95% confidence intervals[CI])**
Progression-free survival (central independent review)	**Primary analysis at November 2010 (median 7.1 month follow-up)**	Median 4.0 months	Median 2.8 months	HR 0.77 (95% CI 0.62 to 0.96)
	Final analysis at February 2013 (median 31.7 month follow-up)	Median 4.2 months	Median 2.8 months	HR 0.84 (95% CI 0.71 to 1.00)
Overall survival (final analysis at February 2013)		Median 12.6 months	Median 10.3 months	HR 0.83 (95% CI 0.70 to 0.99)

Figure 2. Kaplan-Meier curves for progression-free survival in the adenocarcinoma population in the LUME-Lung1 trial (primary analysis [November 2010], central review)

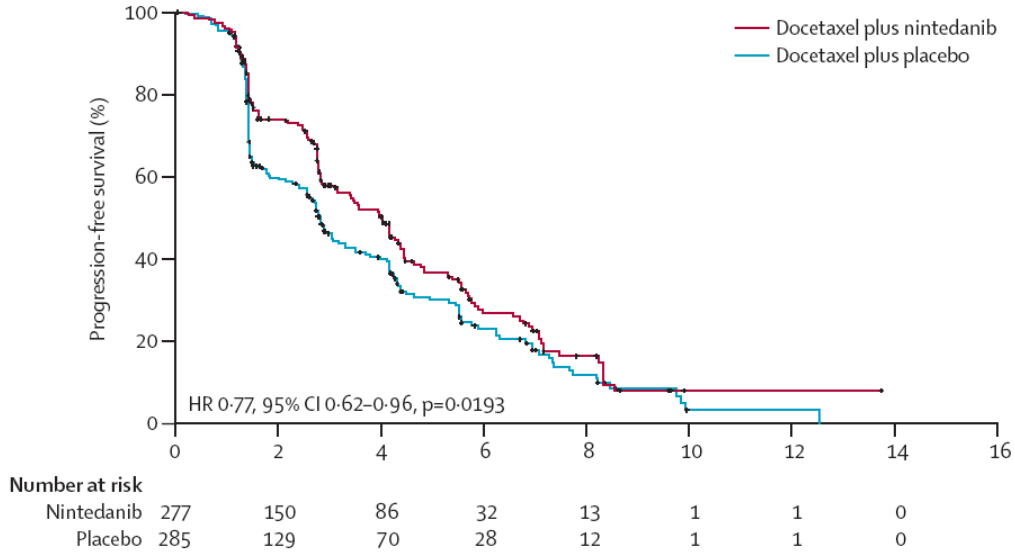


Figure 3. Kaplan-Meier curves for progression-free survival in the adenocarcinoma population in the LUME-Lung 1 trial (follow-up analysis, February 2013)

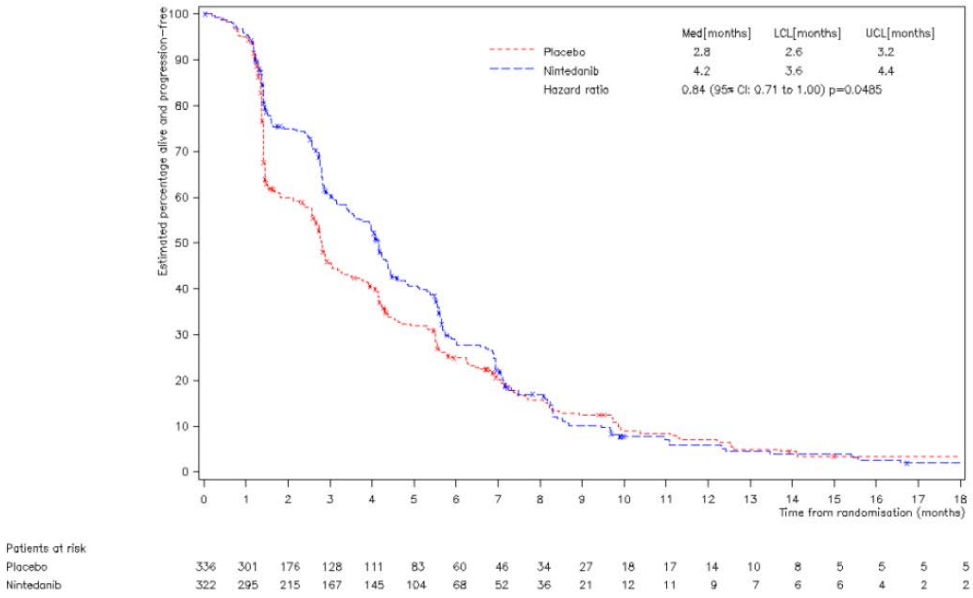


Figure 4. Kaplan-Meier curves for overall survival in the adenocarcinoma population in the LUME-Lung 1 trial (follow-up analysis, February 2013)

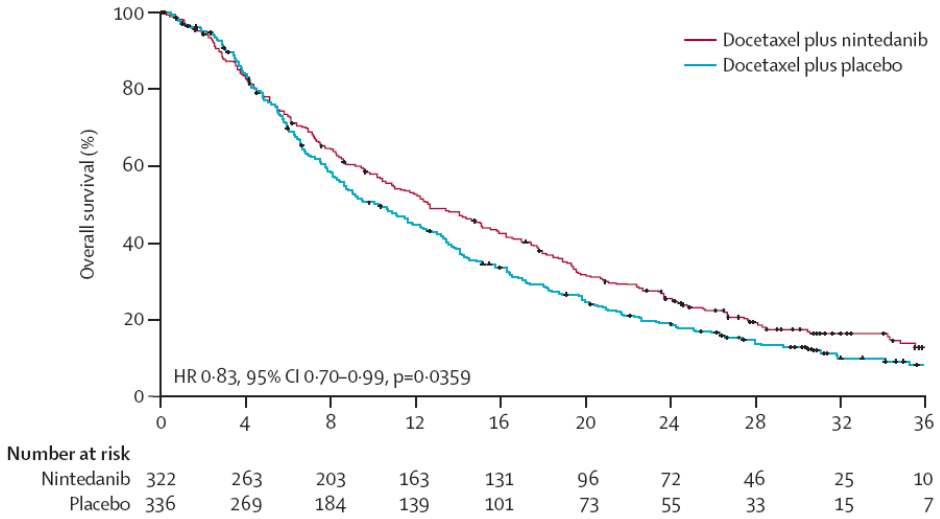


Table 4. Progression –free survival results for the overall trial population in LUME-Lung 1 (cut off November 2010 and February 2013) (see Table 18 and 19, pages 87 and 89 of the company submission)

Outcome		LUME-Lung 1		
		Nintedanib plus docetaxel	**Placebo plus docetaxel**	**Hazard ratio compared with placebo (95% confidence interval [CI])**
Progression-free survival (central independent review)	**Primary analysis at November 2010 (median 7.1 month follow-up)**	Median 3.4 months	Median 2.7 months	0.79 (0.68 to 0.92) p=0.0019
	Final analysis at February 2013 (median 31.7 month follow-up)	Median 3.5 months	Median 2.7 months	0.85 (0.75 to 0.96) p=0.007

4.12 Subgroup analyses were performed at the time of the final overall survival analysis (February 2013). The majority of pre-specified and post-hoc progression-free survival subgroup analyses showed the effect of

nintedanib plus docetaxel to be consistent with the treatment benefit observed in the primary analysis. The only exceptions to this were 2 subgroups (i) more than 9 months since start of first-line treatment and (ii) Asian region where there was a trend in favour of placebo plus docetaxel. The results of the pre-specified and post-hoc overall survival subgroup analyses also showed treatment effects in favour of nintedanib plus docetaxel, supporting the findings of the primary analysis. The only exception to this were 2 baseline characteristics: (i) presence of brain metastases and (ii) below stage lllB disease at diagnoses.

4.13 Health-related quality of life in LUME-Lung 1 was measured at the screening visit, at 21-day intervals during treatment, at the end of treatment and at the first follow-up visit. The investigators used 3 questionnaires: EQ-5D, European Organisation for Research and Treatment of Cancer (EORTC) Quality of Life Questionnaire (EORTC QLQ-C30) and EORTC lung cancer specific supplementary module (EORTC QLQ-LC13). In the LUME-Lung 1 trial, investigators found no differences in global health status, quality of life or self-reported health related quality of life reported for coughing, dyspnoea or pain between the nintedanib plus docetaxel and docetaxel alone groups. Diarrhoea was significantly worse in those treated with nintedanib plus docetaxel.

ERG comments

4.14 The company reported progression-free and overall survival as hazard ratios from Cox proportional hazards models. The ERG commented that the progression-free survival curve for the LUME-Lung 1 trial groups diverge after 6 weeks and then converge and cross after approximately 1 year, thereby violating the proportional hazards assumptions and suggesting that advantage of nintedanib is limited to the first year of treatment. The ERG suggested that the proportional hazards model was not appropriate and that the progression-free and overall survival results should be treated with caution. The ERG noted that the company could

have used alternative approaches to better reflect relative efficacy in the data.

Mixed-treatment comparisons

4.15 As there are no head to head trials comparing nintedanib plus docetaxel with erlotinib, the company carried out a systematic review and mixed treatment comparison. The company identified 9 studies to include in its mixed treatment comparison. The company carried out 3 types of analyses:

- The base-case analyses excluded trials where more than 20% of patients had EGFR positive adenocarcinoma, and trials where chemotherapy was a single comparator. The company included 4 trials in its base-case analyses (see Figure 5 and Table 5).
- Scenario analyses where the company assumed that docetaxel and pemetrexed were equally effective. The company stated that that it used this assumption to allow as many treatments to be compared with nintedanib plus docetaxel as possible. The company included 4 trials in this scenario analyses (see Table 5). For the scenario analysis diagram see Figure 21 on page 113 on the company's submission.
- Sensitivity analyses included trials in which greater than 20% of patients had EGRF-positive adenocarcinoma along with;
  - The trials included in the base-case (sensitivity analysis i)
  - The trials included in the scenario analyses (sensitivity analysis ii).

  The company included 9 studies in the mixed treatment comparison although included only 8 in any given analysis (see Table 5).

Figure 5. Network diagram for mixed treatment comparison base-case analyses assuming equivalence of docetaxel and pemetrexed (see Figure 1, page 39 of ERG report)

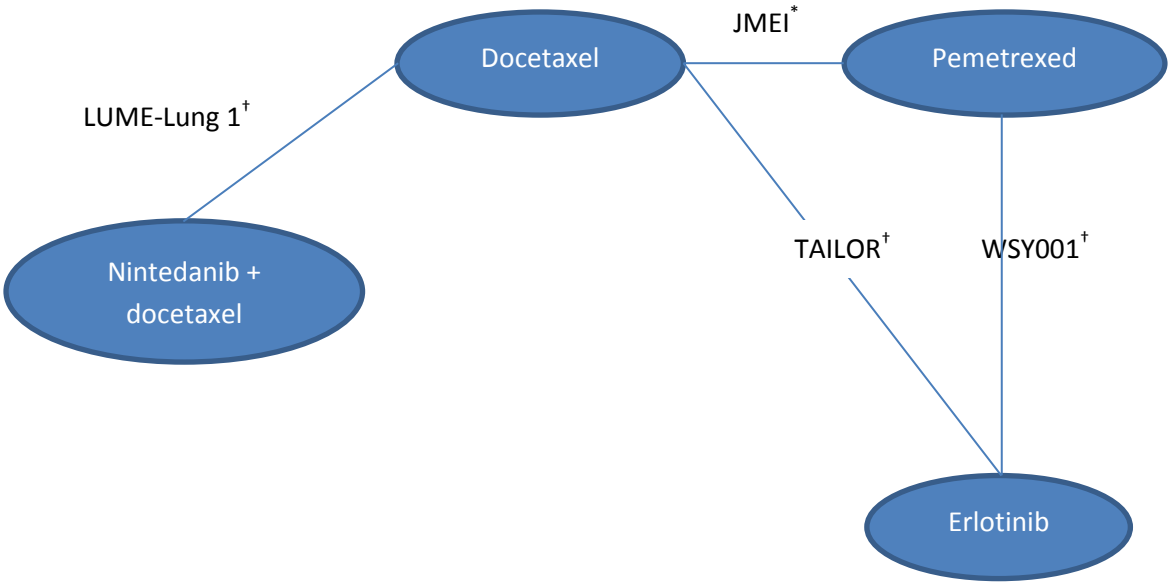


Table 5. Trials included in the mixed treatment comparisons identified by the company (see Table 10, page 40 of ERG report)

Trial name	Intervention	Comparator	Analyses included in
LUME-Lung 1	Nintedanib + docetaxel	Placebo + docetaxel	Base-case, scenario and sensitivity
TAILOR	Erlotinib	Docetaxel	Base-case, scenario and sensitivity
WSY001	Erlotinib	Pemetrexed	Base-case, scenario and sensitivity
JMEI	Pemetrexed	Docetaxel	Base-case and sensitivity
TITAN	Erlotinib	Chemotherapy (docetaxel or pemetrexed)	Scenario and sensitivity
GEF-ERL	Gefitinib	Erlotinib	Sensitivity
KCSG-LU08-01	Gefitinib	Pemetrexed	Sensitivity
V-15-32	Gefitinib	Docetaxel	Sensitivity
S103	Pemetrexed + erlotinib	Pemetrexed or erlotinib	Sensitivity

4.16 The company explained that the rationale for excluding patients with EGFR-positive adenocarcinoma from all but the sensitivity analyses was to enable a comparison between nintedanib plus docetaxel and other tyrosine kinase inhibitors (TKIs) in a population similar to the patient population in LUME-Lung 1. The company analysed the data using fixed- and random-effects Bayesian mixed treatment meta-analyses.

4.17 The baseline characteristics of patients included in the base-case, scenario and sensitivity analyses are provided in Table 6 and Table 7.

Table 6. Adenocarcinoma patient characteristics of trials included in only the MTC base-case and scenario analyses (see Tables 12 and 14, pages 45, 46 and 50 of ERG report).

Trial and arm	Location	No. at baseline	Adenocarcinoma		Age (years)	Wild-type mutations (EGFR-negative) (%)	ECOG PS 0 to 1 (%)	Female (%)	Never smokers (%)
			%	N					
LUME-Lung 1	Europe, Asia,	1314	50.1	658			100.0		
Nintedanib + docetaxel	South Africa	655	49.2	322	Median: 60 Range: 53 to 67	[redacted] [AiC] [redacted] [AiC]	100.0	27.3	35.7*
Placebo + docetaxel		659	51.0	336	Median: 60 Range: 54 to 66	[redacted] [AiC] [redacted] [AiC]	100.0	27.3	34.2*
TAILOR	Italy	219	69.4	152			92.7		
Erlotinib		109	63.3	69	Median: 66 Range: 40 to 81	100	93.6	29.4	17.4
Docetaxel		110	75.5	83	Median: 67 Range: 35 to 83	100	91.7	33.6	27.2
WSY001	China	123	100	123			94.3		
Erlotinib		61	100	61	Median: 54.3 Range: 30 to 74	100	93.4	34.4	24.6
Pemetrexed		62	100	62	Median: 55.1 Range: 33 to 75	100	95.2	37.1	27.4
JMEI	Not	571	52.9	302			86.8*		
Pemetrexed	Reported	283	55.8	158	Median: 57.4* Range not reported	Not reported	84.8*	39.2*	Not reported
Docetaxel		288	50.0	144	Median: 56.7* Range not reported	Not reported	88.9*	34.0*	Not reported
TITAN	International	424	49.5	201			80.0		
Erlotinib		203	47.3	96	Median: 59 Range: 36 to 80	36.9 Indeterminate: 15.8 Missing: 43.3	80.8	20.7	14.8

Trial and arm	Location	No. at baseline	Adenocarcinoma		Age (years)	Wild-type mutations (EGFR-negative) (%)	ECOG PS 0 to 1 (%)	Female (%)	Never smokers (%)
			%	N					
Chemotherapy		221		51.6	114		Median: 59 Range: 22 to 79	33.5 Indeterminate: 16.3 Missing: 45.7	79.2

Table 7. Patient characteristics of trials included in only the MTC sensitivity analyses (see Tables 13 and 15, pages 47 and 51 of ERG report).

Trial and arm	Location	Number at baseline	Adenocarcinoma		Age (years)	Wild-type mutations (EGFR-negative) (%)	ECOG PS 0 to 1 (%)	Female (%)	Never smokers (%)
			%	N					
GEF-ERL	South Korea	96	90.6	87			85.4		
Gefitinib		48	91.7	44	Median: 60 Range: 37 to 83	25.0 Missing: 56.3	85.4	85.4	91.7
Erlotinib		48	89.6	43	Median: 56 Range: 32 to 81	41.7 Missing: 41.7	85.4	85.4	95.8
KCSG-LU08-01	Korea	135[†]	100.0	135			91.1		
Gefitinib		68[†]	100.0	68	Median: 58 Range: 40 to 77	22.1 Missing: 50.0	91.2	85.3	100.0
Pemetrexed		67[†]	100.0	67	Median: 64 Range: 30 to 78	23.9 Missing: 44.8	91.0	85.1	100.0
V-15-32	Japan	489[¥]	77.7	380		5.3 Missing: 88.3	95.7		
Gefitinib		244[¥]	78.4	191	≤64 years: 56.3		95.5	38.4	29.0
Docetaxel		239[¥]	77.0	184	≤64 years: 55.3		95.9	38.1	35.7
S103	Not reported	240	93.8	225		7.9	92.9		

Trial and arm	Location	Number at baseline	Adenocarcinoma		Age (years)	Wild-type mutations (EGFR-negative) (%)	ECOG PS 0 to 1 (%)	Female (%)	Never smokers (%)
			%	N					
						Missing: 82.1			
Erlotinib + pemetrexed		78	92.3	72	Median: 55.8 Range not reported		91.0	74.4	100.0
Erlotinib		82	92.7	76	Median: 53.9 Range not reported		92.7	65.9	100.0
Pemetrexed		80	96.3	77	Median: 55.9 Range not reported		95.0	56.3	100.0

4.18 The results from the mixed treatment comparison base-case analyses for overall and progression-free survival are provided in tables 8 and 9. Results from the Bucher indirect comparisons and the scenario and sensitivity analyses supported the findings from the mixed-treatment comparisons. Table 10 provides a summary of the overall response rate from the mixed treatment comparison base case.

Table 8. Summary of overall survival findings from mixed treatment comparison base-case analysis (taken from ERG report Table 18, page 56).

Treatment	Hazard Ratio (95% CI) to fixed-effects
Nintedanib + docetaxel vs. docetaxel	
Result from MTC	0.83 (0.70 to 0.99)
Result from Bucher indirect comparison	Not applicable
Nintedanib + docetaxel vs. pemetrexed	
Result from MTC	0.82 (0.60 to 1.11)
Result from Bucher indirect comparison	0.90 (0.65 to 1.26)
Nintedanib + docetaxel vs. erlotinib	
Result from MTC	0.64 (0.46 to 0.90)
Result from Bucher indirect comparison	0.56 (0.38 to 0.82)
Deviance information criterion	0.4095

Table 9. Summary of the progression-free survival findings from mixed treatment comparison base-case analysis (taken from ERG report Table 20, page 57)

Treatment	Hazard Ratio (95% CI) to fixed-effects
Nintedanib + docetaxel vs. docetaxel	
Result from MTC	0.77 (0.62 to 0.96)
Result from Bucher indirect comparison	Not applicable
Nintedanib + docetaxel vs. pemetrexed	
Result from MTC	0.84 (0.61 to 1.15)
Result from Bucher indirect comparison	0.93 (0.67 to 1.29)
Nintedanib + docetaxel vs. erlotinib	
Result from MTC	0.70 (0.50 to 1.00)
Result from Bucher indirect comparison	0.58 (0.39 to 0.87)
Deviance information criterion	1.568

Table 10. Summary of the base-case overall response rate from the mixed treatment (taken from ERG report Table 22, page 58)

Treatment	Hazard Ratio (95% CI) to fixed-effects
Nintedanib + docetaxel vs. docetaxel	
Result from MTC	1.33 (0.61 to 2.95)
Result from Bucher indirect comparison	Not applicable
Nintedanib + docetaxel vs. pemetrexed	
Result from MTC	0.98 (0.33 to 2.84)
Result from Bucher indirect comparison	0.98 (0.34 to 2.83)
Nintedanib + docetaxel vs. erlotinib	
Result from MTC	0.33 (0.07 to 1.56)
Result from Bucher indirect comparison	Not applicable
Deviance information criterion	37.47

4.19 In the sensitivity analysis (i) of the mixed treatment comparison for overall response rate, nintedanib plus docetaxel was inferior to erlotinib, gefitinib and erlotinib plus pemetrexed using a fixed-effects or a random-effects model. Sensitivity analysis (ii) found nintedanib plus docetaxel was not significantly different from chemotherapy (docetaxel or pemetrexed) or erlotinib, but was significantly inferior to gefitinib and erlotinib plus pemetrexed.

ERG comments

4.20 The ERG did not identify any additional trials not included by the company. The ERG also considered that conducting Bucher indirect comparisons was a suitable approach to assessing consistency within the network and the reliability of the results.

4.21 The ERG considered it inappropriate to undertake mixed-treatment comparisons for the following reasons:

- Erlotinib is not an appropriate comparator for the population of patients for whom nintedanib plus docetaxel will be considered appropriate (that is those patients with an ECOG status of 0-1, and who are EGFR-TK mutation-negative, see Table 1).

- LUME-Lung 1 is the only trial in which any patients received pemetrexed as a first-line treatment and pemetrexed is used as a first-line treatment in clinical practice in England.
- The data from LUME-Lung 1 violate the proportional hazards assumption for progression-free or overall survival. As LUME-Lung 1 alone provides evidence for nintedanib plus docetaxel, any comparison with this trial will be associated with uncertainty.
- Differences in trial and patient characteristics (see Tables 6 and 7) suggest there is heterogeneity across trials in the mixed treatment comparison.

Adverse effects of treatment

4.22 In LUME-Lung 1, diarrhoea (43.4% compared with 24.6%), nausea (28.4% compared with 17.7%) and vomiting (19.4% compared with 12.3%) occurred more often with nintedanib plus docetaxel then with docetaxel alone. Deaths from adverse events, not attributed to disease progression, were more common in the nintedanib plus docetaxel (6.3%) than the placebo plus docetaxel groups (2.4%) although there was a longer median duration of both the nintedanib and docetaxel treatments (4.2 months and 5 cycles) than with the placebo plus docetaxel treatments (3.0 months and 4 cycles). The number of grade 3 or greater adverse events and grade 3 or greater serious adverse events were greater in the nintedanib plus docetaxel group (75.9% and 31.3%) than in the placebo plus docetaxel group (68.5% and 26.6%).

4.23 To compare adverse effects of nintedanib with chemotherapeutic regimens other than docetaxel, the company compiled data on fatigue, nausea and diarrhoea. These were the only safety outcomes reported in a consistent format in more than 1 trial. The company also stated that because few trials reported these outcomes, and because of the low incidence of adverse events, it compared nintedanib plus docetaxel with other treatments using the sensitivity analysis where the company assumed docetaxel and pemetrexed were equally effective (see sections 4.15 and 4.27). In the mixed-treatment comparison of adverse events

LUME-Lung 1 did not connect with the other studies. The results suggested that patients receiving nintedanib plus docetaxel were significantly more likely to develop any grade diarrhoea compared with docetaxel or pemetrexed, but not compared with erlotinib. Patients receiving nintedanib plus docetaxel were significantly more likely to develop any grade nausea compared with docetaxel/pemetrexed or erlotinib. The risk of fatigue was similar for all comparisons.

ERG's comments

4.24 The ERG noted that in LUME-Lung 1 trial the median number of cycles in the nintedanib plus docetaxel treatment arm (5) was higher than the maximum number of cycles provided in practice in England (4). The ERG did not consider that the greater number of cycles of docetaxel received by patients treated with nintedanib was likely to have been a confounder since, as reported by the National Confidential Enquiry into Patient Outcome and Death, most patients with life threatening toxicity tend to experience fatal adverse events during the first cycle of treatment.

4.25 The ERG also noted that nintedanib plus docetaxel caused more grade 3 or greater adverse and serious adverse events than docetaxel alone. The ERG explained that grade 3 adverse events can lead to drug discontinuation and hospitalisation, and grade 2 adverse events, in particular vomiting and diarrhoea, can lead to lowering of the nintedanib dose.

4.26 Regarding death as an adverse event, the ERG considered that the number of deaths related to adverse events was small, but agreed with the company it should monitor adverse event related deaths in the future. Patients in the nintedanib plus docetaxel group were more likely to die in the progressive disease state than patients in the control group. The ERG considered that this could relate to being treated longer, but then rejected this, noting that most patient with life threatening toxicity tend to

experience fatal adverse events in the first cycle of treatment which was similar between the 2 treatment groups (98.1% compared with 98.7%).

4.27 The ERG disagreed with the company's assumption that pemetrexed and docetaxel were equally tolerable. The ERG was not aware of evidence supporting this assumption in patients with adenocarcinoma. The ERG noted that the WSY001 trial, included in the mixed treatment comparison of adverse events by the company, was conducted in China. The ERG suggested that the difference in co-morbidities, smoking history and pharmacokinetics between these populations may mean the adverse events are not generalisable to England. The ERG interpreted the data from these trials as showing that patients tolerate erlotinib better than nintedanib plus docetaxel or docetaxel alone.

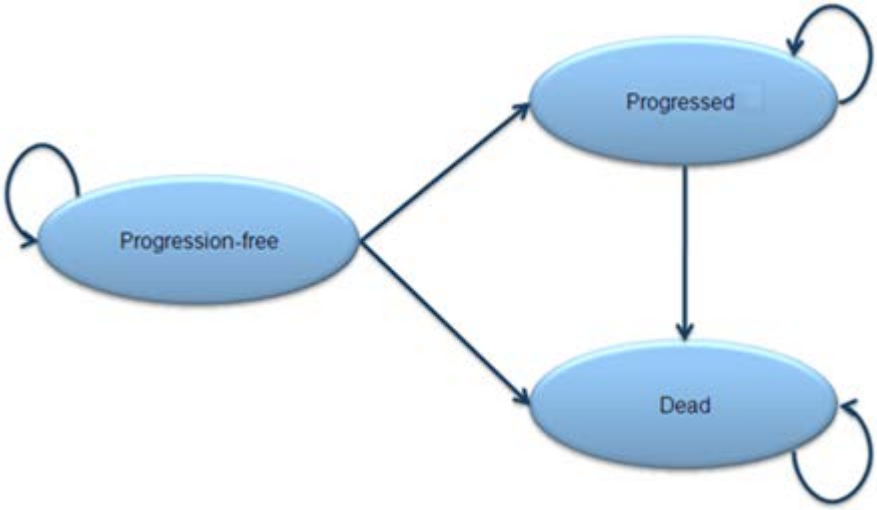
5 Cost-effectiveness evidence

Model structure

5.1 The company provided a partitioned survival Markov model containing 3 health states; progression-free (on or off treatment), progressed disease and death (see Figure 6). The company modelled 3-weekly cycle lengths, a half-cycle correction and a time horizon of 15 years. All costs and outcomes were discounted by 3.5% and the company stated that all costs were from the NHS and Personal Social Services perspective, although the company included only NHS costs in the model. In the company's base-case analysis, it compared nintedanib plus docetaxel with docetaxel alone. In the company's secondary analysis, it compared nintedanib plus docetaxel with erlotinib. The company did not consider erlotinib to be a suitable comparator, because if patients were fit enough (ECOG performance status 0 to 1) to receive to receive docetaxel, they would not receive erlotinib. The model includes people with locally advanced, metastatic or locally recurrent adenocarcinoma who progressed.

5.2 All patients enter the model at the progression-free state and stay in the same state or progress to a worse health state. The company estimated the proportion of people in the progressed disease state as the difference between overall survival and progression-free survival. The company took the data for overall survival and progression-free survival from the LUME-Lung 1 trial and from the parametric curves that the company generated from the trial data. The company assumed that patients receive best supportive care on stopping second-line treatment, although some people in the progressed disease state can have subsequent treatments (5% erlotinib, 25% platinum doublet therapy and 70% best supportive care). The company included the cost of subsequent treatments in the model but made no assumptions about their efficacy.

Figure 6. Diagram of the company's model taken from company submission (Figure 24, page 181 of the company submission)



ERG's comments

5.3 The ERG commented that the company's model was generally appropriate, but that there were a number of issues which affected the results generated in the model. These are discussed in more detail in section 5.14.

Model details

5.4 The company took Kaplan-Meir curves for overall survival and progression-free survival from the LUME-Lung 1 trial, and then used them to estimate beyond the end of the trial. In order to extrapolate data beyond the trial, data for progression-free survival and overall survival data were analysed using parametric survival models. These parametric survival curves were fitted using 2 approaches:

- Joint models including data from both treatment groups with a term for treatment and using the same distributions for each group.
- Separate modelling were statistical models were fitted to each randomised treatment group separately.

The company tested the 'fit' of the curves using Akaike information criteria (AIC). The intercept and scale parameters of the separately fitted curves indicated that the curves should not be forced into the same model, therefore separate curves were selected for progression-free survival and overall survival. The log-normal model had the lowest AIC among the separate progression-free survival fits, and the Weibull had the lowest AIC among the separate proportional hazard models for progression-free survival; therefore, these were selected to model progression-free survival. The log-logistic model had the lowest AIC among the separately fitted overall survival models, and the Weibull model had the lowest AIC among the separate proportional hazard models for overall survival; therefore, these were selected to model the overall survival data. The company stated that it tested the validity of the data running the results by a group of 'key opinion leaders' (clinicians) and comparing the data to data from the British National Lung Cancer Audit (LUCADA) and from the American Surveillance, Epidemiology and End Result (SEER).

5.5 The company had collected health-related quality of life data in the LUME-Lung 1 trial using EQ-5D questionnaires which it used in a longitudinal model to adjust for certain baseline characteristics including ECOG score,

prior treatment with bevacizumab, presence of brain metastases, health status and key adverse events. In the progression-free survival health state, the company estimated utility values from week 0 to 30 in 3-week intervals without a treatment term. The company extrapolated the trend it observed up to week 30 to provide data beyond this time point which it incorporated into its base-case. To estimate utility values for the progressed disease state, the company used utility values from the LUME-Lung 1 trial. Utility values for progression-free survival and progressed disease from the literature (Chouaid et al.) were also tested during the sensitivity analyses in the model and considered utility values from non-small cell lung cancer patients who were being treated in the UK, Europe, Canada, Australia and Turkey. The model also incorporated adverse events and the impact on health-related quality of life using decrements in utility associated with each adverse event. The company acknowledged that the model may have double counted disutility as people may have more than one adverse event.

5.6 In the model, the company assumed that patients take 2 100 mg capsules of nintedanib twice daily. The company modelled an option of patients taking 150 mg capsules. The price of both formulations is likely to be the same at £2,151. In the model nintedanib plus docetaxel is given for a minimum of 4 cycles before nintedanib can be administered alone. The model included no administration cost associated with nintedanib, but a cost of £155 for docetaxel. Intravenous docetaxel was modelled at a concentration of 75 mg/m^2 on day 1 of a 21-day cycle. For the comparison of nintedanib plus docetaxel with erlotinib, a 30 tablet pack of erlotinib was £1631.53 (MIMS list price [2013]) even though erlotinib has a confidential patient access scheme. The company assumed that the cost of best supportive care was £403.63 per 3 week cycle.

Company's base-case results and sensitivity analysis

Nintedanib plus docetaxel compared with docetaxel alone

5.7 The company's base-case analysis compared nintedanib plus docetaxel with docetaxel alone. In the base-case the ICER for nintedanib plus docetaxel compared with docetaxel was £50,677 per QALY gained (see Table 11).

Table 11. The Company's base-case cost-effectiveness results for nintedanib plus docetaxel compared with docetaxel alone

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£) vs baseline (QALYs)	ICER (£) incremental (QALYs)
Nintedanib + docetaxel	[redacted] [CiC]	[redacted] [CiC]	[redacted] [CiC]	-	-	-	-	-
Docetaxel	[redacted] [CiC]	[redacted] [CiC]	[redacted] [CiC]	£11,051	0.33	0.22	£50,776	£50,776
Abbreviations: QALY, Quality adjusted life year; ICER, incremental cost effectiveness ratio; LYG, life years gained								

Company's deterministic sensitivity analyses

5.8 The company carried out a range of deterministic sensitivity analyses for the base-case of nintedanib plus docetaxel compared with docetaxel alone, which included hazard ratios for progression-free survival, hazard ratios for overall survival, utility values for progressed disease, model costs for progressed disease (nintedanib plus docetaxel compared with best supportive care), risk of discontinuig nintedanib and docetaxel per cycle and percentage of patients switching to best supportive care. Of these, change in the utility value of progressed disease for nintedanib plus docetaxel and also for docetaxel alone had the greatest effect on cost effectiveness. For further details see figure 33, page 273 of the company's submission.

Company's probabilistic sensitivity analyses

5.9 The company carried out 5000 iterations of the cost-effectiveness model to derive the mean ICERs for nintedanib plus docetaxel compared with

docetaxel alone and erlotinib. The result for nintedanib plus docetaxel compared with docetaxel alone showed that nintedanib plus docetaxel had a 2% probability of being cost-effective at the £30,000 per QALY gained threshold and a 50% chance of being cost-effective at the £50,000 per QALY gained threshold. The probabilistic sensitivity analysis for nintedanib plus docetaxel compared with erlotinib showed that nintedanib plus docetaxel had a 65% probability of being cost-effective at the £30,000 per QALY gained threshold and a 94% chance at the £50,000 per QALY gained threshold.

Company's scenarios

5.10 The company commented that the survival modelling was a key driver of the cost effectiveness analyses and therefore the company undertook survival modelling scenarios using the Weibull distribution and Kaplan-Meier curves using the LUCADA or the SEER results. Table 12 presents the results of the company's survival modelling scenarios.

Table 12. Results from modelling scenario analyses (taken from company submission Table 135, page 250)

Progression-free survival	Overall survival	Incremental Lys	Incremental costs	Incremental QALYs	ICER
Separate – Lognormal (base-case)	Separate – Loglogistic (base-case)	0.33	£11,051	0.22	£50,776
Separate - Weibull	Separate – Weibull	0.22	£9,852	0.14	£69,884
KM Curve	KM Curve	0.11	£9,425	0.08	£119,209
KM Curve - used until time horizon	Mixed: KM & SEER-Lognormal	0.27	£10,304	0.18	£56,769
KM Curve - used until time horizon	Mixed: KM & LUCADA-Lognormal	0.26	£10,245	0.17	£58,660
KM Curve - used until time horizon	Mixed curves: KM & Separate Loglogistic	0.34	£10,637	0.22	£48,264
KM Curve - used until time horizon	Mixed curves: KM & Separate Weibull	0.23	£10,071	0.15	£65,274
Abbreviations: QALY, Quality adjusted life year; ICER, incremental cost effectiveness ratio; LYs, life years					

5.11 The company undertook a number of other scenario analyses which included resource use, utility scenarios, utility values and time horizon. The results are shown in Table 13.

Table13. Scenario analyses results (taken from Tables 137-140, pages 281 and 282 of the company's submission)

Scenario	ICER for nintedanib plus docetaxel compared with docetaxel (lognormal, loglogistic)
Base-case	£50,776
Resources	£52,692
Utility scenarios (LOCF for PFS)	£51,496
Chouaid trial for PFS and PD	£65,408
Time horizon (based on LUME-Lung1 trial)	£86,023
Time horizon (3 years)	£98,119
Time horizon (5 years)	£70,951
Time horizon (10 years)	£55,132
Abbreviations: Inc., incremental; QALY, Quality adjusted life year; ICER, incremental cost effectiveness ratio; PFS, progression-free survival; LOCF, last observation carried forward; PD, progressive disease	

ERG's comments

5.12 The ERG commented that to provide overall and progression-free survival data from the LUME-Lung 1 trial over a lifetime, the company had fitted a variety of parametric functions to the available trial data and used these in the model to predict the results beyond those available from the trial. The ERG considered the company's approach to be flawed in several ways. According to the ERG:

- The main reason for curve fitting is to anticipate what will happen to the minority of patients who remain 'at risk' at the time of the data cut. However, the majority of patients had died, progressed or stopped treatment. Extrapolating in this situation can cause bias against survivors still at risk and can lead to fitting inappropriate functions which generate misleading projections.
- The company fitted parametric functions based on descriptive data sets including SEER and LUCADA but it was not possible for the ERG to

assess whether this approach was valid. The company did not provide references for the specific data sets used, nor present sufficiently detailed explanations of the data employed (see section 4.13).

- The company provided only 1 trial to support its submission but replaced a large part of the data with a fitted model which could have disguised underlying 'disease dynamic' and added uncertainty.
- To model survival, the company used standard statistical functions which lacked a logical or empirical basis for representing the biology of the disease and instead chose them for convenience.

5.13 The ERG was concerned about the company's use of data from SEER and LUCADA to calculate overall survival for nintedanib plus docetaxel because the company did not provide references or relevant details for the data. The ERG had to infer from the text that the SEER results were related to all-cause mortality from the date of Stage 4 diagnosis. For the LUCADA data, the ERG appreciated that the data were related to second-line chemotherapy, but the ERG had no information on first-line treatments. The ERG commented that it was difficult to assess whether the company's chosen parametric survival functions were valid.

ERG exploratory analyses for nintedanib plus docetaxel compared with docetaxel

5.14 The ERG identified 11 aspects of the company's base-case model that involved errors in data analysis, parameter values or methodology. The ERG corrected these to estimate the ICER, but still considered that the model generated uncertainty in overall survival, progression-free survival and time to treatment. The ERG undertook 11 different amendments to the company's base-case. The amendments were as follows;

5.15 The company's base-case comparison of nintedanib plus docetaxel compared with docetaxel alone indicated an undiscounted overall survival gain of 4.7 months. The ERG noted that only 15% of this gain could be attributed to the pre-progression phase. The ERG stated that this is

unusual since in locally advanced and metastatic cancers the benefit from treatment normally occurs in the progression-free stage when patients receive active treatment. To confirm this, the ERG carried out its own analysis using the data for overall survival and progression survival from the trial, and found that after 300 days, a simple linear trend for both groups was observed. This indicates that a simple exponential projective model can be used, and the ERG calculated a long-term hazard ratio of 0.83 for overall survival, in favour of nintedanib plus docetaxel. To confirm this, the ERG produced a cumulative hazard chart which suggested that patients in LUME-Lung 1 who survived beyond disease progression continued to gain survival benefit associated with treatment. The ERG estimated overall survival by calculating the area under the Kaplan Meier curves then projected survival using the exponential trends. The ERG estimated mean overall survival in the docetaxel treatment arm as 453.0 days (14.9 months) whereas for the nintedanib plus docetaxel treatment arm it was 545.7 days (17.9 months) meaning a net survival of 92.7 days (3.05 months). The ERG commented that this result was considerably lower than the company's estimate of overall survival gain of 4.7 months. Replacing the company's preferred overall survival with the ERG's result increased the ICER to £68,587 per QALY gained (see Table 14). For further details of the ERG's estimation of overall survival, see section 5.5.2, pages 87-89 of the ERG report.

5.16 The ERG noted that the company's model base-case comparison of nintedanib plus docetaxel compared with docetaxel indicated a gain in (undiscounted) progression-free survival of 28.6 days, based on calibrating a LogNormal hazard distribution to each trial group and applying these to represent patient experience until all patients have died or suffered disease progression. Examination of the progression-free survival temporal profile indicated that although the addition of nintedanib to docetaxel therapy generates a short-term delay in disease progression for some patients (such as when the progression-free survival curves begin to separate), subsequently this advantage progressively reduces

until the progression-free survival experience of patients in the 2 trial groups is almost equal. Here, the extent of advantage in mean progression-free survival can be readily estimated directly from the Kaplan-Meier analysis results by comparing the area under the curve estimates up to the point when the curves converge. The ERG identified that the curves converged at day 375 and the difference in the area under the curve at this time was 36.4 days, which suggested that the company's model had underestimated progression-free survival (28.6 days). The ERG incorporated its own result into the company's model and noted that a common long-term exponential model was appropriate from day 375 onwards. This increased the ICER to £52,445 per QALY gained (see Table 14). For further details of the ERG's estimation of progression-free survival, see section 5.5.3, pages 90-91 of the ERG report.

5.17 The ERG used a similar approach, as that used for calculating the effect of the ERG's progression-free survival estimates, to estimate duration of treatment in the 2 groups of patients in the LUME-Lung 1 trial which increased the discounted cost per patient and the incremental cost per patient increase by 2.2% in both groups. This increased the ICER to £51,930 per QALY gained (see Table 14).

5.18 The ERG commented that the company costed both nintedanib plus docetaxel and docetaxel alone using average number of patients receiving treatment across each cycle. The ERG commented that adjusting mid-cycle is not accurate for docetaxel or for nintedanib because patients receive treatment on the first day of a 3-week cycle. The error underestimated the quantity and cost of drugs used in the trial. The ERG's correction of this error increased the ICER to £53,839 per QALY gained (see Table 14).

5.19 The ERG commented that the company calculated the average cost per dose of docetaxel using body surface area relevant to the UK population, but did not take into account the sex of the patients. The company also

only costed the full 75 mg/m^2 dose rather than the reduced dose of 60 mg/m^2. The ERG considered it more accurate to cost the reduced dose, and then create a weighted average based on the proportions of the 2 doses recorded in the trial. The ERG considered that the nintedanib capsules would likely be dispensed with docetaxel so any missed dosing unlikely to have an effect on the dispensing pattern. Therefore the ERG considered a reduction in cost through a randomised dose intensity index, from trial data, to be inappropriate. The ERG re-estimated the overall average cost per dose of docetaxel using separate subgroups for men and women and also re-estimated the randomised dose index multiplier to match the balance of full and reduced doses. The ERG estimated an overall mean cost for nintedanib treatment per cycle using the LUME-Lung 1 trial data and this caused the ICER to increase to £52,587 per QALY gained (see Table 14).

5.20 The cost of treating the adverse event of febrile neutropenia was included in the company's model at £2012.10 per patient affected. The ERG noted that this is substantially lower than the figure estimated by the NICE Decision Support Unit in 2007 and the updated figure used in the on-going multiple technology appraisal for 'erlotinib and gefitinib for treating non-small-cell lung cancer that has progressed following prior chemotherapy (Review of TA162 and TA175)', which used £5,240.40 per episode and a mean cost per patient of £7,352.54 (assuming 1.4 episodes per patient). Using these revised costs, the ICER increased to £51,372 per QALY gained (see Table 14).

5.21 The ERG also noted that there were discrepancies in monitoring costs between patients who were progression-free. The company assigned monitoring costs of £188 per cycle to patients in the nintedanib plus docetaxel group who were progression-free and receiving active treatment and a cost of £205 per cycle to those receiving docetaxel alone when the only difference is self-administration of the nintedanib capsules. When the

ERG reallocated costs the ICER increased to £51,140 per QALY gained (see Table 14).

5.22 In the opinion of the ERG, the company modelled discounting incorrectly basing the discounting on the 3-weekly cycle rather than annually. The ERG's amendment decreased the ICER to £50,532 per QALY gained (see Table 14).

5.23 The main adverse events in LUME-Lung 1 trial were stage 3 or 4 diarrhoea and fatigue. The company indicated that the disutility for diarrhoea is low (-0.04) whereas for fatigue it was much higher (-0.21). The ERG also noted that the company indicated a statistically significant difference between effect sizes in the 2 treatment groups with a disutility of -0.326 for the nintedanib plus docetaxel group and -0.101 for the docetaxel alone group. The ERG suggested that fatigue was a more serious side effect for those receiving the dual therapy. The company used an average disutility for the 2 treatment groups whereas the ERG applied a disutility to the 2 groups separately. The ERG's amendment resulted in an ICER of £50,830 per QALY gained (see Table 14).

5.24 The cost of care for patients who had finished active treatment, but remained in a stable condition differed between the model and the figures supplied to the company by clinicians. In the model the company assumed that these patients accrued the costs of receiving palliative nursing care every week and a bone scan every 3 weeks. In the ERG's opinion, this reflected an error which significantly reduces the care costs of patients in a stable condition after second-line treatment. The ERG's amendment resulted in an ICER of £53,470 per QALY gained (see Table 14).

5.25 The ERG noted that the company's' model followed the protocol used in the LUME-Lung 1 trial which allowed patients to have unlimited docetaxel treatment. The ERG explained that in the UK patients are restricted to 4 cycles of docetaxel because of unacceptable adverse events. Although

the company's model allowed the number of cycles to be restricted, the ERG found an error which limited the cycles to 5 rather than to 4. When the ERG restricted the cycles to 4, this affected only the drug acquisition and administration costs, but not adverse events. This reduced the base-case incremental cost per patient by 5.4% and reduced the ICER to £48,060 per QALY gained (see Table 14).

5.26 The ERG provided an ICER which incorporated all its amendments simultaneoulsy to produce an ICER for nintedanib plus docetaxel compared with docetaxel alone of £85,292 per QALY gained. The ERG also provided an ICER which included all amendments excluding analyses of the number of cycles of docetaxel. This produced an ICER of £82,995 per QALY gained (see Table 14).

Table 14. ERG exploratory analyses for nintedanib plus docetaxel compared with docetaxel alone (see Table 40, page 103 of the ERG report)

Scenario	Nintedanib plus docetaxel		Docetaxel		Incremental		ICER
	Total cost	Total QALY	Total cost	Total QALY	cost	QALY	Cost per QALY
Company's base-case	[redacted] [CiC]	[redacted] [CiC]	[redacted] [CiC]	[redacted] [CiC]	£11,051	0.218	£50,776
(1) ERG OS	[redacted] [CiC]	[redacted] [CiC]	[redacted] [CiC]	[redacted] [CiC]	£10,497	0.153	£68,587
(2) ERG PFS estimates	[redacted] [CiC]	[redacted] [CiC]	[redacted] [CiC]	[redacted] [CiC]	£11,527	0.220	£52,445
(3) ERG ToT estimates	[redacted] [CiC]	[redacted] [CiC]	[redacted] [CiC]	[redacted] [CiC]	£11,298	0.218	£51,930
(4) Mid-cycle adjustment	[redacted] [CiC]	[redacted] [CiC]	[redacted] [CiC]	[redacted] [CiC]	£11,717	0.218	£53,839
(5) Cost of treatment doses	[redacted] [CiC]	[redacted] [CiC]	[redacted] [CiC]	[redacted] [CiC]	£11,445	0.218	£52,587
(6) Febrile neutropenia cost	[redacted] [CiC]	[redacted] [CiC]	[redacted] [CiC]	[redacted] [CiC]	£11,180	0.218	£51,372
(7) Monitoring cost	[redacted] [CiC]	[redacted] [CiC]	[redacted] [CiC]	[redacted] [CiC]	£11,130	0.218	£51,140
(8) Discounting method	[redacted] [CiC]	[redacted] [CiC]	[redacted] [CiC]	[redacted] [CiC]	£11,189	0.221	£50,532
(9) Disutility of fatigue	[redacted] [CiC]	[redacted] [CiC]	[redacted] [CiC]	[redacted] [CiC]	£11,051	0.217	£50,830
(10) Stable disease costs	[redacted] [CiC]	[redacted] [CiC]	[redacted] [CiC]	[redacted] [CiC]	£11,637	0.218	£53,470
(11) Docetaxel 4	[redacted] [CiC]	[redacted] [CiC]	[redacted] [CiC]	[redacted] [CiC]	£10,452		£48,060

or less cycles						0.217	
ERG base-case with first 10 revisions	[redacted] [CiC]	[redacted] [CiC]	[redacted] [CiC]	[redacted] [CiC]	£13,087	0.158	£82,995
ERG base-case with all 11 revisions	[redacted] [CiC]	[redacted] [CiC]	[redacted] [CiC]	[redacted] [CiC]	£13,437	0.158	£85,292
Abbreviations: QALY, Quality adjusted life year; ICER, incremental cost effectiveness ratio, OS, overall survival; PFS, progression-free survival; ToT, time to treat							

Company's secondary analysis for nintedanib plus docetaxel compared with erlotinib

5.27 The progression-free survival and overall survival curves for erlotinib were not available so the company derived them by applying the hazard ratios for all the comparisons in the mixed treatment comparison to the overall survival and progression-free survival of nintedanib plus docetaxel. The company considered that hazard ratios can only be used if the survival distribution satisfies the proportional hazard assumptions. Therefore, in the model, the company used a Weibull distribution to evaluate erlotinib for both overall survival and progression free survival.

5.28 The results of the company's secondary analysis for nintedanib plus docetaxel compared with erlotinib are presented in Table 15.

Table 15. The Company's secondary cost-effectiveness results: nintedanib plus docetaxel compared with erlotinib (see Table 30, page 273 of the company's submission)

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£) vs baseline (QALYs)	ICER (£) increment al (QALYs)
Nintedanib + docetaxel	[redacted] [CiC]	[redacted] [CiC]	[redacted] [CiC]	-	-	-	-	-
Erlotinib	[redacted] [CiC]	[redacted] [CiC]	[redacted] [CiC]	£7,571	0.43	0.28	£27,008	£27,008
Abbreviations: QALY, Quality adjusted life year; ICER, incremental cost effectiveness ratio; LYG, life years gained								

Company's probabilistic sensitivity analyses for nintedanib plus docetaxel compared with erlotinib

5.29 The company carried out probabilistic sensitivity analysis for nintedanib plus docetaxel compared with erlotinib showed that nintedanib plus docetaxel had a 65% probability of being cost-effective at the £30,000 per QALY gained threshold and a 94% chance at the £50,000 per QALY gained threshold. The comparison between the deterministic and probabilistic results for nintedanib plus docetaxel compared with erlotinib in shown in Table 16.

Table 16. Comparison of ICERs obtained from deterministic and probabilistic sensitivity analyses for nintedanib plus docetaxel compared with erlotinib (taken from Table 134, page 278 of the company submission)

	Incremental cost	Incremental QALY	ICER
Deterministic Values	£7,571	0.28	£27,008
Average value for PSA	£7,518	0.27	£27,484
Abbreviations: QALY, Quality adjusted life year; ICER, incremental cost effectiveness ratio; PSA; probabilistic sensitivity analysis			

Company scenario analysis for nintedanib plus docetaxel compared with erlotinib

5.30 The company undertook a scenario analysis using the results from the network meta-analysis scenario analysis rather than the network meta-analysis base-case. The results of the analysis are provided in Table 17.

Table 17. Results of indirect comparison scenarios for nintedanib plus docetaxel compared with erlotinib (taken from Table 13, page 280 of the company's submission)

Progression-free survival hazard ratio	Overall survival hazard ratio	Incremental costs	Incremental QALYs	ICER
NMA Base-case network: 0.70	NMA Base-case network: 0.64	£7,571	0.28	£27,008
NMA Scenario Analysis network, Fixed-effect model: 0.68	NMA Scenario Analysis network, Fixed-effect model: 0.74	£6,952	0.20	£34,509
Abbreviations: QALY, Quality adjusted life year; ICER, incremental cost effectiveness ratio; NMA, network meta-analysis				

5.31 The company undertook a number of other scenario analyses which included resource use, utility scenarios, utility values and time horizon,. The result are shown in Table 18.

Table 18. Scenario analyses results (taken from Tables 137 and 139 in company submission)

Scenario	ICER for nintedanib plus docetaxel compared with erlotinib (Weibull)
Base-case	£27,008
Resources	£25,301
Utility scenarios (LOCF for PFS)	£26,961
Chouaid study for PFS and PD	£33,464
Time horizon (based on LUME-Lung1 study)	£29,744
Time horizon (3 years)	£31,816
Time horizon (5 years)	£27,740
Time horizon (10 years)	£27,013
Abbreviations: Inc., incremental; QALY, Quality adjusted life year; ICER, incremental cost effectiveness ratio; PFS, progression-free survival; LOCF, last observation carried forward; PD, progressive disease	

5.32 The company applied a range of discounts to the list price of erlotinib as it was not aware of the size of discount available to the NHS through the agreed confidential patient access scheme for erlotinib. The results are shown in Table 19.

Table 19. Impact of discount applied to the list price of erlotinib (taken from Table 142 in company submission)

	Discount Applied to list price of erlotinib						
Erlotinib	**Base-case (0% discount)**	**5%**	**10%**	**15%**	**20%**	**25%**	**30%**
ICER	£27,008	£27,934	£28,866	£29,797	£30,729	£31,660	£32,592
Erlotinib	**35%**	**40%**	**45%**	**50%**	**55%**	**60%**	**65%**
ICER	£33,524	£34,455	£35,387	£36,318	£37,250	£38,182	£39,113
Erlotinib	**70%**	**75%**	**80%**	**85%**	**90%**	**95%**	
ICER	£40,045	£40,977	£41,908	£42,840	£43,771	£44,703	

ERG's comments and exploratory analyses for nintedanib plus docetaxel compared with erlotinib

5.33 The ERG did not consider the company's comparison of nintedanib plus docetaxel with erlotinib to be appropriate because the data used by the company for time to treatments were based on the mean number of cycles of erlotinib taken from a previous appraisal (Erlotinib for the treatment of non-small-cell lung cancer, TA162) whereas this appraisal had used indirect trial data which may have overestimated time to treatment. The ERG were unable to rectify the issue without access to patient level data for the studies used in the mixed treatment comparison and could not estimate the drug acquisition costs in the company's model.

5.34 To calculate overall survival and progression-free survival for erlotinib the ERG commented that the meta-analysis of time-to-treatment data must incorporate some conditions:

- Within each trial, the assumptions of proportional hazards must apply.
- Between trials featuring the treatment at nodes in the mixed treatment comparison, treatment outcomes should be equivalent (that is proportional hazards and very similar outcomes at all time points).
- Any parametric survival function propagated through the network must comply with proportional hazard assumptions.

The company used a Weibull function even though this did not give the best match for the LUME-Lung 1 overall survival data for nintedanib plus docetaxel. If the criteria outlined by the ERG had been met, the company should have adjusted the Weibull curve by an overall hazard ratio (0.64 for nintedanib plus docetaxel compared with erlotinib in overall survival) which was consistent with 2 of the trials included in the mixed treatment comparison. However, the ERG commented that the proportional hazards assumption was seriously violated in the erlotinib trials included in the mixed treatment comparison. This indicated that the estimated overall survival data were inconsistent within the network, and that the Weibull

data from LUME-Lung 1, when added to the network, did not generate the same outcome patterns seen in the other trials. In the opinion of the ERG, this added doubt to both the overall survival estimate for erlotinib and the use of a Weibull parametric form. The ERG was unable to assess fully the estimates of progression-free survival for erlotinib but suspected the same issues would apply.

5.35 The ERG carried out 7 of the 11 amendments, it had identified when analysing nintedanib plus docetaxel compared with docetaxel alone (see Table 14), on the ICER for nintedanib plus docetaxel compared with erlotinib. The ERG also took into account the assumed patient access scheme discounts for erlotinib on the ICER (see Table 20). However, the ERG still concluded that it did not consider erlotinib to be a suitable comparator.

Table 20. Cost-effectiveness results for nintedanib plus docetaxel compared with erlotinib incorporating the ERGs amendments and possible discounts on the list price of erlotinib (taken from Table 42, page 105 of ERG report).

Model scenario & ERG revisions	Patient access scheme discount for erlotinib										
	0%	5%	10%	15%	20%	25%	30%	35%	40%	45%	50%
Company's base-case	**£27,008**	**£27,939**	**£28,870**	**£29,802**	**£30,733**	**£31,664**	**£32,596**	**£33,527**	**£34,458**	**£35,390**	**£36,321**
Mid-cycle adjustment	£27,878	£28,902	£29,926	£30,950	£31,975	£32,999	£34,023	£35,047	£36,071	£37,095	£38,119
Cost of treatment	£28,275	£29,206	£30,138	£31,069	£32,000	£32,932	£33,863	£34,794	£35,726	£36,657	£37,588
Febrile neutropenia	£28,173	£29,104	£30,035	£30,967	£31,898	£32,830	£33,761	£34,692	£35,624	£36,555	£37,486
Discounting method	£26,927	£27,851	£28,775	£29,699	£30,623	£31,547	£32,471	£33,395	£34,319	£35,243	£36,167
Disutility of fatigue	£27,020	£27,951	£28,883	£29,815	£30,747	£31,678	£32,610	£33,542	£34,474	£35,405	£36,337
Stable disease	£27,027	£27,958	£28,890	£29,821	£30,752	£31,684	£32,615	£33,546	£34,478	£35,409	£36,340
Docetaxel ≤4 cycles	£24,975	£25,897	£26,820	£27,742	£28,664	£29,587	£30,509	£31,431	£32,354	£33,276	£34,198
Base-case + revisions 4-11	**£28,307**	**£29,314**	**£30,320**	**£31,327**	**£32,334**	**£33,341**	**£34,348**	**£35,354**	**£36,361**	**£37,368**	**£38,375**

Innovation

5.36 Justifications provided by the company for considering nintedanib plus docetaxel to be innovative:

- Nintedanib plus docetaxel would provide an alternative second-line treatment option for adenocarcinoma patients. This combination would be the first to offer a significant and clinically meaningful overall survival benefit for second-line adenocarcinoma patients compared with an active ingredient in a phase 3 trial.

5.37 A patient and carer organisation also considered nintedanib to be a new and innovative therapy for non-small cell lung cancer.

6 End-of-life considerations

Table 21. End-of-life considerations from the company (taken from page 288 of company submission) and ERG.

Criterion	Data available
The treatment is indicated for patients with a short life expectancy, normally less than 24 months	Median overall survival of patients on docetaxel monotherapy is 10.23 months and the mean overall survival is 15.96 months.
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 median overall survival at final analysis in LUME-Lung 1 trial for median extension over docetaxel monotherapy was 2.3 months. Using the company's economic base-case assumptions the mean extension over docetaxel monotherapy was 3.96 months and the mean extension over erlotinib was 5.16 months. The ERG's exploratory base-case analysis calculated that the mean extension over docetaxel monotherapy was 3.05 months.
The treatment is licensed or otherwise indicated for small patient populations	The company indicated that the total population for nintedanib plus docetaxel in England is less than 800 people

7 Equality issues

7.1 No potential equality issues were identified during the draft scope consultation. Scoping workshop attendees noted that the pivotal studies for nintedanib included people whose disease had progressed following first-line, but did not include people whose disease progressed following maintenance therapy. This could be a potential equality issue given maintenance therapy is now used in clinical practice for some patients following first-line induction therapy. The population in the scope is 'Adults with locally advanced, metastatic or locally recurrent non-small cell lung cancer of adenocarcinoma tumour histology that has progressed after first-line chemotherapy' and therefore would include this group. However, NICE technology appraisal guidance will only be issued in accordance with the marketing authorisation and the clinical evidence presented during the appraisal.

7.2 Nintedanib does not currently have a UK marketing authorisation for previously treated NSCLC but has received a positive CHMP opinion on 25th September 2014 in combination with docetaxel for the treatment of adult patients with locally advanced, metastatic or locally recurrent non-small cell lung cancer (NSCLC) of adenocarcinoma tumour histology after first-line chemotherapy'. CHMP opinion is based on LUME-Lung 1 trial, which did not include maintenance therapy. As a result, this is not an equality issue that can be addressed by the recommendations for this appraisal.

8 Authors

Caroline Hall

Technical Lead

Nicola Hay

Technical Adviser

with input from the Lead Team (Mark Chapman, Peter Norrie and Danielle Preedy).

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal (STA)

Submitted by Boehringer Ingelheim Ltd

Nintedanib in combination with docetaxel for the treatment of adult patients with locally advanced, metastatic or recurrent non-small cell lung cancer (NSCLC) of adenocarcinoma tumour histology after first-line chemotherapy

August 2014

Contents

List of tables

List of figures

List of abbreviations

ABPI	Association of the British Pharmaceutical Industry
ACD	Appraisal consultation document
ACS	American Cancer Society
AE	Adverse event
AESI	Adverse event of special interest
AIK	Akaike information criteria
ALK	Anaplastic lymphoma kinase
ALT	Alanine aminotransferase
ASCO	American Society of Clinical Oncology
AST	Aspartate aminotransferase
ATP	Adenosine triphosphate
AUC	Area under the curve
BD	Twice daily
BGR	Brooks-Gelman-Rubin
BIC	Bayesian information criteria
British National Formulary	British National Formulary
BSA	Body surface area
BSC	Best supportive care
CBC	Complete blood count
CC	Completeness of cytoreduction
CHMP	Committee for Medicinal Products for Human Use
CI	Confidence interval
CR	Complete response
CRD	Centre for Reviews and Dissemination
CRF	Case report form
CRO	Contract research organisation
CT	Computed/computerised tomography
CTC	Common toxicity criteria
CTCAE	Common terminology criteria for adverse events

CTR	Clinical trial report
DALY	Disability-adjusted life year
DH	Department of Health
DHFR	Dihydrofolate reductase
DIC	Deviance information criteria
DMC	Data monitoring committee
DSU	Decision support unit
ECCO	European Cancer Organisation
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
EE	External Expert
EED	Economic evaluation database
EGFR	Estimated glomerular filtration rate
EMA	European Medicines Agency
EORTC QLQ LC	European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire (Lung Cancer Module)
EPAR	European Public Assessment Report
ERG	Evidence Review Group
ESMO	European Society for Medical Oncology
ESTRO	European Society for Radiotherapy and Oncology
FAC	Final appraisal determination
FBC	Full blood count
FDA	Food and Drug Administration
FGF	Fibroblast growth factor
FGFR	Fibroblast growth factor receptor
FISH	Fluorescent in-situ hybridisation
GCP	Good clinical practice
GGT	Gamma-glutamyl transpeptidase
GI	Gastrointestinal
GP	General practitioner
HEED	Health Economic Evaluation Database
HES	Hospital Episode Statistics
HQIP	Healthcare Quality Improvement Partnership
HR	Hazard ratio
HRQL	Health-related Quality of Life
HTA	Health Technology Assessment
ICER	Incremental cost-effectiveness ratio
ICH	International Conference on Harmonisation
ICU	Intensive care unit

IFU	Information for use
IQR	Interquartile range
IRC	Independent review committee
ITC	Indirect treatment comparison
ITT	Intention-to-treat
IV	Intravenous
IVRS/IWRS	Interactive voice/web response system
LCL	Lower control limit
LDH	Lactate dehydrogenase
LOCF	Last observation carried forward
LOR	Log-odds ratio
LUCADA	National Lung Cancer Audit Data Set
LY	Life years
LYG	Life years gained
MRI	Magnetic resonance imaging
MS	Manufacturer's submission
MTC	Mixed treatment comparison
NHS	National Health Service
NMA	Network meta-analysis
NSCLC	Non-small-cell lung cancer
ORR	Objective response rate
OR	Odds ratio
OS	Overall survival
OWSA	One-way deterministic sensitivity analysis
PAS	Patient access scheme
PDGF	Platelet-derived growth factor
PDGFR	Platelet-derived growth factor receptor
PET	Positron emission tomography
PF	Progression-free
PFS	Progression-free survival
PHM	Proportional Hazard Model
PI	Patient information
PPS	Post-progression state
PR	Partial response
PRO	Patient reported outcome
PS	Performance status
PSA	Probabilistic sensitivity analysis
PSS	Personal social services

PSSRU	Personalised Social Services Research Unit
QALY	Quality-adjusted life year
QLQ	Quality of life questionnaire
QOL	Quality of life
RCT	Randomised controlled trial
RET	Rearranged during transfection
RR	Relative risk
RT	Radiation therapy
SAP	Statistical analysis plan
SD	Standard deviation
SE	Standard error
SEER	Surveillance, Epidemiology, and End Results
SLD	Sum of longest diameter
SLR	Systematic literature review
SMC	Scottish Medicines Consortium
SMQ	Standardised MedDRA Queries
SOP	Standard operating procedure
SPC	Summary of Product Characteristics
STA	Single technology appraisal
TK	Tyrosine kinase
TKI	Tyrosine kinase inhibitor
TS	Thymidylate synthase
TSAP	Trial statistical analysis plan
UCL	Upper control limit
VAS	Visual analogue scale
VEGF	Vascular endothelial growth factor
VEGFR	Vascular endothelial growth factor receptor
WBC	White blood cell
WHO	World Health Organisation
WTP	Willingness-to-pay

Instructions for manufacturers and sponsors

This is the specification for submission of evidence to the National Institute for Health and Care Excellence (NICE) as part of the single technology appraisal (STA) process. It shows manufacturers and sponsors what information NICE requires and the format in which it should be presented. NICE acknowledges that for medical devices manufacturers particular sections might not be as relevant as they are for pharmaceuticals manufacturers. When possible the specification will refer to requirements for medical devices, but if it hasn't done so, manufacturers or sponsors of medical devices should respond to the best of their ability in the context of the question being addressed.

Use of the specification and completion of appendices 1 to 13 (sections 10.1 to 10.13) are mandatory (when applicable), and the format should be followed whenever possible. Reasons for not following this format must be clearly stated. Sections that are not considered relevant should be marked 'N/A' and a reason given for this response. The specification should be completed with reference to the NICE document 'Guide to the methods of technology appraisal' (www.nice.org.uk), particularly with regard to the 'reference case'. Users should see NICE's 'Guide to the single technology appraisal (STA) process' (www.nice.org.uk) for further details on some of the procedural topics referred to only briefly here.

If a submission is based on preliminary regulatory recommendations, the manufacturer or sponsor must advise NICE immediately of any variation between the preliminary and final approval.

A submission should be as brief and informative as possible. It is expected that the main body of the submission will not usually exceed **100 pages excluding the pages covered by the template**. The submission should be sent to NICE electronically in Word or a compatible format, and not as a PDF file.

The submission must be a stand-alone document. Additional appendices may only be used for supplementary explanatory information that exceeds the level of detail requested, but that is considered to be relevant to the submission. Appendices are not normally presented to the Appraisal Committee. Any additional appendices should be clearly referenced in the body of the submission and should not be used for core information that has been requested in the specification. For example, it is not acceptable to attach a key study as an appendix and to complete the clinical-effectiveness section with ‘see appendix X’. Clinical trial reports and protocols should not be submitted, but must be made available on request.

Trials should be identified by the first author or trial ID, rather than by relying on numerical referencing alone (for example, ‘Trial 123/Jones et al.[126]’ rather than ‘One trial[126]’).

For information on submitting cost-effectiveness analysis models, disclosure of information and equality and diversity, users should see ‘Related procedures for evidence submission’, section 11.

If a patient access scheme is to be included in the submission, please refer to the patient access scheme submission template available on request. Please submit both documents and ensure consistency between them.

Executive summary

Please provide an executive summary that summarises the key sections of the submission. All statements should be directly relevant to the decision problem, be evidence-based when possible and clearly reference the relevant section of the submission. The summary should cover the following items.

- **The UK approved name, brand name, marketing status and principal mechanism of action of the proposed technology.**

Nintedanib, marketed as Vargatef, is a potent, orally-administered small molecule triple angiokinase inhibitor targeting three receptor classes that have a key role in angiogenesis and tumour growth: vascular endothelial growth factor receptors (VEGFR) 1-3, fibroblast growth factor receptors (FGFR) 1-3, and platelet-derived growth factor receptors (PDGFR) α and β(1-3). Additionally, receptor kinases of RET, FLT3, and the Src family are also inhibited(1-3).

Positive opinion for nintedanib is expected in [redacted], and marketing authorisation is expected in [redacted].

- **The formulation(s), strength(s), pack size(s), maximum quantity(ies), anticipated frequency of any repeat courses of treatment and acquisition cost.**

Nintedanib is available as soft gelatine capsules in 150mg (60 capsules per pack) and 100mg (120 capsules per pack) sizes. Patients are expected to take two 100mg capsules twice a day. A dose reduction to one 150mg capsule twice daily, and a further dose reduction to one 100mg capsule twice daily, is also available in the event of prolonged adverse events (AEs). Patients are expected to continue treatment continuously until disease progression or intolerable AEs.

Anticipated NHS list price per 30 day pack is £2151.10.

- **The indication(s) and any restriction(s).**

Vargatef is indicated in combination with docetaxel for the treatment of adult patients with locally advanced, metastatic or recurrent non-small cell lung cancer (NSCLC) of adenocarcinoma tumour histology after first-line chemotherapy.

- **The main comparator(s).**

As an established treatment option in England and Wales, docetaxel monotherapy is considered as the primary comparator for the second-line treatment of locally advanced or metastatic NSCLC in patients who have relapsed after previous chemotherapy.

Within its licenced indication, erlotinib is also recommended as a second-line option in England and Wales as an alternative to docetaxel monotherapy. Therefore, erlotinib is considered as an additional comparator. Please note that in an advisory board the opinion of all 5 leading clinicians was that patients fit enough for treatment with docetaxel would receive docetaxel rather than erlotinib (see section 7.3.5).

- **Whether the key clinical evidence in the submission comes from head-to-head randomised controlled trials (RCTs), from an indirect and/or mixed treatment comparison, or from non-randomised studies.**

The key clinical evidence for the comparison vs docetaxel monotherapy comes from the LUME Lung 1 trial (a phase III randomised controlled trial).

The comparison with the additional comparator erlotinib required an indirect comparison to be performed as no head to head trials were available.

- **The main results of the RCTs and any relevant non-RCT evidence.**

LUME-Lung 1

Between Dec 23, 2008, and Feb 9, 2011, 655 patients were randomly assigned to receive docetaxel plus nintedanib and 659 to receive docetaxel plus placebo. The primary analysis

was done after a median follow-up of 7.1 months (interquartile range [IQR] 3.8—11.0). Progression-free survival (PFS) was significantly improved in the docetaxel plus nintedanib group compared with the docetaxel plus placebo group (median 3.4 months [95% confidence interval, CI, 2.9—3.9] vs 2.7 months [2.6—2.8]; hazard ratio [HR] 0.79 [95% CI 0.68—0.92], p=0.0019)(3, 4). The median PFS was also significantly prolonged in patients with adenocarcinoma histology who were treated with docetaxel plus nintedanib compared to docetaxel plus placebo (4.0 vs 2.8 months respectively; HR 0.77 [95% CI 0.62-0.96], p=0.0193). After a median follow-up of 31.7 months (IQR 27.8—36.1), overall survival (OS) was significantly improved for patients with adenocarcinoma histology (322 patients in the docetaxel plus nintedanib group and 336 in the docetaxel plus placebo group; median OS 12.6 months [95% CI 10.6—15.1] vs 10.3 months [95% CI 8.6—12.2]; HR 0.83 [95% CI 0.70—0.99], p=0.0359), but not in the total study population (median 10.1 months [95% CI 8.8—11.2] vs 9.1 months [8.4—10.4]; HR 0.94, 95% CI 0.83—1.05, p=0.2720)(3, 5).

Grade 3 or worse AEs that were more common in the docetaxel plus nintedanib group than in the docetaxel plus placebo group were diarrhoea (43 [6.6%] of 652 vs 17 [2.6%] of 655), reversible increases in alanine aminotransferase ([ALT] 51 [7.8%] vs six [0.9%]), and reversible increases in aspartate aminotransferase ([AST] 22 [3.4%] vs three [0.5%]). 35 patients in the docetaxel plus nintedanib group and 25 in the docetaxel plus placebo group died of AEs possibly unrelated to disease progression; the most common of these events were sepsis (five with docetaxel plus nintedanib vs one with docetaxel plus placebo), pneumonia (two vs seven), respiratory failure (four vs none), and pulmonary embolism (none vs three)(3, 5). In the adeno carcinoma group, the proportion of patients with AEs grade ≥3 was higher in the nintedanib plus docetaxel arm (75.9%) than in the placebo plus docetaxel arm (68.5%). The proportion of patients with SAEs was however comparable across arms (34.7% and 32.1% for nintedanib plus docetaxel and placebo plus docetaxel, respectively)(5).

Indirect treatment comparison (ITC) vs erlotinib

The main HR results of the ITC are shown in Table 1 below.

Table 1: HRs of PFS and OS for nintedanib plus docetaxel versus erlotinib

Comparison	Model Base-case Analysis	Model Sensitivity Analysis	
	NMA Base-case Analysis (fixed effects)	NMA Scenario Analysis (fixed effects)	NMA Scenario Analysis (random effects)
OS (HR 95% CrIs)			
Nintedanib + docetaxel vs erlotinib	**0.64 [0.46, 0.90]**	0.74 [0.57, 0.96]	0.74 [0.40, 1.35]
PFS (HR 95% CrIs)			
Nintedanib + docetaxel vs erlotinib	**0.70 [0.50, 1.00]**	0.68 [0.49, 0.95]	0.68 [0.35, 1.35]

CrsI = Credible intervals; OS = Overall survival; PFS = Progression-free survival; HR = Hazard Ratio

Source: see section 6.7

- In relation to the economic evaluation, details of:
  - the type of economic evaluation and justification for the approach used.
  - the pivotal assumptions underlying the model/analysis.
  - the mean costs, outcomes and incremental cost-effectiveness rations (ICERs) from the evaluation

A cost-utility analysis was undertaken to assess the cost-effectiveness of nintedanib across its anticipated licensed indication.

The economic model is based predominantly on evidence from LUME-Lung 1 and the ITC. The model is a disease-state cohort model which utilises the partitioned survival method to determine the proportion of patients in each of the three health states in each model cycle (progression-free [PF] disease, progressed disease [PD], and death). Both the model structure and health states are characteristic of modelling in metastatic oncology and have been used in previous NICE single technology appraisals (STAs) and multiple technology appraisals (MTAs)(6-8). The model has been designed for the UK, and both the model structure and parameterisation aims to reflect UK clinical practice.

The sensitivity analysis in section 7.7 shows that the key drivers behind the results are the assumptions around OS and post-progression health-related quality of life (HRQL) and resource use.

End of Life Criteria

Nintedanib plus docetaxel in second-line treatment of NSCLC of adenocarcinoma histology fulfils the 'End of life' criteria.

- Patients with advanced NSCLC have a short life expectancy of less than 24 months on average. Using the extrapolated results from the LUME Lung 1 trial data implemented in the cost effectiveness model, the median OS of patients on docetaxel monotherapy (current standard of care) is 10.23 months and the mean OS is 15.96 months.
- The total eligible population for nintedanib plus docetaxel is 745 (see section 8.1)
- Extension to life due to nintedanib plus docetaxel vs docetaxel monotherapy in the target population with the base-case assumptions within the model is a mean of 3.96 months. The extension in OS over erlotinib is a mean of 5.16 months.

- **Tabulation of the base-case results as follows:**

The results from the cost-effectiveness analysis are summarised in Table 2 and Table 3 below. Note that the base-case for the comparison vs docetaxel monotherapy uses a lognormal extrapolation for PFS and a loglogistic for OS extrapolation, as these were the best statistical fit and were validated by clinicians and external data from the SEER and LUCADA databases (section 7.3.5). The comparison vs erlotinib uses Weibull extrapolations as the curves for erlotinib are derived from HRs from the ITC, and require survival models which do not violate the proportional hazards assumptions. Weibull extrapolations underestimate the OS of the cohort.

Table 2: Distributions used – OS: Log-logistic; PFS: Log-normal

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£) versus baseline (QALYs)	ICER (£) incremental (QALYs)
Nintedanib + Docetaxel	[redacted]	[redacted]	[redacted]	-	-	-	-	-
Docetaxel	[redacted]	[redacted]	[redacted]	£10,932	0.33	0.22	£50,234	£50,234

ICER = incremental cost-effectiveness ratio; LYG = life years gained; QALYs = quality-adjusted life years

Table 3: Distributions used – OS: Weibull Distributions; PFS – Weibull Survival

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£) versus baseline (QALYs)	ICER (£) incremental (QALYs)
Nintedanib + Docetaxel	[redacted]	[redacted]	[redacted]	-	-	-	-	-
Erlotinib	[redacted]	[redacted]	[redacted]	£7,425	0.43	0.28	£26,488	£26,488

ICER = incremental cost-effectiveness ratio; LYG = life years gained; QALYs = quality-adjusted life years

Section A – Decision problem

Manufacturers and sponsors will be requested to submit section A in advance of the full submission (for details on timelines, see the NICE document 'Guide to the single technology appraisal (STA) process' – www.nice.org.uk). A (draft) summary of product characteristics (SPC) for pharmaceuticals or information for use (IFU) for devices, a (draft) assessment report produced by the regulatory authorities (for example, the European Public Assessment Report [EPAR]), and a (draft) technical manual for devices should be provided (see section 10.1, appendix 1).

1 Description of technology under assessment

1.1 Give the brand name, approved name and, when appropriate, therapeutic class. For devices, provide details of any different versions of the same device

- Brand name: Vargatef
- Approved name: Nintedanib
- Therapeutic class: Angiogenesis inhibitor

1.2 What is the principal mechanism of action of the technology

Nintedanib is a potent, orally-administered small molecule triple angiokinase inhibitor targeting three receptor classes that have a key role in angiogenesis and tumour growth: vascular endothelial growth factor receptors (VEGFR) 1-3, fibroblast growth factor receptors (FGFR) 1-3, and platelet-derived growth factor receptors (PDGFR) α and β(1-3). Additionally, receptor kinases of RET, FLT3, and the Src family are also inhibited(1-3). Growing scientific evidence shows that these three receptor classes play an important role in the formation and maintenance of new blood vessels (angiogenesis).(9-11) VEGF and its receptor VEGFR-2 are crucial for the formation of new tumour vessels (12, 13), and there is preclinical evidence to suggest that FGF and PDGF, and their associated receptors, contribute to tumour angiogenesis(14, 15). Recent data has also identified FGF-receptor signaling as a possible

escape mechanism for tumour angiogenesis when the VEGF pathway is disrupted(16). Furthermore preclinical models show that PDGFRα activation is important in human lung cancer(17), and that FGFR amplification and mutation are frequently seen in human tumour cells(18), implying that inhibition of these pathways may have a direct anti-tumour effect on those malignant cells which over express PDGFR and/or FGFR.

Angiogenesis inhibitors, such as nintedanib interfere with steps in the angiogenesis signalling cascade therefore impacting tumour growth and spread(9-11). Therefore, suppression of neo-angiogenesis via inhibition of VEGFR is a promising strategy for the treatment of human solid tumours, and the simultaneous targeting of all three pathways may be more effective than inhibition of angiogenesis via the VEGF pathway alone.

Preclinical studies with nintedanib have shown sustained (>30 hours) blockade of VEGFR2 in vitro, and delay or arrest of tumour growth in xenograft models of human solid tumours(1, 3). In phase I and II clinical trials, nintedanib showed a manageable safety profile and anti-tumour activity in patients with solid tumours, including NSCLC. Limited drug-drug interactions based on its pharmacokinetic profile and absence of interactions with CYP450 enzymes allows combination of nintedanib with cytotoxic chemotherapies, such as docetaxel or pemetrexed(19, 20).

1.3 Does the technology have a UK marketing authorisation/CE marking for the indications detailed in this submission? If so, give the date on which authorisation was received. If not, state current UK regulatory status, with relevant dates (for example, date of application and/or expected approval dates).

Nintedanib does not currently have a UK Marketing Authorisation. A Marketing Authorisation Application was submitted to the European Medicines Agency (EMA) on 30 September 2013 and Marketing Authorisation is currently anticipated in [redacted].

1.4 Describe the main issues discussed by the regulatory organisation (preferably by referring to the [draft] assessment report [for example, the EPAR]). If appropriate, state any special conditions attached to the marketing authorisation (for example, exceptional circumstances/conditions to the marketing authorisation).

Not applicable at this stage.

1.5 What are the (anticipated) indication(s) in the UK? For devices, provide the (anticipated) CE marking, including the indication for use.

If approved by the EMA (expected [redacted] as per section 1.3), Vargatef is indicated in combination with docetaxel for the treatment of adult patients with locally advanced, metastatic or recurrent NSCLC of adenocarcinoma tumour histology after first-line chemotherapy.

Nintedanib is registered as a pharmaceutical and therefore does not carry a CE mark.

1.6 Please provide details of all completed and ongoing studies from which additional evidence is likely to be available in the next 12 months for the indication being appraised.

There is one phase III randomised controlled trial (RCT) (LUME-Lung 1) that supports the use of nintedanib in this indication. Additional supporting information also comes from the phase III clinical trial LUME-Lung 2, which was solely used to inform the pre-specified statistical analysis of LUME-Lung 1 (please refer to section 6.3 for more detail).

LUME-Lung 1 (NCT00805194)(3)

A multicentre, randomised, double-blind, phase III trial designed to investigate the efficacy and safety of oral nintedanib plus standard docetaxel therapy compared to placebo plus standard docetaxel therapy in patients with stage IIIB/IV or recurrent NSCLC after failure of first-line chemotherapy.

LUME-Lung 2 (NCT00806819)(21)

A multicentre, randomised, double-blind, phase III trial designed to investigate the efficacy and safety of oral nintedanib plus standard pemetrexed therapy compared to placebo plus standard pemetrexed therapy in patients with stage IIIB/IV or recurrent non-squamous NSCLC after failure of first-line chemotherapy.

There are currently no ongoing nintedanib studies that are relevant to this indication.

1.7 If the technology has not been launched, please supply the anticipated date of availability in the UK.

As per the response to Question 1.3, our current estimation is that nintedanib will become available in [redacted], provided marketing authorisation is granted in [redacted]

1.8 Does the technology have regulatory approval outside the UK? If so, please provide details.

No, see Section 1.3.

1.9 Is the technology subject to any other form of health technology assessment in the UK? If so, what is the timescale for completion?

Yes, Boehringer Ingelheim Ltd. intends to make a full submission to the Scottish Medicines Consortium (SMC) in this indication on [redacted]. It is anticipated that advice will be issued to NHS Scotland in [redacted] and published on the SMC website in [redacted].

1.10 For pharmaceuticals, please complete the table below. If the unit cost of the pharmaceutical is not yet known, provide details of the anticipated unit cost, including the range of possible unit costs.

Table 4: Unit costs of technology being appraised

Pharmaceutical formulation	Soft gelatine capsule
Acquisition cost (excluding VAT)	£2151.10 pack
Method of administration	Oral. Capsules are to be swallowed whole with water, and taken with food.
Doses	200mg twice daily (BD) with the option to dose reduce in case of AEs to 150mg BD in a first dose reduction step, and 100mg BD if a second dose reduction is required (according to the protocol-defined dose reduction scheme). No dose increase is assumed after a dose reduction.
Dosing frequency	Administered twice daily, approximately every 12 hours.
Average length of a course of treatment	Treatment should be given continuously until tumour progression or unacceptable AEs. The median duration of nintedanib treatment in patients of adenocarcinoma tumour histology in the pivotal LUME-Lung 1 clinical trial was 4.2 months. The median number of docetaxel cycles received in the nintedanib arm was 5.
Average cost of a course of treatment	£1,505.70 per 21-day cycle
Anticipated average interval between courses of treatments	Nintedanib is administered continuously until disease progression or undue toxicity. Patients are therefore only expected to undergo one course of nintedanib treatment. Patients may however temporarily interrupt nintedanib treatment to recover from AEs, as per the nintedanib dose reduction and AE management recommendations.
Anticipated number of repeat courses of treatments	Nintedanib is administered continuously until disease progression or undue toxicity. Patients are therefore only expected to undergo one continuous course of nintedanib treatment. Patients may however temporarily interrupt nintedanib treatment to recover from AEs, as per the nintedanib dose reduction and AE management recommendations.
Dose adjustments	Two dose reductions are permitted with nintedanib, in case of AEs: from a starting dose of 200mg twice daily to 150mg twice daily in a first dose reduction step, and, if necessary, to 100mg twice daily in a second dose reduction step.

1.11 For devices, please provide the list price and average selling price. If the unit cost of the device is not yet known, provide details of the anticipated unit cost, including the range of possible unit costs.

Not applicable.

1.12 Are there additional tests or investigations needed for selection, or particular administration requirements for this technology?

In order to be eligible for nintedanib treatment, patients will need to have stage IIIB/IV or recurrent NSCLC of confirmed adenocarcinoma tumour histology.

Nintedanib needs to be administered in combination with docetaxel, for a minimum of four cycles of combination therapy, before it can be administered as monotherapy. The usual docetaxel administration requirements, including administration of pre-medications, and associated laboratory investigations, as per the docetaxel SPC and local clinical practice, will therefore apply.

1.13 Is there a need for monitoring of patients over and above usual clinical practice for this technology?

Hepatic transaminases, alkaline phosphatase and bilirubin levels will need to be monitored after the start of nintedanib therapy. The monitoring should occur periodically, i.e. at the beginning of each treatment cycle during nintedanib plus docetaxel combination therapy. Additional monitoring may be required in case of AEs

1.14 What other therapies, if any, are likely to be administered at the same time as the intervention as part of a course of treatment?

Nintedanib is administered alongside intravenous docetaxel 75mg/m^2. The docetaxel dose can be reduced to 60mg/m^2 as per the docetaxel SmPC and standard clinical practice. Docetaxel is administered on day 1 of each 21 day cycle. Nintedanib and docetaxel combination therapy needs to be given for a minimum of four cycles before nintedanib can be administered as monotherapy.

2 Context

In this background section the manufacturer or sponsor should contextualise the evidence relating to the decision problem.

2.1 Please provide a brief overview of the disease or condition for which the technology is being used. Include details of the underlying course of the disease.

Lung cancer is the second most common cancer in the UK; there are around 41,500 new cases diagnosed each year, with 35,406 new cases in England and Wales in 2010, and more than one in five cancer deaths (22%) in the UK are from lung cancer(22). Smoking causes more than 8 in 10 lung cancers in the UK(23).

The disease encompasses a complex family of neoplasms arising from the major bronchi or from the distant airway bronchioles and alveoli. It has two major classes: NSCLC, the most common type, accounting for 85% to 90% of cases; and small cell lung cancer. Adenocarcinoma is the most common histological sub-type of NSCLC(24). At diagnosis, 10 to 15% of patients have locally advanced cancer, i.e. stage IIIB and 40% of patients have metastatic cancer i.e. stage IV(25, 26). Patients with NSCLC have a poor prognosis that has not changed significantly in the past decades. Moreover, patients with stage IIIB and stage IV NSCLC have the lowest 5-year survival rate, at 5% and 1%, respectively (24, 27-29).

The Disease Course

Lung cancer does not usually cause noticeable symptoms until it has spread through much of the lungs or into other parts of the body. This is known as advanced or metastatic lung cancer. This means that the outlook for lung cancer is poor compared with other types of cancer (see section 2.3 for estimated life expectancy)(30).

The type of treatment for locally advanced or metastatic lung cancer depends on several factors, including:(31)

- Tumour histology
- The presence or absence of actionable mutations (i.e. epidermal growth factor receptor [EGFR] or anaplastic lymphoma kinase [ALK])
- The type of lung cancer (non-small cell or small cell)
- The size and position of the cancer
- How far advanced the cancer is (the stage)
- Overall health of the patient

Patients with advanced or metastatic cancer without actionable mutations usually receive platinum doublet chemotherapy in the first-line setting (pemetrexed plus cisplatin for adenocarcinoma and gemcitabine plus cisplatin for squamous cell NSCLC)(31).

Whilst the benefit of first-line chemotherapy patients with NSCLC with a good performance status is well established, approximately 30% to 50% of NSCLC patients will receive second-line treatment(29, 32, 33). The major goal of second-line treatment is to prolong life without worsening HRQL. There are a number of new therapies that target patients with relatively rare mutations (e.g. EGFR), but patients with adenocarcinomas and without actionable mutations who progress following first-line chemotherapy have limited therapy options. Following failure of first-line chemotherapy, treatment options are limited to docetaxel monotherapy or erlotinib(31, 34).

2.2 Please provide the number of patients covered by this particular therapeutic indication in the marketing authorisation and also including all therapeutic indications for the technology, or for which the technology is otherwise indicated, in England and Wales and provide the source of the data.

Based on the predicted population figures for England and Wales, it is estimated that there will be a total of 3,936 second-line stage IIIb/IV NSCLC patients with adenocarcinoma histology for each year from 2014 to 2018(33-36). Based on internal estimates, it is predicted that approximately 78.6% of these patients will be treated with first-line chemotherapy, of which 24.1% of patients will progress after first-line therapy and be eligible for second-

line(35). As a result, a total of 745 patients are expected to be eligible for second-line treatment of stage IIIb/IV NSCLC with Eastern Cooperative Oncology Group (ECOG) performance status (PS) 0-1 of the adenocarcinoma sub-type in England and Wales. As there is no population growth assumed, the eligible population remains constant from 2014 to 2018. For more details regarding the calculation of the population eligible for second-line treatment, please refer to section 8.

2.3 Please provide information about the life expectancy of people with the disease in England and Wales and provide the source of the data.

In the most recent cancer survival publication by the Office of National Statistics, the 1- and 5-year survival rates of lung cancer patients in England, diagnosed between 2006 and 2010, were reported as 31.6% and 9.8% respectively(36). Similar survival rates were found in a separate study(37) conducted by the International Cancer Benchmarking Partnership, in patients diagnosed between 1995 and 2007. In this study, the one- and five-year survival rates of lung cancer patients in England were reported to be 29.7% and 8.7% respectively between 2005 and 2007. This study also found the corresponding survival rates in Wales to be 28.5% and 9.0%(37). Moreover, patients with stage IIIB and stage IV NSCLC have the lowest 5-year survival rate, at 5% and 1%, respectively (24, 27-29).

2.4 Please give details of any relevant NICE guidance or protocols for the condition for which the technology is being used. Specify whether any specific subgroups were addressed.

NICE clinical guideline 121 – Lung cancer: the diagnosis and treatment of lung cancer (31).

NICE quality standard 17 – Lung cancer for adults (38).

NICE technology appraisal (TA)162 – Erlotinib for the treatment of NSCLC (34).

NICE pathway – Lung cancer (39).

2.5 Please present the clinical pathway of care that depicts the context of the proposed use of the technology. Explain how the new technology may change the existing pathway. If a relevant NICE clinical guideline has been published, the response to this question should be consistent with the guideline and any differences should be explained.

Current clinical pathway of care for advanced or metastatic NSCLC

First-line treatment

1. Erlotinib(40)

- Erlotinib is recommended as an option for the first-line treatment of people with locally advanced or metastatic NSCLC if:
- they test positive for the epidermal growth factor receptor tyrosine kinase (EGFR-TK) mutation and
- the manufacturer provides erlotinib at the discounted price agreed under the patient access scheme (as revised in 2012).

2. Gefitinib(41)

- Gefitinib is recommended as an option for the first-line treatment of people with locally advanced or metastatic NSCLC if:
- they test positive for the EGFR-TK mutation and
- the manufacturer provides gefitinib at the fixed price agreed under the patient access scheme.

3. Pemetrexed(42)

- Pemetrexed in combination with cisplatin is recommended as an option for the first-line treatment of patients with locally advanced or metastatic NSCLC only if the histology of the tumour has been confirmed as adenocarcinoma or large-cell carcinoma.
- People who are currently being treated with pemetrexed for NSCLC but who do not meet the criteria above should have the option to continue their therapy until they and their clinicians consider it appropriate to stop.

4. Afatinib(43)

- Afatinib is recommended as an option, within its marketing authorisation, for treating adults with locally advanced or metastatic NSCLC only if:
- the tumour tests positive for the EGFR-TK mutation and
- the person has not previously had an EGFR-TK inhibitor and
- the manufacturer provides afatinib with the discount agreed in the patient access scheme.

Second-line treatment

1. Docetaxel monotherapy(31)

- Docetaxel monotherapy can be considered for second-line treatment of locally advanced or metastatic NSCLC when cancer has relapsed after previous chemotherapy.

2. Erlotinib(34)

- Erlotinib is recommended, within its licensed indication, as an alternative to docetaxel as a second-line treatment option for patients with NSCLC only on the basis that it is provided by the manufacturer at an overall treatment cost (including administration, AEs and monitoring costs) equal to that of docetaxel.

3rd and subsequent lines of therapy

Currently, there are no NICE-recommended technologies for 3rd and subsequent lines of treatment of patients with advanced or metastatic NSCLC.

Maintenance therapy

NICE recommends pemetrexed as a possible maintenance treatment for some people with NSCLC. A patient should be eligible to have pemetrexed if all of the following apply (44):

- locally advanced or metastatic NSCLC
- squamous cell carcinoma is not the main type of cancer
- not received pemetrexed and cisplatin together as a first-line treatment
- condition did not worsen immediately after the patient received platinum-based chemotherapy together with gemcitabine, paclitaxel or docetaxel.

Changes to the clinical pathway of care for advanced or metastatic NSCLC

Nintedanib fits well in the existing clinical pathway and can complement docetaxel treatment as an effective second-line option for patients with locally advanced/metastatic or recurrent NSCLC of adenocarcinoma tumour histology, previously treated with one line of chemotherapy.

2.6 Please describe any issues relating to current clinical practice, including any variations or uncertainty about best practice.

TA162(34) is currently undergoing an appraisal consultation; the appraisal consultation document from February 2014 states that the appraisal committee's preliminary recommendations are(45):

- Erlotinib is recommended as an option for treating locally advanced or metastatic NSCLC in people who have received non-targeted chemotherapy because of delayed confirmation that their tumour is EGFR-TK mutation-positive, only if the manufacturer provides erlotinib with the discount agreed in the patient access scheme.
- Erlotinib is not recommended for treating locally advanced or metastatic NSCLC in people with EGFR-TK mutation-negative tumours after the failure of at least 1 prior non-targeted chemotherapy regimen.
- Erlotinib is recommended as an option for treating locally advanced or metastatic NSCLC that has progressed after chemotherapy in people with tumours of unknown EGFR-TK mutation status, only if:
    - the result of a EGFR-TK mutation diagnostic test is unobtainable because of an inadequate tissue sample or poor quality DNA and
    - the tumour is very likely to be EGFR-TK mutation-positive based on patient characteristics and
    - the person's disease responds to the first 2 cycles of treatment with erlotinib and
    - the manufacturer provides erlotinib with the discount agreed in the patient access scheme.

Since then, this draft guidance has been withdrawn, however the 3rd NICE Appraisal Committee meeting for erlotinib in second-line NSCLC was held on the 8th July. The result of this meeting will be announced shortly.

Nintedanib would provide a treatment option for adenocarcinoma patients in the second-line. This option would provide available treatment in addition to docetaxel.

2.7 Please identify the main comparator(s) and justify their selection.

As an established treatment option in England and Wales, docetaxel monotherapy is considered as the primary comparator for the second-line treatment of locally advanced or metastatic NSCLC in patients whom have relapsed after previous chemotherapy.

Within its licenced indication, erlotinib is also recommended as a second-line option in England and Wales as an alternative to docetaxel monotherapy. Therefore, erlotinib is considered as an additional comparator. Please note that in an advisory board on the 10th April 2014, the opinion of all five leading UK clinicians was that patients fit enough for treatment with docetaxel would receive docetaxel rather than erlotinib (see section 7.3.5).

2.8 Please list therapies that may be prescribed to manage adverse reactions associated with the technology being appraised.

In general, there is no specific medicine that has to be given in conjunction with nintedanib plus docetaxel therapy in order to manage AEs. However, most AEs associated with the treatment are consistent with the known safety profile of the drug (diarrhoea, nausea, vomiting, and ALT/AST increase). These can be treated with dose reduction of nintedanib, dose interruption and/or symptomatic treatment according to standard clinical practice.

2.9 Please identify the main resource use to the NHS associated with the technology being appraised. Describe the location of care, staff usage, administration costs, monitoring and tests. Provide details of data sources used to inform resource estimates and values.

The main resource use associated with the use of nintedanib is its acquisition cost. It does not require resource of any other kind in terms of administration, monitoring or tests over and above routine clinical practice.

2.10 Does the technology require additional infrastructure to be put in place?

No additional infrastructure is required.

3 Equality

NICE is committed to promoting equality of opportunity, eliminating unlawful discrimination and fostering good relations between people with particular protected characteristics and others. For further information, please see the NICE website (www.nice.org.uk/aboutnice/howwework/NICEEqualityScheme.jsp).

3.1 Identification of equality issues

3.1.1 Please let us know if you think that this appraisal:

- could exclude from full consideration any people protected by the equality legislation who fall within the patient population for which [the treatment(s)] is/are/will be licensed;
- could lead to recommendations that have a different impact on people protected by the equality legislation than on the wider population, e.g. by making it more difficult in practice for a specific group to access the technology
- could lead to recommendations that have any adverse impact on people with a particular disability or disabilities

  Please provide us with any evidence that would enable the Committee to identify and consider such impacts.

Boehringer Ingelheim Ltd. does not believe that nintedanib will be associated with any equality issues.

3.1.2 How has the analysis addressed these issues?

Not applicable.

4 Innovation

4.1.1 Discuss whether and how you consider the technology to be innovative in its potential to make a significant and substantial impact on health-related benefits, and whether and how the technology is a 'step-change' in the management of the condition.

Over the previous decade, no phase III study with any combination of agents has demonstrated an OS benefit versus an active comparator in second-line stage IIIB/IV NSCLC patients with adenocarcinoma histology(46-53). Thus, there is a high unmet need to improve the treatment options for these patients, in the second-line setting. It has been postulated that an effective strategy for this class of drugs may be the use of an agent targeting more than one angiogenic pathway, in combination with chemotherapy(54, 55). Moreover, anti-angiogenic agents may be particularly effective in patients with adenocarcinoma histology, which has been characterised as having higher levels of microvessel density compared to other NSCLC histological subtypes(56, 57).

As a unique, oral triple angiokinase inhibitor simultaneously acting on three receptor classes that have a role in angiogenesis (VEGFR, PDGFR and FGFR), nintedanib has the potential to offer important clinical benefits across a broad range of cancers. In the UK, nintedanib fits into the care pathway as an add-on therapy to docetaxel for the treatment of patients with locally advanced, metastatic or recurrent NSCLC of adenocarcinoma histology, who have progressed after one prior line of chemotherapy. Docetaxel is currently recommended by NICE Clinical Guideline 121 as a second-line therapy for patients with locally advanced or metastatic NSCLC in whom relapse has occurred after previous chemotherapy(31). However, there is no NICE recommended add-on therapy to second-line docetaxel(31). Erlotinib is recommended, within its licenced indication, as an alternative to second-line docetaxel treatment only on the basis that it is provided by the manufacturer at an overall treatment cost (including administration, AEs and monitoring costs) equal to that of docetaxel(34).

Nintedanib, in addition to docetaxel, would provide an alternative treatment option for adenocarcinoma patients in the second-line. This treatment combination would be the first treatment to demonstrate a significant and clinically meaningful OS benefit for second-line adenocarcinoma patients versus an active agent in a phase III clinical trial.(3)

4.1.2 Discuss whether and how you consider that the use of the technology can result in any potential significant and substantial health-related benefits that are unlikely to be included in the quality-adjusted life year (QALY) calculation.

As the cost effectiveness of nintedanib has been calculated from a payer perspective, the impact on carers has not been included in the quality-adjusted life year (QALY). The impact of prolonged PFS may be expected to result in an improvement in carers' quality of life (QoL); this would not be captured in the QALY.

Other aspects of the patient experience such as the psychological impact of an extension in OS are also expected to result in improved QoL. These are unlikely to be fully captured in the QALY, and as a result the increase in QALYs resulting from nintedanib + docetaxel vs. comparators is likely to be a conservative assumption.

4.1.3 Please identify the data you have used to make these judgements, to enable the Appraisal Committee to take account of these benefits.

The submission does not contain numerical values on the non-QALY benefits detailed above.

5 Statement of the decision problem

In this section the manufacturer or sponsor should specify the decision problem that the submission addresses. The decision problem should be derived from the final scope issued by NICE and should state the key parameters that the information in the evidence submission will address.

Table 5: Decision problem addressed in this submission

	Final scope issued by NICE	**Decision problem addressed in the submission**	**Rationale if different from the scope**
Population	Patients with locally advanced, metastatic or recurrent NSCLC of adenocarcinoma tumour histology after first-line chemotherapy.	As in final scope.	Not applicable
Intervention	Nintedanib	As in final scope.	Not applicable

	Final scope issued by NICE	Decision problem addressed in the submission	Rationale if different from the scope
Comparator(s)	Docetaxel monotherapy Erlotinib	As in final scope, docetaxel monotherapy is the comparator for the primary analysis. Comparisons versus erlotinib are also presented as secondary analyses. No other agents are licenced or routinely used for this indication (pemetrexed is licensed but not NICE approved). Therefore, no other comparisons are presented.	Not applicable
Outcomes	The outcome measures to be considered include: • PFS by central independent review • OS • best tumour response (according to modified RECIST v1.0 criteria) • HRQL measured by standard questionnaires (health status self-assessment questionnaire: EQ-5D) • AEs of treatment	As in final scope. Each of these outcomes is considered.	Not applicable
Economic analysis	Cost effectiveness of treatments should be expressed in terms of incremental cost per QALY. 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 and NHS and Personal Social Services perspective.	As in final scope. Results are expressed in terms of incremental cost per QALY gained. Various time horizons are presented with lifetime (15 years) being that of the primary analysis (appropriate for a condition such as lung cancer, with low survival rates). Costs are considered from the NHS and PSS perspective	Not applicable
Subgroups to be considered	None	Not applicable	Not applicable

	Final scope issued by NICE	Decision problem addressed in the submission	Rationale if different from the scope
Special considerations, including issues related to equity or equality		Not applicable	

Section B – Clinical and cost effectiveness

When estimating clinical and cost effectiveness, particular emphasis should be given to adhering to the 'reference case' (see the NICE document 'Guide to the methods of technology appraisal' – www.nice.org.uk). Reasons for deviating from the reference case should be clearly explained. Particularly important features of the reference case include those listed in the table below.

Element of health technology assessment	Reference case	Section in 'Guide to the methods of technology appraisal'
Defining the decision problem	The scope developed by NICE	5.2.5 and 5.2.6
Comparator(s)	Therapies routinely used in the NHS, including technologies regarded as current best practice	5.2.5 and 5.2.6
Perspective costs	NHS and PSS	5.2.7 to 5.2.10
Perspective benefits	All health effects on individuals	5.2.7 to 5.2.10
Type of economic evaluation	Cost-effectiveness analysis	5.2.11 and 5.2.12
Synthesis of evidence on outcomes	Based on a systematic review	5.3
Measure of health effects	QALYs	5.4
Source of data for measurement of HRQL	Reported directly by patients and carers	5.4
Source of preference data for valuation of changes in HRQL	Representative sample of the public	5.4
Discount rate	An annual rate of 3.5% on both costs and health effects	5.6
Equity weighting	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	5.12
HRQL, health-related quality of life; NHS, National Health Service; PSS, personal social services; QALY(s), quality-adjusted life year(s)		

6 Clinical evidence

Manufacturers and sponsors are requested to present clinical evidence for their technology in the following sections. This section should be read in conjunction with NICE's 'Guide to the methods of technology appraisal', sections 3 and 5.3.1 to 5.3.8.

- **The pivotal phase III clinical trial LUME-Lung 1 investigated the efficacy and safety of nintedanib in combination with docetaxel compared with placebo plus docetaxel in patients with advanced, metastatic or recurrent NSCLC after failure of first-line chemotherapy(3)**
- **In LUME-Lung 1, statistically significant improvements were observed for the primary endpoint of 'centrally assessed PFS' in all patients, regardless of histology, and for the key secondary endpoint of OS in patients with tumours of adenocarcinoma histology(3)**
- **Second-line therapy with nintedanib in combination with docetaxel significantly prolonged median PFS in patients with tumours of adenocarcinoma histology (4.0 vs 2.8 months; HR 0.77, p=0.0193), compared with docetaxel alone(3, 58)**
- **Median OS was significantly prolonged in patients with tumours of adenocarcinoma histology who received second-line treatment with nintedanib in combination with docetaxel, compared to docetaxel alone (12.6 vs 10.3 months; HR 0.83,p=0.0359)(3)**
- **In adenocarcinoma patients the disease control rate was significantly improved in the nintedanib plus docetaxel arm, compared with placebo plus docetaxel (60.2% and 44.0% respectively; odds ratio [OR] 1.93, p<0.0001). The objective response rate was comparable across arms(3)**
- **The addition of nintedanib to docetaxel resulted in a slightly higher incidence of treatment-related AEs, and AEs of CTCAE ≥Grade 3, compared with the placebo plus docetaxel arm. The discontinuation rate was also comparable across arms, suggesting that the AEs were manageable. In addition, the incidence of AEs commonly associated with anti-angiogenic compounds was low and comparable across arms(5)**
- **The significant OS benefit observed in adenocarcinoma patients with the addition of nintedanib to docetaxel therapy was achieved with no detrimental effect on patient self-reported HRQL(59)**

6.1 Identification of studies

6.1.1 Describe the strategies used to retrieve relevant clinical data, both from the published literature and from unpublished data that may be held by the manufacturer or sponsor. The methods used should be justified with reference to the decision problem. Sufficient detail should be provided to enable the methods to be reproduced, and the rationale for any inclusion and exclusion criteria used should be provided. Exact details of the search strategy used should be provided in section 10.2, appendix 2.

A combined, single systematic review of the literature was conducted to identify all potentially relevant published and non-published RCTs on investigating the efficacy and safety of second-line treatments for patients with NSCLC, the economic and humanistic burden of NSCLC, and economic evaluations of second-line treatments for NSCLC. The following key indexed-databases were systematically searched:

- Clinical Efficacy and Safety Review: MEDLINE and MEDLINE R-IN PROCESS (via PubMed), EMBASE, and Cochrane Library (Central and Cochrane Reviews)
- Humanistic Review: MEDLINE, EMBASE, and Cochrane Library (NHS EED)
- Economic Models Review: MEDLINE, EMBASE, Cochrane Library (NHS EED), Health Economic Evaluation Database (HEED) and EconLit.

Across all topics, the bibliography lists of relevant systematic literature reviews (SLRs) were manually searched, as were the following 'grey' literature sources:

- Clinicaltrials.gov
- American Society of Clinical Oncology (ASCO) conference proceedings for 2011-2014
- European Society for Medical Oncology (ESMO) conference proceedings for 2011-2014
- National Guidelines Clearinghouse.

These sources were selected as being those that were most likely to have relevant data on the topics of interest, and therefore offered the most efficient way of identifying data to support the analysis of clinical and cost-effectiveness of nintedanib plus docetaxel. The search strategies for the different topics are reported in Appendix 10.2 (clinical), Appendix 10.10 (economic) and Appendix 10.12 (humanistic).

The database searches were last performed on 28 February 2014. The citation lists of relevant systematic reviews published since 2009 were also examined to identify other relevant studies.

6.2 Study selection

6.2.1 *Describe the inclusion and exclusion selection criteria, language restrictions and the study selection process. A justification should be provided to ensure that the rationale is transparent. A suggested format is provided below.*

The abstracts obtained from the database search were each examined manually by two researchers applying a set of predefined inclusion criteria described below. Following this, a random sample of excluded abstracts was checked for accuracy by a third researcher to confirm the exclusion decisions. Any discrepancy in the decision to include or exclude a study was reviewed by and resolved between researchers. The full-text articles for abstracts deemed potentially relevant during this first level of screening were retrieved in order to confirm their inclusion in the review. All full-text publications were independently reviewed by two researchers, with all disagreements being resolved by consensus.

The results of this search are used to complete this section, in which clinical evidence for nintedanib is presented. Search criteria for the humanistic review and economic models review can be found in Appendix 10.12.4 and Appendix 10.10.4., respectively.

Table 6: Clinical efficacy and safety review: inclusion and exclusion criteria

	Inclusion criteria	**Exclusion criteria**
Population	**Relapsed or refractory NSCLC (RR NSCLC)** Adults with histologically or cytologically confirmed, locally advanced and/or metastatic NSCLC of stage IIIB or IV (according to American Joint Committee on Cancers) or recurrent NSCLC (all histologies): • Squamous-cell carcinoma • Adenocarcinoma • Large cell carcinoma	Any patient population other than RR NSCLC
Interventions	Any second-line pharmacological treatment for RR NSCLC • Monotherapy • Combination chemotherapy	Patients who were treatment-naïve, had received more than first-line therapy, or had received only non-pharmacological interventions
Outcomes	Relevant outcomes for full-text inclusion: • OS and PFS • Time to relapse • Time to death • AEs (all Grades and Grade 3 to 4)	No outcomes of interest

	Inclusion criteria	Exclusion criteria
	• Withdrawals • Mean dose and number of cycles of therapy received	
Study design	Randomised controlled trials (RCTs) only	Not an RCT (e.g. observational)
Language restrictions	Any language‡	
Date	2000 onwards*	Prior to 2000*
Country	Any	None

‡ Non-English-language publications were identified for the efficacy review but none met the inclusion criteria.

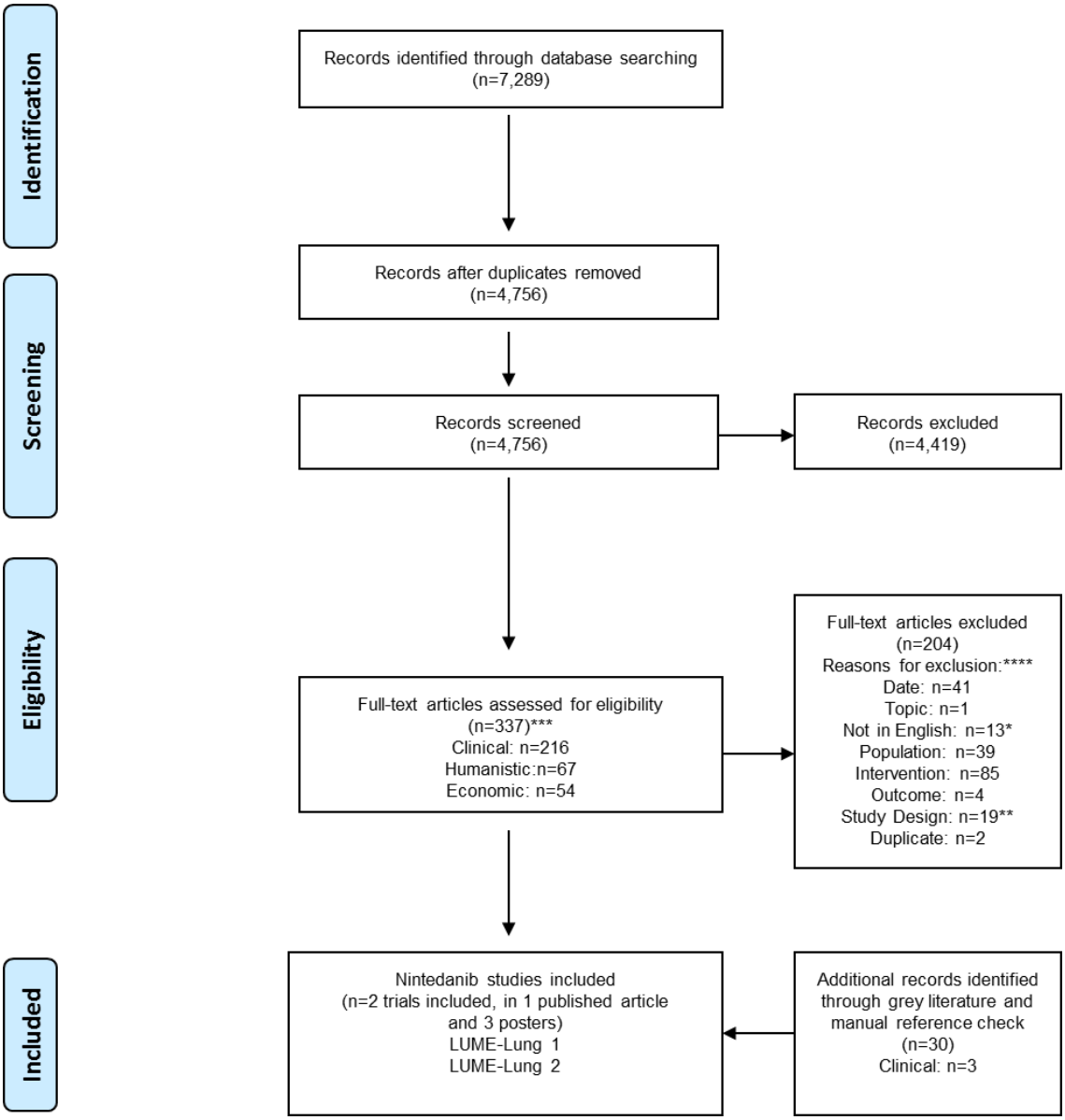
*Abstracts published prior to the year 2011 and systematic reviews published prior to the year 2009 were excluded.

6.2.2 *A flow diagram of the numbers of studies included and excluded at each stage should be provided using a validated statement for reporting systematic reviews and meta-analyses such as the QUOROM statement flow diagram (www.consort-statement.org/?o=1065). The total number of studies in the statement should equal the total number of studies listed in section 6.2.4.*

The inclusion criteria of the search were made wide enough to enable the identification of relevant studies investigating any of the interventions licensed for the treatment of locally advanced or metastatic NSCLC. However, as this section presents nintedanib clinical evidence only, all non-nintedanib studies were subsequently excluded from the results of the search.

The search of the literature yielded 7,289 citations. De-duplication resulted in the removal of 2,533 overlapping citations. Following screening of the remaining 4,756 studies, 4,419 studies were excluded. Full text was obtained for the remaining 337 studies. Following the application of exclusion criteria, most notably the requirement for nintedanib administration, two trials remained (LUME-Lung 1 and LUME-Lung 2). The flow of studies in the systematic literature review is presented in Figure 1.

Figure 1: PRISMA flow diagram for systematic literature reviews of nintedanib clinical studies



* No relevant non-English language articles were identified for the efficacy review. The humanistic and economic burden reviews excluded studies not published in English, to focus on studies of most relevance to the UK setting.
** The reference lists of the systematic reviews were assessed for additional relevant studies; no additional studies were identified.
***Some publications report on more than one topic; these counts do not reflect studies reporting more than one topic.
****More than one reason for exclusion may have applied per study

6.2.3 When data from a single RCT have been drawn from more than one source (for example, a poster and a published report) and/or when trials are linked (for example, an open-label extension to an RCT), this should be made clear.

Information on the pivotal RCT for nintedanib, LUME-Lung 1, has been drawn from the following documents:

- Reck et al (2014) Lancet Oncology – LUME-Lung 1 study publication(3)
- BI, Data on file. LUME-Lung 1 clinical trial report (CTR) – final OS analysis(5)
- BI, Data on file. LUME-Lung 1 CTR – primary PFS analysis(4)
- BI, Data on file. LUME-Lung 1 trial statistical analysis plan (TSAP)(60)
- BI, Data on file. LUME-LUNG 1 TSAP addendum(61)
- Novello et al (2013) (poster presented at WCLC 2013 meeting)(59)
- Kaiser et al (2013) (poster presented at European Cancer Congress 2013)(62)
- Reck et al (2013) (presentation at ASCO 2013)(58)
- BI, Data on file. Summary of clinical efficacy(63)
- BI, Data on file. LUME-Lung 1 CTR – final OS analysis appendix(64)

Complete list of relevant RCTs

6.2.4 Provide details of all RCTs that compare the intervention with other therapies (including placebo) in the relevant patient group. The list must be complete and will be validated by independent searches conducted by the Evidence Review Group. This should be presented in tabular form. A suggested format is presented below.

This submission is based on clinical data from the pivotal trial LUME-Lung 1. The study was a phase III, international, multicentre, randomised, double-blind, placebo-controlled, parallel-group trial investigating the efficacy and safety of nintedanib in combination with docetaxel compared to placebo plus docetaxel in patients with advanced, metastatic (stage IIIB/IV) or recurrent NSCLC after failure of first-line chemotherapy.

It should be noted that although the LUME-Lung 1 trial included patients with stage IIIB/IV or recurrent NSCLC of all histological sub-types (i.e. adenocarcinoma, squamous cell carcinoma, large cell carcinoma, combination and unspecified histology), the focus of this submission will be on patients with adenocarcinoma tumour histology, in accordance with the anticipated marketing authorisation for nintedanib in the EU(3).

In addition, a second international, multicentre, randomised, double-blind, placebo-controlled phase III trial, LUME-Lung 2, was also conducted. In LUME-Lung 2, patients with advanced, metastatic (stage IIIB/IV) or recurrent NSCLC of non-squamous histology (adenocarcinoma, large cell carcinoma and unspecified non-squamous histology) who progressed after first-line treatment with chemotherapy received either nintedanib or matching placebo, in combination with pemetrexed. The study design and the study endpoints were comparable to those in LUME-Lung 1(21).

Based on a pre-planned futility analysis of investigator-assessed PFS by an external Data Monitoring Committee (DMC), the LUME-Lung 2 study was stopped after randomising 713/1,300 planned patients on 18 June 2011(61). To better understand the futility outcome of LUME-Lung 2 and to identify a patient population that would benefit from treatment with nintedanib, further detailed analyses were performed(62). Prognostic baseline variables were initially identified in the placebo arm of the LUME-Lung 2 study. The interaction of identified prognostic variables with treatment was then explored to identify variables that were also predictive of a nintedanib treatment benefit. This was done using centrally assessed PFS data and interim OS data obtained at the time of the primary analysis of the LUME-Lung 1 and LUME-Lung 2 trials.

An inverse relationship between the length of time since start of first-line therapy and the treatment effect of nintedanib plus second-line chemotherapy was shown for PFS and OS; the shorter the time from start of first-line therapy to randomisation, the better the treatment effect. To categorise the continuous variable ‘time since start of first-line therapy’, a cut-off of 9 months was chosen based on the width of the 95% CI and the time when the upper boundary of the 95% CI approached a HR of 1 (T<9m)(62).

This hypothesis was to be validated using final OS data from the LUME-Lung 1 trial. The LUME-Lung 1 statistical analysis plan was therefore amended, prior to database lock and unblinding of data for the final OS analysis. Additional details on the hypothesis generation using LUME-Lung 2 data can be seen in Section 6.3.6. The phase III studies in the nintedanib NSCLC clinical development programme are summarised in Table 7.

Table 7: List of relevant RCTs

Trial no. (acronym)	Intervention	Comparator	Population	Primary study ref.
LUME-Lung 1 **(1199.13; (NCT00805194)(3)**	Nintedanib 200mg twice daily, orally, on days 2 to 21 of a 21-day cycle in combination with docetaxel 75mg/m^2 IV on day 1 of a 21-day cycle	Matched placebo twice daily on days 2 to 21 of a 21-day cycle in combination with docetaxel 75mg/m^2 IV on day 1 of a 21-day cycle	• n=1,314 • Histologically or cytologically confirmed stage IIIB or IV, or recurrent NSCLC with relapse or failure of 1 prior first-line chemotherapy • ECOG PS of 0 or 1 • ≥18 years	*Reck M et al. Lancet Oncol 2014;15(2):143-155*
LUME-Lung 2 **(1199.14; NCT00806819)(21)**	Nintedanib 200mg twice daily, orally, on days 2 to 21 of a 21-day cycle in combination with pemetrexed 500mg/m^2 IV on day 1 of a 21-day cycle	Matched placebo twice daily on days 2 to 21 of a 21-day cycle in combination with pemetrexed 500mg/m^2 IV on day 1 of a 21-day cycle	• n=713 • Histologically or cytologically confirmed stage IIIB or IV, or recurrent NSCLC of non-squamous histology (adenocarcinoma, large cell carcinoma and unspecified non-squamous), with relapse or failure of 1 prior first-line chemotherapy • ECOG PS of 0 or 1 • ≥18 years	*Hanna et al. 2013 ASCO abstract*

ECOG = Eastern Cooperative Oncology Group; NSCLC = non-small lung cancer; PS = performance status

6.2.5 *Please highlight which of the RCTs identified above compares the intervention directly with the appropriate comparator(s) with reference to the decision problem. If there are none, please state this.*

The LUME-Lung 1 study compared nintedanib against placebo in combination with a standard second-line NSCLC treatment option (i.e. docetaxel)(3). Docetaxel is an established treatment option in England and Wales. It is recommended by NICE as monotherapy if second-line treatment is appropriate for patients with locally advanced or metastatic NSCLC in whom relapse has occurred after previous chemotherapy(31).

6.2.6 X*When studies identified above have been excluded from further discussion, a justification should be provided to ensure that the rationale for doing so is transparent. For example, when studies have been identified but there is no access to the level of trial data required, this should be indicated.*

The LUME-Lung 2 study, summarised in Table 7, generated hypotheses that informed the statistical analysis of LUME-Lung 1. Details of the hypothesis generation and subsequent analysis are presented in Section 6.3.6. No other data from this trial is relevant to the decision problem as the comparator in this trial is not relevant to clinical practice in England and Wales. Pemetrexed monotherapy is not recommended by NICE for the treatment of locally advanced or metastatic NSCLC in patients who have had prior chemotherapy(65).

List of relevant non-RCTs

6.2.7 *Please provide details of any non-RCTs (for example experimental and observational data) that are considered relevant to the decision problem and a justification for their inclusion. Full details should be provided in section 6.8 and key details should be presented in a table; the following is a suggested format.*

No relevant non-RCT data have been identified, and none are therefore included in this submission.

6.3 Summary of methodology of relevant RCTs

6.3.1 As a minimum, the summary should include information on the RCT(s) under the subheadings listed in this section. Items 2 to 14 of the CONSORT checklist should be provided, as well as a CONSORT flow diagram of patient numbers (www.consort-statement.org). It is expected that all key aspects of methodology will be in the public domain; if a manufacturer or sponsor wishes to submit aspects of the methodology in confidence, prior agreement must be requested from NICE. When there is more than one RCT, the information should be tabulated.

Scientific background/study rationale

In the UK, lung cancer is the second most common type of cancer(23) and the most common cause of cancer death(66). Data for 2012 from the National Lung Cancer Audit report show that histologically confirmed NSCLC accounted for 63.2% of all lung cancers in England and Wales(67). This equated to 20,881 cases in 2012 that were submitted for the audit(67). Adenocarcinoma is the most prevalent histological subtype of NSCLC cases (39.6% in English Cancer Networks in 2010) (68).

Current NICE clinical guidelines recommend docetaxel for the second-line treatment of locally advanced or metastatic NSCLC(31). Erlotinib is also recommended as a second-line treatment option, within its licensed indication, as an alternative to docetaxel(34). Docetaxel was first recommended as a treatment option in the second-line setting in 2001,(69) while erlotinib was first recommended in 2008(34).

An analysis based on 120,745 patients with NSCLC diagnosed in England between 1 January 2004 and 31 December 2010 has found that, despite the emergence of new targeted treatments such as erlotinib, survival among patients with NSCLC (across all lines of treatment) has remained static. The proportion of NSCLC patients surviving for more than 1 year was 34.5% in 2004/2005 and 34.0% in 2010(70). Over the previous decade, no phase III study with any combination of agents has demonstrated an OS benefit versus an active comparator in second-line NSCLC patients with purely adenocarcinoma histology(46-49, 51-53, 71-80). Thus, there is a high unmet need to improve the treatment options for these patients, in the second-line setting.

Angiogenesis is an essential process in healthy individuals that can be utilised by tumours, including NSCLC, for the supply of oxygen and other nutrients that are necessary for growth and metastasis(54). Angiogenesis is enabled by interactions between growth factors with their cognate angiokinase receptors, examples include:

- Vascular endothelial growth factor (VEGF), which binds to and activates VEGFR 1–3(81). VEGFR-2 is considered to be the crucial receptor involved in the formation as well as the maintenance of tumour vasculature, including in NSCLC tumours(17).
- Platelet-derived growth factor (PDGF), which binds to and activates PDGFR, and has been shown to be important in NSCLC tumours(17, 82).
- Fibroblast growth factor (FGF), which binds to and activates FGFR 1–3(83).

The monoclonal antibody bevacizumab, which inhibits angiogenesis by targeting only VEGF, has demonstrated efficacy in first-line NSCLC patients in combination with chemotherapy(84, 85). A triple angiokinase inhibitor in combination with chemotherapy which inhibits VEGF, PDGF and FGF and therefore targets more than one angiogenic pathway, has the potential to improve the therapeutic outcomes for patients with NSCLC(54, 55). Moreover, anti-angiogenic agents may be particularly effective in patients with adenocarcinoma histology, which has been characterised as having higher levels of microvessel density compared to other NSCLC histological subtypes(56, 57). X Nintedanib is an oral, triple anti-angiogenesis agent that inhibits FGFR and VEGFR in endothelial cells; PDGFR in pericytes; and FGFR and PDGFR in smooth muscle cells. Inhibition of the activity of these receptors by nintedanib leads to reduced cell proliferation and to apoptosis in vitro and inhibition of blood vessel formation within the tumour in xenograft models (including lung cancer models), leading to reduced vessel density in tumours and ultimately tumour growth inhibition(1, 86). Inhibition of these receptors may also interfere with autocrine and paracrine stimulation of tumour angiogenesis via activation loops utilised by perivascular cells such as pericytes and vascular smooth muscle cells(1). On the molecular level, nintedanib is thought to inhibit the signalling cascade mediating angiogenesis by binding to the adenosine triphosphate (ATP) binding pocket of the receptor kinase domain, thus interfering with cross-phosphorylation of the receptor homodimers and their subsequent activation(1).

In phase I trials nintedanib has displayed an acceptable tolerability profile and promising tumour response in a variety of solid tumours, including NSCLC(87-91). Phase I trials have investigated nintedanib as monotherapy and in combination with pemetrexed, docetaxel, paclitaxel/carboplatin or the FOLFOX6 regimen(87-91). The maximum tolerated dose of nintedanib in combination with

chemotherapy was established as 200mg in two phase I trials in second-line NSCLC patients(89, 90). Gastrointestinal disorders, liver enzyme elevations and fatigue were the most frequent AEs in these patients(89).

A double-blind, randomised phase II trial investigated nintedanib monotherapy (200mg or 150mg bid) in 73 patients with locally advanced and/or metastatic NSCLC who had failed first- or second-line platinum-based chemotherapy(2). Median PFS based on investigator assessment was 6.9 weeks and median OS was 21.9 weeks, while disease control was achieved in 46% of patients(2). There was no statistically significant difference in efficacy between doses. However, PFS was longer in patients with baseline ECOG 0–1 than in those with ECOG 2 (11.6 vs 6 weeks; HR = 3.2, p=0.0002)(2). The pattern of AEs was similar to that observed in phase I trials, with gastrointestinal disorders and liver enzyme increases the most frequently reported events(2). Based on the tolerability profile and efficacy signals in patients with advanced/metastatic or recurrent NSCLC in the phase I and II trials, the pivotal phase III LUME-Lung 1 trial was initiated.

Docetaxel

Docetaxel is an anti-neoplastic agent which acts by promoting the assembly of tubulin into stable microtubules and inhibiting their disassembly. It has been shown in vitro to disrupt the cell's microtubule network which is essential for mitosis(92). In addition, it has been shown to be a potent inhibitor of angiogenesis in vitro and in vivo through inhibition of endothelial cell migration and microvessel formation(93, 94). Docetaxel is approved for the treatment of a number of cancers and recommended by NICE for consideration where second-line treatment is appropriate for patients with locally advanced or metastatic NSCLC when relapse has occurred after prior chemotherapy(31, 69). The most commonly reported AEs associated with docetaxel monotherapy are: neutropenia, anaemia, alopecia, nausea, vomiting, stomatitis, diarrhoea and asthenia(92).

Additive or synergistic effects with nintedanib and docetaxel combinations have been demonstrated in preclinical models(5), which may be due to the independent pharmacological mechanisms by which each drug inhibits tumour growth and angiogenesis(1, 93, 94). These additive or synergistic effects may be expected in the treatment of NSCLC patients(4).

A phase I trial, combining nintedanib, docetaxel (75mg/m^2) and prednisone (5 mg bid) in patients with hormone refractory prostate cancer has shown no indication that nintedanib exacerbates the AEs commonly associated with docetaxel treatment and there has been no indication of clinically

significant pharmacokinetic interactions(4). Moreover, the phase III LUME-Lung 2 trial of nintedanib in combination with chemotherapy (pemetrexed) has shown that nintedanib has a manageable safety profile and improves PFS compared with chemotherapy alone, in an analysis conducted after early study termination(21).

Study objectives

The pivotal trial LUME-Lung 1 (clinical trial number: NCT00805194) was an international, multicentre, randomised, double-blind, phase III trial investigating the efficacy and safety of oral nintedanib in combination with standard docetaxel therapy compared to placebo plus standard docetaxel therapy in patients with stage IIIB/IV recurrent NSCLC after failure of first-line chemotherapy. A summary of the LUME-Lung 1 trial methodology is provided in Table 8 and is discussed in detail in the following sections.

Study hypothesis

- H_0: The PFS time for patients treated with nintedanib plus docetaxel is equal to the PFS time for patients treated with placebo plus docetaxel((5).
- H_1: The PFS time for patients treated with nintedanib plus docetaxel is longer than for patients treated with placebo plus docetaxel(5).

Methods

6.3.2 Describe the RCT(s) design (for example, duration, degree and method of blinding, and randomisation) and interventions. Include details of length of follow-up and timing of assessments. The following tables provide a suggested format for when there is more than one RCT.

A summary of the methodology in the LUME-Lung 1 trial can be seen below in Table 8(3).

Table 8: Summary of methodology of the LUME-Lung 1 trial

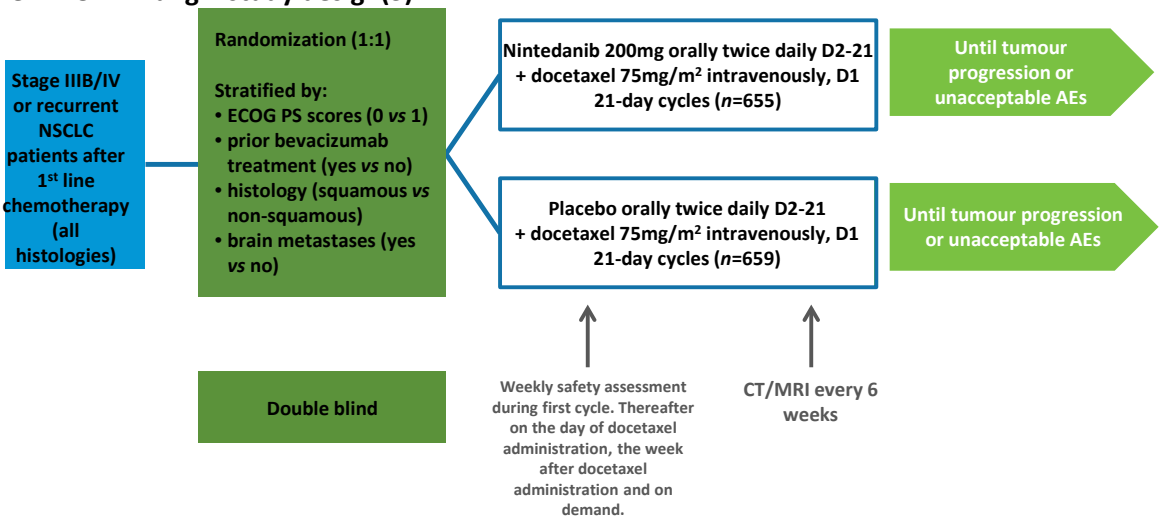
Trial no. (acronym)	LUME-Lung 1
Location(3)	211 locations in 27 countries: Austria, Belarus, Belgium, Bulgaria, China, Croatia, Czech Republic, Denmark, France, Georgian Republic, Germany, Greece, India, Israel, Italy, Lithuania, Poland, Portugal, Romania, Russia, Slovakia, South Korea, South Africa, Spain, Switzerland, Ukraine, United Kingdom
Design(3)	Phase III multi-centre, randomised, parallel-group, double-blind, placebo-controlled RCT comparing the efficacy and safety of nintedanib + docetaxel to placebo + docetaxel in patients with advanced, metastatic (stage IIIB/IV) or recurrent NSCLC after failure of first-line chemotherapy.
Duration of study(3)	• 23 December 2008 – 15 February 2013 (data cut-off date)
Method of randomisation(3)	Patients were randomised in a 1:1 ratio to nintedanib + docetaxel or placebo + docetaxel. Treatment assignment was made using a third-party phone or web-based randomisation involving the use of an interactive voice/web response system (IVRS/IWRS). Randomisation was done in blocks of four per country for administrative reasons. Within each country randomisation was stratified by ECOG performance (0 vs 1), previous bevacizumab treatment (yes vs no), histology (squamous vs non-squamous) and presence of brain metastases (yes vs no). The randomisation lists were provided by a separate group within Boehringer Ingelheim, the Clinical Trial Support Group, using a validated randomisation number generating system. Patients and investigators were blinded to assignment, and no individuals directly involved in the conduct or analysis of the study had access to treatment allocation until final database lock.
Method of blinding(5)	Neither the patient nor the investigator was informed of treatment allocation. All personnel of Boehringer Ingelheim and the appointed CRO who were involved in the conduct of the trial were unaware of the treatment allocation of patients until final database lock. The primary analysis of PFS in this trial was performed when 713 patients had experienced a PFS event as determined by central independent review. A data snapshot was taken and the analysis was performed by Boehringer Ingelheim personnel who were not involved in the further conduct of the ongoing trial. All personnel involved in these analyses and the preparation of the CTR, including the authors, reviewers, and approvals of the CTR, signed an appropriate confidentiality agreement. All personnel who were involved in the further conduct of the trial, as well as the investigators (with the exception of the Coordinating Investigator) and all patients remained blinded regarding patient treatment allocation.
Intervention(s) (n =)† and comparator(s) (n =)†(3)	• Nintedanib + docetaxel (n=655) Nintedanib 200mg twice daily, orally, on days 2 to 21 of a 21-day cycle in combination with docetaxel 75mg/m^2 IV on day 1 of a 21-day cycle. Two dose-reductions were permitted for nintedanib (from 200 to 150mg twice daily and from 150 to 100mg twice daily). . One dose-reduction was permitted for docetaxel (from 75 to 60mg/m^2). Details of the dose reduction scheme can be seen in Table 11. • Matched placebo + docetaxel (n=659) Matched placebo twice daily on days 2 to 21 of a 21-day cycle in combination with docetaxel 75mg/m^2 IV on day 1 of a 21-day cycle. One dose-reduction was permitted for docetaxel (from 75 to 60mg/m^2). Details of the dose reduction scheme can be seen in Table 11. Continuous treatment until disease progression or unacceptable AEs Nintedanib/placebo monotherapy allowed in patients who received ≥4 cycles of combination therapy Docetaxel monotherapy allowed in patients who experienced unacceptable nintedanib-related AEs

Trial no. (acronym)	LUME-Lung 1
Primary outcomes (including scoring methods and timings of assessments)(3)	• PFS by central independent review, using modified RECIST (version 1.0) criteria. Tumour assessments were performed at baseline (within 4 weeks of randomisation), and every six weeks after first docetaxel administration.
Secondary outcomes (including scoring methods and timings of assessments)(3)	• OS (key secondary endpoint) • PFS by investigator review • Tumour response by central independent review and investigator assessment, according to modified RECIST (version 1.0) criteria, including: o Confirmed objective response o Disease control o Time to confirmed objective response o Duration of confirmed objective response o Duration of disease control o Change in tumour size • Clinical improvement • HRQL • Pharmacokinetics • Safety and tolerability, AEs classified according to CTCAE version 3.0 (recorded during study period and follow-up), and changes in safety laboratory parameters. Safety assessments were performed on a weekly basis during the first cycle, on the day of docetaxel administration thereafter, and on demand. Blood samples were taken for laboratory analyses on a weekly basis throughout the first cycle. Thereafter, blood samples were taken on the day of, and the week after docetaxel administration.
Changes in the conduct of the trial or the planned analysis(61)	The analyses in the LUME-Lung 1 trial were extended beyond the original specifications of the statistical analysis plan to validate findings from a hypothesis-generating analysis of the independent LUME-Lung 2 study. This extension to the TSAP was introduced following unblinding of the trial for the primary PFS analysis, but prior to database lock for the final OS analysis. The extension to the TSAP was signed on 23 Jan 2013.
Duration of follow-up	Follow-up until death or lost to follow-up. Median follow-up was 7.1 months (interquartile range: 3.8-11.0) at the time of the primary PFS analysis and 31.7 months (interquartile range: 27.8-36.1 months) at the time of the final OS analysis(3). Follow-up visits were to be performed after the end of treatment with nintedanib/placebo in combination with standard docetaxel therapy, after monotherapy with docetaxel in cases where nintedanib/placebo had been discontinued, or after monotherapy with nintedanib/placebo in cases where docetaxel had been discontinued. Follow-up visits were to be performed every 6 to 8 weeks until the patient died or was lost to follow-up(5).

AEs = adverse events; CTCAE = common terminology criteria for adverse events; HRQL = health related quality of life; OS =overall survival; PFS = progression-free survival; TSAP = trial statistical analysis plan
† Randomised number

A diagrammatical representation of the LUME-Lung 1 study design can be seen in Figure 2(3).

Figure 2: LUME-Lung 1 study design(3)



CT = computed tomography; ECOG PS = Eastern Cooperative Oncology Group performance status; MRI = magnetic resonance imaging; NSCLC = non-small cell lung carcinomaX

Participants

6.3.3 Provide details of the eligibility criteria (inclusion and exclusion) for the trial. The following table provides a suggested format for the eligibility criteria for when there is more than one RCT. Highlight any differences between the trials.

To be eligible for LUME-Lung 1, NSCLC patients (all histologies) had to have failed or relapsed on one prior first-line chemotherapy. In the case of recurrent disease, one additional prior regimen was allowed for adjuvant, neoadjuvant, or neoadjuvant plus adjuvant therapy(3). A full list of inclusion and exclusion criteria is provided in Table 9 .

Table 9: Inclusion and exclusion criteria (full list) for selection of the trial population in LUME-Lung 1(3)

LUME-Lung 1	
Inclusion criteria	Male or female patient aged 18 years or older Histologically or cytologically confirmed, locally advanced and/or metastatic NSCLC of stage IIIB or IV or recurrent NSCLC Relapse or failure of one first-line prior chemotherapy At least one target tumour lesion that has not been irradiated within the past 3 months and that can accurately be measured Life expectancy of at least 3 months ECOG PS of 0 or 1 Patient has given written informed consent
Exclusion criteria	More than one prior chemotherapy regimen for advanced and/or metastatic or recurrent NSCLC More than one chemotherapy treatment regimen (either neoadjuvant or adjuvant or neoadjuvant +

LUME-Lung 1	
	adjuvant) prior to first-line chemotherapy
	Previous therapy with other VEGFR inhibitors (other than bevacizumab) or docetaxel for treatment of NSCLC
	Persistence of clinically relevant therapy related toxicities from previous chemotherapy and/or radiotherapy
	Treatment with other investigational drugs or other anti-cancer therapy, or treatment in another clinical trial within the past 4 weeks before start of therapy or concomitantly with this trial
	Radiotherapy (except extremities and brain) within the past 3 months prior to baseline imaging
	Active brain metastases or leptomeningeal disease
	Radiographical evidence of cavitary or necrotic tumours
	Centrally located tumours with radiographical evidence (CT or MRI) of local invasion of major blood vessels
	History of clinically significant haemoptysis within the past 3 months
	Therapeutic anticoagulation (except low dose heparin) or antiplatelet therapy
	History of major thrombotic or clinically relevant major bleeding event in the past 6 months
	Known inherited predisposition to bleeding or thrombosis
	Significant cardiovascular diseases
	Inadequate safety laboratory parameters
	Significant weight loss (>10 %) within the past 6 weeks
	Current peripheral neuropathy greater than CTCAE grade 2 except due to trauma
	Pre-existing ascites and/or clinically significant pleural effusion
	Major injuries and/or surgery within the past 10 days prior to randomisation with incomplete wound healing
	Serious infections requiring systemic antibiotic therapy
	Decompensated diabetes mellitus or other contraindication to high-dose corticosteroid therapy
	Gastrointestinal disorders or abnormalities that would interfere with absorption of the study drug
	Active or chronic hepatitis C and/or B infection
	Serious illness or concomitant non-oncological disease or laboratory abnormality that may increase the risk associated with study participation or study drug administration
	Patients who are sexually active and unwilling to use a medically acceptable method of contraception during the trial and for at least 12 months after end of active therapy
	Pregnancy or breast feeding
	Psychological, familial, sociological, or geographical factors potentially hampering compliance with the study protocol and follow-up schedule
	Patients unable to comply with the protocol
	Active alcohol or drug abuse
	Other malignancy within the past 3 years other than basal cell skin cancer, or carcinoma *in situ* of the cervix
	Any contraindications for therapy with docetaxel
	History of severe hypersensitivity reactions to docetaxel or other drugs formulated with polysorbate 80 (Tween 80)
	Hypersensitivity to nintedanib and/or the excipients of the trial drugs
	Hypersensitivity to contrast media

CT = computerised (or computed) tomography, CTCAE = Common Toxicity Criteria for Adverse Events. ECOG PS = Eastern Cooperative Oncology Group Performance Status; MRI = magnetic resonance imaging, NSCLC = non-small-cell lung cancer, VEGFR = vascular endothelial growth factor receptor

6.3.4 *Describe the patient characteristics at baseline. Highlight any differences between study groups. The following table provides a suggested format for the presentation of baseline patient characteristics for when there is more than one RCT.*

In the LUME-Lung 1 intention-to-treat (ITT) population, patient demographics were balanced across treatment arms. Of the 1,314 randomised patients, 759 patients (57.8%) had non-squamous cell carcinoma and 555 patients (42.2%) had squamous cell carcinoma. Of the patients with non-squamous cell carcinoma, 658 patients had adenocarcinoma(3).

The number of patients with adenocarcinoma, as well as their baseline characteristics such as sex, age, race, smoking status and ECOG performance status was balanced across the two treatment arms. Approximately 50% of all patients enrolled in each arm had confirmed adenocarcinoma histology (49.2% in the nintedanib plus docetaxel and 51.0% in the placebo plus docetaxel arm)(3) Overall, 62.5% of adenocarcinoma patients were male. Their mean age was 58.5 years. The majority of patients were Caucasian (76.9%), 21.7% were Asian, 70.4% of patients had an ECOG performance status of 1, and 7.4% of patients had brain metastases at baseline. 95.9% of adenocarcinoma patients had received platinum-based first-line therapy and 6.8% received bevacizumab treatment prior to entry into LUME-Lung 1) (Table 10)(3, 5).

Information on EGFR-TK and ALK mutation status was not systematically collected for patients enrolled in LUME-Lung 1. At the time of study initiation, the influence of EGFR-TK status was still being researched, and testing for these mutations was not routine clinical practice(34). Only last year, in 2013 NICE issued guidance recommending testing for EGFR-TK mutations with one of five different methods(95).

Table 10: Patient demographics in LUME-Lung 1 (adenocarcinoma population)(5)

Parameter		**Patients with adenocarcinoma**	
		Nintedanib + docetaxel (*n*=322)	**Placebo + docetaxel (*n*=336)**
Sex, *n* (%)	Male Female	203 (63.0%) 119 (37.0%)	208 (61.9%) 128 (38.1%)
Age, years	Mean (StD)	58.5 (10.1)	58.6 (9.5)
Race, *n* (%)	Asian White Other	65 (20.2%) 253 (78.6%) 4 (1.2%)	78 (23.2%) 253 (75.3%) 5 (1.5%)
Smoking status, *n* (%)	Never smoked Ex-smoker Current smoker	115 (35.7%) 151 (46.9%) 56 (17.4%)	115 (34.2%) 162 (48.2%) 59 (17.6%)
ECOG performance status, *n* (%)	0 1[†]	96 (29.8%) 226 (70.2%)	99 (29.5%) 237 (70.5%)
Prior first-line therapy	Platinum-based therapy	308 (95.7%)	323 (96.1%)
	Non-platinum-based therapy	10 (3.1%)	10 (3.0%)
Prior bevacizumab, *n* (%)	Yes No	24 (7.5%) 298 (92.5%)	21 (6.3%) 315 (93.8%)
Brain metastases at study entry, *n* (%)	Present Absent	26 (8.1%) 296 (91.9%)	23 (6.8%) 313 (93.2%)
Post study therapy	Any systemic therapy Any chemotherapy Pemetrexed Docetaxel	179 (55.6%) 123 (38.2%) 52 (16.1%) 15 (4.7%)	188 (56.0%) 136 (40.5%) 62 (18.5%) 13 (3.9%)

Parameter		Patients with adenocarcinoma	
		Nintedanib + docetaxel (*n*=322)	**Placebo + docetaxel (*n*=336)**
	Other chemotherapy	90 (28.0%)	101 (30.1%)
	EGFR TK inhibitor	98 (30.4%)	105 (31.3%)
	Anti-angiogenesis agent	6 (1.9%)	2 (0.6%)
	Investigational agent	18 (5.6%)	5 (1.5%)

ECOG = Eastern Cooperative Oncology Group; EGFR-TK TK = epidermal growth factor receptor tyrosine kinase; StD = standard deviation

† Including one patient in the nintedanib arm who had an ECOG PS of 2 at screening and at randomisation (i.e. at baseline)

Interventions

Patients were randomised (1:1) to nintedanib 200mg twice daily, orally, or matching placebo, on days 2 to 21 of a 21-day cycle in combination with IV docetaxel 75mg/m^2 given on day 1 of a 21-day cycle (Figure 2)(3). Patients continued with nintedanib/placebo treatment until tumour progression or unacceptable AEs(3).

Dose reduction scheme

In case of AEs related to the study drug, up to two nintedanib dose reductions were permitted, first to 150mg twice daily and then to 100mg twice daily (Table 11)(3). The initial docetaxel dose could also be reduced from 75mg/m^2 to 60mg/m^2, according to label recommendations(92). Details of the dose reduction scheme for patients receiving nintedanib or placebo plus docetaxel combination therapy can be seen below in Table 11(4).

Table 11: Dose-reduction schemes for nintedanib/placebo and docetaxel combination therapy(4)

Adverse event	Nintedanib/placebo	Docetaxel
Haematological and drug related non-haematological AEs (excluding liver enzyme increases, diarrhoea, nausea and vomiting)		
Neutropenia CTCA grade 4 for >7 days	No dose reduction	Dose reduction
Febrile neutropenia	No dose reduction	Dose reduction
Cumulative cutaneous reactions	No dose reduction	Dose reduction
Peripheral neurotoxicity CTCAE grade 2	No dose reduction	Dose reduction
Non-haematological AEs CTCAE grade ≥3 (except diarrhoea, nausea, vomiting, isolated increase of GGT, ALT, AST)	Dose reduction	Dose reduction
Liver enzyme increases		
AST or ALT elevations of CTCAE grade 2 in conjunction with bilirubin elevations of CTCAE grade ≥1, or AST or ALT elevations of CTCAE grade ≥3		
1st episode	Dose reduction	No dose reduction
2nd episode	Dose reduction	Dose reduction
3rd episode	Stop treatment	Stop treatment
Diarrhoea, nausea or vomiting despite adequate supportive treatment		
Vomiting of CTCAE grade ≥2 or nausea of CTCAE grade ≥3 within 3 days after docetaxel therapy		
1st episode	No dose reduction	No dose reduction
2nd episode	Dose reduction	Dose reduction
3rd episode	Dose reduction	No dose reduction
4th episode	Stop treatment	Stop treatment
Vomiting of CTCAE grade ≥2 or nausea of CTCAE grade ≥3 starting >3 days after docetaxel therapy		
1st episode	Dose reduction	No dose reduction

Adverse event	Nintedanib/placebo	Docetaxel
2nd episode	Dose reduction	No dose reduction
3rd episode	Stop treatment	No dose reduction
Diarrhoea of CTCAE grade 2 for >7 consecutive days		
1st episode	Dose reduction	No dose reduction
2nd episode	Dose reduction	Dose reduction
3rd episode	Stop treatment	Stop treatment
Diarrhoea of CTCAE grade ≥3		
1st episode	Dose reduction	No dose reduction
2nd episode	Dose reduction	Dose reduction
3rd episode	Stop treatment	Stop treatment

AEs = adverse events; ALT = alanine transaminase; AST = aspartate transaminase; CTCAE = Common Terminology Criteria for Adverse Events; GGT = gamma-glutamyl transpeptidase

Docetaxel had to be discontinued in case of CTCAE grade ≥3 peripheral neuropathy, severe hypersensitivity or an AE requiring a second dose reduction. Patients who discontinued docetaxel for reasons other than progression could continue with nintedanib/placebo monotherapy provided they had received ≥4 cycles of combination treatment. The maximum number of docetaxel cycles that patients could receive was not restricted. Nintedanib had to be discontinued in case of additional AE episodes requiring a third dose reduction. Patients who discontinued nintedanib/placebo due to intolerable AEs could continue standard-dose docetaxel monotherapy(3).

Details of the dose reduction scheme for patients receiving monotherapy with nintedanib or placebo can be seen in Table 12(4).

Table 12: Dose reduction scheme for monotherapy with nintedanib or placebo

Adverse event	Nintedanib/placebo subsequent treatment
Non-haematological or haematological AEs of CTCAE grade ≥3 (except diarrhoea, nausea, vomiting, isolated increase of GGT, ALT, AST)	Dose reduction
AST or ALT elevations of CTCAE grade 2 in conjunction with bilirubin increases of CTCAE grade ≥1, or AST or ALT elevations of CTCAE grade ≥3	Dose reduction
Vomiting of CTCAE grade ≥2 or nausea of CTCAE grade ≥3 despite supportive care	Dose reduction
Diarrhoea of CTCAE grade 2 for >7 consecutive days despite supportive care	Dose reduction
Diarrhoea of CTCAE grade ≥3 despite supportive care1	Dose reduction

AEs = adverse events; ALT = alanine transaminase; AST = aspartate transaminase; CTCAE = Common Terminology Criteria for Adverse Events; GGT = gamma-glutamyl transpeptidase

Outcomes

6.3.5 *Provide details of the outcomes investigated and the measures used to assess those outcomes. Indicate which outcomes were specified in the trial protocol as primary or secondary, and whether they are relevant with reference to the decision problem. This should include therapeutic outcomes, as well as patient-related outcomes such as assessment of health-related quality of life (HRQL), and any arrangements to measure compliance. Data provided should be from pre-specified outcomes rather than post-hoc analyses. When appropriate, also provide evidence of reliability or validity, and current status of the measure (such as use within UK clinical practice). The following table provides a suggested format for presenting primary and secondary outcomes when there is more than one RCT.*

The primary endpoint in the LUME-Lung 1 trial was PFS based on central independent assessment. The key secondary endpoint was OS. Other secondary endpoints included investigator-assessed PFS, tumour response, clinical improvement, patient reported QoL, pharmacokinetics and safety and tolerability. A detailed description of all study endpoints is presented in Table 13(5).

EMA guidance on the evaluation of anticancer medicinal products in man (EMA/CHMP/205/95/Rev.4) states that these are acceptable criteria for demonstration of clinical benefit(96). Moreover, the guideline states that convincingly demonstrated favourable effects on survival are, from both a clinical and methodological perspective, the most persuasive outcome of a clinical trial(96).

Table 13: Summary of key endpoints and assessments in LUME-Lung 1

Endpoint/ assessment	Details	Reliability/validity/ current use in clinical practice
Primary endpoint		
PFS(5)	• PFS by central review, using modified RECIST (version 1.0) criteria. Tumour assessments performed at baseline (within 4 weeks of randomisation), and every 6 weeks after first docetaxel administration • PFS was defined as time from date of randomisation to date of disease progression, or to date of death, whichever occurred earlier • Disease progression was defined as: o new lesions, including new lesions in a previously irradiated field o an unequivocal increase in a tumour within a previously irradiated field o an increase in sum of longest diameter (SLD) of the target lesions of 20% from nadir (lowest value measured since treatment started) • Patients who experienced a 30% reduction from baseline in SLD of target lesions and a single instance of a 20% increase in SLD from nadir were considered as having progressed • The primary PFS analysis considered all data collected until the cut-off date for the efficacy analysis, which was the date of the 713th PFS event	• Well-established primary endpoint in oncology clinical trials to assess the efficacy of a drug (97) • Superior to time to tumour progression as it does not censor patients who die from any cause(97) • Unaffected by post-progression treatment administration(97) • Recommended by the EMA as an endpoint in oncology clinical trials(96)
Secondary endpoints		
OS(5)	• OS was the key secondary endpoint • OS was defined as the time from date of randomisation to date of death (irrespective of cause of death). Patients who stopped active trial treatment were followed until death or lost to follow-up	• Meaningful clinical outcome for determining the efficacy of interventions which extend survival(97) • Well validated endpoint due to its objectivity and benefit to patients(97) • Recommended by the EMA as the most persuasive endpoint in oncology clinical trials(96) • However, can be confounded by crossover following disease progression and by causes of mortality unrelated to cancer(97)

Endpoint/ assessment	Details	Reliability/validity/ current use in clinical practice
PFS by investigator review(5)	• PFS by investigator review	• Well-established primary endpoint in oncology clinical trials to assess the efficacy of a drug (97) • Considered a relevant measure of patients benefit by the EMA(96)
Tumour response evaluation(5)	• Tumour response by central independent review and investigator assessment, according to modified RECIST (version 1.0) criteria was assessed at baseline (within 4 weeks of randomisation) and every 6 weeks after first docetaxel administration, and categorised into one of the following categories: o complete response (CR) – disappearance of all target lesions and non-target lesions o partial response (PR) – at least a 30% decrease in the SLD of target lesions, taking as reference the baseline SLD o stable disease (SD) – neither sufficient shrinkage of target lesions to qualify for PR nor sufficient increase to qualify as PD; persistence of one or more non-target lesions o progressive disease (PD): • new lesions, including new lesions in a previously irradiated field • an unequivocal increase in a tumour within a previously irradiated field • an increase in SLD of the target lesions of 20% from nadir (lowest value measured since treatment started) o unknown (UNK) Evaluation of tumour response was based on radiological tumour assessments (CT or MRI) • Tumour images were centrally reviewed by a panel of independent radiologists. Following radiological review, all patient information was presented to an oncologist. The radiologists and the oncologist were blinded to treatment • Best overall response:(60) o represents the best response a patient has had during their time in the study up until progression, last evaluable assessment in the absence of progression or the start of subsequent anti-cancer therapy. o for patients whose progression event is death, best objective response will be calculated based on data up until the last evaluable RECIST assessment prior to death. • Confirmed objective response o A patient was considered to have a confirmed objective response if a CR or PR was confirmed by imaging no earlier than 28 days after the first occurrence of the response • Disease control o Disease control was defined as a best overall response of CR, PR, or SD recorded at least 6 weeks after the date of randomisation	• Recognised measurement of therapeutic efficacy in oncology trials which may be indicative of clinical benefit(97) • Recommended by the EMA as a secondary endpoint or exploratory analysis(96)

Endpoint/ assessment	Details	Reliability/validity/ current use in clinical practice

	• Time to confirmed objective response
	  o Time from randomisation to first documented confirmed response (CR or PR) recorded at least 6 weeks after the date of randomisation
	• Duration of confirmed objective response
	  o Time from first documented confirmed response [CR or PR] to progression, or death in the absence of progression
	• Duration of disease control
	  o Time from randomisation to progression, or death in the absence of progression (whichever occurs earlier) amongst patients with disease control
	• Change in tumour size
	  o The best change in size (i.e. SLD) of target lesions from baseline was analysed. The maximum SLD decrease from baseline (or the minimum increase in SLD for patients with no reduction in target lesion size) was considered as the best change of the target lesion size in a patient	

Clinical improvement(5)	• Clinical improvement quantified the maintenance of body weight and ECOG PS, by measuring the time from randomisation to deterioration in body weight of more than 10% from baseline, and/or increase in ECOG performance score of at least 1 category from baseline, whichever occurred earlier. Patients who died without prior deterioration were considered as having deteriorated at the time of death. • Clinical improvement was analysed until end-of treatment only	• Clinically relevant outcomes for both clinicians and patients

QoL(5, 59)	• HRQL was measured at the screening visit, at 21-day intervals during treatment, at the end of active treatment, and at the first follow-up visit by the following standardised self-assessment questionnaires:
	  o EQ-5D health status self-assessment questionnaire
	  o EORTC Quality of Life Questionnaire (EORTC QLQ-C30)
	  o EORTC lung cancer specific supplementary module (EORTC QLQ-LC13)
	• The EQ-5D includes the following two questionnaires, which were analysed descriptively:
	  o Five dimensions of health (mobility, self-care, usual activities, pain/discomfort and anxiety/depression), which are analysed descriptively. Each dimension comprised three levels (no problems, some problems, severe problems)
	  o A visual analogue scale (VAS) recorded the respondents self-rated health status on a vertical graduated (0-100) scale
	• The EORTC QLQ-C30 questionnaire includes a global health status/HRQL scale, 5 functional scales, 3 symptom scales, and 6 single items to assess dyspnoea, insomnia, appetite loss, constipation, diarrhoea and financial difficulties. The QLQ-LC13 supplementary module was designed to be used by patients receiving chemotherapy or radiotherapy. It incorporates a multi-item scale to assess dyspnoea, and a series of single items to assess pain, coughing sore mouth, dysphagia, peripheral neuropathy, alopecia and haemoptysis.
	• The main HRQL endpoints were the time to deterioration for cough (QLQ-LC13, question 1), dyspnoea (QLQ-LC13, questions 3
		• Clinically relevant outcome for both clinicians and patients
		• Recommended by the EMA as a secondary endpoint or exploratory analysis(96)
		• The QLQ-C30, QLQ-LC13 and EQ-5D have been shown to be reliable and valid measures of the QoL of cancer patients and are commonly used in clinical trials(98, 99)

Endpoint/ assessment	Details	Reliability/validity/ current use in clinical practice
	to 5) and pain (QLQ-C30, Questions 9 and 19) and were evaluated as follows: o Distribution of patients with improved, stable, or worsened scores. Improvement was defined as scores that improve by ≥10 points (0-100 point scale) at any time during study. Worsening was defined as a worsening in EORTC scores of ≥ 10 points at any time in patients with no improvement. Otherwise, a patient was considered stable. o Time to deterioration: defined as time from randomisation to the first 10-point increase (i.e. worsening) from baseline score	
Pharmacokinetics(5)	• Pharmacokinetics of nintedanib and of its clinical relevant metabolites BIBF1202 and BIBF1202 glucuronide were determined from blood samples taken at Visit 2 of Treatment Course 2 and 3; both prior to and after the administration of nintedanib.	
Safety(5)	• Incidence and intensity of AEs according to the CTCAE version 3.0 • Changes in safety laboratory parameters • The safety analysis included data collected until the safety cut-off date	• CTCAE version 3.0 criteria are the current, standard assessment of safety

CR = complete response; CT = computed tomography; CTCAE = Common Terminology Criteria for Adverse Events; ECOG = Eastern Cooperative Oncology Group; EORTC QLQ LC = European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire (Lung Cancer Module); EMA = European Medicines Agency; EQ-5D = European Quality of Life-5 Dimensions; HRQL = health related quality of life; MRI = Magnetic resonance imaging; OS =overall survival; PD = progressive disease; PFS = progression-free survival; PR = partial response; PRO = patient reported outcome; RECIST = Response Evaluation Criteria in Solid Tumours; SD = stable disease; SLD = sum of longest diameters; VAS = visual analogue scale

Efficacy assessments

A baseline scan was to be performed within 4 weeks prior to initiation of study treatment to fully assess the extent of the tumour disease. Baseline scans were to include computed tomography (CT) or magnetic resonance imaging (MRI) of the brain, and CT or MRI of the chest with inclusion of the whole liver and both adrenal glands. Bone scans at baseline were to be performed in patients with known bone metastases and in case of clinical suspicion of previously unknown bone metastases. One to 10 target lesions (up to a maximum of 5 per organ) were to be identified at baseline. Post-baseline tumour imaging was to be repeated every 6 weeks after the first administration of docetaxel. The same method of assessment and the same technique had to be used to characterise each reported lesion at baseline and during the trial (except for brain metastases, for which CT or MRI could be used). This imaging schedule followed the 'Minimum clinical recommendations for diagnosis, treatment and follow-up of NSCLC' of the European Society for Medical Oncology [ESMO](100).

Safety assessments

Patient safety in LUME-Lung 1 was primarily assessed based on the occurrence of AEs and changes in safety laboratory parameters, but ECGs and vital signs were also recorded.

After inclusion into the trial, the patient's condition was assessed (e.g. documentation of medical history and concomitant diagnoses and diseases) and this was used as baseline for subsequent comparisons. All AEs occurring during the course of the trial (i.e. from signing the informed consent until 28 days after the end of study treatment, and beyond for serious AEs that were considered related to study drug, the trial, or trial procedures) were to be collected, documented, and reported to the sponsor by the investigator. Patients were to spontaneously report any AEs, their date of onset and end. In case of AEs, patients were monitored more frequently until recovery(4).

Clinical laboratory examinations to evaluate safety included haematology, biochemistry, and coagulation parameters. Blood samples for the assessment of these were to be collected at all scheduled visits. Blood samples were taken for laboratory analyses on a weekly basis throughout the first cycle, and on the day of and the week after docetaxel administration thereafter(4).

Other safety assessments included a general physical examination, which was to be performed at screening, Visit 1 of each treatment course and at the end-of-treatment visit, and assessments of height, weight, ECOG performance score, a 12-lead ECG, and a measurement of blood pressure, pulse rate (after 2 minutes rest), and body temperature, all to be performed at pre-specified timepoints during the trial(4).

Statistical analysis and definition of study groups

6.3.6 State the primary hypothesis or hypotheses under consideration and the statistical analysis used for testing hypotheses. Also provide details of the power of the study and a description of sample size calculation, including rationale and assumptions. Provide details of how the analysis took account of patients who withdrew (for example, a description of the intention-to-treat analysis undertaken, including censoring methods; whether a per-protocol analysis was undertaken). The following table

provides a suggested format for presenting the statistical analyses in the trials when there is more than one RCT.

The statistical analysis plan for LUME-Lung 1 is summarised in Table 14 and is described in detail in the following sections.

Table 14: Summary of statistical analyses in LUME-Lung 1

Trial no. (acronym)	**LUME-Lung 1**
Hypothesis objective	• H_0: The PFS time for patients treated with nintedanib + docetaxel is equal to the PFS time for patients treated with placebo + docetaxel(3) • H_1: The PFS time for patients treated with nintedanib + docetaxel is longer than for patients treated with placebo + docetaxel(3)
Statistical analysis	Initial statistical analysis plan: The primary PFS analysis was to be performed when 713 patients had experienced a centrally assessed PFS event (cut-off date 2 Nov 2010) to detect a HR of 0·78 with 90% power(3). The stratified log-rank test was used to test for the effect of nintedanib at the 2-sided alpha-level of 0.05(3). The log-rank test included the four stratification factors used at randomisation: baseline ECOG PS (0 vs 1), prior treatment with bevacizumab, tumour histology (squamous vs non-squamous) and presence of brain metastases at baseline(5). At this time, a protocol-defined interim analysis of OS was also to be performed. The primary analysis was based on the ITT population(3). • The final analysis of OS was performed when 1,151 patients had died. This would provide 80% power to detect a HR of 0.85 with the use of a stratified log-rank test and a two-look Lan-DeMets group sequential design with an O'Brien-Fleming-type boundary at a two-sided cumulative 5% level of significance(3). • At the time of the final OS analysis a follow-up analysis of all available PFS events was also performed (cut-off date 15 Feb 2013)(3). Futility analysis: • A pre-planned futility analysis was to be performed by the independent DMC after approximately 50% of the PFS events needed for the primary PFS analysis had occurred (~356 events), for the purpose of advising the sponsor as to whether or not the study should continue as planned(3). The sponsor was blinded to the results of this analysis(3). Although PFS by central independent review was the primary endpoint, PFS as assessed by the investigator was used for the futility analysis because of the logistical complexity and the time it took to complete the central independent review of patients' imaging data(5). Amended statistical analysis plan: The planned analyses were extended beyond the original specifications in order to validate findings from a hypothesis-generating analysis of the independent LUME-Lung 2 study. OS (the key secondary endpoint) was to be tested in a hierarchical fashion, where the first step was to test OS in adenocarcinoma patients who had progressed within 9 months of starting first-line therapy (i.e. the T<9m adenocarcinoma population), followed by OS testing in the adenocarcinoma population, and finally in the overall trial population, regardless of histology. Each of these three hypotheses could only be tested if the previous null hypothesis in the testing sequence had been rejected. Further details on the hierarchical testing for OS are provided in Section 6.3.7(61)
Sample size, power calculation	• The sample size in LUME-Lung 1 was calculated based on the assumption that nintedanib + docetaxel will increase the median PFS by approximately 28% to 32%, compared with placebo + docetaxel. A median PFS of 4 months (for patients with a PS of 0 or 1) was assumed for patients treated with docetaxel alone. To detect an underlying treatment difference of 27.5% or 1.1 months, (i.e. PFS of 4.0 *vs* 5.1 months in placebo + docetaxel and nintedanib + docetaxel treatment arms, respectively), the study would require 713 centrally assessed PFS events to

Trial no. (acronym)	LUME-Lung 1
	have 90% statistical power to show a significant improvement in median PFS with a HR of 0.7843(4). 1,151 deaths would provide a statistical power of 80% to detect an 18% increase in median OS with nintedanib + docetaxel compared to placebo + docetaxel, with an HR of 0.8475. It should be noted that the magnitude and pattern of the effect of any third-line or later treatment after progression could have obscured the treatment effect on OS(4). • Sample size calculations were performed using the EAST-5 software using the log-rank test, and excluding the interim analysis(4).
Data management, patient withdrawals	• The trial was conducted in compliance with the clinical trial protocol and its amendment, the principles laid down in the Declaration of Helsinki, local laws, in accordance with the ICH Harmonised Tripartite Guideline for GCP, and the sponsor's and CRO's SOPs(5). Commitments to conduct the study in accordance with the protocol and with GCP were obtained from investigators by their signing of the clinical trial protocol(5). • Patients who interrupted nintedanib/placebo therapy for ≥14 consecutive days, except for temporary discontinuation of treatment due to AEs, were considered non-compliant. Patients were asked to return all unused nintedanib/placebo capsules (including reserve medication) at the next scheduled visit. The investigator or the investigator's deputy were instructed to assess whether the patient had taken the medication according to the clinical trial protocol(5). • In accordance with the Declaration of Helsinki, a patient was to be withdrawn from the trial if he or she withdrew consent. Patients were free to discontinue their participation in the trial at any time and for any reason. In such cases, the patient was asked to participate in an end-of-treatment investigation and all data collected until the time of withdrawal were included in the final analyses. No further follow-up was performed if the patient did not agree(5). A patient was to be withdrawn from active treatment (combination therapy, as well as monotherapy with nintedanib/placebo in patients who had discontinued docetaxel therapy, or monotherapy with docetaxel in patients who had discontinued nintedanib treatment) if any of the following applied:(4) o The patient requested discontinuation of active treatment but agreed to be followed up. o The patient was no longer able to receive active treatment (e.g. due to AEs, pregnancy, surgery, concomitant diagnoses, concomitant therapies or for administrative reasons). The investigator was permitted to stop a patient's treatment if the patient was no longer able to attend trial visits, e.g. due to worsening of disease. o Treatment could be stopped for an individual patient upon agreement of the sponsor and the investigator if eligibility criteria were violated or the patient failed to comply with the protocol. o Further dose reductions were considered necessary but were not permitted according to the clinical trial protocol. o Patients who were eligible for monotherapy with nintedanib/placebo but who had received fewer than 4 courses of combination treatment (nintedanib/placebo + docetaxel) o Patients with radiologically documented progressive disease o Patients who were administered restricted concomitant medications

AEs = adverse events; CRO = contract research organisation; DMC = data monitoring committee; ECOG = Eastern Cooperative Oncology Group; GCP = good clinical practice; HR = hazard ratio; ICH = International Conference on Harmonisation; ITT = intention to treat; OS = overall survival; PFS = progression-free survival; PS = performance status; SOP = standard operating procedure

Censoring rules for the primary PFS analysis

Patients without a PFS event prior to the efficacy cut-off date were censored at the date of the last evaluable tumour imaging. Further censoring rules for PFS are summarised in Table 15. The same

censoring rules were applied to other analyses of PFS as well as to other efficacy endpoints (if not stated otherwise)(4).

Table 15: Censoring rules for the determination of PFS(4)

Situation	Outcome (event or censored)	Date of PFS or censoring
No baseline tumour assessment	censored	Date of randomisation
Progressed according to central imaging (no missed radiological assessments)	event	Date of PD determined by central review
Non-PD from central review[1], death before next scheduled assessment	event	Date of death
One missed assessment, death or progression after date of missed assessment, but before or at the second scheduled assessment	event	Date of PD or death
Non-PD from central review[1], more than 1 consecutive missed assessment, death or progression after date of second missed assessment	censored	Date of last imaging before missed assessment
New anticancer therapies before progression or death[2]	censored	Date of last imaging before new anticancer medication
Death before the scheduled date of first imaging	event	Date of death
No imaging performed post-baseline, patient died between first and second scheduled assessments	event	Date of death
No imaging performed post-baseline, patient died after second scheduled assessment	censored	Date of randomisation
No imaging performed post-baseline, vital status is unknown or patient is known to be alive	censored	Date of randomisation
Alive and not progressed according to central review (no missed radiological assessments)	censored	Date of last imaging

CRF = case report form; PD = progressive disease; PFS = progression-free survival

[1]This applies to the last assessment at which non-PD (SD or better) was assessed

[2]Subsequent anticancer therapies collected on the 'Additional anticancer therapy' of the CRF underwent medical review by the sponsor to identify anticancer therapies. Only anticancer drugs (including investigational drugs) and surgery with the verbatim terms 'lobectomy' or 'pneumonectomy' were included in the censoring algorithm. The following therapies were not considered as subsequent anticancer therapies for the purpose of censoring: monotherapy with non-study docetaxel, supportive care (e.g. biphosphonates), radiotherapy, and other therapies that were no anticancer therapies or lacked clear evidence of anticancer activities (e.g. herbal therapies)

Patients without an event prior to the cut-off were censored at the date they were last known to be alive. Further censoring rules had been specified and are summarised in Table 16. Patients might have refused to be followed for progression/survival. However, if the date of death for such a patient became available from a cancer registry or another public source, this date was used for the derivation of OS(5).

Table 16: Censoring rules for the determination of OS(5)

Situation	Outcome (event or censored)	Date of death or censoring
Patient did not receive combination therapy	censored	Date of randomisation
Patient died and date of death is known	event	Date of death
Patient died and date of death is unknown	censored	Last date when the patient was known to be alive
Patient alive	censored	Date of last contact
Unknown[1]	censored	Last date when the patient was known to be alive

[1]Including patients who were lost to follow-up, with no vital status information available

Sensitivity analyses

PFS sensitivity analyses

Four pre-planned sensitivity analyses were undertaken to assess the robustness of statistical model assumptions and study conduct (i.e. image collection) of the primary PFS analysis. Sensitivity analysis performed for the primary PFS analysis included a Cox proportional hazards model fitting the four stratification factors as covariates, a stepwise variable selection method to identify covariates that might be relevant to efficacy, an analysis replacing actual tumour imaging dates with the originally scheduled dates of radiological assessments, and a sensitivity analysis using an interval-censoring approach(4).

Sensitivity analysis 1

A Cox proportional hazards model, fitting the 4 stratification factors used at randomisation as covariates, was used to test the effect of nintedanib *vs* placebo at the 2-sided level of 0.05. This model assumed that the hazards were proportional on an overall basis, i.e. that the underlying shape of the survival curve was the same for each stratum and treatment combination and that survival curves were proportional. Graphical methods were used to investigate the assumption of proportionality of hazards in the above model in an exploratory way(4).

Sensitivity analysis 2

Exploratory analyses were performed to identify covariates that might be relevant to efficacy. A stepwise variable selection method was used to obtain the best fitting model to test the effect of nintedanib vs placebo at the nominal 2-sided level of 0.05. The critical value for inclusion and exclusion from the model was significance at the 10% level .Treatment was included in all stages of the model selection process. The following covariates were included in the modelling process:(4)

- brain metastases at baseline

- prior treatment with bevacizumab
- sex
- body- surface area
- region (Asia *vs* non-Asia)
- age
- race
- best response to first-line chemotherapy
- stage at diagnosis
- duration of first-line chemotherapy
- time to first progression
- liver metastases at baseline
- smoking status
- time since first histological diagnosis
- time since last chemotherapy
- presence of ipsilateral metastases in the lung at baseline
- presence of contralateral metastases in the lung at baseline
- bone metastases at baseline
- adrenal metastases at baseline
- sum of target lesions at baseline.

Factors were excluded from the final model if they did not improve the model fit according to a pre-defined algorithm(4).

*Sensitivity analysis 3X*If there was a systematic deviation in the timing of tumour imagings, the treatment effect of nintedanib vs placebo for the primary PFS endpoint might have been biased. This means that an observed treatment effect could have been due to tumour imaging being conducted earlier in one of the treatment groups, as opposed to a true underlying treatment effect. A sensitivity analysis was performed, replacing actual imaging dates with the originally scheduled dates of radiological assessments. This analysis was conducted for all time points of assessment, and was performed using the same statistical method as for the primary PFS analysis however, instead of Breslow's method, the exact method for the handling of tied observations was used according to Hertz-Picciotto(4).

Sensitivity analysis 4

According to the trial design (i.e. with examination of patients for disease progression at regular, protocol-defined intervals) it was unlikely that the exact date of a patient's disease progression was observed. A sensitivity analysis was carried out for the primary endpoint PFS, using an interval censoring approach. For this analysis it was assumed that the exact date of progression was between the last tumour imaging not showing disease progression and the first imaging documenting progression. For death without documented disease progression, progression was assumed to have occurred between the last imaging date before death without disease progression and the date of death(4).

OS sensitivity analyses

Two sensitivity analyses were performed for OS to assess the robustness of statistical model assumptions. One model included an analysis of OS using a Cox proportional hazards model with the 4 stratification factors used at randomisation as covariates (in analogy to the analysis performed for PFS), and the second model included both the stratification factors and the baseline sum of the longest diameters (SLD) of the target lesions (mm) as covariates(5).

6.3.7 Provide details of any subgroup analyses that were undertaken and specify the rationale and whether they were pre-planned or post-hoc.

Subgroup analyses

The consistency of the treatment effect of nintedanib *vs* placebo on the primary endpoint was investigated for a number of demographic and baseline characteristics. Subgroup analyses were performed for the efficacy endpoints PFS based on central review and OS. Subgroups were predefined in the TSAP except for 'geographical region', 'best response to first-line therapy', 'sum of longest diameters at baseline' and 'time since first-line therapy' which were added post hoc. Demographic and baseline characteristics which were analysed included:(5)

- baseline ECOG PS (0 vs 1)
- presence of brain metastases at baseline (yes vs no)
- prior treatment with bevacizumab (yes vs no)
- sex (male, female)
- age (<65years, ≥65 years)

- race (Asian vs non-Asian patients; information was derived from the race categories as documented on the CRF)
- smoking status (never smoked vs currently smokes/ex smoker)
- geographical region (Asia, Europe, South Africa; based on country of enrolment)
- best response to first-line therapy (CR/PR/SD, PD, unknown/missing/NA)
- sum of longest diameters at baseline (<7.5cm vs ≥7.5cm)
- time since first-line therapy (<9 months vs ≥9 months)

In addition to the subgroups listed above, the efficacy of nintedanib vs placebo for PFS and OS was investigated for the following baseline characteristics:(5)

- presence of liver metastases (yes vs no)
- disease stage at diagnosis (<IIIB/IV, IIIB, IV)
- concomitant therapy with biphosphonates at baseline (yes vs no)
- presence of adrenal metastases (yes vs no)
- number of metastatic organs at baseline (≤2 metastatic organs, >2 metastatic organs, not centrally reviewed)
- lactate dehydrogenase (LDH) level at baseline (LDH ≤1, LDH >1)

HRs were produced in order to investigate the consistency of the treatment effect for each level of pre-defined baseline characteristics. HRs were obtained from models fitted for each level of the baseline covariate e.g. for the baseline characteristics of sex, one model was produced for males and females. All models were stratified by the stratification factors used in randomisation, and were fitted using identical methodology. However, in cases where the stratification factor is the baseline covariate that was investigated, this was not included in the strata statement of the models. Patients from strata combinations with no events did not contribute to the stratified test(3).

In order to provide a statistical framework for interpretation of the consistency of the treatment effect, interaction p-values were created. The interaction p-value formally tested the hypothesis of whether the HR (treatment effect) was different in the two levels of the baseline characteristic. Interaction p-values were created using a modelling procedure that assumed proportionality on a global basis (within and between strata). Models were fitted to include the factors used to stratify the randomisation as covariates. Models were fitted with and without treatment by covariate interactions and the models compared using the log likelihood ratio statistic(3).

An exploratory analysis of the independent LUME-Lung 2 trial and of the interim LUME-Lung 1 trial data identified a subgroup of patients with tumours of adenocarcinoma histology and <9 months since start of first-line therapy that appeared to derive a clinical benefit from nintedanib in combination with chemotherapy(62). In order to validate this, a hierarchical testing strategy was introduced prospectively into the LUME-Lung 1 trial, by an extension to the trial statistical analysis plan (TSAP). This was done prior to the database lock for the final OS analysis, but after the primary PFS analysis and interim OS analysis had been done.

To minimise any potential bias resulting from the interim evaluation of OS at the time of the primary PFS analysis and to ensure the integrity of the ongoing LUME-Lung 1 trial, the interim analysis of OS was performed by a limited group of individuals who were not involved in overseeing the day-to-day conduct of the study. These individuals were held to strict confidentiality. The study team responsible for data collection and day-to-day operation of the clinical trial remained blinded. The sponsor also decided not to include the OS data in the Clinical Trial Report for the primary PFS analysis of the LUME-Lung 1 trial. In addition, the sponsor decided not to publish any of the results of analyses of the LUME-Lung 1 and 2 data before the read out for final OS of the LUME-Lung 1 trial(3).

Please note that, although the analysis of OS in patients of adenocarcinoma histology was prespecified, the PFS analysis in this patient population was conducted retrospectively.

Hypothesis generation using LUME-Lung 2 data

LUME-Lung 2 was an international, multicentre, randomised, double-blind, placebo-controlled phase III trial, where patients with advanced, metastatic (stage IIIB/IV) or recurrent NSCLC of non-squamous histology (adenocarcinoma, large cell carcinoma and unspecified non-squamous histology) who progressed after first-line treatment with chemotherapy received either nintedanib or matching placebo in combination with pemetrexed. The study design and the study endpoints were comparable to those in LUME-Lung 1(21).

A pre-planned interim futility analysis of investigator assessed PFS events in the LUME-Lung 2 trial, conducted by an independent DMC, indicated that the primary endpoint (PFS) would likely not be met (however, no safety concerns were identified)(21). Due to the DMC's recommendation, the study was stopped on 18 June 2011(61). A subsequent analysis showed that the primary endpoint of

centrally reviewed PFS was met and that the addition of nintedanib to pemetrexed resulted in a statistically significant PFS prolongation compared to placebo (4.4 *vs* 3.6 months, respectively; HR: 0.83; 95% CI: 0.70-0.99; p=0.0435)(21).

Following the halt of the study, the DMC recommended that data from LUME-Lung 2 should be analysed to better understand the futility outcome and to identify a patient population that would benefit from treatment with nintedanib(62).

The approach followed the principles of a prognostic and predictive enrichment concept(101). Any identified clinical biomarker was required to be both prognostic and predictive, with a consistently observed treatment effect for both investigator- and centrally assessed PFS, and OS. Prognostic baseline variables were thus initially identified in the placebo arm of the LUME-Lung 2 study. The interaction of identified prognostic variables with treatment was then explored to identify variables that were also predictive of a nintedanib treatment benefit. This was done using centrally assessed PFS data and interim OS data obtained at the time of the primary analysis of the LUME-Lung 1 and LUME-Lung 2 trials. Finally, these hypotheses were to be validated using final OS data from the LUME-Lung 1 trial. To ensure comparability across studies, the clinical biomarker was confirmed and validated using data from non-squamous patients only(5).

Time since start of first-line therapy was identified as the only prognostic clinical biomarker that was also predictive for the treatment effect of nintedanib in combination with pemetrexed or docetaxel in the LUME-Lung 2 and LUME-Lung 1 studies(62). An inverse relationship between the length of time since start of first-line therapy and the treatment effect of nintedanib plus second-line chemotherapy was shown for PFS and OS; the shorter the time from start of first-line therapy to randomisation, the better the treatment effect. The effect observed in non-squamous patients was primarily driven by patients with adenocarcinoma. To categorise the continuous variable 'time since start of first-line therapy', a cut-off of 9 months was chosen based on the width of the 95% CI and the time when the upper boundary of the 95% CI approached a HR of 1 (T<9m)(5).

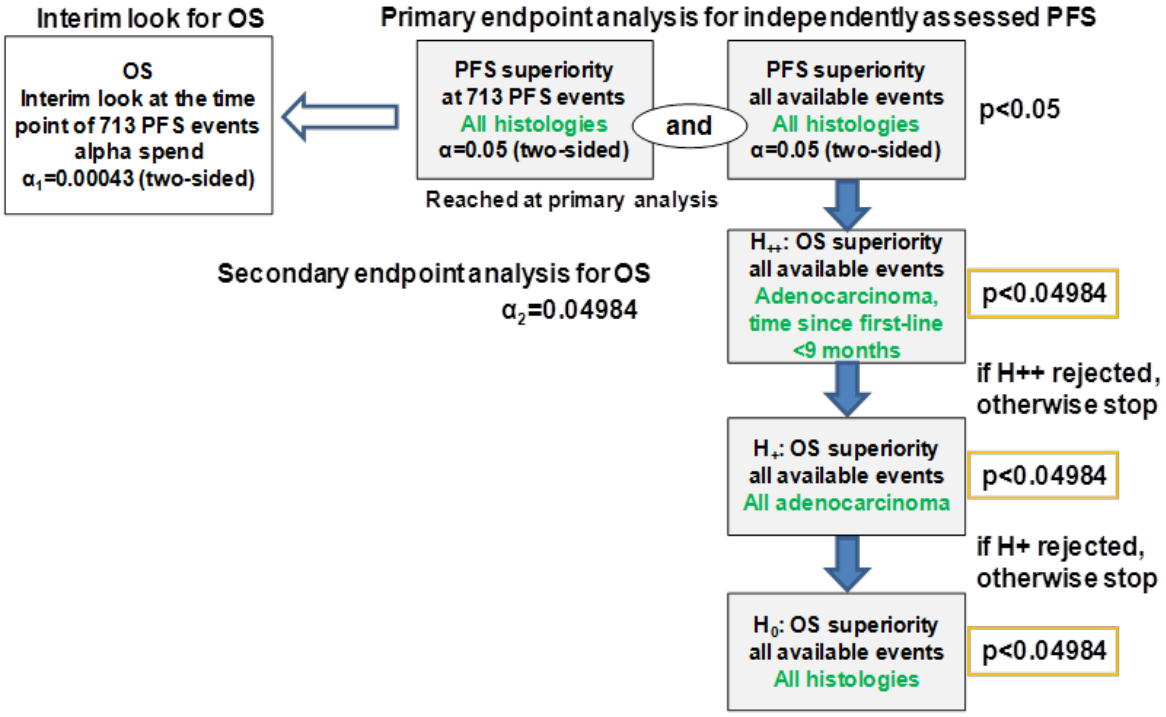
Hypothesis validation using LUME-Lung 1 data

The statistical analysis plan of LUME-Lung 1 was amended before database lock and un-blinding of data for the final OS analysis, in order to validate the hypothesis-generating findings from the primary analyses of LUME-Lung 2 and LUME-Lung 1 data (see Section 6.3.6). A hierarchical procedure was applied to control the type I error rate when analysing the key secondary endpoint of

OS. Formal statistical testing for OS in LUME-Lung 1 was therefore only allowed if the difference in the primary endpoint PFS was significant and confirmed by a further PFS analysis at the time of final OS analysis. OS was then to be tested in a pre-specified stepwise fixed-sequence order of statistical hypotheses: first in adenocarcinoma patients whose disease had progressed within 9 months of starting their first-line therapy (i.e. the T<9m population), followed by all adenocarcinoma patients, and finally the overall trial population, regardless of histology (Figure 3). Each of the three hypotheses could only be tested at the pre-specified alpha level if the previous null hypothesis in the testing sequence had been rejected(5).

The overall alpha level for the OS analysis followed a Lan-DeMets spending function with O'Brien-Fleming shape parameter to preserve an overall two-sided alpha-level of 0.05(5). Therefore, the exact overall alpha level depended on the number of deaths that had occurred at the interim OS analysis (423 deaths) and the number of deaths that occurred by the time of the final OS analysis (1,121 deaths). Due to the final number of OS events, the alpha level used for the fixed sequence-testing procedure described above was 0.04984(63).

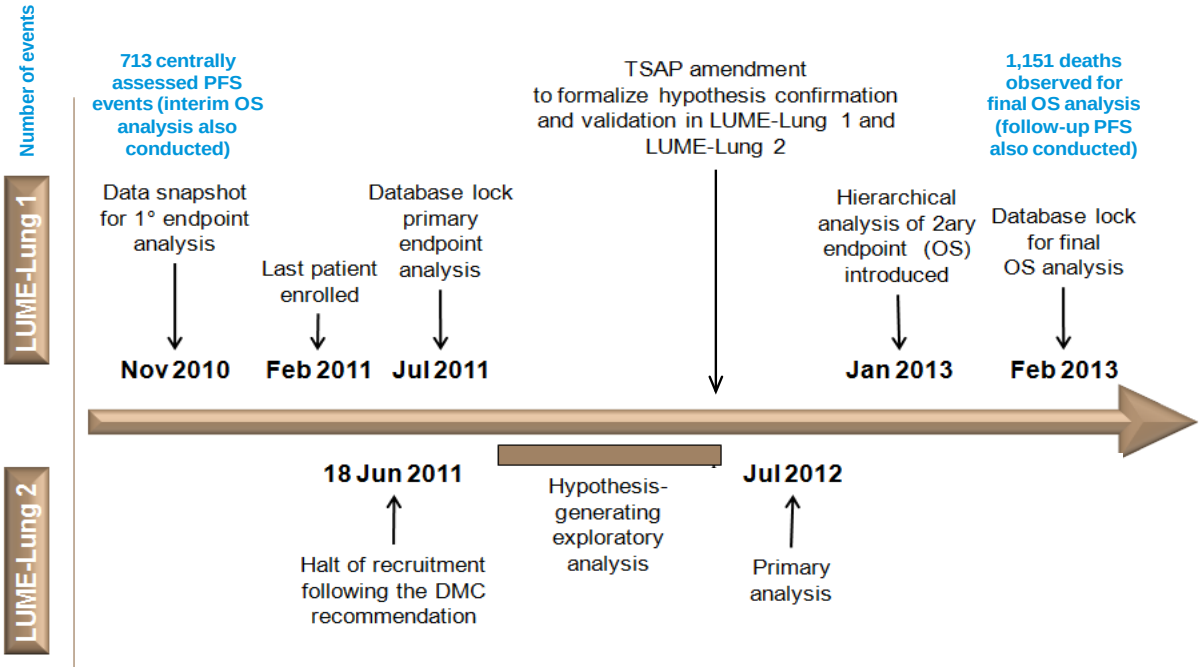
Figure 3: Analysis strategy for the key secondary endpoint of OS in LUME-Lung 1 via subgroup hierarchical testing(63)



OS = overall survival; PFS = progression-free survival

The timing of the amendment to the trial statistical analysis plan is presented in Figure 4.

Figure 4: Data analysis timeline in LUME-Lung 1 and LUME-Lung-2(63)



DMC = data monitoring committee; OS = overall survival; TSAP = Trial statistical analysis plan

Participant flow

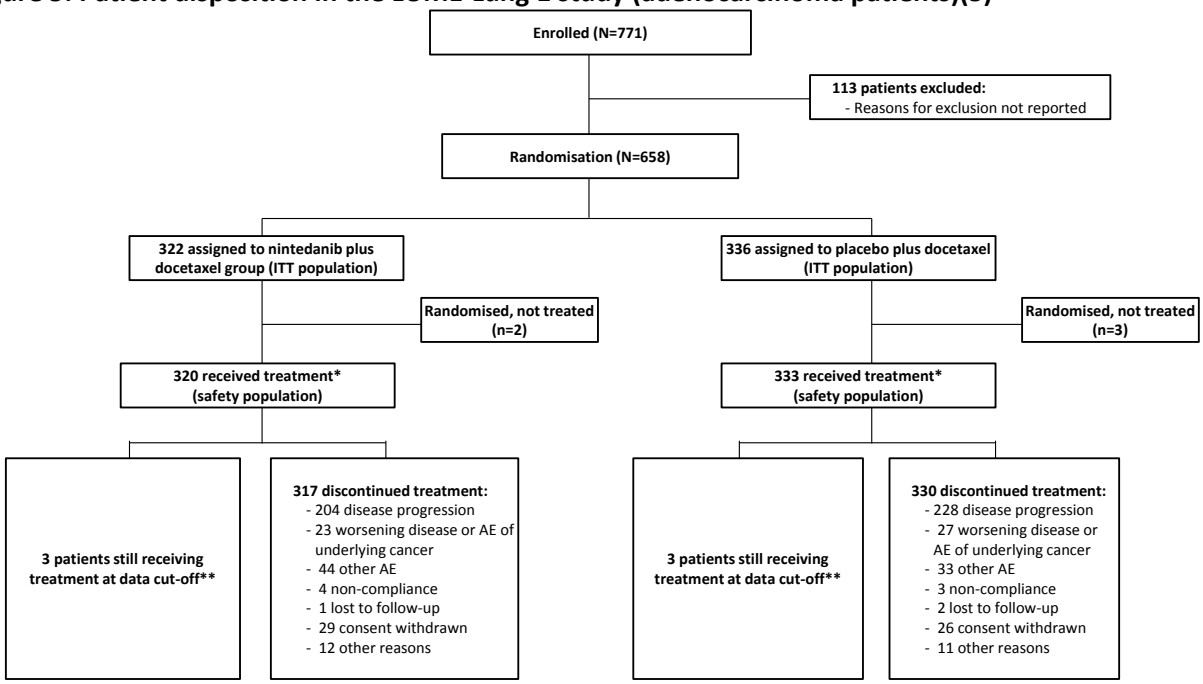
6.3.8 Provide details of the numbers of patients who were eligible to enter the RCT(s), randomised, and allocated to each treatment. Provide details of, and the rationale for, patients who crossed over treatment groups and/or were lost to follow-up or withdrew from the RCT. This information should be presented as a CONSORT flow chart.

In the LUME-Lung 1 study a total of 771 patients with adenocarcinoma were screened. Of these patients, 658 were randomised to treatment with nintedanib plus docetaxel (322 patients) or placebo plus docetaxel (336 patients)(5). Of the 658 randomised patients, five never received treatment(5).

At the data cut-off (15 February 2013) three patients in the nintedanib group and three in the placebo group were still receiving treatment(5). A total of 99.1%% patients had permanently discontinued all study medication. The reasons for permanent discontinuation of last study

treatment were comparable between both treatment groups. The most frequent reason was disease progression (66.8%). Other commonly reported reasons included AEs (11.9%) and worsening disease or AE attributable to underlying cancer disease (7.7%). A summary of the disposition of patients in the LUME-Lung 1 study can be seen in Table 10(5).

Figure 5: Patient disposition in the LUME-Lung 1 study (adenocarcinoma patients)(5)



*This patient set includes all patients who received at least one dose of study medication (Docetaxel and/or Nintedanib / Placebo).
** Refers to CTR cut−off date

6.4 Critical appraisal of relevant RCTs

6.4.1 The validity of the results of an individual study will depend on the robustness of its overall design and execution, and its relevance to the decision problem. Each study that meets the criteria for inclusion should therefore be critically appraised. Whenever possible, the criteria for assessing published studies should be used to assess the validity of unpublished and part-published studies. The critical appraisal will be validated by the ERG. The following are the minimum criteria for assessment of risk of bias in RCTs, but the list is not exhaustive.

- *Was the method used to generate random allocations adequate?*

- *Was the allocation adequately concealed?*
- *Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?*
- *Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?*
- *Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?*
- *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?*

6.4.2 Please provide as an appendix a complete quality assessment for each RCT. See section 10.3, appendix 3 for a suggested format.

6.4.3 If there is more than one RCT, tabulate a summary of the responses applied to each of the critical appraisal criteria. A suggested format for the quality assessment results is shown below.

Critical appraisal of LUME-Lung 1 can be seen in Table 17 and Appendix 10.3 provides a complete quality assessment for the RCT.

Table 17: Quality assessment results for RCT(5)

Trial no. (acronym)	**LUME-Lung 1**
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?	Yes
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

6.5 Results of the relevant RCTs

6.5.1 Provide the results for all relevant outcome measure(s) pertinent to the decision problem. Data from intention-to-treat analyses should be presented whenever possible and a definition of the included patients provided. If patients have been excluded from the analysis, the rationale for this should be given. ***If there is more than one RCT, tabulate the responses.***

6.5.2 The information may be presented graphically to supplement text and tabulated data. If appropriate, please present graphs such as Kaplan–Meier plots.

6.5.3 For each outcome for each included RCT, the following information should be provided.

- *The unit of measurement.*
- *The size of the effect; for dichotomous outcomes, the results ideally should be expressed as both relative risks (or odds ratios) and risk (or rate) differences. For time-to-event analysis, the hazard ratio is an equivalent statistic. Both absolute and relative data should be presented.*
- *A 95% confidence interval.*

- *Number of participants in each group included in each analysis and whether the analysis was by 'intention to treat'. State the results in absolute numbers when feasible.*
- *When interim RCT data are quoted, this should be clearly stated, along with the point at which data were taken and the time remaining until completion of that RCT. Analytical adjustments should be described to cater for the interim nature of the data.*
- *Other relevant data that may assist in interpretation of the results may be included, such as adherence to medication and/or study protocol.*
- *Discuss and justify definitions of any clinically important differences.*
- *Report any other analyses performed, including subgroup analysis and adjusted analyses, indicating those pre-specified and those exploratory.*

Data for all endpoints were collected, analysed, and reported for the overall trial population(3). Where available, only data for patients with adenocarcinoma histology are presented in this submission, as this is the relevant indication sought from the regulatory authorities.

Results of the primary endpoint analysis (PFS based on a central independent review) for the overall trial population and OS for the T<9m population are also presented, as these were a part of the fixed-sequence hierarchical OS statistical analysis(3):

- PFS (primary and follow-up analysis): overall trial population (all histologies)
- OS: T<9m adenocarcinoma sub-group
- OS: all adenocarcinoma patients
- OS: overall trial population (all histologies – not presented)

Statistical analysis for OS was only permitted if the result for the primary endpoint PFS was significant at $p<0.05$ for both the primary analysis (with 713 patients) and follow-up analysis (all available PFS events at the time of the final OS analysis). Each of the subsequent steps in the OS testing sequence had to reach significance at $p<0.04984$ to justify the next level of test. Therefore, each of the three OS analyses could only be carried out if the previous null hypothesis in the testing sequence had been rejected (see also Section 6.3.7)(3, 60).

Pharmacokinetic data collected as a secondary outcome is not relevant to the decision problem and is not presented in this submission.

In the adenocarcinoma population, the median duration of treatment was 4.3 months (range 0.10 months to 41.53 months) in the nintedanib plus docetaxel arm and 3.0 months (range 0.07 months to 31.10 months) in the placebo plus docetaxel arm(5). The mean dose intensity of nintedanib/placebo in adenocarcinoma patients was 91.2% for nintedanib plus docetaxel and 93.8% for placebo plus docetaxel(5). The median number of docetaxel cycles was higher in the nintedanib plus docetaxel arm than in the docetaxel plus placebo arm (5 vs 4 cycles, respectively)(5).

Primary endpoint: PFS

The primary analysis of PFS was to be performed when 713 patients had died or experienced disease progression as determined by central independent review. The total number of patients with a PFS event was 710 on 1 November 2010, and 714 on 2 November 2010. Therefore, 2 November 2010 was identified as the efficacy cut-off date. Median follow-up was 7.1 months (interquartile range 3.8–11.0 months) at the time at the primary PFS analysis(3).

The LUME-Lung 1 study met its primary endpoint based on a central independent review of 714 PFS events (Table 18)(3):

- In the overall trial population (HR: 0.79, 95% CI: 0.68-0.92, p=0.0019 for the primary analysis and HR: 0.85, 95% CI: 0.75-0.96, p=0.0070 for the follow-up analysis)
- In the adenocarcinoma population (HR: 0.77, 95% CI: 0.62-0.96, p=0.0193 for the primary analysis and HR: 0.84 95% CI: 0.71-1.00, p=0.0485 for the follow-up analysis). This analysis was conducted retrospectively.

The primary and follow-up analyses of centrally assessed PFS for these patient populations are summarised in Table 18 and Table 19(3).

Primary analysis

Table 18: Primary analysis of centrally assessed PFS in LUME-Lung 1(3)

	Nintedanib + docetaxel (median)[1]	Placebo docetaxel (median)[1]	HR *vs* placebo arm (95% CI)[2]	P-value	Risk reduction
PFS in overall trial population	3.4 months	2.7 months	0.79 (0.68-0.92)	0.0019	21%
PFS in adenocarcinoma population[3]	4.0 months(58)	2.8 months(58)	0.77 (0.62-0.96)	0.0193	23%

HR = hazard ratio; PFS = progression-free survival

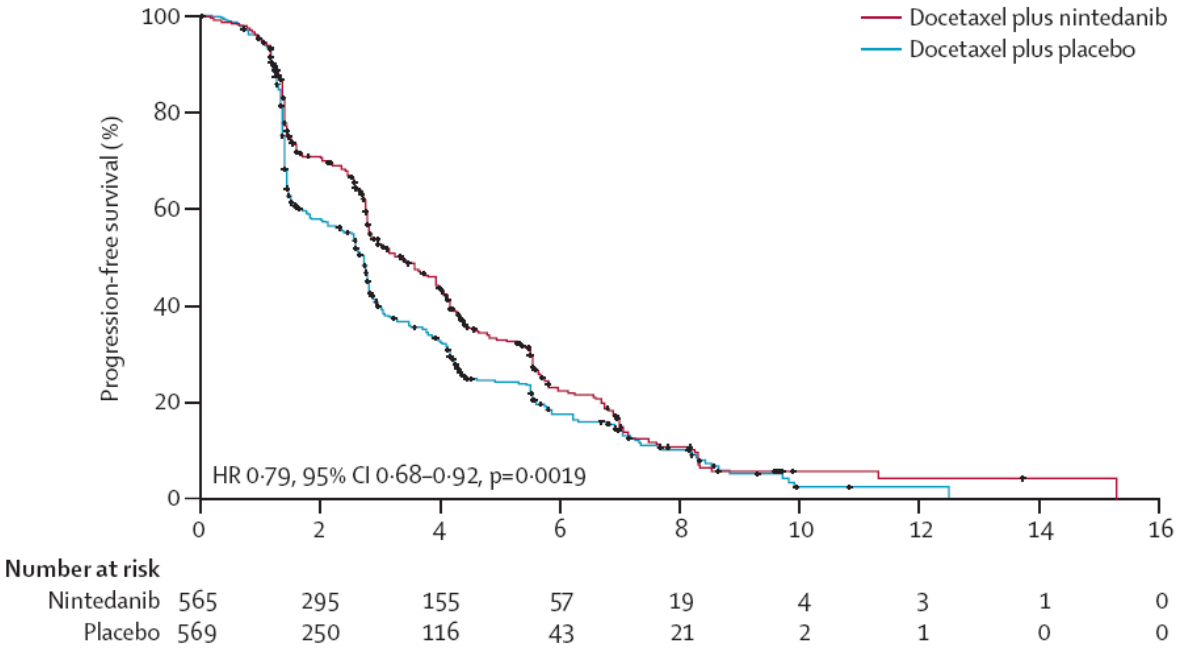
[1] Based on unadjusted Kaplan-Meier estimates for each treatment arm

[2] A proportional hazards model stratified by three factors (ECOG PS at baseline, presence of brain metastases at baseline, prior bevacizumab therapy) was used to derive the HR, 95% CI, and *P*-value (corresponding to the stratified log-rank test *P*-value)

[3] Analysis conducted retrospectively

XIn the primary analysis of centrally assessed PFS in the overall trial population, the addition of nintedanib to docetaxel significantly prolonged median PFS compared with placebo and docetaxel (3.4 months *vs* 2.7 months; HR 0.79; 95% CI 0.68-0.92; p=0.0019)(3). The Kaplan-Meier curves separated after the first pre-planned tumour imaging (i.e. after 6 weeks), when about 70% of the patients were estimated to be progression-free (Figure 6)(3). These results were confirmed in the follow-up PFS analysis done at the time of the final OS analysis (Table 19)(3).

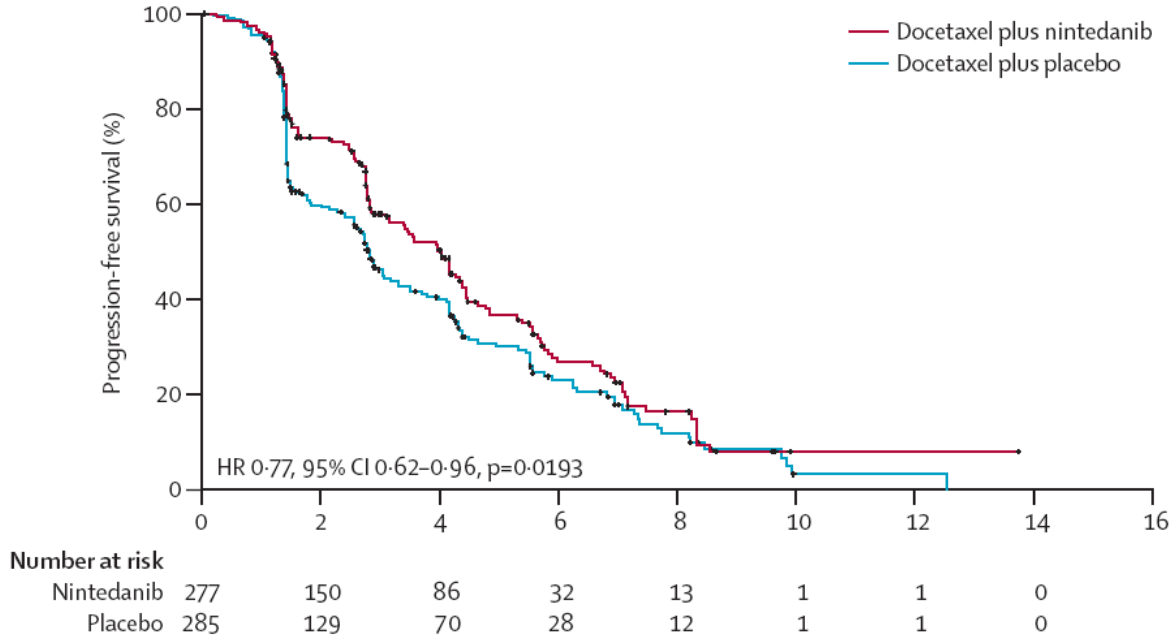
Figure 6: Probability rate of PFS in the overall trial population in LUME-Lung 1 (primary analysis, central review)(3)



CI = confidence interval; HR =hazard ratio

In the primary analysis of centrally assessed PFS in patients with adenocarcinoma histology, , which was conducted retrospectively, the addition of nintedanib to docetaxel significantly prolonged median PFS compared with placebo and docetaxel (4.0 months *vs* 2.8 months; HR 0.77; 95% CI 0.62-0.96; p=0.0193, Figure 7, Table 18)(3, 58). The Kaplan-Meier curves for PFS (primary analysis) in adenocarcinoma patients separated after 6 weeks and remained separated for the major part of the observation period (Figure 7). Results for this patient population were confirmed in the PFS follow-up analysis (4.2 months vs 2.8 months; HR 0.84; 95% CI 0.71-1.00, Table 19)(3).

Figure 7: Probability rate of PFS in the adenocarcinoma population in LUME-Lung 1 (primary analysis, central review)(3)



CI = confidence interval; HR =hazard ratio

Follow-up analysis

The data obtained in the primary PFS analyses for all patient populations were confirmed in the follow-up analyses (Table 19)(3). Median follow-up was 31.7 months (interquartile range: 27.8-36.1 months) at the time of the final OS analysis(3).

Table 19: Follow-up analysis of centrally assessed PFS conducted at a time of final OS analysis in LUME-Lung 1(3)

	Nintedanib + docetaxel (median)[1]	Placebo + docetaxel (median)[1]	HR *vs* placebo arm (95% CI)[2]	P-value	Risk reduction
PFS in the overall trial population	3.5 months	2.7 months	0.85 (0.75-0.96)	0.0070	15%
PFS in adenocarcinoma population[3]	4.2 months	2.8 months	0.84 (0.71-1.00)	0.0485	16%

HR = hazard ratio; PFS = progression-free survival

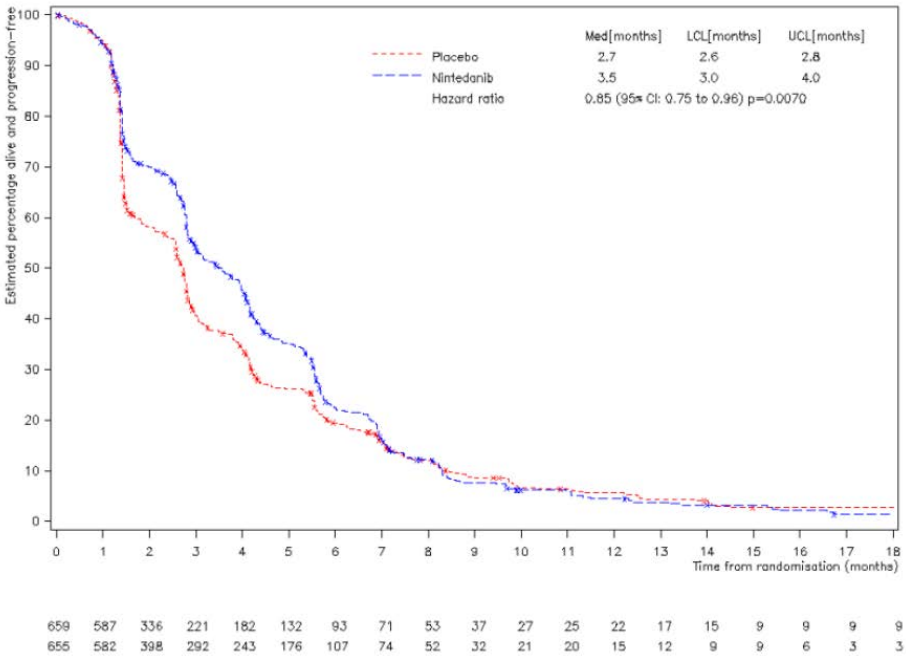
[1] Based on unadjusted Kaplan-Meier estimates for each treatment arm

[2] A proportional hazards model stratified by three factors (ECOG PS at baseline, presence of brain metastases at baseline, prior bevacizumab therapy) was used to derive the HR, 95% CI, and *P*-value (corresponding to the stratified log-rank test *P*-value)

[3] Analysis conducted retrospectively

In the overall trial population the median PFS at the time of the follow-up analysis was 3.5 months in the nintedanib plus docetaxel arm and 2.7 months in the placebo plus docetaxel arm (HR 0.85; 0.75-0.96; p=0.0070, Figure 8, Table 19)(3).

Figure 8: Probability rate of PFS in the overall trial population in LUME-Lung 1 (follow-up analysis)(5)

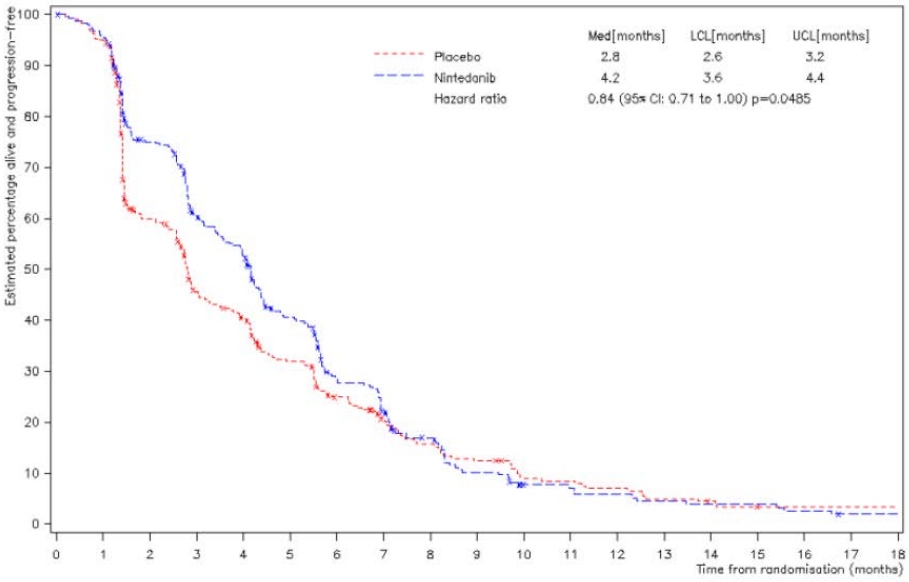


CI = confidence interval; LCL = lower control limit; Med = median; UCL = upper control limit

In the adenocarcinoma population the median PFS at the time of the follow-up analysis was 4.2 months in the nintedanib plus docetaxel arm compared to 2.8 months in the placebo plus docetaxel

arm (HR 0.84; 95% CI 0.71-1.00, p=0.0485, Figure 9, Table 19)(3). This analysis was conducted retrospectively.

Figure 9: Probability rate of PFS in the adenocarcinoma population in LUME-Lung 1 (follow-up analysis)(5)



CI = confidence interval; LCL = lower control limit; Med = median; UCL = upper control limit

Secondary endpoint: OS

The key secondary endpoint OS was analysed according to a pre-defined hierarchical statistical analysis plan, in a pre-specified step-wise order(3):

(1) OS: T<9m adenocarcinoma sub-group

(2) OS: all adenocarcinoma patients

(3) OS: overall trial population (all histologies) – not presented

Statistical testing for OS was only allowed if the difference in the primary endpoint PFS was significant both at the time of the primary analysis and follow-up analysis. Each of the subsequent OS analyses could only be carried out if the previous null hypothesis in the testing sequence had been rejected (see also Section 6.3.7).

As the null hypotheses for the primary endpoint (PFS at the time of primary and follow-up analyses in the overall trial population) had been rejected, the hierarchical OS testing proceeded in the pre-defined order specified above (steps 1-3) (Table 20). In step 1, the T<9m adenocarcinoma population

in LUME-Lung 1, treatment with nintedanib plus docetaxel significantly prolonged OS(3). Median OS was 3 months longer in the nintedanib arm than in the placebo arm (10.9 months and 7.9 months respectively, [HR 0.75; 95% CI 0.60-0.92; p=0.0073]). Analysis of OS in all adenocarcinoma patients (step 2) was therefore permitted. No further data on the T<9m adenocarcinoma subgroup are presented, as the decision problem of this submission specifies all second-line NSCLC patients with adenocarcinoma.

Data for OS in the overall trial population are not presented, as nintedanib is not indicated in this population, and as this population was the last to be tested according to the hierarchical OS testing strategy (see Figure 3).

Table 20: OS in LUME-Lung 1 in the adenocarcinoma population(3)

	Nintedanib + docetaxel (median)[1]	**Placebo + docetaxel (median)[1]**	**HR *vs* placebo arm (95% CI)[2]**	**P-value**
OS in adenocarcinoma population, final analysis	12.6 months	10.3 months	0.83 (0.70-0.99)	**0.0359**

[1] Based on unadjusted Kaplan-Meier estimates for each treatment arm

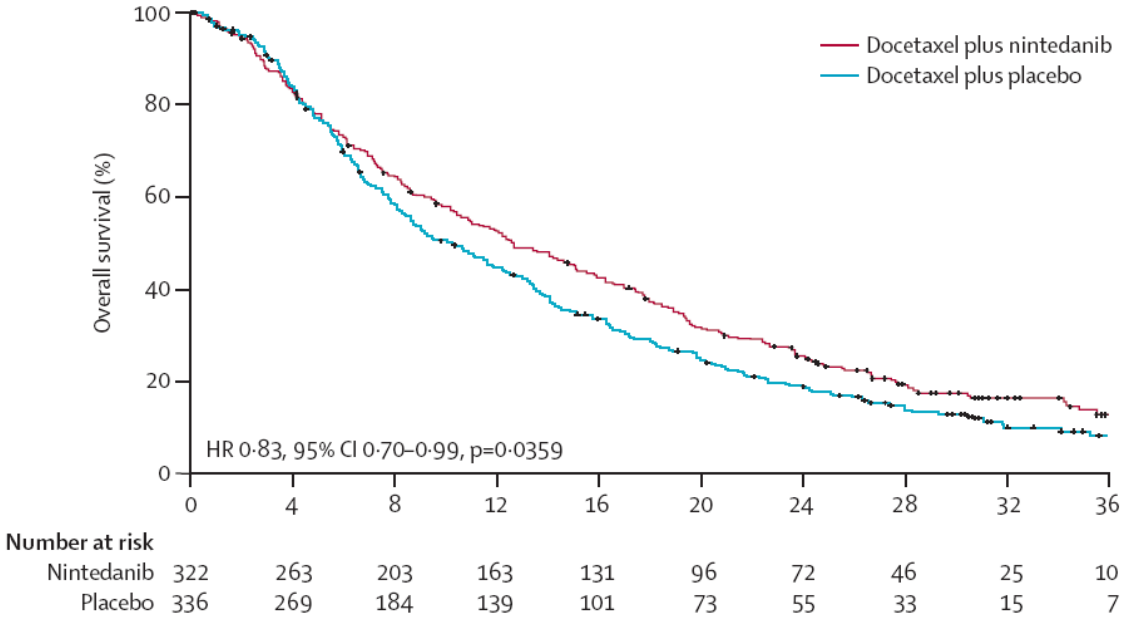
[2] A proportional hazards model stratified by three factors (ECOG PS at baseline, presence of brain metastases at baseline, prior bevacizumab therapy) was used to derive the HR, 95% CI, and P-value (corresponding to the stratified log-rank test *P*-value)

[3] T<9m adenocarcinoma population = subgroup of patients with adenocarcinoma who progressed during or shortly after the first-line therapy and who were enrolled in the trial less than 9 months since the start of the first-line therapy

For the adenocarcinoma population in LUME-Lung 1, the median OS was 12.6 months *vs* 10.3 months in the nintedanib plus docetaxel arm and the placebo plus docetaxel arm, respectively (Table 20)(3). This difference of 2.3 months was statistically significant (HR 0.83; 95% CI 0.70-0.99; p=0.0359). The Kaplan-Meier curves for OS separated after 6 months and remained apart over the entire observation period of approximately 36 months (Figure 10)(3).

In the adenocarcinoma population, one-year survival was 52.7% (95% CI 46.8-57.9) in the nintedanib plus docetaxel group compared with 44.7% (95% CI 38.9-49.8) in the placebo plus docetaxel group. Two-year survival was 25.7% (95% CI 20.5-30.2) in the nintedanib plus docetaxel group compared with 19.1% (95% CI 14.4-23.2) in the placebo plus docetaxel group(3).

Figure 10: Probability of OS in the adenocarcinoma population in LUME-Lung 1(3)



CI = confidence interval; HR = hazard ratio

PFS by investigator review (primary analysis)

Similar to PFS analyses based on central independent review, the addition of nintedanib to docetaxel showed a significant improvement in investigator assessed median PFS in the adenocarcinoma patient population, compared with placebo plus docetaxel (HR 0.78; 95% CI 0.62-0.97; p=0.0246)(3).

Tumour response based on central independent review

A summary of the tumour response assessment based on central independent review is presented below(5).

In the adenocarcinoma patient population, significantly more patients achieved disease control (60.2%) when treated with nintedanib plus docetaxel compared to those treated with placebo plus docetaxel (44.0%): odds ratio (OR, by logistic regression adjusted for baseline ECOG performance status) for disease control in patients with adenocarcinoma was 1.93 (95% CI 1.42-2.64; p<0.0001) (Table 21)(3). There was no significant difference in objective tumour response rates between the nintedanib and placebo groups (4.7% vs 3.6%; OR 1.32 (0.61-2.93), p=0.4770, Table 21)(3). The median duration of response in this patient population was 4.9 months for the nintedanib plus docetaxel arm and 4.3 months in the placebo plus docetaxel arm. The median time to confirmed objective response was 1.6 months for nintedanib plus docetaxel and 5.1 months for placebo plus docetaxel(Table 21)(5).

Table 21: Tumour response and disease control according to modified RECIST version 1.0 in the adenocarcinoma population in LUME-Lung 1 trial (by central independent review at the time of final OS analysis)(3, 5)

	Adenocarcinoma population		
	Nintedanib + placebo (n=322)	**Placebo + docetaxel (n=336)**	**Odds ratio[1] (95% CI)**
Patients with objective tumour response, n (%)(3)	15 (4.7)	12 (3.6)	1.32 (0.61-2.93), p=0.4770
Complete response, n (%)	0	0	-
Partial response, n (%)	15 (4.7)	12 (3.6)	-
Unconfirmed complete/partial response n (%)	10 (3.1)	7 (2.1)	-
Median duration of confirmed objective response (months)	4.9	4.3	-
Median time to confirmed objective response (months)	1.6	5.1	-
Stable disease[2], n (%)	179 (55.6)	136 (40.5)	-
Patients with disease control[3], n (%)(3)	194 (60.2)	148 (44.0)	1.93 (1.42-2.64); p<0.0001
Median duration of disease control (months)	5.7	6.3	-
Progressive disease[4], n (%)	87 (27.0)	147 (43.8)	-
Other[5], n (%)	41 (12.7)	41 (12.2)	-

ECOG PS = Eastern Cooperative Oncology Group performance status; SD = stable disease; PD = progressive disease

[1] Odds ratios were obtained from logistic regression model adjusted for baseline ECOG PS

[2] SD was assumed if a follow-up imaging indicated SD at least once and at least 6 weeks after randomisation (i.e. at or after Day 43).

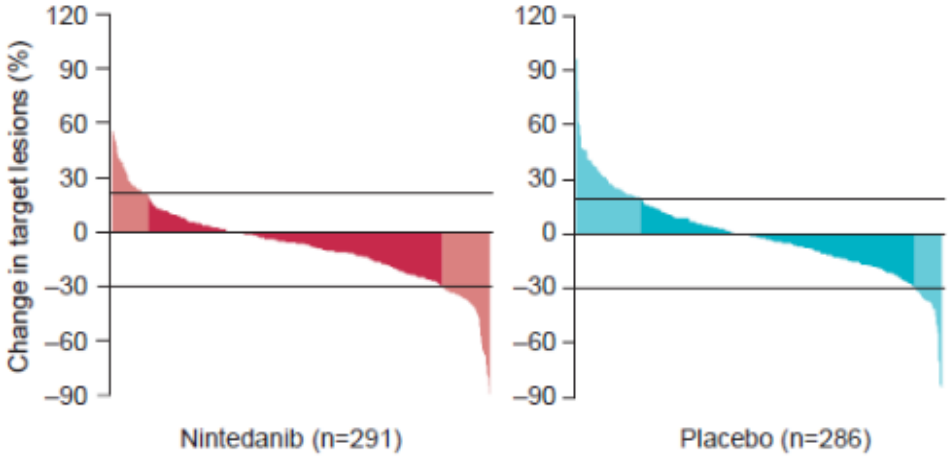
[3] A patient was considered to have disease control if he/she had a best objective response of SD or better.

[4] Including patients with SD from a radiological imaging earlier than Day 43 followed by PD

[5] Including patients with SD from a radiological imaging earlier than Day 43 followed by a non-evaluable response

At the time of the final analysis, there was a statistically significant difference in the adjusted mean of the best percentage change in sum of the longest diameters of target lesions from baseline in adenocarcinoma patients (-7.76% [95% CI -10.25-(-5.26%)] vs -0.97% [95% CI -3.48-(+1.55%)], respectively; p=0.0002, Figure 11)(3).

Figure 11: Best percentage change in the sum of the longest diameters of the target lesions in patients with adenocarcinoma histology(3)



Clinical Improvement

Clinical improvement was only reported for the overall study population, data are unavailable for the adenocarcinoma histology subgroup. Clinical improvement was defined as the prolongation of time to deterioration in body weight (first occurrence of a decrease from baseline of more than 10%) and/or first increase of one category of the ECOG performance score from baseline, whichever occurs earlier. A summary of the clinical improvement analysis for the overall trial population can be seen in Table 22(5).

Table 22: Clinical improvement in the LUME-Lung 1 study (overall trial population)(5)

	Nintedanib + docetaxel	Placebo + docetaxel	HR vs placebo arm (95% CI)[3]	P-value
Patients with a deterioration event[1], n (%)	303 (46.3)	272 (41.3)	-	-
Median time to deterioration in body weight and/or ECOG PS in the overall trial population[2]	5.9 months	5.2 months	1.03 (0.87-1.21)	0.7282

ECOG PS = Eastern Cooperative Oncology Group performance status; CI = confidence interval

[1] Deterioration of body weight and/or ECOG PS

[2] Based on unadjusted Kaplan-Meier estimates for each treatment arm

[3] A proportional hazards model stratified by three factors (ECOG PS at baseline, presence of brain metastases at baseline, prior bevacizumab therapy) was used to derive the HR, 95% CI, and p-value (corresponding to the stratified log-rank test p-value)

The proportion of patients with a clinical deterioration event was similar in both treatment arms; 46.3% in the nintedanib plus docetaxel arm compared with 41.3% in the placebo plus docetaxel arm(5). The median time to deterioration was 5.9 months in the nintedanib plus docetaxel group compared with 5.2 months in the placebo plus docetaxel group (HR 1.03; 95% CI 0.87-1.21; p=0.7282) (Table 22)(5). In the majority of patients, the deterioration was an increase in ECOG PS, and the proportion of patients experiencing a deterioration in performance status was similar between groups (nintedanib plus docetaxel: 33.9% vs placebo plus docetaxel: 32.9%)(5).

HRQL

In this submission, HRQL data are presented for the adenocarcinoma population. Over 80% of patients completed the questionnaires over the first 20 cycles, and approximately 70% of patients completed the questionnaires at the end of treatment visit(59). HRQL scores for cough, dyspnoea and pain at baseline were balanced across the two treatment arms, in the overall trial population (Table 23). HRQL scores at baseline are not available for the adenocarcinoma patient population(59).

Table 23: HRQL scores at baseline for cough, dyspnoea and pain (overall trial population)(59)

	Nintedanib + docetaxel, (*n*=610)			Placebo + docetaxel, (*n*=612)		
	n	Mean score	SD	n	Mean score	SD
Cough	607	39.6	27.0	610	35.9	26.4
Dyspnoea	598	29.8	20.5	605	28.3	20.4
Pain	610	27.0	26.9	612	27.6	26.5
Global health status/QoL	609	61.2	19.9	606	62.3	19.9

SD = standard deviation, QoL = quality of life

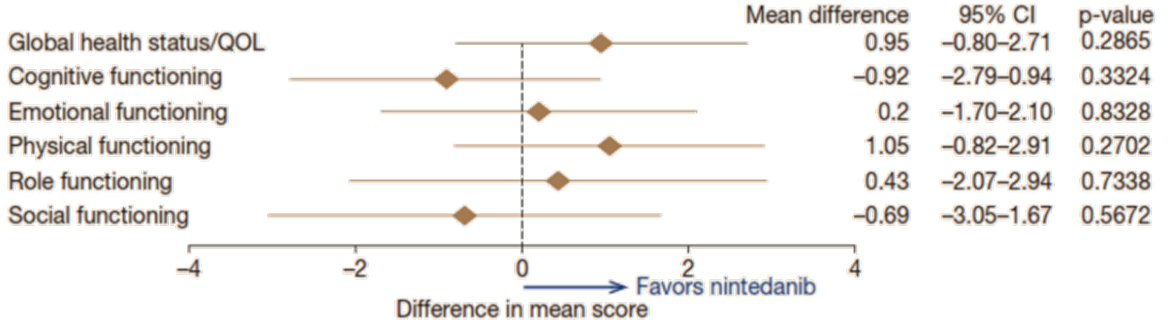
Longitudinal analysis

Treatment with nintedanib and docetaxel did not result in a change in Global health status/QoL in patients with adenocarcinoma (HR=0.86, 95% CI: 0.71–1.05)(59).

The longitudinal model analysis of differences in mean global health status/QoL and functional scales, found no significant difference between treatment groups in any aspect of the scales in patients with adenocarcinoma histology (Figure 12)(59).

Despite not reaching statistical significance, global health status/QoL, emotional functioning, role functioning and physical functioning favoured nintedanib plus docetaxel compared with placebo plus docetaxel (Figure 12)(59).

Figure 12: Longitudinal model analysis of differences in mean global health status and functional scales in patients with adenocarcinoma histology(59)

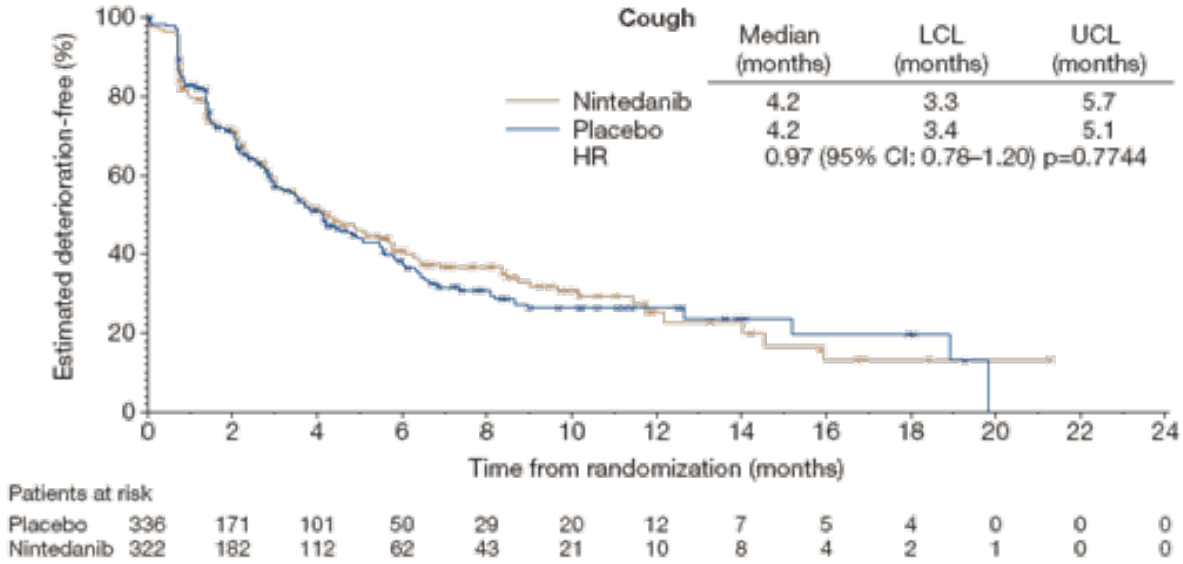


CI = confidence interval; QOL = quality of life

Time to deterioration in the adenocarcinoma population

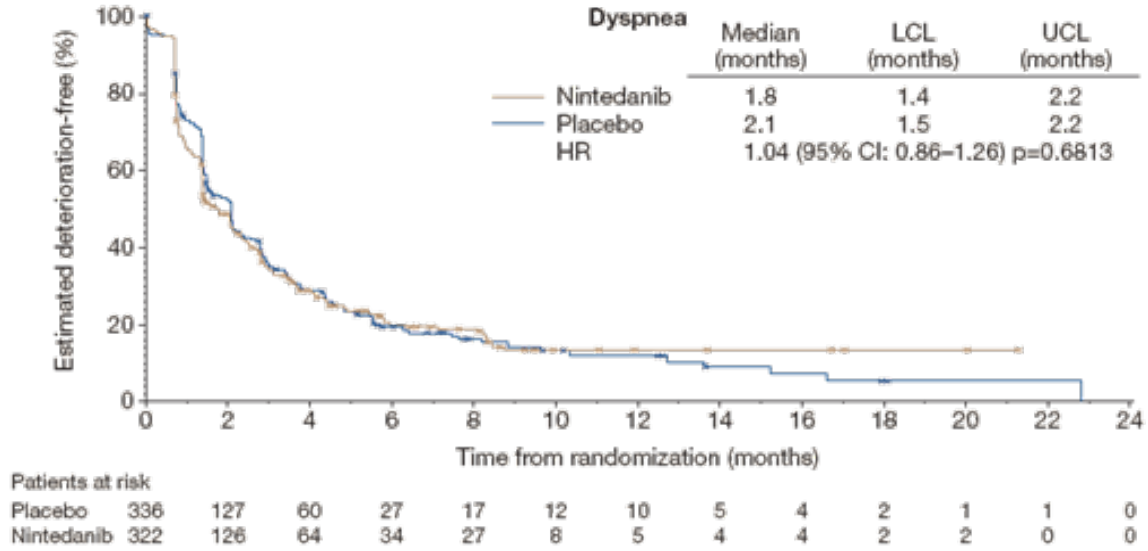
In the adenocarcinoma population, there was no significant difference in time to deterioration for the pre-specified symptoms of cough, dyspnoea and pain between the nintedanib plus docetaxel and placebo plus docetaxel arms (Figure 13, Figure 14 and Figure 15)(59).

Figure 13: Time to deterioration of cough in LUME-Lung 1 (adenocarcinoma population)(59)



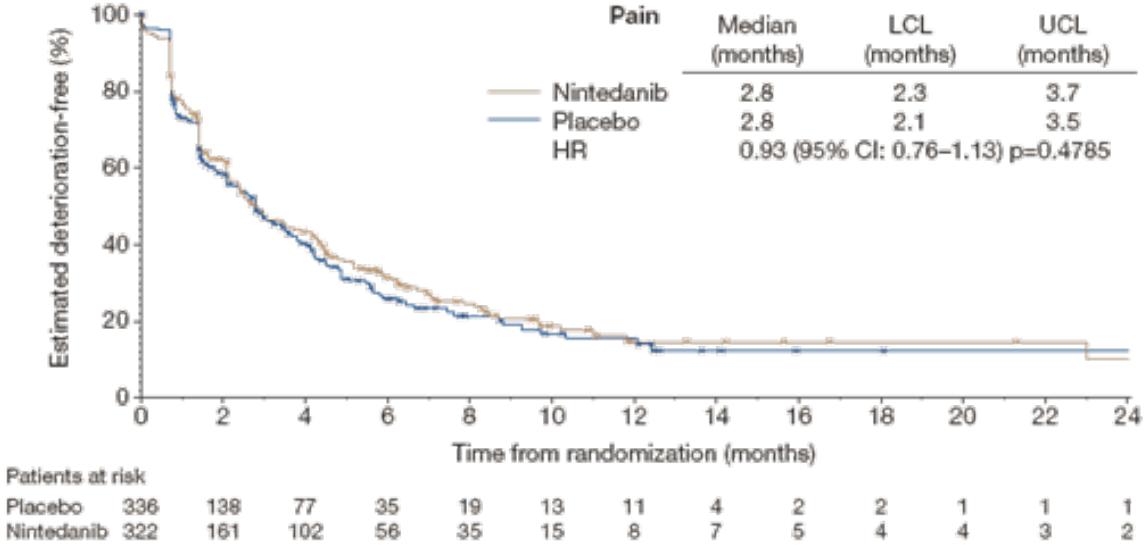
CI = confidence interval; HR = hazard ratio; LCL = lower control limit; Med = median; UCL = upper control limit

Figure 14: Time to deterioration of dyspnoea in LUME-Lung 1 (adenocarcinoma population)(59)



CI = confidence interval; HR = hazard ratio; LCL = lower control limit; Med = median; UCL = upper control limit

Figure 15: Time to deterioration of pain in LUME-Lung 1 (adenocarcinoma population)(59)



CI = confidence interval; HR = hazard ratio; LCL = lower control limit; Med = median; UCL = upper control limit

In the longitudinal analysis of differences in each of the three pre-specified symptom scores for cough, dyspnoea and pain captured by EORTC QLQ-C30 and QLQ-LC13, no statistically significant difference was observed between nintedanib plus docetaxel and placebo plus docetaxel for the adenocarcinoma population. Although significance was not achieved, nintedanib-treated patients achieved numerically better cough and pain scores than placebo-treated patients (Figure 16)(59):

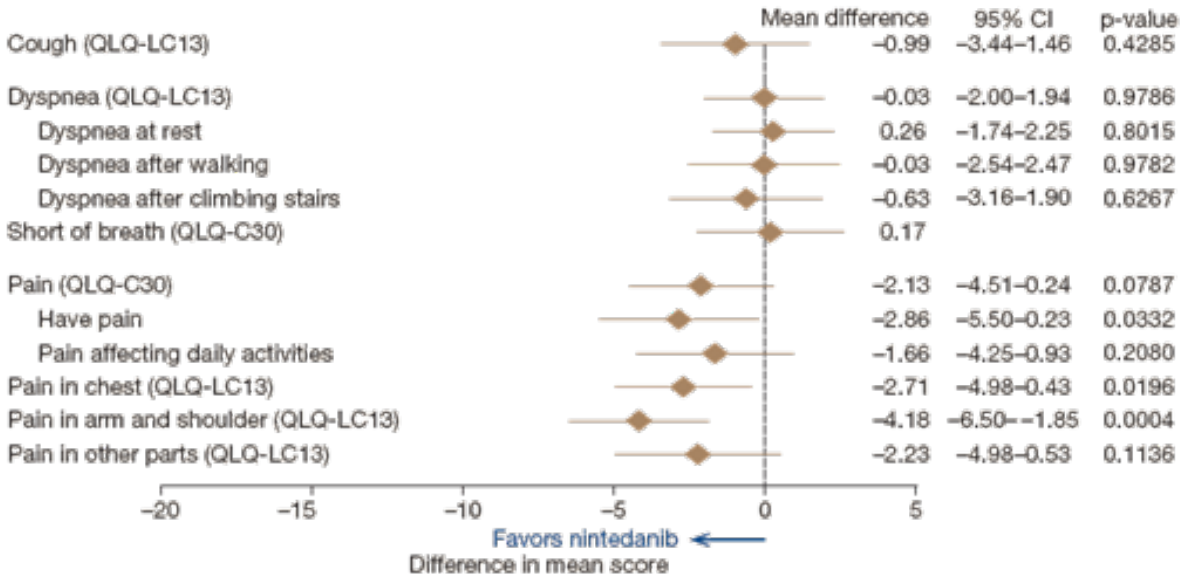
- cough (mean difference: -0.99 [-3.44, 1.46] p=0.4285)
- pain (mean difference: -2.13, [-4.51, 0. 24] p=0.0787)
- dyspnoea (mean difference: -0.03 [-2.00-1.94] p=0.9786)

However significant differences were observed favouring treatment with nintedanib vs placebo for three individual pain items (Figure 16)(59):

- 'have pain' (mean difference -2.86, [-5.50, 0.23] p=0.0332)
- 'pain in chest' (mean difference -2.71, [-4.98, 0. 43] p=0.0196)
- 'pain in arm and shoulder' (mean difference -4.18, [-6.50, -1.85] p=0.0004)

These data suggest that, while the overall scores for pain were not significantly different between treatments, the addition of nintedanib to docetaxel chemotherapy relieved some aspects of pain compared with placebo plus docetaxel(59).

Figure 16: Differences in mean scores over time for measures of cough, dyspnoea, and pain in adenocarcinoma patients(59)



CI = confidence interval; QOL = quality of life

The time to deterioration for diarrhoea was significantly worsened in adenocarcinoma patients receiving nintedanib plus docetaxel compared to patients receiving placebo plus docetaxel (HR=1.86, 95% CI: 1.51–2.30, p<0.05, Table 24)(59). There was no significant difference between groups for nausea and vomiting, or appetite loss (59).

Table 24:Time to deterioration of nausea and vomiting, appetite loss and diarrhoea in patients with adenocarcinoma(59)

Symptom	Patients with adenocarcinoma HR (95% CI)
Nausea and vomiting	1.23 (1.00-1.51)
Appetite loss	1.13 (0.92-1.38)
Diarrhoea	1.86 (1.51-2.30)*

*p<0.05

HR = hazard ratio

Sensitivity analysis

Two predefined sensitivity analyses of OS were performed. The first analysis used a proportional hazards model including the stratification factors used at randomisation as covariates. Tumour volume and tumour size are known prognostic factors for NSCLC patients and a large tumour burden at baseline is predictive of shorter survival(4). Therefore, a second predefined sensitivity analysis of OS was performed using a proportional hazards model including the stratification factors additionally adjusting by the sum of longest diameters (SLD) of target lesions at baseline(5).

For patients with adenocarcinoma the results of the OS analysis using a proportional hazards model, including ECOG PS, brain metastasis, and prior bevacizumab treatment (HR 0.83, 95% CI 0.70-0.98; p=0.0295) were similar to the main OS analysis for patients with adenocarcinoma (HR 0.83; 95% CI 0.70-0.99; p=0.0359). There was little difference in HR when baseline SLD was included in the model; the HR was 0.83 (95% CI 0.70-0.99) without SLD in the model compared to a HR of 0.81 (95% CI 0.69-0.97) when SLD was included in the model(5).

No sensitivity analysis of PFS in the adenocarcinoma population was performed.

Subgroup analysis of primary and secondary endpoint

At the time of the final OS analysis, subgroup analyses were performed for the stratification factor squamous vs non-squamous as defined in the TSAP. In addition, subgroup analyses also focused on(5):

- baseline ECOG PS (categories: 0, 1)
- presence of brain metastases at baseline (yes, no)
- prior bevacizumab treatment (yes, no)
- sex (male, female)
- age (<65years, ≥65 years)
- race (Asian vs non-Asian patients; information was derived from the

- race (categories as documented on the CRF)
- smoking status (never smoked vs currently smokes/ex-smoker)
- geographical region (Asia, Europe, South Africa; based on a patient's country of enrolment)[1]
- best response to first-line therapy (CR/PR/SD, PD, unknown/missing/NA)
- sum of longest diameters at baseline (<7.5 cm, ≥7.5 cm)
- time since first-line therapy (<9 months, ≥9 months)

The efficacy of nintedanib vs placebo for PFS and OS was also analysed for the following baseline characteristics:

- presence of liver metastases (yes, no)
- disease stage at diagnosis (<IIIB/IV, IIIB, IV)
- concomitant therapy with biphosphonates at baseline (yes, no)
- presence of adrenal metastases (yes, no)
- number of metastatic organs at baseline (≤2 metastatic organs, >2 metastatic organs, not centrally reviewed)
- lactate dehydrogenase (LDH) level at baseline (baseline LDH ≤1, baseline LDH >1)

Results for subgroup analyses in the adenocarcinoma patient population are presented below.

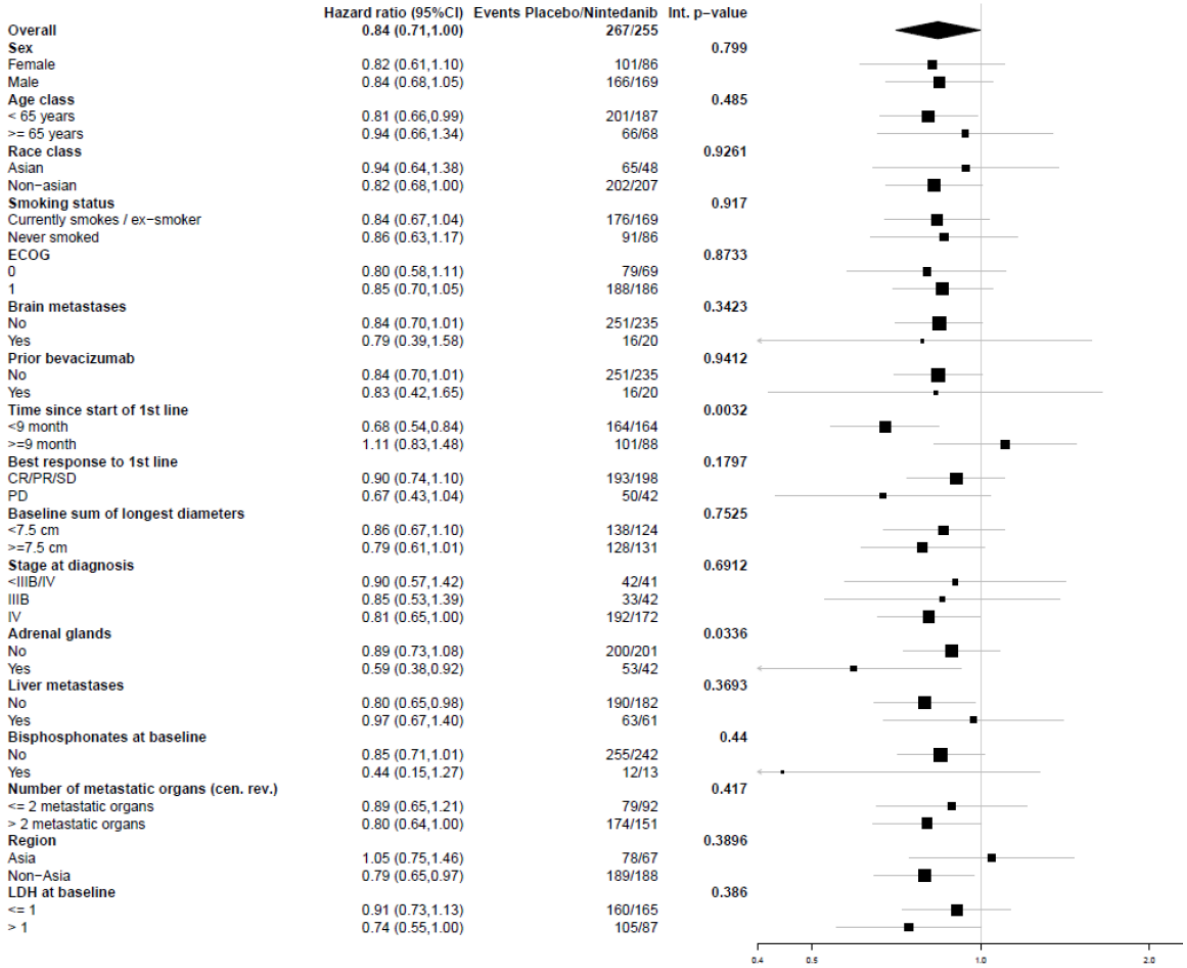
The treatment effect of nintedanib on centrally reviewed PFS, in patients of adenocarcinoma histology, was consistent across most of the analysed baseline characteristics (p-values ≥0.1, Figure 17). A significant interaction between baseline characteristic and treatment was observed for(5):

- 'time since start of first-line therapy' (p=0.0032)
- metastases in 'adrenal glands' (p=0.0336)

These findings indicate that adenocarcinoma patients who progressed sooner after the start of their first-line therapy and those with adrenal gland metastases were more likely to benefit from treatment with nintedanib plus docetaxel (Figure 17)(5).

[1] Subgroups were predefined in the TSAP except for 'geographical region', 'sum of longest diameters at baseline', 'best response to first-line therapy' and 'time since first-line therapy' which were added post hoc.

Figure 17: HR of PFS by baseline characteristics, based on central independent review – patients with adenocarcinoma(5)



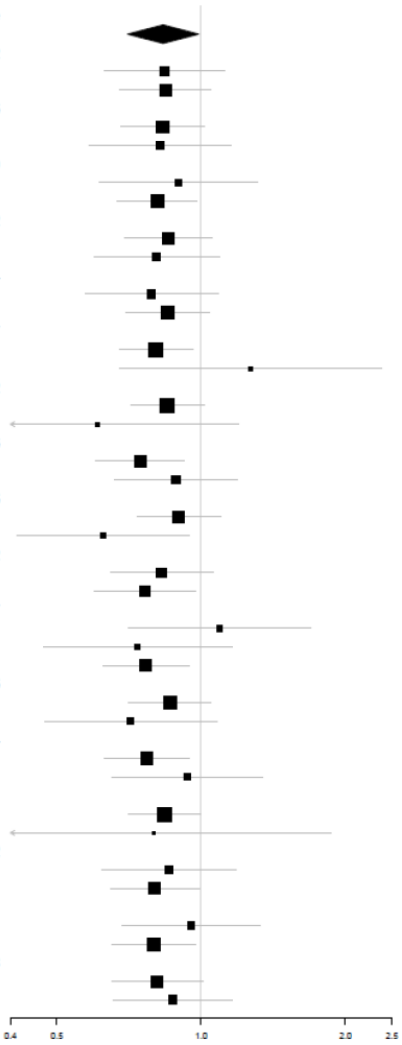
CR = complete response; ECOG = Eastern Cooperative Oncology Group; PD = progressive disease; PFS = central independent review at time of primary analysis; PR = partial response; OS = at time of final analysis; SD = stable disease

P-values ≥0.1 indicate no statistical evidence that the treatment effect varied between subgroups

The treatment effect of nintedanib plus docetaxel on OS, in patients of adenocarcinoma histology, was consistent across most of the analysed baseline characteristics (p-values ≥0.1, Figure 18). Significant interaction between baseline characteristic and treatment was observed for 'best response to first-line treatment' (p=0.0766), suggesting that there may be a more pronounced treatment benefit for those patients whose best response to first-line therapy was progressive disease. However the sample size of this sub-group was small (n=117)(5).

Figure 18: HR of OS by baseline characteristics – patients with adenocarcinoma(5)

	Hazard ratio (95%CI)	Events Placebo/Nintedanib	Int. p−value
Overall	**0.83 (0.70,0.99)**	**276/259**	
Sex			**0.7942**
Female	0.84 (0.63,1.12)	104/94	
Male	0.84 (0.68,1.05)	172/165	
Age class			**0.7556**
< 65 years	0.83 (0.68,1.02)	197/191	
>= 65 years	0.82 (0.58,1.15)	79/68	
Race class			**0.5319**
Asian	0.89 (0.61,1.31)	62/49	
Non−asian	0.81 (0.67,0.98)	214/210	
Smoking status			**0.5112**
Currently smokes / ex−smoker	0.86 (0.69,1.06)	185/175	
Never smoked	0.81 (0.60,1.09)	91/84	
ECOG			**0.7854**
0	0.79 (0.57,1.08)	79/74	
1	0.85 (0.70,1.04)	197/185	
Brain metastases			**0.1247**
No	0.80 (0.67,0.96)	258/235	
Yes	1.27 (0.67,2.38)	18/24	
Prior bevacizumab			**0.2412**
No	0.85 (0.71,1.01)	257/242	
Yes	0.61 (0.31,1.20)	19/17	
Time since start of 1st line			**0.4186**
<9 month	0.75 (0.60,0.92)	172/173	
>=9 month	0.89 (0.66,1.19)	101/83	
Best response to 1st line			**0.0766**
CR/PR/SD	0.90 (0.73,1.10)	191/197	
PD	0.62 (0.41,0.94)	58/45	
Baseline sum of longest diameters			**0.4802**
<7.5 cm	0.83 (0.65,1.06)	139/125	
>=7.5 cm	0.76 (0.60,0.97)	136/134	
Stage at diagnosis			**0.1767**
<IIIB/IV	1.09 (0.70,1.70)	42/45	
IIIB	0.74 (0.47,1.16)	36/44	
IV	0.77 (0.62,0.94)	198/170	
Adrenal glands			**0.4066**
No	0.86 (0.71,1.05)	201/198	
Yes	0.71 (0.47,1.08)	58/49	
Liver metastases			**0.244**
No	0.77 (0.63,0.95)	193/183	
Yes	0.94 (0.65,1.35)	66/64	
Bisphosphonates at baseline			**0.6501**
No	0.84 (0.70,1.00)	263/245	
Yes	0.80 (0.34,1.87)	13/14	
Number of metastatic organs (cen. rev.)			**0.3772**
<= 2 metastatic organs	0.86 (0.62,1.19)	74/85	
> 2 metastatic organs	0.80 (0.65,0.99)	185/162	
Region			**0.2731**
Asia	0.95 (0.68,1.33)	76/67	
Non−Asia	0.79 (0.65,0.97)	200/192	
LDH at baseline			**0.6878**
<= 1	0.81 (0.65,1.01)	170/159	
> 1	0.87 (0.66,1.16)	102/96	



CR = complete response; ECOG = Eastern Cooperative Oncology Group; PD = progressive disease; PFS = central independent review at time of primary analysis; PR = partial response; OS = at time of final analysis; SD = stable disease

P-values ≥0.1 indicate no statistical evidence that the treatment effect varied between subgroups

6.6 Meta-analysis

When more than one study is available and the methodology is comparable, a meta-analysis should be undertaken. This section should be read in conjunction with NICE's 'Guide to the methods of technology appraisal', sections 5.3.9 to 5.3.12.

6.6.1 The following steps should be used as a minimum when presenting a meta-analysis.

- *Perform a statistical assessment of heterogeneity. If the visual presentation and/or the statistical test indicate that the RCT results are heterogeneous, try to provide an explanation for the heterogeneity.*
- *Statistically combine (pool) the results for both relative risk reduction and absolute risk reduction using both the fixed effects and random effects models (giving four combinations in all).*
- *Provide an adequate description of the methods of statistical combination and justify their choice.*
- *Undertake sensitivity analysis when appropriate.*
- *Tabulate and/or graphically display the individual and combined results (such as through the use of forest plots).*

6.6.2 If a meta-analysis is not considered appropriate, a rationale should be given and a qualitative overview provided. The overview should summarise the overall results of the individual studies with reference to their critical appraisal.

No head-to-head randomised clinical trials were found that provided evidence of the efficacy and safety of nintedanib plus docetaxel versus pemetrexed, erlotinib, or gefitinib in the second-line treatment of adenocarcinoma of the lung. This information could only be obtained indirectly using statistical methods. On this basis, no direct meta-analysis was undertaken, and instead a mixed-treatment comparison (MTC) incorporating a network meta-analysis was formulated. Details of this analysis are provided in Section 6.7.

6.6.3 *If any of the relevant RCTs listed in response to section 6.2.4 (Complete list of relevant RCTs) are excluded from the meta-analysis, the reasons*

for doing so should be explained. The impact that each exclusion has on the overall meta-analysis should be explored.

Not applicable (see Section 6.6.2).

6.7 Indirect and mixed treatment comparisons

Data from head–to–head RCTs should be presented in the reference-case analysis, if available. If data from head–to–head RCTs are not available, indirect treatment comparison methods should be used. This section should be read in conjunction with NICE's 'Guide to the methods of technology appraisal', sections 5.3.13 to 5.3.22.

6.7.1 *Describe the strategies used to retrieve relevant clinical data on the comparators and common references both from the published literature and from unpublished data. The methods used should be justified with reference to the decision problem. Sufficient detail should be provided to enable the methods to be reproduced, and the rationale for any inclusion and exclusion criteria used should be provided. Exact details of the search strategy used should be provided in section 10.4, appendix 4*.

As described in Sections 6.6.1, Section 6.6.2 and Appendix 10.2, a systematic review of the literature was conducted to identify all potentially relevant published and non-published RCTs investigating the efficacy and safety of second-line treatments for patients with NSCLC. The literature search was conducted in MEDLINE, EMBASE, the Cochrane Central Register of Controlled Trials (CENTRAL) and Cochrane Reviews, as well as a search of recent conference proceedings from the American Society of Clinical Oncology (ASCO) and the European Society for Medical Oncology (ESMO). These sources were selected from all potential databases and conference websites as being those most likely to provide studies of acceptable quality relevant to the efficacy of NSCLC treatments.

The objective of the efficacy literature review was to identify all available randomized controlled trials, meta-analyses or systematic reviews reporting the clinical efficacy of pharmacological interventions for the second-line treatment of NSCLC using methodology published by the Centre of Reviews and Dissemination (CRD), as recommended by NICE(102).

The search strategy developed consisted of three groups of search strings, designed to identify:

- the appropriate patient population (patients with NSCLC)
- the appropriate treatments (second-line treatments)
- the appropriate study types (RCTs, systematic reviews, meta-analyses)

Separate versions of the search strings were created to conform to the different indexing terms and syntax requirements of the different databases searched. These have been provided in Appendix 10.2.4. The search was limited to articles with abstracts (to be able to identify which articles were relevant) published from 2000 onwards (to identify studies relevant to current clinical practice) and excluded any non-human studies and non-systematic (that is, narrative) reviews. Systematic reviews published in the last four years (2010–2014) were included and the reference lists of these reviews were checked to capture any trials not identified through our literature review sources. The efficacy review search was not limited to English-language publications; however, no non-English language publications were identified that met the inclusion criteria for this review.

Articles identified by the searches of the different databases were combined and duplicates removed. All abstracts were independently reviewed by two scientists with any disagreements resolved independently by a third scientist. The full text was retrieved for all articles considered to be potentially relevant on abstract screening. Two scientists reviewed each full text publication and any discrepancies were independently resolved by a third researcher. The inclusion and exclusion criteria applied to evaluate each article, based on their titles, abstracts, and subsequently full-text publication, were organised in the "PICOS" format (i.e. Population, intervention, comparator, outcomes, and study design) and are summarised below in Table 25.

Table 25: Inclusion and exclusion criteria for inclusion in MTC

	Inclusion criteria	Exclusion criteria	Rationale
Population	Relapsed or refractory NSCLC (RR NSCLC) Adults with histologically or cytologically confirmed, locally advanced and/or metastatic NSCLC of stage IIIB or IV (according to American Joint Committee on Cancers) or recurrent NSCLC (all histologies, including patients with mixed histology): Squamous-cell carcinoma Adenocarcinoma Large cell carcinoma Additional inclusion criteria applied during feasibility assessment: Study must report data for adenocarcinoma subgroup, or 75% or more of participants should have adenocarcinoma	Studies not assessing patients with locally advanced or metastatic, stage IIIB, or IV/recurrent NSCLC. Additional exclusion criteria applied during feasibility assessment: Study does not report data for an adenocarcinoma subgroup Fewer than 75% of participants overall had adenocarcinoma	The patient population evaluated in our MTC matches the population for which nintedanib is being considered for approval.
Interventions	Any second-line pharmacological treatment for RR NSCLC: Monotherapy Combination therapy with other pharmacological agents Additional inclusion criteria applied during feasibility assessment: Intervention should be licensed for use as second-line treatment for NSCLC	Trials evaluating non-second-line treatment (e.g., first-, third- or subsequent-line therapy) without subgroup data provided for second-line treatment only Dose comparison studies without a placebo or control arm Studies evaluating maintenance treatment	To evaluate nintedanib versus currently available licensed interventions for the second-line treatment of relapsed or refractory (RR) NSCLC.
Comparators	Any pharmacotherapy or no treatment: Other second-line pharmacological treatment Usual care/no additional intervention Placebo	None in addition to the above criteria	To compare included interventions with common comparators currently available for the second-line treatment of RR NSCLC, as well as usual care/no intervention and placebo.
Outcomes	Outcomes relevant to clinical efficacy and safety which were reported in the LUME-Lung 1 study, including: OS PFS OR AEs Additional inclusion criteria applied during feasibility assessment: Study must report relevant data from at least one outcome that	Study protocols without outcome data presented Studies with only patient baseline characteristics reported	We considered outcomes for which an MTC comparing nintedanib + docetaxel with other second-line treatments was feasible, and only included studies with published results for these outcomes.

	Inclusion criteria	Exclusion criteria	Rationale
	has been reported for other studies, thus enabling a comparison across treatments		
Study design	Randomised controlled trials (RCTs) only	Non-RCTs Pooled analyses of RCTs	RCTs provide the highest quality clinical trial data.
Language restrictions	Any language		To minimise bias, RCTs published in languages other than English were included in the search, but no relevant non-English language papers were identified
Date	2000 onwards If a study is an abstract only (for example, from a conference), it was only included if it was published in 2011 or onwards	Primary studies published prior to 2000, systematic literature reviews published before 2010 and conference abstracts published prior to 2011 were also excluded	Limiting the review to studies published from 2000 enabled us to focus on the latest trials evaluating the second-line treatment of NSCLC that reflect current clinical practice and patient populations. Conference abstracts were limited to those presented in 2011 onwards, as full text publications of earlier abstracts reporting on studies of a high quality would be expected to have been published. Systematic reviews were limited to those published in the previous 4 years, as these were used only to identify additional relevant primary research papers and therefore needed to be as up-to-date as possible.

6.7.2 Please follow the instructions specified in sections 6.1 to 6.5 for the identification, selection and methodology of the trials, quality assessment and the presentation of results. Provide in section 10.5, appendix 5, a complete quality assessment for each comparator RCT identified.

The efficacy review search was completed on 28 February, 2014 and yielded a total of 4,966 unique abstracts (see Figure 19). The abstracts were independently screened by two reviewers with any discrepancies independently resolved by a third reviewer, applying a set of predefined inclusion criteria as described below. Full-text publications of 334 abstracts deemed potentially relevant during this first level of review were then retrieved and reviewed independently by two researchers. Of these full-text articles, 61 primary studies—published in 67 articles—reported on second-line treatment of relapsed or refractory NSCLC of any histology. The citation lists of systematic literature reviews were also reviewed in order to identify relevant trials. A total of 16 additional articles were

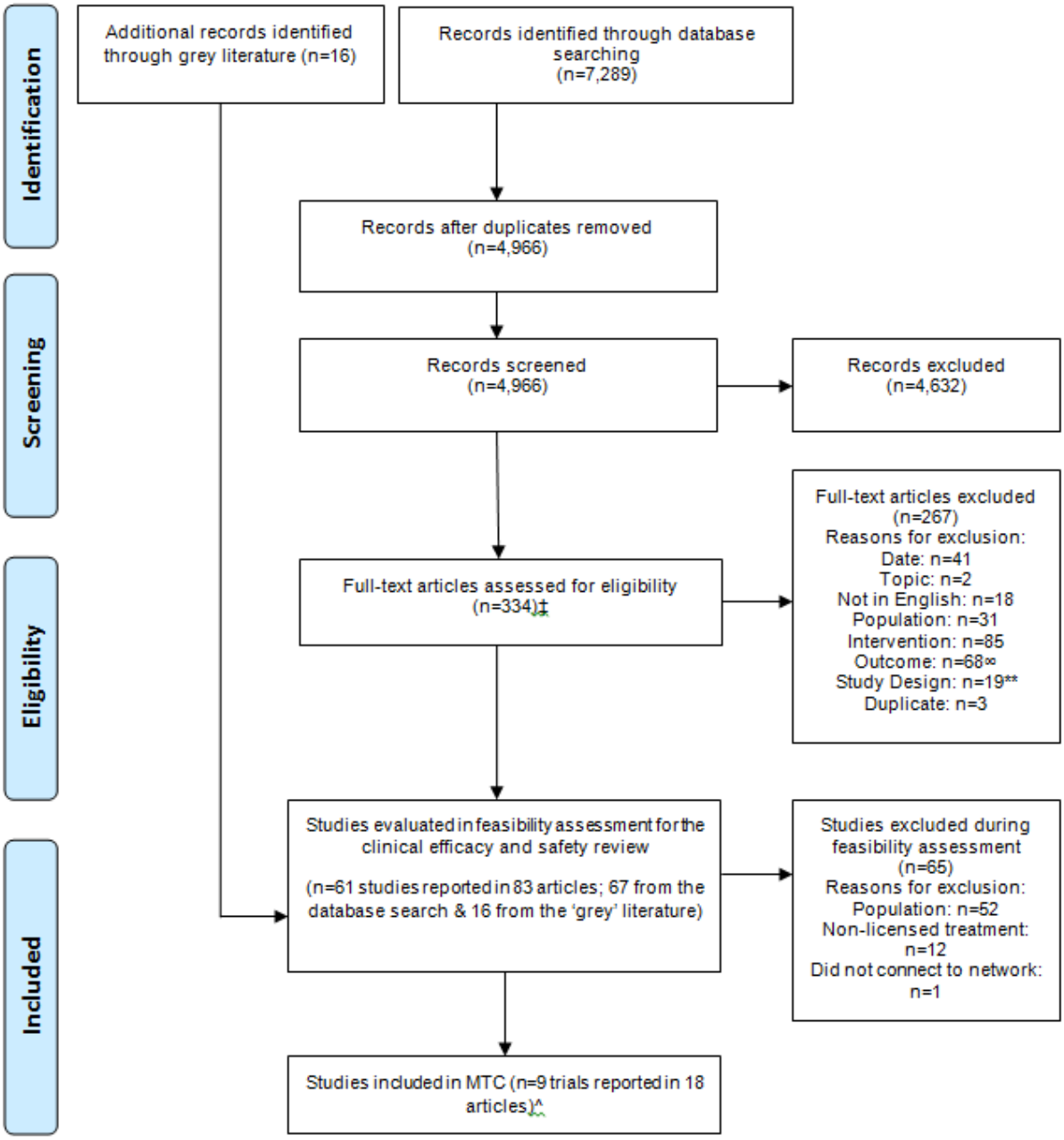
included upon evaluation of the bibliographies of existing systematic reviews published since 2011 and the search of 'grey' literature sites.

In order to select a population that is as close as possible to the histological indication for nintedanib, without unduly restricting the evidence base available for the MTC, we selected only those trials that contained at least 75% of participants with adenocarcinoma, or those that reported results specifically for the adenocarcinoma patient sub-group.

Of the studies that met the original criteria for the clinical efficacy and safety review, nine trials reported relevant outcomes for the second-line treatment of patients of whom at least 75% had adenocarcinoma and were treated with currently licensed drugs[2]. The results are presented as a PRIMSA flow diagram (Figure 19).

[2] While the combination of pemetrexed plus erlotinib is not currently approved, the individual treatments provided as monotherapy are licensed and so studies with this combination treatment could therefore be included in the MTC.

Figure 19: PRISMA flow diagram for systematic literature review on efficacy of second-line treatments for NSCLC (adenocarcinoma)



** The reference lists of the systematic reviews (n=4) were examined for any additional relevant studies; no additional studies were identified.

‡ Full-text articles were retrieved for the wider SLR, which included studies on humanistic and economic outcomes in addition to trials on clinical efficacy and safety.

∞ 25 studies were included in the economic review and 40 in the humanistic review.

^ Four trials were included in the base-case analysis, and four were added in a scenario analysis. As the studies in each network varied slightly, a total of nine studies were included across all analyses.

Data extraction was performed on all studies that met the final inclusion criteria and that could be joined in a network (see Table 25). All data were extracted by one researcher and validated by a second to ensure accuracy of data reporting. These data were analysed in fixed- and random-effects Bayesian MTC meta-analyses, as described below (see Section 10.4.8) Bucher indirect comparisons were also run wherever the data permitted, as a way of confirming the conclusions of the MTC.

Outcomes of interest were those reported for nintedanib plus docetaxel in the LUME-Lung 1 study, and included:

- OS
- PFS
- Objective response rate (ORR)
- AEs

An assessment of the quality of the studies was conducted using guidance from the Centre for Reviews and Dissemination (CRD) and can be found in Section 10.5.

6.7.3 *Provide a summary of the trials used to conduct the indirect comparison. A suggested format is presented below. Network diagrams may be an additional valuable form of presentation*.

Trials included in the base-case analyses

The base-case analysis analysed treatment comparators separately, with no pooling of treatments that could be considered standard chemotherapy. It also excluded studies that targeted patients with EGFR-TK mutations or that had a higher percentage than would be expected in the overall NSCLC adenocarcinoma population (>20% of patients with the mutation at baseline)[3].

The review identified the following four trials in patients treated with second-line pharmacotherapy for adenocarcinoma of the lung that were suitable for analysis in the base-case scenario. Studies that provided an active treatment arm with placebo versions of the comparator were not distinguished from other studies that did not provide a placebo.

1. Reck et al., 2013 (LUME-Lung 1)(3)
2. Hanna et al., 2004 (JMEI)(48, 53)
3. Garassino et al., 2013 (TAILOR)(76)

[3] Based on data from Gerber DE, Gandhi L and Costa DB. Management and future directions in non-small cell lung cancer with known activating mutations. Available at: http://meetinglibrary.asco.org/content/11400353-144

4. Li et al., 2014 (WSY001)(103)

Trials included in the scenario analyses

Docetaxel and pemetrexed considered to be equivalent efficacy

To ensure as many treatments as possible could be compared to nintedanib plus docetaxel, and to validate the conclusions of the MTC, we ran a scenario analysis. This based on the assumption that docetaxel and pemetrexed were considered as treatments with equivalent efficacy. The assumption was judged to be reasonable, given that these drugs were considered to be interchangeable in one of the studies identified by the review (TITAN)(46). In this trial, the comparator was "standard chemotherapy", which was a non-randomised choice of docetaxel or pemetrexed, selected at the physician's discretion. In this analysis, any treatment arm that was docetaxel, pemetrexed, or a non-randomised choice of either of these drugs, were pooled into one treatment group. As with the base-case network, studies that provided an active treatment arm with placebo versions of the comparator were not distinguished from other studies that did not provide a placebo.

Studies included in the scenario analysis that assumed equivalence of docetaxel and pemetrexed were as follows:

1. Reck et al., 2013 (LUME-Lung 1)(3)
2. Garassino et al., 2013 (TAILOR)(76)
3. Li et al., 2014 (WSY001)(103)
4. Ciuleanu et al., 2012 (TITAN)(46)

Trials included in sensitivity analyses

EGFR-TK mutation status

Among the trials identified in the review, there was notable variation in patient characteristics, particularly in regards to EGFR-TK mutation status, gender, and smoking history, as demonstrated in Table 26. It has been demonstrated in a recent meta-analysis(104) that the tyrosine kinase inhibitors (TKIs) gefitinib and erlotinib show an advantage among patients with EGFR-TK mutations and, conversely, that standard chemotherapy is superior to these TKIs in patients with EGFR-TK wild-type patients. The base-case analyses excluded studies that had specifically or indirectly selected patients with EGFR-TK mutations, such as by including patients selected on the basis of clinical characteristics associated with a higher prevalence of EGFR-TK mutations (e.g. patients who had never smoked). This was to allow a comparison between nintedanib plus docetaxel and other TKIs in a population close to that of LUME-Lung trial participants, where EGFR-TK mutation status was

unknown or likely to be predominantly wild-type. Of note, the base-case analyses therefore included studies that had selected only patients with wild-type EGFR-TK status. However, to minimise bias in the reporting of the results, and to validate the results of the base-case analyses, additional sensitivity analyses were run for both the base-case network and the scenario analysis assuming equivalent efficacy of docetaxel and pemetrexed, where additional studies that had selected for patients with EGFR-TK mutations were included. These were the following four trials:

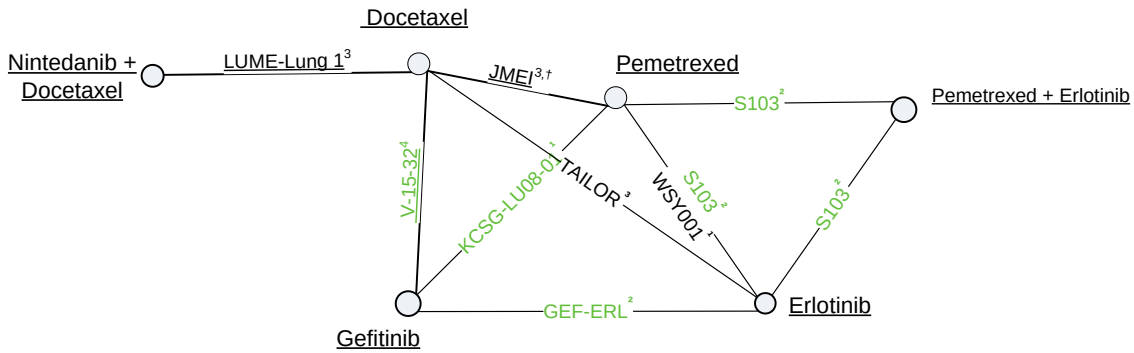
1. Sun et al., 2012 (KCSG-LU08-01)(105)
2. Maruyama et al., 2008 (V-15-32)(106)
3. Lee et al., 2013 (S103)(107)
4. Kim et al., 2012 (GEF-ERL)(108)

The KCSG-LU08-01 trial(105) compared gefitinib with pemetrexed in patients from South Korea with characteristics associated with higher incidence of EGFR-TK mutations (for example, all patients were required to be never-smokers). This was also true of the S103 trial(107), which also enrolled never-smokers, and in which 56% of a subgroup of 43 tested patients (22% of the study population) had EGFR-TK mutations. V-15-32(106) and GEF-ERL (108) were also eliminated from the base-case analyses. In V-15-32, 54% of a subgroup of 57 tested patients (12% of the study population) had an EGFR-TK mutation-positive status. GEF-ERL also recruited a large proportion of never-smokers, with 35% of a subgroup of 49 tested patients (18% of the study population) having EGFR-TK mutations.

Network diagrams

The network diagram presented below in Figure 20 represents the comparators considered in the base-case analyses. The network diagram for the scenario analyies where docetaxel and pemetrexed were assumed to have comparable efficacy is shown below in Figure 21. In each case, the analyses were run with and without the studies selectively including patients with EGFR-TK mutations, which are highlighted in green within each network diagram.

Figure 20: Base-case analysis diagram



[1] 100% adenocarcinoma

[2] >80% adenocarcinoma

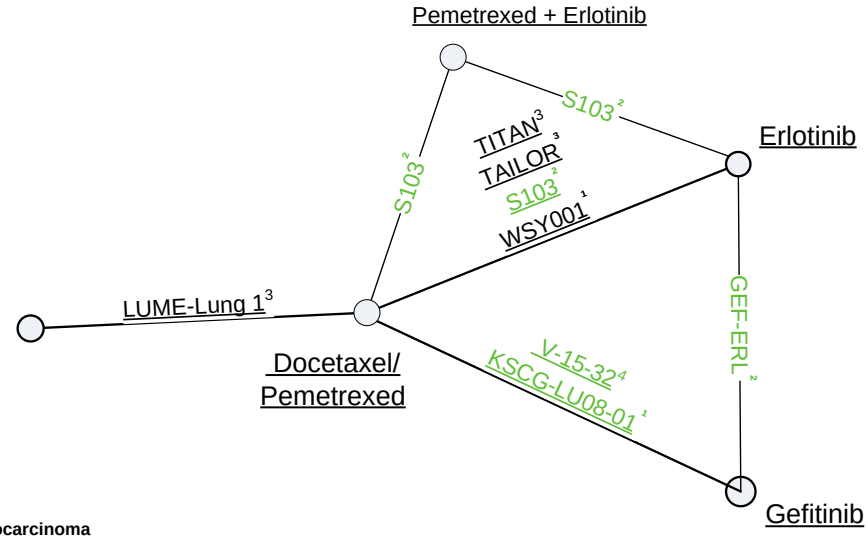
[3] Subgroup with 100% adenocarcinoma

[4] 78% adenocarcinoma

† Abstracts/grey literature only

Trial removed in base-case analysis due to high likelihood or prevalence of EGFR mutations

Figure 21: Scenario analysis diagram assuming equivalence of docetaxel and pemetrexed



[1] 100% adenocarcinoma

[2] >80% adenocarcinoma

[3] Subgroup with 100% adenocarcinoma

[4] 78% adenocarcinoma

† Abstracts/grey literature only

Trial removed in base-case analysis due to high likelihood or prevalence of EGFR mutations

The baseline characteristics of patients included in the base-case and scenario analyses are reported below in Table 26, and those included in the sensitivity analyses are reported in Table 27.

Table 26: Patient variation: trials included in the base-case and scenario analyses

Reference and location	Inclusion and exclusion criteria	Sample size	Proportion with adenocarcinoma	Age (years)	Gender (% female)	Prior mutations	Smoking history
Trials common to both the base-case and scenario analyses							
Reck et al., 2014 (LUME-Lung 1) Europe, Asia, South Africa	*Inclusion Criteria* • Histologically or cytologically confirmedstage IIIB-IV or recurrent NSCLC of any histology, following relapse or failure of one previous first-line chemotherapy (in the case of recurrent disease one additional previous regimen was allowed for adjuvant, neoadjuvant, or neoadjuvant + adjuvant therapy) • ECOG PS 0-1 *Exclusion Criteria* • Prior docetaxel or VEGF/VEGFR inhibitor (other than bevacizumab) usage • Radiographic evidence of cavitary or necrotic tumours, centrally located tumours with radiographic evidence (CT or MRI) of local invasion of major blood vessels, or a recent history (<3 months) of clinically significant haemoptysis or a major thrombotic or clinically relevant major bleeding event in the past 6 months	*Number Randomised & Evaluated at Baseline* **Overall:** 1,314 **Nintedanib + docetaxel:** 655 **Placebo + docetaxel**: 659	Subgroup with adenocarcinoma only **N:** 658 (50.1%) **Nintedanib + docetaxel:** 322 (49.2%) **Placebo + docetaxel:** 336 (51.0%)	**Nintedanib + docetaxel (overall population)** Median: 60 Range: 53-67 **Placebo + docetaxel (overall population):** Median: 60 Range: 54-66	**Nintedanib + docetaxel (overall population):** 27.3% (179/655) **Placebo + docetaxel (overall population):** 27.3% (180/659)	NR	*Never-smokers* **Nintedanib + docetaxel (overall population):** 25.2% (165/655) **Placebo + docetaxel (overall population):** 24.4% (161/659)
Garassino et al., 2013 (TAILOR) Italy	*Inclusion Criteria* • Patients with wild-type EGFR-TK advanced NSCLC, who had recurrence or progression after failing platinum-based chemotherapy • ECOG PS ≤2 *Exclusion Criteria* • Previous treatment with taxanes or	*Number Randomised & Evaluated at Baseline* **Overall:** 222 (219 included in ITT analysis) **Erlotinib:** 112 overall (109 included in ITT analysis) **Docetaxel:** 110 overall (110 included in ITT analysis)	Subgroup with adenocarcinoma only **N:** 152 (69.4%) **Erlotinib:** 69 (63.3%) **Docetaxel:** 83 (75.5%)	**Erlotinib (overall population):** Median: 66 Range: 40-81 **Docetaxel (overall population):** Median: 67 Range: 35-83	**Erlotinib (overall population):** 29.4% (32/109) **Docetaxel (overall population):** 33.6% (37/110)	*EGFR mutation* **Erlotinib:** Wild-type: 100% (109/109) **Docetaxel:** Wild-type: 100% (110/110)	*Never smokers* **Erlotinib (overall population):** 17% (19/109) **Docetaxel (overall population):** 27% (30/110)

Reference and location	Inclusion and exclusion criteria	Sample size	Proportion with adenocarcinoma	Age (years)	Gender (% female)	Prior mutations	Smoking history
	anti-EGFR drugs						
Li et al., 2014 (WSY001) China	*Inclusion Criteria* • Upper age cut-off of 75 years • Pathologically or cytologically confirmed stage IIIB or IV lung adenocarcinoma or postoperative recurrent lung adenocarcinoma incurable by surgery or radiotherapy • within 6 months of neoadjuvant or adjuvant chemotherapy • ECOG PS 0-2 • EGFR wild-type and EGFR-TK FISH-positive disease *Exclusion Criteria* • Prior treatment with EGFR-TK TKIs or pemetrexed • symptomatic brain metastases	*Number Randomised & Evaluated at Baseline* **Overall:** 123 **Erlotinib:** 61 **Pemetrexed:** 62	100% adenocarcinoma patients	**Erlotinib:** Median: 54.3 Range: 30–74 **Pemetrexed:** Median: 55.1 Range: 33–75	**Erlotinib (overall population):** 34.4% (21/61) **Pemetrexed (overall population):** 37.1% (23/62)	*EGFR mutation* **Erlotinib:** Wild-type: 100% (61/61) **Pemetrexed:** Wild-type: 100% (62/62)	*Never-smokers* **Erlotinib:** 24.6% (15/61) **Pemetrexed:** 27.4% (17/62)
Trial in only the base-case analysis							
Hanna et al., 2004 (JMEI); Scagliotti et al., 2009 NR	*Inclusion Criteria* • Histologically or cytologically confirmed stage III or IV NSCLC not amendable to curative therapy • Received treatment with only one prior chemotherapy for advanced disease (one prior additional therapy allowed for neoadjuvant, adjuvant, or neoadjuvant + adjuvant therapy) • ECOG PS 0-2 *Exclusion Criteria* • Patients with prior docetaxel or pemetrexed treatment • CTC ≥grade 3 peripheral neuropathy • An inability to interrupt nonsteroidal anti-inflammatory drugs • Uncontrolled pleural effusions,	*Number Randomised & Evaluated at Baseline* **Overall:** 571 **Pemetrexed:** 283 **Docetaxel:** 288	Subgroup with adenocarcinoma only **N:** 302 (52.9%) **Pemetrexed:** 158 (55.8%) **Docetaxel:** 144 (50%)	**Pemetrexed**: Median: 57.4 years (adenocarcinoma subgroup) Range: NR **Docetaxel:** Median: 56.7 years (adenocarcinoma subgroup) Range: NR	**Pemetrexed (adenocarcinoma subgroup):** 39% (62/158) **Docetaxel (adenocarcinoma subgroup):** 34% (49/144)	NR	NR

Reference and location	Inclusion and exclusion criteria	Sample size	Proportion with adenocarcinoma	Age (years)	Gender (% female)	Prior mutations	Smoking history
	symptomatic or uncontrolled brain metastases, or significant weight loss (≥ 10% body weight in the preceding 6 weeks) were ineligible.						
Trial in only the scenario analysis							
Ciuleanu et al., 2012 (TITAN) International	*Inclusion Criteria* • Histologically documented locally advanced, recurrent, or metastatic NSCLC • Disease progression while receiving four cycles of a standard first-line platinum-based chemotherapy doublet (representing a population with poor prognosis) • ECOG PS 0-2 *Exclusion Criteria* • Previous exposure to anti-human-EGFR-directed drugs or drugs directed at pemetrexed molecular targets (i.e., thymidylate synthase and dihydrofolate reductase inhibitors) • Prior chemotherapy or systemic anti-neoplastic therapy other than the permitted platinum-based regimens • Uncontrolled or untreated brain metastasis • Spinal cord compression or other malignancies within the past 5 years (except carcinoma in situ)	*Number Randomised & Evaluated at Baseline* **Overall:** 424 **Erlotinib:** 203 **Chemotherapy:** 221	Subgroup with adenocarcinoma only **N:** 210 (49.5%) **Erlotinib:** 96 (47.3%) **Chemotherapy:** 114 (51.6%)	**Erlotinib (overall population):** Median: 59 years Range: 36–80 years **Chemotherapy (overall population):** Median: 59 years Range: 22–79 years	**Erlotinib (overall population):** 20.7% (42/203) **Chemotherapy (overall population):** 27.6% (61/221)	*EGFR mutation* **Erlotinib:** Activating mutation: 3.4% (7/203) Other mutation (including resistance mutation): <1% (1/203) Wild-type: 36.9% (75/203) Indeterminate: 15.8% (32/203) Missing: 43.3% (88/203) **Chemotherapy:** Activating mutation: 1.8% (4/221) Other mutation (including resistance mutation): 2.7% (6/221)	*Never-smokers* **Erlotinib:** 14.8% (30/203) **Standard chemotherapy (docetaxel or pemetrexed):** 19.9% (44/221)

Reference and location	Inclusion and exclusion criteria	Sample size	Proportion with adenocarcinoma	Age (years)	Gender (% female)	Prior mutations	Smoking history
						Wild-type: 33.5% (74/221) Indeterminate: 16.3% (36/221) Missing: 45.7% (101/221)	

Table 27: Patient variation: trials included in sensitivity analyses

Reference and location	Inclusion and exclusion criteria	Sample size	Proportion with adenocarcinoma	Age (years)	Gender (% female)	Prior mutations	Smoking history
Kim et al., 2012 (GEF-ERL) South Korea	*Inclusion Criteria* • Histologically confirmed stage IIIb or IV NSCLC including recurrent or metastatic disease following failure of first-line chemotherapy • WHO performance status of 0-2 • Presence of either an activating EGFR-TK mutation, or two of three clinical factors associated with higher incidence of EGFR-TK mutations. • Brain metastasis permitted if treated at least 4 weeks before entry and clinically stable without steroid treatment for 1 week *Exclusion Criteria* • Previous treatment with EGFR-TK signalling inhibitors and radiation therapy within the preceding 4 weeks	*Number Randomised & Evaluated at Baseline* **Overall:** 96 **Gefitinib:** 48 **Erlotinib:** 48	75% or more with adenocarcinoma **N:** 87 (90.6%) **Gefitinib**: 44 (91.7%) **Erlotinib**: 43 (89.6%)	**Gefitinib (overall population):** Median: 60 Range: 37–83 **Erlotinib (overall population):** Median: 56 Range: 32–81	**Gefitinib (overall population):** 85.4% (41/48) **Erlotinib (overall population):** 85.4% (41/48)	*EGFR activating mutation* **Gefitinib:** 42.9% (9/21) **Erlotinib:** 28.6% (8/28)	*Never-smokers* **Gefitinib:** 91.7% (44/48) **Erlotinib:** 95.8% (46/48)

Reference and location	Inclusion and exclusion criteria	Sample size	Proportion with adenocarcinoma	Age (years)	Gender (% female)	Prior mutations	Smoking history
Sun et al., 2012 (KCSG-LU08-01) Korea	*Inclusion Criteria* • Histologically or cytologically confirmed pulmonary adenocarcinoma that progressed after just 1 previous platinum-based chemotherapy regimen for advanced disease (stage NR) • Never-smoked (a total of ≤100 cigarettes in their lifetime) • ECOG PS 0-2 *Exclusion Criteria* • Patients with prior EGFR-TK TKI or pemetrexed treatment • Symptomatic or uncontrolled brain metastases were ineligible.	*Number Randomised* **Overall:** 141 **Gefitinib:** 71 **Pemetrexed:** 70 *Number treated, and analysed for efficacy and safety* **Overall:** 135 **Gefitinib:** 68 **Pemetrexed:** 67	100% adenocarcinoma patients	**Gefitinib:** Median: 58 Range: 40-77 **Pemetrexed**: Median: 64 Range: 30-78	**Gefitinib:** 85.3% (58/68) **Pemetrexed:** 85.1% (57/67)	*EGFR mutation* **Gefitinib:** Activating mutation: 23.5% (16/68) Other mutation: 4.4% (3/68) Wild-type: 22.1% (15/68) Unknown mutation status: 50% (34/68) **Pemetrexed:** Activating mutation: 25.4% (17/67) Other mutation: 6.0% (4/67) Wild-type: 23.9% (16/67) Unknown mutation status: 44.8% (30/67)	*Never-smokers* **Gefitinib:** 100% (68/68) **Pemetrexed:** 100% (67/67)
Lee et al., 2013 (S103) NR	*Inclusion Criteria* • Histologically or cytologically confirmed, locally advanced or metastatic non-squamous NSCLC following failure of first-line chemotherapy regimen • ECOG PS 0-2 • Only never-smoking patients	*Number Randomised & Evaluated at Baseline* **Overall:** 247 (240 non-squamous – Q-ITT populaton) **Erlotinib + pemetrexed:** 78 **Erlotinib:** 82	75% or more with adenocarcinoma **N:** 225 (93.8%) **Erlotinib + pemetrexed:** 72 (92.3%) **Erlotinib:** 76 (92.7%)	**Erlotinib + pemetrexed (overall population):** Median: 55.8 Range: NR **Erlotinib**	**Erlotinib + pemetrexed (overall population:** 74.4% (58/78) **Erlotinib (overall population):** 65.9% (54/82)	*EGFR mutation*: Mutant: 55.8% (24/43) Wild-type: 44.2% (19/43)	*Never smokers* **Erlotinib + pemetrexed:** 100% (78/78) **Erlotinib:** 100% (82/82)

Reference and location	Inclusion and exclusion criteria	Sample size	Proportion with adenocarcinoma	Age (years)	Gender (% female)	Prior mutations	Smoking history
	(<100 lifetime cigarettes) were eligible. *Exclusion Criteria* • Prior exposure to agents directed at the human EGFR-TK axis or at pemetrexed molecular targets (e.g. TS or DHFR inhibitors) • Brain metastasis (unless treated and stable after radiotherapy ≥2 weeks) • Concurrent administration of any other antitumour therapy.	**Pemetrexed:** 80	**Pemetrexed:** 77 (96.3%)	**(overall population):** Median: 53.9 Range: NR **Pemetrexed (overall population):** Median: 55.9 Range: NR	**Pemetrexed (overall population):** 56.3% (45/80)		**Pemetrexed:** 100% (80/80)
Maruyama et al., 2008 (V-15-32) Japan	*Inclusion Criteria* • Histologically or cytologically confirmed stage IIIB or IV NSCLC not amenable to curative surgery or radiotherapy, or postoperative recurrent NSCLC • Failure of prior treatment with one or two chemotherapy regimens (≥1 platinum-based regimen) • WHO PS 0 to 2 • Protocol amendment allowed recruitment of patients without measurable lesions *Exclusion Criteria* • NR	*Number Randomised & Evaluated at Baseline (2nd line only: 84.4% of the total study population)* **Overall:** 413 **Gefitinib:** 212 **Docetaxel:** 201	75% or more with adenocarcinoma **N:** 380 (92.0%) **Gefitinib:** 192 (78.4%) **Pemetrexed/Docetaxel:** 188 (77.0%)	*NR specifically for second-line subgroup* **Gefitinib (overall population):** ≤64 years: 138 (56.3%) **Docetaxel (overall population):** ≤64 years: 135 (55.3%)	*NR specifically for second-line subgroup.* **Gefitinib (overall population):** 38.4% (94/245) **Docetaxel (overall population):** 38.1% (93/244) **Gefitinib:**	*EGFR mutation NR specifically for 2nd line sub-group.* Mutant (overall population): 54.4% (31/57)	*Never-smokers NR specifically for second-line subgroup.* **Gefitinib (overall population):** 29.0% (71/245) **Docetaxel (overall population):** 35.7% (87/244)

Baseline severity of NSCLC: all included trials

The clinical trials identified in our systematic literature review had similar inclusion criteria and, therefore similar baseline characteristics for Eastern Cooperative Oncology Group (ECOG) performance status. Most studies had an inclusion criterion of an ECOG performance status of 0 to 2, with one trial reporting a status of 0 to 1 (LUME-Lung 1(1)), and two studies using a World Health Organisation (WHO) performance status of 0-2 to select patients, both only included in the sensitivity analysis (GEF-ERL(108); V-15-32(106)). At baseline, the proportion of patients in the base-case analyses who had an ECOG performance status of 2 ranged between 0% and 9%, with two trials that were only included in the sensitivity analyses (GEF-ERL(108) and TITAN(46)) reporting a higher proportion of patients with higher ECOG grades and therefore more severe disease (14.6% and 21%, respectively).

Interventions

The interventions and comparators included in trials analysed in the base-case MTC and sensitivity analyses are listed in Table 28. The treatment regimens used among the trials were similar with the exception of two trials. V-15-32(106) varied from the other studies on docetaxel as it was administered every three weeks as a one-hour intravenous infusion of 60 mg/m^2 (the approved dosage in Japan). In addition, in the TAILOR study(76) there was also the option of administering docetaxel as a one-hour infusion of 35 mg/m^2 on days 1, 8, and 15 of a 28-day cycle.

Table 28: Treatment and comparator arms among all included trials

Treatment*	Number of trials	Trial names	Dosage
Nintedanib + docetaxel	1	**LUME-Lung 1(3)**	Docetaxel 75 mg/m^2 IV on day 1 + nintedanib 200 mg twice daily orally on days 2-21
Pemetrexed + erlotinib	1	S103(107)	Pemetrexed 500 mg/m^2 IV on day 1 + erlotinib 150 mg per day orally on days 2-14 of a 21 day cycle
Pemetrexed	4	KCSG-LU08-01(105) **JMEI(48, 53)** **WSY001(103)** S103(107)	Pemetrexed 500 mg/m^2 IV on day 1 of a 21-day cycle Pemetrexed 500 mg/m^2 IV on day 1 of a 21-day cycle Pemetrexed 500 mg/m^2 IV on day 1 of a 21-day cycle Pemetrexed 500 mg/m^2 IV on day 1 of a 21-day cycle
Docetaxel	4	**LUME-Lung 1(3)** V-15-32(106) **JMEI(48, 53)** **TAILOR(76)**	Docetaxel 75 mg/m^2 IV on day 1 of a 21-day cycle Docetaxel 60 mg/m^2 IV once every 21 days Docetaxel 75 mg/m^2 IV on day 1 of a 21-day cycle Docetaxel 75 mg/m^2 IV once every 21 days, or docetaxel 35 mg/m^2 IV on days 1, 8 and 15, every 28 days
Standard chemotherapy (docetaxel or pemetrexed)	1	TITAN(46)	Standard docetaxel or pemetrexed dosing schedule at investigator's discretion
Erlotinib	5	TITAN(46) GEF-ERL(108) **WSY001(103)** S103(107) **TAILOR(76)**	Erlotinib 150 mg per day orally Erlotinib 150 mg per day orally Erlotinib 150 mg per day orally Erlotinib 150 mg per day orally Erlotinib 150 mg per day orally
Gefitinib	3	GEF-ERL(108) KCSG-LU08-01(105) V-15-32(106)	Gefitinib 250 mg per day orally on a 28-day cycle Gefitinib 250 mg per day orally on a 21-day cycle Gefitinib 250 mg per day orally

*We have made no distinction between treatments that were provided with or without placebo.

Trials in the base-case analyses are shown in **bold**.

As noted in the table above, in the base-case analyses, one study compared an active intervention to pemetrexed alone (WSY001(103)); two studies compared an active intervention with docetaxel (LUME-Lung 1(3); TAILOR(76)), and one study compared docetaxel with pemetrexed (JMEI trial(48, 53)).

In the scenario analyses, one trial (TITAN(46)) randomised patients to a 'standard chemotherapy' treatment arm, where patients received either docetaxel or pemetrexed, selected by the physician

as a non-randomised treatment choice. A further two studies compared an active intervention to pemetrexed alone (KCSG-LU08-01(105); S103(107)); and one study compared an active intervention with docetaxel (V-15-32(106)).

Three trials, one in the scenario analysis and two in the sensitivity analyses, explicitly stated that patients were permitted to cross over to the comparator treatment. These are summarised below in Table 29.

Table 29: Information on trials that allowed for crossover

Trial	Cross-over percentage
Scenario analysis	
TITAN(46)	7% of docetaxel-treated patients and 5% of pemetrexed-treated patients crossed over to the alternative standard chemotherapy arm. Cross-over rates were not reported for patients randomised to erlotinib
Sensitivity analysis	
KCSG-LU08-01(105)	69.8% of gefitinib-treated patients and 65.1% of pemetrexed-treated patients crossed over to the comparator treatment arm
V-15-32(106)	36% of gefitinib-treated patients and 53% of docetaxel-treated patients crossed over to the comparator treatment arm

Any additional treatments used

Only one trial, reported only in the sensitivity analyses, explicitly reported the use of concomitant treatment (KCSG-LU08-01(105)). Patients on the pemetrexed arm received oral folic acid (1 mg) daily and a vitamin B12 injection (1000 µg) every nine weeks, beginning one week before the first dose and continuing until three weeks after the last dose of study treatment. Patients on the pemetrexed arm were also prescribed dexamethasone (4 mg orally twice daily the day before, the day of, and the day after pemetrexed) as a prophylactic measure against skin rash.

Outcomes measured

The trials that report on each outcome of interest are represented in Table 30.

Table 30: Outcomes reported among the included trials

Trial	OS	PFS	ORR	Any grade AE: fatigue	Any grade AE: nausea	Any grade AE: diarrhoea	Grade 3+ AE: fatigue	Grade 3+ AE: nausea
Included in base-case analysis								
LUME-Lung 1 (3)	X‡,1	X‡,1	X‡	X‡	X‡	X‡	X‡	X‡
JMEI (48, 53)	X‡,2	X‡,2	X*‡,2					
WSY001 (12)	X	X	X	X	X	X	X#	X#
TAILOR (76)	X‡	X‡						
Included in scenario analyses								
TITAN (46)	X†‡							
Included in sensitivity analyses								
KCSG-LU08-01 (105)	X	X	X	X	X	X	X#	X#
V-15-32 (106)	X3	X3						
GEF-ERL (108)	X	X	X	X#	X	X	X#	X#
S103 (107)	X	X	X					

X: This outcome was reported for the study
1: For the LUME-Lung 1 trial adjusted OS, PFS, and ORR data for the adenocarcinoma subgroup are available from Clinical Trial Report No.1199.13
2: For the JMEI trial OS, PFS, and ORR data for the adenocarcinoma subgroup are available in the related Scagliotti et al. 2009 publicaion
3: The V-15-32 trial enrolled patients eligible for second- or third-line treatment and outcomes are reported for the group as a whole. Since >80% of patients in each treatment arm received second-line treatment, we have included this publication in the MTC.
† TITAN was only applicable to network 1.
#: One or more treatment arms had zero event rates, so the trial cannot be analysed for this outcome
‡: Only subgroup data for adenocarcinoma group was analysed

The outcomes for nintedanib plus docetaxel that could be compared with other treatments in the base-case analysis and tested in the sensitivity analyses were as follows:

- OS, months (HR; 95% CI)
- PFS, months (HR; 95% CI)
- ORR (number of patients)

Safety outcomes were only reported in a consistent format in more than one trial for fatigue, nausea and diarrhoea. However, because of the small number of trials reporting these outcomes, and because of low event rates in those trials that did report these outcomes in an equivalent way, it was only possible to compare nintedanib plus docetaxel with other treatments using the sensitivity analysis that assumed equivalent efficacy of docetaxel and pemetrexed. In the base-case analysis,

the LUME-Lung 1 trial did not connect with the other trials reporting these safety data (KCSG-LU08-01 and WSY001, Table 30).

Unadjusted data was preferentially analysed in the MTC, as it was more commonly reported; however, adjusted data was used whenever unadjusted data was unavailable. If HRs were not reported, but Kaplan-Meier curves for OS and/or PFS were presented in the publications, data were extracted to calculate HRs using the Parmar method(109). In this methodology, the log HR was estimated for each non-overlapping interval from the HR extracted from Kaplan-Meier curves and combined in a stratified way across intervals to obtain an overall log HR for each trial.

Time points and follow-up durations

Specific time points were not evaluated, as outcomes at any study endpoint were considered. Outcomes analysed in the MTC were evaluated at the end of the study, and end-of-study relative effects were assumed to be independent of follow-up time. Median follow-up time ranged from 7.5 to 33 months (Table 31).

Table 31: Median follow-up duration of included trials

Trial name	Median follow-up duration
GEF-ERL(108)	16.3 months
JMEI(48, 53)	7.5 months
KCSG-LU08-01(105)	15.9 months
LUME-Lung 1(3)	31.7 months
S103(107)	Median not reported, scheduled follow-up was for 18 months
TAILOR(76)	33 months
TITAN(46)	24.8 months (chemotherapy arm) - 27.9 months (erlotinib arm)
V-15-32(106)	21 months
WSY001(103)	14.7 months

Methodology

All included studies randomised patients to treatments arms, with allocation concealment clearly reported and care providers, participants and assessors blinded to treatment allocation in only two of the four trials, both in the base-case analysis (LUME-Lung 1 and TAILOR).

All the trials had patients with similar prognostic factors at the start of the trial. One trial (KCSG-LU08) contained unexpected imbalances in dropout rates between groups and it was unclear

whether this had occurred in two additional trials (GEF-ERL and LUME-Lung 1). Details of the study methodology are provided in Section 10.5.

6.7.4 For the selected trials, provide a summary of the data used in the analysis.

The results of the trials included in the MTC are outlined below, with efficacy results for the trials in the base-case and scenario analyses in Table 32 and those in the sensitivity analyses Table 33, and safety results for the base-case and scenario analyses in Table 34 and the sensitivity analysis trials in Table 35.

Table 32: Efficacy results: Trials included in base-case and scenario analyses

Analysis		Base-case only		Included in both base-case and scenario analysis						Scenario analysis only	
Outcomes		**JMEI¥**		**LUME-Lung 1‡**		**WSY001**		**TAILOR**		**TITAN**	
Treatment arm		**PEM**	**DOC**	**NIN + DOC**	**DOC + PBO**	**ERL**	**ERL**	**ERL**	**PEM**	**ERL**	**DOC/PEM**
N efficacy		158	144	322	336	61	96	96	83	96	114
Unadjusted OS	HR	NR		0.83		1.01		0.67		0.95	
	95% CI or p-value	NR		(0.7, 0.99);p= 0.0359		(0.66, 1.54);p= 0.97		(0.48, 0.95); reported as significant		(0.7, 1.29);p= NR	
Adjusted OS	HR	0.92		0.81		NR		NR		NR	
	95% CI or p-value	(0.69, 1.22); p= 0.551		(0.69, 0.97);p= 0.0186 (two-sided)		NR		NR		NR	
	Variables adjusted for	NR		ECOG PS at baseline, prior bevacizumab treatment, presence of brain metastases at baseline		NR		NR		NR	
Unadjusted PFS	HR	NR		0.77		0.92		0.76		NR	
	95% CI or p-value	NR		(0.62, 0.96);p= 0.0193		(0.62, 1.37); p= 0.683		(0.54, 1.05); p= NR		NR	
Adjusted PFS	HR	0.83		0.84		NR		NR		NR	
	95% CI or p-value	(0.65, 1.06); p= 0.135		(0.71, 1); p= 0.0485 (two-sided)		NR		NR		NR	
	Variables adjusted for	NR		NR		NR		NR		NR	
Response	**Criteria**	Southwest Oncology Group Criteria		RECIST		RECIST		NR		RECIST	
Objective response	**Definition**	CR, PR*		Objective tumour response (CR+PR)		PR+CR		NR		Overall response	
ORR	**N evaluated**	158	144	322	336	61	62	NR	NR	NR	NR
	N			15	12	12	5	NR	NR	NR	NR
	%	12.8	9.9	4.7	3.6	19.7	8.1	NR	NR	NR	NR

‡: For the LUME-Lung 1 trial adjusted OS, PFS, and ORR data for the adenocarcinoma subgroup are available from Clinical Trial Report No.1199.13

¥: For the JMEI trial OS, PFS, and ORR data for the adenocarcinoma subgroup are available in the related Scagliotti et al. 2009 publication(53)

* Complete response: complete disappearance of all measurable and evaluable disease; Partial response: ≥50% decrease in the sum of products of perpendicular diameters of all measurable lesions

Table 33: Efficacy results: trials included in sensitivity analyses

Outcomes		**GEF-ERL**		**KCSG-LU08-01**		**V-15-32**		**S103**		
Treatment arm		**GEF**	**ERL**	**GEF**	**PEM+PBO**	**GEF**	**DOC**	**ERL+PEM**	**ERL**	**PEM**
N efficacy		48	48	68	67	212*	201*	78	82	80
Unadjusted OS	HR	0.47#		0.8		1.12		ERL+PEM vs ERL:1.08 ERL+PEM vs PEM: 0.75 ERL vs PEM: 1.44		
	95% CI or p-value	(0.22, 0.99)#		(0.5, 1.3);p= 0.37		(0.89, 1.40);p= 0.330		ERL+PEM vs ERL: (0.69, 1.67); p= 0.747 ERL+PEM vs PEM: (0.49, 1.13);p= 0.168 ERL vs PEM: (0.94, 2.21); p= 0.094		
Adjusted OS	HR	NR		0.83		1.01		NR	NR	NR
	95% CI or p-value	NR		(0.5, 1.38); p= NR		(0.80, 1.27); 0.914		NR	NR	NR
	Variables adjusted for	NR		NR		Gender, ECOG PS, tumour type, smoking history, prior chemotherapy regimen, age		NR	NR	NR
Unadjusted PFS	HR	1.17#		0.54		0.9		ERL+PEM vs ERL: 0.57 ERL+PEM vs PEM: 0.58 ERL vs PEM: 0.99		
	95% CI or p-value	(0.81, 1.7)#		(0.37, 0.79); p= 0.0006		(0.72, 1.12); p= 0.335		ERL+PEM vs ERL: (0.4, 0.81); p= 0.002 ERL+PEM vs PEM: (0.39, 0.85); p= 0.005 ERL vs PEM: (0.70, 1.40); p= 0.959		
Adjusted PFS	HR	NR		0.53		0.81		NR	NR	NR
	95% CI or p-value	NR		(0.36, 0.80); p= NR		(0.65, 1.02); p= 0.077		NR	NR	NR
	Variables adjusted for	NR		Age, sex, ECOG PS		Gender, ECOG PS, tumour type, smoking history, prior chemotherapy regimen, age		NR	NR	NR
Response	**Criteria**	RECIST		RECIST		NR		RECIST		

Outcomes		GEF-ERL		KCSG-LU08-01		V-15-32		S103		
Treatment arm		GEF	ERL	GEF	PEM+PBO	GEF	DOC	ERL+PEM	ERL	PEM
Objective response	Definition	Overall response rate		O (not further specified)		NR		CR + PR		
ORR	N Evaluated	48	48	68	67	NR	NR	78	82	80
	N	23	19	NR	NR	NR	NR	34	24	8
	%	47.9	39.6	45.6	28.4	NR	NR	44.7	29.3	10

*The V-15-32 trial enrolled patients eligible for second- or third-line treatment, and outcomes are reported only for the combined treatment lines. Since >80% of patients in each treatment arm were eligible for second-line treatment, we have included this publication in the MTC. The N values reported are for the second-line population alone.

HR and 95%CI data from GEF-ERL were derived from Kaplan-Meier charts in the primary publication using the Parmar method.

Table 34: Safety results: trials included in base-case and scenario analyses

Analysis		Base-case only		Included in both base-case and scenario analysis						Scenario analysis only	
Outcomes		JMEI		LUME-Lung 1		WSY001		TAILOR		TITAN	
Treatment arm		PEM	DOC	NIN+DOC	DOC+PBO	ERL	PEM	ERL	DOC	ERL	ERL+PEM
N randomised		283	288	655	659	61	62	69	83	203	221
N evaluated for safety		158	144	320	333	61	62	NR	NR	NR	NR
Any grade AE: fatigue	N	NR	NR	99	98	12	16	NR	NR	NR	NR
	%	NR	NR	30.9	29.4	19.7	25.8	NR	NR	NR	NR
Any grade AE: nausea	N	NR	NR	91	59	1	15	NR	NR	NR	NR
	%	NR	NR	28.4	17.7	1.6	24.2	NR	NR	NR	NR
Any grade AE: diarrhoea	N	NR	NR	139	82	10	2	NR	NR	NR	NR
	%	NR	NR	43.4	24.6	16.4	3.2	NR	NR	NR	NR
Grade 3+ fatigue	N	NR	NR	15	14	0	0	NR	NR	NR	NR
	%	NR	NR	4.7	4.2	0	0	NR	NR	NR	NR
Grade 3+ nausea	N	NR	NR	3	2	0	2	NR	NR	NR	NR
	%	NR	NR	0.9	0.6	0	3.2	NR	NR	NR	NR

Table 35: Safety results: trials included in sensitivity analyses

Outcomes		GEF-ERL		KCSG-LU08-01		V-15-32		S103		
Treatment arm		GEF	ERL	GEF	PEM+PBO	ERL	PEM	ERL+PEM	ERL	PEM
N randomised		48	48	71	70	212	201	78	82	80
N evaluated for safety		48	48	68	67	NR	NR	75	82	76
Any grade AE: Fatigue	N	0	8	15	14	NR	NR	NR	NR	NR
	%	0	16.7	22.1	20.9	NR	NR	NR	NR	NR
Any grade AE: nausea	N	3	2	11	11	NR	NR	NR	NR	NR
	%	6.3	4.2	16.2	16.4	NR	NR	NR	NR	NR
Any grade AE: diarrhoea	N	16	17	18	3	NR	NR	NR	NR	NR
	%	33.4	35.5	26.5	4.5	NR	NR	NR	NR	NR
Grade 3+ fatigue	N	0	0	0	0	NR	NR	NR	NR	NR
	%	0	0	0	0	NR	NR	NR	NR	NR
Grade 3+ nausea	N	0	0	0	0	NR	NR	NR	NR	NR
	%	0	0	0	0	NR	NR	NR	NR	NR

*The V-15-32 trial enrolled patients eligible for second- or third-line treatment, and outcomes are reported only for the combined treatment lines. Since >80% of patients in each treatment arm were eligible for second-line treatment, we have included this publication in the MTC. The N values reported are for the second-line population alone.

6.7.5 Please provide a clear description of the indirect/mixed treatment comparison methodology. Supply any programming language in a separate appendix.

The data were analysed in fixed- and random-effects Bayesian MTC meta-analyses using OpenBUGS and Bucher indirect treatment comparison, as described below.

Unadjusted data was preferentially analysed in the MTC over adjusted data, as it was more commonly reported; however, adjusted data was used whenever unadjusted data was unavailable. Outcomes analysed in the MTC were evaluated at the end-of-study, and end-of-study relative effects were assumed to be independent of follow-up time (of note: median follow-up time ranged from 7.5 to 33 months).

The statistical approach is provided in Section 10.4.8, while the programming language is in Section 10.4.9.

6.7.6 *Please present the results of the analysis*.

The base-case MTC results show that nintedanib plus docetaxel is statistically superior to docetaxel and erlotinib and has a non-statistically significant advantage over pemetrexed for improving both OS and PFS using a Bayesian MTC fixed-effects model. There were no significant differences between treatments for OR. For the base-case analyses, nintedanib plus docetaxel was most likely to be the best treatment for improving OS and PFS.

In the base case sensitivity analysis, when studies that selected patients with EGFR-TK mutations are added to the network, nintedanib plus docetaxel is significantly more effective than docetaxel, erlotinib and gefitinib for prolonging OS and is significantly more effective than docetaxel, pemetrexed and erlotinib for prolonging PFS. However, the wide credible intervals mean that the Bayesian MTC is no longer statistically significant using a random-effects model.

Scenario analyses assuming the equivalence of docetaxel and pemetrexed revealed that nintedanib plus docetaxel is statistically superior to docetaxel/pemetrexed and erlotinib at

improving both OS and PFS using a Bayesian MTC fixed-effects model, although this was not found for the random-effects model due to the wide credible intervals.

XThese results are summarised below in Table 36 and Table 37 and presented in more detail in the following sections.

XTable 36: Summary of results for OS

OS	Base-case analysis	Scenario analysis assuming equivalence of docetaxel and pemetrexed
	Nintedanib + docetaxel vs	
Base-case analysis	**> Docetaxel** **> Erlotinib** **≥ Pemetrexed**	> Docetaxel/pemetrexed > Erlotinib
Sensitivity analyses with addition of trials selecting patients with EGFR-TK mutations	> Docetaxel ≥ Pemetrexed > Erlotinib >Gefitinib ~ Erlotinib + pemetrexed	> Docetaxel/pemetrexed ≥ Erlotinib > Gefitinib ≤ Erlotinib + pemetrexed

Key:

> indicates that nintedanib plus docetaxel showed a statistically significant advantage to comparator;

≥ indicates an advantage that was not statistically significant;

~ indicates that the comparison was non-significant and very close to 1.0 (0.85 to 1.18) suggesting similarity;

≤ indicates that nintedanib plus docetaxel demonstrated a trend for disadvantage to a comparator that was not statistically significant.

All comparisons are for fixed-effects models.

Table 37: Summary of results for PFS

PFS	Base-case analysis	Scenario analysis assuming equivalence of docetaxel and pemetrexed
	Nintedanib + docetaxel vs	
Base-case analysis	**> Docetaxel** **> Erlotinib** **≥ Pemetrexed**	> Docetaxel/pemetrexed > Erlotinib
Sensitivity analyses with addition of trials selecting patients with EGFR-TK mutations	> Docetaxel > Pemetrexed > Erlotinib ~ Gefitinib ≤ Erlotinib + pemetrexed	> Docetaxel/pemetrexed > Erlotinib ~ Gefitinib ≤ Erlotinib + pemetrexed

Key:

> indicates that nintedanib plus docetaxel showed a statistically significant advantage to comparator;

≥ indicates an advantage that was not statistically significant;

~ indicates that the comparison was non-significant and very close to 1.0 (0.85 to 1.18) suggesting similarity;

≤ indicates that nintedanib plus docetaxel demonstrated a trend for disadvantage to a comparator that was not statistically significant.

All comparisons are for fixed-effects models.

Summary tables of the comparative efficacy of treatments from the MTC and Bucher indirect comparisons are reported in each of the following Results sections. Forest plots for each of these analyses are reported in Appendix 10.5.2, with links to the relevant forest plot in each section of the Results.

Detailed results presented by each network for each outcome are summarised in the sections below.

Base-case analyses

OS

As demonstrated in Table 38, in a fixed-effects model nintedanib plus docetaxel has a statistically significant advantage in prolonging OS compared with patients who received docetaxel alone or erlotinib (Figure 39). There was, however, no statistically significant difference when compared with pemetrexed.

Nintedanib plus docetaxel remained statistically superior to docetaxel in the fixed-effects sensitivity analyses that included trials with a high likelihood of EGFR-TK mutations, but also showed a statistically significant advantage over pemetrexed (Figure 41). For the new comparisons permitted in the sensitivity analysis, there was a statistically significant advantage for nintedanib plus docetaxel versus gefitinib, but no significant difference versus erlotinib plus pemetrexed. In the random-effects model, no comparisons achieved statistical significance (Figure 42).

Table 38: Summary of base-case analysis for nintedanib plus docetaxel versus all comparators for OS

Base-case analysis	OS (HR, 95% CrIs)		
	Base-case analysis	Sensitivity analysis adding trials selecting patients with EGFR-TK mutations	
	Fixed effects	Fixed effects	Random effects
Nintedanib + docetaxel vs docetaxel	**0.83 [0.70, 0.99]**	**0.83 [0.70, 0.99]**	0.83 [0.48, 1.44]
Nintedanib + docetaxel vs pemetrexed	0.82 [0.60, 1.11]	**0.72 [0.54, 0.95]**	0.69 [0.34, 1.37]
	0.90 [0.65, 1.26]†	0.90 [0.65, 1.26]†	0.90 [0.65, 1.26]†
Nintedanib + docetaxel vs erlotinib	**0.64 [0.46, 0.90]**	**0.73 [0.54, 0.99]**	0.74 [0.38, 1.54]
	0.56 [0.38, 0.82]†	**0.56 [0.38, 0.82]†**	0.56 [0.38, 0.82]†
Nintedanib + docetaxel vs erlotinib + pemetrexed	----	0.96 [0.58, 1.59]	0.93 [0.35, 2.38]
Nintedanib + docetaxel vs gefitinib	----	**0.71 [0.54, 0.94]**	0.70 [0.34, 1.40]
Sqrt(tau)	----	----	0.2359
Deviance information criterion	0.4095	7.059	4.212

Notes: Results are from MTC unless otherwise indicated. The results from the base-case analysis do not feature the random-effect model as there were no instances of two trials with the same comparison

†Indicates results from Bucher indirect comparison.

In the base-case analyses, the Bucher indirect comparisons were similar to those found in the MTC (Figure 39 and Figure 40). However in the sensitivity fixed-effects analysis, nintedanib plus docetaxel was no longer significantly superior to pemetrexed (Figure 41 and Figure 42).

As demonstrated in Table 39, for the base-case network nintedanib plus docetaxel had the greatest probability of being the best treatment in prolonging OS in the base-case analysis, followed by pemetrexed, then docetaxel, with erlotinib having the lowest probability of being the best treatment. Nintedanib plus docetaxel also had the highest probability of being the best treatment in the sensitivity analysis, followed by erlotinib plus pemetrexed.

XTable 39: Probabilities of each treatment being the best at improving OS: base-case analysis

	Base-case analysis	Sensitivity analysis adding trials selecting patients with EGFR-TK mutations
Nintedanib + docetaxel	70.44%	49.2%
Docetaxel	9.81%	5.62%
Pemetrexed	16.42%	0.60%
Erlotinib	3.33%	4.69%
Erlotinib + pemetrexed	NA	37.17%
Gefitinib	NA	2.72%

PFS

As demonstrated in Table 40 (Figure 47), the addition of nintedanib to docetaxel significantly prolonged PFS compared with docetaxel alone in base-case analysis using a fixed-effects model. Nintedanib plus docetaxel also showed a statistically significant advantage over erlotinib (Figure 47), but there was no significant difference compared with pemetrexed.

When the trials of populations with a high likelihood of EGFR-TK mutations were added to the network in a sensitivity analysis, nintedanib plus docetaxel was statistically superior in improving PFS compared with docetaxel, pemetrexed, and erlotinib in the fixed-effects model (Figure 49). The difference between nintedanib plus docetaxel and erlotinib remained statistically significant using a random effects model, but the other comparisons were no longer significantly different under this model due to the wider credible intervals (Figure 50). Nintedanib plus docetaxel displayed similar efficacy to gefitinib and no significant difference compared with erlotinib plus pemetrexed.

Table 40: Summary of base-case analysis for nintedanib plus docetaxel versus all comparators: PFS

Base-case analysis	PFS (HR, 95% CrIs)		
	Base-case analysis	Sensitivity analysis adding trials selecting patients with EGFR-TK mutations	
	Fixed effects	Fixed effects	Random effects
Nintedanib + docetaxel vs docetaxel	**0.77 [0.62, 0.96]**	**0.77 [0.62, 0.96]**	0.77 [0.45, 1.31]
Nintedanib + docetaxel vs pemetrexed	0.84 [0.61, 1.15]	**0.75 [0.56, 0.99]**	0.71 [0.36, 1.36]
	0.93 [0.67, 1.29]†	0.93 [0.67, 1.29]†	0.93 [0.67, 1.29]†
Nintedanib + docetaxel vs erlotinib	**0.70 [0.50, 1.00]¥**	**0.72 [0.53, 0.98]**	0.71 [0.36, 1.39]
	0.58 [0.39, 0.87]†	**0.58 [0.39, 0.87]†**	**0.58 [0.39, 0.87]†**
Nintedanib + docetaxel vs erlotinib + pemetrexed	----	1.28 [0.79, 2.09]	1.23 [0.49, 2.95]
Nintedanib + docetaxel vs gefitinib	----	0.95 [0.71, 1.27]	0.96 [0.49, 1.88]
Sqrt(tau)	----	----	0.2135
Deviance information criterion	1.568	3.625	0.9259

Notes: Results are from MTC unless otherwise indicated. The results from the base-case analysis do not feature the random-effect model as there were no instances of two trials with the same comparison.

† Indicates results from Bucher indirect comparison.

¥ The estimate for the upper bound of the 95% CrI was 0.9958, making the result statistically significant.

As with the MTC, the results from the Bucher indirect comparison showed a significantly longer PFS with nintedanib plus docetaxel versus erlotinib in both the base-case (Figure 47) and sensitivity analyses (Figure 49). However, unlike the MTC, the Bucher analysis found no

significant difference between nintedanib plus docetaxel and pemetrexed in the sensitivity analysis using a fixed-effects model (Figure 49).

Table 41: Base-case analysis probabilities of each treatment being the best: PFS

	Base-case analysis	Sensitivity analysis adding trials selecting patients with EGFR-TK mutations
Nintedanib + docetaxel	69.69%	25.01%
Docetaxel	5.01%	0.41%
Pemetrexed	18.53%	0.09%
Erlotinib	6.77%	0.35%
Erlotinib + pemetrexed	----	61.99%
Gefitinib	----	12.15%

As demonstrated in Table 41, for the base-case analysis, nintedanib plus docetaxel had the greatest probability of being the best treatment in improving PFS in the base-case analysis, followed by pemetrexed, with erlotinib and docetaxel having low probabilities of being the best treatment. However in the sensitivity analysis, which included trials with a high likelihood of having patients with EGFR-TK mutations, erlotinib plus pemetrexed had the greatest probably of being the best treatment, with nintedanib plus docetaxel ranked second best, followed by gefitinib. All other treatments were associated with a less than 1% probability of being the best treatment.

Objective response

As demonstrated in Table 42 (Figure 55), there was no significant difference in objective response between nintedanib plus docetaxel compared with docetaxel, pemetrexed or erlotinib in the base-case analysis using a fixed-effects model.

When the trials of populations with a high likelihood of EGFR-TK mutations were added to the network in a sensitivity analysis, nintedanib plus docetaxel remained not statistically different from docetaxel or pemetrexed at improving objective response using fixed or random effects. However, using fixed effects models, nintedanib plus docetaxel was statistically inferior to erlotinib, gefitinib and erlotinib plus pemetrexed (Figure 55), although the wider credible intervals with the random effects model meant that the difference was no longer statistically significant and Figure 57).

Table 42: Summary of base-case analysis for nintedanib plus docetaxel versus all comparators: objective response

Base-case analysis	OR (Odds Ratio, 95% CrIs)		
	Base-case analysis	Sensitivity analysis adding trials selecting patients with EGFR-TK mutations	
	Fixed effects	Fixed effects	Random effects
Nintedanib + docetaxel vs docetaxel	1.33 (0.61 – 2.95)	1.33 (0.61 – 2.94)	1.36 (0.4 – 4.49)
Nintedanib + docetaxel vs pemetrexed	0.98 (0.33 – 2.84)	0.97 (0.33 – 2.81)	1.0 (0.18 – 5.28)
	0.98 (0.34 – 2.83) †	0.98 (0.34 – 2.83) †	0.98 (0.34 – 2.83) †
Nintedanib + docetaxel vs erlotinib	0.33 (0.07 – 1.56)	**0.27 (0.08 – 0.92)**	0.28 (0.04 – 1.84)
Nintedanib + docetaxel vs erlotinib + pemetrexed	---	**0.14 (0.04 – 0.51)**	0.14 (0.02 – 1.10)
Nintedanib + docetaxel vs gefitinib	---	**0.18 (0.05 – 0.63)**	0.19 (0.03 – 1.29)
Sqrt(tau)	---	---	0.30
Deviance information criterion	37.47	78.55	79.88

Notes: Results are from MTC unless otherwise indicated. The results from the base-case analysis do not feature the random-effect model as there were no instances of two trials with the same comparison.
† Indicates results from Bucher indirect comparison.

Safety outcomes

The safety outcomes of any grade fatigue, nausea and diarrhoea were only able to be analysed as part of the sensitivity analysis where docetaxel and pemetrexed were assumed to be of comparable efficacy. These analyses are reported in Section 6.7.8. Although the LUME-Lung 1 trial reported additional safety outcomes, including grade 3+ fatigue and nausea, these outcomes could not be compared as either no other linked trial reported equivalent data, or the event rates in one or more of the treatment arms were zero.

6.7.7 *Please provide the statistical assessment of heterogeneity undertaken. The degree of, and the reasons for, heterogeneity should be explored as fully as possible*.

Both fixed and random effects models were investigated for the network meta-analyses.

Heterogeneity in results for OS

For the base-case analyses, we attempted to minimise heterogeneity across trials by restricting the network to those trials that did not select participants who were likely to have EGFR-TK mutations. However, the small number of trials in the network meant that we were

only able to compare fixed versus random effects models for the sensitivity analyses of OS, where these trials selecting for patients with EGFR-TK mutations were included. The differences in point estimates of effect size were very small (HR differences of up to 0.03) and the main difference was in the width of the 95% credible intervals, which meant that only the fixed effects analyses had statistical significance for nintedanib plus docetaxel compared with docetaxel, erlotinib and gefitinib.

The base-case analysis showed some inconsistency for OS effect sizes when direct and indirect evidence was compared for pemetrexed versus docetaxel, erlotinib versus docetaxel and erlotinib versus pemetrexed. There may be a variety of reasons contributing to this heterogeneity, which may include baseline EGFR-TK and FISH mutation status: the WSY001 and TAILOR studies recruited only EGFR-TK wild-type patients, and WSY001 also recruited only FISH-positive participants, whereas the LUME-Lung 1 and JMEI studies did not report EGFR-TK status (Table 43).

Table 43: Direct and indirect evidence: base-case analysis of OS (fixed effects, Base-case analysis without studies selecting for EGFR-TK mutations)

Comparison	HRs [95% CIs]	Studies	Source
Pemetrexed vs docetaxel	0.92 [0.69 - 1.22]	JMEI	Direct
Pemetrexed vs docetaxel	1.48 [0.86 - 2.55]	TAILOR _VS_ WSY001	Indirect
Erlotinib vs docetaxel	1.49 [1.06 - 2.10]	TAILOR	Direct
Erlotinib vs docetaxel	0.93 [0.56 - 1.55]	JMEI _VS_ WSY001	Indirect
Erlotinib vs pemetrexed	1.01 [0.66 - 1.54]	WSY001	Direct
Erlotinib vs pemetrexed	1.62 [1.04 - 2.53]	JMEI _VS_ TAILOR	Indirect
Nintedanib + docetaxel vs docetaxel	0.83 [0.70 - 0.99]	LUME-Lung 1	Direct
Nintedanib + docetaxel vs pemetrexed	0.90 [0.65 - 1.26]	JMEI vs LUME-Lung 1	Indirect
Nintedanib + docetaxel vs erlotinib	0.56 [0.38 - 0.82]	TAILOR _VS_ LUME-Lung 1	Indirect

Heterogeneity in PFS effect sizes

As with the OS analyses, for the base-case analysis, we were only able to compare fixed versus random effects models of PFS when trials that selected for EGFR-TK mutations were added to the network in a sensitivity analysis. The differences in point estimates of effect size were again very small (HR differences of up to 0.04) and the main difference was in the width of the 95% credible intervals, which meant that, although the comparison with erlotinib remained statistically significant using fixed or random effects models, only the fixed effects analyses were statistically significant for nintedanib plus docetaxel compared with docetaxel or pemetrexed.

The base-case analysis also showed some inconsistency for PFS effect sizes when direct and indirect evidence was compared for pemetrexed versus docetaxel, erlotinib versus docetaxel and erlotinib versus pemetrexed (Table 44). As with OS, this heterogeneity may be at least partly explained by differences in EGFR-TK and other mutation rates across studies.

Table 44: Direct and indirect evidence: base-case analysis for PFS (fixed effects, excluding studies selecting for EGFR-TK mutation)

Comparison	HRs [95% CIs]	Studies	Source
Pemetrexed vs docetaxel	0.83 [0.65 - 1.06]	JMEI	Direct
Pemetrexed vs docetaxel	1.43 [0.85 - 2.40]	TAILOR _VS_ WSY001	Indirect
Erlotinib vs docetaxel	1.32 [0.94 - 1.83]	TAILOR	Direct
Erlotinib vs docetaxel	0.76 [0.48 - 1.22]	JMEI _VS_ WSY001	Indirect
Erlotinib vs pemetrexed	0.92 [0.62 - 1.37]	WSY001	Direct
Erlotinib vs pemetrexed	1.59 [1.05 - 2.40]	JMEI _VS_ TAILOR	Indirect
Nintedanib + docetaxel vs docetaxel	0.77 [0.62 - 0.96]	LUME-Lung 1	Direct
Nintedanib + docetaxel vs pemetrexed	0.93 [0.67 - 1.29]	JMEI vs LUME-Lung 1	Indirect
Nintedanib + docetaxel vs erlotinib	0.59 [0.39 - 0.87]	TAILOR _VS_ LUME-Lung 1	Indirect

6.7.8 *If there is doubt about the relevance of a particular trial, please present separate sensitivity analyses in which these trials are excluded*.

Scenario analyses where docetaxel and pemetrexed are assumed to have equivalent efficacy

The MTC results for the scenario analysis network in which docetaxel and pemetrexed were assumed to be equivalent show that nintedanib plus docetaxel is statistically superior to docetaxel/pemetrexed and erlotinib at improving both OS and PFS using a Bayesian MTC

fixed-effects model, although the wide credible intervals mean that the Bayesian MTC is no longer statistically significant using a random-effects model. Nintedanib plus docetaxel was more effective than docetaxel/pemetrexed at increasing objective response but was not significantly different from erlotinib for this outcome.

The main treatments varied in their safety profiles. Nintedanib plus docetaxel was associated with higher risks of nausea than docetaxel/pemetrexed or erlotinib; erlotinib was associated with the lowest risks of nausea. In the scenario analysis that did not consider EGFR-TK status, nintedanib plus docetaxel had a significantly lower risk of diarrhoea than either erlotinib or gefitinib. However, nintedanib plus docetaxel was still associated with a higher risk of diarrhoea than docetaxel/pemetrexed. Nintedanib plus docetaxel was associated with a similar risk of fatigue compared with erlotinib or docetaxel/pemetrexed, but gefitinib was associated with a lower risk of fatigue than all other comparators.

Overall survival

As summarised in Table 45, the scenario analysis of trials that did not select patients primarily with EGFR-TK mutations, using the network that assumes equivalent efficacy of docetaxel and pemetrexed, shows that nintedanib plus docetaxel led to statistically significantly longer OS than docetaxel/pemetrexed and erlotinib using a Bayesian MTC fixed-effects model (Figure 43). When using a random-effects model, the wide credible intervals mean that the favourable trend seen in the Bayesian MTC is no longer statistically significant when nintedanib plus docetaxel is compared with docetaxel/pemetrexed or erlotinib.

Table 45: Summary of sensitivity analysis for nintedanib plus docetaxel versus all comparators: OS

Sensitivity analysis	OS (HR, 95% CrIs)			
	Scenario analysis		Sensitivity analysis adding trials selecting patients with EGFR-TK mutations	
	Fixed effects	Random effects	Fixed effects	Random effects
Nintedanib + docetaxel vs docetaxel/pemetrexed	**0.83 [0.70, 0.99]**	0.83 [0.50, 1.38]	**0.83 [0.70, 0.99]**	0.83 [0.48, 1.42]
Nintedanib + docetaxel vs erlotinib	**0.74 [0.57, 0.96]**	0.74 [0.40, 1.35]	0.82 [0.64, 1.05]	0.84 [0.46, 1.59]
	0.74 [0.57, 0.96][†]	0.74 [0.53, 1.02][†]	0.80 [0.63, 1.03][†]	0.82 [0.58, 1.17][†]
Nintedanib + docetaxel vs gefitinib	----	----	**0.73 [0.56, 0.96]**	0.75 [0.38, 1.46]
Nintedanib + docetaxel vs pemetrexed + erlotinib	----	----	1.11 [0.70, 1.74]	1.10 [0.47, 2.55]
Sqrt(tau)	----	0.1951	----	0.23

Sensitivity analysis	OS (HR, 95% CrIs)			
	Scenario analysis		Sensitivity analysis adding trials selecting patients with EGFR-TK mutations	
	Fixed effects	Random effects	Fixed effects	Random effects
Deviance information criterion	0.3104	-0.2181	6.755	4.189

Notes: Results are from MTC unless otherwise indicated. Statistically significant comparisons are shown in **bold**.
† Indicates results from Bucher indirect comparison.

In the additional sensitivity analysis, when trials of populations with a high likelihood of EGFR-TK mutations were added to the MTC, a comparison between nintedanib plus docetaxel and gefitinib and pemetrexed plus erlotinib was available in addition to those comparisons found in the base-case. As with the scenario analysis for this network, nintedanib plus docetaxel was significantly superior to docetaxel/pemetrexed for OS using fixed-effects, but the difference was not significant using random-effects. Nintedanib plus docetaxel was statistically superior to gefitinib using fixed-effects, but there were no significant differences compared with erlotinib, or pemetrexed plus erlotinib (Figure 45 and Figure 46).

In addition to the MTC, Bucher indirect comparisons were performed for nintedanib plus docetaxel compared with erlotinib and found similar results to the MTC; however, the Bucher indirect comparisons showed narrower credible intervals for the base-case (Figure 43 and Figure 44) and sensitivity analyses (Figure 45 and Figure 46).

As demonstrated in Table 46, nintedanib plus docetaxel had the greatest probability of being the best treatment for improving OS in the scenario analysis assuming equivalence of docetaxel and pemetrexed (using data from the random-effects model), followed by docetaxel/pemetrexed and erlotinib. Results differed in the additional sensitivity analysis when studies selecting patients with EGFR-TK mutations were added to the network. In this analysis, erlotinib plus pemetrexed demonstrated the greatest probability of being the best treatment, followed by nintedanib plus docetaxel, with docetaxel/pemetrexed having the lowest probability of being the best treatment.

Table 46: Probabilities of each treatment being the best: OS

	Scenario analysis	Sensitivity analysis adding trials selecting patients with EGFR-TK mutations
Nintedanib + docetaxel	78.95%	34.21%
Docetaxel/pemetrexed	13.65%	1.20%

Erlotinib	7.40%	6.79%
Gefitinib	----	3.40%
Erlotinib + pemetrexed	----	54.39%

Progression Free Survival

When docetaxel and pemetrexed are assumed to have equal efficacy, nintedanib plus docetaxel is significantly better at prolonging PFS than both docetaxel/pemetrexed and erlotinib, using a fixed-effects model (Table 47, Figure 51). As expected with using a random-effects model, the wide credible intervals mean that neither of the comparisons showed statistical significance (Figure 52).

Table 47: Summary of scenario analysis of nintedanib plus docetaxel versus all comparators for PFS

Sensitivity analysis	PFS (HR, 95% CrIs)			
	Scenario analysis		Sensitivity analysis adding trials selecting patients with EGFR-TK mutations	
	Fixed effects	Random effects	Fixed effects	Random effects
Nintedanib + docetaxel vs docetaxel/pemetrexed	**0.77 [0.62, 0.96]**	0.77 [0.45, 1.30]	**0.77 [0.62, 0.96]**	0.77 [0.47, 1.27]
Nintedanib + docetaxel vs erlotinib	**0.68 [0.49, 0.95]**	0.68 [0.35, 1.35]	**0.74 [0.55, 0.98]**	0.75 [0.42, 1.34]
	0.68 [0.48, 0.95]†	0.69 [0.46, 1.04]†	**0.71 [0.53, 0.96]†**	**0.71 [0.53, 0.97]†**
Nintedanib + docetaxel vs gefitinib	----	----	0.96 [0.72, 1.27]	0.99 [0.55, 1.83]
Nintedanib + docetaxel vs erlotinib + pemetrexed	----	----	1.33 [0.85, 2.07]	1.33 [0.61, 2.85]
Sqrt(tau)	----	0.1953	----	0.1825
Deviance information criterion	0.2461	0.1108	1.787	1.041

Notes: Results are from MTC unless otherwise indicated.
† Indicates results from Bucher indirect comparison.

When the trials of populations with a high likelihood of EGFR-TK mutations were added to the network in an additional sensitivity analysis, nintedanib plus docetaxel continued to be statistically superior to docetaxel/pemetrexed and erlotinib in improving PFS using a fixed-effects model but not when using a random-effects model. For the new comparisons available, in both the fixed- and random-effects models of the MTC, there were no significant differences between nintedanib plus docetaxel and gefitinib or erlotinib plus pemetrexed (Figure 53 and Figure 54).

The Bucher indirect comparison of nintedanib plus docetaxel versus erlotinib had similar point estimates to the MTC, but the narrower credible intervals meant that the benefit of nintedanib plus docetaxel was statistically significant under fixed effects models for both the scenario analysis and the sensitivity analysis with trials that selected for EGFR-TK mutations, and was also significant with a random effects model in this sensitivity analysis.

Table 48: Scenario analysis probabilities of each treatment being the best: PFS

	Scenario analysis	Sensitivity analysis adding trials selecting patients with EGFR-TK mutations
Nintedanib + docetaxel	83.57%	16.42%
Docetaxel/pemetrexed	8.75%	00.04%
Erlotinib	7.67%	0.30%
Gefitinib	----	10.99%
Erlotinib + pemetrexed	---	72.23%

As demonstrated in Table 48, nintedanib plus docetaxel had the highest probability of being the best treatment in the scenario analysis for this network, followed by docetaxel/pemetrexed, with erlotinib having the lowest probability of being the best treatment. In the additional sensitivity analysis, when the trials targeting populations with a high likelihood of EGFR-TK mutations were added, erlotinib plus pemetrexed had the highest probability of being the best treatment in prolonging PFS, followed by nintedanib plus docetaxel and gefitinib. The probability of erlotinib or standard chemotherapy being the best treatment was very low.

Objective response

The size of the network for objective response limits analysis to a fixed-effects model only. As seen in Table 49, nintedanib plus docetaxel shows no significant difference in objective response compared with docetaxel/pemetrexed or erlotinib (Figure 58).

Table 49: Summary of scenario analysis for nintedanib plus docetaxel versus all comparators: objective response

Sensitivity analysis	Objective response (Odd Ratios, 95% CrIs)	
	Scenario analysis	Sensitivity analysis
	Fixed effects	Fixed effects
Nintedanib + docetaxel vs docetaxel/pemetrexed	1.33 [0.61, 2.96]	1.33 [0.61, 2.99]
Nintedanib + docetaxel vs erlotinib	0.45 [0.11, 1.77]	0.37 [0.14, 1.01]
	0.47 [0.12, 1.83]†	0.40 [0.14, 1.11]†

Nintedanib + docetaxel vs gefitinib	----	**0.25 [0.09, 0.70]**
Nintedanib + docetaxel vs erlotinib + pemetrexed	----	**0.19 [0.07, 0.57]**
Deviance information criterion	24.37	65.49

Notes: Results are from MTC unless otherwise indicated. † Indicates results from Bucher indirect comparison.

When trials of populations with a high likelihood of EGFR-TK mutations were added to the network in an additional sensitivity analysis, similar results were found compared with the core analysis for this network. Nintedanib plus docetaxel was not significantly different from docetaxel/pemetrexed or erlotinib but was significantly inferior to gefitinib and erlotinib plus pemetrexed (Table 49, Figure 59).

As with the MTC, Bucher indirect comparisons of nintedanib plus docetaxel were not significantly different from erlotinib (Table 49, Figure 58 and Figure 59).

Safety

Safety outcomes for any grade fatigue, nausea and diarrhoea were only analysable as part of the sensitivity analysis where docetaxel and pemetrexed were assumed to be of comparable efficacy. Although the LUME-Lung 1 trial reported additional safety outcomes, including grade 3+ fatigue and nausea, these outcomes could not be compared as either no other linked trial reported equivalent data, or the event rates in one or more of the treatment arms were zero. As each treatment arm in the network had only one trial, only fixed effects models could be used for the safety analyses.

Any grade adverse event: diarrhoea

For safety outcomes, analyses could only be conducted for the scenario analysis that assumed equivalent efficacy and tolerability of docetaxel and pemetrexed. Using a Bayesian MTC fixed-effects model (Figure 60), results suggest that patients taking nintedanib plus docetaxel were significantly more likely to develop any grade diarrhoea compared with docetaxel/pemetrexed. Compared with erlotinib, however, there was no significantly increased risk of diarrhoea with nintedanib plus docetaxel (Table 50).

Table 50: Summary of scenario analysis for nintedanib plus docetaxel versus all comparators: any grade diarrhoea

Sensitivity analysis	**Any grade diarrhoea (odds ratios, 95% CrIs)**	
	Scenario analysis	**Sensitivity analysis**
	Fixed effects	**Fixed effects**

Nintedanib + docetaxel vs docetaxel/pemetrexed	**2.35 [1.68, 3.28]**	**2.36 [1.69, 3.31]**
Nintedanib + docetaxel vs erlotinib	0.34 [0.04, 1.54]	**0.31 [0.09, 0.93]**
	0.40 [0.04, 1.54]†	0.40 [0.08, 1.98]†
Nintedanib + docetaxel vs gefitinib	----	**0.32 [0.10, 0.90]**
Deviance information criterion	26.74	48.74

Notes: Results are from MTC unless otherwise indicated. † Indicates results from Bucher indirect comparison.

When trials targeting populations with a high likelihood of EGFR-TK mutations were added to the network in an additional sensitivity analysis, the increased risk of diarrhoea with nintedanib plus docetaxel remained compared with docetaxel/pemetrexed. However, risk with nintedanib plus docetaxel was significantly lower than with erlotinib or gefitinib (Table 50, Figure 61).

Compared with the MTC, the Bucher analyses found similar results for nintedanib plus docetaxel compared with erlotinib in the base-case analysis (Figure 60). However, the advantage found for nintedanib plus docetaxel compared with erlotinib was no longer statistically significant in the Bucher indirect comparison for the sensitivity analysis (Table 50, Figure 61).

Any grade adverse event: fatigue

Using Bayesian MTC fixed-effects models (Table 51, Figure 62), the risk of fatigue was similar for all comparisons in the scenario analysis and additional sensitivity analyses when trials targeting populations with a high likelihood of EGFR-TK mutations were added to the network (Figure 63).

Table 51: Summary of scenario analysis for nintedanib plus docetaxel versus all comparators: any grade fatigue

Sensitivity analysis	**Any grade fatigue (odds ratios, 95% CrIs)**	
	Scenario analysis	**Sensitivity analysis**
	Fixed effects	**Fixed effects**
Nintedanib + docetaxel vs docetaxel/pemetrexed	1.07 [0.77, 1.50]	1.08 [0.77, 1.50]
Nintedanib + docetaxel vs erlotinib	1.54 [0.61, 3.96]	0.92 [0.39, 2.11]
	1.53 [0.61, 3.80]†	1.53 [0.61, 3.80]†
Nintedanib + docetaxel vs gefitinib	----	1.63 [0.72, 3.75]
Deviance information criterion	28.61	56.14

Notes: Results are from MTC unless otherwise indicated. † Indicates results from Bucher indirect comparison.

Bucher indirect comparisons between nintedanib plus docetaxel and erlotinib agreed with the base-case MTC analysis and found no significant difference in the risk of fatigue (Table 51, Figure 62, Figure 63).

Any grade adverse event: nausea

Using Bayesian MTC fixed-effects models (Table 52, Figure 64), patients taking nintedanib plus docetaxel were significantly more likely to develop any grade nausea compared with docetaxel/pemetrexed or erlotinib.

Table 52: Summary of scenario analysis for nintedanib plus docetaxel versus all comparators: any grade nausea

Sensitivity analysis	Any grade nausea (odds ratios, 95% CrIs)	
	Scenario analysis	Sensitivity analysis
	Fixed effects	Fixed effects
Nintedanib + docetaxel vs docetaxel/pemetrexed (MTC)	**1.85 [1.27, 2.68]**	**1.85 [1.28, 2.69]**
Nintedanib + docetaxel vs erlotinib (MTC)	**50.65 [7.78, 1380.22]**	**15.83 [4.54, 78.73]**
	35.34 [4.36, 286.63]†	**35.34 [4.36, 286.63]†**
Nintedanib + docetaxel vs gefitinib (MTC)	----	2.46 [0.93, 6.52]
Deviance information criterion	26.03	48.81

Notes: Results are from MTC unless otherwise indicated. Bold: results from MTC were statistically significant.
† Indicates results from Bucher indirect comparison.

When trials targeting populations with a high likelihood of EGFR-TK mutations were added to the analysis, results were similar although the difference between nintedanib plus docetaxel and erlotinib was greatly reduced in the MTC (Figure 648, Table 52). The difference between nintedanib plus docetaxel and gefitinib was not statistically significant.

Bucher indirect comparisons gave similar results compared with the MTC analyses, with a significantly higher risk of nausea with nintedanib plus docetaxel versus erlotinib for both the base-case and sensitivity analyses (Figure 64 and Figure 65).

6.7.9 *Please discuss any heterogeneity between results of pairwise comparisons and inconsistencies between the direct and indirect evidence on the technologies*.

Please refer to Section 6.7.7.

6.8 Non-RCT evidence

Non-RCT, both experimental and observational, evidence will be required, not just for those situations in which RCTs are unavailable, but also to supplement information from RCTs when they are available. This section should be read in conjunction with NICE's 'Guide to the methods of technology appraisal', sections 3.2.8 to 3.2.10.

6.8.1 If non-RCT evidence is considered (see section 6.2.7), please repeat the instructions specified in sections 6.1 to 6.5 for the identification, selection and methodology of the trials, and the presentation of results. For the quality assessments of non-RCTs, use an appropriate and validated quality assessment instrument. Key aspects of quality to be considered can be found in 'Systematic reviews: CRD's guidance for undertaking reviews in health care' (www.york.ac.uk/inst/crd). Exact details of the search strategy used and a complete quality assessment for each trial should be provided in sections 10.6 and 10.7, appendices 6 and 7.

Not applicable. No RCT evidence is included in this submission.

6.9 Adverse events

This section should provide information on the adverse events experienced with the technology in relation to the decision problem. Evidence from comparative RCTs and regulatory summaries is preferred; however, findings from non-comparative trials may sometimes be relevant. For example, post-marketing surveillance data may demonstrate that the technology shows a relative lack of adverse events commonly associated with the comparator, or the occurrence of adverse events is not significantly associated with other treatments.

6.9.1 If any of the main trials are designed primarily to assess safety outcomes (for example, they are powered to detect significant differences between treatments with respect to the incidence of an adverse event), please repeat the instructions specified in sections 6.1 to 6.5 for the identification, selection, methodology and

quality of the trials, and the presentation of results. Examples for search strategies for specific adverse effects and/or generic adverse-effect terms and key aspects of quality criteria for adverse-effects data can found in 'Systematic reviews: CRD's guidance for undertaking reviews in health care' (www.york.ac.uk/inst/crd). Exact details of the search strategy used and a complete quality assessment for each trial should be provided in sections 10.8 and 10.9, appendices 8 and 9.

The LUME-Lung 1 study included safety and tolerability as a secondary endpoint. There are no RCTs of nintedanib with safety and tolerability as the primary outcome.

6.9.2 Please provide details of all important adverse events for each intervention group. For each group, give the number with the adverse event, the number in the group and the percentage with the event. Then present the relative risk and risk difference and associated 95% confidence intervals for each adverse event. A suggested format is shown below.

In a similar manner to efficacy outcomes, safety data are presented for patients with adenocarcinoma tumour histology, as this is the relevant indication being sought from the regulatory authorities. Safety data for the overall patient population are not presented here, but were reflective of that seen in the subgroup of patients with adenocarcinoma(3).

Treatment exposure

In the adenocarcinoma patient population, the median duration of nintedanib/placebo treatment was 4.2 months (range 0.1 months to 41.5 months) in the nintedanib plus docetaxel arm, and 3.0 months (range 0.1 months to 31.1 months) in the docetaxel plus placebo arm (Table 53). The mean dose intensity of nintedanib/placebo was 91.2% in the nintedanib plus docetaxel arm and 93.8% in the docetaxel plus placebo arm(5).

The addition of nintedanib did not impact on the median duration of docetaxel treatment. The median number of docetaxel cycles in adenocarcinoma patients was higher in the nintedanib plus docetaxel arm (5 cycles, range 1 to 45) than in the placebo plus docetaxel

arm (4 cycles, range 1 to 42; Table 53)(5). The overall mean dose intensity of docetaxel was similar in each treatment arm; 98.1% in the nintedanib plus docetaxel arm, and 98.7% in the placebo plus docetaxel arm(5).

Table 53: Treatment exposure in LUME-Lung 1 in the adenocarcinoma population(5)

	Nintedanib plus docetaxel	Placebo plus docetaxel
Nintedanib/placebo		
Median duration of nintedanib/placebo treatment (range)	4.2 months (0.10 to 41.53)	3.0 months (0.07 to 31.10)
Mean dose intensity of nintedanib/placebo (%, SD)	91.2 (15.0)	93.8 (13.3)
Docetaxel		
Number of docetaxel courses (median, range)	5.0 (1 to 45)	4.0 (1 to 41)
Mean overall dose intensity of docetaxel (%, SD)	98.1 (4.5)	98.7 (3.7)

Dose reduction and dose interruption

For patients with adenocarcinoma, dose reductions were more frequent in the nintedanib arm than in the placebo arm. The proportion of adenocarcinoma patients with at least one dose reduction of nintedanib or placebo was 21.9% and 6.6%, respectively (Table 54)(5). However, these numerical differences between the treatment arms may be influenced by the greater treatment duration in the nintedanib plus docetaxel arm (4.2 months) compared with the placebo plus docetaxel arm (3.0 months)(5).

For most patients with adenocarcinoma, a single dose reduction was sufficient to manage AEs (nintedanib: 17.2%, placebo: 6.6%). A second dose reduction was necessary in 4.7% and 0% patients in the nintedanib and placebo arms, respectively(5).

Table 54: Dose reductions in LUME-Lung 1 in the adenocarcinoma population(5)

	Nintedanib plus docetaxel	Placebo plus docetaxel
Nintedanib/placebo		
At least 1 dose reduction of nintedanib/placebo (n, %)	70 (21.9)	22 (6.6)
Docetaxel		
Dose reduction of docetaxel (n, %)	54 (16.9)	41 (12.3)

Nintedanib dose reductions due to the most commonly reported AEs were required in only a small proportion of adenocarcinoma patients, though these rates were higher in the

nintedanib plus docetaxel arm compared with the placebo plus docetaxel arm (Table 55)(64).

Table 55: AEs leading to dose reduction of nintedanib/placebo during LUME-Lung 1 in >1% of the adenocarcinoma population(64)

Adverse event	Nintedanib plus docetaxel (n, %)	Placebo plus docetaxel (n, %)
Total requiring dose reduction	69 (21.6)	22 (6.6)
Diarrhoea	26 (8.1)	11 (3.3)
ALT increased	25 (7.8)	2 (0.6)
AST increased	12 (3.8)	0
Vomiting	7 (2.2)	2 (0.6)
Nausea	4 (1.3)	1 (0.3)

ALT = Alanine aminotransferase; AST = Aspartate transaminase

During the trial, treatment with nintedanib/placebo was interrupted in the event of diarrhoea, nausea, vomiting, ALT/AST increase, or non-haematological drug-related AEs of pre-specified severity. Treatment could be subsequently restarted, at a reduced dose, as long as the events causing the interruption had recovered to pre-dose values or to a CTCAE grade which allowed further therapy(4).

Nintedanib/placebo treatment interruptions were more common in patients in the nintedanib arm (52.2%) than in the placebo arm (41.4%)(5). Treatment interruptions for >14 consecutive days were also more common in the nintedanib arm (nintedanib: 10.0%, placebo: 6.6%, Table 56)(5). The difference in median treatment duration with nintedanib compared to placebo (4.2 months 3 months, respectively) needs to be taken into consideration as the longer exposure in the nintedanib arm may have contributed to the higher rate of AEs.

Table 56: Treatment interruptions in LUME-Lung 1 in the adenocarcinoma population(5)

	Nintedanib plus docetaxel	Placebo plus docetaxel
At least 1 temporary interruption of nintedanib/placebo (n, %)	167 (52.2)	137 (41.4)
At least 1 temporary interruption of nintedanib/placebo >14 consecutive days (n, %)	32 (10.0)	22 (6.6)

Treatment discontinuation in LUME-Lung 1

The proportion of adenocarcinoma patients who permanently discontinued therapy in LUME-Lung 1 was 99.1% in both treatment arms. The most common reason for

discontinuation of therapy was disease progression (64.4% vs 69.1%) for the nintedanib and placebo groups, respectively (Table 57)(5). In adenocarcinoma patients, AEs leading to permanent discontinuation of study medication were numerically higher in the nintedanib arm than in the placebo arm for all AEs overall (20.9% nintedanib and 17.7% placebo)(5).

Table 57: Patient disposition in LUME-Lung 1 in the adenocarcinoma population (5)

	Nintedanib plus docetaxel (n, %)	Placebo plus docetaxel (n, %)
On treatment	3 (0.9)	3 (0.9)
Discontinued permanently[1]	317 (99.1)	330 (99.1)
Reason for permanent discontinuation		
Progressive disease	204 (64.4)	228 (69.1)
All patients who discontinued due to an AE	67 (20.9)	59 (17.7)
Worsening or AE of underlying cancer disease	23 (7.3)	27 (8.2)
Other AE[2]	44 (13.9)	33 (10.0)
Noncompliance with protocol	4 (1.3)	3 (0.9)
Lost to follow-up	1 (0.3)	2 (0.6)
Patient refusal to continue study medication	29 (9.1)	26 (7.9)
Other	12 (3.8)	11 (3.3)

[1] Defined as permanent discontinuation of all components of the study medication (nintedanib or placebo and chemotherapy, if given).

[2] Other AEs than 'Worsening or AE of underlying cancer disease'

A higher proportion of adenocarcinoma patients had to permanently discontinue treatment in the nintedanib arm compared with the placebo arm due to increased ALT (1.6% nintedanib; 0% placebo) and increased AST (1.3% nintedanib; 0.3% placebo). However fewer patients in the nintedanib plus docetaxel arm discontinued treatment due to dyspnoea, compared to the placebo plus docetaxel arm (1.3% and 3.3% respectively). Otherwise, the difference between treatment arms for individual AEs leading to discontinuation was <1% (Table 58). In the nintedanib treatment arm, the discontinuation rates due to the most commonly reported AEs were low (diarrhoea: 0.9%, nausea: 0.3%, vomiting: 0.6%), and comparable to the placebo arm (diarrhoea: 0.3%, nausea: 0%, vomiting: 0%)(5).

Table 58: AEs leading to permanent discontinuation of last study treatment in ≥1% of patients with adenocarcinoma in the nintedanib plus docetaxel arm in LUME-Lung 1(5)

Adverse event	Nintedanib plus docetaxel (n, %)	Placebo plus docetaxel (n, %)
All patients who discontinued due to an AE	67 (20.9)	59 (17.7)
ALT increased	5 (1.6)	0
Malignant neoplasm progression	5 (1.6)	5 (1.5)

Adverse event	Nintedanib plus docetaxel (n, %)	Placebo plus docetaxel (n, %)
AST increased	4 (1.3)	1 (0.3)
Dyspnoea	4 (1.3)	11 (3.3)

AEs = adverse events; ALT = Alanine aminotransferase; AST = Aspartate transaminase

Overview of AEs in LUME-Lung 1

In the adenocarcinoma population, the proportion of patients who experienced any AE was similar between the nintedanib plus docetaxel group (96.3%) and placebo plus docetaxel group (94.3%)(5). Drug-related AEs were more frequent with nintedanib plus docetaxel (81.3%) than with placebo plus docetaxel (72.4%). Similarly, the proportion of patients with AEs grade ≥3 was higher in the nintedanib plus docetaxel arm (75.9%) than in the placebo plus docetaxel arm (68.5%). The proportion of patients with SAEs was however comparable across arms (34.7% and 32.1% for nintedanib plus docetaxel and placebo plus docetaxel, respectively)(5). A summary of AEs in the adenocarcinoma population can be seen in Table 59(3).

Table 59: Summary of AEs in LUME-Lung 1 in the adenocarcinoma population(3, 5)

	Nintedanib plus docetaxel n (%)	Placebo plus docetaxel n (%)
Patients with AEs(3)	308 (96.3)	314 (94.3)
Drug-related AEs[1](5)	260 (81.3)	241 (72.4)
AEs leading to dose reduction of nintedanib placebo[1](3)	69 (21.6)	22 (6.6)
AEs leading to dose reduction of docetaxel(3)	53 (16.6)	41 (12.3)
AEs leading to permanent discontinuation[2](3)	67 (20.9)	59 (17.7)
SAEs(3)	111 (34.7)	107 (32.1)
Fatal AEs(3)	56 (17.5)	32 (9.6)
Fatal AEs not attributed to PD(3)	20 (6.3)	8 (2.4)
Fatal AEs attributed to PD[3](3)	36 (11.3)	24 (7.2)
Highest CTCAE grade ≥3(3)	243 (75.9)	228 (68.5)

AEs = adverse events; CTCAE = Common Terminology Criteria for Adverse Events; PD = progressive disease; SAEs = serious adverse events

[1] As judged by the investigator

[2] AEs leading to discontinuation of last study medication i.e. AEs leading to permanent treatment discontinuation

[3] Attribution to PD by the investigator, as documented on the Case Report Form

Fatal AEs

The overall incidence of fatal AEs in the adenocarcinoma population was 17.5% in the nintedanib plus docetaxel arm and 9.6% in the placebo plus docetaxel arm(5). AEs leading to death were attributed by the investigator to progression of the underlying disease in 11.3%

of patients in the nintedanib arm, and 7.2% of patients in the placebo arm(5). Moreover, when considering only those fatal events with a start and end date in the on-treatment period, the time-adjusted incidence rate ratio was 1.54 (95% CI 0.96-2.46)(64). The number of patients with AEs leading to death that were considered drug-related was low in patients with adenocarcinoma (1.9% vs 0.3% in the nintedanib plus docetaxel and the placebo plus docetaxel arm, respectively)(5).

The number of patients who died within the first 6 weeks after the start of treatment was balanced between treatment arms in patients with adenocarcinoma (nintedanib plus docetaxel: 4% vs placebo plus docetaxel: 3.6%), indicating that the combination therapy with nintedanib and docetaxel had no acute toxicity(64).

A summary of AEs leading to death in ≥2 patients in the adenocarcinoma population can be seen in Table 60.

Table 60: Summary of fatal AEs in LUME-Lung 1 in the adenocarcinoma population(5)

	Nintedanib plus docetaxel n (%)	**Placebo plus docetaxel n (%)**
Fatal AEs	56 (17.5)	32 (9.6)
Malignant neoplasm progression	12 (3.8)	7 (2.1)
Dyspnoea	6 (1.9)	7 (2.1)
General physical health deterioration	5 (1.6)	3 (0.9)
Respiratory failure	5 (1.6)	1 (0.3)
Sepsis	3 (0.9)	0 (0)
Chest pain	2 (0.6)	0 (0)
Metastases to meninges	2 (0.6)	0 (0)
Multi-organ failure	2 (0.6)	0 (0)
Pneumonia	2 (0.6)	1 (0.3)
Fatal AEs not attributed to PD	20 (6.3)	8 (2.4)
Fatal AEs attributed to PD[1]	36 (11.3)	24 (7.2)

AEs = adverse events

Attribution to PD by the investigator, as documented on the Case Report Form

The number of patients with fatal AEs considered drug-related by the investigator was low in both arms, but was slightly higher in the nintedanib plus docetaxel arm than in the placebo plus docetaxel arm (1.9% vs 0.3%, Table 61)(5).

Table 61: Fatal AEs considered drug-related by the investigator in the adenocarcinoma population(5)

	Nintedanib plus docetaxel n (%)	Placebo plus docetaxel n (%)
Fatal AEs	6 (1.9)	1 (0.3)
Sepsis	2 (0.6)	0
Dehydration	1 (0.3)	0
Diverticulum intestinale[1]	1 (0.3)	0
Ischaemic stroke	1 (0.3)	0
Large intestine perforation[1]	1 (0.3)	0
Neutropenic infection	1 (0.3)	0
Dyspnoea	0	1 (0.3)

[1] One patient experienced more than 1 fatal AE considered drug-related (patient with large intestine perforation and diverticulum intestinale)

Most commonly reported AEs in LUME-Lung 1

The number of patients who experienced a drug-related AE or CTCAE grade ≥3 AE was higher in the nintedanib plus docetaxel arm, compared to the placebo plus docetaxel arm (Table 62). However, these data must be seen in context with the longer treatment duration in the nintedanib plus docetaxel arm (Table 53)(5).

AEs occurring at an incidence of ≥5% in either treatment arm can be seen in Table 62. AEs which occurred more frequently in the nintedanib group (≥5% difference) included:(3)

- Diarrhoea (any grade: 43.4% vs 24.6%; grade ≥3 6.3% vs 3.6%)
  - The risk ratio for any grade diarrhoea in patients with adenocarcinoma was 1.76 (95% CI; 1.41–2.21), risk difference 0.19 (95% CI; 0.12–0.26).(110)
  - The risk ratio for grade ≥3 diarrhoea in patients with adenocarcinoma was 1.73 (95% CI; 0.86–3.49), risk difference 0.03 (95% CI; -0.01-0.06).(110)
- Increased ALT (any grade: 37.8% vs 9.3%; grade ≥3 11.6% vs 0.9%)
  - The risk ratio for any grade increased ALT in patients with adenocarcinoma was 4.06 (95% CI; 2.82–5.84), risk difference 0.29 (95% CI; 0.22–0.35).(110)
  - The risk ratio for grade ≥3 increased ALT in patients with adenocarcinoma was 12.83 (95% CI; 4.00–41.21), risk difference 0.11 (95% CI; 0.07–0.14).(110)
- Increased AST (any grade: 30.3% vs 7.2%; grade ≥3 4.1% vs 0.6%)
  - The risk ratio for any grade increased AST in patients with adenocarcinoma was 4.21 (95% CI; 2.76–6.40), risk difference 0.23 (95% CI; 0.17–0.29).(110)

- The risk ratio for grade ≥3 increased AST in patients with adenocarcinoma was 6.76 (95% CI; 1.54–29.74), risk difference 0.03 (95% CI; 0.01–0.06).(110)

- Nausea (any grade: 28.4% vs 17.7%; grade ≥3 0.9% vs 0.6%)
- Decreased appetite (any grade 23.4% vs 15.6%; grade ≥3 1.3% vs 1.5%
- Vomiting (any grade 19.4% vs 12.3%; grade ≥3 1.3% vs 0.6%)

The low rate of permanent nintedanib treatment discontinuations due to commonly reported AEs (Table 58), suggests that these were largely manageable by dose reductions, treatment interruption and/or symptomatic therapy(5).

Table 62: AEs occurring with an incidence >5% in either treatment arm in the adenocarcinoma population(3)

	Nintedanib plus docetaxel, n (%)		Placebo plus docetaxel, n (%)	
	Any grade n (%)	Grade ≥3 n (%)	Any grade n (%)	Grade ≥3 n (%)
Patients with AEs	308 (96.3)	243 (75.9)	314 (94.3)	228 (68.5)
Diarrhoea	139 (43.4)	20 (6.3)	82 (24.6)	12 (3.6)
Neutrophil count decrease	131 (40.9)	116 (36.3)	135 (40.5)	116 (34.8)
ALT increased	121 (37.8)	37 (11.6)	31 (9.3)	3 (0.9)
Fatigue	99 (30.9)	15 (4.7)	98 (29.4)	14 (4.2)
AST increased	97 (30.3)	13 (4.1)	24 (7.2)	2 (0.6)
Nausea	91 (28.4)	3 (0.9)	59 (17.7)	2 (0.6)
WBC decreased	89 (27.8)	63 (19.7)	94 (28.2)	61 (18.3)
Decreased appetite	75 (23.4)	4 (1.3)	52 (15.6)	5 (1.5)
Vomiting	62 (19.4)	4 (1.3)	41 (12.3)	2 (0.6)
Alopecia	56 (17.5)	1 (0.3)	68 (20.4)	0 (0)
Dyspnoea	54 (16.9)	15 (4.7)	52 (15.6)	20 (6.0)
Neutropenia	44 (13.8)	38 (11.9)	51 (15.3)	45 (13.5)
Cough	42 (13.1)	3 (0.9)	63 (18.9)	2 (0.6)
Pyrexia	39 (12.2)	2 (0.6)	47 (14.1)	1 (0.3)
Stomatitis	36 (11.3)	4 (1.3)	26 (7.8)	1 (0.3)
Haemoglobin decreased	35 (10.9)	3 (0.9)	46 (13.8)	7 (2.1)
Constipation	22 (6.9)	0 (0)	39 (11.7)	1 (0.3)

AEs = adverse events; ALT = alanine transaminase; AST = aspartate transaminase; WBC = white blood cell

Serious AEs in LUME-Lung 1

Serious AEs (SAEs) in the adenocarcinoma population were similar between treatment groups (34.7% for nintedanib plus docetaxel vs 32.1% for placebo plus docetaxel)(3). The incidence of grade ≥3 SAEs was also comparable. The SAEs reported more frequently (≥1% difference) in the nintedanib plus docetaxel arm compared with the placebo plus docetaxel arm included febrile neutropenia (5.6% vs 1.8%), malignant neoplasm progression (3.8% vs 2.4%), atrial fibrillation (1.3% vs 0.0%), asthenia (1.6% vs 0.6%), respiratory failure (1.6% vs 0.3%) and sepsis (1.3% vs 0.3%,

Table 63)(5).

Table 63: SAEs occurring with an incidence ≥1% in either treatment arm in the adenocarcinoma population(5)

	Nintedanib plus docetaxel, n (%)		Placebo plus docetaxel, n (%)	
	Any grade, n (%)	Grade ≥3, n (%)	Any grade, n (%)	Grade ≥3, n (%)
Patients with SAEs	111 (34.7)	100 (31.3)	107 (32.1)	92 (27.6)
Febrile neutropenia	18 (5.6)	18 (5.6)	6 (1.8)	6 (1.8)
Malignant neoplasm progression	12 (3.8)	12 (3.8)	8 (2.4)	7 (2.1)
Dyspnoea	9 (2.8)	8 (2.5)	18 (5.4)	16 (4.8)
Pneumonia	9 (2.8)	7 (2.2)	12 (3.6)	6 (1.8)
Diarrhoea	6 (1.9)	5 (1.6)	7 (2.1)	6 (1.8)
General physical health deterioration	6 (1.9)	6 (1.9)	5 (1.5)	4 (1.2)
Neutropenia	6 (1.9)	5 (1.6)	11 (3.3)	11 (3.3)
Asthenia	5 (1.6)	4 (1.3)	2 (0.6)	1 (0.3)
Respiratory failure	5 (1.6)	5 (1.6)	1 (0.3)	1 (0.3)
Vomiting	5 (1.6)	2 (0.6)	4 (1.2)	2 (0.6)
Atrial fibrillation	4 (1.3)	3 (0.9)	0 (0)	0 (0)
Chest pain	4 (1.3)	3 (0.9)	6 (1.8)	5 (1.5)
Pleural effusion	4 (1.3)	4 (1.3)	6 (1.8)	4 (1.2)
Sepsis	4 (1.3)	4 (1.3)	1 (0.3)	1 (0.3)
Pyrexia	2 (0.6)	0 (0)	4 (1.2)	0 (0)

SAEs = serious adverse events

AEs of special interest in LUME-Lung 1

AEs of special interest (AESIs) were categorised by pooling Medical Dictionary for Drug Regulatory Activities (MEdDRA) terms and by using Standardised MedDRA Queries (SMQs).

AESIs were defined prior to the lock of the database for the primary PFS analysis. The following AESIs were analysed:(5)

- Listed AEs/possible side effects of nintedanib – diarrhoea, liver-enzyme elevations, nausea, vomiting, abdominal pain, fatigue, dehydration.
- Potential class effects of VEGFR inhibitors – perforations (gastrointestinal and non-gastrointestinal), bleeding, (including respiratory bleeding), thromboembolism (venous arterial), hypertension.
- Potential association/complication of AEs – association/complication of diarrhoea and vomiting with dehydration and renal failure, association/complication of liver-enzyme elevations with liver failure and hepatitis.
- Potential interaction with concomitant chemotherapy – mucositis, peripheral neuropathies, myelotoxicity (neutropenia, thrombocytopenia, anaemia, infections, febrile neutropenia, pneumonia, sepsis).
- AEs listed with other angiogenesis inhibitors (such as bevacizumab, sorafenib and dasatinib)(111-113) – pulmonary hypertension, osteonecrosis, ovarian failure, hypothyroidism, skin disorders (hand-foot syndrome, rash, cutaneous serious skin reactions).
- Cardiac events – cardiac arrest, cardiac failure, sudden death, cardiac arrhythmias, myocardial infarction.
- Other AEs of interest – interstitial lung disease, photosensitivity.

In the adenocarcinoma population of LUME-Lung 1, AESIs occurred at a frequency of 92.2% (65.9% CTCAE grade ≥3) in the nintedanib plus docetaxel arm, compared with 87.4% (58.3% CTCAE grade ≥3) in the placebo plus docetaxel arm(5).

AESIs related to nintedanib

The AESIs possibly related to nintedanib occurring with an incidence of more than ≥5% in the nintedanib plus docetaxel arm compared with the placebo plus docetaxel arm were diarrhoea (43.4% vs 24.6%), liver-related investigations (42.8% vs 14.7%), nausea 28.4% vs 17.7%), and vomiting (19.4% vs 12.3%, Table 64). Most of these AESIs were of grade <3(5).

Table 64: Summary of AESIs possibly related to nintedanib in LUME-Lung 1 in the adenocarcinoma population(5)

	Nintedanib plus docetaxel, n (%)		Placebo plus docetaxel, n (%)	
	Any grade, n (%)	Grade ≥3, n (%)	Any grade, n (%)	Grade ≥3, n (%)
Patients with any AESIs	295 (92.2)	211 (65.9)	291 (87.4)	194 (58.3)
Liver-related investigations[1]	137 (42.8)	49 (15.3)	49 (14.7)	6 (1.8)
Liver-related investigations (specific)[2]	130 (40.6)	43 (13.4)	42 (12.6)	5 (1.5)
Diarrhoea	139 (43.4)	20 (6.3)	82 (24.6)	12 (3.6)
Nausea	91 (28.4)	3 (0.9)	59 (17.7)	2 (0.6)
Vomiting	62 (19.4)	4 (1.3)	41 (12.3)	2 (0.6)
Fatigue	127 (39.7)	24 (7.5)	123 (36.9)	16 (4.8)
Abdominal pain	32 (10.0)	2 (0.6)	28 (8.4)	1 (0.3)
Dehydration	6 (1.9)	2 (0.6)	0 (0)	0 (0)

AESIs = Adverse events of special interest

[1] The SMQ term liver related investigations included the following preferred terms: alanine aminotransferase increased, aspartate aminotransferase increased, blood alkaline phosphatase increased, gamma−glutamyltransferase increased, hyperbilirubinaemia, hypoalbuminaemia, blood bilirubin increased, hepatic enzyme increased, ascites, hepatic function abnormal, hepatic pain, transaminases increased, bilirubin conjugated increased, liver function test abnormal

[2] Liver related investigations (specific) used a special search category to identify the following preferred terms: alanine aminotransferase increased, aspartate aminotransferase increased, hyperbilirubinaemia, blood bilirubin increased, hepatic enzyme increased, hepatic function abnormal, transaminases increased, liver function test abnormal

Note: Some events contribute to more than one special interest category. Patients with such AEs were counted in each of the AESI categories but were counted only once in the overall number of patients with AESI

The risk ratio for any grade liver-related investigations (specific) in patients with adenocarcinoma was 3.22 (95% CI; 2.36–4.40), risk difference 0.28 (95% CI; 0.22–0.34). The risk ratio for grade ≥3 liver-related investigations (specific) in patients with adenocarcinoma was 8.95 (95% CI; 3.59–22.31), risk difference 0.12 (95% CI; 0.08–0.16).(110)

AESIs related to VEGFR inhibitor class effects

The incidence of class effects typically associated with anti-angiogenic agents, such as hypertension, bleeding, perforation and thromboembolism, in patients with tumours of adenocarcinoma histology, was low and largely balanced across treatment groups(5). None of these AESIs occurred at an incidence of more than ≥5% in the nintedanib plus docetaxel arm compared with the placebo plus docetaxel arm. The most notable difference between the nintedanib and placebo treatment arms was observed for any grade hypertension, with a higher frequency in the nintedanib plus docetaxel arm (3.4% vs 0.6%; Table 65). However, the incidence of grade ≥3 hypertension was balanced across arms (0.9% vs 0.6%)(5).

Table 65: Summary of AESIs related to VEGFR inhibitor class effects in LUME-Lung 1 in the adenocarcinoma population(5)

	Nintedanib plus docetaxel, n (%)		Placebo plus docetaxel, n (%)	
	Any grade, n (%)	Grade ≥3, n (%)	Any grade, n (%)	Grade ≥3, n (%)
Patients with any AESIs	295 (92.2)	211 (65.9)	291 (87.4)	194 (58.3)
Perforation				
Gastrointestinal	1 (0.3)	1 (0.3)	1 (0.3)	1 (0.3)
Non-gastrointestinal	4 (1.3)	1 (0.3)	1 (0.3)	1 (0.3)
Hypertension	11 (3.4)	1 (0.3)	2 (0.6)	1 (0.3)
Bleeding	35 (10.9)	4 (1.3)	37 (11.1)	5 (1.5)
Respiratory bleeding	15 (4.7)	2 (0.6)	20 (6.0)	3 (0.9)
Thromboembolic events	17 (5.3)	8 (2.5)	18 (5.4)	11 (3.3)
Arterial thromboembolism	3 (0.9)	3 (0.9)	7 (2.1)	3 (0.9)
Venous thromboembolism	9 (2.8)	3 (0.9)	4 (1.2)	2 (0.6)

AESIs = Adverse events of special interest; VEGFR = vascular endothelial growth factor receptor
Some events contribute to more than one special interest category. Patients with such AEs were counted in each of the AESI categories but were counted only once in the overall number of patients with AESI

The risk ratio for any grade gastrointestinal perforation in patients with adenocarcinoma was 1.04 (95% CI; 0.07–16.57), risk difference 0.00 (95% CI; -0.01–0.01). The risk ratio for grade ≥3 gastrointestinal perforation in patients with adenocarcinoma was 1.04 (95% CI; 0.07–16.57), risk difference 0.00 (95% CI; -0.01–0.01).(110)

The risk ratio for any grade non-gastrointestinal perforation in patients with adenocarcinoma was 4.16 (95% CI; 0.47–37.04), risk difference 0.01 (95% CI; -0.00–0.02). The risk ratio for grade ≥3 non-gastrointestinal perforation in patients with adenocarcinoma was 1.04 (95% CI; 0.07–16.57), risk difference 0.00 (95% CI; -0.01–0.01).(110)

XThe risk ratio for any grade bleeding in patients with adenocarcinoma was 0.98 (95% CI; 0.64–1.52), risk difference -0.00 (95% CI; -0.05–0.05). The risk ratio for grade ≥3 bleeding in patients with adenocarcinoma was 0.83 (95% CI; 0.23–3.07), risk difference -0.00 (95% CI; -0.02–0.02).(110)

XThe risk ratio for any grade venous thromboembolism in patients with adenocarcinoma was 2.34 (95% CI; 0.73–7.53), risk difference 0.02 (95% CI; -0.01–0.04). The risk ratio for grade ≥3 venous thromboembolism in patients with adenocarcinoma was 1.56 (95% CI; 0.26–9.28), risk difference 0.00 (95% CI; -0.01–0.02).(110)

The risk ratio for any grade aterial thromboembolism in patients with adenocarcinoma was 0.45 (95% CI; 0.12–1.71), risk difference -0.01 (95% CI; -0.03–0.01). The risk ratio for grade ≥3 arterial thromboembolism in patients with adenocarcinoma was 1.04 (95% CI; 0.21–5.12), risk difference 0.00 (95% CI; -0.01–0.02).(110)

AESIs based on potential associations/complications of AEs

The incidence of AESIs based on potential association/complications of AEs was balanced across treatment arms. None of these AESIs occurred at an incidence of more than ≥5% in the nintedanib plus docetaxel arm compared with the placebo plus docetaxel arm. The only AESI to occur at a slightly higher rate in the nintedanib plus docetaxel arm compared with the placebo plus docetaxel arm was any grade dehydration (1.9% vs 0%, respectively; Table 66). However, the incidence of grade ≥3 dehydration was balanced across arms (0.6 % vs 0%)(5).

Table 66: Summary of AESIs based on potential association/complications of AEs in LUME-Lung 1 in the adenocarcinoma population(5)

	Nintedanib plus docetaxel, n (%)		Placebo plus docetaxel, n (%)	
	Any grade, n (%)	Grade ≥3, n (%)	Any grade, n (%)	Grade ≥3, n (%)
Patients with any AESIs	295 (92.2)	211 (65.9)	291 (87.4)	194 (58.3)
Dehydration	6 (1.9)	2 (0.6)	0 (0)	0 (0)
Hepatic failure	3 (0.9)	3 (0.9)	1 (0.3)	0 (0)
Renal failure	3 (0.9)	1 (0.3)	1 (0.3)	0 (0)

AESIs = Adverse events of special interest
Some events contribute to more than one special interest category. Patients with such AEs were counted in each of the AESI categories but were counted only once in the overall number of patients with AESI

The risk ratio for any grade hepatic failure in patients with adenocarcinoma was 3.12 (95% CI; 0.33–29.86), risk difference 0.01 (95% CI; -0.01–0.02). Data for the risk ratio for grade ≥3 hepatic failure was not available due to a lack of events, the risk difference was 0.01 (−0.00, 0.02).(110)

AESIs related to potential interaction with concomitant chemotherapy

The incidence of AESIs related to potential interaction with concomitant chemotherapy was balanced across treatment arms. The only AESI to occur more frequently (≥5% difference) in the nintedanib plus docetaxel arm compared with the placebo plus docetaxel arm was any grade mucositis (16.6% nintedanib vs 11.4% placebo) (Table 67). However, the incidence of grade ≥3 mucositis was balanced across arms (1.3 % vs 0.6%)(5).

Table 67: Summary of AESIs related to potential interaction with concomitant chemotherapy in LUME-Lung 1 in the adenocarcinoma population(5)

	Nintedanib plus docetaxel, n (%)		Placebo plus docetaxel, n (%)	
	Any grade, n (%)	Grade ≥3, n (%)	Any grade, n (%)	Grade ≥3, n (%)
Patients with any AESIs	295 (92.2)	211 (65.9)	291 (87.4)	194 (58.3)
Neutropenia	176 (55.0)	161 (50.3)	178 (53.5)	155 (46.5)
Febrile neutropenia	24 (7.5)	23 (7.2)	15 (4.5)	15 (4.5)
Infection	84 (26.3)	21 (6.6)	73 (21.9)	18 (5.4)
Pneumonia	33 (10.3)	13 (4.1)	37 (11.1)	13 (3.9)
Sepsis	4 (1.3)	4 (1.3)	2 (0.6)	2 (0.6)
Thrombocytopenia	16 (5.0)	4 (1.3)	12 (3.6)	4 (1.2)
Anaemia	51 (15.9)	8 (2.5)	65 (19.5)	10 (3.0)
Peripheral neuropathies	61 (19.1)	9 (2.8)	55 (16.5)	2 (0.6)
Mucositis	53 (16.6)	4 (1.3)	38 (11.4)	2 (0.6)

AESIs = Adverse events of special interest
Some events contribute to more than one special interest category. Patients with such AEs were counted in each of the AESI categories but were counted only once in the overall number of patients with AESI

The risk ratio for any grade neutropenia in patients with adenocarcinoma was 1.03 (95% CI; 0.89–1.18), risk difference 0.02 (95% CI; -0.06–0.09). The risk ratio for grade ≥3 neutropenia in patients with adenocarcinoma was 1.08 (95% CI; 0.92–1.27), risk difference 0.04 (95% CI; -0.04–0.11).(110)

The risk ratio for any grade febrile neutropenia in patients with adenocarcinoma was 1.67 (95% CI; 0.89–3.12), risk difference 0.03 (95% CI; -0.01–0.07). The risk ratio for grade ≥3 febrile neutropenia in patients with adenocarcinoma was 1.60 (95% CI; 0.85–3.00), risk difference 0.03 (95% CI; -0.01–0.06).(110)

The risk ratio for any grade sepsis in patients with adenocarcinoma was 2.08 (95% CI; 0.38–11.28), risk difference 0.01 (95% CI; -0.01–0.02).(110)

AESIs selected based on competitor labelling

The incidence of selected AESIs based on competitor labelling was generally balanced between treatment arms (Table 68)(5). The only AESI occurring more frequently (≥5% difference) in the nintedanib plus docetaxel arm compared with the placebo plus docetaxel arm was any grade cutaneous skin reactions (15.6% vs 10.5%). The number of patients experiencing any grade rash was slightly higher in the nintedanib plus docetaxel arm (12.5% vs 8.7%) (Table 68). The incidence of both grade ≥3 cutanous skin reactions and grade ≥3 rash was however balanced across arms (1.3 % vs 0.6% for cutaneous skin reactions and 0.3% and 0% for rash)(5).

Table 68: Summary of AESIs related to known effects of other VEGFR inhibitors in LUME-Lung 1 in the adenocarcinoma population(5)

	Nintedanib plus docetaxel, n (%)		Placebo plus docetaxel, n (%)	
	Any grade, n (%)	Grade ≥3, n (%)	Any grade, n (%)	Grade ≥3, n (%)
Patients with any AESIs	295 (92.2)	211 (65.9)	291 (87.4)	194 (58.3)
Rash	40 (12.5)	1 (0.3)	29 (8.7)	0 (0)
Cutaneous serious skin reactions	50 (15.6)	4 (1.3)	35 (10.5)	2 (0.6)
Hand-foot syndrome	1 (0.3)	0 (0)	1 (0.3)	0 (0)
Pulmonary hypertension	0 (0)	0 (0)	0 (0)	0 (0)
Osteonecrosis	1 (0.3)	0 (0)	0 (0)	0 (0)
Hypothyroidism	1 (0.3)	0 (0)	0 (0)	0 (0)

AESIs = Adverse events of special interest
Some events contribute to more than one special interest category. Patients with such AEs were counted in each of the AESI categories but were counted only once in the overall number of patients with AESIs

AESIs related to cardiac events

No AESIs related to cardiac events occurred at an incidence of more than ≥5% in the nintedanib plus docetaxel arm compared with the placebo plus docetaxel arm. Any grade cardiac arrhythmias occurred at a slightly higher incidence in the nintedanib plus docetaxel treatment arm compared with the placebo plus docetaxel arm (11.6% vs 7.5%, respectively). However, the incidence of grade ≥3 cardiac arrhythmias was balanced across arms (2.2 % vs 1.5%). The incidence of other AESIs related to cardiac events was comparable between treatment arms (Table 69)(5).

Table 69: Summary of AESIs related to cardiac events in LUME-Lung 1 in the adenocarcinoma population(5)

	Nintedanib plus docetaxel, n (%)		Placebo plus docetaxel, n (%)	
	Any grade, n (%)	Grade ≥3, n (%)	Any grade, n (%)	Grade ≥3, n (%)
Patients with any AESIs	295 (92.2)	211 (65.9)	291 (87.4)	194 (58.3)
Cardiac failure	25 (7.8)	2 (0.6)	22 (6.6)	2 (0.6)
Cardiac failure (tailored)	0 (0)	0 (0)	2 (0.6)	2 (0.6)
Cardiac arrest	2 (0.6)	2 (0.6)	1 (0.3)	1 (0.3)
Cardiac arrhythmias	37 (11.6)	7 (2.2)	25 (7.5)	5 (1.5)
Myocardial infarction	4 (1.3)	1 (0.3)	4 (1.2)	3 (0.9)

AESIs = Adverse events of special interest
Some events contribute to more than one special interest category. Patients with such AEs were counted in each of the AESI categories but were counted only once in the overall number of patients with AESI

The risk ratio for any grade myocardial infarction in patients with adenocarcinoma was 1.04 (95% CI; 0.26–4.13), risk difference 0.00 (95% CI; -0.02–0.02). The risk ratio for grade ≥3 myocardial infarction in patients with adenocarcinoma was 0.35 (95% CI; 0.04–3.32), risk difference -0.01 (95% CI; -0.02–0.01).(110)

Other AESIs

The incidence of other AESIs in the adenocarcinoma population was comparable across treatment groups (Table 70)(5).

Table 70: Summary of other AESIs in LUME-Lung 1 in the adenocarcinoma population(5)

	Nintedanib plus docetaxel, n (%)		Placebo plus docetaxel, n (%)	
	Any grade, n (%)	Grade ≥3, n (%)	Any grade, n (%)	Grade ≥3, n (%)
Patients with any AESIs	295 (92.2)	211 (65.9)	291 (87.4)	194 (58.3)
Interstitial lung disease	4 (1.3)	0 (0)	1 (0.3)	1 (0.3)
Photosensitivity conditions	1 (0.3)	0 (0)	2 (0.6)	0 (0)
Anaphylactic reaction	0 (0)	0 (0)	1 (0.3)	1 (0.3)

AESIs = Adverse events of special interest
Some events contribute to more than one special interest category. Patients with such AEs were counted in each of the AESI categories but were counted only once in the overall number of patients with AESI

6.9.3 Give a brief overview of the safety of the technology in relation to the decision problem.

In the LUME-Lung 1 trial, adenocarcinoma patients in the nintedanib plus docetaxel arm received treatment for 1.2 months longer, on average, compared with placebo plus docetaxel (4.2 months vs 3.0 months)(5). The addition of nintedanib did not negatively impact the median duration of docetaxel administration, with a median of 5 cycles being administered in the nintedanib arm and 4 cycles in the placebo arm(5). Safety data should be seen in the context of the longer time on treatment and exposure to the study drugs in the nintedanib plus docetaxel arm compared with the placebo plus docetaxel arm.

Dose reductions (21.9% vs 6.6%) and dose interruptions (52.2% vs 41.4%) were each more common in the nintedanib plus docetaxel arm compared with the placebo plus docetaxel arm(5). Nintedanib dose reductions due to the most frequently reported AEs (diarrhoea, nausea and vomiting) were required in only a small proportion of adenocarcinoma patients, though these rates were higher in the nintedanib plus docetaxel arm (diarrhoea: 8.1%, nausea: 1.3% and vomiting: 2.2%) compared with the placebo plus docetaxel arm (diarrhoea: 3.3%, nausea: 0.3% and vomiting: 0.6%)(5).

AEs leading to permanent discontinuation of study medication were higher in the nintedanib plus docetaxel arm than in the placebo plus docetaxel arm for all AEs overall (20.9% nintedanib and 17.7% placebo)(3).

Treatment with nintedanib in combination with docetaxel did lead to additional AEs compared with docetaxel treatment alone. These AEs were consistent with the known safety profile for the drug(3). Diarrhoea (43.4%), nausea (28.4%) and vomiting (19.4%) were the most common GI AEs among adenocarcinoma patients treated with nintedanib plus docetaxel(3). Typically, these were mild (grade ≤2) and led to a permanent discontinuation of nintedanib in ≤1% of patients, indicating that they were successfully managed by dose reduction, dose interruption and/or symptomatic treatment. Other commonly reported AEs associated with nintedanib treatment included ALT/AST increase (37.8% and 30.3% respectively)(3). These were generally reversible and led to permanent nintedanib discontinuation in <2% of patients(5).

The proportion of adenocarcinoma patients with SAEs was comparable across treatment arms (34.7% and 32.1% for nintedanib plus docetaxel and placebo plus docetaxel, respectively)(3). The SAEs reported more frequently (≥1% difference) in the nintedanib plus docetaxel arm compared with the placebo plus docetaxel arm included febrile neutropenia (5.6% vs 1.8%), malignant neoplasm progression (3.8% vs 2.4%), asthenia (1.6% vs 0.6%), respiratory failure (1.6% vs 0.3%), vomiting (1.6% vs 1.2%), atrial fibrillation (1.3% vs 0%) and sepsis (1.3% vs 0.3%)(5).

The number of adenocarcinoma patients with AEs leading to death was 17.5% in the nintedanib arm and 9.6% in the placebo arm(3). However, the number considered drug-related was low, but was slightly higher in the nintedanib plus docetaxel arm than in the placebo plus docetaxel arm (1.9% vs 0.3% in the nintedanib plus docetaxel and the placebo plus docetaxel arm respectively)(5). Moreover, when considering only those fatal events with a start and end date in the on-treatment period, the time-adjusted incidence rate ratio was 1.54 (95% CI 0.96-2.46)(5).

AESIs included those AEs that are potential class effects of VEGFR inhibitors, such as perforations (gastrointestinal and non-gastrointestinal), bleeding, (including respiratory bleeding), thromboembolism (venous arterial), hypertension. The incidence of each of these was <6% in the nintedanib plus docetaxel arm (except any grade bleeding; 10.9%). None of these AESIs occurred at an incidence of more than ≥5% in the nintedanib plus docetaxel arm compared with the placebo plus docetaxel(5).

6.10 Interpretation of clinical evidence

6.10.1 Please provide a statement of principal findings from the clinical evidence highlighting the clinical benefit and harms from the technology.

In the Phase III study, LUME-Lung 1, treatment with nintedanib 200 mg twice daily (bid) plus docetaxel improved centrally assessed progression-free survival (PFS; primary endpoint) versus placebo plus docetaxel (HR=0.79; 95% CI: 0.68–0.92; p=0.002) for second-line NSCLC patients. Examination of OS, the key secondary endpoint, showed that nintedanib significantly prolonged survival in patients with adenocarcinoma histology (HR=0.83; 95% CI: 0.70–0.99; p=0.036)(3).

Self-reported QoL assessments by EORTC questionnaires revealed that no significant differences in cough, dyspnea or pain were observed in patients receiving nintedanib plus docetaxel compared with placebo plus docetaxel(59). There were trends towards improvements in TTD for global health status/QoL in patients with adenocarcinoma. QoL scores for nausea and vomiting, appetite loss and diarrhea were worsened in patients who received nintedanib plus docetaxel compared with those who received placebo plus docetaxel. Global health status/QoL remained unchanged in adenocarcinoma patients (HR=0.8695% CI: 0.71–1.05). Therefore the improvements seen in terms of PFS and OS in the adenocarcinoma patients were achieved without substantial alterations in self-reported QoL(59).

As detailed in Section 6.9 in the adenocarcinoma patient population of LUME-Lung 1 treatment-related AEs were more frequent with nintedanib plus docetaxel (81.3%) than with placebo plus docetaxel (72.4%). Similarly, the proportion of patients with AEs grade ≥3 was higher in the nintedanib plus docetaxel arm (75.9%) than in the placebo plus docetaxel arm (68.5%). SAEs, however, were comparable across both arms (34.7% for nintedanib plus docetaxel and 32.1% for placebo plus docetaxel, respectively)(5). The number of patients with AEs leading to death that were considered drug-related was low in patients with adenocarcinoma (1.9% vs 0.3% in the nintedanib plus docetaxel and the placebo plus docetaxel arm, respectively)(5).

While nintedanib demonstrated a significant and clinically meaningful OS benefit in adenocarcinoma patients, more non-PD fatal AEs occurred in the nintedanib arm. This

finding is confounded by the fact that (1) the extent of exposure was longer on nintedanib plus docetaxel compared to docetaxel alone and (2) the analysis focusing on the on-treatment fatal AEs resulted in a skewed view of the deaths that occurred during the study.

The imbalance in fatal AEs is due to an increased number of patients experiencing fatal sepsis and respiratory failure. The higher exposure to docetaxel may have contributed, at least in part, to the higher incidence of fatal AEs of sepsis caused by neutropenia in the nintedanib arm through the known myelotoxic effect of docetaxel. Consequently neutropenia and sepsis are considered possible side effects of nintedanib therapy in combination with docetaxel and are regarded as important identified risks for future monitoring and ongoing safety surveillance. The term respiratory failure was used by the investigator to document a pathophysiological endpoint (terminal status of the patient) rather than a respiratory disease entity per se. Further review of PD and non-PD deaths occurring during the entire observation period revealed no other safety pattern suggestive of nintedanib associated toxicities(5).

AEs that are potential class effects of VEGFR inhibitors, e.g. perforations (gastrointestinal and non-gastrointestinal), bleeding, (including respiratory bleeding), thromboembolism (venous arterial), hypertension were reported in <6% of patients in the nintedanib plus docetaxel arm (except bleeding; 10.9%). None of these AESIs occurred at an incidence of ≥5% in the nintedanib plus docetaxel arm compared with the placebo plus docetaxel arm(5).

6.10.2 Please provide a summary of the strengths and limitations of the clinical-evidence base of the intervention.

LUME-Lung 1 recruited patients with PS 0 and 1 to maximise fitness for docetaxel administration. This may be a limitation in the generalisability of the clinical evidence for patients with poorer performance status or those not fit for docetaxel.

Ideally, identification of biomarker-defined subgroups of patients for which there is a treatment by biomarker interaction will maximise the benefit to toxicity ratio for even moderately effective therapies. However, as described in the literature, efforts to identify biomarkers predictive of benefit for anti-angiogenic therapies have met with limited success for NSCLC patients thus far(114).

Aside from histology (non-squamous for bevacizumab) no prospectively validated genetic or biologic marker has been identified during the last decade of clinical research for the treatment effect of anti-angiogenic compounds for lung cancer patients. Nintedanib is an anti-angiogenic agent; potential biomarkers for efficacy of nintedanib have been investigated in preclinical models using standard human tumour xenografts grown subcutaneously in nude mice, but with limited success. One factor that is important to consider from these models is that because anti-angiogenic therapy targets primarily the tumour stroma derived from the murine host, it is difficult to translate these data to the cancer patient setting. Additionally, as a highly preserved host function, angiogenesis is not subject to the same genetic variability that is observed for tumour-related molecular markers(114).

Given the lack of known and valid molecular biomarkers to predict response to anti-angiogenic therapy, analysis of data from LUME-Lung 1 and 2 focused on the identification of clinical markers predictive of clinical benefit in response to nintedanib treatment. Time since start of first-line therapy was identified as a predictive variable for PFS and OS in adenocarcinoma patients. A cut-off of 9 months since start of first-line line therapy defines a population of patients with poor prognosis who show significant benefit from the addition of nintedanib to standard second-line chemotherapy(114).

Treatment algorithms for NSCLC patients with adenocarcinoma histology have changed in recent years and maintenance treatment with pemetrexed after first-line has been shown to significantly improve survival in this patient population. In LUME-Lung 1 patients who received pemetrexed as maintenance therapy prior to enrolment into the trial was low. This is not unexpected since recruitment into the study ended in February 2011 and the registration of pemetrexed as maintenance treatment was approved in September 2011. Although this may be viewed as a potential limitation of this study, it nevertheless is representative of clinical practice at the time. The on-study frequency of patients who received pemetrexed as maintenance therapy in first-line was balanced between the arms; 14 patients in the placebo arm compared with 13 patients in the nintedanib arm(114).

As shown in Table 71, the data suggest that treatment with nintedanib in combination with docetaxel would lead to an improvement in OS regardless of whether patients did or did not receive maintenance therapy with pemetrexed (HR: 0.78, 95% CI: 0.30-2.07 for patients treated with pemetrexed maintenance vs HR: 0.84, 95% CI: 0.70-1.00, for patients not treated with pemetrexed maintenance). The interaction p-value was not significant. The median OS was 12.8 months on the placebo arm to 18.9 months on the nintedanib arm for patients who were treated with pemetrexed maintenance. This is in line with the data for patients who had pemetrexed as first line treatment(114).

While it is noted that the number of patients with pemetrexed maintenance treatment available for analysis is too low to allow definitive conclusions, there is no evidence of any less activity in patients that received pemetrexed in combination with a platinum agent followed by pemetrexed maintenance(114).

Table 71: Overall survival by pemetrexed maintenance therapy in first-line – randomised set (LUME-Lung 1) adenocarcinoma patients – final OS snapshot(114)

	No maintenance pemetrexed in first-line		**Maintenance pemetrexed in first -line**	
Final OS analysis snapshot	**Placebo**	**Nintedanib**	**Placebo**	**Nintedanib**
Patients, n (%)	322 (100.0)	309 (100.0)	14 (100.0)	13 (100.0)
Patients with OS event, n (%)	266 (82.6)	250 (80.9)	10 (71.4)	9 (69.2)
Median*OS (months)	10.0	12.6	12.8	18.9
HR#(95% CI)	0.84 (0.70,1.00)		0.78 (0.30,2.07)	
Interaction between treatment and subgroup variable^	**0.7162**			

OS = Overall survival.
* Medians are calculated from an unadjusted Kaplan–Meier curve for each treatment arm.
# If HR is below 1 then favours nintedanib. Hazard Ratio and confidence interval obtained from a proportional–hazards model stratified by baseline ECOG PS (0 vs. 1), brain metastases at baseline (yes vs. no) and prior treatment with bevacizumab (yes vs. no).
^ Test of interaction derived by fitting a proportional hazards model with and without treatment by taxane in first-line interaction and comparing the difference in log–likelihoods.
One patient (135301) has a baseline ECOG PS of 2.

The clinical evidence favours the use of nintedanib in combination with docetaxel in second-line adenocarcinoma compared to docetaxel alone in this patient group who have great unmet medical need.

6.10.3 Please provide a brief statement of the relevance of the evidence base to the decision problem. Include a discussion of the relevance of the outcomes assessed in clinical trials to the clinical benefits experienced by patients in practice.

There is a high unmet medical need to improve the treatment options for patients with advanced NSCLC in the second-line setting. Over the last decade of clinical research, no statistically significant prolongation of OS has been reported in the literature in the second-line treatment setting for any of the major tumour histologies in lung cancer versus an active comparator.

Within these histologies, no therapeutic improvements have been achieved for patients most refractory to first-line therapy. In this setting of advanced second-line NSCLC patients, the treatment effect seen with addition of nintedanib to docetaxel in the adenocarcinoma patients is of clinical relevance. The treatment benefit was consistent across most of the predefined subgroups.

In line with these results, related study endpoints such as PFS by investigator, disease control rate and change in tumour size showed significant improvement. The final OS analysis in the pre-defined population of adenocarcinoma patients showed a statistically significant improvement in OS which translated into a 17% reduction in the risk of death, a median OS improvement of 2.3 months and a significant increase in the one-year and two-year survival rates. The robustness of the treatment effect of nintedanib on PFS and OS in the adenocarcinoma patients was confirmed by analyses of subgroups defined by demographic and baseline characteristics (e.g. ECOG PS, previous treatment with bevacizumab, first-line pemetrexed or taxanes, sex, age, race, smoking status, geographical region, and best response to prior anticancer therapy) which demonstrated a consistent treatment benefit across these patient subgroups.

Considering patients with tumours of adenocarcinoma histology, a statistically significant improvement in PFS in favour of the nintedanib plus docetaxel arm has been shown. At the time of the primary analysis, improvement in median PFS was 4.0 months in the nintedanib plus docetaxel arm vs 2.8 months in the placebo plus docetaxel arm. This improvement in PFS in adenocarcinoma patients is also observed at the time of the follow-up PFS analysis with a HR of 0.84, confirming the clinically meaningful benefit observed at the time of the pre-defined primary endpoint analysis. Furthermore, the prolongation of median PFS (4.2 months for the nintedanib arm vs. 2.8 months for the placebo arm), represents a 50% improvement with the addition of nintedanib to docetaxel. This represents the highest and statistically significant median PFS improvement as assessed by central independent review for second-line NSCLC adenocarcinoma patients thus far.

In addition, a statistically significant and clinically meaningful median improvement of 2.3 (12.6 vs 10.3) months was seen for the secondary endpoint OS with a HR of 0.83(5). This was supported by statistically significant improvement of other secondary endpoints, i.e. disease control (OR 1.93, $p<0.0001$) and change in tumour size ($p=0.0002$), and no difference in deterioration in HRQL compared to the control arm(5, 59). As such, the combination of nintedanib and docetaxel in the LUME-Lung 1 study is one of the first treatments to extend OS beyond a year in the second-line treatment setting. Median OS of more than one year for patients with adenocarcinoma has not been previously reported in the literature in the second-line treatment of NSCLC; median OS values of more than one year have only been reported previously in the first-line setting.

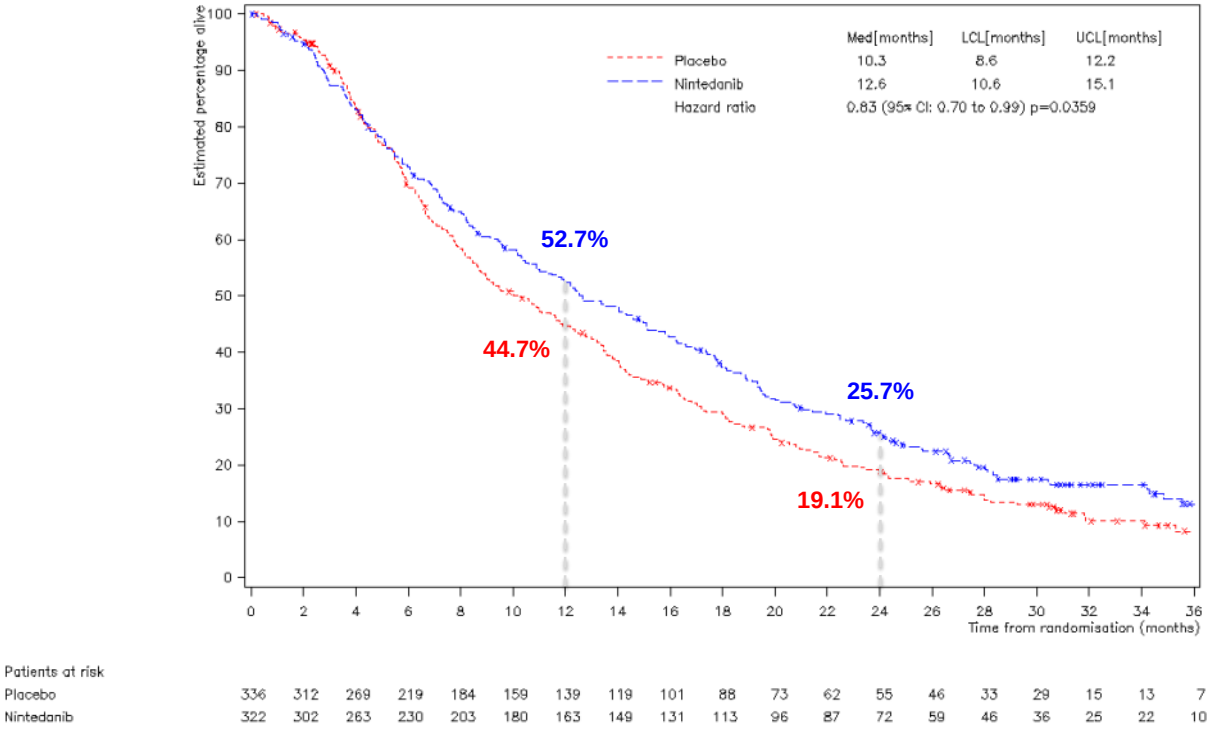
The 2.3 month improvement with nintedanib plus docetaxel compared to placebo plus docetaxel represents a 22% improvement in median OS. Of note, at the 75% percentile, the magnitude of OS improvement increased by 4.3 months from 19.9 months in the placebo arm to 24.2 months in the nintedanib arm(5).

The 2.3 month improvement in median OS is relevant when evaluated in the context of the recently published ASCO guidelines on clinically meaningful improvements in NSCLC developed by the ASCO Cancer Research Committee. While these guidelines do not distinguish between first- and second-line treatment, and while the authors make it clear that the target gains are aspirational and highlight the future promise of yet-to-be developed predictive biomarkers to select the appropriate patient populations, it is clear that the magnitude of treatment benefit observed for the adenocarcinoma population

treated with nintedanib in combination with docetaxel is unprecedented in the second-line setting of NSCLC.

As shown in Figure 22 below, the Kaplan Meier OS curves started to separate early and remained separated over the whole observation period leading to a significant improvement in the one year and two year survival rates. At 12 months, the estimated survival rates were 52.7% in the nintedanib- vs. 44.7% in the placebo arm (p= 0.044) and 25.7% vs. 19.1% (p= 0.051), at 24 months(5). As such, LUME LUNG 1 is the first trial to observe significant increases in 12 and 24 month survival rates in the second-line treatment of NSCLC. Importantly, subsequent treatments were balanced between the treatment arms and thus, could not have influenced the OS outcome.

Figure 22: Probability of OS in the adenocarcinoma population in LUME-Lung 1(5)



From a clinical perspective, it is also important that the treatment effect resulting from the addition of nintedanib to docetaxel is not the result of an underperforming control arm. This is substantiated by the median OS in the placebo plus docetaxel arm which was comparable to, or longer than, in historical trials investigating docetaxel in the second-line setting e.g. 10 months in the ZODIAC trial(49) , 7.9 months in the JMEI trial(48), or 7.5 and 5.7 months, respectively, in TAX 317(72) and TAX 320(71).

In addition, prolongation of OS was achieved without detriment on the overall HRQL in this palliative setting adding to the clinical significance of the OS findings.

The robustness of the OS treatment benefit of nintedanib plus docetaxel in adenocarcinoma patients and the clinical relevance with respect to other patient populations was confirmed by analyses of subgroups defined by demographic and baseline characteristics. The study was conducted at sites in the UK and is therefore representative of this patient population

6.10.4 Identify any factors that may influence the external validity of study results to patients in routine clinical practice; for example, how the technology was used in the trial, issues relating to the conduct of the trial compared with clinical practice, or the choice of eligible patients. State any criteria that would be used in clinical practice to select patients for whom treatment would be suitable based on the evidence submitted. What proportion of the evidence base is for the dose(s) given in the SPC?

The main factors are discussed in Section 6.10.2. In summary:

- Dose and dose reductions within the study results will be consistent with the SPC.
- Adult patients would be selected based on adenocarcinoma histology
- PS of patients greater than 1 were not included in the study for the reasons described in Section 6.10.2

7 Cost effectiveness

7.1 Published cost-effectiveness evaluations

Identification of studies

7.1.1 Describe the strategies used to retrieve relevant cost-effectiveness studies from the published literature and from unpublished data held by the manufacturer or sponsor. The methods used should be justified with reference to the decision problem. Sufficient detail should be provided to enable the methods to be reproduced, and the rationale for any inclusion and exclusion criteria used should be provided. The search strategy used should be provided as in section 10.10, appendix 10.

The objective of this search was to identify any previously published cost-effectiveness studies that could address the decision problem. This was done using a systematic review. The format of this systematic review has been reported to be in line with NICE's STA requirements as outlined in the NICE STA template. To date, NICE's requirements are the most stringent with respect to systematic review requirements.

The scope of this systematic review is to review all available published data on economic evaluations of second-line therapies for locally advanced or metastatic NSCLC that could inform a HTA submission, based on BI's second-line comparative trials of nintedanib.

The methodology followed is explained in Section 6.1 and Appendix 10.11. A single systematic literature review was performed for the clinical, cost-effectiveness, resource use and cost data, as well as studies reporting utility scores for health states within the model.

The cost-effectiveness studies' inclusion and exclusion criteria are shown in Table 72

Table 72: Cost-effectiveness Inclusion and Exclusion Criteria

	Inclusion Criteria	Exclusion Criteria
Population	RR NSCLC (receiving second-line chemotherapy or relapsed/refractory to first-line chemotherapy)	Any patient population other than RR NSCLC
Interventions	Any second-line chemotherapy for RR NSCLC: • Monotherapy • Combination therapy with other chemotherapy Other interventions that are considered standard care in the patient population that will be relevant to the economic model	Patients who were treatment-naïve or had received more than first-line therapy
Outcomes	Economic models: • Cost-utility analyses • Cost-effectiveness analyses • Cost-benefit analyses • Cost-minimisation analyses	No outcomes of interest included
Study design	Economic models: Economic studies	Not an economic model
Language restrictions	English language‡	Non-English language
Date	Economic models: 2002 onwards	Prior to the year 2002*
Country	Any	None

Quality assessment

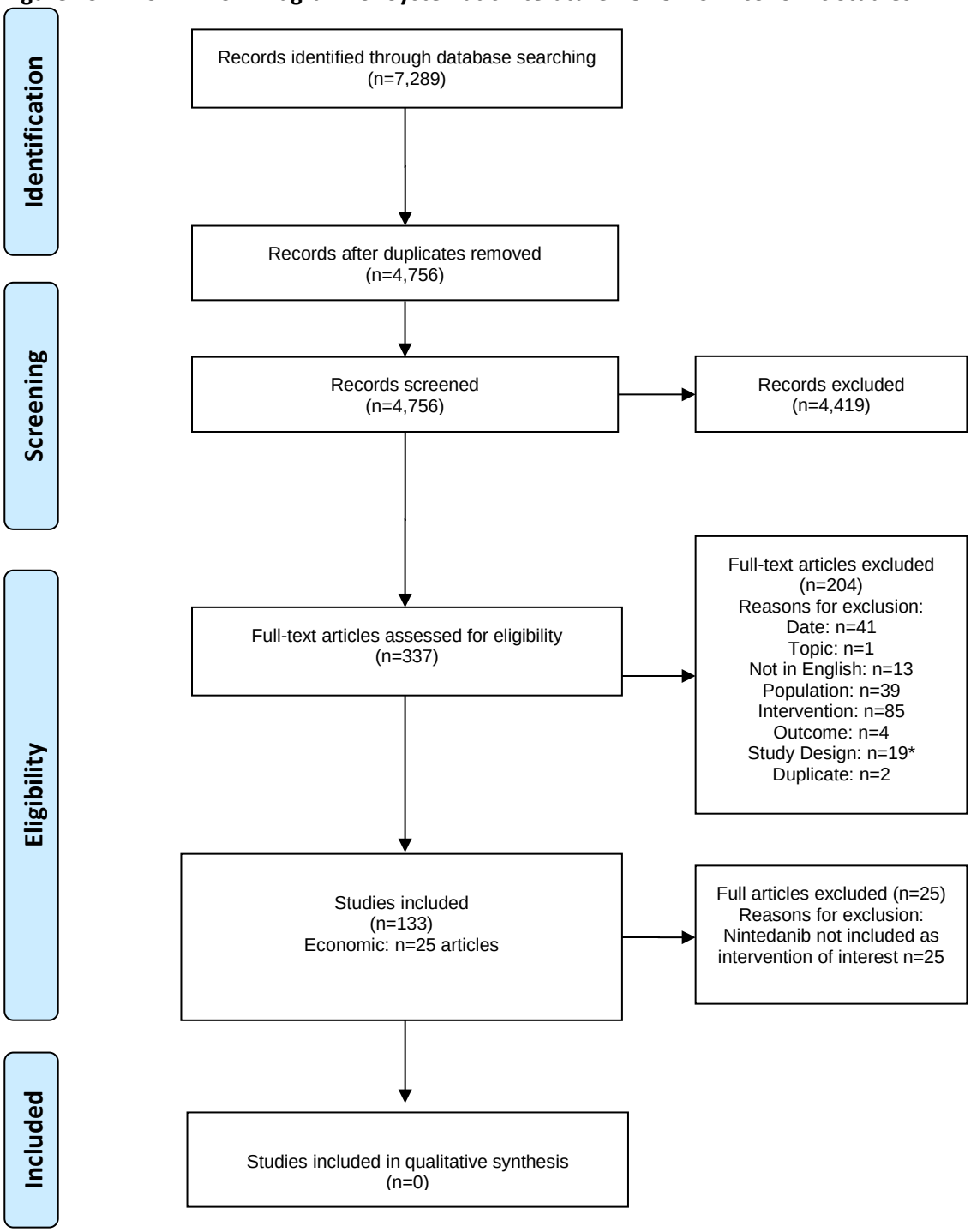
A quality assessment score was derived from that of Drummond (1997) to assess the quality of included economic models (115). The Drummond criteria were created to support the generation of high-quality, rigorous economic evaluations. They involve using a total of 36 questions to assess three broad areas of the studies, namely: study design; data collection; and analysis and interpretation of results. The result of the assessment process is a summary quality score on which models are judged to be either: high (++), moderate (+) or low quality (-). High-quality studies are considered to report clearly on almost all of the Drummond quality criteria questions, while studies of low quality do not report on most items. In this review, only studies in full-text form underwent a quality assessment because of the lack of details available for assessment in abstracts and posters.

Results

The search of the literature yielded 7,289 citations. De-duplication resulted in the removal of 2,533 overlapping citations. Following screening of the remaining 4,756 studies, 4,419 studies were excluded. Full text was obtained for the remaining 337 studies; 41 were excluded due to an incorrect date; 1 due to incorrect topic; 13 because the study was not in English; 39 because of incorrect population; 85 because the intervention did not match the original search criteria; 4 because of incorrect outcome; 19 because of incorrect study design and 2 because they were duplicates (total excluded = 204).

Of the remaining 133 studies, 25 were economic studies; however none of these included nintedanib as the intervention of interest. For this reason, the remaining 25 studies were excluded. The flow of studies in the systematic literature review is presented in Figure 23. Note this figure shows the initial search results of the entire literature review (as described in Section 6.1, Appendix 10.2 and Appendix 10.11), including searches of economic, resource use, utility and clinical searches as this was a combined search. The flow then demonstrates how studies were included or excluded according to the criteria relevant to the search of interest.

Figure 23: PRISMA Flow Diagram for Systematic Literature Review on Economic Studies



* The reference lists of the systematic reviews were assessed for additional relevant studies; no additional studies were identified.

As no studies relevant to the decision problem were identified, a de novo economic evaluation was required.

Description of identified studies

7.1.2 Provide a brief overview of each study, stating the aims, methods, results and relevance to decision-making in England and Wales. Each study's results should be interpreted in light of a critical appraisal of its methodology. When studies have been identified and not included, justification for this should be provided. If more than one study is identified, please present in a table as suggested below.

NA (see Section 7.1.1)

7.1.3 Please provide a complete quality assessment for each cost-effectiveness study identified. Use an appropriate and validated instrument, such as those of Drummond and Jefferson (1996)[4] or Philips et al. (2004)[5]. For a suggested format based on Drummond and Jefferson (1996), please see section 10.11, appendix 11.

NA (see Section 7.1.1)

[4] Drummond MF, Jefferson TO (1996) Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. British Medical Journal 313 (7052): 275–83.

[5] Philips Z, Ginnelly L, Sculpher M, et al. (2004) Quality assessment in decision-analytic models: a suggested checklist (Appendix 3). In: Review of guidelines for good practice in decision-analytic modelling in health technology assessment. Health Technology Assessment 8: 36.

7.2 De novo analysis

Patients

7.2.1 What patient group(s) is(are) included in the economic evaluation? Do they reflect the licensed indication/CE marking or the population from the trials in sections 1.3 and 6.3.3, respectively? If not, how and why are there differences? What are the implications of this for the relevance of the evidence base to the specification of the decision problem? For example, the population in the economic model is more restrictive than that described in the (draft) SPC/IFU and included in the trials.

A Marketing Authorisation Application was submitted to the EMA in October 2013 for the approval of nintedanib in combination with docetaxel for the treatment of patients with locally advanced, metastatic, or recurrent NSCLC of adenocarcinoma tumour histology after first-line chemotherapy.

The application was based on the findings of the LUME-Lung 1 trial. Thus, the model population was based on this trial and included patients with the adenocarcinoma type of locally advanced and/or metastatic, stage IIIB–IV or recurrent NSCLC who failed after first-line chemotherapy.

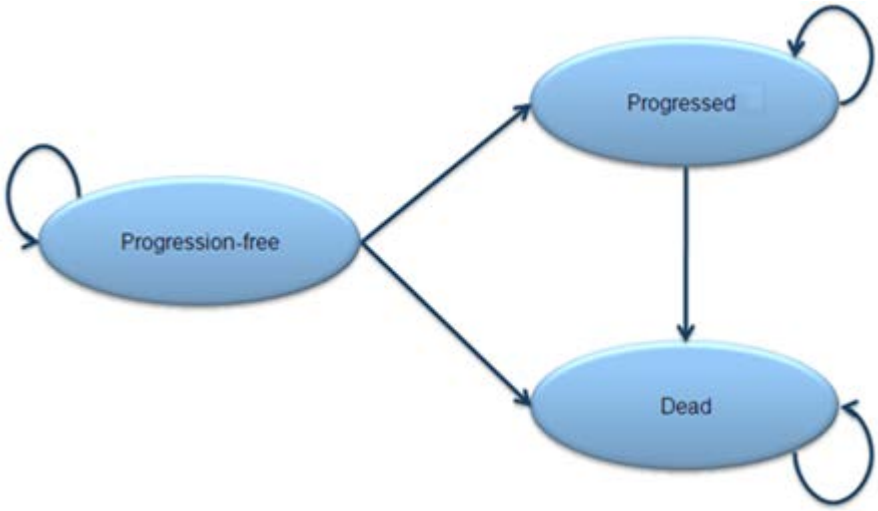
Model structure

7.2.2 Please provide a diagrammatical representation of the model you have chosen.

A partitioned survival Markov model (Figure 24) was developed in Microsoft Excel® using a three-week cycle length with 3 health states including:

- Progression-free (on or off treatment) (PF)
- Progressed disease (PD)
- Death (D)

Figure 24: Model structure



7.2.3 Please justify the chosen structure in line with the clinical pathway of care identified in section 2.5.

The model structure is in line with the clinical pathway of care in NSCLC treatment. Both the model structure and health states are characteristic of modelling in metastatic oncology and have been used in previous NICE STAs and MTAs (6-8).

The model uses the partitioned survival (also known as area under the curve or AUC) method to determine the proportion of patients in each of the three health states during each model cycle. The proportion of patients in the progressive disease state is estimated as the difference between OS and PFS. Estimates of OS and PFS in the model are based on the progression-free and OS data from LUME-Lung 1 and the corresponding parametric survival models.

Each health state (PF and PD) is associated with a cost and a health-related utility to estimate QALY over the time horizon of the analysis. The cycle length in the model is three weeks, which allows adequate granularity when assessing progression and survival. QALYs in the treatment arms are estimated as the sum of AUCs for the PF state and PD states, weighted by the respective health related utilities. Costs relating to health state management (excluding treatment costs or costs relating to AEs) are also introduced into the model by weighting the respective areas under the curve by the health state management costs for the PF and PD health states.

Cost and utility reduction due to AEs are applied in the model based on the estimated proportions of patients suffering from AEs in each treatment arm. The impact of AEs on health outcomes (QALY) is calculated using information on the duration of AEs and their impact on health-related utility.

7.2.4 Please define what the health states in the model are meant to capture.

The PFS state represents the period patients' cancer does not worsen according to the Resist Criteria used in LUME Lung 1 whilst receiving active treatment. Patients in the PFS health state experience a relatively high QoL prior to disease progression. The PD state involves the worsening of the disease during which time patients suffer a relatively poorer QoL. These health states are characteristic of those used in the modelling of metastatic oncology.

7.2.5 How does the model structure capture the main aspects of the condition for patients and clinicians as identified in section 2 (Context)? What was the underlying disease progression implemented in the model? Or what treatment was assumed to reflect underlying disease progression? Please cross-reference to section 2.1.

NSCLC is characterised by rapid progression and high mortality rates. The three-state model used in this submission reflects this disease course and is the usual method for modelling patients with metastatic cancer (7, 116).

The model is appropriate for the course of disease outlined in Section 2, and patients experience disease progression that can be affected by therapeutic interventions. Once patients have locally advanced or metastatic second-line NSCLC their treatment options are limited to docetaxel, erlotinib or nintedanib plus docetaxel. This is demonstrated in the model whereby patients are treated in the progression-free state and progress when their treatment fails. Patients in both the PF and PD states may die, and this is a transition option included in the model.

7.2.6 Please provide a table containing the following information and any additional features of the model not previously reported. A suggested format is presented below.

Table 73: Key features of analysis

Factor	Chosen values	Justification	Reference
Time horizon	15 years	Set to cover the lifetime of the patients, in order to fully incorporate the costs and health outcomes of NSCLC.	NICE 2013(102)
Cycle length	3 weeks	Allows an adequate granularity when assessing progression and survival.	NICE 2013 (102)
Half-cycle correction	Yes	Mitigate bias due to cycle length	NICE 2013 (102)
Were health effects measured in QALYs; if not, what was used?	QALYs	NICE Reference Case	NICE 2013 (102)
Discount of 3.5% for utilities and costs	Yes	NICE Reference Case	NICE 2013 (102)
Perspective (NHS/PSS)	Yes	NICE Reference Case	NICE 2013 (102)
Number of patients per cohort	1	To estimate cost and outcomes per patient	NA
Days per monthly cycle	30.42	= 365.25 / 12	NA
Days per year	365.25	NA	NA
NHS= National Health Service; PSS = personal social services; QALYs = quality-adjusted life years			

Technology

7.2.7 Are the intervention and comparator(s) implemented in the model as per their marketing authorisations/CE marking and doses as stated in sections 1.3 and 1.5? If not, how and why are there differences? What are the implications of this for the relevance of the evidence base to the specified decision problem?

As discussed in Section 7.2.1, the indication outlined as expected in the marketing authorisation in Section 1.3 to 1.5 is modelled as to be consistent with the available data.

In clinical practice, the doses may be adjusted by the clinicians, in the same way that they were during the clinical trials. As the model is based on the clinical trial data, the outcomes from the model would be expected to mirror the clinical trial data and hence clinical practice.

Note: Although erlotinib is included as a comparator within the model, it is considered a secondary comparator in this submission. This decision was made based on feedback from an advisory board held on the 10th April 2014 with five leading lung cancer clinicians (see Section 7.3.5). All five clinicians agreed that patients likely to receive erlotinib will be a different patient group to those receiving either docetaxel monotherapy or nintedanib plus docetaxel. Recent studies such as the TAILOR study(76) have shown that erlotinib is likely to be inferior to docetaxel in patients with EGFR wild-type tumours. This has led to practise within the NSCLC community that any patient of PS 0-1 should currently receive docetaxel. The clinicians agreed that as the patients treated with erlotinib are a different patient population, erlotinib is not a relevant comparator in this economic evaluation.

7.2.8 Please note that the following question refers to clinical continuation rules and not patient access schemes. Has a treatment continuation rule been assumed? If the rule is not stated in the (draft) SPC/IFU, this should be presented as a separate scenario by considering it as an additional treatment strategy alongside the base-case interventions and comparators. Consideration should be given to the following.

- The costs and health consequences of factors as a result of implementing the continuation rule (for example, any additional monitoring required).
- The robustness and plausibility of the endpoint on which the rule is based.
- Whether the 'response' criteria defined in the rule can be reasonably achieved.
- The appropriateness and robustness of the time at which response is measured.

- Whether the rule can be incorporated into routine clinical practice.
- Whether the rule is likely to predict those patients for whom the technology is particularly cost effective.
- Issues with respect to withdrawal of treatment from non-responders and other equity considerations.

The model assumes that patients are treated continuously with nintedanib plus docetaxel, docetaxel monotherapy or erlotinib until disease progression or treatment discontinuation for any other reason.

7.3 Clinical parameters and variables

When relevant, answers to the following questions should be derived from, and be consistent with, the clinical-evidence section of the submission (section 6). Cross-references should be provided. If alternative sources of evidence have been used, the method of identification, selection and synthesis should be provided as well as a justification for the approach.

7.3.1 Please demonstrate how the clinical data were implemented into the model.

Modelling Effectiveness

Kaplan-Meier survival curves for OS and PFS for nintedanib with docetaxel and docetaxel monotherapy were available from the LUME-Lung 1 trial. These curves show the proportion of patients in the three health states (no progression, progression, dead) at each time point. These data were incorporated into the cost-effectiveness model by using full parametric approximation of the raw data in the base-case. In the sensitivity analyses, Kaplan-Meier data from the clinical trial were used to model OS (until at least 5% of trial patients are still at risk) and were extrapolated using parametric function as a tail to the KW data to provide a lifetime time horizon.

Survival data in LUME-Lung 1 were fairly 'mature', the Kaplan Meier curves reached about 2% for PFS and 5% for OS. The proportion of censored patients was similar in both treatment arms (for PFS 20.5% versus 20.8%, for OS 17.9% versus 19.6% for docetaxel monotherapy versus nintedanib with docetaxel, respectively). Nevertheless, in order to facilitate

extrapolation of trial data beyond the trial time horizon, indirect comparison, and probabilistic analyses, OS and PFS data were analysed using parametric survival models.

The parameterisations, along with calculation of confidence intervals (95%), variance-covariance matrices for the use in uncertainty analysis, and goodness of fit statistics were generated by the statistical services consultancy contracted by BI, and were based on statistical analyses conducted on the data from the LUME-Lung 1 trial (the details are presented in health economics statistical analysis plans)(117, 118).

Parametric survival curves were fitted on PFS and OS Kaplan-Meier curves using two approaches: 1) 'Joint' models – statistical models including data for both treatment groups, with a term for treatment, and 2) 'Separate' models – statistical models that were fitted to each randomised treatment arm separately. Distributions fitted included exponential, Weibull, Gompertz, log-logistic, log-normal, and generalized gamma. Generalized gamma distribution has three parameters and is therefore more flexible and often presents the best fit. However, it is not straightforward to implement in Excel. Weibull and lognormal are special cases of the gamma distribution, and gamma was considered to help the choice between a Weibull and log-normal distribution. Since simpler functions were found to be good fits the gamma was not implemented.

Note that since erlotinib is compared to via a HR generated from an indirect comparison, when generating results in the model for nintedanib plus docetaxel versus erlotinib, it is necessary to use a proportional hazard model; log-normal and log-logistic cannot be used.

Choice of statistical model

To assess "goodness of fit", Akaike information criteria (AIC) and Bayesian information criteria (BIC) were calculated for each functional form. In addition, linear diagnostic plots (119) were generated for all parametric distributions.

Parameters of the separate distributions, Kaplan-Meier curves and diagnostic plots were examined to check the proportionality of hazard and assess if joint or separate models should be used. In general, the models with the lower value of AIC or sum of AICs were preferred. However, visual inspection of diagnostic tests was always checked if they suggested other distributions to be best.

When the best fit was not a proportional hazard model (exponential, Gompertz or Weibull), the best proportional model was also implemented within the model for scenario analysis, to enable inclusion of comparators not in the trial.

Goodness-of-fit statistics for PFS and OS, along with the recommendations on the most appropriate fits for the overall adenocarcinoma patients, are displayed in Table 74. Diagnostic plots are reported in Appendix 10.16.

Table 74: Goodness-of-fit Measures for IRC PFS and OS for the Overall Adenocarcinoma Patients

Distribution	Treatment	Separate Models		Joint Models	
		AIC	BIC	AIC	BIC
PFS for the overall adenocarcinoma patients					
Best fit: Log-normal function					
Best fit (PH model): Weibull function					
Exponential	Nin+Doc	760.36	758.13	1612.99	1615.97
	Doc	857.16	854.98		
Gompertz	Nin+Doc	754.08	755.63	1614.46	1621.93
	Doc	859.16	860.79		
Log-logistic	Nin+Doc	714.48	716.03	1509.3	1516.77
	Doc	799.98	801.61		
Log-normal	Nin+Doc	***711.61***	***713.16***	***1496.52***	***1503.99***
	Doc	***790.88***	***792.51***		
Weibull	Nin+Doc	***721.69***	***723.24***	***1575.76***	***1583.23***
	Doc	***850.82***	***852.45***		
OS for the overall adenocarcinoma patients					
Best fit: Log-logistic function					
Best fit (PH model): Weibull function					
Exponential	Nin+Doc	916.24	914.01	1819.39	1822.37
	Doc	905.78	903.6		
Gompertz	Nin+Doc	916.63	918.18	1817.62	1825.09
	Doc	905.36	906.99		
Log-logistic	Nin+Doc	***905.58***	***907.13***	***1778.76***	***1786.23***
	Doc	***875.28***	***876.91***		
Log-normal	Nin+Doc	910.88	912.43	1786.21	1793.68
	Doc	876.18	877.81		
Weibull	Nin+Doc	***911.45***	***913.00***	***1800.42***	***1807.89***
	Doc	***892.41***	***894.04***		

AIC = Akaike information criteria; BIC = Bayesian information criteria; Doc = docetaxel; Nin+Doc = A combination treatment of nintedanib and docetaxel; OS = Overall survival; PFS = Progression-free survival

Cells with light grey background are the best fits

Cells with light red background are the best proportional hazard fits

Note that although "goodness of fit" based on AICs indicated that joint models were appropriate, the intercept and scale parameters of the separately fitted curves indicated that the curves should not be forced into the same model, thus separate curves were selected for OS and PFS. The log-logistic model had the lowest AIC among the separately fitted OS models, and the Weibull model had the lowest AIC among the separate proportional hazard models for OS; therefore, these were selected to model the OS data. The log-normal model had the lowest AIC among the separate PFS fits, and the Weibull had the lowest AIC among the separate proportional hazard models for PFS; therefore, these were selected to model PFS. The resulting survival curves of adenocarcinoma patients are presented in Figure 25 and Figure 26. Other models were also allowed in the model for sensitivity analysis (see Section 7.6.1).

Figure 25: PFS Curves – Adenocarcinoma Patients

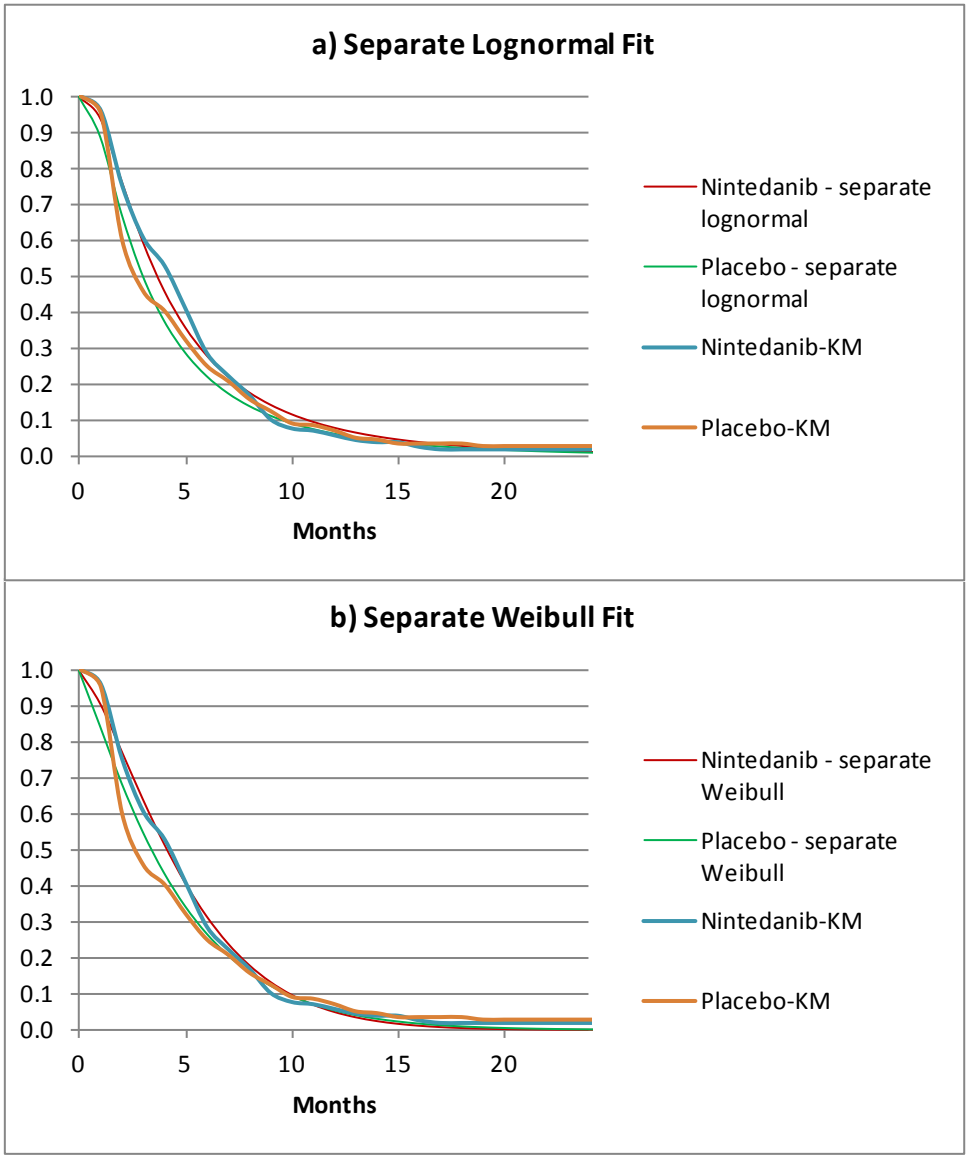
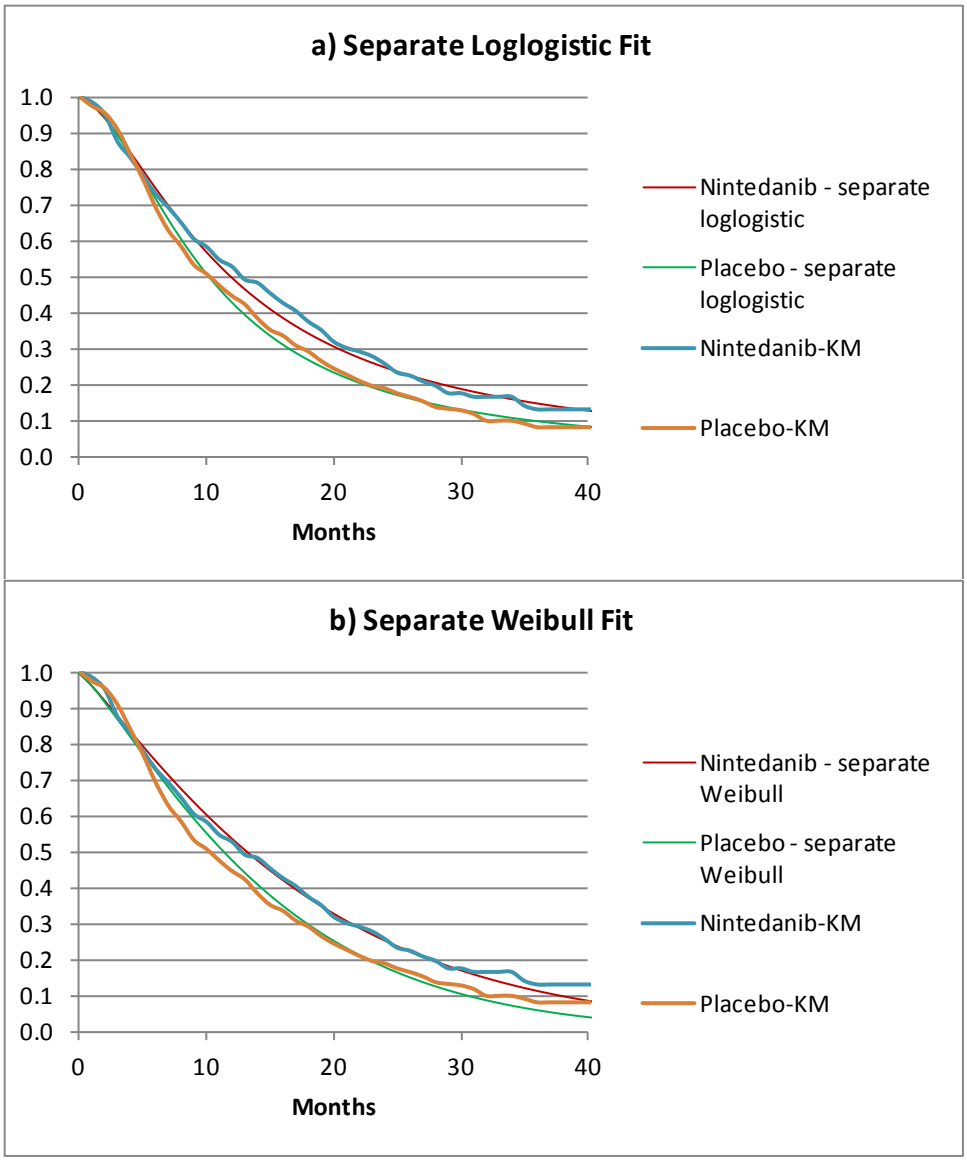


Figure 26: OS Curves – Adenocarcinoma Patients



In addition, the long-term extrapolation of trial data was validated with a group of UK clinicians and against data from the Surveillance, Epidemiology, and End Results (SEER) program using the SEER*Stat software, as well as against data from the National Lung Cancer Audit Database (LUCADA). Details of this validation can be found in Appendix 10.14 and Appendix 10.15.

Survivals implemented in the model

Survival modelling options programmed into the cost-effectiveness model are displayed in Table 75. In the base-case, the analysis used separate models for PFS and OS, with log-normal distribution for the PFS and log-logistic distribution for the OS. Impact on the cost-effectiveness results using other survival options will be explored in the sensitivity analysis.

Table 75: Survival Estimation Models Employed in the Model

PFS	OS
Separate model – Log-normal (base-case)	Separate model – Log-logistic (base-case)
Separate model – Weibull	Separate model – Weibull
Kaplan-Meier curve*	Kaplan-Meier curve*
	Kaplan-Meier curve & SEER lognormal†
	Kaplan-Meier curve & Separate Log-logistic†
	Kaplan-Meier curve & Separate Weibull†
	Kaplan-Meier curve & LUCADA lognormal†

OS = Overall survival; PFS = Progression-free survival

* With this option, the model does not extrapolate the PFS/OS with the use of parametric models but it uses the Kaplan-Meier curves for PFS/OS obtained directly from the LUME-Lung 1 trial. Note that this option only applies for nintedanib + docetaxel and docetaxel monotherapy.

† With this option, the Kaplan-Meier curves from the LUME-Lung 1 trial are used for the estimation of OS until patient number at risk drops down to 5% of original patients, afterwards parametric models are used.

Efficacy data for erlotinib

As the Kaplan Meier curves on OS and PFS for erlotinib were not available, model inputs on OS and PFS for erlotinib was derived by applying HRs (i.e., versus nintedanib plus docetaxel) obtained from the mixed treatment comparisons to the OS and PFS of nintedanib plus docetaxel. Note that HRs can only be used if survival distribution is a proportional hazard model (PHM) such as exponential, Weibull, or Gompertz. Thus, in the model, erlotinib can be evaluated only if Weibull distribution is selected for both OS and PFS.

HRs for erlotinib obtained from the network-meta analysis are shown in Table 76. The model base-case analysis utilized data from the NMA Base-case network – fixed effects model. A fixed effects model was chosen because there was one trial per comparison.

Table 76: HRs of PFS and OS for Nintedanib plus Docetaxel versus Erlotinib

Comparison	**OS (HR, 95% CrIs)**		
	Model Base-case Analysis	**Model Sensitivity Analysis**	
	NMA Base-case Analysis (fixed effects)	NMA Scenario Analysis (fixed effects)	NMA Scenario Analysis (random effects)
Nintedanib + docetaxel vs erlotinib	**0.64 [0.46, 0.90]**	0.74 [0.57, 0.96]	0.74 [0.40, 1.35]
PFS (HR 95% CrIs)			
Nintedanib + docetaxel vs erlotinib	**0.70 [0.50, 1.00]**	0.68 [0.49, 0.95]	0.68 [0.35, 1.35]

Incidence and duration of AEs

Risks of AEs for nintedanib in combination with docetaxel and docetaxel monotherapy were obtained from the LUME-Lung 1 trial CTR (5). AEs for erlotinib were obtained from the NICE TA162 (34) (for fatigue, febrile neutropenia, infection, nausea and vomiting, neutropenia, rash and grade 3 and 4 diarrhoea) and a published study (for grade 2 diarrhoea) (108). Risks of AEs for each model cycle were calculated using mean time on treatment as reported from these studies (i.e., 5.52 months for nintedanib in combination with docetaxel, 3.52 months for docetaxel monotherapy, equal to 5.1 cycles, and 4.11 months for erlotinib).

The list of AEs used by the model includes grade 3 and 4 AEs, with the exception of diarrhoea, for which grade 2 is also included. Based on the opinion of the clinical EE (see Section 7.3.5), grade 2 diarrhoea can have significant impact on resource use, and thus, was taken into consideration in the analysis. The AEs were included if they occurred in more than 5% of the cases or if they were recommended by EEs due to their importance, both in terms of costs and effects.

Overall frequencies of AEs over the duration of the respective trials are shown in Table 77. These are converted to a cycle probability, based on the duration of the trial, and are applied in the model for all patients who are still progression-free and are still on treatment (i.e. did not discontinue due to AEs). This method of calculation assumes that AEs happen any time while on treatment, with a constant hazard, i.e. some may emerge earlier but others may result as drug use is accumulated. The resulting proportion of patients with AE per cycle is 12.8% on nintedanib plus docetaxel, 12.3% on docetaxel, and 8% on erlotinib.

The proportion of patients having AEs in each model cycle accrued costs related to the management of the AEs. Briefly, for the 12.8% of patients still on treatment in any given model cycle for nintedanib plus docetaxel, a cycle cost of AEs are assigned that are calculated as the weighted average of the various AEs related to the drug.

Impact on HRQL was modelled as utility decrement associated with each type of AE and was assumed to have an impact for a period of one model cycle (i.e., three weeks). This assumption was validated during the Advisory Board (Section 7.3.5). This is likely to be a conservative assumption, because clinicians noted that patients who present with a symptomatic AE will be treated and most AEs should really be resolved within a matter of

some days, with the exception of fatigue and potentially, mild but ongoing diarrhoea (see Section 7.3.5).

Table 77: Frequencies of AEs for Adenocarcinoma Patients during the Treatment Period*

AE	Nintedanib + Docetaxel	Docetaxel	Erlotinib
ALT increase	10.3%	0.6%	0.0%
Anaemia	2.5%	3.0%	0.0%
AST increase	4.1%	0.6%	0.0%
Diarrhoea – Grade 2	28.8%	12.9%	6.3%
Diarrhoea – Grade 3 and 4	5.3%	3.0%	6.0%
Fatigue	2.2%	1.8%	19.0%
Febrile neutropenia	7.2%	4.5%	0.0%
Infection	6.6%	5.4%	2.0%
Nausea and vomiting	1.5%	0.6%	3.0%
Neutropenia	9.1%	12.0%	0.0%
Rash	0.3%	0.0%	9.0%
Thrombocytopenia	1.3%	1.2%	0.0%
WBC count decreased	15.9%	14.7%	0.0%

ALT = Alanine aminotransferase; AST = Aspartate aminotransferase; WBC = White blood cell

*Treatment durations are 5.52 months for nintedanib in combination with docetaxel, 3.52 months for docetaxel monotherapy and 4.11 months for erlotinib

Source: LUME-Lung 1(5); NICE STA for erlotinib(34); except for liver-related toxicities: FDA PI for erlotinib.

7.3.2 Demonstrate how the transition probabilities were calculated from the clinical data. If appropriate, provide the transition matrix, details of the transformation of clinical outcomes or other details here.

The model includes three health states, progression-free (PF), progressive disease (PD), and death. The proportion of patients in each of the three health states during each model cycle is determined by the AUC, or partitioned survival approach, based on parametric survival models for progression-free and OS. Therefore the model does not use transition probabilities and includes no transition matrix.

The proportion of the model cohort in each health state each model cycle is calculated by partitioning the area under the OS curve into the proportion of patients in PF and the proportion of patients in PD. For each cycle, the proportion of patients in the PD health state is defined as the difference between the OS and PFS for that cycle (OS – PD). The total time

spent in the PF and PD health states for the time horizon considered by the model is the sum of the time spent in each health state over the time horizon of the analysis.

Progressive disease is represented in the model by a single health state. However, in order to reflect the progression and treatment of patients after progression, assumptions are made with regard to the likely patient treatment after progression. In the model, the patients in a progressed health state may have received subsequent active treatments or best supportive care (BSC). However, the impact of the subsequent therapy on OS was not included in the model, and thus choice of subsequent therapy only had an implication on costs. It was assumed that patients would remain on the subsequent treatment or BSC from progression until death. Because of this, treatment switch or discontinuation in the third line was not allowed in the model (i.e., time on subsequent treatment or BSC = time in progressed state). The base-case assumed that 5% of the patients received erlotinib, about 25% received a platinum-based combination therapy ("platinum doublet") and 70% received BSC post-progression, based discussion with external experts (EEs) at an Advisory Board Meeting on the 10th April 2014 (see Section 7.3.5 for advisory board and Table 78 for subsequent therapy).

Table 78: Subsequent Therapy – Base-case

Variable	Value	Source
Treatment switch due to progression		
Docetaxel	0.0%	EE input
Erlotinib	5.0%	EE input
Pemetrexed	0.0%	EE input
Placeholder	0.0%	EE input
Platinum doublet	25.0%	EE input
BSC	70.0%	EE input

BSC = Best supportive care; EE = External expert' SE = Standard error;

Source: Discussion at Advisory Board meeting, in April 2014.

7.3.3 Is there evidence that (transition) probabilities should vary over time for the condition or disease? If so, has this been included in the evaluation? If there is evidence that this is the case, but it has not been included, provide an explanation of why it has been excluded.

As discussed in Section 7.3.2, the model does not use transition probabilities because it assumes the AUC approach. However, the time-dependent aspects of NSCLC are captured in the model through the incorporation of trial-based parametric survival models describing PFS and OS.

7.3.4 Were intermediate outcome measures linked to final outcomes (for example, was a change in a surrogate outcome linked to a final clinical outcome)? If so, how was this relationship estimated, what sources of evidence were used, and what other evidence is there to support it?

No surrogate markers were used in the model.

7.3.5 If clinical experts assessed the applicability of values available or estimated any values, please provide the following details[6]:

- the criteria for selecting the experts
- the number of experts approached
- the number of experts who participated
- declaration of potential conflict(s) of interest from each expert or medical specialist whose opinion was sought
- the background information provided and its consistency with the totality of the evidence provided in the submission
- the method used to collect the opinions
- the medium used to collect opinions (for example, was information gathered by direct interview, telephone interview or self-administered questionnaire?)
- the questions asked
- whether iteration was used in the collation of opinions and if so, how it was used (for example, the Delphi technique).

[6] Adapted from Pharmaceutical Benefits Advisory Committee (2008) Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee (Version 4.3). Canberra: Pharmaceutical Benefits Advisory Committee.

Boehringer Ingelheim selected clinical experts to review assumptions within the submission on the basis they were widely published and were involved in clinical trials and guidelines and guidance development.

Initial interview with Dr [redacted]

Evidera, a health economics consultancy involved in model development approached one clinical expert to review the assumptions. The clinician was Dr [redacted] (BSc PhD FRCP, Consultant Medical Oncologist, Royal Marsden Hospital and Honorary Clinical Senior Lecturer, Imperial College, London). Dr [redacted] is familiar with cost-effectiveness analyses as he was involved with previous NICE Technology Appraisals of other lung cancer products. He was selected on the basis of his familiarity with both the therapy area and NICE HTA submissions. Dr [redacted] has received honoraria from Evidera for his participation and received payment from BI earlier for participating in Advisory Board meetings.

Two one-hour telephone interviews were conducted with Dr [redacted] after model conceptualisation and trial data review, but prior to full model implementation. No background information was provided prior to this, however, Dr [redacted] participated in advisory boards organised by Boehringer Ingelheim and he was clearly familiar with the design and the analysis plan of the LUME Lung 1 trial. During this interview the clinical assumptions of the model were checked and resource use for regular monitoring was asked. Further email clarifications were sought on three occasions. Questions and answers from this interview are provided in Appendix 10.18.1. The full questionnaire results can be seen in Appendix 10.18.2.

All of the recommendations from Dr [redacted] were addressed in the analysis.

In addition, once the model was developed, an advisory board with five UK clinicians was organised to check clinical face validity of final inputs and the survival extrapolation of the clinical trial data beyond the time horizon of the LUME Lung 1 trial. Dr [redacted] was not part of the advisory board in order to minimise bias. This advisory board is described below and the outputs are described in Appendix 10.19.

Due to inconsistency of the total costs estimates for treating AEs with all previous cost estimates in NICE submissions, further input was sought from the participating experts. The questionnaire developed was sent to participants via email. Only one clinical expert filled in the questionnaire for the advisory board and had follow up questions. After clarification of

these questions by Evidera, Dr [redacted] revised some of his estimates and all AEs were discussed in person and the resource use data for AE management were finalised. Total costs resulting from this exercise were applied in the cost-effectiveness model. The full questionnaire results can be seen in Appendix 10.18.

Advisory board with 5 clinicians

Additionally, Boehringer Ingelheim approached five clinical experts to review the assumptions as part of an advisory board held on the 10th April 2014, and all five attending and gave their opinions. The notes from this advisory board were written up and agreed upon by all of the clinicians.

The clinicians were:

- James Spicer - Consultant in Medical Oncology at Guy's and St Thomas' Hospitals, London
- Marianne Nicolson - Consultant Medical Oncologist, Aberdeen Royal Infirmary
- Yvonne Summers - Consultant Medical Oncologist, The Christie Hospital NHS Trust & University Hospital South Manchester
- [redacted] - Consultant Clinical Oncologist, Velindre Cancer Centre, Cardiff
- Tim Benepal - Consultant Medical Oncologist, St George's Hospital, London

Clinicians were aware that the advisory board was to discuss aspects of the nintedanib for NSCLC HTA submission, and they were aware of the LUME Lung 1 trial. During the advisory board, the clinical assumptions of the model were checked and discussed amongst all clinicians. The details of the discussion held at this meeting are presented in Appendix 10.19.

Summary of selected values

7.3.6 Please provide a list of all variables included in the cost-effectiveness analysis, detailing the values used, range (distribution) and source. Provide cross-references to other parts of the submission. Please present in a table, as suggested below.

Table 79: Summary of variables applied in the economic model

Category	Variable	Value	Variance	Reference	Source
General settings	Patient population	All adenocarcinoma patients	NA	Section 7.2.6	NA
	Time horizon	15 years (lifetime)	NA	Section 7.2.6	NICE 2013(102)
	Model cycle length (weeks)	3	None	Section 7.2.6	-

Category	Variable		Value	Variance	Reference	Source
	Discount rate for costs		3.5%	0%-6%	Section 7.2.6	NICE 2013(102)
	Discount rate for health		3.5%	0%-6%	Section 7.2.6	NICE 2013(102)
	Average body surface area (BSA) of patients at baseline (m^2)		1.855	None	Section 7.2.6	Sacco et al. 2010(120)
	Number of patients per cohort		1	None	Section 7.2.6	NA
	Days per monthly cycle		30.44	None	Section 7.2.6	NA
	Days per year		365.25	NA	Section 7.2.6	NA
Efficacy	Survival estimation models employed	PFS	Separate – LogNormal	NA	Section 7.3.1	Trial data (fit on LUME-Lung 1 data + Best fit based on AIC and BIC)
		OS	Separate – LogLogistic	NA	Section 7.3.1	Trial data (fit on LUME-Lung 1 data + Best fit based on AIC and BIC)
	HR – PFS	nintedanib + docetaxel vs docetaxel	0.77	0.62-0.96 (CI)	Section 6.7.4	MTC (Base-case, Fixed effect)
		ninte + doce vs erlotinib	0.70	0.50-1.00 (CI)	Section 6.7.4	MTC (Base-case, Fixed effect)
	HR – OS	ninte + doce vs docetaxel	0.83	0.70-0.99 (CI)	Section 6.7.4	MTC (Base-case, Fixed effect)
		ninte + doce vs erlotinib	0.64	0.46-0.90 (CI)	Section 6.7.4	MTC (Base-case, Fixed effect)
Treatment dis-continuation	Cycle probability of discontinuing nintedanib		12.5%	SE = 1.2%	None	Calculated from LUME-Lung 1(5)
	Cycle probability of discontinuing docetaxel (while taken in combination with nintedanib)		17.5%	SE = 1.8%	None	Calculated from LUME-Lung 1(5)
	Cycle probability of discontinuing docetaxel (as monotherapy)		19.6%	SE = 2.0%	None	Calculated from LUME-Lung 1(5)
	Cycle probability of discontinuing erlotinib		16.8%	SE = 1.7%	None	Calculated from NICE TA162(34)
	Mean time on treatment (months)	Nintedanib	5.53	SE = 0.29	None	LUME-Lung 1 CSR Table 15.3.2.4.1:5
		Docetaxel in combination with nintedanib	3.93	SE=0.18	None	Calculated from LUME-Lung 1 CSR Table 12.1.2.3:1(5)
		Docetaxel as monotherapy	3.52	SE = 0.17	None	Calculated from LUME-Lung 1 CSR Table 12.1.2.3:1(5)
		Erlotinib	4.10	NA	None	Calculated from NICE TA162(34)
Treatment switch due to progression	Proportion of patients switching to docetaxel		0%	SE = 0%	Section 7.3.5	EE opinion
	Proportion of patients switching to erlotinib		5%	SE = 0.26% (= (10%/1.96) of mean)	Section 7.3.5	EE opinion
	Proportion of patients switching to pemetrexed		0%	SE = 0%	Section 7.3.5	EE opinion
	Proportion of patients switching to platinum doublet therapy		25%	SE = 1.28% (= (10%/1.96) of mean)	Section 7.3.5	EE opinion
	Proportion of patients switching to BSC		70%	SE = 3.57% (= (10%/1.96) of mean)	Section 7.3.5	EE opinion
	Average duration of third-line active treatment (months)		3.30	-	None	BI, data on file

Category	Variable		Value	Variance	Reference	Source
Drug costs of active therapies	Wastage included for IV treatments		Yes	NA	None	NA
	Wastage included for drugs administered orally		No	NA	None	NA
	Monthly cost of nintedanib		£2151.10	None	Section 1.10	BI
	Docetaxel pack prices	20 mg (vial size = 0.5 ml)	£6.57	None	Section 7.5.5	http://cmu.dh.gov.uk/electronic-market-information-tool-emit/ Accessed 24 Apr 2014
		20 mg (vial size = 1 ml)	£6.42	None	Section 7.5.5	http://cmu.dh.gov.uk/electronic-market-information-tool-emit/ Accessed 24 Apr 2014
		80 mg (vial size = 2 ml)	£44.45	None	Section 7.5.5	http://cmu.dh.gov.uk/electronic-market-information-tool-emit/ Accessed 24 Apr 2014
		80 mg (vial size = 4 ml)	£21.23	None	Section 7.5.5	http://cmu.dh.gov.uk/electronic-market-information-tool-emit/ Accessed 24 Apr 2014
		140 mg	£34.29	None	Section 7.5.5	http://cmu.dh.gov.uk/electronic-market-information-tool-emit/ Accessed 24 Apr 2014
		160 mg	£47.30	None	Section 7.5.5	http://cmu.dh.gov.uk/electronic-market-information-tool-emit/ Accessed 24 Apr 2014
	nintedanib	Units per administration	400 mg	None	Section 7.5.5	LUME-Lung 1(5)
		Price per mg	£0.18	None	Section 7.5.5	Calculated
		Dose intensity	91.2%	SE = 0.84%	Section 7.5.5	LUME-Lung 1 CSR Table 12.1.1.3:1(5)
		Administrations per model cycle	21	None	Section 7.5.5	-
		Cycle drug and administration costs	£1353.52	None	Section 7.5.5	Calculated
	docetaxel in combination with nintedanib	Units per administration	75 mg/m^2	None	Section 7.5.5	LUME-Lung 1(5)
		Dose intensity	98.1%	SE = 0.25%	Section 7.5.5	LUME-Lung 1 CSR Table 12.1.2.3:1(5)
		Administrations per model cycle	1	None	Section 7.5.5	LUME-Lung 1(5)
		Cost of administration	£155	None	Section 7.5.5	National Schedule of Reference Costs Year: 2012–13 – NHS trusts and NHS foundation trusts – Chemotherapy – Outpatient. HRG Code: SB12Z. Outpatient – Deliver simple Parenteral Chemotherapy at first attendance(121)
		Cycle drug and administration costs	£195.08	None	Section 7.5.5	Calculated
	docetaxel as monotherapy	Units per administration	75 mg/m^2	None	Section 7.5.5	LUME-Lung 1(5)
		Dose intensity	98.7%	SE = 0.2%	Section 7.5.5	LUME-Lung 1 CSR Table 12.1.2.3:1(5)
		Administrations per model cycle	1	None	Section 7.5.5	LUME-Lung 1(5)
		Cost of administration	£155	None	Section 7.5.5	National Schedule of Reference Costs Year: 2012–13 – NHS trusts and NHS foundation trusts –

Category	Variable		Value	Variance	Reference	Source
						Chemotherapy(121)
		Cycle drug and administration costs	£195.33	None	Section 7.5.5	Calculated
	erlotinib	Units per administration	150 mg	None	Section 7.5.5	NICE TA162(34)
		Price per mg	£0.36	None	Section 7.5.5	Calculated from British National Formulary October 2013.(122) Accessed 17 Oct 2013
		Dose intensity	92.0%	None	Section 7.5.5	NICE TA162(34)
		Administrations per model cycle	21	None	Section 7.5.5	NICE TA162(34)
		Cycle drug and administration costs	£1050.71	None	Section 7.5.5	Calculated
	Cycle drug and administration costs of platinum doublet		£701.19	-	Section 7.5.5	Section 7.5.5
Other drug costs	Loperamide	Units in a pill/vial	2 mg	NA	Section 7.5.5	British National Formulary January 2014.(123) Accessed January 2014
		Pills/vials in a pack	30	NA	Section 7.5.5	British National Formulary January 2014.(123) Accessed January 2014
		Price per pack	£1.03	None	Section 7.5.5	British National Formulary January 2014.(123) Accessed January 2014
		Dose intensity	100%	None	Section 7.5.5	Assumption
	Codeine phosphate	Units in a pill/vial	30 mg	NA	Section 7.5.5	British National Formulary January 2014.(123) Accessed January 2014
		Pills/vials in a pack	28	NA	Section 7.5.5	British National Formulary January 2014.(123) Accessed January 2014
		Price per pack	£1.40	None	Section 7.5.5	British National Formulary January 2014.(123) Accessed January 2014
		Dose intensity	100%	None	Section 7.5.5	Assumption
	Octreotide	Units in a pill/vial	500 mcg/mL	NA	Section 7.5.5	British National Formulary January 2014.(123) Accessed January 2014
		Pills/vials in a pack	1	NA	Section 7.5.5	British National Formulary January 2014.(123) Accessed January 2014
		Price per pack	£27.09	None	Section 7.5.5	British National Formulary January 2014.(123) Accessed January 2014
	Budesonide	Units in a pill/vial	10 mcg/metered spray	NA	Section 7.5.5	British National Formulary January 2014.(123) Accessed January 2014
		Pills/vials in a pack	100	NA	Section 7.5.5	British National Formulary January 2014.(123) Accessed January 2014
		Price per pack	£5.90	None	Section 7.5.5	British National Formulary January 2014.(123) Accessed January 2014
		Dose intensity	100%	None	Section 7.5.5	Assumption January 2014
	Co-amoxiclav 500/125mg	Units in a pill/vial	1	NA	Section 7.5.5	British National Formulary January 2014.(123) Accessed January 2014

Category	Variable		Value	Variance	Reference	Source
		Pills/vials in a pack	21	NA	Section 7.5.5	British National Formulary January 2014.(123) Accessed January 2014
		Price per pack	£2.48	None	Section 7.5.5	British National Formulary January 2014.(123) Accessed January 2014
		Dose intensity	100%	None	Section 7.5.5	Assumption
	Carboplatin	Units in a pill/vial	150 mg/ml	NA	Section 7.5.5	British National Formulary.(123) Accessed 14 May 2014
		Pills/vials in a pack	1	NA	Section 7.5.5	British National Formulary.(123) Accessed 14 May 2014
		Price per pack	£50.00	None	Section 7.5.5	British National Formulary.(123) Accessed 14 May 2014
		Dose intensity	100%	None	Section 7.5.5	Assumption
		Administrations per model cycle	1	None	Section 7.5.5	-
		Cost of administration	£155	None	Section 7.5.5	NHS trusts and NHS foundation trusts – Chemotherapy – Outpatient. HRG Code: SB12Z. Outpatient – Deliver simple Parenteral Chemotherapy at first attendance(121)
	Vinorelbine	Units in a pill/vial	50 mg/m^2	NA	Section 7.5.5	British National Formulary.(123) Accessed 14 May 2014
		Pills/vials in a pack	1	NA	Section 7.5.5	British National Formulary.(123) Accessed 14 May 2014
		Price per pack	£139.00	None	Section 7.5.5	British National Formulary.(123) Accessed 14 May 2014
		Dose intensity	100%	None	Section 7.5.5	Assumption
		Administrations per model cycle	3	None	Section 7.5.5	-
		Cost of administration	£155	None	Section 7.5.5	NHS trusts and NHS foundation trusts – Chemotherapy – Outpatient. HRG Code: SB12Z. Outpatient – Deliver simple Parenteral Chemotherapy at first attendance(121)
End of life costs	End of life costs		£0	SE = 0	Section 7.3.5	EE opinion
Unit costs	Healthcare professional visit	Routine physician consultation/GP (monitoring)	£63.0	None	Section 7.5.6	PSSRU. Unit Costs of Health & Social Care 2012. Compiled by L. Curtis(124)
		Oncologist specialist visit (specialised monitoring)	£139.0	None	Section 7.5.6	National Schedule of Reference Costs Year: 2012–13 – NHS trusts and NHS foundation trusts – Consultant Led Outpatient Attendances. Currency Code: WF01A - 370. Non-Admitted Face to Face Attendance, Follow-up – Medical Oncology(121)
		Hepatologist specialist visit	£200.0	None	Section 7.5.6	National Schedule of Reference Costs Year: 2012–13 – NHS trusts and NHS foundation trusts – Consultant Led Outpatient

Category	Variable		Value	Variance	Reference	Source
						Attendances. Currency Code: WF01A - 306. Non-Admitted Face to Face Attendance, Follow-up – Hepatology(121)
		Gastroenterologist specialist visit	£123.0	None	Section 7.5.6	National Schedule of Reference Costs Year: 2012–13 – NHS trusts and NHS foundation trusts – Consultant Led Outpatient Attendances. Currency Code: WF01A - 301. Non-Admitted Face to Face Attendance, Follow-up – Gastroenterology(121)
		Palliative care nurse	£70.0	None	Section 7.5.6	National Schedule of Reference Costs Year: 2012–13 – NHS trusts and NHS foundation trusts – Community Health Services – Nursing. Currency Code: N10AF. Specialist Nursing – Cancer Related, Adult, Face to face(121)
		Radiation oncologist	£121.0	None	Section 7.5.6	National Schedule of Reference Costs Year: 2012–13 – NHS trusts and NHS foundation trusts – Consultant Led Outpatient Attendances. HRG Code: WF01A - 800. Non-Admitted Face to Face Attendance, Follow-up – Clinical Oncology (Previously Radiotherapy)(121)
		Surgeon visit	£119.0	None	Section 7.5.6	National Schedule of Reference Costs Year: 2012–13 – NHS trusts and NHS foundation trusts – Consultant Led Outpatient Attendances. Currency Code: WF01A - 100. Non-Admitted Face to Face Attendance, Follow-up – General Surgery(121)
		Nurse visit	£38.0	None	Section 7.5.6	NHS trusts and NHS foundation trusts – Community Health Services – Nursing. Currency Code: N02AF. District Nurse, Adult, Face to face(121)
		Nurse home visit	£70.0	None	Section 7.5.6	PSSRU. Unit Costs of Health & Social Care 2013. Compiled by L. Curtis(124)
		Physician home visit	£292.0	None	Section 7.5.6	PSSRU. Unit Costs of Health & Social Care 2013. Compiled by L. Curtis(124)
		A&E visit	£115	None	Section 7.5.6	National Schedule of Reference Costs - Year 2012-13 - NHS trusts and NHS foundation trusts - AE Weighted Average National Cost(121)
		Other visits	£110.5	None	Section 7.5.6	Average of GP, oncologist, radiologist, and surgeon visits
	Procedures	Radiotherapy – Inpatient	£195.0	None	Section 7.5.6	National Schedule of Reference Costs Year: 2012–13 – NHS trusts and NHS

Category	Variable		Value	Variance	Reference	Source
						foundation trusts – Radiotherapy. HRG Code: SC23Z. Inpatient – Deliver a fraction of complex treatment on a megavoltage machine(121)
		Radiotherapy – Outpatient	£121.0	None	Section 7.5.6	National Schedule of Reference Costs Year: 2012–13 – NHS trusts and NHS foundation trusts – Radiotherapy. HRG Code: SC23Z. Outpatient – Deliver a fraction of complex treatment on a megavoltage machine(121)
		Blood transfusion – Inpatient	£1,121.0	None	Section 7.5.6	National Schedule of Reference Costs Year: 2012–13 – NHS trusts and NHS foundation trusts – Elective Inpatients. HRG Code: SA13A. Single Plasma Exchange, Leucophoresis or Red Cell Exchange, 19 years and over(121)
		Blood transfusion – Outpatient	£167.0	None	Section 7.5.6	National Schedule of Reference Costs Year: 2012–13 – NHS trusts and NHS foundation trusts – Procedures in Outpatients. HRG Code: SA13A. Single Plasma Exchange, Leucophoresis or Red Cell Exchange, 19 years and over(121)
		Oxygen	£13.4	None	Section 7.5.6	http://www.ppa.org.uk/edt/December_2012/mindex.htm
		Oxygen assessment	£171.0	None	Section 7.5.6	National Schedule of Reference Costs Year: 2012–13 – NHS trusts and NHS foundation trusts – Procedures in Outpatients. HRG Code: DZ38Z. Oxygen Assessment and Monitoring(121)
		CT scan	£90.0	None	Section 7.5.6	National Schedule of Reference Costs Year: 2012–13 – NHS trusts and NHS foundation trusts – Diagnostic Imaging. Weighted Average National Cost. HRG Code: RA08A. CT Scan, one area, no contrast, 19 yrs and over(121)
		Chest X-ray	£28.0	None	Section 7.5.6	National Schedule of Reference Costs Year: 2012–13 – NHS trusts and NHS foundation trusts – Directly Accessed Diagnostic Services. HRG Code: DAPF. Direct Access Plain Film(121)
		MRI	£204.0	None	Section 7.5.6	National Schedule of Reference Costs Year: 2012–13 – NHS trusts and NHS foundation trusts – Diagnostic Imaging. Weighted Average National Cost. HRG Code: RA03Z. Magnetic Resonance

Category	Variable		Value	Variance	Reference	Source
						Imaging Scan, one area, pre and post contrast(121)
		PET	£282.0	None	Section 7.5.6	National Schedule of Reference Costs Year: 2012–13 – NHS trusts and NHS foundation trusts – Diagnostic Imaging. Weighted Average National Cost. HRG Code: RA40Z. Nuclear Medicine, Category 6(121)
		Other imaging	£159.0	None	Section 7.5.6	Assumption: average of CT, X-ray, MRI, PET, and bone scan
		FBC	£3.0	None	Section 7.5.6	National Schedule of Reference Costs Year: 2012–13 – NHS trusts and NHS foundation trusts – Directly Accessed Pathology Services. HRG Code: DAPS05. Haematology(121)
		Electrolytes	£4.0	None	Section 7.5.6	National Schedule of Reference Costs Year: 2012–13 – NHS trusts and NHS foundation trusts – Directly Accessed Pathology Services. HRG Code: DAPS04. Clinical Biochemistry, x4 to include 4 tests(121)
		Liver function / LFT	£7.0	None	Section 7.5.6	National Schedule of Reference Costs Year: 2012–13 – NHS trusts and NHS foundation trusts – Directly Accessed Pathology Services. HRG Code: DAPS04. Clinical Biochemistry, x7 to include 7 tests(121)
		Renal function	£10.0	None	Section 7.5.6	National Schedule of Reference Costs Year: 2012–13 – NHS trusts and NHS foundation trusts – Directly Accessed Pathology Services. HRG Code: DAPS04. Clinical Biochemistry, x10 to include 10 tests(121)
		Calcium	£1.0	None	Section 7.5.6	National Schedule of Reference Costs Year: 2012–13 – NHS trusts and NHS foundation trusts – Directly Accessed Pathology Services. HRG Code: DAPS04. Clinical Biochemistry, x1(121)
		Colonoscopy	£309.5	None	Section 7.5.6	National Schedule of Reference Costs Year: 2012–13 – NHS trusts and NHS foundation trusts – Procedures in Outpatients. HRG Code: FZ51Z, FZ52Z. Diagnostic Colonoscopy, 19 years and over; and Diagnostic Colonoscopy with Biopsy, 19 years and over. Average(121)
		Stool cultures	£7.0	None	Section 7.5.6	National Schedule of Reference Costs Year: 2012–

Category	Variable		Value	Variance	Reference	Source
						13 – NHS trusts and NHS foundation trusts – Directly Accessed Pathology Services. HRG Code: DAPS07. Microbiology(121)
		Ultrasound	£57.0	None	Section 7.5.6	National Schedule of Reference Costs Year: 2012–13 – NHS trusts and NHS foundation trusts – Diagnostic Imaging. HRG Code: RA23Z, RA24Z. Ultrasound Scan, less than 20 minutes; and 20 minutes and over. Average(121)
		99Tc bone scintigraphy scan	£191.0	None	Section 7.5.6	National Schedule of Reference Costs Year: 2012–13 – NHS trusts and NHS foundation trusts – Diagnostic Imaging. Weighted Average National Cost. HRG Code: RA36Z. Nuclear Medicine, Category 2(121)
		Chemistry panel	£8.0	None	Section 7.5.6	National Schedule of Reference Costs Year: 2012–13 – NHS trusts and NHS foundation trusts – Directly Accessed Pathology Services. HRG Code: DAPS04. Clinical Biochemistry, x8 to include 8 tests(121)
		Coagulation test	£3.0	None	Section 7.5.6	National Schedule of Reference Costs Year: 2012–13 – NHS trusts and NHS foundation trusts – Directly Accessed Pathology Services. HRG Code: DAPS05. Haematology(121)
		U and E	£5.0	None	Section 7.5.6	National Schedule of Reference Costs Year: 2012–13 – NHS trusts and NHS foundation trusts – Directly Accessed Pathology Services. HRG Code: DAPS04. Clinical Biochemistry, x5 to include 5 tests(121)
		Other lab tests	£5.2	None	Section 7.5.6	Assumption: average of FBC/CBC, LFT, chemistry panel, coagulation test, U & E tests
	Hospitalisation costs per stay		£2,001	None	Section 7.5.6	
	Hospitalisation costs of AEs	ALT increase	£2,128	None	Section 7.5.7	National Schedule of Reference Costs Year: 2012–13 – NHS trusts and NHS foundation trusts – Non-Elective Inpatients (Long Stay). Weighted Average National Cost. HRG Code: GC17A–H, GC17J, GC17K. Non-Malignant Hepatobiliary or Pancreatic Disorders(121)
		Anaemia	£2,559	None	Section 7.5.7	National Schedule of Reference Costs Year: 2012–13 – NHS trusts and NHS foundation trusts – Non-

Category	Variable		Value	Variance	Reference	Source
						Elective Inpatients (Long Stay). Weighted Average National Cost. HRG Code: SA01G–H, SA01J, SA01K. Acquired Pure Red Cell Aplasia or Other Aplastic Anaemia(121)
		Diarrhoea grade 1 and 2	£434	None	Section 7.5.7	National Schedule of Reference Costs Year: 2012–13 – NHS trusts and NHS foundation trusts – Day Cases. Weighted Average National Cost. HRG Code: FZ36M–N, FZ36P–Q. Gastrointestinal Infections, without Interventions(121)
		Diarrhoea grade 3 and 4	£2,067	None	Section 7.5.7	National Schedule of Reference Costs Year: 2012–13 – NHS trusts and NHS foundation trusts – Non-Elective Inpatients (Long Stay). Weighted Average National Cost. HRG Code: FZ36G–H, FZ36J–M, FZ36P–Q. Gastrointestinal Infections(121)
		Fatigue	£2,559	None	Section 7.5.7	Assumption: same as anemia
		Febrile neutropenia	£2,339	None	Section 7.5.7	National Schedule of Reference Costs Year: 2012–13 – NHS trusts and NHS foundation trusts – Non-Elective Inpatients (Long Stay). Weighted Average National Cost. HRG Code: SA35A–E. Agranulocytosis(121)
		Infection	£2,574	None	Section 7.5.7	National Schedule of Reference Costs Year: 2012–13 – NHS trusts and NHS foundation trusts – Non-Elective Inpatients (Long Stay). Weighted Average National Cost. HRG Code: WA03A–C. Septicaemia(121)
		Nausea and vomiting	£1,998	None	Section 7.5.7	National Schedule of Reference Costs Year: 2012–13 – NHS trusts and NHS foundation trusts – Non-Elective Inpatients (Long Stay). Weighted Average National Cost. HRG Code: FZ91A–H, FZ91J–M. Non-Malignant Gastrointestinal Tract Disorders(121)
		Neutropenia, Thrombocytopenia, WBC count decreased	£549	None	Section 7.5.7	National Schedule of Reference Costs Year: 2012–13 – NHS trusts and NHS foundation trusts – Non-Elective Inpatients (Short Stay). Weighted Average National Cost. HRG Code: SA35A–E. Agranulocytosis(121)
		Rash	£2,385	None	Section 7.5.7	National Schedule of Reference Costs Year: 2012–13 – NHS trusts and NHS foundation trusts – Non-Elective Inpatients (Long

Category	Variable		Value	Variance	Reference	Source
						Stay). Weighted Average National Cost. HRG Code: JD07A–H, JD07J-K. Skin Disorders(121)
Monitoring costs	Source of monitoring costs		Monitoring costs/resource use data based on EE interview	NA	Section 7.5.6	NA
	Monitoring costs for nintedanib + docetaxel patients	In PF on active treatment	£187.84	SE = 9.58(= (10%/1.96) of mean)	Section 7.5.6	Calculated from British National Formulary and EE opinion
		In PF on BSC	£460.75	SE = 23.51 (= (10%/1.96) of mean)	Section 7.5.6	Calculated from British National Formulary and EE opinion
		In PD on active treatment	£98.46	SE = 5.02(= (10%/1.96) of mean)	Section 7.5.6	Calculated from British National Formulary and EE opinion
		In PD on BSC	£406.63	SE = 20.75 (= (10%/1.96) of mean)	Section 7.5.6	Calculated from British National Formulary and EE opinion
		At time of progression	£126.00	SE = 6.43 (= (10%/1.96) of mean)	Section 7.5.6	Calculated from British National Formulary and EE opinion
	Monitoring costs for docetaxel patients	In PF on active treatment	£205.22	SE = 10.47 (= (10%/1.96) of mean)	Section 7.5.6	Calculated from British National Formulary and EE opinion
		In PF on BSC	£460.75	SE = 23.51 (= (10%/1.96) of mean)	Section 7.5.6	Calculated from British National Formulary and EE opinion
		In PD on active treatment	£98.46	SE = 5.02 (= (10%/1.96) of mean)	Section 7.5.6	Calculated from British National Formulary and EE opinion
		In PD on BSC	£406.63	SE =20.75(= (10%/1.96) of mean)	Section 7.5.6	Calculated from British National Formulary and EE opinion
		At time of progression	£126.00	SE = 6.43 (= (10%/1.96) of mean)	Section 7.5.6	Calculated from British National Formulary and EE opinion
	Monitoring costs for erlotinib patients	In PF on active treatment	£101.43	SE = 5.18 (= (10%/1.96) of mean)	Section 7.5.6	Calculated from British National Formulary and EE opinion
		In PF on BSC	£460.75	SE = 23.51 (= (10%/1.96) of mean)	Section 7.5.6	Calculated from British National Formulary and EE opinion

Category	Variable		Value	Variance	Reference	Source
		In PD on active treatment	£98.46	SE =5.02 (= (10%/1.96) of mean)	Section 7.5.6	Calculated from British National Formulary and EE opinion
		In PD on BSC	£406.63	SE = 20.75 (= (10%/1.96) of mean)	Section 7.5.6	Calculated from British National Formulary and EE opinion
		At time of progression	£126.00	SE = 6.43(= (10%/1.96) of mean)	Section 7.5.6	Calculated from British National Formulary and EE opinion
Frequency of AEs	nintedanib + docetaxel	ALT increase	10.3%	None	Section 10.17	LUME- Lung 1 CSR table 12.2.2.2.3:1 (Drug related grade 3/4/5)(5)
		Anaemia	2.5%	None	Section 10.17	LUME- Lung 1 CSR table 12.2.2.4:1 (AEs of special interest grade 3/4/5)(5)
		AST increase	4.1%	None	Section 10.17	LUME- Lung 1 CSR table 12.2.2.2.3:1 (Drug related grade 3/4/5)(5)
		Diarrhoea – Grade 2	28.8%	None	Section 10.17	LUME- Lung 1 CSR table 12.2.2.2.3:1 (Drug related grade 3/4/5)(5)
		Diarrhoea – Grade 3 and 4	5.3%	None	Section 10.17	LUME- Lung 1 CSR table 12.2.2.2.3:1 (Drug related grade 3/4/5)(5)
		Fatigue	2.2%	None	Section 10.17	LUME- Lung 1 CSR table 12.2.2.2.3:1 (Drug related grade 3/4/5)(5)
		Febrile neutropenia	7.2%	None	Section 10.17	LUME- Lung 1 CSR table 12.2.2.4:1 (AEs of special interest grade 3/4/5)(5)
		Infection	6.6%	None	Section 10.17	LUME- Lung 1 CSR table 12.2.2.4:1 (AEs of special interest grade 3/4/5)(5)
		Nausea and vomiting	1.5%	None	Section 10.17	LUME- Lung 1 CSR table 12.2.2.2.3:1 (Drug related grade 3/4/5(5)
		Neutropenia	9.1%	None	Section 10.17	LUME- Lung 1 CSR table 12.2.2.2.3:1 (Drug related grade 3/4/5)(5)
		Rash	0.3%	None	Section 10.17	LUME- Lung 1 CSR table 12.2.2.4:1 (AEs of special interest grade 3/4/5)(5)
		Thrombocytopenia	1.3%	None	Section 10.17	LUME- Lung 1 CSR table 12.2.2.4:1 (AEs of special interest grade 3/4/5)(5)
		WBC count decreased	15.9%	None	Section 10.17	LUME- Lung 1 CSR table 12.2.2.2.3:1 (Drug related grade 3/4/5)(5)
	docetaxel	ALT increase	0.6%	None	Section 10.17	LUME- Lung 1 CSR table 12.2.2.2.3:1 (Drug related grade 3/4/5)(5) }
		Anaemia	3.0%	None	Section 10.17	LUME- Lung 1 CSR table 12.2.2.4:1 (AEs of special interest grade 3/4/5)(5)
		AST increase	0.6%	None	Section 10.17	LUME- Lung 1 CSR table 12.2.2.2.3:1 (Drug related grade 3/4/5)(5)
		Diarrhoea – Grade 2	12.9%	None	Section 10.17	LUME- Lung 1 CSR table 12.2.2.2.3:1 (Drug related grade 3/4/5)(5)
		Diarrhoea – Grade 3 and 4	3.0%	None	Section 10.17	LUME- Lung 1 CSR table 12.2.2.2.3:1 (Drug related grade 3/4/5)(5)

Category	Variable		Value	Variance	Reference	Source
		Fatigue	1.8%	None	Section 10.17	LUME- Lung 1 CSR table 12.2.2.2.3:1 (Drug related grade 3/4/5)(5)
		Febrile neutropenia	4.5%	None	Section 10.17	LUME- Lung 1 CSR table 12.2.2.4:1 (AEs of special interest grade 3/4/5)(5)
		Infection	5.4%	None	Section 10.17	LUME- Lung 1 CSR table 12.2.2.4:1 (AEs of special interest grade 3/4/5)(5)
		Nausea and vomiting	0.6%	None	Section 10.17	LUME- Lung 1 CSR table 12.2.2.2.3:1 (Drug related grade 3/4/5)(5)
		Neutropenia	12.0%	None	Section 10.17	LUME- Lung 1 CSR table 12.2.2.2.3:1 (Drug related grade 3/4/5)(5)
		Rash	0.0%	None	Section 10.17	LUME- Lung 1 CSR table 12.2.2.4:1 (AEs of special interest grade 3/4/5)(5)
		Thrombocytopenia	1.2%	None	Section 10.17	LUME- Lung 1 CSR table 12.2.2.4:1 (AEs of special interest grade 3/4/5)(5)
		WBC count decreased	14.7%	None	Section 10.17	LUME- Lung 1 CSR table 12.2.2.2.3:1 (Drug related grade 3/4/5)(5)
	erlotinib	ALT increase	0.0%	None	Section 10.17	FDA PI, text below Table 3(125)
		Anaemia	0.0%	None	Section 10.17	FDA PI, Table 3(125)
		AST increase	0.0%	None	Section 10.17	FDA PI, text below Table 3(125)
		Diarrhoea – Grade 2	6.3%	None	Section 10.17	Kim et al. 2012(108)
		Diarrhoea – Grade 3 and 4	6.0%	None	Section 10.17	NICE TA162(34)
		Fatigue	19.0%	None	Section 10.17	NICE TA162(34)
		Febrile neutropenia	0.0%	None	Section 10.17	NICE TA162(34)
		Infection	2.0%	None	Section 10.17	NICE TA162(34)
		Nausea and vomiting	3.0%	None	Section 10.17	NICE TA162(34)
		Neutropenia	0.0%	None	Section 10.17	NICE TA162(34)
		Rash	9.0%	None	Section 10.17	NICE TA162(34)
		Thrombocytopenia	0.0%	None	Section 10.17	FDA PI, Table 3(125)
		WBC count decreased	0.0%	None	Section 10.17	FDA PI, Table 3(125)
AE costs	Cycle cost of each AE	ALT increase	£587	None	Section 7.5.7	Calculated from EE opinion
		Anaemia	£978	None	Section 7.5.7	Calculated from EE opinion
		AST increase	£336	None	Section 7.5.7	
		Diarrhoea – Grade 2	£250	None	Section 7.5.7	Calculated from EE opinion
		Diarrhoea – Grade 3 and 4	£1,796	None	Section 7.5.7	Calculated from EE opinion
		Fatigue	£370	None	Section 7.5.7	Calculated from EE opinion
		Febrile neutropenia	£2,012	None	Section 7.5.7	Calculated from EE opinion
		Infection	£2,181	None	Section 7.5.7	Calculated from EE opinion
		Nausea and vomiting	£1,919	None	Section 7.5.7	Calculated from EE opinion

Category	Variable		Value	Variance	Reference	Source
		Neutropenia	£346	None	Section 7.5.7	Calculated from EE opinion
		Rash	£639	None	Section 7.5.7	Calculated from EE opinion
		Thrombocytopenia	£422	None	Section 7.5.7	Calculated from EE opinion
		WBC count decreased	£423	None	Section 7.5.7	Calculated from EE opinion
	Cycle costs of AEs per treatment	nintedanib + docetaxel	£91.20	SE = 4.65 (= (10%/1.96) of mean)	Section 7.5.7	British National Formulary, EE opinion and LUME-Lung 1
		docetaxel	£92.09	SE = 4.70 (= (10%/1.96) of mean)	Section 7.5.7	British National Formulary, EE opinion and LUME-Lung 1
		erlotinib	£61.41	SE = 3.13 (= (10%/1.96) of mean)	Section 7.5.7	British National Formulary, EE opinion, NICE TA162, Kim et al. 2012 and FDA PI
Utility	Projection method after week 30		LUME-Lung 1 with linear trendline post week 30	NA	Section 7.4.3	NA
	PF nintedanib + docetaxel and docetaxel – week 0		0.710	SE = 0.01	Section 7.4.3	LUME-Lung 1, Table 17.1(5)
	PF nintedanib + docetaxel and docetaxel – week 3		0.721	SE = 0.01	Section 7.4.3	LUME-Lung 1, Table 17.1(5)
	PF nintedanib + docetaxel and docetaxel – week 6		0.707	SE = 0.01	Section 7.4.3	LUME-Lung 1, Table 17.1(5)
	PF nintedanib + docetaxel and docetaxel – week 9		0.699	SE = 0.01	Section 7.4.3	LUME-Lung 1, Table 17.1(5)
	PF nintedanib + docetaxel and docetaxel – week 12		0.692	SE = 0.01	Section 7.4.3	LUME-Lung 1, Table 17.1(5)
	PF nintedanib + docetaxel and docetaxel – week 15		0.687	SE = 0.01	Section 7.4.3	LUME-Lung 1, Table 17.1(5)
	PF nintedanib + docetaxel and docetaxel – week 18		0.682	SE = 0.01	Section 7.4.3	LUME-Lung 1, Table 17.1(5)
	PF nintedanib + docetaxel and docetaxel – week 21		0.677	SE = 0.02	Section 7.4.3	LUME-Lung 1, Table 17.1(5)
	PF nintedanib + docetaxel and docetaxel – week 24		0.671	SE = 0.02	Section 7.4.3	LUME-Lung 1, Table 17.1(5)
	PF nintedanib + docetaxel and docetaxel – week 27		0.666	SE = 0.02	Section 7.4.3	LUME-Lung 1, Table 17.1(5)
	PF nintedanib + docetaxel and docetaxel – week 30		0.661	SE = 0.02	Section 7.4.3	LUME-Lung 1, Table 17.1(5)
	PF nintedanib + docetaxel and docetaxel after week 30		-0.0057x + 0.7227	-	Section 7.4.3	Calculated assumption, x is time from model start
	PD nintedanib + docetaxel		0.638	SE = 0.01	Section 7.4.3	LUME-Lung 1, Table 17.1(5)
	PD docetaxel		0.638	SE = 0.01	Section 7.4.3	LUME-Lung 1, Table 17.1(5)
	PD erlotinib		0.638	SE = 0.01	Section 7.4.3	LUME-Lung 1, Table 17.1(5)
Disutility	Disutilities due to AEs	ALT increase	-0.05	SE = 0.01	Section 7.4.8	Assumption
		Anaemia	-0.07	SE = 0.01	Section 7.4.8	NICE TA192 (Nafees et al. 2008)(41, 126)
		AST increase	0.00	SE = 0.00	Section 7.4.8	Assumption
		Diarrhoea – Grade 2	-0.02	SE = 0.00	Section 7.4.8	Assumption: half of the grade 3 and 4 disutility
		Diarrhoea – Grade 3 and 4	-0.04	SE = 0.05	Section 7.4.8	Data on file, Table 18.1
		Fatigue	-0.21	SE = 0.03	Section	Data on file, Table 18.1

Category	Variable		Value	Variance	Reference	Source
					7.4.8	
		Febrile neutropenia	-0.09	SE = 0.01	Section 7.4.8	NICE TA192 (Nafees et al. 2008)(41, 126)
		Infection	-0.05	SE = 0.01	Section 7.4.8	Assumption
		Nausea and vomiting	-0.05	SE = 0.00	Section 7.4.8	NICE TA192 (Nafees et al. 2008)(41, 126)
		Neutropenia	-0.09	SE = 0.01	Section 7.4.8	NICE TA192 (Nafees et al. 2008)(41, 126)
		Rash	-0.03	SE = 0.00	Section 7.4.8	NICE TA192 (Nafees et al. 2008)(41, 126)
		Thrombocytopenia	-0.05	SE = 0.00	Section 7.4.8	NICE TA181(42)
		WBC count decreased	-0.05	SE = 0.01	Section 7.4.8	Assumption
	Average disutilities due to AEs	nintedanib + docetaxel	-0.049	SE = 0.01	Section 7.4.8	Calculated as weighted average of disutilities and per cycle frequencies of AEs
		docetaxel	-0.059	SE = 0.01	Section 7.4.8	Calculated as weighted average of disutilities and per cycle frequencies of AEs
		erlotinib	-0.110	SE = 0.01	Section 7.4.8	Calculated as weighted average of disutilities and per cycle frequencies of AEs

7.3.7 Are costs and clinical outcomes extrapolated beyond the trial follow-up period(s)? If so, what are the assumptions that underpin this extrapolation and how are they justified? In particular, what assumption was used about the longer term difference in effectiveness between the intervention and its comparator? For the extrapolation of clinical outcomes, please present graphs of any curve fittings to Kaplan–Meier plots.

Data for OS and PFS were extrapolated beyond the trial period due to the immature data.

Please see response to Section 7.3.1 for details of the extrapolation methods and graphs of curve fittings to Kaplan-Meier plots.

7.3.8 Provide a list of all assumptions in the de novo economic model and a justification for each assumption.

Model Assumptions

- The time horizon of 15 years was assumed to be a lifetime (the time when 99% of the patients in the nintedanib plus docetaxel arm were dead).
- The rate of future events was assumed to be independent of the events that occurred during previous cycles.

- A patient's history was not taken into account – those in the progressed health state were treated irrespective of their prior treatment options.
- Half-cycle correction was applied by taking the average number of patients in the previous and the current cycles in the different health states.

Assumptions Regarding Efficacy and Treatment Duration

- The efficacy data from the LUME LUNG 1 trial were applicable to the patient population in the UK and the outcome differences observed in the trial also translated to this population. LUME LUNG 1 was a multicentre trial with the UK as one of the participating countries. There was no reason to believe patients in the UK would respond differently to nintedanib.
- The adenocarcinoma population from the LUME Lung1 trial was the base-case, in line with licensing for nintedanib.
- For the primary comparator, PFS, OS, and treatment discontinuation observed in the treatment groups over the follow-up period of the LUME Lung1 trial could be extrapolated to the modelled time horizon, with the help of separately fitted log-logistic distributions for OS and separately fitted lognormal distribution for PFS, and exponential distribution for treatment discontinuation. Alternative distributions were explored in sensitivity analyses.
- Treatment discontinuation had the same probability throughout the time horizon.
- Patients on nintedanib plus docetaxel arm may discontinue only one treatment of docetaxel or nintedanib, or may discontinue together.

Assumptions Regarding Costs

- The premedications of the docetaxel and nintedanib plus docetaxel arms were similar, so they were not included to the model.
- The concomitant medication of docetaxel and nintedanib plus docetaxel was assumed to be similar, so they were not included in the model.
- The assumption was made that the end-of-life period was one month before death, irrespective of the treatment arm the patient is on. Note that in the base-case end of life costs were set to zero.
- Monitoring costs on BSC and on third-line therapy were assumed to be the same, independent of previous treatments, although patients discontinuing docetaxel had a chest x-ray every two to three months.

- The composition of third-line treatments was assumed to be the same after each second-line therapy.
- In the calculation of the medication costs, the model assumed a body surface area of 1.86 m^2 based on data from Sacco 2010 (120).

Assumptions Regarding AEs

- AEs were incorporated only for second-line treatment options.
- The management cost and the disutility associated with the individual AEs depended on neither the health state the patient was in nor the type of treatment administered.
- The rate of AEs was assumed to be constant over the time horizon. In clinical practice, however, AEs were likely to be experienced at different stages of treatment, particularly at initiation.
- In the analyses for erlotinib, the same AEs were/will be taken into account as for the primary comparator docetaxel, using the number of occurrences from the literature.

Assumptions Regarding Utilities

- Utility values were assumed to depend only on the health state a given patient was in (PF or PD) and on the patient experiencing an AE (disutilities), but not the treatment arm – only to the extent AE incidence was different.

7.4 Measurement and valuation of health effects

This section should be read in conjunction with NICE's 'Guide to the methods of technology appraisal', section 5.4.

The HRQL impact of adverse events should still be explored regardless of whether they are included in cost-effectiveness analysis.

All parameters used to estimate cost effectiveness should be presented clearly in tabular form and include details of data sources. For continuous variables, mean values should be presented and used in the analyses. For all variables, measures of precision should be detailed.

Patient experience

7.4.1 Please outline the aspects of the condition that most affect patients' quality of life.

Advanced NSCLC is a progressive disease with the majority of patients dying as a result of having it. The spread of the tumour may directly affect patients' QoL, leading to symptoms of cough, breathlessness and chest pain. Spread of the tumour systemically may lead to deterioration in global health status and activities involved in everyday life and as a result, a decline in HRQL outcomes such as role function, emotional, cognitive, social and physical functions. There may also be an increase in fatigue, nausea and vomiting, as well as a reduction in appetite.

7.4.2 Please describe how a patient's HRQL is likely to change over the course of the condition.

See Section 7.4.1.

HRQL data derived from clinical trials

7.4.3 If HRQL data were collected in the clinical trials identified in section 6 (Clinical evidence), please comment on whether the HRQL data are consistent with the reference case. The following are suggested elements for consideration, but the list is not exhaustive.

- Method of elicitation.
- Method of valuation.
- Point when measurements were made.
- Consistency with reference case.
- Appropriateness for cost-effectiveness analysis.
- Results with confidence intervals.

PF and Post-progression Utilities

QoL data were collected in the LUME-Lung 1 trial using the EQ-5D instrument. Data from the LUME-Lung 1 trial were analysed using a longitudinal model adjusting for baseline Eastern Cooperative Oncology Group (ECOG) score, prior treatment with bevacizumab, presence of brain metastases, controlling for health status (progression-free or progressed), and key AEs(127). The analysis estimated utility values over time for PF patients from week 0 to 30 in

three-week intervals – without a treatment term. For the post-progression period, although utility values over time could be generated similarly to those for the PF state, mean utilities for the post-progression period were used in the model to accommodate the memory-less feature of the Markov approach; model inputs on utilities are displayed in Table 80.

Table 80: Utilities for PF and Post-progression States

Nintedanib + Docetaxel and Docetaxel – Pooled	PF, Without AEs	
	Mean	SE
Week 0	0.710	0.01
Week 3	0.721	0.01
Week 6	0.707	0.01
Week 9	0.699	0.01
Week 12	0.692	0.01
Week 15	0.687	0.01
Week 18	0.682	0.01
Week 21	0.677	0.02
Week 24	0.671	0.02
Week 27	0.666	0.02
Week 30	0.661	0.02
Treatment arm	**Progressed Health State**	
Nintedanib + docetaxel	0.64	0.01
Docetaxel	0.64	0.01

AE = Adverse event; PF = Progression-free; SE = standard error

There are two utility calculation options built into the model for the utility extrapolation for the PF state while on active second-line therapy after week 30:

1. Last observation carried forward (LOCF)

   The LOCF option assumes that beyond week 30, the utility of PF patients is equal to the utility at week 30.

2. Linear trend line

   Alternatively, linear trend line is fitted to the utility data. The equation of the line is -0.0057 * time (in cycles) + 0.7227. This trend line is used for the calculation of utilities beyond week 30, until it drops down below the utility of post-progression health state. The trendline was fitted from cycle 3 – when utilities started to decrease.

The linear trend line was used in the base-case analysis as it allows modelling of continuing change in utility in the PF state beyond the trial data.

In addition, the model also includes an option to implement the utilities from a prospective cross-sectional patient survey in a real-world setting (128), the details of which can be found in Table 82 in Section 7.4.6.

To be in line with the NICE reference case(102) (which specifies that utilities should be evaluated by the EQ-5D as measured by patients), the base-case utilities are from the LUME-Lung 1 trial. An assumption of linear extrapolation of trend observed until week 30 for the PF health state is employed in the base-case.

Mapping

7.4.4 If mapping was used to transform any of the utilities or quality-of-life data in clinical trials, please provide the following information.

- Which tool was mapped from and onto what other tool? For example, SF-36 to EQ-5D.
- Details of the methodology used.
- Details of validation of the mapping technique.

As EQ-5D was directly measured, no mapping was required.

HRQL studies

7.4.5 Please provide a systematic search of HRQL data. Consider published and unpublished studies, including any original research commissioned for this technology. Provide the rationale for terms used in the search strategy and any inclusion and exclusion criteria used. The search strategy used should be provided in section 10.12, appendix 12.

The utility search was developed to identify studies reporting the utility in the three stages of the Markov model: PF disease, PD and death. The disutility due to AEs was also captured. Limits for the database search included searching only for items with abstracts, on humans, and published from 2002 onwards. The search was also designed to exclude non-systematic reviews.

The methodology followed is explained in Section 6.1 and Appendix 10.12. A single systematic literature review was performed for the clinical, cost-effectiveness, resource use and cost data, as well as studies reporting utility scores for health states within the model.

The procedures for study selection in the humanistic review were similar to those described above for all studies. Studies were included in the humanistic review if they met the criteria outlined in Table 81.

Table 81: Humanistic Review: Inclusion and Exclusion Criteria

	Inclusion Criteria	Exclusion Criteria
Population	**NSCLC**	Any patient population other than NSCLC
Interventions	Where relevant: • All relevant chemotherapy for RR NSCLC (all lines of therapy): o Monotherapy o Combination chemotherapy • Other interventions that are considered standard care in the patient population that will be relevant to the economic model	NA
Outcomes	• Humanistic outcomes related to the patient population (may or may not be related to any intervention), from real-world observational studies: Utilities (EQ-5D) • Other HRQL outcomes	No outcomes of interest included
Study design	Any	NA
Language restrictions	English language	Non-English language
Date	2002 onwards*	Prior to the year 2002*
Country	Any	None

*Abstracts published prior to the year 2011 and systematic reviews published prior to the year 2009 were excluded.

Results

The search of the literature yielded 7,289 citations. De-duplication resulted in the removal of 2,533 overlapping citations. Following screening of the remaining 4,756 studies, 4,419 studies were excluded. Full text was obtained for the remaining 337 studies; 41 were excluded due to an incorrect date; 1 due to incorrect topic; 13 because the study was not in English; 39 because of incorrect population; 85 because the intervention did not match the original search criteria; 4 because of incorrect outcome; 19 because of incorrect study design and 2 because they were duplicates (total excluded = 204).

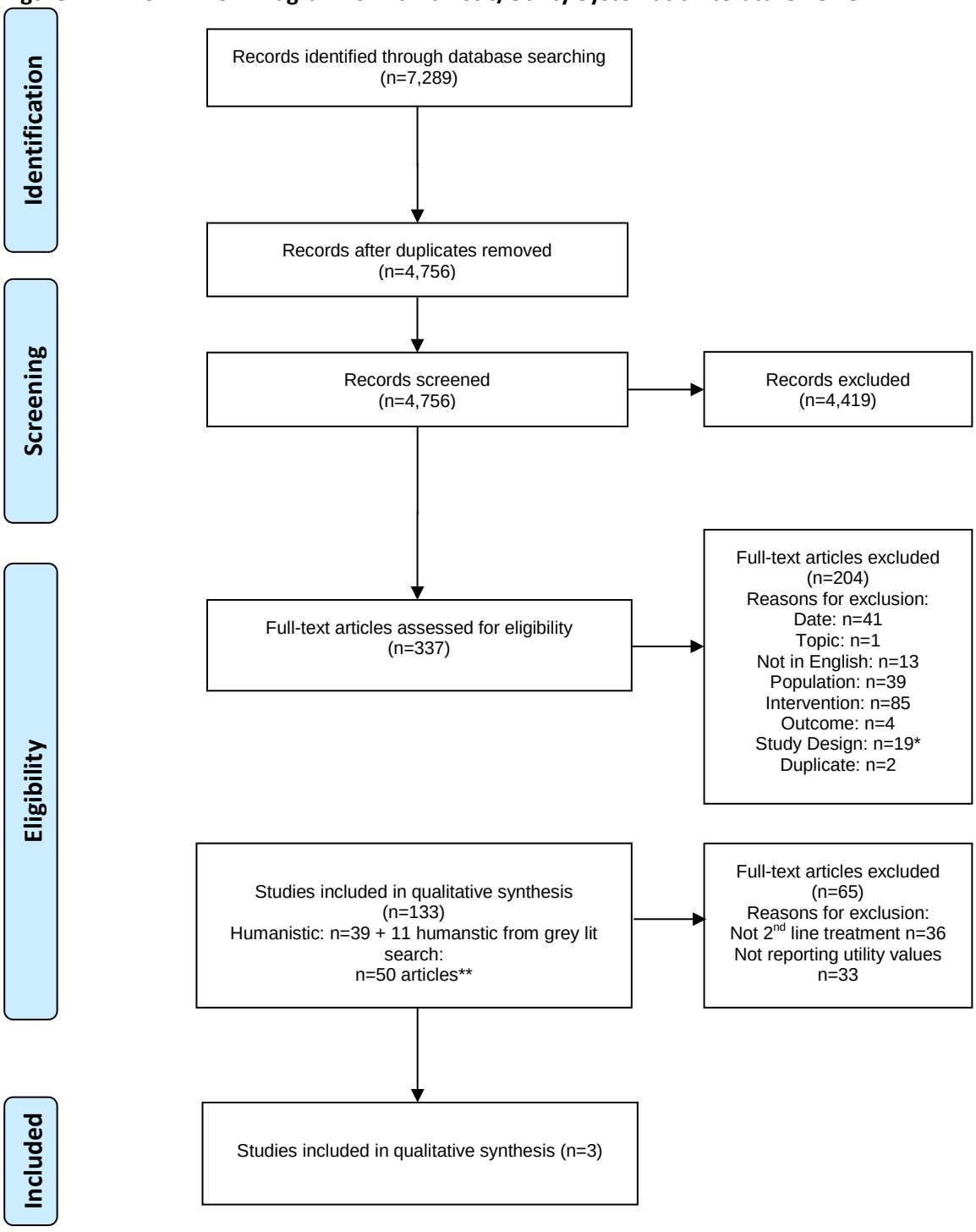
Of the remaining 133 studies, 39 reported QoL outcomes.

An additional 11 articles reporting QoL outcomes were retrieved from 'grey' literature sources, and an additional 18 papers containing QoL data from the clinical efficacy and safety topic were included in this section of the review. This gave a total of 68 studies providing data on HRQL in patients with NSCLC.

Of these 68 studies, 29 reported HRQL in patients receiving second-line treatment, and of these one study reported utilities (126) and two studies reported corresponding EQ-5D values (129, 130) (Table 82 in Section 7.4.6). The flow of studies in the systematic literature review is presented in Figure 27. Note this figure shows the initial search results of the entire literature review (as described in Section 6.1 and Appendix 10.2), including searches of economic, resource use, utility and clinical searches as this was a combined search. The flow then demonstrates how studies were included or excluded according to the criteria relevant to the search of interest.

In addition to the full text papers reviewed according to the 3 studies identified in the PRISMA diagram below (Figure 27), Chouaid et al. (2013)(128) was identified as a relevant study (produced by Boehringer Ingelheim) containing HRQL and utility in patients with advanced NSCLC. The data from this study was extracted alongside the other 3 studies identifies, and is presented in Table 82 in Section 7.4.6.

Figure 27: PRISMA Flow Diagram for Humanistic/Utility Systematic Literature Review



* The reference lists of the systematic reviews were assessed for additional relevant studies; no additional studies were identified.

** An additional 18 papers containing quality-of-life data from the clinical efficacy and safety topic were included in the humanistic review

7.4.6 Provide details of the studies in which HRQL is measured. Include the following, but note that the list is not exhaustive.

- Population in which health effects were measured.

- Information on recruitment.
- Interventions and comparators.
- Sample size.
- Response rates.
- Description of health states.
- AEs.
- Appropriateness of health states given condition and treatment pathway.
- Method of elicitation.
- Method of valuation.
- Mapping.
- Uncertainty around values.
- Consistency with reference case.
- Appropriateness for cost-effectiveness analysis.
- Results with confidence intervals.
- Appropriateness of the study for cost-effectiveness analysis.

Of the three studies reported in Table 82 which were included from the PRISMA diagram above (Figure 27) two were considered not appropriate in this submission for the reasons given in the table. One study(126) was considered relevant to this submission, collecting utility scores for NSCLC health states and toxicities commonly associated with chemotherapy treatments from a random sample representative of the UK general adult population. However, this study recorded utilities for either stable disease or treatment response states rather than in a pre-progression state, and utilities were derived from either standard gamble or visual analogue scales given to the general population rather than the EQ-5D given to patients. Research has suggested that values from patients are more appropriate as they are based on experience rather than on preferences as in the case of the general public being asked to value a hypothetical state of health(131, 132).

For this reason the data from Chouaid et al 2013 (128) was implemented in the model as a sensitivity analysis for the utilities for pre-progression and post-progression states. Although it reports utilities recorded from patients in Europe, Canada, Australia and Turkey as well as the UK, Chouaid et al 2013 (128) used the EQ-5D with relevant patients to obtain utilities for

the states used in the cost effectiveness model for this submission, and was the most recent study analysed.

Note: For the post-progression state in this sensitivity analysis, a conservative assumption was used; the utility is assumed to be equal to the third/fourth line progressive disease state. In reality, the patients in the model are more likely to also include patients from the second-line progressive disease and third/fourth line PF states, both of which have higher utilities than the third/fourth line progressive disease state.

Table 82: Utility studies reporting health states for NSCLC

Study	Population	Recruitment	Sample size and response	Description of health states, & appropriateness	AEs reported in study	Methods of elicitation, valuation and mapping	Results with CIs	Appropriateness for cost-effectiveness analysis	Appropriateness to this submission
Blackhall et al. 2013(130)	Patients with advanced ALK-positive NSCLC randomised to receive second-line treatment with crizotinib or standard chemotherapy from various countries	Participants were recruited from various countries including the United Kingdom, as part of a randomised control trial (PROFILE 1007 study)	343 (responder rates were ≥85% in each group)	Higher scores (range 0–100) indicated higher symptom severity or better functioning/QOL	Constipation Diarrhoea	EORTC QLQ-C30 and QLQ-LC13	EQ-5D Baseline Crizotinib Group: Mean (SD) = 0.72 (0.25) EQ-5D Baseline Chemotherapy Group: Mean (SD) = 0.69 (0.26) VAS Baseline Crizotinib Group: Mean (SD) = 64.09 (21.04) VAS Baseline Chemotherapy Group: Mean (SD) = 66.76 (20.74) EQ-5D After Treatment Crizotinib Group: Mean (SD) = 0.82 (0.01) EQ-5D After Treatment Chemotherapy Group: Mean (SD) = 0.73 (0.02); P<0.001 A significantly greater overall improvement from baseline was observed in VAS scores in the crizotinib arm compared with chemotherapy (4.68 vs −6.06; P< 0.001).	Not appropriate – an international study available only in abstract form. Eq-5D or EQ-VAS data only reported by treatment arm at baseline and end of second-line treatment; no utility values reported for specific AEs or for progression versus PF.	Not appropriate

Study	Population	Recruitment	Sample size and response	Description of health states, & appropriateness	AEs reported in study	Methods of elicitation, valuation and mapping	Results with CIs	Appropriateness for cost-effectiveness analysis	Appropriateness to this submission
Nafees et al. 2008(126)	Random sample representative of the UK general adult population	Participants were recruited from the Greater London area through a volunteer database, advertisements and a study recruitment website.	100	Progressive disease Treatment response Treatment response (with diarrhoea) Treatment response (with fatigue) Treatment response (with febrile neutropenia) Treatment response (with hair loss) Treatment response (with nausea/vomiting) Treatment response (with neutropenia) Treatment response (with rash) Stable disease Stable disease (with diarrhoea) Stable disease (with fatigue) Stable disease (with febrile neutropenia) Stable disease (with hair loss) Stable disease (with nausea/vomiting) Stable disease (with neutropenia) Stable disease (with rash)	Diarrhoea Fatigue Febrile neutropenia Hair loss Nausea/vomiting Neutropenia Rash	Standard Gamble and Visual Analogue Scale	Progressive disease: 0.473 Treatment response: 0.673 Treatment response (with diarrhoea): 0.626 Treatment response (with fatigue): 0.599 Treatment response (with febrile neutropenia): 0.582 Treatment response (with hair loss): 0.628 Treatment response (with nausea/vomiting): 0.624 Treatment response (with neutropenia): 0.583 Treatment response (with rash): 0.640 Stable disease: 0.653 Stable disease (with diarrhoea): 0.606 Stable disease (with fatigue): 0.580 Stable disease (with febrile neutropenia): 0.563 Stable disease (with hair loss): 0.608 Stable disease (with nausea/vomiting): 0.605	Appropriate	Not appropriate- Utility measured from general population rather than patients

Study	Population	Recruitment	Sample size and response	Description of health states, & appropriateness	AEs reported in study	Methods of elicitation, valuation and mapping	Results with CIs	Appropriateness for cost-effectiveness analysis	Appropriateness to this submission
							Stable disease (with neutropenia):0.563 Stable disease (with rash): 0.621		
Schuette et al. 2012(129)	Patients with stage III/IV NSCLC receiving second-line pemetrexed treatment in routine clinical practice in Germany and Austria.	Participants were recruited from in- and outpatient settings in 102 hospitals and practices in Germany and Austria, as part of a prospective, non-interventional phase IV study.	521 (231 returned EQ-5D and 225 returned the VAS)	EQ-5D: Mobility, Self Care, Usual Activities, Pain/Discomfort, Anxiety and Depression. Rated on scale- 1= Some problems, 2= Moderate Problems, 3= Extreme problems VAS: Rated on a scale of 0 to 100- 0 indicating worst	Fatigue/asthenia Neutropenia Nausea Febrile neutropenia Rash/desquamation Stomatitis/pharyngitis Mucositis Vomiting Diarrhoea	EQ-5D and Visual Analogue Scale	EQ-5D Baseline: Mean (SD) = 0.66 (0.256) EQ-5D 2nd Treatment Cycle: Mean increase (SD) = 0.02 (0.214); P<0.003 EQ-5D 6th Treatment Cycle: Mean increase (SD) = 0.11 (0.228); P<0.001 VAS Baseline: Mean (SD) = 59.3 (17.80) VAS 2nd Treatment	Appropriate	Not appropriate – included only patients in Germany and Austria.

Study	Population	Recruitment	Sample size and response	Description of health states, & appropriateness	AEs reported in study	Methods of elicitation, valuation and mapping	Results with CIs	Appropriateness for cost-effectiveness analysis	Appropriateness to this submission
				health state imaginable, 100 indicating best health state imaginable			Cycle: Means increase (SD) = 3.3 (12.58); P<0.001 VAS 6th Treatment Cycle: Means increase (SD) = 12.8 (17.62); P<0.001		
Chouaid et al 2013(128)	Patients with ECOG status 0-2, receiving first, second or third/fourth line pharmacoth erapy or BSC	Patients were enrolled prospectively at a total at 25 hospitals in Australia, Belgium, Canada, France, Italy, Turkey, The Netherlands, Sweden and the UK.	319 enrolled with 56 excluded.	Patients were stratified into predefined health states according to line of therapy and disease status (PF [PF]/progressive disease [PD]). There was no specific question in the survey asking which line of treatment the patient was currently undergoing, consequently, the line of treatment variable was derived from information on the number		EQ-5D	All patients (N= 263) Mean utility (SD) 95% CI =0.66 (0.29) 0.62-0.69 First-line PF (N=115) 0.71 (0.24) 0.67-0.76 First-line PD (N=26) 0.67 (0.2) 0.59-0.75 Second-line PF (N=47) 0.74 (0.18) 0.68-0.80 Second-line PD (N=17) 0.59 (0.34) 0.42-0.77 Third/fourth line PF (N=25) 0.62 (0.29) 0.49-0.74	Appropriate	Appropriate

Study	Population	Recruitment	Sample size and response	Description of health states, & appropriateness	AEs reported in study	Methods of elicitation, valuation and mapping	Results with CIs	Appropriateness for cost-effectiveness analysis	Appropriateness to this submission
				of previous lines received.			Third/fourth line PD (N=21) 0.46 (0.38) 0.28-0.63		

7.4.7 Please highlight any key differences between the values derived from the literature search and those reported in or mapped from the clinical trials.

The model uses utility values derived from LUME Lung 1 in the base-case (see Table 80 in Section 7.4.3). Utility values from the literature are also tested within the model. Table 83 shows the utility values explored within the model.

Table 83: Utilities from Chouaid et al. (2013)(128)

	Mean (SE)
PF	0.74 (0.03)
Post-progression	0.46 (0.08)

PF = Progression-free; SE = Standard error

The PF utilities from the LUME Lung 1 trial are generally lower than the PF utilities from Chouaid et al 2013(128), however the utilities from the trial represent the utilities of patients from the trial taken at regular intervals. This provides a strong source of accurate information for the PF utilities.

The post-progression utilty from the LUME Lung 1 trial is higher than that from Choauid et al 2013(128), however the post-progression state in Table 83, is a conservative assumption; the utility is assumed to be equal to the third/fourth line progressive disease state. In reality, the patients in the model are more likely to also include patients from the second-line progressive disease and third/fourth line PF states, both of which have higher utilities than the third/fourth line progressive disease state.

Adverse events

7.4.8 Please describe how adverse events have an impact on HRQL.

The model also included the impact of AEs on HRQL where utility decrement associated with each AE was applied for a period of one model cycle (assumption-based). Disutilities due to AEs are presented in Table 84.

Advisory board members highlighted that some patients may experience multiple AEs at the same time (e.g. fatigue along with anaemia) (Section 7.3.5). As a result, the model may have double counted disutilities; this conservative approach was used in the base-case.

Table 84: Disutilities associated with AEs

Adverse Event	Disutility	Sources
ALT increased	-0.05	Assumption
Anemia	-0.07	Nafees et al. 2008(126)
Diarrhoea - grade 2	-0.02	Assumption: half of the disutility for grade 3/4 diarrhoea
Diarrhoea - grade ¾	-0.04	BI Data on file, Table 18.1(127)
Fatigue	-0.21	BI Data on file, Table 18.1(127)
Febrile neutropenia	-0.09	NICE TA192(41), Nafees et al. 2008(126)
Infection	-0.05	Assumption
Liver-related investigations	-0.05	Assumption
Nausea and vomiting	-0.05	Nafees et al. 2008(126)
Neutropenia	-0.09	Nafees et al. 2008(126)
Neutrophil count decreased	-0.09	Assumption: same as disutility of neutropenia
Rash	-0.33	Nafees et al. 2008(126)
Thrombocytopenia	-0.05	NICE TA181(42)
WBC count decreased	-0.05	Assumption

Quality-of-life data used in cost-effectiveness analysis

7.4.9 Please summarise the values you have chosen for your cost-effectiveness analysis in the following table, referencing values obtained in sections 7.4.3 to 7.4.8. Justify the choice of utility values, giving consideration to the reference case.

See Section 7.4.3 Table 80 for utilities used within the model and explanation; and Section 7.4.8, Table 84 for disutilities applied for AEs and explanation. For the variance of utility values used, see Section 7.3.6, Table 79.

7.4.10 If clinical experts assessed the applicability of values available or estimated any values, please provide the following details[7]:

- the criteria for selecting the experts
- the number of experts approached
- the number of experts who participated
- declaration of potential conflict(s) of interest from each expert or medical specialist whose opinion was sought

[7] Adapted from Pharmaceutical Benefits Advisory Committee (2008) Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee (Version 4.3). Canberra: Pharmaceutical Benefits Advisory Committee.

- the background information provided and its consistency with the totality of the evidence provided in the submission
- the method used to collect the opinions
- the medium used to collect opinions (for example, was information gathered by direct interview, telephone interview or self-administered questionnaire?)
- the questions asked
- whether iteration was used in the collation of opinions and if so, how it was used (for example, the Delphi technique).

Although the utility values used were not specifically critiqued by clinicians, a discussion of patient reported outcomes was conducted during the advisory board described in Section 7.3.5.

7.4.11 Please define what a patient experiences in the health states in terms of HRQL. Is it constant or does it cover potential variances?

The HRQL of patients in the PF state from week 0 to 30 was estimated in three-week intervals and modelled according to data from LUME-Lung 1. Beyond this point, the change in patient HRQL in the PF state is based upon a linear trend line following change from cycle 3.

In addition, changed in HRQL also occur for the following reasons:

- When a patient's diseases progresses
- Experience of AEs:
- Death

Mean utilities from LUME-Lung 1 were used for the post-progression period in the model to accommodate the memory-less feature of the Markov approach. Model inputs on utilities are displayed in Table 80 in Section 7.4.3).

The decrease in QoL of carers has not been captured.

7.4.12 Were any health effects identified in the literature or clinical trials excluded from the analysis? If so, why were they excluded?

Anorexia was removed from the model as no disutility for it could be identified and cost impact was small.

7.4.13 If appropriate, what was the baseline quality of life assumed in the analysis if different from health states? Were quality-of-life events taken from this baseline?

The baseline QoL was the same as the QoL of patients in PF disease state. This is appropriate as this is the starting point for the analysis and the cohort.

7.4.14 Please clarify whether HRQL is assumed to be constant over time. If not, provide details of how HRQL changes with time.

Please see the response to Section 7.4.11.

7.4.15 Have the values in sections 7.4.3 to 7.4.8 been amended? If so, please describe how and why they have been altered and the methodology.

No changes.

7.5 Resource identification, measurement and valuation

This section should be read in conjunction with NICE's 'Guide to the methods of technology appraisal', section 5.5.

All parameters used to estimate cost effectiveness should be presented clearly in a table and include details of data sources. For continuous variables, mean values should be presented and used in the analyses. For all variables, measures of precision should be detailed.

NHS costs

7.5.1 Please describe how the clinical management of the condition is currently costed in the NHS in terms of reference costs and the payment by results (PbR) tariff. Provide the relevant Healthcare Resource Groups (HRG) and PbR codes and justify their selection. Please consider in reference to section 2.

Table 85: National Schedule of Reference Costs Year: 2012–13 – NHS trusts and NHS foundation trusts – Elective Inpatients. Weighted Average National Cost. HRG Data(121)

Currency Code	Currency Description
DZ17E	Respiratory Neoplasms with CC Score 11+
DZ17F	Respiratory Neoplasms with CC Score 8-10
DZ17G	Respiratory Neoplasms with CC Score 5-7
DZ17H	Respiratory Neoplasms with CC Score 3-4
DZ17J	Respiratory Neoplasms with CC Score 1-2
DZ17K	Respiratory Neoplasms with CC Score 0

Table 86: National Schedule of Reference Costs Year: 2012–13 – NHS trusts and NHS foundation trusts – Chemotherapy – Outpatient. HRG Data(121)

Currency Code	Currency Description
SB12Z	Outpatient – Deliver simple Parenteral Chemotherapy at first attendance

These codes were chosen as they represent the disease of interest and the appropriate inpatient and outpatient data, including relevant resource use associated with chemotherapy in England and Wales.

7.5.2 Please describe whether NHS reference costs or PbR tariffs are appropriate for costing the intervention being appraised.

As discussed in Section 7.5.1, the appropriate HRG costing was used for this model.

Resource identification, measurement and valuation studies

7.5.3 Please provide a systematic search of relevant resource data for the UK. Include a search strategy and inclusion criteria, and consider published and unpublished studies. The search strategy used should be provided as in section 10.13, appendix 13. If the systematic search yields limited UK-specific data, the search strategy may be extended to capture data from non-UK sources. Please give the following details of included studies:

- country of study
- date of study
- applicability to UK clinical practice
- cost valuations used in study
- costs for use in economic analysis
- technology costs.

The objective of this search was to identify costs and resource use associated with NSCLC. The search centred on previously published cost-effectiveness models and followed the methodology explained in section 6.1. A single systematic literature review was performed for the clinical, cost-effectiveness, utility scores and resource use and cost data.

The resource use search inclusion and exclusion criteria are shown in Table 87.

Table 87: Cost-effectiveness Inclusion and Exclusion Criteria

	Inclusion Criteria	Exclusion Criteria
Population	**RR NSCLC** (receiving second-line chemotherapy or relapsed/refractory to first-line chemotherapy)	Any patient population other than RR NSCLC
Interventions	Any second-line chemotherapy for RR NSCLC: • Monotherapy • Combination therapy with other chemotherapy Other interventions that are considered standard care in the patient population that will be relevant to the economic model	Patients who were treatment-naïve or had received more than first-line therapy
Outcomes	Economic models: • Cost-utility analyses • Cost-effectiveness analyses • Cost-benefit analyses • Cost-minimisation analyses	No outcomes of interest included
Study design	Economic models: Economic studies	Not an economic model
Language restrictions	English language	Non-English language
Date	Economic models: 2002 onwards	Prior to the year 2002*
Country	UK	Not UK

*Abstracts published prior to the year 2011 and systematic reviews published prior to the year 2009 were excluded.

Quality assessment

A quality assessment score was derived from that of Drummond (1997) to assess the quality of included economic models(133).The Drummond criteria were created to support the generation of high-quality, rigorous economic evaluations. They involve using a total of 36 questions to assess three broad areas of the studies, namely: study design; data collection; and analysis and interpretation of results. The result of the assessment process is a summary quality score on which models are judged to be either: high (++), moderate (+) or low quality (-). High-quality studies are considered to report clearly on almost all of the Drummond quality criteria questions, while studies of low quality do not report on most items. In this

review, only studies in full-text form underwent a quality assessment because of the lack of details available for assessment in abstracts and posters.

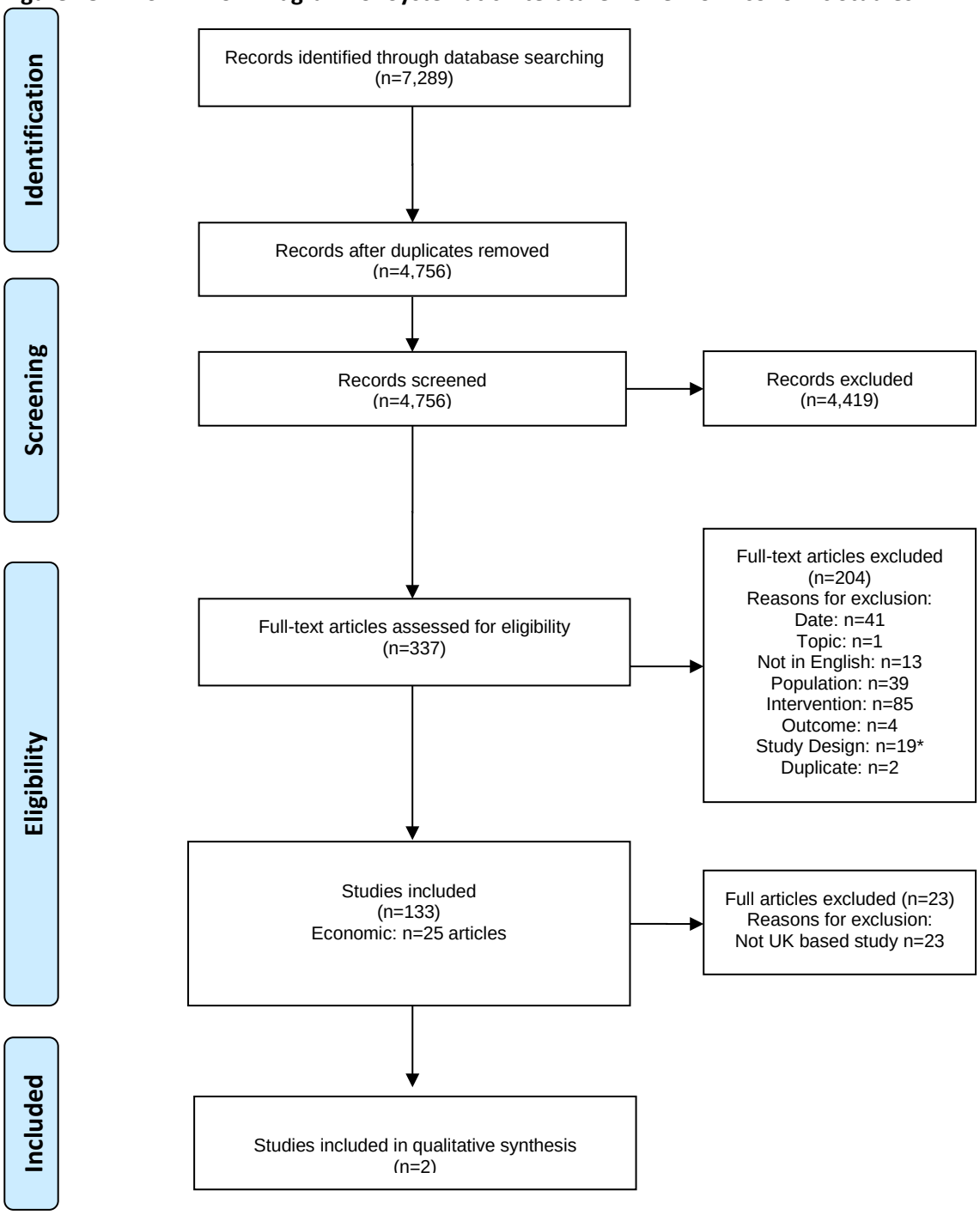
Results

The search of the literature yielded 7,289 citations. De-duplication resulted in the removal of 2,533 overlapping citations. Following screening of the remaining 4,756 studies, 4,419 studies were excluded. Full text was obtained for the remaining 337 studies; 41 were excluded due to an incorrect date; 1 due to incorrect topic; 13 because the study was not in English; 39 because of incorrect population; 85 because the intervention did not match the original search criteria; 4 because of incorrect outcome; 19 because of incorrect study design and 2 because they were duplicates (total excluded = 204).

Of the remaining 133 studies, 25 were economic studies, 2 of which were UK based studies. The flow of studies in the systematic literature review is presented in Figure 28. Note this figure shows the initial search results of the entire literature review (as described in section 7.5.3), including searches of economic, resource use, utility and clinical searches as this was a combined search. The flow then demonstrates how studies were included or excluded according to the criteria relevant to the search of interest.

Data was extracted from the two UK based studies and is reported in Table 88.

Figure 28: PRISMA Flow Diagram for Systematic Literature Review on Economic Studies



* The reference lists of the systematic reviews were assessed for additional relevant studies; no additional studies were identified.

Table 88: Resource use and cost data extracted from SLR of cost effectiveness models

Study	Country	Date	Applicability to UK	Cost valuations used	Cost/resource data	Applicability to this evaluation
Holmes, J., D. Dunlop, et al. (2004). "A cost-effectiveness analysis of docetaxel in the second-line treatment of non-small cell lung cancer." Pharmacoeconomics 22(9): 581-589.(134)	UK	Cost year 2000/2001 Published 2004	Applicable	Bottom-up	**Docetaxel related costs:** Treatment cost: £4,338 (range: £3,438-£5,238) Administration cost: £77 (range: £61-£93) Co-drug cost: £17 (£13-£20) Net costs per patient: £4,432 (£3,512-£5,351)	Not applicable-limited data provided
Lewis, G., M. Peake, et al. (2010). "Cost-effectiveness of erlotinib versus docetaxel for second-line treatment of advanced NSCLC in the United Kingdom." J Int Med Res 38(1): 9-21.(135)	UK	Cost year not recorded Published 2010	Applicable	Not recorded	**Total drug cost:** Erlotinib: £6,796 Docetaxel: £4,656 Difference: £2,140 **Total cost of progression-free health states:** Erlotinib: £1,482 Docetaxel: £1,201 Difference: £281 **Total cost of disease progression** Erlotinib: £5,309 Docetaxel: £6,151 Difference: £-842 **Total drug administration cost** Erlotinib: £0 Docetaxel: £1,188 Difference: £-1,188 **Total AE cost** Erlotinib: £143 Docetaxel: £760 Difference: £-617 **Total cost** Erlotinib: £13,730 Docetaxel: £13,956 Difference: £=226	Not applicable-limited data

In addition to the search described above, the National Institute for Health and Care Excellence was searched to identify the key HTA submissions for submissions for NSCLC.

Results of HTA search:

- National Institute for Health and Care Excellence (2006). Pemetrexed for the treatment of relapsed NSCLC (TA124).(65) (Table 89).
- National Institute for Health and Care Excellence (2008). Erlotinib for the treatment of relapsed NSCLC (TA162)(34) (Table 90).
- National Institute for Health and Care Excellence (2010). Gefitinib for the first-line treatment of locally advanced or metastatic NSCLC (TA192).(41) (Table 91).
- National Institute for Health and Care Excellence (2012). Erlotinib for the first-line treatment of locally advanced or metastatic EGFR-TK mutation-positive NSCLC (TA258).(40) (Table 92).
- National Institute for Health and Care Excellence (2014). Afatinib for treating EGFR mutation-positive locally advanced or metastatic NSCLC (TA310).(43) (Table 93).
- National Institute for Health and Care Excellence (2013). Crizotinib for previously treated NSCLC associated with an anaplastic lymphoma kinase fusion gene.(136) (Table 94).
- National Institute for Health and Care Excellence (2009). Pemetrexed for the first-line treatment of NSCLC (TA181).(42) (Table 95).

Table 89: Resource use and costs used in TA124: pemetrexed for the first-line treatment of NSCLC

Resource use and costs utilised in the company model	Unit cost*	Source
Chemotherapy agents		
Pemetrexed	£800 per 500mg vial	British National Formulary 2006
Docetaxel 0.5ml – 20mg	£162.75	British National Formulary 2006
Docetaxel 2ml – 80mg	£534.75	British National Formulary 2006
BSC	£2,158	Lees 2002
Pre-medications		
Dexamethasone	£42.30	British National Formulary 2006
Folic acid	£2.24	British National Formulary 2006
Vitamin B12	£2.46	British National Formulary 2006
Piriton	£0.19	British National Formulary 2006
Paracetamol	£0.31	British National Formulary 2006
AE-related treatments		
Blood transfusion – whole	£125.07	National Blood Bank
Blood transfusion – platelets	£206.34	National Blood Bank
Blood transfusion – standard red cells	£124.80	National Blood Bank

Resource use and costs utilised in the company model	Unit cost*	Source
Steroid cream (Betnovate)	£3.34	British National Formulary 2006
Lomotil	£1.63	British National Formulary 2006
Domperidone	£2.47	British National Formulary 2006
Haemoglobin levels	£3.04	NHS Reference Costs
Electrolytes	£1.65	NHS Reference Costs
Blood cultures	£3.04	NHS Reference Costs
Stool cultures	£6.59	NHS Reference Costs
Complete blood cell count	£3.04	NHS Reference Costs
Differential white blood cell count	£3.04	NHS Reference Costs
Platelet count	£3.04	NHS Reference Costs
Liver function tests	£1.65	NHS Reference Costs
Treatment for febrile neutropenia	£3,860.30	Holmes et al., (2004)
1 day in hospital: chemotherapy with a respiratory system primary diagnosis – non-elective admission	£250.19	NHS Reference costs
Administration time		
Clinic time (1 hour) D98: Chemotherapy with a respiratory system primary diagnosis	£62.91	NHS Reference costs
Palliative care costs	£3,236	NICE (2004)

Table 90: Resource use and costs used in TA162: erlotinib for the second-line treatment of NSCLC

Model variables	Value	Source
Patient Survival		
Costs		
Erlotinib	£54.38 per day	British National Formulary
Docetaxel	£1,023 per cycle	British National Formulary
PFS	£327 per month	Expert panel (resource use); schedule of reference costs and PSSRU (unit costs)
PPS	£988 per month	Expert panel (resource use); schedule of reference costs, British National Formulary and PSSRU 2004 (unit costs)
Docetaxel drug administration	£202 per month	Expert panel (resource use); not stated (unit costs)
Cost per episode of rash	£117	Expert panel (resource use); schedule of reference costs and British National Formulary (unit costs)
Cost per episode of anorexia	£119	Expert panel (resource use); schedule of reference costs and British National Formulary (unit costs)
Cost per episode of diarrhoea	£237	Expert panel (resource use); schedule of reference costs, British National

Model variables	Value	Source
		Formulary and PSSRU 2004 (unit costs)
Cost per episode of nausea	£240	Expert panel (resource use); schedule of reference costs, British National Formulary and PSSRU 2004 (unit costs)
Cost per episode of infection	£1227	Expert panel (resource use); schedule of reference costs (unit costs)
Cost per episode of stomatitis	£188	Expert panel (resource use); schedule of reference costs (unit costs)
Cost per episode of neutropenia	£375	Expert panel (resource use) and schedule of reference costs and British National Formulary (unit costs)
Cost per episode of fatigue	£19	Expert panel (resource use); schedule of reference costs and British National Formulary (unit costs)
Cost per episode of neuropathy	£18	Expert panel (resource use); schedule of reference costs and PSSRU 2004 (unit costs)

Table 91: Resource use and costs specified in TA192: Gefitinib for the first-line treatment of locally advanced or metastatic NSCLC

Model variable	Value	Source
Costs		
Gefitinib (single fixed payment per patient)	Marked out as commercial in confidence	AstraZeneca Commercial in Confidence
EGFR mutation test (per test)	Marked out as commercial in confidence	Lab 21 Commercial Contract
Gefitinib patient monitoring (per month)	£86	Reference costs (2009/08)
Drug acquisition gem/carb (per cycle)	£999	British National Formulary (2009), Dictionary of Medicines and Devices
Drug acquisition pac/carb (per cycle)	£1,489	British National Formulary (2009)
Drug acquisition vin/cis (per cycle)	£403	British National Formulary (2009)
Drug acquisition gem/cis (per cycle)	£795	British National Formulary (2009), Dictionary of Medicines and Devices
Administration gem/carb (per cycle)	£307	Reference costs (2007/08)
Administration pac/carb (per cycle)	£153	Reference costs (2007/08)
Administration vin/cis (per cycle)	£527	Reference costs (2007/08)
Administration gem/cis (per cycle)	£527	Reference costs (2007/08)
Drug acquisition g-CSF (per patient treated)	£1,284	British National Formulary (2009)
Grade 3 /4 neutropenia	£92.80	ERG Addendum (2007)
Grade 3 /4 febrile neutropenia	£2,286	ERG Addendum (2007)
Grade 3 /4 fatigue	£39	Eli Lilly (2009)
Grade 3 /4 nausea and vomiting	£701	Eli Lilly (2009)
Grade 3 /4 diarrhoea	£867	Eli Lilly (2009)
Grade 3 /4 rash	£117	Roche (2006)
Grade 3 /4 anaemia	£615	Eli Lilly (2009)
NHS patient transport service (per journey)	£28	Reference costs (2007/08)
BSC (per cycle)	£600	Clegg (2002)
2nd line therapy followed by BSC (per	£1,022	ERG report (2006)

Model variable	Value	Source
cycle)		

Table 92: Resource use and costs specified in TA258: Erlotinib for the first-line treatment of locally advanced or metastatic EGFR-TK mutation-positive NSCLC

Drug costs			
Cost	**Value**	**95% CI**	**Source**
Pharmacy costs per pack of erlotinib/gefitinib dispensed	£13	£6.63 to £19.37	MS 6.5.5.3
Erlotinib drug costs	30 x 150 mg = £1,631.53 30 x 100 mg = £1,324.14 30 x 25 mg – 50 mg = £378.33	Not applicable	British National Formulary 62
Gefitinib PAS fixed cost payment	£12,200	Not applicable	MS 6.5.5.1.2
Gefitinib PAS administration cost	£70 set up cost per pt £34 per month (ongoing)	Gamma distribution applied under assumption standard error was a quarter of base-case value	MS 6.5.5.3
Care Costs			
Health states	**Included elements**	**Value**	**Source**
Monthly PFS BSC cost (including monitoring)	Supportive care plus CT assessment of response every three months	£181.46	MS 6.5.6
Monthly PD BSC cost	Supportive care plus CT assessment of response every three months whilst on 2nd line treatment (estimate based upon SATURN RCT in NICE TA227)	£160.06	MS 6.5.6
Terminal phase BSC	Supportive care	£2,588.25	MS 6.5.6
AEs			
Rash		£116	Roche 2006 cited in Brown et al 2009 (NICE TA192 ERG report)
Diarrhoea		£867	Eli Lilly 2009 cited in Brown et al 2009 (NICE TA192 ERG report)

Table 93: Resource use and costs specified in TA310: afatinib for treating EGFRmutation-positive locally advanced or metastatic NSCLC

Drug acquisition costs			
	Unit	**British National Formulary cost per pack**	**Cost per month**
Erlotinib	30 x 150 mg	£1,631.53	£1,654.19
Gefitinib	30 x 250 mg	£2,167.71	£12,200 on receipt of third pack
Afatinib	28 x 40 mg	£2,167.71	£2,197.82
Docetaxel	250 mg/ml, 4ml vial	£534.75	£1,549.25
Drug administration costs			
	Afatinib, erlotinib and gefitinib	**Docetaxel**	**Reference**
Introductory cost	£163		DH 2013
Monthly administration cost (SB14Z)		£302.41	DH 2013
Gefitinib PAS			
PAS set up cost	£70		Roche, 2011
PAS administration cost	£34		Roche, 2011
Health state costs			

	Included elements	Value/month	Source
First-line PFS	Outpatient visits (CT scan, MRI scan, surgical procedure, ultrasound, x-ray, radiotherapy, GP, specialist, nurse, occupational therapist, physiotherapy); Unscheduled hospitalisations (ICU visit, emergency room visit)	£220	LUX-Lung 3
Second-line PFS	Not reported	£362	Lewis et al 2010
Third-line/progressive disease	Outpatient visits (GP, specialist, nurse, occupations therapist, physiotherapist); Outpatient interventions (blood transfusion, CT scan, infusion, MRI scan, physical therapy, respiratory therapy, surgical procedure, ultrasound, x-ray, radiotherapy); Unscheduled hospital stay (ICU visit, emergency room)	£418	LUX-Lung 1
AE costs			
	Cost	**Source**	
Diarrhoea	Marked as commercial in confidence	Resource use data extracted from LUX-Lung 3	
Rash/acne	Marked as commercial in confidence	Resource use data extracted from LUX-Lung 3	
Fatigue	Marked as commercial in confidence	Resource use data extracted from LUX-Lung 1	
Anaemia	Marked as commercial in confidence	Eli Lilly and Co 2008	
Neutropenia	Marked as commercial in confidence	Eli Lilly and Co 2008	

Table 94: Resource use and costs specified in TA296: crizotinib for previously treated NSCLC associated with an anaplastic lymphoma kinase fusion gene

Health State	Description	Resources required	Frequency	Unit cost	Reference
PF	Patients are receiving BSC (no active treatment) and tumour has not yet progressed	Outpatient visit	0.75 visits per month	£123	Expert panel (resource use); Schedule of Reference Costs; NHS Trusts and PCTs combined Outpatient Attendances Data - 370 medical oncology (unit costs)
		GP visit	10% of patients	£36	Expert panel (resource use); PSSRU Per clinic consultation lasting 17.2 minutes without qualification costs (unit costs)
		Cancer nurse	20% of patients 1 per month	£57	Expert panel (resource use); Schedule of Reference Costs; nurse cancer relate adult face to face CN201AF (unit costs)
		Complete blood count	0.75 per month	£3.36	Expert panel (resource use); National Schedule of Reference Costs Year : 2010-11 - NHS Trusts and PCTs combined Direct Access: Pathology Services; DAP823 (unit costs)
		Biochemistry	0.75 per month	£1.26	Expert panel (resource use); National Schedule of Reference Costs Year : 2010-11 - NHS Trusts and PCTs combined Direct

Health State	Description	Resources required	Frequency	Unit cost	Reference
					Access: Pathology Services; DAP841 (unit costs)
		CT scan	30% patients 0.75 per month	£160	Expert panel (resource use); National Schedule of Reference Costs Year : 2010-11 - NHS Trusts and PCTs combined Direct Access: Pathology Services; RA13Z (unit costs)
		Chest X-ray	0.75 per month	£129	Expert panel (resource use); National Schedule of Reference Costs Year : 2010-11 - Category 2 investigation with category 1 treatment VB08Z (unit costs)
Total cost per month, PF					**£241.44**
Progressed disease	Patients have experienced disease progression and are no longer receiving active therapy	Oncology visit	1 visit	Medical oncologist: £123	Expert panel (resource use); Schedule of Reference Costs; NHS Trusts and PCTs combined Outpatient Attendances Data - 370 medical oncology (unit costs)
		Cancer nurse	10% patients (1 visit)	£57	Expert panel (resource use); Schedule of Reference Costs; nurse cancer relate adult face to face CN201AF (unit costs)
		GP visit	28% patients (1 visit)	£36	Expert panel (resource use); PSSRU Per clinic consultation lasting 17.2 minutes without qualification costs (unit costs)
		Complete blood count	All patients, 1 per month	£3.36	Expert panel (resource use); National Schedule of Reference Costs Year : 2010-11 - NHS Trusts and PCTs combined Direct Access: Pathology Services; DAP823 (unit costs)
		Biochemistry	All patients, 1 per month	£1.26	Expert panel (resource use); National Schedule of Reference Costs Year : 2010-11 - NHS Trusts and PCTs combined Direct Access: Pathology Services; DAP841 (unit costs)
		CT scan	5% of patients, 0.75 per month	£151	Expert panel (resource use); National Schedule of Reference Costs Year : 2010-11 - NHS Trusts and PCTs combined Direct Access: Pathology Services; RA13Z (unit costs)
		X-ray	30% of patients, 0.75 per month	£129	Expert panel (resource use); National Schedule of Reference Costs Year : 2010-11 - Category 2 investigation with category 1 treatment (unit costs) VB08Z
Total cost per month, Progressed Disease					**£178.09**
Death		Palliative care	Cost applied only once	£3,923	Coyle et al (1999)
Total cost, Death					**£3,923**

Table 95: Resource use and costs specified in TA 181: Pemetrexed for the first-line treatment of NSCLC

Chemotherapy unit costs (British National Formulary 55, 2008)			
	Unit cost per vial	**Dose**	**Cost per dose**
Chemotherapy			
Pemetrexed (100mg vial)	£160.00		
Pemetrexed (500mg vial)	£800.00	500mg/m^2	£1,440.00
Gemcitabine (200mh vial)	£32.55	1250mh/m^2	£390.62
Gemcitanibe (1000mg vial)	£162.76		
Docetaxel (20mg vial)	£162.75	75mg/m^2	£1,023.00
Docetaxel (80mg vial)	£534.75		
Platinum			

Chemotherapy unit costs (British National Formulary 55, 2008)			
Cisplatin (50mg vial)	£25.37	75mg/m^2	£75.59
Cisplatin (100mg vial)	50.22		
Carboplatin (50mg vial)	£22.04	AUC=5 (500mg per cycle)	£190.89
Carboplatin (150mg vial)	£56.29		
Carboplatin (450mg vial)	£168.85		
Carboplatin (600mg vial)	£260.00		
	Mean cost per patient per cycle	**Mean number of cycles per patient**	**Mean total cost per patient**
Pem/cis	£1,440 + £75.59	3.80	£5,759.24
Gem/cis	(£390.62 x 2) + £75.59	3.81	£3,264.52
Gem/carbo	(£390.62 x 2) + £190.89	3.75	£3,645.49
Doc/cis	£1,023 + £75.59	3.79	£4,163.66
Concomitant therapy (British National Formulary 55, 2008)			
Premedication			**Unit cost**
Dexamethasone			£2.39
Folic Acid			£1.65
Vitamin B_{12}			£2.46
Piriton			£1.62
Paracetamol			£1.59
Pharmaceutical products			
Lomotil			£1.63
Domperidone			£2.35
Administration costs (DH, 2008)			
HRG code	**HRG label**	**Unit cost**	
		Outpatients	**Inpatients**
SB12Z	Deliver simple parenteral chemotherapy at first attendance	£170	£309
SB13Z	Deliver more complex parenteral chemotherapy at first attendance	£104	£298
SB14Z	Deliver complex chemotherapy including prolonged infusional treatment at first attendance	£179	£430
SB15Z	Deliver subsequent elements of a chemotherapy cycle	£189	£255
AE hospital resource utilisation (Duran et al, 2008)			
Adverse event			Unit cost
Neutropenia			£330.93
Nausea and vomiting			£700.79
Fatigue			£38.90
Diarrhoea			£867.12
Anaemia			£615.04
Thrombocytopenia			£314.69
BSC and terminal care costs			
Per cancer death, applied to every patient in the last three months of life			£2,686

7.5.4 If clinical experts assessed the applicability of values available or estimated any values, please provide the following details[8]:

- the criteria for selecting the experts
- the number of experts approached
- the number of experts who participated
- declaration of potential conflict(s) of interest from each expert or medical specialist whose opinion was sought
- the background information provided and its consistency with the totality of the evidence provided in the submission
- the method used to collect the opinions
- the medium used to collect opinions (for example, was information gathered by direct interview, telephone interview or self-administered questionnaire?)
- the questions asked
- whether iteration was used in the collation of opinions and if so, how it was used (for example, the Delphi technique).

See section 7.3.5.

Intervention and comparators' costs

7.5.5 Please summarise the cost of each treatment in the following table. Cross-reference to other sections of the submission; for example, drugs costs should be cross-referenced to sections 1.10 and 1.11. Provide a rationale for the choice of values used in the cost-effectiveness model discussed in section 7.2.2.

Treatment Cost

The costing of drug treatments was based on the following dosing schedules:

- Nintedanib: 200 mg twice daily
- Docetaxel: 75 mg/m^2, once every three weeks
- Erlotinib: 150 mg daily
- Platinum doublet therapy (relevant for 3rd line, Carboplatin / Vinorelbine):

[8] Adapted from Pharmaceutical Benefits Advisory Committee (2008) Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee (Version 4.3). Canberra: Pharmaceutical Benefits Advisory Committee.

- Carboplatin: 750 mg every 3 weeks, assuming AUC=5
- Vinorelbine: 30 mg/m^2 weekly

The costing of drug treatment was based on the dosing schedules shown in Table 96. The acquisition cost used to calculate the cycle costs for nintedanib plus docetaxel can be found in section 1.10. Unit costs were based on the British National Formulary October 2013 (Accessed 17 October 2013). BSC was assumed to have zero cost in the base-case.

Table 96: Drug Costs

Drug	Units per Administration	Price per Unit	Route	Administrations per Cycle	Administration Cost	Costs per Cycle*
Nintedanib	400 mg	£0.18	Oral	21	-	£1,354
Docetaxel in combination with nintedanib	75 mg/m^2	£5.68	IV	1	£155	£196
Docetaxel	75 mg/m^2	£5.68	IV	1	£155	£196
Erlotinib	150 mg	£0.36	Oral	21	-	£1,051
Carboplatin	750 mg	£0.33	IV	1	£155	£250
Vinorelbine	30 mg/m^2	£2.78	IV	3		£465

*Taken into account dose intensity as detailed in Table 97.

IV = Intravenous

Dose Intensity and Treatment Duration

Adjustments in drug costs due to change in dose intensity and treatment discontinuation as observed in the LUME-Lung 1 trial were included in the model. This ensured that model outcomes on drug costs reflected the actual drug exposure/dose intensity representing the efficacy and safety data being employed in the model. Note that there was a discrepancy in the sets of population analyses between the drug exposure and dose intensity data versus the efficacy data. Drug exposure and dose intensity data obtained from the LUME-Lung 1 trial were based on the 'treated' set (i.e., those who received at least one dose of study medication [chemotherapy and/or nintedanib/placebo]), while OS and PFS data were based on the 'randomised' set (i.e., all randomised patients, regardless of whether or not they have received treatment). Thus, the model slightly overestimated drug costs as some patients were never started on the treatment. However, this should not have a significant impact, as the population size in each data set was only slightly different (i.e., 'treated' versus 'randomised' set; nintedanib arm 320 versus 322; placebo arm 333 versus 336).

The average dose intensity for each treatment (Table 97) was multiplied to drug costs per cycle (Table 96

Table 96) to reflect the actual dosage used. Regarding treatment discontinuation, the average treatment duration for nintedanib and the average number of docetaxel courses as reported in LUME-Lung 1 trial were used to estimate treatment discontinuation risk per model cycle. The proportion of patients determined to have treatment discontinuations in each cycle no longer accrued drug costs. Note that changes in dose intensity or treatment discontinuation inputs only affected outcomes on drug costs but not the clinical outcomes (i.e., OS, PFS, AEs).

Dose intensity with carboplatin plus vinorelbine was taken to be 100%.

Table 97: Dose Intensity

Treatment/Patient population	Mean	SE*	Sources
Nintedanib			
Adenocarcinoma	91.20%	0.84%	CSR(5)
Docetaxel in combination with nintedanib			
Adenocarcinoma	98.10%	0.25%	CSR(5)
Docetaxel monotherapy			
Adenocarcinoma	98.70%	0.20%	CSR(5)
Erlotinib	92%		NICE TA 162(137)

CSR = Clinical study report; SE = Standard error

* SEs are calculated based on standard deviation and N (number of patients) data of LUME-Lung 1

Health-state costs

7.5.6 Please summarise, if appropriate, the costs included in each health state. Cross-reference to other sections of the submission for the resource costs. Provide a rationale for the choice of values used in the cost-effectiveness model. The health states should refer to the states in section 7.2.4.

There is little published literature that explores the detailed resource use resulting from NSCLC or other metastatic cancer treatment. To estimate the treatment patterns in NSCLC, a resource use questionnaire was constructed, which formed the basis of the interview with an oncologist who specialised in the treatment of patients with lung cancer exclusively, and with experience in NICE HTAs: Dr [redacted] (UK). The questionnaire included questions on all relevant details for the management of patients with NSCLC, including inpatient and outpatient treatment of patients with NSCLC in different health states (stable, during second-line treatment, at the time of progression and after progression [on active treatment or on BSC]). The answers to these questions were tabulated and reviewed with the oncologist. The full results of the questionnaire can be seen in Appendix 10.18. Then, each resource was assigned a unit cost to calculate the total cost of each section per comparator. The current base-case model included responses from Dr [redacted] as he was able to provide a detailed breakdown of the resource use of patients in each state within the model.

Monitoring Costs

The following questions relating to resource use were asked separately in the different health states:

- Routine follow-up: Type and frequency of physician visit, lab tests, radiological scans
- Treatment at time of progression: Hospitalisations, physician visits, lab tests, radiological scans, procedures used
- Resources used during BSC/palliative care: Initial tests, procedures, hospitalisations, physician visits, lab tests, radiological scans, procedures

Dr [redacted] noted that resource use for monitoring during the PF state depended on the type of treatment administered, so the resource use of pre-progression health state was separated by treatments in the model.

Unit costs obtained from NHS(121) were applied to the resource use responses obtained from the EE interview; monitoring costs per model cycle were derived, as shown in Table 98. The resource use data obtained from the EE interview are displayed in Table 99 to Table 105.

Furthermore, an additional option for monitoring costs was allowed in the model for sensitivity analysis, including the monitoring costs based on those in the afatinib NICE submission(43) (Table 98).

In the afatinib submission, the disease management costs assigned to the second-line treatment PF health state were extracted from a literature review. The cost is derived from a UK based study of patients receiving 2nd line treatment for NSCLC(135).

For the 'monitoring in progressed health state on BSC' the data from the afatinib submission regarding third line progressive disease health state specific cost is from phase 3 RCT data. The disease management resource use, weighted costs and total cost are reported in Table 109 of the afatinib submission to NICE(43), and the resource use was taken from the H-SAP from LUX-Lung 1 (138).

Table 98: Monitoring Costs per Cycle

	Base-case	Sensitivity Analysis
	Costs per cycle based on EE interview	**Costs per cycle based on afatinib submission**
Monitoring in stable health state on active treatment – Nintedanib + docetaxel	£188 (details inTable 99)	£250
Monitoring in stable health state on active treatment – Docetaxel	£205 (details in Table 100)	£250
Monitoring in stable health state on active treatment – Erlotinib	£101 (details in Table 101)	£250
Monitoring in stable health state on BSC	£461 (details in Table 102)	£250
Monitoring in progressed health state on active treatment	£98 (details in Table 103)	£288
Monitoring in progressed health state on BSC	£407 (details in Table 104)	£288
Monitoring at time of progression –- One off cost	£126 (details in Table 105)	£0

BSC = Best supportive care; EE = External expert.

Table 99: Detailed Resource Use based on EE Interview in Stable Health State – Nintedanib plus Docetaxel

	Nintedanib + Docetaxel					
	Proportion of Patients	Frequency per Model Cycle		Unit Cost (£)	Costs per Model Cycle	Source
					Mean (£)	
Monitoring						
Healthcare professional visit						
GP visit				£63.0	£0.0	
Palliative care				£70.0	£0.0	
Clinical assessment of patients						
Oncologist visit	100.0%	1.00		£139.0	£139.0	EE
Radiotherapy (brain)				£121.0	£0.0	
Radiotherapy (bone)				£121.0	£0.0	
Laboratory tests						
CT scan (thorax or abdominal)	100.0%	0.28		£90.0	£24.8	EE
CT scan (brain)				£90.0	£0.0	EE
Full blood test	100.0%	1.00		£3.0	£3.0	EE
Electrolytes	100.0%	1.00		£4.0	£4.0	EE
Liver function	100.0%	1.00		£7.0	£7.0	EE
Renal function	100.0%	1.00		£10.0	£10.0	EE
99Tc bone scintigraphy scan				£191.0	£0.0	EE
X-ray		0.50		£28.0	£0.0	

CT = Computed tomography; GP = General practitioner; EE = External expert;

Table 100: Detailed Resource Use based on EE Interview in Stable Health State –Docetaxel

	Docetaxel					
	Proportion of Patients	Frequency per Model Cycle		Unit cost (£)	Costs per Model Cycle	Source
					Mean (£)	
Monitoring						
Healthcare professional visit (routine GP)						
GP visit	100.0%	0.28		£63.0	£17.4	EE
Palliative care				£70.0	£0.0	
Clinical assessment (specialist – Oncologist)						
Oncologist visit	100.0%	1.00		£139.0	£139.0	EE
Radiotherapy				£121.0	£0.0	
Laboratory tests (in hospital)						
CT scan (thorax or abdominal)	100.0%	0.28		£90.0	£24.8	EE
CT scan (brain)				£90.0	£0.0	EE
Full blood test	100.0%	1.00		£3.0	£3.0	EE
Electrolytes	100.0%	1.00		£4.0	£4.0	EE
Liver function	100.0%	1.00		£7.0	£7.0	EE
Renal function	100.0%	1.00		£10.0	£10.0	EE
Hospitalisation						
Inpatient care – Per stay	0.0%	0.00		£2,001.0	£0.0	

CT = Computed tomography; GP = General practitioner; EE = External expert

Table 101: Detailed Resource Use based on EE Interview in Stable Health State – Erlotinib

	Erlotinib						
	Proportion of Patients	Frequency per Model Cycle		Unit Cost (£)	Costs per Model Cycle		Source
					Mean (£)		
Monitoring							
Healthcare professional visit (routine GP)							
GP visit				£63.0	£0.0		EE
Palliative care				£70.0	£0.0		
Clinical assessment (specialist – Oncologist)							
Oncologist visit	100.0%	0.46		£139.0	£63.9		EE
Radiotherapy				£121.0	£0.0		
Laboratory tests (in hospital)							
CT scan (thorax or abdominal)	100.0%	0.28		£90.0	£24.8		EE
CT scan (brain)				£90.0	£0.0		EE
Full blood test	100.0%	1.00		£3.0	£3.0		EE

	Erlotinib						
	Proportion of Patients	Frequency per Model Cycle		Unit Cost (£)	Costs per Model Cycle		Source
					Mean (£)		
Electrolytes	100.0%	0.46		£4.0	£1.8		EE
Liver function	100.0%	0.46		£7.0	£3.2		EE
Renal function	100.0%	0.46		£10.0	£4.6		EE
Hospitalisation							
Inpatient care – Per stay	0.0%	0.00		£2,001.0	£0.0		

CT = Computed tomography; GP = General practitioner; EE = External expert

Table 102: Detailed Resource Use based on EE Interview in Stable Health State – BSC

	BSC						
	Proportion of Patients	Frequency per Model Cycle		Unit cost (£)	Costs per Model Cycle		Source
					Mean (£)		
Monitoring							
Health care professional visit (routine GP)							
GP visit	100.0%	0.06		£63.0	£3.6		EE
Palliative care	100.0%	3.00		£70.0	£210.0		EE
Clinical assessment (specialist – Oncologist)							
Oncologist visit							
Radiotherapy (brain)	20.0%	1.00		£121.0	£24.2		EE
Radiotherapy (bone)	20.0%	1.00		£121.0	£24.2		EE
Laboratory tests (in hospital)							
CT scan (thorax or abdominal)				£90.0	£0.0		
CT scan (brain)				£90.0	£0.0		
Full blood test				£3.0	£0.0		
Electrolytes				£4.0	£0.0		
Liver function				£7.0	£0.0		
Renal function				£10.0	£0.0		EE
99Tc bone scintigraphy scan	100.0%	1.00		£191.0	£191.0		
Chest X-ray	100.0%	0.28		£28.0	£7.7		EE

BSC = Best supportive care; CT = Computed tomography; GP = General practitioner; EE = External expert

Table 103: Detailed Resource Use based on EE interview in Progressed Health State – Active Treatment

	All Active treatments						
	Proportion of Patients	Frequency per Model Cycle		Unit Cost (£)	Costs per Model Cycle Mean (£)		Source
Monitoring							
Healthcare professional visit							
GP visit	0.0%	0.0%		£63.0	£0.0		EE
Palliative care	0.0%	0.0%		£70.0	£0.0		EE
Clinical assessment of patients							
Oncologist visit	100.0%	46.0%		£139.0	£63.9		EE
Radiotherapy	0.0%	0.0%		£121.0	£0.0		EE
Laboratory tests							
CT scan (thorax or abdominal)	100.0%	27.6%		£90.0	£24.8		EE
CT scan (brain)	0.0%	0.0%		£90.0	£0.0		EE
Full blood test	100.0%	100.0%		£3.0	£3.0		EE
Electrolytes	100.0%	46.0%		£4.0	£1.8		EE
Liver function	100.0%	46.0%		£7.0	£3.2		EE
Renal function	100.0%	46.0%		£10.0	£4.6		EE
Blood transfusion				£167.0	£0.0		EE
Oxygen				£13.4	£0.0		EE
99Tc bone scintigraphy scan				£191.0	£0.0		EE
X-ray				£28.0	£0.0		EE
Hospitalisation							
Inpatient care – Per stay	0.0%	0.0%		£2,001.0			

CT = Computed tomography; GP = General practitioner; EE = External expert

Table 104: Detailed Resource Use based on EE interview in Progressed Health State – BSC

	BSC						
	Proportion of Patients	Frequency per Model Cycle		Unit Cost (£)	Costs per Model Cycle Mean (£)		Source
Monitoring							
Healthcare professional visit							
GP visit				£63.0	£0.0		EE
Palliative care	100.0%	3.00		£70.0	£210.0		EE
Clinical assessment of patients							
Oncologist visit				£139.0	£0.0		EE
Radiotherapy	50.0%	1.00		£121.0	£60.5		EE

	BSC					
	Proportion of Patients	Frequency per Model Cycle		Unit Cost (£)	Costs per Model Cycle Mean (£)	Source
Laboratory tests						
CT scan (thorax or abdominal)				£90.0	£0.0	EE
CT scan (brain)				£90.0	£0.0	EE
Full blood test				£3.0	£0.0	EE
Electrolytes				£4.0	£0.0	EE
Liver function				£7.0	£0.0	EE
Renal function				£10.0	£0.0	EE
Blood transfusion	50.0%	1.00		£167.0	£83.5	EE
Oxygen	50.0%	1.00		£13.4	£6.7	EE
99Tc bone scintigraphy scan	20.0%	1.00		£191.0	£38.2	EE
X-ray	100.0%	0.28		£28.0	£7.7	EE
Hospitalisation						
Inpatient care – Per stay				£2,001.0	£0.0	EE

CT = Computed tomography; GP = General practitioner; EE = External expert;

Table 105: Detailed Resource Use based on EE Interview in at Time of Progression – All Active Treatments

	All Active Treatments					
	Proportion of Patients	Number of Procedures		Unit Cost (£)	Costs per Model Cycle Mean (£)	Source
Monitoring						
Healthcare professional visit						
GP visit				£63.0	£0.0	
Palliative care				£70.0	£0.0	
Clinical assessment of patients						
Oncologist visit				£139.0	£0.0	
Radiotherapy				£121.0	£0.0	
Laboratory tests						
CT scan (thorax or abdominal)	100.0%	1.00		£90.0	£90.0	EE
CT scan (brain)	40.0%	1.00		£90.0	£36.0	EE
Full blood test				£3.0	£0.0	
Hospitalisation						
Inpatient care – Per stay				£2,001.0	£0.0	

CT = Computed tomography; GP = General practitioner; EE = External expert;

Unit costs of visit procedures and laboratory tests were derived mainly from National Schedule of Reference Costs 2012–2013(121) and some visit costs were from the Personal Social Services Research Unit (PSSRU). The unit costs used are in Table 106.

Table 106: Unit Costs

	Unit Costs £	Source
Healthcare Professional Visit		
Routine physician consultation/GP (monitoring)	£63.0	PSSRU. Unit Costs of Health & Social Care 2012. Compiled by L. Curtis
Oncologist specialist visit (specialised monitoring)	£139.0	National Schedule of Reference Costs 2012–2013 for NHS trusts and NHS foundation trusts (306 Hepatology – Consultant-led Outpatient Attendances Face to-face Follow-up Visit)
Hepatologist specialist visit	£200.0	National Schedule of Reference Costs 2012–2013 for NHS trusts and NHS foundation trusts (370 Medical Oncology – Consultant-led Outpatient Attendances Face to-face Follow-up Visit)
Gastroenterologist specialist visit	£123.0	National Schedule of Reference Costs 2012–2013 for NHS trusts and NHS foundation trusts (370 Medical Oncology – Consultant-led Outpatient Attendances Face to-face Follow-up Visit)
Palliative care nurse	£70.0	National Schedule of Reference Costs 2012–2013 for NHS trusts and NHS foundation trusts – Community Health Services – Nursing
Radiation oncologist	£121.0	National Schedule of Reference Costs v Year 2012–2013 – NHS trusts and NHS foundation trusts. Consultant-led Outpatient Attendances. Non-admitted face to face.
Surgeon visit	£119.0	National Schedule of Reference Costs – Year 2012–2013 – NHS trusts and NHS foundation trusts. Consultant-led Outpatient Attendances. Non-admitted face to face.
Nurse visit	£38.0	National Schedule of Reference Costs – Year 2012–2013 – NHS trusts and NHS foundation trusts. Community Health Services – Nursing.
Nurse home visit	£70.0	PSSRU. Unit Costs of Health & Social Care 2013. Compiled by Lesley Curtis.
Physician home visit	£292.0	PSSRU. Unit Costs of Health & Social Care 2013. Compiled by Lesley Curtis.
Other visits	£110.5	Average of GP, oncologist, radiologist, and surgeon visits
Procedures		
Radiotherapy – Inpatient	£195.0	National Schedule of Reference Costs 2012–2013 for NHS trusts and NHS foundation trusts (370 Medical Oncology – Consultant-led Outpatient Attendances Face-to-face Follow-up Visit)
Radiotherapy – Outpatient	£121.0	National Schedule of Reference Costs 2012–2013 for NHS trusts and NHS foundation trusts (370 Medical Oncology – Consultant-led Outpatient Attendances Face-to-face Follow-up Visit)
Blood transfusion – Inpatient	£1,121.0	National Schedule of Reference Costs – Year 2012–2013 – NHS trusts and NHS foundation trusts – Elective Inpatients
Blood transfusion – Outpatient	£167.0	National Schedule of Reference Costs – Year 2012–2013 – NHS trusts and NHS foundation trusts – Procedures in Outpatients
Oxygen	£13.4	http://www.ppa.org.uk/edt/December_2012/mindex.htm
Oxygen assessment	£171.0	National Schedule of Reference Costs – Year 2012–2013 – NHS trusts and NHS foundation trusts – Procedures in Outpatients
CT scan	£90.0	
Chest X-ray	£28.0	National Schedule of Reference Costs – Year 2012–2013 – NHS trusts and NHS foundation trusts – Procedures in Outpatients
MRI	£204.0	National Schedule of Reference Costs – Year 2012–2013 – NHS trusts and NHS foundation trusts – Diagnostic Imaging – Outpatients
PET	£282.0	National Schedule of Reference Costs – Year 2012–2013 – NHS trusts and NHS foundation trusts – Diagnostic Imaging – Outpatients
Other imaging	£159.0	Assumption: average of other imaging

	Unit Costs £	Source
FBC	£3.0	National Schedule of Reference Costs – Year 2012–2013 – NHS trusts and NHS foundation trusts – Diagnostic Imaging. Weighted Average National Cost. (CT scan, one area, no contrast, 19 yrs and over.)
Electrolytes	£4.0	National Schedule of Reference Costs – Year 2012–2013 – NHS trusts and NHS foundation trusts – Diagnostic Imaging. Weighted Average National Cost.
Liver function	£7.0	National Schedule of Reference Costs – Year 2012–2013 – NHS trusts and NHS foundation trusts – Diagnostic Imaging. Weighted Average National Cost.
Renal function	£10.0	National Schedule of Reference Costs – Year 2012–2013 – NHS trusts and NHS foundation trusts – Diagnostic Imaging. Weighted Average National Cost.
Calcium	£1.0	National Schedule of Reference Costs – Year 2012–2013 – NHS trusts and NHS foundation trusts – Diagnostic Imaging. Weighted Average National Cost.
Colonoscopy	£309.5	National Schedule of Reference Costs – Year 2012–2013 – NHS trusts and NHS foundation trusts – Procedures in Outpatients. Average of with and without biopsy.
Stool cultures	£7.0	National Schedule of Reference Costs – Year 2012–2013 – NHS trusts and NHS foundation trusts – Directly Accessed Pathology Services
Ultrasound	£57.0	National Schedule of Reference Costs – Year 2012–2013 – NHS trusts and NHS foundation trusts – Diagnostic Imaging – Outpatients. Average of less and over 20 minutes.
99Tc bone scintigraphy scan	£191.0	National Schedule of Reference Costs – Year 2012–2013 – NHS trusts and NHS foundation trusts – Diagnostic Imaging – Outpatients
LFT	£7.0	National Schedule of Reference Costs – Year 2012–2013 – NHS trusts and NHS foundation trusts – Directly Accessed Pathology Services. Clinical Biochemistry and Haematology.
Chemistry panel	£8.0	National Schedule of Reference Costs – Year 2012–2013 – NHS trusts and NHS foundation trusts – Directly Accessed Pathology Services. Clinical Biochemistry and Haematology.
Coagulation test	£3.0	National Schedule of Reference Costs – Year 2012–2013 – NHS trusts and NHS foundation trusts – Directly Accessed Pathology Services. Clinical Biochemistry and Haematology.
U and E	£5.0	National Schedule of Reference Costs – Year 2012–2013 – NHS trusts and NHS foundation trusts – Directly Accessed Pathology Services. Clinical Biochemistry and Haematology.
Other lab tests	£5.2	Average of FBC/CBC, LFT, chemistry panel, coagulation test, U and E tests

CBC = Complete blood count; FBC = Full blood count; GP = General practitioner; LFT = Liver function test; MRI = Magnetic resonance imaging; NHS = National Health Service; PET = Positron emission tomography; PSSRU = Personal Social Services Research Unit; U and E = Urea and electrolytes; WBC = White blood cell

Adverse-event costs

7.5.7 Please summarise the costs for each adverse event listed in section 6.9 (Adverse events). These should include the costs of therapies identified in sections 2.7 and 2.8. Cross-reference to other sections of the submission for the resource costs. Provide a rationale for the choice of values used in the cost-effectiveness model discussed in section 7.2.2.

AE management costs were provided by a UK EE, Dr. [redacted], from Cardiff, as mentioned in Section 7.3.5. He provided data on inpatient and outpatient treatment of all listed grade 3 and 4 toxicities and

grade 2 diarrhoea. The following questions were asked separately for grade 3 and 4 toxicities and grade 2 diarrhoea:

- What percentage of patients was hospitalised?
- What percentage of patients required an outpatient visit? How many outpatient visits were required? What specialty (urologist, cardiologist, oncologist)?
- What procedures were used? What percentage of patients underwent the procedure?
- What medications were used? What percentage of patients required this medication? What was the dosage used and what was the length of treatment?

Total cost of treatment for each toxicity event was calculated, considering the percentage of patients requiring the resource and published unit costs. The assumption was made that patients who were hospitalised would incur the cost of the hospitalisation; however, additional outpatient costs (physician visits, procedure, and medication costs) might be necessary before or after the hospitalisation. These additional costs were also incorporated in the calculations.

Costs for inpatient hospitalisations were taken from the Health Care National Schedule of Reference Costs - Year 2012-13(121). Outpatient costs were taken from the same source or from the PSSRU(124). The cost of each AE is summarised in Table 107. Details of the resource use are described in more detail in Appendix 10.17.

Table 107. Cost of AEs

	Cost of AEs
ALT increased	£587
Anaemia	£978
AST increased	£336
Diarrhoea – Grade 1 and 2	£250
Diarrhoea – Grade 3 and 4	£1,796
Fatigue	£370
Febrile neutropenia	£2,012
Infection	£2,181
Nausea and vomiting	£1,919
Neutropenia	£346
Rash	£639
Thrombocytopenia	£422
WBC count decreased	£423

AE = Adverse event; ALT = Alanine aminotransferase; AST = Aspartate aminotransferase; WBC = White blood cell

Source: Calculation based on EE input.

Miscellaneous costs

7.5.8 Please describe any additional costs that have not been covered anywhere else (for example, PSS costs). If none, please state.

None

7.6 Sensitivity analysis

This section should be read in conjunction with NICE's 'Guide to the methods of technology appraisal', sections 5.1.11, 5.8, and 5.9.4 to 5.9.12.

Sensitivity analysis should be used to explore uncertainty around the structural assumptions used in the analysis. Analysis of a representative range of plausible scenarios should be presented and each alternative analysis should present separate results.

The uncertainty around the appropriate selection of data sources should be dealt with through sensitivity analysis. This will include uncertainty about the choice of sources for parameter values. Such sources of uncertainty should be explored through sensitivity analyses, preferably using probabilistic methods of analysis.

All inputs used in the analysis will be estimated with a degree of imprecision. Probabilistic sensitivity analysis (PSA) is preferred for translating the imprecision in all input variables into a measure of decision uncertainty in the cost effectiveness of the options being compared.

For technologies whose final price/acquisition cost has not been confirmed, sensitivity analysis should be conducted over a plausible range of prices.

7.6.1 Has the uncertainty around structural assumptions been investigated? Provide details of how this was investigated, including a description of the alternative scenarios in the analysis.

Scenario analyses were conducted. The Scenarios evaluated were as follows:

- Discount rate:
    - Discount rates: 0% for both costs and health effects
    - Discount rates: 6% for both costs and health effects
    - Discount rates: 0% for costs and 6% for health effects
    - Discount rates: 6% for costs and 0% for health effects
- Efficacy measures:
    - PFS extrapolated with Weibull logistic distribution fitted separately for the two treatment arms

    - OS extrapolated with the help of a Weibull distribution fitted separately for the two treatment arms
    - Use of Kaplan-Meier curves instead of distribution until the end of the trial follow-up period (the distributions selected for the base-case were used after the follow-up period)
    - Use of Kaplan-Meier curves instead of distribution set to zero for PFS and extrapolated using SEER data for OS
    - Use of Kaplan-Meier curves instead of distribution set to zero for PFS and extrapolated using LUCADA data for OS
- Source of utility inputs
    - Utility extrapolation for PFS set to LOCF
    - Published literature(128)
- Source of resource use data
    - Afatinib submission (LUX-Lung trial data)
- Time horizon: Five years to 15 years
- Exploratory analysis with the indirect comparison
    - Erlotinib HR scenarios
- Maximum number of docetaxel cycles = 4 (EE suggestion)
- Cost effectiveness vs erlotinib at a range of discounts from its list price

7.6.2 Which variables were subject to deterministic sensitivity analysis? How were they varied and what was the rationale for this? If any parameters or variables listed in section 7.3.6 (Summary of selected values) were omitted from sensitivity analysis, please provide the rationale.

To identify key model parameters, a one-way deterministic sensitivity analyses (OWSA) was conducted using extreme values for all model parameters. Those extreme values corresponded to the reference case estimates ± 20%. Results of the deterministic one-way sensitivity analyses were plotted as a tornado diagram based on the impact of the variable on the incremental net benefit using £50,000/QALY as a threshold. The impact of the five most influential variables on the ICER was also examined. Table 108 shows the variables investigated in the OWSA.

Table 108: Upper and lower bounds of variables tested in the OWSA (Taken from Sheet 'Sensitivity Analysis Tornado' in the Excel model)

Variable	Cell Reference	Base value	Lower value	Upper value
Discount rate for costs/year	discCost	0.035	2.80%	4.20%
Discount rate for health/year	discHealth	0.035	2.80%	4.20%
Discount rate for costs/cycle	discCc	0.002	0.16%	0.24%
Discount rate for health/cycle	discHc	0.002	0.16%	0.24%
BSA	BSA	1.855	1.4843	2.2265
HRs PFS Nintedanib+docetaxel vs Erlotinib	HR_PFS_Comp3	0.700	0.3534	1.0466
HRs PFS Nintedanib+docetaxel vs Pemetrexed	HR_PFS_Comp4	0.840	0.5230	1.1570
HRs PFS Nintedanib+docetaxel vs Placeholder	HR_PFS_Comp5	1.000	1.0000	1.0000
HRs OS Nintedanib+docetaxel vs Erlotinib	HR_OS_Comp3	0.640	0.3044	0.9756
HRs OS Nintedanib+docetaxel vs Pemetrexed	HR_OS_Comp4	0.820	0.5124	1.1276
Discontinuation risk per cycle for nintedanib - Nintedanib+docetaxel	rDiscontinuation_nine_Comp1	0.125	0.1004	0.1493
Discontinuation risk per cycle for docetaxel - Nintedanib+docetaxel	rDiscontinuation_doce_Comp1	0.175	0.1510	0.1999
Discontinuation rate per cycle for Docetaxel	rDiscontinuation_Comp2	0.196	0.1617	0.2305
Discontinuation rate per cycle for Erlotinib	rDiscontinuation_Comp3	0.168	0.1296	0.2065
Drug and administration cost for nintedanib in Comp1	cDrugAdmin_nine_Comp1	[redacted]	[redacted]	[redacted]
Drug and administration cost for docetaxel in Comp1	cDrugAdmin_doxa_Comp1	[redacted]	[redacted]	[redacted]
Drug and administration cost for Docetaxel in comp 2	cDrugAdmin_Comp2	195.327	195.3265	195.3265
Drug and administration cost for Erlotinib	cDrugAdmin_Comp3	1,050.705	1050.7053	1050.7053
% patients switching to Erlotinib	SwitchP_Comp3	0.050	0.0450	0.0550
% patients switching to BSC	SwitchP_BSC	0.700	0.6300	0.7700
MM cost in PF for Nintedanib+docetaxel - AT	cMM_Stable_AT_Comp1	187.838	169.0537	206.6219
MM cost in PF for Nintedanib+docetaxel - BSC	cMM_Stable_BSC_Comp1	460.749	414.6737	506.8253
MM cost in PD for Nintedanib+docetaxel - AT	cMM_Progr_AT_Comp1	98.461	88.6149	108.3075
MM cost in PD for Nintedanib+docetaxel - BSC	cMM_Progr_BSC_Comp1	406.627	365.9638	447.2908
MM cost at time of progression for Nintedanib+docetaxel	cMM_timeofProgr_Comp1	126.000	113.3998	138.6002
MM cost in PF for Docetaxel - AT	cMM_Stable_AT_Comp2	205.224	184.7014	225.7470
MM cost in PF for Docetaxel - BSC	cMM_Stable_BSC_Comp2	460.749	414.6737	506.8253
MM cost in PD for Docetaxel - AT	cMM_Progr_AT_Comp2	98.461	88.6149	108.3075
MM cost in PD for Docetaxel - BSC	cMM_Progr_BSC_Comp2	406.627	365.9638	447.2908
MM cost at time of progression for Docetaxel	cMM_timeofProgr_Comp2	126.000	113.3998	138.6002
MM cost in PF for Erlotinib - AT	cMM_Stable_AT_Comp3	101.431	91.2879	111.5745
MM cost in PF for Erlotinib - BSC	cMM_Stable_BSC_Comp3	460.749	414.6737	506.8253
MM cost in PD for Erlotinib - AT	cMM_Progr_AT_Comp3	98.461	88.6149	108.3075
MM cost in PD for Erlotinib - BSC	cMM_Progr_BSC_Comp3	406.627	365.9638	447.2908
MM cost at time of progression for Erlotinib	cMM_timeofProgr_Comp3	126.000	113.3998	138.6002
Cycle cost for AE Nintedanib+docetaxel	cAE_Comp1	91.198	82.0784	100.3184
Cycle cost for AE Docetaxel	cAE_Comp2	92.094	82.8843	101.3034
Cycle cost for AE Erlotinib	cAE_Comp3	61.413	55.2712	67.5540
Utility of PF Nintedanib+docetaxel week 0	u_PF_w0	0.710	0.6904	0.7296
Utility of PF Nintedanib+docetaxel week 3	u_PF_w3	0.721	0.6994	0.7426
Utility of PF Nintedanib+docetaxel week 6	u_PF_w6	0.707	0.6835	0.7305
Utility of PF Nintedanib+docetaxel week 9	u_PF_w9	0.699	0.6735	0.7245
Utility of PF Nintedanib+docetaxel week 12	u_PF_w12	0.692	0.6685	0.7155

Utility of PF Nintedanib+docetaxel week 15	u_PF_w15	0.687	0.6615	0.7125
Utility PF Nintedanib+docetaxel week 18	u_PF_w18	0.682	0.6565	0.7075
Utility PF Nintedanib+docetaxel week 21	u_PF_w21	0.677	0.6476	0.7064
Utility PF Nintedanib+docetaxel week 24	u_PF_w24	0.671	0.6396	0.7024
Utility PF Nintedanib+docetaxel week 27	u_PF_w27	0.666	0.6307	0.7013
Utility PF Nintedanib+docetaxel week 30	u_PF_Comp1_w30	0.661	0.6218	0.7002
Utility PD Nintedanib+docetaxel	u_Progr_Comp1	0.638	0.6125	0.6635
Utility PD Docetaxel	u_Progr_Comp2	0.638	0.6125	0.6635
Utility PD Erlotinib	u_Progr_Comp3	0.638	0.6125	0.6635
Disutility AE Nintedanib+docetaxel	u_Aedisutility_Comp1	(0.049)	-0.0682	-0.0290
Disutility AE Docetaxel	u_Aedisutility_Comp2	(0.059)	-0.0790	-0.0398
Disutility AE Erlotinib	u_Aedisutility_Comp3	(0.110)	-0.1296	-0.0904

7.6.3 Was PSA undertaken? If not, why not? If it was, the distributions and their sources should be clearly stated if different from those in section 7.3.6, including the derivation and value of 'priors'. If any parameters or variables were omitted from sensitivity analysis, please provide the rationale for the omission(s).

To test the robustness of the results with respect to uncertainty in the model input parameters, a probabilistic sensitivity analysis was performed using a second-order Monte Carlo simulation. In this analysis, each parameter (costs and outcomes) was assigned a probability distribution, and cost-effectiveness results associated with simultaneously selecting random values from those distributions were generated. The uncertainty in the survival probabilities was represented through the uncertainty in the survival function parameters represented by the joint variance-covariance matrix of these parameters together, including the treatment coefficients(139). HRs are the ratio of hazard in two groups, and the standard statistical approach to estimating variance and CIs for such ratios is to assume normality on the log scale. Therefore, uncertainty in HRs for PFS and OS estimated from external sources (and not from patient-level data) was represented using log-normal distributions according to the means and 95%. Since utilities were also constricted on the interval zero to one, they were varied according to beta-distributions based on the means and standard deviations reported in the analysis of the LUME-Lung 1 trial. Healthcare resource use parameters (e.g., number of physician visits, length of stay) and costs were assumed to follow gamma distributions. Resource use counts followed discrete Poisson-distributions, whose conjugate distribution to describe the mean was the gamma distribution. The gamma distribution is also usually a good candidate to represent uncertainty in costs, because costs are constrained on the interval zero to positive infinity, and are often highly skewed. Since there was no information on the

variability of some of these parameters, their standard deviation was assumed to equal 10% of the mean.

Acquisition costs of treatment drugs were not varied as they were considered certain. The Monte Carlo simulation was run on a total of 5,000 iterations. Results of the probabilistic analysis were plotted on the cost-effectiveness plane (Figure 29) and were used to calculate cost-effectiveness acceptability curves.

Figure 29: Cost-effectiveness Plane

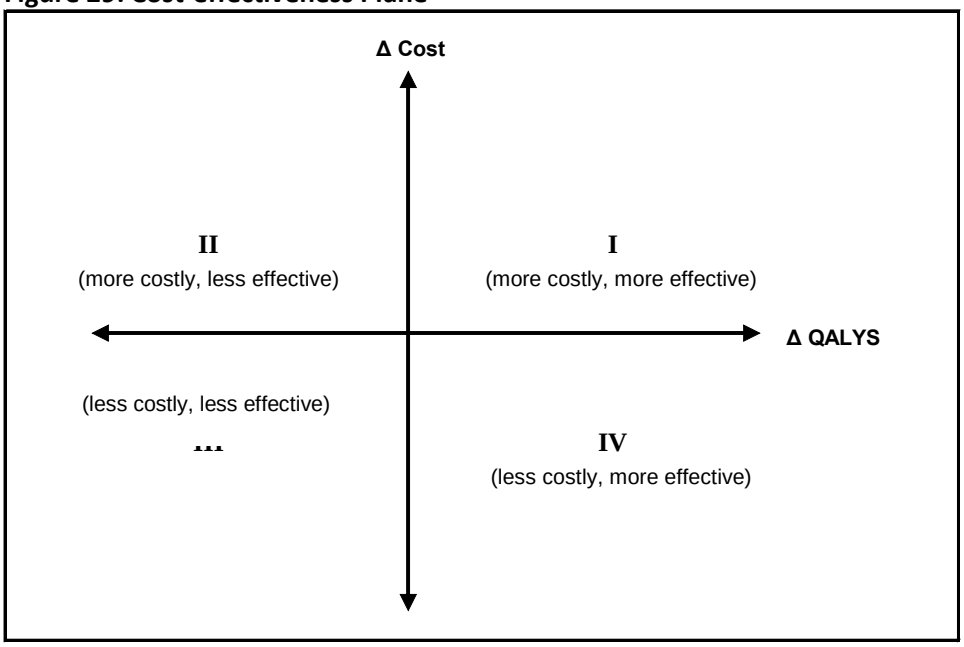


Table 109: Parameters of distribution used for extrapolation of PFS data – Separate LogNormal distribution (base-case)

Nintedanib + docetaxel	All adenocarcinoma patients	Intercept	1.30774	
		Scale	0.83857	
		VC matrix	0.002498	0.000181
			0.000181	0.001386
Docetaxel	All adenocarcinoma patients	Intercept	1.10629	
		Scale	0.90433	
		VC matrix	0.002768	0.000191
			0.000191	0.001564

Table 110: Parameters of distribution used for extrapolation of OS data – Separate LogLogistic distribution (base-case)

Nintedanib + docetaxel	All adenocarcinoma patients	Intercept	2.48427	
		Scale	0.6305	
		VC matrix	0.003859	0.000069
			0.000069	0.001061
Docetaxel	All adenocarcinoma patients	Intercept	2.32495	
		Scale	0.56332	
		VC matrix	0.003031	0.000057
			0.000057	0.00078

Table 111: Efficacy inputs for PSA sampled using LogNormal distribution

Variable	Expected mean	LN mean	SE	Lower CI	Upper CI
HR for PFS (nintedanib + docetaxel vs docetaxel) [Used only at scenario analysis]	0.77	-0.26	0.11	0.62	0.96
HR for PFS (nintedanib + docetaxel vs erlotinib)	0.70	-0.36	0.18	0.5	1.00
HR for OS (nintedanib + docetaxel vs docetaxel) [Used only at scenario analysis]	0.83	-0.19	0.09	0.70	0.99
HR for OS (nintedanib + docetaxel vs erlotinibl)	0.64	-0.45	0.17	0.46	0.90

Table 112: Discontinuation inputs for PSA sampled using Beta distribution

Variable	Expected mean	SE	Alpha	Beta
Discontinuation risk per cycle for nintedanib – nintedanib + docetaxel	0.125	0.01	87.39	612.56
Discontinuation risk per cycle for docetaxel – nintedanib + docetaxel	0.175	0.02	82.28	386.70
Discontinuation risk per cycle for docetaxel – docetaxel monotherapy	0.196	0.02	80.20	328.80
Discontinuation risk per cycle for erlotinib	0.168	0.02	83.02	410.93

Table 113: Drug and administration cost inputs - not included in PSA

Variable	Expected mean	SE	Alpha	Beta
Drug and administration cost per cycle for nintedanib – nintedanib + docetaxel	1353.52	0.00	-	-
Drug and administration cost per cycle for docetaxel – nintedanib + docetaxel	195.08	0.00	-	-
Drug and administration cost per cycle for docetaxel – docetaxel monotherapy	195.33	0.00	-	-
Drug and administration cost per cycle for erlotinib	1050.71	0.00	-	-

Table 114: Treatment switch due to progression inputs for PSA sampled using Dirichlet distribution

Variable	Expected mean	SE	Alpha	Beta
Proportion of patients switching to docetaxel	0.00	0.00	-	-
Proportion of patients switching to erlotinib	0.05	0.003	384.15	0.0001
Proportion of patients switching to pemetrexed	0.00	0.00	-	-
Proportion of patients switching to platinum doublet	0.25	0.013	384.15	0.0007
Proportion of patients switching to BSC	0.70	0.036	384.15	0.0018

Table 115: Monitoring cost inputs for PSA sampled using Gamma distribution

Variable		Expected mean	SE	Alpha	Beta
Monitoring costs for nintedanib + docetaxel patients	In stable health state on active treatment	187.84	9.58	384.15	0.49
	In stable health state on BSC	460.75	23.51	384.15	1.20
	In progressed health state on active treatment	98.46	5.02	384.15	0.26
	In progressed health state on BSC	406.63	20.75	384.15	1.06
	At time of progression	126.00	6.43	384.15	0.33
Monitoring costs for docetaxel patients	In stable health state on active treatment	205.22	10.47	384.15	0.53
	In stable health state on BSC	460.75	23.51	384.15	1.20
	In progressed health state on active treatment	98.46	5.02	384.15	0.26
	In progressed health state on BSC	406.63	20.75	384.15	1.06
	At time of progression	126.00	6.43	384.15	0.33
Monitoring costs for erlotinib patients	In stable health state on active treatment	101.43	5.18	384.15	0.26
	In stable health state on BSC	460.75	23.51	384.15	1.20
	In progressed health state on active treatment	98.46	5.02	384.15	0.26
	In progressed health state on BSC	406.63	20.75	384.15	1.06
	At time of progression	126.00	6.43	384.15	0.33

Table 116: End of life cost input - not included in PSA

Variable	Expected mean	SE	Alpha	Beta
Cycle cost of end of life	0.00	0.00	-	-

Table 117: Cost of AEs per treatment inputs for PSA sampled using Gamma distribution

Variable	Expected mean	SE	Alpha	Beta
Cycle cost of AEs for nintedanib + docetaxel	91.20	4.65	384.15	0.24
Cycle cost of AEs for docetaxel	92.09	4.70	384.15	0.24
Cycle cost of AEs for erlotinib	61.41	3.13	384.15	0.16

Table 118: Utility inputs for PSA sampled using Beta distribution

Variable		Expected mean	SE	Alpha	Beta
Utility in PF – nintedanib + docetaxel and docetaxel monotherapy	Week 0	0.710	0.01	1461.18	596.82
	Week 3	0.721	0.01	1197.92	463.55
	Week 6	0.707	0.01	1016.35	421.20
	Week 9	0.699	0.01	869.53	374.43
	Week 12	0.692	0.01	1023.54	455.57
	Week 15	0.687	0.01	873.43	397.94
	Week 18	0.682	0.01	874.52	407.77
	Week 21	0.677	0.02	657.28	313.59
	Week 24	0.671	0.02	577.96	283.38
	Week 27	0.666	0.02	456.58	228.98
	Week 30	0.661	0.02	369.63	189.57
Utility in PD – nintedanib + docetaxel		0.638	0.01	871.26	494.35
Utility in PD – docetaxel		0.638	0.01	871.26	494.35
Utility in PD – erlotinib		0.638	0.01	871.26	494.35

Table 119: Average disutility due to AEs per treatment inputs for PSA sampled using Beta distribution

Variable	Expected mean	SE	Alpha	Beta
Average disutility due to AEs – nintedanib + docetaxel	-0.049	0.01	24.82	-535.51
Average disutility due to AEs – docetaxel	-0.059	0.01	37.41	-667.41
Average disutility due to AEs – erlotinib	-0.110	0.01	134.40	-1356.33

7.7 Results

Provide details of the results of the analysis. In particular, results should include, but are not limited to, the following.

- Link between clinical- and cost-effectiveness results.
- Costs, QALYs and incremental cost per QALY.
- Disaggregated results such as LYG, costs associated with treatment, costs associated with adverse events, and costs associated with follow-up/subsequent treatment.
- A statement as to whether the results are based on a PSA.
- Cost-effectiveness acceptability curves, including a representation of the cost-effectiveness acceptability frontier.
- Scatter plots on cost-effectiveness quadrants.
- A tabulation of the mean results (costs, QALYs, ICERs), the probability that the treatment is cost effective at thresholds of £20,000–£30,000 per QALY gained and the error probability.

Clinical outcomes from the model

7.7.1 For the outcomes highlighted in the decision problem (see section 5), please provide the corresponding outcomes from the model and compare them with clinically important outcomes such as those reported in clinical trials. Discuss reasons for any differences between modelled and observed results (for example, adjustment for cross-over). Please use the following table format for each comparator with relevant outcomes included.

Table 120: PFS and OS from LUME Lung 1 and cost effectiveness model

	Nintedanib + Docetaxel		Docetaxel		Erlotinib	
Outcome	**LUME-Lung 1 result**	**Model result (Lognormal, Loglogistic)**	**LUME-Lung 1 result**	**Model result (Lognormal, Loglogistic)**	**MTC result**	**Model result (Weibull)**
PFS	Median 3.4 months (95% CI 2.9 - 3.9)	Median 3.71 months Mean 5.16 months	Median 2.7 months (95% CI 2.6 – 2.8)	Median 3.04 months Mean 4.44 months	HR 0.70 (95% CI 0.50 – 1.00)	HR 0.70 (95% CI 0.50 – 1.00)
OS	Median 12.6 months (95% CI 10.6 – 15.1)	Median 12.00 Mean 19.92 months	Median 10.3 months (95% CI 8.6 – 12.2)	Median 10.23 Mean 15.96 months	HR 0.64 (95% CI 0.46 – 0.90)	HR 0.64 (95% CI 0.46 – 0.90)

The results from the model fall within the CIs from the LUME-Lung trials, providing confidence that the model is able to reproduce the course of the disease to an acceptable degree of certainty.

Note that the Weibull extrapolation underestimates OS, however it provided the best statistical fit of the models which did not violate the proportional hazards assumption. This was needed as the PFS and OS curves for erlotinib are generated by hazards ratios from the indirect treatment comparison.

7.7.2 Please provide (if appropriate) the proportion of the cohort in the health state over time (Markov trace) for each state, supplying one for each comparator.

The Markov model contains three states: Progressive Disease (PD), PFS and Death. The Markov trace for each of these three states is shown for nintedanib in Figure 30, for docetaxel in Figure 31 and erlotinib in Figure 32. These are for the second-line indication using the base-case assumptions.

Figure 30: Markov trace for nintedanib plus docetaxel 2nd line treatment using base-case assumptions (PFS: Separate-LogNormal and OS: LogLogistic)

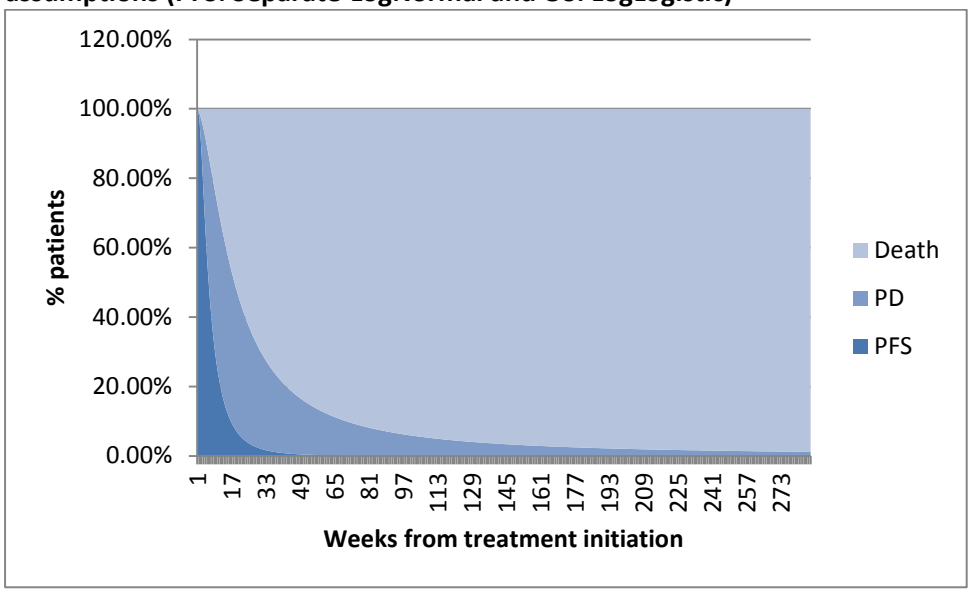


Figure 31: Markov trace for docetaxel 2nd line treatment using base-case assumptions (PFS: Separate-LogNormal and OS: LogLogistic)

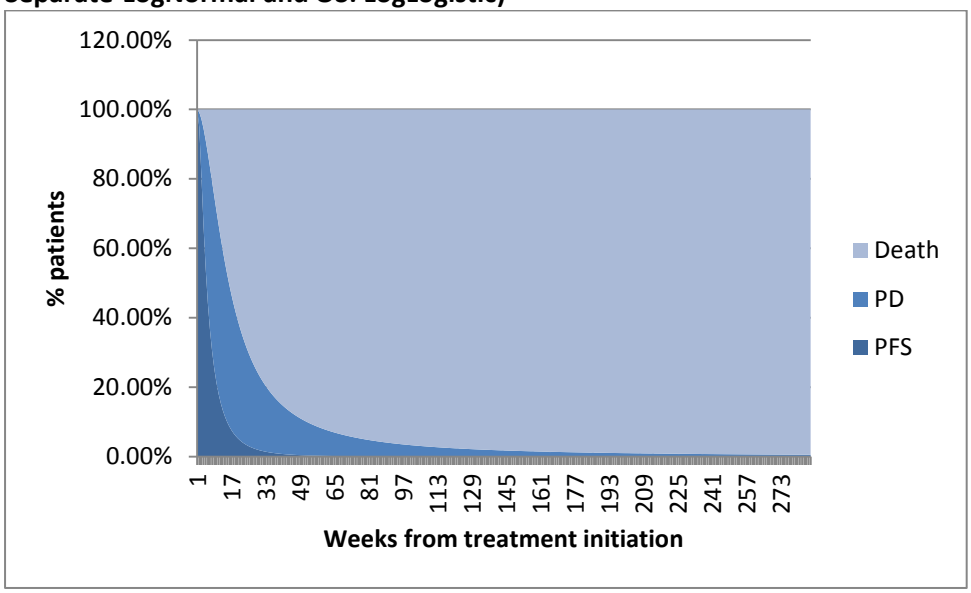
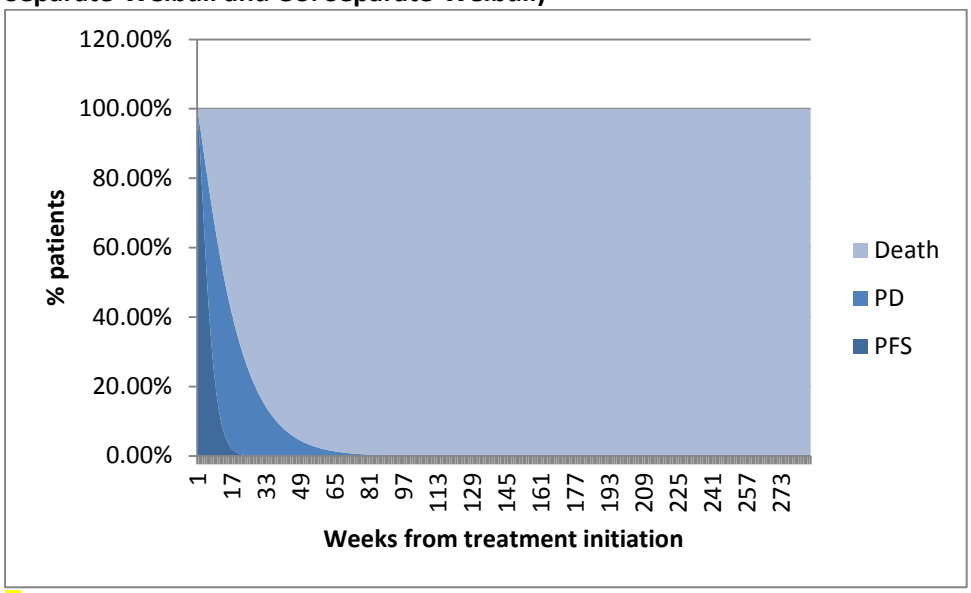


Figure 32: Markov trace for erlotinib 2nd line treatment using base-case assumptions (PFS: Separate-Weibull and OS: Separate-Weibull)



X

It should be noted that the majority of the cohort has died before the 15 year time horizon for the model, indicating that the maximum time horizon used is adequate.

7.7.3 Please provide details of how the model assumes QALYs accrued over time. For example, Markov traces can be used to demonstrate QALYs accrued in each health state over time.

The Markov model has three health states: Progressive Disease (PD), PFS and Death. Each live health state (PFS, PD) is associated with a health-related utility to estimate the QALY over the time horizon of the analysis. The cycle length in the model is 3 weeks and patients transition between states at each cycle. This component of the QALY is calculated per cycle based on the distribution of the cohort across the health states and the utility associated with being in the health state.

Utility reduction (disutility) due to AEs are applied in the model based on the estimates proportions of patients suffering from AEs in each treatment arm and are considered to occur during the PF health state. The impact of AEs on health outcomes (QALY) is calculated using the information on the duration of AEs and their impact on health-related utility on a monthly basis.

The model's default time horizon is 15 years. This has been set to cover the lifetime of the patients and fully incorporate the health outcomes of NSCLC. No discounting is required during the first year of the model; after year 1, discounting is applied at 3.5% per annum to QALYs.

7.7.4 Please indicate the life years and QALYs accrued for each clinical outcome listed for each comparator. For outcomes that are a combination of other states, please present disaggregated results. For example:

Table 121 and Table 122 illustrate the model outputs by clinical outcomes for the nintedanib plus docetaxel and docetaxel arms, respectively, using the log-normal distribution for PFS and the log-logistic distribution for OS. In addition, Table 123 and Table 124 show the model outputs by clinical outcomes for nintedanib plus docetaxel and erlotinib, respectively, using Weibull survival estimates for both PFS and OS.

Table 121: Model outputs by clinical outcomes – undiscounted (due to "accrued" in text) – nintedanib plus docetaxel arm, Lognormal/Loglogistic survivals

Outcome	LY	QALY	Cost (£)
PF outcomes	0.44	[redacted]	[redacted]
Post-progression outcomes	1.37	[redacted]	[redacted]
At time of progression outcomes	-	[redacted]	[redacted]
Overall outcomes	1.81	[redacted]	[redacted]
ALT increase	-	[redacted]	[redacted]
Anemia	-	[redacted]	[redacted]
AST increase	-	[redacted]	[redacted]
Diarrhoea - grade 2	-	[redacted]	[redacted]
Diarrhoea - grade 3/4	-	[redacted]	[redacted]
Fatigue	-	[redacted]	[redacted]
Febrile neutropenia	-	[redacted]	[redacted]
Infection	-	[redacted]	[redacted]
Nausea and vomiting	-	[redacted]	[redacted]
Neutropenia	-	[redacted]	[redacted]
Rash	-	[redacted]	[redacted]
Thrombocytopenia	-	[redacted]	[redacted]
WBC count decrease	-	[redacted]	[redacted]
LY, life years; PF, progression-free; QALY, quality-adjusted life year			

Table 122: Model outputs by clinical outcomes – undiscounted (due to "accrued" in text) - docetaxel arm, Lognormal/Loglogistic survivals

Outcome	LY	QALY	Cost (£)
PF outcomes	0.38	[redacted]	[redacted]
Post-progression outcomes	1.04	[redacted]	[redacted]
At time of progression outcomes	-	[redacted]	[redacted]
Overall outcomes	1.42	[redacted]	[redacted]
ALT increase	-	[redacted]	[redacted]
Anemia	-	[redacted]	[redacted]
AST increase	-	[redacted]	[redacted]
Diarrhoea - grade 2	-	[redacted]	[redacted]
Diarrhoea - grade 3/4	-	[redacted]	[redacted]
Fatigue	-	[redacted]	[redacted]
Febrile neutropenia	-	[redacted]	[redacted]
Infection	-	[redacted]	[redacted]
Nausea and vomiting	-	[redacted]	[redacted]
Neutropenia	-	[redacted]	[redacted]
Rash	-	[redacted]	[redacted]
Thrombocytopenia	-	[redacted]	[redacted]
WBC count decrease	-	[redacted]	[redacted]
LY, life years; PF, progression-free; QALY, quality-adjusted life year			

Table 123: Model outputs by clinical outcomes – undiscounted (due to "accrued" in text) - nintedanib+docetaxel arm, Weibull survivals

Outcome	LY	QALY	Cost (£)
PF outcomes	0.41	[redacted]	[redacted]
Post-progression outcomes	1.03	[redacted]	[redacted]
At time of progression outcomes	-	[redacted]	[redacted]
Overall outcomes	1.44	[redacted]	[redacted]
ALT increase	-	[redacted]	[redacted]
Anemia	-	[redacted]	[redacted]
AST increase	-	[redacted]	[redacted]
Diarrhoea - grade 2	-	[redacted]	[redacted]
Diarrhoea - grade 3/4	-	[redacted]	[redacted]
Fatigue	-	[redacted]	[redacted]
Febrile neutropenia	-	[redacted]	[redacted]
Infection	-	[redacted]	[redacted]
Nausea and vomiting	-	[redacted]	[redacted]
Neutropenia	-	[redacted]	[redacted]
Rash	-	[redacted]	[redacted]
Thrombocytopenia	-	[redacted]	[redacted]
WBC count decrease	-	[redacted]	[redacted]
LY, life years; PF, progression-free; QALY, quality-adjusted life year			

Table 124: Model outputs by clinical outcomes – undiscounted (due to "accrued" in text) - erlotinib arm, Weibull survivals

Outcome	LY	QALY	Cost (£)
PF outcomes	0.32	[redacted]	[redacted]
Post-progression outcomes	0.66	[redacted]	[redacted]
At time of progression outcomes	-	[redacted]	[redacted]
Overall outcomes	0.98	[redacted]	[redacted]
ALT increase	-	[redacted]	[redacted]
Anemia	-	[redacted]	[redacted]
AST increase	-	[redacted]	[redacted]
Diarrhoea - grade 2	-	[redacted]	[redacted]
Diarrhoea - grade 3/4	-	[redacted]	[redacted]
Fatigue	-	[redacted]	[redacted]
Febrile neutropenia	-	[redacted]	[redacted]
Infection	-	[redacted]	[redacted]
Nausea and vomiting	-	[redacted]	[redacted]
Neutropenia	-	[redacted]	[redacted]
Rash	-	[redacted]	[redacted]
Thrombocytopenia	-	[redacted]	[redacted]
WBC count decrease	-	[redacted]	[redacted]
LY, life years; PF, progression-free; QALY, quality-adjusted life year			

7.7.5 Please provide details of the disaggregated incremental QALYs and costs by health state, and of resource use predicted by the model by category of cost. Suggested formats are presented below.

Table 125: Summary of QALY gain by health state – discounted – Lognormal/Loglogistic survivals

Health state	QALY Nintedanib+docetaxel	QALY docetaxel	Increment	Absolute increment	% absolute increment
PF state	[redacted]	[redacted]	0.04	0.04	[redacted]
Post-progression state	[redacted]	[redacted]	0.17	0.17	[redacted]
Total	[redacted]	[redacted]	0.21	0.21	[redacted]
PF, progression-free; QALY, quality-adjusted life year Adapted from Pharmaceutical Benefits Advisory Committee (2008) Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee (Version 4.3). Canberra: Pharmaceutical Benefits Advisory Committee					

Table 126: Summary of QALY gain by health state – discounted – Weibull survivals

Health state	QALY Nintedanib+docetaxel	QALY erlotinib	Increment	Absolute increment	% absolute increment
PF state	[redacted]	[redacted]	0.06	0.06	[redacted]
Post-progression state	[redacted]	[redacted]	0.22	0.22	[redacted]
Total	[redacted]	[redacted]	0.28	0.28	[redacted]

PF, progression-free; QALY, quality-adjusted life year

Adapted from Pharmaceutical Benefits Advisory Committee (2008) Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee (Version 4.3). Canberra: Pharmaceutical Benefits Advisory Committee

Table 127: Summary of costs by health state – discounted – Lognormal/Loglogistic survivals

Health state	Cost nintedanib+docetaxel	Cost docetaxel	Increment	Absolute increment	% absolute increment
PF state	[redacted]	[redacted]	£9,547	£9,547	[redacted]
Post-progression state	[redacted]	[redacted]	£1,504	£1,504	[redacted]
At time of progression	[redacted]	[redacted]	£0	£0	[redacted]
Total	[redacted]	[redacted]	£11,051	£11,051	[redacted]

PF, progression-free

Adapted from Pharmaceutical Benefits Advisory Committee (2008) Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee (Version 4.3). Canberra: Pharmaceutical Benefits Advisory Committee

Table 128: Summary of costs by health state – discounted – Weibull survivals

Health state	Cost nintedanib+docetaxel	Cost erlotinib	Increment	Absolute increment	% absolute increment
PF state	[redacted]	[redacted]	£5,716	£5,716	[redacted]
Post-progression state	[redacted]	[redacted]	£1,855	£1,855	[redacted]
At time of progression	[redacted]	[redacted]	£0	£0	[redacted]
Total	[redacted]	[redacted]	£7,571	£7,571	[redacted]

PF, progression-free

Adapted from Pharmaceutical Benefits Advisory Committee (2008) Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee (Version 4.3). Canberra: Pharmaceutical Benefits Advisory Committee

Base-case analysis

7.7.6 Please present your results in the following table. List interventions and comparator(s) from least to most expensive and present ICERs in comparison with baseline (usually standard care) and then incremental analysis ranking technologies in terms of dominance and extended dominance.

Table 129 and Table 130 present the total discounted efficacy, cost outcomes by treatment, and incremental outcomes available in the adenocarcinoma population with OS log-logistic and PFS log-normal (base-case), and OS and PFS Weibull distribution (for vs erlotinib), respectively. As shown in Table 129, in the adenocarcinoma patient population, nintedanib plus docetaxel had higher total average per-patient lifetime cost compared to docetaxel [redacted]) and higher in all three efficacy outcomes ([redacted]) with log-logistic/log-normal survival distributions. This resulted in an incremental cost effectiveness ratio (ICER of £50,776 per QALY). When using Weibull distributions for both OS and PFS where erlotinib was allowed to be included in the analysis via use of HRs, the results (Table 130) showed that nintedanib plus docetaxel had a higher total cost per patient than erlotinib[redacted]). The same trend was observed in the effectiveness outcomes where nintedanib plus docetaxel had the highest LYs and QALYs compared to erlotinib [redacted]). This resulted in an ICER of £27,008 per QALY.

Note that the Weibull extrapolation underestimates OS, however it provided the best statistical fit of the models which did not violate the proportional hazards assumptions. This was needed as the PFS and OS curves for erlotinib are generated by hazards ratios from the indirect treatment comparison.

Table 129: Distributions used – OS: Log-logistic; PFS: Log-normal

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£) versus baseline (QALYs)	ICER (£) incremental (QALYs)
Nintedanib + Docetaxel	[redacted]	[redacted]	[redacted]	-	-	-	-	-
Docetaxel	[redacted]	[redacted]	[redacted]	£11,051	0.33	0.22	£50,776	£50,776
ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years								

Table 130: Distributions used – OS: Weibull Distributions; PFS – Weibull Survival

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£) versus baseline (QALYs)	ICER (£) incremental (QALYs)
Nintedanib + Docetaxel	[redacted]	[redacted]	[redacted]	-	-	-	-	-
Erlotinib	[redacted]	[redacted]	[redacted]	£7,571	0.43	0.28	£27,008	£27,008
ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years								

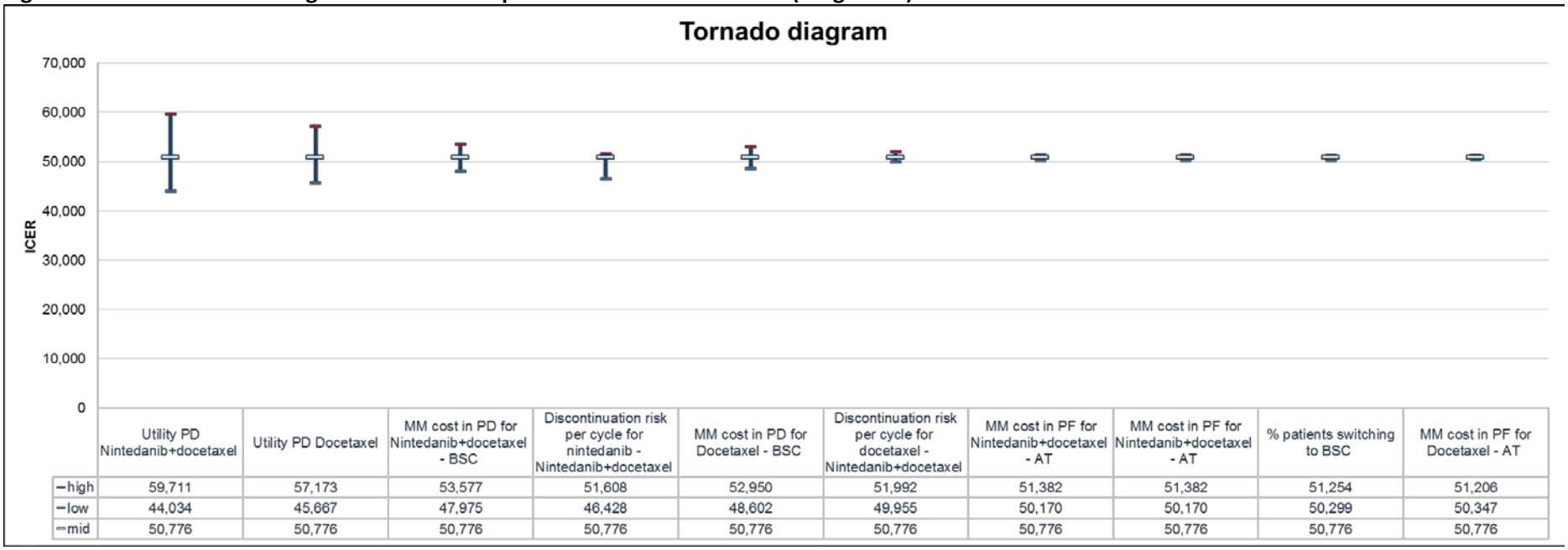
Sensitivity analyses

7.7.7 Please present results of deterministic sensitivity analysis. Consider the use of tornado diagrams.

The ten variables which have the largest impact on the base-case incremental cost-effectiveness ratio are shown in Figure 33 and Figure 34 for the comparison of nintedanib plus docetaxel versus docetaxel and versus erlotinib, respectively.

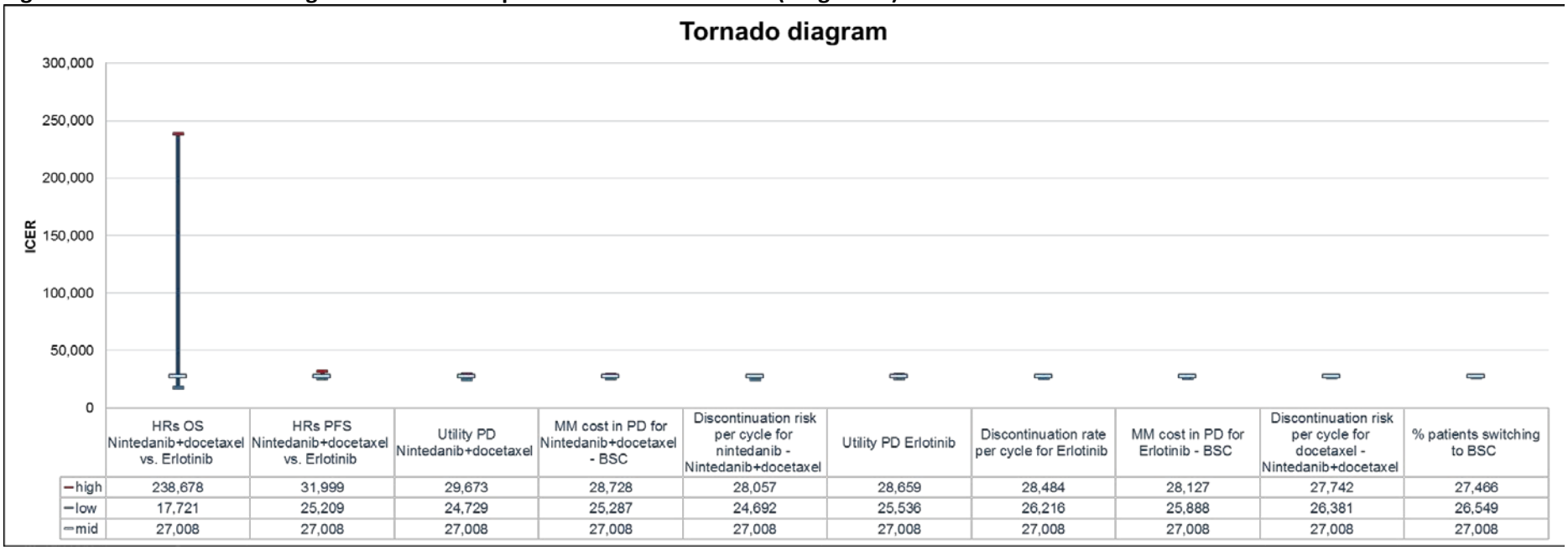
The tornado diagrams show that the ICER was most sensitive to univariate change in utility values after progression. However, it is unlikely that patients progressing on one or the other treatments had such very different utility values after progression. Cost of BSC had a small impact as well as discontinuation risk for either component of the combination therapies. All other variables, including AE related costs or disutilities had very minimal impact, ceteris paribus, on the ICER results in terms of cost/QALYs. For the comparison of nintedanib to erlotinib, the HR for OS was the single most influential variable that appeared in the tornado diagram. Other parameters had similar effect on ICER than in the chemotherapy comparison.

Figure 33: OWSA Tornado Diagram – Nintedanib plus Docetaxel vs Docetaxel (range 20%)



	Utility PD Nintedanib+docetaxel	Utility PD Docetaxel	MM cost in PD for Nintedanib+docetaxel - BSC	Discontinuation risk per cycle for nintedanib - Nintedanib+docetaxel	MM cost in PD for Docetaxel - BSC	Discontinuation risk per cycle for docetaxel - Nintedanib+docetaxel	MM cost in PF for Nintedanib+docetaxel - AT	MM cost in PF for Nintedanib+docetaxel - AT	% patients switching to BSC	MM cost in PF for Docetaxel - AT
high	59,711	57,173	53,577	51,608	52,950	51,992	51,382	51,382	51,254	51,206
low	44,034	45,667	47,975	46,428	48,602	49,955	50,170	50,170	50,299	50,347
mid	50,776	50,776	50,776	50,776	50,776	50,776	50,776	50,776	50,776	50,776

Figure 34: OWSA Tornado Diagram – Nintedanib plus Docetaxel vs Erlotinib (range 20%)



	HRs OS Nintedanib+docetaxel vs. Erlotinib	HRs PFS Nintedanib+docetaxel vs. Erlotinib	Utility PD Nintedanib+docetaxel	MM cost in PD for Nintedanib+docetaxel - BSC	Discontinuation risk per cycle for nintedanib - Nintedanib+docetaxel	Utility PD Erlotinib	Discontinuation rate per cycle for Erlotinib	MM cost in PD for Erlotinib - BSC	Discontinuation risk per cycle for docetaxel - Nintedanib+docetaxel	% patients switching to BSC
high	238,678	31,999	29,673	28,728	28,057	28,659	28,484	28,127	27,742	27,466
low	17,721	25,209	24,729	25,287	24,692	25,536	26,216	25,888	26,381	26,549
mid	27,008	27,008	27,008	27,008	27,008	27,008	27,008	27,008	27,008	27,008

7.7.8 Please present the results of a PSA, and include scatter plots and cost-effectiveness acceptability curves.

The probability of cost-effectiveness, at different willingness to pay thresholds for the comparison of nintedanib plus docetaxel against docetaxel in the second-line setting is shown in Table 131.

Table 131: Probability of cost-effectiveness in the 2nd line setting

Intervention	Comparator	£30,000 per QALY WTP	£50,000 per QALY WTP
Nintedanib + docetaxel	Docetaxel	2%	50%

The comparison between the deterministic and probabilistic results for nintedanib plus docetaxel versus docetaxel in shown in Table 132

Table 132: Comparison of ICERs obtained from deterministic and probabilistic sensitivity analyses for nintedanib plus docetaxel versus docetaxel

	Incremental cost	Incremental QALY	ICER
Deterministic Values	£11,051	0.22	£50,776
Average value for PSA	£10,916	0.22	£49,965

The cost-effectiveness scatter plot and acceptability curves for nintedanib plus docetaxel versus docetaxel are displayed in Figure 35 and Figure 36, respectively.

Figure 35: Incremental cost-effectiveness scatterplot for nintedanib + docetaxel versus docetaxel

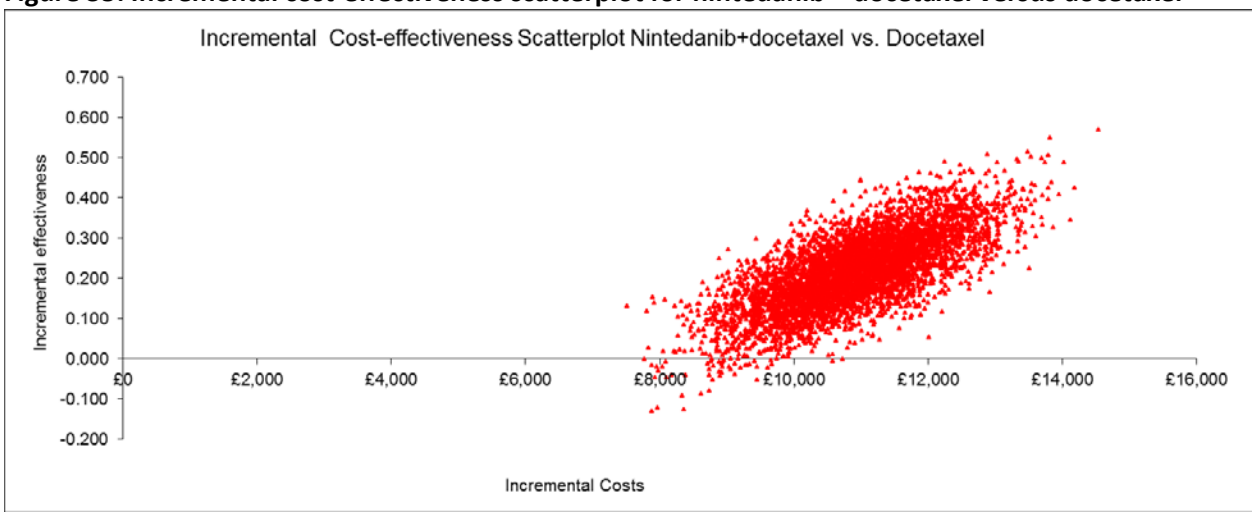
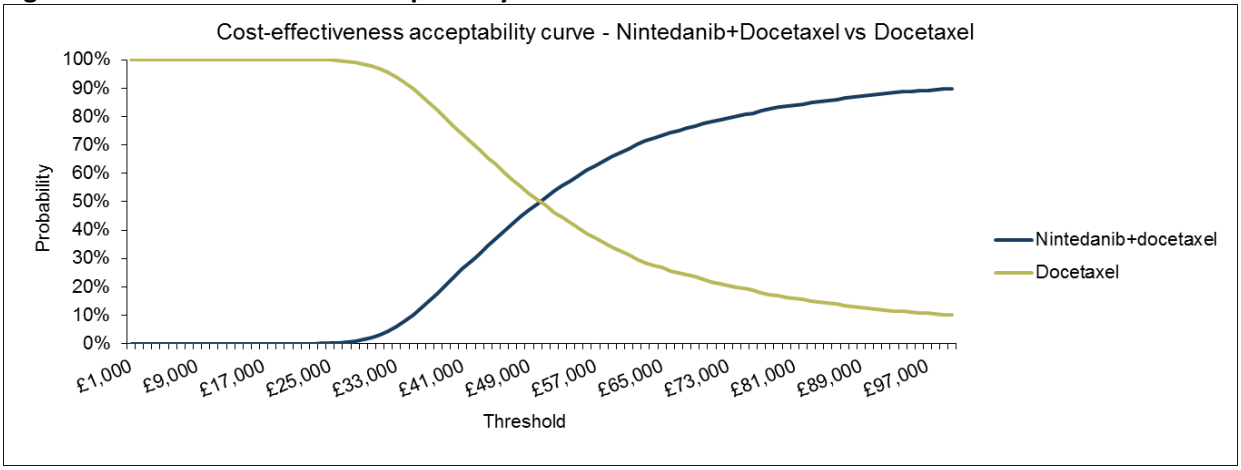


Figure 36: Cost effectiveness acceptability curve for nintedanib +docetaxel versus docetaxel



The probability of cost-effectiveness, at different willingness to pay thresholds for the comparison of nintedanib plus docetaxel against erlotinib in the second-line setting is shown in Table 133.

Table 133: Probability of cost-effectiveness in the 2nd line setting: Weibull

Intervention	Comparator	£30,000 per QALY WTP	£50,000 per QALY WTP
Nintedanib + docetaxel	Erlotinib	65%	94%

The comparison between the deterministic and probabilistic results for nintedanib plus docetaxel versus erlotinib in shown in Table 134.

Table 134: Comparison of ICERs obtained from deterministic and probabilistic sensitivity analyses for nintedanib + docetaxel versus erlotinib

	Incremental cost	Incremental QALY	ICER
Deterministic Values	£7,571	0.28	£27,008
Average value for PSA	£7,518	0.27	£27,484

The cost-effectiveness planes and acceptability curves for nintedanib plus docetaxel versus erlotinib are displayed in Figure 37 and Figure 38, respectively.

Figure 37: Cost-effectiveness plane for nintedanib plus docetaxel versus erlotinib

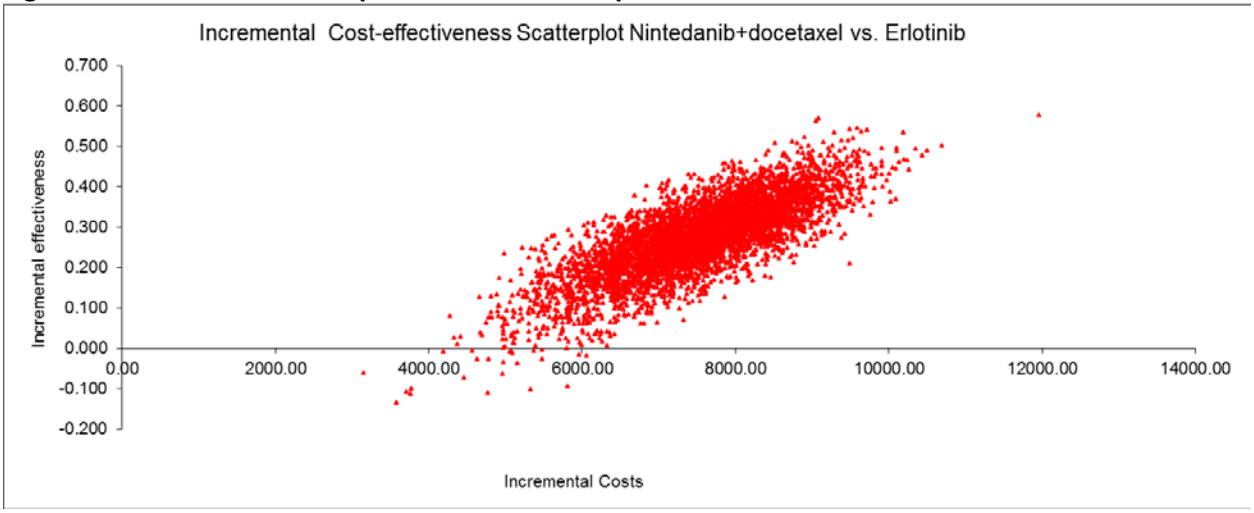
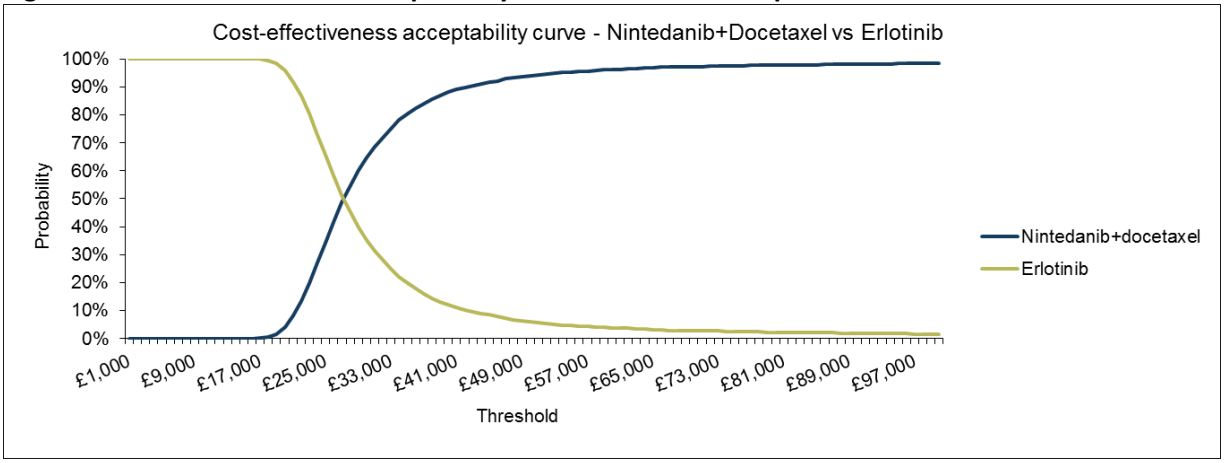


Figure 38: Cost effectiveness acceptability curve for nintedanib plus docetaxel versus erlotinib



7.7.9 Please present the results of scenario analysis. Include details of structural sensitivity analysis.

Survival Modelling Scenarios

Methods of survival modelling are a critical element of all cost-effectiveness analyses, and this holds true for this model. Using the Weibull distribution increases the ICER; however the Weibull distribution is not an accurate representation of the course of the disease for the patient population. It underestimates OS; a point agreed upon by all five clinicians at the advisory board (see section 7.3.5).

Using the Kaplan-Meier curves and SEER data or Kaplan-Meier curves and LUCADA data did not change the ICERs much, supporting our base-case assumption.

Using the Kaplan-Meier curves only to the time they are available for PFS does not change the ICER much as the PFS trial data was fairly mature. Using the Kaplan-Meier curves only to the time where they are available for PFS and OS increases the ICER by limiting any benefits that may accrue after the time that was captured by the trial, and does not accurately represent the course of the disease for the entire patient population.

Note that in all OS scenarios except when using Weibull extrapolation or no extrapolation of OS data, nintedanib + docetaxel extends OS by over 3 months compared with docetaxel monotherapy, and therefore meets the end of life criteria.

The various survival modelling scenarios for the comparison of nintedanib plus docetaxel versus docetaxel are shown below in Table 135.

Table 135: Survival Modelling Scenarios

PFS	OS	Incremental LYs	Incremental costs	Incremental QALYs	ICER
Separate – Lognormal (base-case)	Separate – Loglogistic (base-case)	0.33	£11,051	0.22	£50,776
Separate - Weibull	Separate - Weibull	0.22	£9,852	0.14	£69,884
KM Curve	KM Curve	0.11	£9,425	0.08	£119,209
KM Curve - used until time horizon	Mixed: KM & SEER-Lognormal	0.27	£10,304	0.18	£56,769
KM Curve - used until time horizon	Mixed: KM & LUCADA-Lognormal	0.26	£10,245	0.17	£58,660
KM Curve - used until time horizon	Mixed curves: KM & Separate Loglogistic	0.34	£10,637	0.22	£48,264
KM Curve - used until time horizon	Mixed curvex: KM & Separate Weibull	0.23	£10,071	0.15	£65,274

ICER = incremental cost-effectiveness ratio; KM = Kaplan-Meier; LY = life years; OS = overall survival; PFS = progression-free survival; QALYs = quality-adjusted life years

Indirect Comparison Scenario

The effect of varying the indirect comparison scenario for the comparison of nintedanib plus docetaxel against erlotinib is illustrated in Table 136. Using results from the NMA Scenario Analysis network instead of the NMA Base-case Analysis network increased the ICER of nintedanib plus docetaxel compared to erlotinib.

Table 136: Indirect Comparison Scenario – Erlotinib, Adenocarcinoma

PFS HR	OS HR	Incremental costs	Incremental QALYs	ICER
NMA Base-case network: 0.70	NMA Base-case network: 0.64	£7,571	0.28	£27,008
NMA Scenario Analysis network, Fixed-effect	NMA Scenario Analysis network, Fixed-effect	£6,952	0.20	£34,509

model: 0.68	model: 0.74			
NMA Scenario Analysis network, Random-effect model: 0.68	NMA Scenario Analysis network, Random-effect model: 0.74	£6,952	0.20	£34,509

HR = hazard ratio; ICER = incremental cost-effectiveness ratio; OS = overall survival; NMA = network meta-analysis; PFS = progression-free survival; QALYs = quality-adjusted life years

Resource Use Scenarios

Table 137 lists the effect of using various resource use scenarios on the ICERs. Switching from using data from EEs to the numbers from the afatinib submission does not have a large impact on ICERs.

Table 137: Impact of Resource Use Scenarios

Scenarios	ICER (£/QALY) Nindetanib + Docetaxel versus:	
	Docetaxel (Lognormal, Loglogistic)	Erlotinib (Weibull)
Base-case	£50,776	£27,008
Afatinib Submission	£52,692	£25,301

ICER = incremental cost-effectiveness ratio; QALYs = quality-adjusted life years

Utility Scenarios

Table 138 shows the effect of various utility scenarios on the ICER. Changing the method of applying the trial-based utility analysis beyond 30 weeks was not influential: LOCF approach results were very similar to the base-case of linear trendline approach. However, applying values published by Chouaid et al (2013)(128), which included values for progressive disease which were lower than the data derived from the trial, increased ICERs substantially. The impact was most pronounced in the docetaxel monotherapy comparison. Note that the values from the Chouaid paper implemented within the model are likely to be conservative; the post-progression utility is assumed to be equal to the third/fourth line progressive disease state. In reality, the patients in the model are more likely to also include patients from the second-line progressive disease and third/fourth line PF states, both of which have higher utilities than the third/fourth line progressive disease state.

Table 138: Impact of Utility Scenarios

Scenarios	ICER (£/QALY) Nindetanib + Docetaxel versus:	
	Docetaxel (Lognormal, Loglogistic)	Erlotinib (Weibull)
Base-case	£50,776	£27,008
LOCF for PFS	£51,496	£26,961
Chouaid (2013) for both PFS and PD	£65,408	£33,464

ICER = incremental cost-effectiveness ratio; LOCF = last-observation carried forward; PFS = progression-free survival; QALYs = quality-adjusted life years

Time Horizon Variation

Table 139 shows the impact of varying the time horizon on the ICERs. Lifetime was about 15 years (1.16% of patients still alive on nintedanib plus docetaxel and 0.52% still alive on docetaxel

monotherapy). Changes in the results when time horizon was varied from 10 to 15 years were fairly small; by this time, the proportion of patients dead was 0.9875 on docetaxel monotherapy, so the effect of the additional years on the results was small. Decreasing the time horizon increased the ICER, because while most costs were presented earlier in time and were still incorporated, the OS gain could not be fully taken into account; shortening the time horizon is likely to underestimate the OS benefits and as a result a higher ICER is produced.

Table 139: Impact of time horizon

Time Horizon	ICER (£/QALY), Nintedanib + Docetaxel versus:	
	Docetaxel (Lognormal, Loglogistic)	Erlotinib (Weibull)
Based on LUME-Lung 1 trial	£86,023	£29,744
3 years	£98,119	£31,816
5 years	£70,951	£27,740
10 years	£55,132	£27,013
15 years	£50,776	£27,008

ICER = incremental cost-effectiveness ratio; QALYs = quality-adjusted life years

Impact of Discount Rate Variation (Costs and Benefits)

Table 140 and Table 141 illustrate the effect of varying the discount rate on the ICERs. Zero percent and 6% discount rates were evaluated separately for costs and benefits, in addition to the 3.5% rate. Higher discount rates led to lower ICERs however the changes were fairly small.

Table 140: Impact of Discount Rates: Log-normal, Log-logistic

Cost per QALY gained (ICER) nintedanib + docetaxel vs docetaxel		Discount rate for costs		
		0%	3.5%	6%
Discount rate for benefits	0%	£45,176	£43,390	£42,322
	3.5%	£52,866	£50,776	£49,526
	6%	£58,474	£56,163	£54,780

ICER = incremental cost-effectiveness ratio; QALYs = quality-adjusted life years

Table 141: Impact of Discount Rates: Weibull Survivals

Cost per QALY gained (ICER) nintedanib + docetaxel vs erlotinib		Discount rate for costs		
		0%	3.5%	6%
Discount rate for benefits	0%	£25,978	£25,121	£24,565
	3.5%	£27,928	£27,008	£26,410
	6%	£29,329	£28,362	£27,735

ICER = incremental cost-effectiveness ratio; QALYs = quality-adjusted life years

Impact of Discount Rate Variation (Erlotinib PAS)

As Boehringer Ingelheim does not have access to the net price of erlotinib used in practise, scenario analyses were performed by reducing the list price for erlotinib in £5% increments from a 5% discount to a 95% discount. With a 95% discount the ICER for nintedanib plus docetaxel vs erlotinib was £44,183.

Table 142: Impact of Erlotinib discount rate on ICER: Weibull survivals

	Discount Applied to Drug Cost						
Erlotinib	**Base-case (0% discount)**	**5%**	**10%**	**15%**	**20%**	**25%**	**30%**
Cost per pack	£1,632	£1,550.40	£1,468.80	£1,387.20	£1,305.60	£1,224	£1,142.40
ICER	£27,008	£27,934	£28,866	£29,797	£30,729	£31,660	£32,592
Erlotinib	**35%**	**40%**	**45%**	**50%**	**55%**	**60%**	**65%**
Cost per pack	£1,060.80	£979.20	£897.60	£816	£734.4	£652.80	£571.20
ICER	£33,524	£34,455	£35,387	£36,318	£37,250	£38,182	£39,113
Erlotinib	**70%**	**75%**	**80%**	**85%**	**90%**	**95%**	
Cost per pack	£489.60	£408	£326.40	£244.80	£163.20	£81.6	
ICER	£40,045	£40,977	£41,908	£42,840	£43,771	£44,703	

ICER = incremental cost-effectiveness ratio

7.7.10 What were the main findings of each of the sensitivity analyses?

PSA was carried out using 5,000 iterations of the cost-effectiveness model. The incremental cost-effectiveness scatter plots are presented in section 7.7.8. When nintedanib plus docetaxel is compared to docetaxel monotherapy (Figure 35), the points are tightly scattered, with the majority of point lying in the north-east quadrant, representing points where nintedanib plus docetaxel is more effective but more costly. When nintedanib plus docetaxel is compared to erlotinib (Figure 37), the majority of the PSA outcome points are also in the north east-quadrant.

The corresponding cost-effectiveness acceptability curves are shown in Figure 36 and Figure 38. When compared to docetaxel monotherapy, nintedanib plus docetaxel has a 2% probability of being cost-effective at the £30,000 per QALY threshold and a 50% chance of being cost-effective at a £50,000 per QALY threshold (Table 131). Nintedanib plus docetaxel has a 68% probability of being cost-effective at the £30,000 per QALY threshold and a 95% probability of being cost-effective at a £50,000 per QALY threshold vs erlotinib. (Table 133).

In the UK setting the base-case ICER is £50,234 per QALY with docetaxel monotherapy as the comparator, and £26,488 per QALY with erlotinib as comparator.

Second line treatment with nintedanib plus docetaxel extends PFS and OS time versus both comparators considered in the model.

The base-case ICER's in the first-line setting are sensitive to changes in the PFS and OS HRs and to the costs and utilities associated with the post-progression state. Therefore the assumptions on the patient treatment pathway and health outcomes between progression and death have an impact on cost-effectiveness. Health related utility used in the base-case is from the clinical trial and therefore should provide an accurate representation of the health state of patients. Detailed resource use was provided from a EE (Section 7.3.5) and the sensitivity analysis using data from the afatnib submission (140) derived from trial data (LUX Lung trials (138)), gave a similar ICER to the base-case, suggesting the values used are sound.

7.7.11 What are the key drivers of the cost-effectiveness?

The key drivers of the cost effectiveness results are the OS extrapolation method and the post-progression costs and utilities.

7.8 Validation

7.8.1 Please describe the methods used to validate and quality assure the model. Provide references to the results produced and cross-reference to evidence identified in the clinical, quality of life and resources sections.

A number of steps were taken to ensure that the analysis was validated. These include:

- External review by leading UK clinical expert (see Section 7.3.5)
    - The model structure was developed in conjunction with leading clinicians. This clinical validation serves to ensure that the model adheres to the clinical course of the disease and is reflective of current clinical practice.
- Sensitivity analysis outlined in Section 7.6.
- Validation by model developers
    - Apart from the interviews with the UK clinical experts (discussed in Section 7.3.5), a senior modeller within the model developers organisation (with no involvement in the afatinib model's development) perform a detailed QA check on the model.
- Validation by manufacturer

- This involved increasing and decreasing various parameters or changing assumptions in the model and then monitoring the impact on outputs. If the outputs were unexpected, further checks were made to determine whether this was the result of an error in the model.

7.9 Subgroup analysis

For many technologies, the capacity to benefit from treatment will differ for patients with differing characteristics. This should be explored as part of the reference-case analysis by providing separate estimates of clinical and cost effectiveness for each relevant subgroup of patients.

This section should be read in conjunction with NICE's 'Guide to the methods of technology appraisal', section 5.10.

Types of subgroups that are not considered relevant are those based solely on the following factors.

- Individual utilities for health states and patient preference.
- Subgroups based solely on differential treatment costs for individuals according to their social characteristics.
- Subgroups specified in relation to the costs of providing treatment in different geographical locations within the UK (for example, when the costs of facilities available for providing the technology vary according to location).

7.9.1 Please specify whether analysis of subgroups was undertaken and how these subgroups were identified. Were they identified on the basis of an a priori expectation of differential clinical or cost effectiveness because of known, biologically plausible, mechanisms, social characteristics or other clearly justified factors? Cross-reference the response to section 6.3.7.

No subgroup analysis was undertaken.

7.9.2 Please clearly define the characteristics of patients in the subgroup.

Not applicable.

7.9.3 Please describe how the statistical analysis was undertaken.

Not applicable.

7.9.4 What were the results of the subgroup analysis/analyses, if conducted? Please present results in a similar table as in section 7.7.6 (Base-case analysis).

Not applicable.

7.9.5 Were any obvious subgroups not considered? If so, which ones, and why were they not considered? Please refer to the subgroups identified in the decision problem in section 5.

No.

7.10 Interpretation of economic evidence

7.10.1 Are the results from this economic evaluation consistent with the published economic literature? If not, why do the results from this evaluation differ, and why should the results in the submission be given more credence than those in the published literature?

This is the first economic evaluation of nintedanib in this indication. Therefore there are no published studies with which to draw comparison.

7.10.2 Is the economic evaluation relevant to all groups of patients who could potentially use the technology as identified in the decision problem in section 5?

Yes. The economic evaluation covers the relevant patient group.

7.10.3 What are the main strengths and weaknesses of the evaluation? How might these affect the interpretation of the results?

The main strength of the economic evaluation is high-quality data that underpins it from the LUME Lung 1 trial. This pivotal trial provides a wealth of robust clinical and HRQL data for the modelled outcomes for the principle comparison of interest. This strength provides confidence in the results of the analysis.

The well-known limitations of a state transition Markov structure also apply; however, as the modelled problem can be simplified without major assumptions regarding metastatic cancers, this was an acceptable approach to use. The model structure was developed in conjunction with leading clinicians. This clinical validation serves to ensure that the model adheres to the clinical course of the disease and is reflective of current clinical practice.

Due to the relatively short duration of the LUME-Lung 1 trial, survival functions had to be extrapolated to 15 years after the follow-up period in the study, which increased uncertainty in the model results. To address this uncertainty, multiple sensitivity analyses were incorporated, and the long-term mortality was compared to mortality in the UK general population to ensure that OS incorporated non-NSCLC specific mortality. The extrapolated OS curves lay well below the general population survival curve. Furthermore, extrapolated data were compared to data from adenocarcinoma patients in SEER and LUCADA as well as discussed with a panel of EEs.

The LUME-Lung 1 trial had only two treatment arms; therefore indirect comparison had to be performed to incorporate the additional comparator erlotinib. The indirect comparison performed by Evidera included multiple scenarios, based on various assumptions, given that trials usually include a mix of NSCLC histologies and EGFR-TK mutation status was not always reported.

The additional comparators are incorporated with the help of HR which implies that the shape of the survival curves for the additional comparators are the same as the shape of the nintedanib + docetaxel curves. In this analysis, to be able to incorporate the HRs, distributions with proportional hazard assumptions, such as Weibull had to be assumed appropriate – despite evidence that Weibull may underestimate OS.

Resource use in the model was derived from a detailed interview with one EE (section 7.3.5), however a sensitivity analysis using resource use data and costing from a recent submission for afatanib, using resource use directly from a phase III trial was carried out (section 7.7.9).

Extensive sensitivity analysis and validation of the model was undertaken to ensure that analysis was robust.

End of Life Criteria

Nintedanib plus docetaxel in second-line treatment of NSCLC of adenocarcinoma histology fulfils the 'End of life' criteria.

- Patients with advanced NSCLC have a short life expectancy of less than 24 months on average. Using the extrapolated results from the LUME Lung 1 trial data implemented in the cost effectiveness model, the median OS of patients on docetaxel monotherapy (current standard of care) is 10.23 months and the mean OS is 15.96 months.
- The total eligible population for nintedanib plus docetaxel is 745 (see section 8.1)
- Extension to life due to nintedanib plus docetaxel vs docetaxel monotherapy in the target population with the base-case assumptions within the model is a mean of 3.96 months. The extension in OS over erlotinib is a mean of 5.16 months.

7.10.4 What further analyses could be undertaken to enhance the robustness/completeness of the results?

Additional data regarding the resource use would be useful to validate the values used within the model.

Section C – Implementation

8 Assessment of factors relevant to the NHS and other parties

The purpose of this section is to provide an analysis of any factors relevant to the NHS and other parties that may fall outside the remit of the assessments of clinical effectiveness and cost effectiveness. This will allow the subsequent evaluation of the budget impact analysis. Such factors might include issues relating to service organisation and provision, resource allocation and equity, societal or ethical issues, plus any impact on patients or carers.

8.1 How many patients are eligible for treatment in England and Wales? Present results for the full marketing authorisation/CE marking and for any subgroups considered. Also present results for the subsequent 5 years.

The estimated epidemiology of NSCLC and treatment rates with current modalities in England and Wales from November 2014 to November 2018 is presented in Table 143.

The population estimates for England and Wales are obtained from the Office of National Statistics population projections for the end of 2013 for each country (141). The prevalence of NSCLC in England is derived from Hospital Episode Statistics (HES) data, and the prevalence in Wales is assumed to be the same as that for England(142). The National Lung Cancer Audit Report in 2012 estimated that 83.3% of patients with lung cancer are expected to have NSCLC (interpreted as lung cancer excluding small-cell and mesothelioma) (143). The Audit also reports that 24.2% of English NSCLC patients and 25.1% of Welsh NCSLS patients are anticipated to have stage IIIb/IV and PS0-1(143).

In order to estimate the number of patients eligible for treatment, the proportion of stage IIIb/IV NSCLC patients with PS0-1 who have a confirmed histological diagnosis of adenocarcinoma sub-type is calculated as 39.6% of patients with stage IIIb/IV NSCLC with PS0-1 (68). Internal market share estimates predict that approximately 78.6% of the resultant subgroup of patients will be treated with first-line chemotherapy, of which 24.1% of patients will progress after first-line therapy and be eligible for second-line treatment (35). As a result, a total of 745 patients are expected to be eligible

for second-line treatment of stage IIIb/IV NSCLC with PS0-1 of the adenocarcinoma sub-type. As there is no population growth assumed, the eligible population remains constant from 2014 to 2018.

Table 143: Estimated number of patients eligible for second-line treatment

	England		Wales		Total patient numbers
	Proportion	Number of patients	Proportion	Number of patients	
Population		53,563,021		3,048,120	56,611,141
Patients with lung cancer	0.0870%	46,618.51	0.0870%	2,652.93	49,271
Patients with NSCLC	83.3%	38,833.22	83.3%	2,209.89	41,043
Patients with Stage IIIb/IV NSCLC and PS0-1	24.2%	9,385.62	25.1%	554.08	9,940
NSCLC patients with adenocarcinoma sub-type	39.6%	3,716.70	39.6%	219.42	3,936
Treatment Eligibility-Adenocarcinoma					
1st-line	78.6%	2,919.79	78.6%	172.37	3,092
2nd-line	24.1%	703.16	24.1%	41.51	745

8.2 What assumption(s) were made about current treatment options and uptake of technologies?

The budget impact model assumes that the OS of each patient for each treatment can be split into two treatment phases: active second-line treatment and BSC in second-line following active treatment. The assumptions for the length of time each patient spends in each treatment phase for each treatment option are depicted in more detail in Table 144.

Table 144: Total length of treatment, length of active treatment and length of BSC assumptions

	Total treatment length in second-line	Length of active treatment in second-line	Length of BSC in second-line following active treatment
	OS (months)	PFS (months)	OS - PFS (months)
Erlotinib	7.75	2.37	5.38
Docetaxel	10.30	2.80	7.50
Nintedanib + Docetaxel	12.60	4.00	8.60

The average total second-line treatment length experienced by each patient is given by the OS accorded by the respective treatment options (3, 144). The average length of active treatment in second-line by each patient is given by the median PFS accorded by the respective treatment options (3, 144). It is assumed that active second-line treatment will be discontinued once a patient's lung cancer is observed to have progressed. Since there is currently no NICE-approved third-line treatment, it is assumed that patients will receive BSC following the discontinuation of active second-line treatment.

8.3 What assumption(s) were made about market share (when relevant)?

Within the share of 745 patients eligible for second-line treatment, a proportional uptake of nintedanib in combination with docetaxel is envisaged under two scenarios: the existing treatment scenario (without nintedanib) and a new treatment scenario (with nintedanib in the expected mix of treatments). In the existing treatment scenario, internal data has estimated that the market share for erlotinib is [redacted]%, and the market share of docetaxel is [redacted]% (Table 145). Once nintedanib is introduced into the market, internal market share assumptions predict that the market share of nintedanib will increase from [redacted] under the new treatment scenario (Table 145) (145).

Table 145: Market share assumptions for the existing and new treatment scenarios

	Existing	New Treatment Scenario			
Market Share		**2014/15**	**2015/16**	**2016/17**	**2017/18**
Erlotinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Docetaxel	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Nintedanib + Docetaxel	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

8.4 In addition to technology costs, please consider other significant costs associated with treatment that may be of interest to commissioners (for example, procedure codes and programme budget planning).

No additional costs are expected.

8.5 What unit costs were assumed? How were these calculated? If unit costs used in health economic modelling were not based on national reference costs or the PbR tariff, which HRGs reflected activity?

The drug acquisition and administration costs in this section are identical to those assumed in the cost-effectiveness evaluation in Section 6.10. Table 146 presents a summary of the drug costs per

month and IV administration costs per month for each comparator for both the active second-line treatment phase and the BSC phase. As nintedanib is taken orally, it is not associated with any additional administration costs.

Table 146: Drug acquisition and administration costs per month

	Second-line active treatment			Second-line BSC		
	Cost of Drug (per month)	IV Administration Cost (per month)	Total Cost (per month)	Cost of Drug (per month)	IV Administration Cost (per month)	Total Cost (per month)
Erlotinib	£1,631.53	£0.00	£1,631.53	£418	£0.00	£418
Docetaxel	£49	£221.43	£270.43	£418	£0.00	£418
Nintedanib + Docetaxel	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

The default prices have been determined as follows:

- Erlotinib 150 mg, 30 tablet pack: £1,631.53 (146).
- Docetaxel: £720.10 per cycle (21 days). This converts to a 30 day cost of £1,028.71 (34, 146).
- Nintedanib plus Docetaxel: the list price for nintedanib [redacted]).
- BSC: £418 per month as per the afatinib NICE submission (TA310) and EE (43).
- IV administration costs: Based on the NHS National Schedule of Reference Costs, it costs £330 per cycle (21 days) to deliver complex chemotherapy, including prolonged infusion treatment at first attendance (HRG currency code SB14Z). This converts to a 30 day cost of £471.43 (147).

8.6 Were there any estimates of resource savings? If so, what were they?

There are no additional estimates of resource savings.

8.7 What is the estimated annual budget impact for the NHS in England and Wales

The budget impact is estimated as the number of patients and associated costs for treating those patients according to the assumed market shares in both the existing and new treatment scenario. Table 147 and Table 148 show the number of patients eligible for each comparator treatment and the associated costs for 2014 to 2018 in the existing treatment scenario (without nintedanib) and the new treatment scenario (with nintedanib), respectively. Please note that it is assumed that there are only 5 months in 2014/15, the first year as nintedanib is expected to launch in November 2014.

Table 147: Patient numbers and associated costs in existing treatment scenario

Patients				
	2014/15	**2015/16**	**2016/17**	**2017/18**
Erlotinib	275.53	661.27	661.27	661.27
Docetaxel	34.75	83.40	83.40	83.40
Nintedanib + Docetaxel	-	-	-	-
Total patients	310.28	744.67	744.67	744.67
Costs				
Erlotinib	£1,685,012.82	£4,044,030.76	£4,044,030.76	£4,044,030.76
Docetaxel	£135,259.32	£324,622.36	£324,622.36	£324,622.36
Nintedanib + Docetaxel	-	-	-	-
Total costs	£1,820,272.13	£4,368,653.12	£4,368,653.12	£4,368,653.12

Table 148: Patient numbers and associated costs in new treatment scenario

Patients				
	2014/15	**2015/16**	**2016/17**	**2017/18**
Erlotinib	[redacted]	[redacted]	[redacted]	[redacted]
Docetaxel	[redacted]	[redacted]	[redacted]	[redacted]
Nintedanib + Docetaxel	[redacted]	[redacted]	[redacted]	[redacted]
Total patients	310.28	744.67	744.67	744.67
Costs				
Erlotinib	£1,638,674.96	£3,487,976.53	£2,881,371.91	£2,375,868.07
Docetaxel	£131,539.69	£279,986.79	£231,293.43	£190,715.64
Nintedanib + Docetaxel	[redacted]	[redacted]	[redacted]	[redacted]
Total costs	[redacted]	[redacted]	[redacted]	[redacted]

A comparison of the differences in patient numbers and treatment costs are show in Table 149 below, and a summary of the expected net budget impact are illustrated in Table 150.

Table 149: Treatment differences between existing and new treatment scenarios

Change in Patients				
	2014/15	**2015/16**	**2016/17**	**2017/18**
Erlotinib	[redacted]	[redacted]	[redacted]	[redacted]
Docetaxel	[redacted]	[redacted]	[redacted]	[redacted]
Nintedanib + Docetaxel	[redacted]	[redacted]	[redacted]	[redacted]
Total patients	[redacted]	[redacted]	[redacted]	[redacted]

Erlotinib	[redacted]	[redacted]	[redacted]	[redacted]
Docetaxel	[redacted]	[redacted]	[redacted]	[redacted]
Nintedanib + Docetaxel	[redacted]	[redacted]	[redacted]	[redacted]
Total costs	[redacted]	[redacted]	[redacted]	[redacted]

Table 150: Summary of budget impact

Budget Impact				
	2014/15	**2015/16**	**2016/17**	**2017/18**
Existing Scenario	[redacted]	[redacted]	[redacted]	[redacted]
New Scenario	[redacted]	[redacted]	[redacted]	[redacted]
Change in Costs	[redacted]	[redacted]	[redacted]	[redacted]
Cumulative Cost Impact	[redacted]	[redacted]	[redacted]	[redacted]

8.8 Are there any other opportunities for resource savings or redirection of resources that it has not been possible to quantify?

There are no additional resource savings or redirection of resources expected.

9 References

Please use a recognised referencing style, such as Harvard or Vancouver.

1. Hilberg F, Roth GJ, Krssak M, et al. BIBF 1120: triple angiokinase inhibitor with sustained receptor blockade and good antitumor efficacy. Cancer Res. 2008;68(12):4774-82. Epub 2008/06/19.

2. Reck M, Kaiser R, Eschbach C, et al. A phase II double-blind study to investigate efficacy and safety of two doses of the triple angiokinase inhibitor BIBF 1120 in patients with relapsed advanced non-small-cell lung cancer. Ann Oncol. 2011;22(6):1374-81.

3. Reck M, Kaiser R, Mellemgaard A, 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(2):143-55. Epub 2014/01/15.

4. Boehringer Ingelheim. LUME-Lung 1 CTR (1199.13) (primary PFS analysis). Data on file. ONC 14-19. 2013.

5. Boehringer Ingelheim. LUME-Lung 1 CTR (1199.13) (final OS analysis). Data on file. ONC 14-18. 2013.

6. NICE. Bevacizumab in combination with oxaliplatin and either fluorouracil plus folinic acid or capecitabine for the treatment of metastatic colorectal cancer (TA212). 2010.

7. NICE. Erlotinib monotherapy for maintenance treatment of non-small-cell lung cancer (TA227). 2011.

8. Hoyle M, Crathorne L, Peters J, et al. The clinical effectiveness and cost-effectiveness of cetuximab (mono- or combination chemotherapy), bevacizumab (combination with non-oxaliplatin chemotherapy) and panitumumab (monotherapy) for the treatment of metastatic colorectal cancer after first-line chemotherapy (review of technology appraisal No.150 and part review of technology appraisal No. 118): a systematic review and economic model. Health Technol Assess. 2013;17(14):1-237.

9. Bergers G, Benjamin LE. Tumorigenesis and the angiogenic switch. Nat Rev Cancer. 2003;3(6):401-10.

10. Folkman J. Seminars in Medicine of the Beth Israel Hospital, Boston. Clinical applications of research on angiogenesis. N Engl J Med. 1995;333(26):1757-63.

11. Ellis LM, Hicklin DJ. VEGF-targeted therapy: mechanisms of anti-tumour activity. Nat Rev Cancer. 2008;8(8):579-91.

12. Lichtenberger BM, Tan PK, Niederleithner H, et al. Autocrine VEGF signaling synergizes with EGFR in tumor cells to promote epithelial cancer development. Cell. 2010;140(2):268-79.

13. Tanno S, Ohsaki Y, Nakanishi K, et al. Human small cell lung cancer cells express functional VEGF receptors, VEGFR-2 and VEGFR-3. Lung Cancer. 2004;46(1):11-9.

14. Pietras K, Hanahan D. A multitargeted, metronomic, and maximum-tolerated dose "chemo-switch" regimen is antiangiogenic, producing objective responses and survival benefit in a mouse model of cancer. J Clin Oncol. 2005;23(5):939-52.

15. Eichhorn ME, Kleespies A, Angele MK, et al. Angiogenesis in cancer: molecular mechanisms, clinical impact. Langenbecks Arch Surg. 2007;392(3):371-9.

16. Casanovas O, Hicklin DJ, Bergers G, et al. Drug resistance by evasion of antiangiogenic targeting of VEGF signaling in late-stage pancreatic islet tumors. Cancer Cell. 2005;8(4):299-309. Epub 2005/10/18.

17. Rikova K, Guo A, Zeng Q, et al. Global survey of phosphotyrosine signaling identifies oncogenic kinases in lung cancer. Cell. 2007;131(6):1190-203. Epub 2007/12/18.

18. Kunii K, Davis L, Gorenstein J, et al. FGFR2-amplified gastric cancer cell lines require FGFR2 and Erbb3 signaling for growth and survival. Cancer Res. 2008;68(7):2340-8.

19. Bousquet G, Alexandre J, Le Tourneau C, et al. Phase I study of BIBF 1120 with docetaxel and prednisone in metastatic chemo-naive hormone-refractory prostate cancer patients. Br J Cancer. 2011;105(11):1640-5. Epub 2011/10/27.

20. Ellis PM, Kaiser R, Zhao Y, et al. Phase I open-label study of continuous treatment with BIBF 1120, a triple angiokinase inhibitor, and pemetrexed in pretreated non-small cell lung cancer patients. Clin Cancer Res. 2010;16(10):2881-9.

21. Hanna N, Kaiser R, Sullivan R, et al. LUME-Lung 2: a multicenter, randomized, double-blind, phaseIII study of nintedanib plus pemetrexed versus placebo plus pemetrexed in patients with advanced non-squamous non-small cell lung cancer after failure of first-line chemotherapy. American Society of Clinical Oncology (ASCO). Chicago, USA, May 31–June 4, 20132013.

22. Roy Castle Lung Cancer Foundation. Lung Cancer Facts and Figures, available from http://www.roycastle.org/lung-cancer/Lung-Cancer-Facts-and-Figures. 2014 [cited 2014 January]; Available from: http://www.roycastle.org/lung-cancer/Lung-Cancer-Facts-and-Figures.

23. Cancer Research UK. Lung Cancer Key Facts. 2013 [cited 2014 28th May]; Available from: http://www.cancerresearchuk.org/cancer-info/cancerstats/keyfacts/lung-cancer/.

24. ACS. NSCLC Detailed Guide 2013a: Lung Cancer (Non-Small Cell). 2013 [cited 2013 19 June]; Available from: http://www.cancer.org/cancer/lungcancer-non-smallcell/detailedguide/non-small-cell-lung-cancer-what-is-non-small-cell-lung-cancer

25. Jett JR, Schild SE, Keith RL, et al. Treatment of non-small cell lung cancer, stage IIIB: ACCP evidence-based clinical practice guidelines (2nd edition). Chest. 2007;132(3 Suppl):266S-76S. Epub 2007/10/06.

26. Socinski MA, Crowell R, Hensing TE, et al. Treatment of non-small cell lung cancer, stage IV: ACCP evidence-based clinical practice guidelines (2nd edition). Chest. 2007;132(3 Suppl):277S-89S. Epub 2007/10/06.

27. Degner LF, Sloan JA. Symptom distress in newly diagnosed ambulatory cancer patients and as a predictor of survival in lung cancer. J Pain Symptom Manage. 1995;10(6):423-31. Epub 1995/08/01.

28. Bottomley A, Efficace F, Thomas R, et al. Health-related quality of life in non-small-cell lung cancer: methodologic issues in randomized controlled trials. J Clin Oncol. 2003;21(15):2982-92. Epub 2003/07/30.

29. Maione P, Rossi A, Bareschino MA, et al. Factors driving the choice of the best second-line treatment of advanced NSCLC. Rev Recent Clin Trials. 2011;6(1):44-51. Epub 2010/09/28.

30. NHS choices. Lung cancer, available from http://www.nhs.uk/Conditions/Cancer-of-the-lung/Pages/Introduction.aspx. 2014 [cited 2014 March]; Available from: http://www.nhs.uk/Conditions/Cancer-of-the-lung/Pages/Introduction.aspx.

31. NICE. The diagnosis and treatment of lung cancer (CG121). 2011.

32. Leighl NB. Treatment paradigms for patients with metastatic non-small-cell lung cancer: first-, second-, and third-line. Curr Oncol. 2012;19(Suppl 1):S52-8. Epub 2012/07/13.

33. Stinchcombe TE, Socinski MA. Considerations for second-line therapy of non-small cell lung cancer. Oncologist. 2008;13 Suppl 1:28-36. Epub 2008/03/08.

34. NICE. Erlotinib for the treatment of non-small-cell lung cancer (TA162). 2008.

35. Boehringer Ingelheim. Cytotoxic regimens by line of treatment for the UK. Data On File. ONC 13-01. 2013.

36. Office of National Statistics. Cancer incidence and mortality in the UK 2008-10, available from http://www.ons.gov.uk/ons/rel/cancer-unit/cancer-incidence-and-mortality/2008-2010/stb-cancer-incidence-and-mortality-in-the-united-kindom--2008-2010.html. 2012.

37. Coleman MP, Forman D, Bryant H, et al. Cancer survival in Australia, Canada, Denmark, Norway, Sweden, and the UK, 1995-2007 (the International Cancer Benchmarking Partnership): an analysis of population-based cancer registry data. Lancet. 2011;377(9760):127-38.

38. NICE. Quality standard for lung cancer (QS17). 2012.

39. NICE. NICE Pathways, Lung cancer overview. 2014.

40. NICE. Erlotinib for the first-line treatment of locally advanced or metastatic EGFR-TK mutation-positive non-small-cell lung cancer (TA258). 2012.

41. NICE. Gefitinib for the first-line treatment of locally advanced or metastatic non-small-cell lung cancer (TA192). 2010.

42. NICE. Pemetrexed for the first-line treatment of non-small-cell lung cancer (TA181). 2009.

43. NICE. Afatinib for treating epidermal growth factor receptor mutation-positive locally advanced or metastatic non-small-cell lung cancer (TA310). 2014.

44. NICE. Pemetrexed for the maintenance treatment of non-small-cell lung cancer (TA190). 2010.

45. NICE. Lung cancer (non-small cell) - erlotinib & gefitinib (post chemotherapy) (rev TA162, TA175): appraisal consultation document. 2014.

46. Ciuleanu T, Stelmakh L, Cicenas S, et al. Efficacy and safety of erlotinib versus chemotherapy in second-line treatment of patients with advanced, non-small-cell lung cancer with poor prognosis (TITAN): a randomised multicentre, open-label, phase 3 study. Lancet Oncol. 2012;13(3):300-8.

47. de Boer RH, Arrieta O, Yang CH, et al. Vandetanib plus pemetrexed for the second-line treatment of advanced non-small-cell lung cancer: a randomized, double-blind phase III trial. J Clin Oncol. 2011;29(8):1067-74. Epub 2011/02/02.

48. Hanna N, Shepherd FA, Fossella FV, et al. Randomized phase III trial of pemetrexed versus docetaxel in patients with non-small-cell lung cancer previously treated with chemotherapy. J Clin Oncol. 2004;22(9):1589-97. Epub 2004/05/01.

49. Herbst RS, Sun Y, Eberhardt WEE, et al. Vandetanib plus docetaxel versus docetaxel as second-line treatment for patients with advanced non-small-cell lung cancer (ZODIAC): a double-blind, randomised, phase 3 trial. Lancet Oncol. 2010;11(7):619-26.

50. Herbst RS, Ansari R, Bustin F, et al. Efficacy of bevacizumab plus erlotinib versus erlotinib alone in advanced non-small-cell lung cancer after failure of standard first-line chemotherapy (BeTa): a double-blind, placebo-controlled, phase 3 trial. Lancet. 2011;377(9780):1846-54. Epub 2011/05/31.

51. Kim ES, Hirsh V, Mok T, et al. Gefitinib versus docetaxel in previously treated non-small-cell lung cancer (INTEREST): a randomised phase III trial. The Lancet. 2008;372(9652):1809-18.

52. Natale RB, Thongprasert S, Greco FA, et al. Phase III trial of vandetanib compared with erlotinib in patients with previously treated advanced non-small-cell lung cancer. J Clin Oncol. 2011;29(8):1059-66. Epub 2011/02/02.

53. Scagliotti G, Hanna N, Fossella F, et al. The Differential Efficacy of Pemetrexed According to NSCLC Histology: A Review of Two Phase III Studies. Oncologist. 2009;14(3):253-63.

54. Crino L, Metro G. Therapeutic options targeting angiogenesis in nonsmall cell lung cancer. Eur Respir Rev. 2014;23(131):79-91. Epub 2014/03/05.

55. Keedy VL, Sandler AB. Inhibition of angiogenesis in the treatment of non-small cell lung cancer. Cancer Sci. 2007;98(12):1825-30. Epub 2007/09/26.

56. Kreuter M, Kropff M, Fischaleck A, et al. Prognostic relevance of angiogenesis in stage III NSCLC receiving multimodality treatment. Eur Respir J. 2009;33(6):1383-8. Epub 2009/02/14.

57. Ozbudak IH, Ozbilim G, Kucukosmanoglu I, et al. Vascular endothelial growth factor expression and neovascularization in non--small cell lung carcinoma. Int J Surg Pathol. 2009;17(5):390-5. Epub 2008/10/14.

58. Reck M, Kaiser R, Mellemgaard A, et al., editors. Nintedanib (BIBF 1120) + Docetaxel in NSCLC Patients Progressing after One Prior Chemotherapy Regimen: LUME-Lung 1, a Randomized, Double-blind, Phase III trial. ASCO Annual Meeting; 2013; Chicago, Illinois. USA.

59. Novello S, Kaiser R, Mellemgaard A, et al. Analysis of Patient-Reported Outcomes from the LUME-Lung 1 Trial: A Randomized, Double-Blind, Placebo-Controlled Phase III Study in Second-Line Advanced Non-Small Cell Lung Cancer (NSCLC) Patients. The World Conference on Lung Cancer. Sydney, Australia, October 27–302013.

60. Boehringer Ingelheim. LUME-Lung 1 trial statistical plan (TSAP). Data on file. ONC 14-20. 2011.

61. Boehringer Ingelheim. LUME-Lung-1 trial statistical plan (TSAP) addendum. Data on file. ONC 14-21. 2013.

62. Kaiser R, Barrueco J, Reck M, et al., editors. Identification of a clinical biomarker for second-line combination treatment with nintedanib in adenocarcinoma non-small cell lung cancer patients in two phase III trials. European Cancer Congress 2013 (ECCO-ESMO-ESTRO); 2013; Amsterdam, The Netherlands, 27 September–1 October 2013.

63. Boehringer Ingelheim. Summary of Clinical Efficacy. Data on file. ONC 14-22. 2013.

64. Boehringer Ingelheim. LUME-Lung 1 CTR (1199.13) (final OS analysis). Appendix. Data on file. ONC 14-24. 2013.

65. NICE. Pemetrexed for the treatment of non-small-cell lung cancer (TA124). 2007.

66. Cancer Research UK. Lung Cancer Mortality Statistics. 2013 [cited 2014 28th May]; Available from: http://www.cancerresearchuk.org/cancer-info/cancerstats/types/lung/mortality/.

67. Health and Social Care Information Centre. National Lung Cancer Audit Report. Report for the audit period 2012. 2013.

68. Healthcare Quality Improvement Partnership (HQIP). Improving the Pathological Diagnosis of Lung Cancer: Initial Snapshot of Current Practice. 2012.

69. NICE. Docetaxel, paclitaxel, gemcitabine and vinorelbine for the treatment of non-small cell lung cancer (TA26). 2001.

70. Khakwani A, Rich AL, Powell HA, et al. Lung cancer survival in England: trends in non-small-cell lung cancer survival over the duration of the National Lung Cancer Audit. Br J Cancer. 2013;109(8):2058-65. Epub 2013/09/21.

71. Fossella FV, DeVore R, Kerr RN, et al. Randomized Phase III Trial of Docetaxel Versus Vinorelbine or Ifosfamide in Patients With Advanced Non–Small-Cell Lung Cancer Previously Treated With Platinum-Containing Chemotherapy Regimens. Journal of Clinical Oncology. 2000;18(12):2354-62.

72. Shepherd FA, Dancey J, Ramlau R, et al. Prospective randomized trial of docetaxel versus best supportive care in patients with non-small-cell lung cancer previously treated with platinum-based chemotherapy. J Clin Oncol. 2000;18(10):2095-103. Epub 2000/05/16.

73. Shepherd FA, Rodrigues Pereira J, Ciuleanu T, et al. Erlotinib in previously treated non-small-cell lung cancer. N Engl J Med. 2005;353(2):123-32. Epub 2005/07/15.

74. Thatcher N, Chang A, Parikh P, 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(9496):1527-37. Epub 2005/11/01.

75. Ramlau R, Gorbunova V, Ciuleanu TE, et al. Aflibercept and Docetaxel versus Docetaxel alone after platinum failure in patients with advanced or metastatic non-small-cell lung cancer: a randomized, controlled phase III trial. J Clin Oncol. 2012;30(29):3640-7. Epub 2012/09/12.

76. Garassino MC, Martelli O, Broggini M, 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(10):981-8. Epub 2013/07/26.

77. Scagliotti GV, Krzakowski M, Szczesna A, et al. Sunitinib plus erlotinib versus placebo plus erlotinib in patients with previously treated advanced non-small-cell lung cancer: a phase III trial. J Clin Oncol. 2012;30(17):2070-8. Epub 2012/05/09.

78. Krzakowski M, Ramlau R, Jassem J, et al. Phase III trial comparing vinflunine with docetaxel in second-line advanced non-small-cell lung cancer previously treated with platinum-containing chemotherapy. J Clin Oncol. 2010;28(13):2167-73. Epub 2010/03/31.

79. Ramlau R, Gervais R, Krzakowski M, et al. Phase III study comparing oral topotecan to intravenous docetaxel in patients with pretreated advanced non-small-cell lung cancer. J Clin Oncol. 2006;24(18):2800-7. Epub 2006/05/10.

80. Okano Y, Ando M, Asami K, et al. Randomized phase III trial of erlotinib (E) versus docetaxel (D) as second- or third-line therapy in patients with advanced non-small cell lung cancer (NSCLC) who have wild-type or mutant epidermal growth factor receptor (EGFR): Docetaxel and Erlotinib Lung Cancer Trial (DELTA). ASCO. 2013.

81. Ferrara N, Gerber HP, LeCouter J. The biology of VEGF and its receptors. Nat Med. 2003;9(6):669-76. Epub 2003/06/05.

82. Andrae J, Gallini R, Betsholtz C. Role of platelet-derived growth factors in physiology and medicine. Genes Dev. 2008;22(10):1276-312. Epub 2008/05/17.

83. Korc M, Friesel RE. The role of fibroblast growth factors in tumor growth. Curr Cancer Drug Targets. 2009;9(5):639-51. Epub 2009/06/11.

84. Reck M, von Pawel J, Zatloukal P, et al. Overall survival with cisplatin-gemcitabine and bevacizumab or placebo as first-line therapy for nonsquamous non-small-cell lung cancer: results from a randomised phase III trial (AVAiL). Ann Oncol. 2010;21(9):1804-9. Epub 2010/02/13.

85. Sandler A, Gray R, Perry MC, et al. Paclitaxel-carboplatin alone or with bevacizumab for non-small-cell lung cancer. N Engl J Med. 2006;355(24):2542-50. Epub 2006/12/15.

86. Kutluk Cenik B, Ostapoff KT, Gerber DE, et al. BIBF 1120 (nintedanib), a triple angiokinase inhibitor, induces hypoxia but not EMT and blocks progression of preclinical models of lung and pancreatic cancer. Mol Cancer Ther. 2013;12(6):992-1001. Epub 2013/06/05.

87. Mross K, Stefanic M, Gmehling D, et al. Phase I study of the angiogenesis inhibitor BIBF 1120 in patients with advanced solid tumors. Clin Cancer Res. 2010;16(1):311-9.

88. du Bois A, Huober J, Stopfer P, et al. A phase I open-label dose-escalation study of oral BIBF 1120 combined with standard paclitaxel and carboplatin in patients with advanced gynecological malignancies. Ann Oncol. 2010;21(2):370-5.

89. Hanna N, Ellis, P., Stopfer, P., Shapiro, D., Gyorffy, S.,. A Phase I study of continuous oral treatment with the triple angiokinase inhibitor BIBF 1120 together with pemetrexed in previously treated patients with non-small cell lung cancer: P3-091. 2007.

90. Camidge DR, Conkling, P., Stephenson, J. J., Glassman, P. M., Zhao, Y., Kaiser, R., Stopfer, P.,. Pharmacokinetic (PK) analysis of a phase I study of continuous oral treatment with the angiokinase inhibitor BIBF 1120, in combination with carboplatin and paclitaxel in patients with advanced non-small cell lung cancer (NSCLC). ASCO. 2008.

91. Lee CP, Taylor, N. J., Attard, G., Nathan, P. D., De Bono, J. S., Temple, G., Tang, A., Padhani, A. R., Judson, I. R., Rustin, G. J.,. A phase I study of BIBF 1120, an orally active triple angiokinase inhibitor (VEGFR, PDGFR, FGFR) given continuously to patients with advanced solid tumours, incorporating dynamic contrast enhanced magnetic resonance imaging (DCE-MRI). ASCO. 2006.

92. Sanofi. Taxotere 20mg/1ml concentrate for solution for infusion - EU SmPC. 2013.

93. Hotchkiss KA, Ashton AW, Mahmood R, et al. Inhibition of endothelial cell function in vitro and angiogenesis in vivo by docetaxel (Taxotere): association with impaired repositioning of the microtubule organizing center. Mol Cancer Ther. 2002;1(13):1191-200. Epub 2002/12/14.

94. Vacca A, Ribatti D, Iurlaro M, et al. Docetaxel versus paclitaxel for antiangiogenesis. J Hematother Stem Cell Res. 2002;11(1):103-18. Epub 2002/02/16.

95. NICE. EGFR-TK mutation testing in adults with locally advanced or metastatic non-small-cell lung cancer (DG9). 2013.

96. EMA. Guideline on the evaluation of anticancer medicinal products in man (EMA/CHMP/205/95/Rev.4). 2012.

97. McCain J. The Ongoing Evolution of Endpoints In Oncology. Supplement to Managed Care. 2010;19(5 (Supplement 1)).

98. Bergman B, Aaronson NK, Ahmedzai S, et al. The EORTC QLQ-LC13: a modular supplement to the EORTC Core Quality of Life Questionnaire (QLQ-C30) for use in lung cancer clinical trials. EORTC Study Group on Quality of Life. Eur J Cancer. 1994;30a(5):635-42. Epub 1994/01/01.

99. Trippoli S, Vaiani M, Lucioni C, et al. Quality of life and utility in patients with non-small cell lung cancer. Quality-of-life Study Group of the Master 2 Project in Pharmacoeconomics. Pharmacoeconomics. 2001;19(8):855-63. Epub 2001/10/13.

100. Felip E SR, Pavlidis N. . ESMO minimum clinical recommendations for diagnosis, treatment and follow-up of non-small cell lung cancer (NSCLC). Annals of Oncology. 2005;16 (Suppl 1):I28-I9.

101. Florescu M, Hasan B, Seymour L, et al. A clinical prognostic index for patients treated with erlotinib in National Cancer Institute of Canada Clinical Trials Group study BR.21. J Thorac Oncol. 2008;3(6):590-8. Epub 2008/06/04.

102. NICE. Guide to the methods of technology appraisal2013. Available from: http://publications.nice.org.uk/guide-to-the-methods-of-technology-appraisal-2013-pmg9.

103. Li N, Ou W, Yang H, et al. A randomized phase 2 trial of erlotinib versus pemetrexed as second-line therapy in the treatment of patients with advanced EGFR wild-type and EGFR FISH-positive lung adenocarcinoma. Cancer. 2014;120(9):1379-86. Epub 2014/02/01.

104. Lee JK, Hahn S, Kim DW, et al. Epidermal Growth Factor Receptor Tyrosine Kinase Inhibitors vs Conventional Chemotherapy in Non-Small Cell Lung Cancer Harboring Wild-Type Epidermal Growth Factor Receptor: A Meta-analysis. JAMA. 2014;311(14):1430-7. Epub 2014/04/10.

105. Sun JM, Lee KH, Kim SW, et al. Gefitinib versus pemetrexed as second-line treatment in patients with nonsmall cell lung cancer previously treated with platinum-based chemotherapy (KCSG-LU08-01): an open-label, phase 3 trial. Cancer. 2012;118(24):6234-42.

106. Maruyama R, Nishiwaki Y, Tamura T, 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(26):4244-52. Epub 2008/09/10.

107. Lee DH, Lee JS, Kim SW, et al. Three-arm randomised controlled phase 2 study comparing pemetrexed and erlotinib to either pemetrexed or erlotinib alone as second-line treatment for never-smokers with non-squamous non-small cell lung cancer. Eur J Cancer. 2013;49(15):3111-21.

108. Kim ST, Uhm JE, Lee J, et al. Randomized phase II study of gefitinib versus erlotinib in patients with advanced non-small cell lung cancer who failed previous chemotherapy. Lung Cancer. 2012;75(1):82-8. Epub 2011/06/21.

109. Parmar MK, Torri V, Stewart L. Extracting summary statistics to perform meta-analyses of the published literature for survival endpoints. Stat Med. 1998;17(24):2815-34. Epub 1999/01/28.

110. Boehringer Ingelheim. Risk Management Plan. Data On file. ONC 14-31. 2014.

111. Roche. Avastin (bevacizumab) Summary of Product Characteristics. 2014.

112. Bayer. Nexavar (sorafenib) Summary of Product Characteristics. 2014.

113. Bristol-Myers Squibb Co. Sprycel (dasatinib) Summary of Product Chracteristics. 2013.

114. Boehringer Ingelheim. Data On file. ONC 14-32. 2014.

115. National Information Center on Health Services Research and Health Care Technology. Module 4: An Introduction to the Principles of Critical Appraisal of Health Economic Evaluation Studies. 2003 [updated 17 January 2014; cited 2014 June]; Available from: http://www.nlm.nih.gov/nichsr/edu/healthecon/drummond_list.html.

116. McNamara S, Moore L, Ray J. PCN97 A cost-utility analysis on the use of trastuzumab + anastrozole compared to lapatinib + letrozole, letrozole monotherapy or anastrozole monotherapy in the treatment of HER2+/hormone receptor positive (Hr+) metastatic breast cancer (Mbc) from the perspective of the UK National Health Service (NHS). Value in Health. 2010;13(7):A270.

117. Boehringer Ingelheim. Health Economic Statistical Analysis Plan LUME-lung 1. Data on file. ONC 14-26. 2011.

118. Boehringer Ingelheim. Health Economic Statistical Analysis Plan LUME-lung 1 (addendum). Data on file. ONC 14-27. 2013.

119. Ishak KJ, Kreif N, Benedict A, et al. Overview of parametric survival analysis for health-economic applications. Pharmacoeconomics. 2013;31(8):663-75.

120. Sacco JJ, Botten J, Macbeth F, et al. The average body surface area of adult cancer patients in the UK: a multicentre retrospective study. PLoS One. 2010;5(1):e8933.

121. Department of Health (UK). NHS reference costs 2012 to 2013. 2013 21 November 2013. Report No.

122. Joint Formulary Committee. British National Formulary (BNF 66, September 2013): London: BMJ Group and Pharmaceutical Press 2013.

123. Joint Formulary Committee. British National Formulary (BNF 67, January 2014): London: BMJ Group and Pharmaceutical Press 2014.

124. PSSRU. Unit Costs of Health and Social Care 2012. In: Curtis L, editor. 2012.

125. Roche. Tarceva (erlotinib) tablets, oral - US PI. 2010.

126. Nafees B, Stafford M, Gavriel S, et al. Health state utilities for non small cell lung cancer. Health Qual Life Outcomes. 2008;6:84. Epub 2008/10/23.

127. Boehringer Ingelheim. Health Economic Report for LUME-Lung 1. Data on file. ONC 14-29 2014.

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

129. Schuette W, Tesch H, Buttner H, et al. Second-line treatment of stage III/IV non-small-cell lung cancer (NSCLC) with pemetrexed in routine clinical practice: evaluation of performance status and health-related quality of life. BMC Cancer. 2012;12:14. Epub 2012/01/17.

130. Blackhall FH, Hirsh V, Kim DW, et al. Impact of crizotinib on patient-reported general health status compared with single-agent chemotherapy in a phase III study of advanced ALK-positive non-small cell lung cancer (NSCLC) - Abstract presented at European Cancer Congress 2013. European Journal of Cancer. 2013;49(Supplement 2):S799-S800.

131. Brazier J, Akehurst R, Brennan A, et al. Should patients have a greater role in valuing health states? Appl Health Econ Health Policy. 2005;4(4):201-8.

132. Dolan P, Kahneman D. Interpretations Of Utility And Their Implications For The Valuation Of Health*. The Economic Journal. 2008;118(525):215-34.

133. Drummond MF, O'Brien B, Stoddart GL, et al. Methods for the Economic Evaluation of Health Care Programmes. Oxford: Oxford Univeristy Press; 1997.

134. Holmes J, Dunlop D, Hemmett L, et al. A cost-effectiveness analysis of docetaxel in the second-line treatment of non-small cell lung cancer. Pharmacoeconomics. 2004;22(9):581-9.

135. Lewis G, Peake M, Aultman R, et al. Cost-effectiveness of erlotinib versus docetaxel for second-line treatment of advanced non-small-cell lung cancer in the United Kingdom. J Int Med Res. 2010;38(1):9-21. Epub 2010/03/18.

136. NICE. Crizotinib for previously treated non-small-cell lung cancer associated with an anaplastic lymphoma kinase fusion gene (TA296). 2013.

137. NICE. Erlotinib for the treatment of non-small-cell lung cancer (TA162). ERG report. 2008.

138. Boehringer Ingelheim. Health Economic Report for LUX-Lung 1. Data on file. ONC 14-28. 2011.

139. Briggs A. Probabilistic analysis of cost-effectiveness models: statistical representation of parameter uncertainty. Value Health. 2005;8(1):1-2.

140. NICE. Single Technology appraisal, Submission of evidence by Boehringer Ingelheim on the use of afatinib in adult patients with locally advanced or metastatic non-small cell lung cancer (NSCLC) with activating Epidermal Growth Factor Receptor (EGFR) mutation(s). Available at http://www.nice.org.uk/Guidance/TA310/Evidence. 2013.

141. Office of National Statistics. National Population Projections, 2010-Based Projections. 2011.

142. Health and Social Care Information Centre. Hospital Episode Statistics, Admitted Patient Care - England, 2011-12 (available at: http://www.hscic.gov.uk/catalogue/PUB08288). 2012.

143. Health and Social Care Information Centre. National Lung Cancer Audit Report. Report for the audit period 2011. 2012.

144. Wojtowicz-Praga S, Leon L. Comparative Efficacy and Safety of Erlotinib in Non-Small Cell Lung Cancer (NSCLC) of Squamous Cell and Adenocarcinoma Histology in the Phase III NCIC CTG BR.21 and Saturn (BO18192) Trials. Abstract 1575. 37th Annual Congress of the European Society for Medical Oncology (ESMO). 28 Sep - 2 Oct; Vienna2012.

145. Boehringer Ingelheim. Data On File. ONC 13-18. 2013.

146. MIMS. Tarceva (Erlotinib). 2013.

147. Department of Health (UK). NHS reference costs: 2010-11 reference costs publication (NSRC1 NHS trust reference cost schedules). 2011.

148. Lambert PC, Sutton AJ, Burton PR, et al. How vague is vague? A simulation study of the impact of the use of vague prior distributions in MCMC using WinBUGS. Stat Med. 2005;24(15):2401-28. Epub 2005/07/15.

149. Brooks SP, Gelman A. Alternative methods for monitoring convergence of iterative simulations. J Comput Graph Stat. 1998;7(4):434-55.

150. Gelman A. Inference and monitoring convergence. In: Gilks WR, Spiegelhalter DR, editors. Markov chain Monte Carlo in practice. London: Chapman & Hall; 1996. p. 131-43.

151. Dias S, Welton N, Sutton A, et al. NICE DSU Technical Support Document 2: A Generalised Linear Modelling Framework for Pairwise and Network Meta-Analysis of Randomised Controlled Trials. 2012; Available from: http://www.nicedsu.org.uk.

152. Hirsh V, Blackhall FH. Impact of crizotinib on patient-reported symptoms and quality of life (QoL) compared with single-agent chemotherapy in a phase III study of advanced ALK + non-small cell lung cancer (NSCLC). ASCO Meeting Abstracts. 2013:Abstract No. 8108.

153. Sequist LV, Schuler M, Yamamoto N. Lux-lung 3: symptom control and health-related quality of life results from a randomized phase III study in first-line patients with advanced NSCLC harbouring EGFR mutations. ESMO. 2012:Abstract No. 1229.

154. Gebbia V, Gridelli C, Verusio C, et al. Weekly docetaxel vs. docetaxel-based combination chemotherapy as second-line treatment of advanced non-small-cell lung cancer patients. The DISTAL-2 randomized trial. Lung Cancer. 2009;63(2):251-8. Epub 2008/07/18.

155. Bucher HC, Guyatt GH, Griffith LE, et al. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. J Clin Epidemiol. 1997;50(6):683-91.

156. Boehringer Ingelheim. LUME-Lung 1 Clinical Trial Protocol. Data on file. ONC 14-30. 2008.

10 Related procedures for evidence submission

10.1 Cost-effectiveness models

NICE accepts executable economic models using standard software – that is, Excel, TreeAge Pro, R or WinBUGs. If you plan to submit a model in a non-standard package, NICE should be informed in advance. NICE, in association with the ERG, will investigate whether the requested software is acceptable, and establish if you need to provide NICE and the ERG with temporary licences for the non-standard software for the duration of the appraisal. NICE reserves the right to reject economic models in non-standard software. A fully executable electronic copy of the model must be submitted to NICE with full access to the programming code. Care should be taken to ensure that the submitted versions of the model program and the written content of the evidence submission match.

NICE will need to distribute an executable version of the model to consultees and commentators because it will be used by the Appraisal Committee to assist their decision-making. On distribution of the appraisal consultation document (ACD) or final appraisal determination (FAD), and the evaluation report produced after the first committee meeting, NICE will advise consultees and commentators by letter that the manufacturer or sponsor has developed a model as part of their evidence submission for this technology appraisal. The letter asks consultees to inform NICE if they wish to receive an electronic copy of the model. If a request is received, NICE will release the model as long as it does not contain information that was designated confidential by the model owner, or the confidential material can be redacted by the model owner without producing severe limitations on the functionality of the model. The letter to consultees indicates clearly that NICE will distribute an executable copy, that the model is protected by intellectual property rights, and can be used only for the purposes of commenting on the model's reliability and informing a response to the ACD or FAD.

Manufacturers and sponsors must ensure that all relevant material pertinent to the decision problem has been disclosed to NICE at the time of submission. There will be no subsequent opportunity to submit information unless it has been specifically requested by NICE.

When making a submission, manufacturers and sponsors should check that:

- an electronic copy of the submission has been given to NICE with all confidential information highlighted and underlined
- an executable electronic copy of the economic model has been submitted
- the checklist of confidential information (provided by NICE along with invitation to submit) has been completed and submitted.

10.2 Disclosure of information

To ensure that the appraisal process is as transparent as possible, NICE considers it highly desirable that evidence pivotal to the Appraisal Committee's decisions should be publicly available. NICE recognises that because the appraisal is being undertaken close to the time of regulatory decisions, the status of information may change during the STA process. However, at the point of issuing the FAD or ACD to consultees and commentators, all the evidence seen by the Committee should be available to all consultees and commentators.

Under exceptional circumstances, unpublished evidence is accepted under agreement of confidentiality. Such evidence includes 'commercial in confidence' information and data that are awaiting publication ('academic in confidence'). Further instructions on the specification of confidential information, and its acceptability, can be found in the agreement between the Association of the British Pharmaceutical Industry (ABPI) and NICE (www.nice.org.uk).

When data are 'commercial in confidence' or 'academic in confidence', it is the manufacturer's or sponsor's responsibility to highlight such data clearly, and to provide reasons why they are confidential and the timescale within which they will remain confidential. The checklist of confidential information should be completed: if it is not provided, NICE will assume that there is no confidential information in the submission. It is the responsibility of the manufacturer or sponsor to ensure that the confidential information checklist is kept up to date.

The manufacturer or sponsor must ensure that any confidential information in their evidence submission is clearly underlined and highlighted. NICE is assured that information marked 'academic in confidence' can be presented and discussed during

the public part of the Appraisal Committee meeting. NICE is confident that such public presentation does not affect the subsequent publication of the information, which is the prerequisite allowing for the marking of information as ‘academic in confidence’.

Please therefore underline all confidential information, and separately highlight information that is submitted under X and information submitted under X.

The manufacturer or sponsor will be asked to supply a second version of the submission with any information that is to remain confidential removed. The confidential information should be ‘blacked out’ from this version, taking care to retain the original formatting as far as possible so that it is clear which data have been removed and where from. For further details on how the document should be redacted/stripped, see the checklist of confidential information.

The last opportunity to review the confidential status of information in an STA, before publication by NICE as part of the consultation on the ACD, is 2 weeks before the Appraisal Committee meeting; particularly in terms of ‘academic in confidence’ information. The ‘stripped’ version will be issued to consultees and commentators along with the ACD or FAD, and made available on NICE’s website 5 days later.

It is the responsibility of the manufacturer or sponsor to ensure that the ‘stripped’ version of the submission does not contain any confidential information. NICE will ask manufacturers and sponsors to reconsider restrictions on the release of data if there appears to be no obvious reason for the restrictions, or if such restrictions would make it difficult or impossible for NICE to show the evidential basis for its guidance. Information that has been put into the public domain, anywhere in the world, cannot be marked as confidential.

Confidential information submitted will be made available for review by the ERG and the Appraisal Committee. Confidential information may be distributed to all consultees with the permission of the manufacturer or sponsor. NICE will at all times seek to protect the confidentiality of the information submitted, but nothing will restrict the disclosure of information by NICE that is required by law (including in particular, but without limitation, the Freedom of Information Act 2000).

The Freedom of Information Act 2000, which came into force on 1 January 2005, enables any person to obtain information from public authorities such as NICE. The Act obliges NICE to respond to requests about the recorded information it holds, and it gives people a right of access to that information. This obligation extends to submissions made to NICE. Information that is designated as ‘commercial in confidence’ may be exempt under the Act. On receipt of a request for information, NICE will make every effort to contact the designated company representative to confirm the status of any information previously deemed ‘commercial in confidence’ before making any decision on disclosure.

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

Single Technology Appraisal (STA)

Nintedanib for previously treated locally advanced or metastatic non-small cell lung cancer [ID438]

Dear [redacted]

The Evidence Review Group, Liverpool Reviews and Implementation Group (LRiG), and the technical team at NICE have now had an opportunity to take a look at the submission received on the 11th August 2014 by Boehringer Ingelheim. In general terms they felt that it is well presented and clear. However, the ERG and the NICE technical team would like further clarification relating to the clinical and cost effectiveness data.

Both the ERG and the NICE technical team will be addressing these issues in their reports. Dex 700 cnt

We request you provide a written response to this letter to the Institute by **5pm** on **18th September 2014**. Two versions of this written response should be submitted; one with academic/commercial in confidence information clearly marked and one from which this information is removed.

Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, and all information submitted under 'academic in confidence' in yellow.

If you present data that is not already referenced in the main body of your submission and that data is seen to be academic/commercial in confidence information, please complete the attached checklist for in confidence information.

Please do not 'embed' documents (i.e. PDFs, spreadsheets) within your response as this may result in your information being displaced or unreadable. Any supporting documents should be emailed to us separately as attachments or sent on a CD.

If you have any further queries on the technical issues raised in this letter then please contact Caroline Hall, Technical Lead (caroline.hall@nice.org.uk). Any procedural questions should be addressed to Jeremy Powell, Project Manager (Jeremy.Powell@nice.org.uk) in the first instance.

Yours sincerely

Nicola Hay
Technical Adviser

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

On behalf of Dr Elisabeth George
Associate Director – Appraisals
Centre for Health Technology Evaluation

Encl. checklist for in confidence information

Section A: Clarification on clinical effectiveness data

LUME-Lung 1 trial

A1. **Priority question: EGFR mutation status.** It is noted in Table 10 (company's submission, page 61) that some patients received other chemotherapy post-progression and that this treatment could be most effective in patients who are EGFR mutation-positive. Please clarify if it has been possible to test EGFR status for any of the patients with adenocarcinoma included in LUME-Lung 1. If so, please provide the breakdown by mutation status at baseline separately for both trial arms.

A2. **Priority question: Tumour response**. The clinical trial report (and Table 21, page 93 of the company's submission) reports a total of 44 patients (19 patients in the placebo treatment arm and 25 patients in the nintedanib treatment arm) with confirmed or unconfirmed tumour response in the adenocarcinoma subgroup. Please provide the following data for each responder:

a. Trial arm
b. Type of response (confirmed or unconfirmed)
c. Time to response (days)
d. Duration of response (days)
e. Time to termination of docetaxel therapy (days)
f. Time to termination of nintedanib therapy (days)
g. Time to investigator assessed disease progression (days)
h. Time to death (days)

A3. Please provide clarification for the sample size calculation for LUME-Lung 1. It is not clear from the protocol how big the sample size would need to be to observe the required number of progression-free survival (PFS) events.

A4. Pages 74 and 75 of the company's submission provide a list of covariates to include in the model for PFS "Sensitivity analysis 2". Not all of these covariates are listed (see below) in the trial protocol under "exploration of factors that might be relevant to efficacy" (protocol, pages 64 and 65). Please clarify which of the covariates listed below were specified a priori or post-hoc:

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

a. Brain metastases at baseline
b. Prior treatment with bevacizumab
c. Body- surface area
d. Age
e. Duration of first-line chemotherapy
f. Time to first progression
g. Time since first histological diagnosis
h. Presence of ipsilateral metastases in the lung at baseline
i. Presence of contralateral metastases in the lung at baseline
j. Bone metastases at baseline
k. Adrenal metastases at baseline
l. Sum of target lesions at baseline

Subgroup analyses

A5. Please clarify if the analyses for the following subgroups were specified a priori or post-hoc (company submission, page 77):

a. Presence of liver metastases (yes vs no)
b. Disease stage at diagnosis (<IIIB/IV, IIIB, IV)
c. Concomitant therapy with biphosphonates at baseline (yes vs no)
d. Presence of adrenal metastases (yes vs no)
e. Number of metastatic organs at baseline (≤2 metastatic organs, >2 metastatic organs, not centrally reviewed)
f. Lactate dehydrogenase (LDH) level at baseline (LDH ≤1, LDH >1)

A6. Table 71 (page 170 of the company's submission) presents results by receipt of first-line pemetrexed maintenance therapy. Please clarify whether this analysis was specified a priori or post-hoc.

A7. Please clarify whether there were any other a priori or post-hoc subgroup analyses carried out but not reported in the company's submission.

Mixed treatment comparisons

A6. Tables 39 and 41 of the company's submission (pages 135 and 136 respectively) present the probabilities of each treatment being the most effective. Please clarify if these are presented for the fixed or random effects models for the sensitivity analyses and present the probabilities for both fixed and random effects for these analyses.

A7. Please present the probabilities for both the fixed and random effects models for Tables 46 and 48 of the company's submission (pages 142 and 144 respectively).

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

Section B: Clarification on cost-effectiveness data

B1. **Priority question: Time to event analyses**. Using the trial data as at 15 February 2013 (that is, corresponding to the final overall survival analysis) please provide full Kaplan-Meier analysis output separately for both trial arms as follows:

a. Progression-Free Survival
- Assessment by Investigator
- Adenocarcinoma population
- Conventional censoring

b. Overall Survival
- Adenocarcinoma population
- Conventional censoring

c. Overall Survival (for sensitivity analysis)
- Adenocarcinoma population
- Patients still at risk at data cut should be censored at time of data cut (not when last known to be alive)

d. Post-Progression Survival
- Adenocarcinoma population
- Assessment by Investigator
- Conventional censoring

e. Post-Progression (for sensitivity analysis)
- Adenocarcinoma population
- Assessment by Investigator
- Patients still at risk at data cut should be censored at time of data cut (not when last known to be alive)

f. Time to Off-Treatment
- Docetaxel treatment (placebo arm)
- Adenocarcinoma population
- Conventional censoring

g. Time to Off-Treatment
- Docetaxel treatment (nintedanib arm)
- Adenocarcinoma population
- Conventional censoring

h. Time to Off-Treatment

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

- Nintedanib treatment
- Adenocarcinoma population
- Conventional censoring

Please provide the above information in a tabular form such as in the example from SAS (below) showing for each event time:

a. Time of event from baseline (days)
b. Product-limit estimate of survival proportion
c. Standard error of survival proportion
d. Number of patients failed
e. Number of patients remaining at risk

Example of output (SAS) required from specified Kaplan-Meier analyses
The LIFETEST Procedure

Product-Limit Survival Estimates						
DAYS		Survival	Failure	Survival Standard Error	Number Failed	Number Left
0.000		1.0000	0	0	0	62
1.000		.	.	.	1	61
1.000		0.9677	0.0323	0.0224	2	60
3.000		0.9516	0.0484	0.0273	3	59
7.000		0.9355	0.0645	0.0312	4	58
8.000		.	.	.	5	57
8.000		.	.	.	6	56
8.000		0.8871	0.1129	0.0402	7	55
10.000		0.8710	0.1290	0.0426	8	54
SKIP...		0.8548	0.1452	0.0447	9	53
389.000		0.1010	0.8990	0.0417	52	5
411.000		0.0808	0.9192	0.0379	53	4
467.000		0.0606	0.9394	0.0334	54	3
587.000		0.0404	0.9596	0.0277	55	2
991.000		0.0202	0.9798	0.0199	56	1
999.000		0	1.0000	0	57	0

NICE National Institute for Health and Care Excellence

Level 1A
City Tower
Manchester
M1 4BT
United Kingdom

+44 (0)845 003 7780

B2. **Priority question: Treatment with nintedanib.** Please provide details by cycle of the number of patients in the nintedanib trial arm receiving full or reduced doses, or for whom treatment was missed/suspended for any reason, tabulated as follows:

	No. of patients still 'on treatment'	No. on reduced dose (2x150 mg)	No. on reduced dose (2x100 mg)	No with treatment suspended
Cycle 1				
Cycle 2				
Cycle 3				
….etc				

N.B. Please check that the number still on treatment is equal to the sum of the figures in the other columns

B3. **Priority question: Treatment with docetaxel**. Please provide details by cycle of the number of patients separately for both trial arms receiving full or reduced doses, or for whom treatment was missed/suspended for any reason. Please tabulate as follows:

	No. of patients still 'on treatment'	No. on reduced dose (60 mg/m^2)	No with treatment suspended
Cycle 1			
Cycle 2			
Cycle 3			
….etc			

N.B. Please check that the number still on treatment is equal to sum of the figures in the other columns

B4. **Priority question: Nintedanib packaging.** Please confirm that it is intended that nintedanib will be available in two types of pack, one sufficient for 30 days treatment at full dose (200 mg twice daily), and one sufficient for 30 days treatment at reduced dose (150 mg twice daily). Please confirm that it is intended that both packs will incur the same cost.

Section C: Additional points

C1. If available, please provide a draft European Medicine Agency's European public assessment reports (EPAR) or draft Summaries of Product Characteristics (SmPC) for nintedanib.

Boehringer Ingelheim's response to clarification questions dated 04/09/2014

[redacted] information is highlighted in yellow and underlined.

Section A: Clarification on clinical effectiveness data

LUME-Lung 1 trial

A1. **Priority question: EGFR mutation status.** *It is noted in Table 10 (company's submission, page 61) that some patients received other chemotherapy post-progression and that this treatment could be most effective in patients who are EGFR mutation-positive. Please clarify if it has been possible to test EGFR status for any of the patients with adenocarcinoma included in LUME-Lung 1. If so, please provide the breakdown by mutation status at baseline separately for both trial arms.*

The availability of information on epidermal growth factor receptor (EGFR) mutational status is very limited in the LUME-Lung 1 trial. EGFR biomarker testing was not standard clinical practice at the time the trial was conducted and therefore was not mandated in the LUME-Lung 1 trial. However EGFR mutational status was collected if available. Based on the patients' oncological history, EGFR mutational status was available in 16.9% of the patients randomised in LUME-Lung 1 trial. Of those, 20 patients (1.5%) were positive for EGFR mutations: 12 in the placebo arm and 8 in the nintedanib arm [redacted]. Of these, 16 patients had adenocarcinoma histology: 11 in the placebo arm and 5 in the nintedanib arm [redacted]

A2. **Priority question: Tumour response**. *The clinical trial report (and Table 21, page 93 of the company's submission) reports a total of 44 patients (19 patients in the placebo treatment arm and 25 patients in the nintedanib treatment arm) with confirmed or unconfirmed tumour response in the adenocarcinoma subgroup. Please provide the following data for each responder:*

a. *Trial arm*
b. *Type of response (confirmed or unconfirmed)*
c. *Time to response (days)*
d. *Duration of response (days)*
e. *Time to termination of docetaxel therapy (days)*
f. *Time to termination of nintedanib therapy (days)*
g. *Time to investigator assessed disease progression (days)*
h. *Time to death (days)*

The following tables provide the data (a-h) for both the nintedanib and placebo treatment arms [redacted]

[redacted]

[redacted]

A3. Please provide clarification for the sample size calculation for LUME-Lung 1. It is not clear from the protocol how big the sample size would need to be to observe the required number of progression-free survival (PFS) events.

The number of events rather than the number of patients is the effective sample size in a time-to-event analysis. An extract from section 7.6 of the protocol provides justification of the sample size [redacted]

> It is assumed that BIBF 1120 [nintedanib] in combination with docetaxel will increase median progression free survival by approximately 28-32 % beyond combination treatment of placebo with docetaxel assuming a median PFS of docetaxel of four months in patients with an ECOG performance score of 0 and/or 1. Table 7.6: 1 indicates that including 713 PFS events would provide 90 % power, if the underlying treatment difference were 1.1 month. Seven hundred thirteen PFS events would be expected to occur within approximately 18-24 months, if patients were randomized at a rate of 45-60 patients per month. In addition at the time of the primary PFS analysis more than 400 death events are expected.

In addition, it is stated in the protocol that 1300 patients are needed to observe 1151 deaths for the key secondary endpoint, overall survival (OS) (see extract from section 7.6 of the protocol) [redacted]

> Although the sample size in this trial could provide 80 % power for OS (hazard ratio [HR] = 0.8475), it has to be noted that the magnitude and pattern of the effect of any third line or higher treatment after progression might obscure the treatment effect. To achieve 80 % power for survival, BIBF 1120 [nintedanib] would need to add 18 % to median survival (HR = 0.8475) over docetaxel monotherapy. Table 7.6: 2 indicates that 1151 deaths would provide 80 % power to detect such an increase of 18 %. This number of deaths would be expected to occur within approximately 48 months, if 1300 patients were randomized at a rate of 45 to 60 patients per month.

A4. Pages 74 and 75 of the company's submission provide a list of covariates to include in the model for PFS "Sensitivity analysis 2". Not all of these covariates are listed (see below) in the trial protocol under "exploration of factors that might be relevant to efficacy" (protocol, pages 64 and 65). Please clarify which of the covariates listed below were specified a priori or post-hoc:

a. Brain metastases at baseline
b. Prior treatment with bevacizumab
c. Body- surface area
d. Age
e. Duration of first-line chemotherapy
f. Time to first progression
g. Time since first histological diagnosis
h. Presence of ipsilateral metastases in the lung at baseline
i. Presence of contralateral metastases in the lung at baseline
j. Bone metastases at baseline
k. Adrenal metastases at baseline
l. Sum of target lesions at baseline

The following list gives the overview which of the covariates listed below were specified a priori or post-hoc for "Sensitivity analysis 2" of the primary PFS endpoint for the LUME-Lung 1 trial:

a. Brain metastases at baseline: predefined strata and also for this analysis in interim TSAP LUME-Lung 1 before unblinding of primary PFS data (a priori).

b. Prior treatment with bevacizumab: predefined strata and also for this analysis in interim TSAP LUME-Lung 1 (a priori).

c. Body- surface area: (post-hoc).

d. Age: Predefined in interim TSAP LUME-Lung 1 in subgroup section (a priori).

e. Duration of first-line chemotherapy: (post-hoc).

f. Time to first progression: specified in TSAP amendment (post-hoc).

g. Time since first histological diagnosis: (post-hoc).

h. Presence of ipsilateral metastases in the lung at baseline: (post-hoc).

i. Presence of contralateral metastases in the lung at baseline: (post-hoc).

j. Bone metastases at baseline: (post-hoc).

k. Adrenal metastases at baseline: specified in TSAP amendment (post-hoc).

l. Sum of target lesions at baseline: predefined in interim TSAP LUME-Lung 1 (a priori).

Subgroup analyses

A5. Please clarify if the analyses for the following subgroups were specified a priori or post-hoc (company submission, page 77):

a. Presence of liver metastases (yes vs no)
b. Disease stage at diagnosis (<IIIB/IV, IIIB, IV)
c. Concomitant therapy with biphosphonates at baseline (yes vs no)
d. Presence of adrenal metastases (yes vs no)
e. Number of metastatic organs at baseline (≤2 metastatic organs, >2 metastatic organs, not centrally reviewed)
f. Lactate dehydrogenase (LDH) level at baseline (LDH ≤1, LDH >1)

In section 7.3.1.2 of the clinical trial protocol (CTP) the following variables were pre-specified [redacted]

- Presence of liver metastases (yes vs. no)
- Disease stage at diagnosis (<IIIB/IV vs. IIIB vs. IV)
- Concomitant therapy with biphosphonates at baseline

In the TSAP amendment to formalize hypothesis confirmation and validation in Lume-Lung 1 and Lume Lung 2 the following variables were pre-specified for the final OS analysis of the LUME-Lung 1 [redacted]:

- Presence of adrenal metastases (yes vs. no)
- Number of metastatic organs at baseline
- Lactate dehydrogenase (LDH) level at baseline

A6. Table 71 (page 170 of the company's submission) presents results by receipt of first-line pemetrexed maintenance therapy. Please clarify whether this analysis was specified a priori or post-hoc.

This was a post-hoc analysis requested by the EMA.

A7. Please clarify whether there were any other a priori or post-hoc subgroup analyses carried out but not reported in the company's submission.

There were no additional a priori subgroups and no new endpoints however additional post hoc analyses of the following endpoints have been carried out:

- Objective response and disease control rate
- Responder analysis of health-related quality of life (HRQoL)
- Time to deterioration of HRQoL
- Time to fist onset of adverse events

In addition, analyses of OS and PFS by Region (Europe vs. Asia vs. South Africa) were performed.

Mixed treatment comparisons

A6. Tables 39 and 41 of the company's submission (pages 135 and 136 respectively) present the probabilities of each treatment being the most effective. Please clarify if these are presented for the fixed or random effects models for the sensitivity analyses and present the probabilities for both fixed and random effects for these analyses.

In the submission, Table 39 and Table 41 presented the results using a random-effects model for OS and PFS respectively. A comparison of the random-effects model results and fixed-effects model results for the probability of each treatment being the most effective in terms of improving OS is presented in Table 39a. It can be seen in both the random and fixed effects models that nintedanib plus docetaxel has the greatest probability improving OS the most. Sensitivity analyses demonstrate that the models are sensitive to the inclusion of trials selecting patients with EGFR mutations, however nintedanib plus docetaxel continues to have the greatest probability of improving OS the most.

Table 39a: Comparison of the probability of each treatment being the most effective in terms of improving OS using the random-effects and fixed-effects models.

	Random-effects model		Fixed-effects model	
	Base-case analysis	**Sensitivity analysis***	**Base-case analysis**	**Sensitivity analysis***
Nintedanib + docetaxel	70.44%	49.2%	[redacted]	[redacted]
Docetaxel	9.81%	5.62%	[redacted]	[redacted]
Pemetrexed	16.42%	0.60%	[redacted]	[redacted]
Erlotinib	3.33%	4.69%	[redacted]	[redacted]
Erlotinib + pemetrexed	----	37.17%	[redacted]	[redacted]
Gefitinib	----	2.72%	[redacted]	[redacted]

*Adding trials selecting patients with EGFR mutations.

A comparison of the random and fixed effects model results of the probability of each treatment being the most effective in terms of improving PFS is presented in Table 41a. Both the random and fixed effects models demonstrate that nintedanib plus docetaxel has the greatest probability of prolonging PFS the most in the base-case.

Table 41a: Comparison of the probability of each treatment being the most effective in terms of improving PFS using the random-effects and fixed-effects models.

	Random-effects model		Fixed-effects model	
	Base-case analysis	Sensitivity analysis*	Base-case analysis	Sensitivity analysis*
Nintedanib + docetaxel	69.69%	25.01%	[redacted]	[redacted]
Docetaxel	5.01%	0.41%	[redacted]	[redacted]
Pemetrexed	18.53%	0.09%	[redacted]	[redacted]
Erlotinib	6.77%	0.35%	[redacted]	[redacted]
Erlotinib + pemetrexed	----	61.99%	[redacted]	[redacted]
Gefitinib	----	12.15%	[redacted]	[redacted]

*Adding trials selecting patients with EGFR mutations.

A7. Please present the probabilities for both the fixed and random effects models for Tables 46 and 48 of the company's submission (pages 142 and 144 respectively).

Table 46 and Table 48 in the Boehringer Ingelheim's submission presented results of scenario analysis in which docetaxel and pemetrexed were assumed to be equivalent. A comparison of the random and fixed effects model results of the probability of each treatment improving OS the most in the scenario analyses is provided in Table 46a. Nintedanib plus docetaxel has the greatest probability of being the most effective in terms of improving OS using both fixed-effects and random-effects models.

Table 46a: Comparison of the probability of each treatment being the most effective in terms of improving OS using the random-effects and fixed-effects models.

	Random-effects model		Fixed-effects model	
	Scenario analysis	Sensitivity analysis*	Scenario analysis	Sensitivity analysis*
Nintedanib + docetaxel	78.95%	34.21%	[redacted]	[redacted]
Docetaxel/pemetrexed	13.65%	1.20%	[redacted]	[redacted]
Erlotinib	7.40%	6.79%	[redacted]	[redacted]
Gefitinib	----	3.40%	[redacted]	[redacted]
Erlotinib + pemetrexed	----	54.39%	[redacted]	[redacted]

*Adding trials selecting patients with EGFR mutations.

Scenario analyses (docetaxel and pemetrexed equivalence) from the random-effects and fixed-effects models assessing PFS are presented in Table 48a. As previously demonstrated in the base-case analyses using both random-effects and fixed-effects models, nintedanib plus docetaxel has the greatest probability of improving PFS.

Table 48a: Comparison of scenario analyses of the probability of each treatment being the most effective in terms of improving PFS using the random-effects and fixed-effects models.

	Random-effects model		Fixed-effects model	
	Scenario analysis	**Sensitivity analysis***	**Scenario analysis**	**Sensitivity analysis***
Nintedanib + docetaxel	83.57%	16.42%	[redacted]	[redacted]
Docetaxel/pemetrexed	8.75%	0.04%	[redacted]	[redacted]
Erlotinib	7.67%	0.30%	[redacted]	[redacted]
Gefitinib	----	10.99%	[redacted]	[redacted]
Erlotinib + pemetrexed	---	72.23%	[redacted]	[redacted]

*Adding trials selecting patients with EGFR mutations.

Section B: Clarification on cost-effectiveness data

B1. **Priority question: Time to event analyses**. *Using the trial data as at 15 February 2013 (that is, corresponding to the final overall survival analysis) please provide full Kaplan-Meier analysis output separately for both trial arms as follows:*

a. Progression-Free Survival
- *Assessment by Investigator*
- *Adenocarcinoma population*
- *Conventional censoring*

b. Overall Survival
- *Adenocarcinoma population*
- *Conventional censoring*

c. Overall Survival (for sensitivity analysis)
- *Adenocarcinoma population*
- *Patients still at risk at data cut should be censored at time of data cut (not when last known to be alive)*

d. Post-Progression Survival
- *Adenocarcinoma population*
- *Assessment by Investigator*
- *Conventional censoring*

e. Post-Progression (for sensitivity analysis)
- *Adenocarcinoma population*
- *Assessment by Investigator*
- *Patients still at risk at data cut should be censored at time of data cut (not when last known to be alive)*

f. Time to Off-Treatment

- *Docetaxel treatment (placebo arm)*
- *Adenocarcinoma population*
- *Conventional censoring*

g. Time to Off-Treatment

- *Docetaxel treatment (nintedanib arm)*
- *Adenocarcinoma population*
- *Conventional censoring*

h. Time to Off-Treatment

- *Nintedanib treatment*
- *Adenocarcinoma population*
- *Conventional censoring*

Please provide the above information in a tabular form such as in the example from SAS (below) showing for each event time:

a. Time of event from baseline (days)
b. Product-limit estimate of survival proportion
c. Standard error of survival proportion
d. Number of patients failed
e. Number of patients remaining at risk

Example of output (SAS) required from specified Kaplan-Meier analyses
The LIFETEST Procedure

Product-Limit Survival Estimates						
DAYS		**Survival**	**Failure**	**Survival Standard Error**	**Number Failed**	**Number Left**
0.000		1.0000	0	0	0	62
1.000		.	.	.	1	61
1.000		0.9677	0.0323	0.0224	2	60
3.000		0.9516	0.0484	0.0273	3	59
7.000		0.9355	0.0645	0.0312	4	58
8.000		.	.	.	5	57
8.000		.	.	.	6	56
8.000		0.8871	0.1129	0.0402	7	55
10.000		0.8710	0.1290	0.0426	8	54
SKIP...		0.8548	0.1452	0.0447	9	53
389.000		0.1010	0.8990	0.0417	52	5
411.000		0.0808	0.9192	0.0379	53	4
467.000		0.0606	0.9394	0.0334	54	3
587.000		0.0404	0.9596	0.0277	55	2

Product-Limit Survival Estimates						
DAYS		**Survival**	**Failure**	**Survival Standard Error**	**Number Failed**	**Number Left**
991.000		0.0202	0.9798	0.0199	56	1
999.000		0	1.0000	0	57	0

With regard to document, [redacted], please see refer to the following "Statdoc" sections:

Note: the LUME-Lung 1 trial is referred to as 1199.13.

a. **Progression-Free Survival:** [redacted]
b. **Overall Survival:** [redacted]
c. **Overall Survival (for sensitivity analysis):** [redacted]
d. **Post-Progression Survival:** [redacted]
e. **Post-Progression (for sensitivity analysis):** [redacted]
f. **Time to Off-Treatment:** [redacted]
g. **Time to Off-Treatment:** [redacted]
h. **Time to Off-Treatment:** [redacted]

B2. **Priority question: Treatment with nintedanib.** *Please provide details by cycle of the number of patients in the nintedanib trial arm receiving full or reduced doses, or for whom treatment was missed/suspended for any reason, tabulated as follows:*

	No. of patients still 'on treatment'	*No. on reduced dose (2x150 mg)*	*No. on reduced dose (2x100 mg)*	*No with treatment suspended*
Cycle 1				
Cycle 2				
Cycle 3				
….etc				

N.B. Please check that the number still on treatment is equal to the sum of the figures in the other columns

Please see [redacted]

B3. **Priority question: Treatment with docetaxel**. *Please provide details by cycle of the number of patients separately for both trial arms receiving full or reduced doses, or*

for whom treatment was missed/suspended for any reason. Please tabulate as follows:

	No. of patients still 'on treatment'	*No. on reduced dose (60 mg/m^2)*	*No with treatment suspended*
Cycle 1			
Cycle 2			
Cycle 3			
….etc			

N.B. Please check that the number still on treatment is equal to sum of the figures in the other columns

Please see [redacted]

B4. **Priority question: Nintedanib packaging.** *Please confirm that it is intended that nintedanib will be available in two types of pack, one sufficient for 30 days treatment at full dose (200 mg twice daily), and one sufficient for 30 days treatment at reduced dose (150 mg twice daily). Please confirm that it is intended that both packs will incur the same cost.*

Vargatef will be available in two pack sizes; one contains 120 x 100mg and the other contains 60 x 150mg. The 120 x 100mg pack will allow 2 x 100mg twice daily for 30 days. The 60 x 150mg pack will allow 1 x 150mg twice daily for 30 days. It is intended that both packs will incur the same cost.

Section C: Additional points

C1. *If available, please provide a draft European Medicine Agency's European public assessment reports (EPAR) or draft Summaries of Product Characteristics (SmPC) for nintedanib.*

The draft EPAR is not yet available, however it will be sent to NICE when it becomes available. The draft SmPC is included with this response.

Submission from **Roy Castle Lung Cancer Foundation,** for consideration by NICE, in their review of Nintedanib in the treatment of previously treated locally advanced or metastatic Non Small Cell Lung Cancer (NSCLC), [ID438].

Submitting Organisation

Roy Castle Lung Cancer Foundation is a UK wide lung cancer charity. We fund lung cancer research, tobacco control initiatives and work in lung cancer patient care (information, support and advocacy activity).

The Foundation has contact with patients/carers through its UK wide network of over 50 monthly Lung Cancer Patient Support Groups, online Forums and its Lung Cancer Information Helpline.

Clearly, our patient group members and contacts are a self-selected group, who have taken the step to seek out information or have accessed specialist support services. As most lung cancer sufferers tend to be older, from lower social class groups and with the five year survival being less than 10%, less physically well, we acknowledge that our patients are perhaps not representative of the vast majority of lung cancer patients, who are not so well informed. It is, however, important that the opinions expressed to us, be passed on to NICE, as it considers the place of this product in the management of non small cell lung cancer (NSCLC).

General Points

1. For patients with advanced or metastatic NSCLC, in this second line setting, cure is not a treatment option. Only two second line therapy options are currently NICE approved – Docetaxel and Erlotinib (note, these are currently undergoing a NICE MTA, so, this may change). In this scenario, improving quality of life and even small extensions in duration of life are of considerable significance to the individual and their family.

2. As active treatment options are limited in second line NSCLC and as overall outcomes remain poor, the availability of new choices, offer 'hope' for patients

3. The issue of "inverse weighting for duration of life" must be stressed. When considering the cost of treatment, it is not appropriate, for example, to give the same weighting to the final six months of life as to all other six months of life. It is important for this to be part of any numeric equation, which is looking at cost and quality of life. This point is of crucial importance to patients and relatives in this situation

4. Improvement in symptoms. Patients with advanced or metastatic non small cell lung cancer are often debilitated with multiple and distressing symptoms. Symptoms such as breathlessness are very difficult to manage clinically. Therapies with anti-tumour activity often provide the best option for symptom relief.

5. The potential of improving quality of life brings obvious benefits. These patients, in general, have quite limited life expectancy. It is of paramount importance, both to them and their families, that they are able to function as fully as is possible, for as long as possible.

This Product

1. Oral Preparation. So, it is easily administered.

2. Side effect profile
   In the anecdotal patient experience reported to us, patients report side effects associated with Docetaxel. The addition of Nintedanib seems therefore, to be well tolerated.

3. Nintedanib is a triple angiokinase inhibitor. So, in NSCLC, it represents a new and innovative therapy.

4. As noted above, for this patient group, prognosis is very poor. Thus, even relatively small benefits of extension to life can be disproportionately large for these patients.

Our observations come from a combination of one-to-one discussion with lung cancer patients, published research and our patient information helpline.

In summary

Patients with advanced and metastatic lung cancer, who have progressed after first line therapy, are in a particularly devastating situation. At present, only two NICE recommended anti-cancer options are available (Docetaxel and Erlotinib). Nintedanib presents a new opportunity.

[redacted]**, RCLCF.**

July 2014.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

Nintedanib for previously treated locally advanced or metastatic non-small cell lung cancer

Thank you for agreeing to give us a statement on your organisation's view of the technology and the way it should be used in the NHS.

Healthcare professionals can provide a unique perspective on the technology within the context of current clinical practice which is not typically available from the published literature.

To help you in making your statement, we have provided a template. The questions are there as prompts to guide you. It is not essential that you answer all of them.

Please do not exceed the 8-page limit.

About you

Your name: Dr Thomas Newsom-Davis

Name of your organisation: Chelsea and Westminster Healthcare NHS Fdn Trust

Are you (tick all that apply):

- ✓ a specialist in the treatment of people with the condition for which NICE is considering this technology?
- \- a specialist in the clinical evidence base that is to support the technology (e.g. involved in clinical trials for the technology)?
- \- an employee of a healthcare professional organisation that represents clinicians treating the condition for which NICE is considering the technology? If so, what is your position in the organisation where appropriate (e.g. policy officer, trustee, member etc.)?
- \- other? (please specify)

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

What is the expected place of the technology in current practice?

Lung cancer is the common cause of cancer death in UK, and the majority of lung cancer patients are diagnosed with advanced stage disease. There are several options for 1st line chemotherapy for advanced non-small cell lung cancer (NSCLC), and treatment has been shown to improve both quality of life and survival. Treatment is palliative, however, with all patients relapsing in the months following treatment. The options for 2nd line chemotherapy are more limited and the prognosis in this patient group is poor, with a median overall survival of around 6 months in the clinical trial populations. Despite numerous and varied attempts, the use of novel agents and/or additional of further chemotherapy agents have, until now, failed to significantly improve survival.

Nintedanib has not yet been licensed for use by either the FDA or the EMA. It is not currently available outside the Boehringer Ingelheim Patient Access Scheme. The principal clinical trial data on Nintedanib comes from the LUME-LUNG-1 trial which was reported at the World Conference for Lung Cancer (2013) and published in Lancet Oncology in 2014. This is relates to its use in patients with advanced NSCLC, whose disease has progressed after 1st line chemotherapy. As such, the details below are restricted to this patient population.

Currently there are three drugs licensed and available to treat this patient group: pemetrexed, erlotinib and pemetrexed. Pemetrexed is most commonly used as a first line agent and so is less commonly employed as a second-line agent. Consequently docetaxel and erlotinib are the main two treatment options and until recently were seen as largely equivalent in efficacy although erlotinib has a more favourable side effect profile. However recent data has raised questions about the effectiveness of erlotinib, leading many oncologists to view docetaxel as the only practical choice for patients with advance lung cancer in the 2nd line setting. Although there is some individual and geographical variation in practice, the above views are reasonably representative of the United Kingdom.

This patient group is characterised by older age, lower socio-economic status, greater medical co-morbidities and poor performance status. This is one explanation why only 50% of patients who receive 1st line chemotherapy for advanced lung cancer go on to receive 2nd line treatment. As would be expected, younger patient, those with good performance status, a longer treatment free interval since 1st line treatment, and a histological diagnosis of adenocarcinoma (as opposed to squamous histology) have a more favourable prognosis.

The LUME-LUNG-1 study investigated the use of nintedanib in combination with docetaxel in the 2nd line treatment of NSCLC and demonstrated a statistically significant improvement in the primary end point (progression free survival, PFS) in the whole patient population. Sub-group analysis found that those who derived greatest benefit were patients with an adenocarcinoma and those whose cancer had relapsed within 9 months of their 1st line chemotherapy. Consequently much of the subsequent interest in nintedanib has focussed on these groups.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

All use of nintedanib would be through established secondary care oncology units and centres. It would be prescribed and coordinated through the oncology department. Since the patient would already be receiving docetaxel chemotherapy, and reflecting that nintedanib is an oral medication, no additional facilities or staff would be anticipated for its use.

Clinical Guidelines
There are no clinical guidelines on nintedanib, as it is not yet licensed.

The advantages and disadvantages of the technology

Advantages:
1. The addition of Nintedanib has demonstrated a statistically significant improvement in PFS and overall survival (OS) in NSCLC patients who are receiving 2nd line docetaxel chemotherapy.
2. There are patient sub-groups which appear to derive greater benefit from nintedanib. These include those with an adenocarcinoma (median OS = 10.3 vs. 12.6 months) and adenocarcinoma patients who started nintedanib/docetaxel within 9 months of their 1st line chemotherapy (median OS = 7.9 vs. 10.9 months). The latter represent a patient group with aggressive disease and an otherwise poor prognosis.
3. Histology was a pre-specified subgroup for analysis, whilst the pattern of OS analysis, including stepwise analysis according to adenocarcinoma patients who were treated within 9 months of starting 1st line chemotherapy, was a pre-specified secondary endpoint.
4. Nintedanib is used as an addition to existing, proven chemotherapy which is the standard of care across the United Kingdom.
5. The side effect profile of following the addition of nintedanib is acceptable, with the commonest grade 3 toxicities being diarrhoea and elevated trans-aminases
6. Many of the toxicities previously noted with other anti-angiogenic agents (for example hypertension, proteinuria, haemoptysis, thrombosis) were either not noted, or were mild and reversible.
7. Nintedanib is easy to use as it is an oral medication. There is little additional work for chemotherapy units, specialist nurses or doctors, and no additional burden on 'chemotherapy chair time'.
8. The use of nintedanib will not affect first line chemotherapy choices (since the only other available anti-angiogenic, bevacizumab, is not available for use in the NHS) nor is it likely to impact on subsequent chemotherapy options.
9. There are no additional tests, biomarkers or biopsies required prior to starting nintedanib.

Disadvantages:
1. The OS benefit in the whole patient cohort from LUME-LUNG-1 was statistically significant, but probably not clinically significant (2.7 vs. 3.4 months).
2. Even in the subgroups where particular benefit for nintedanib was noted, the overall survival benefit remains modest at around 3 months. Consequently the clinical benefits cannot be described as dramatic.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

3. Nintedanib carries its own side effect profile. More patient on nintedanib and docetaxel died from treatment related side effects, as compared to those on docetaxel alone (35 vs. 25 patients). One of the major causes of death was sepsis, which is one of the commonest side effects associated with docetaxel, reflecting its myelosuppressive activity.
4. Additional side effects in this patient group are especially unwelcome.
5. There appears to be minimal activity in patients with squamous cell carcinoma, and small cell lung cancer patients were not included in the study. Consequently this is not a treatment that will benefit all lung cancer patients.
6. There is a lack of comprehensive Quality of Life data available to date. This is essential since treatment is being given with palliative intent and maintenance or improvement of Quality of Life is one of the main reasons to initiate a treatment.
7. The clinical trial data remains limited to one Phase 3 registration trial (LUME-LUNG-1). The data from this was analysed after a median follow up of just 7.1 months. The great majority of patients were recruited from Europe and so the role of nintedanib in other patient groups is not known.
8. The use of nintedanib in combination with other chemotherapies is not known. For example, the trial assessing nintedanib with pemetrexed as 2nd line treatment (LUME-LUNG-2 trial), has not yet published its results.

Given the limited use of nintedanib, formal rules on the use of nintedanib have not been developed. It is advisable to start the nintedanib at the same time as the docetaxel, although my own experience demonstrates that it is possible to start it on the 2nd cycle of docetaxel. Stopping nintedanib is straightforward and requires no additional measures.

Although I have used nintedanib, the patient numbers involved were too small to make a conclusion whether this clinical experience matches the findings from the clinical trial. However the circumstances in which the trials were conducted (docetaxel used as a second line treatment after progression of cancer following 1st line chemotherapy) does reflect current UK practice. One remaining issue is whether nintedanib is similarly effective in patients who carry an identifiable driver mutation (such as EGFR or ALK), who have progressed following targeted therapy and the 1st line chemotherapy.

The most important outcomes in lung cancer are OS, PFS and Quality of Life. Additional outcomes such as response rate are also important. Consequently the end-points of LUME-LUNG-1, were appropriate although some might argue that a 2nd line NSCLC study should use OS as its primary outcome. Others would claim that PFS is a reasonable surrogate endpoint, although this is not accepted by all oncologists.

I am not aware of any adverse effects that were not apparent in clinical trials but have come to light subsequently during use through the Patient Access Scheme.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

Equality and Diversity

I cannot see any situations in which this appraisal could exclude people protected by the equality legislation or could lead to recommendations that have a different impact on people protected by the equality legislation.
I base this opinion on the fact that if the drug were made available, it would be used on the basis of its proven clinical activity as demonstrated by the clinical trial data.

Implementation issues

I cannot foresee and widespread issues with implementation. No additional resources such as facilities or equipment would be needed, and all education and training of staff would be achievable in the timeframes stipulated.

STA REPORT

LIVERPOOL REVIEWS AND IMPLEMENTATION GROUP (LRiG)

Nintedanib for previously treated locally advanced or metastatic non-small cell lung cancer

This report was commissioned by the NIHR HTA Programme as project number 13/106/01

Completed 15 October 2014

Contains AIC and CIC

UNIVERSITY OF LIVERPOOL

LIVERPOOL REVIEWS AND IMPLEMENTATION GROUP

A MEMBER OF THE RUSSELL GROUP

Title: Nintedanib for previously treated locally advanced or metastatic non-small cell lung cancer

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

Authors: Nigel Fleeman, Research Fellow (Clinical Effectiveness), LR*i*G, University of Liverpool

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

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

Kerry Dwan, Research Fellow (Medical Statistician), LR*i*G, University of Liverpool

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

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

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

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

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

Correspondence to: Nigel Fleeman, Research Fellow, Liverpool Reviews and Implementation Group, University of Liverpool, Room 2.10, Whelan Building, The Quadrangle, Brownlow Hill, Liverpool L69 3GB

Date completed: 15 October 2014

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 13/106/01

Declared competing interests of the authors: None

Acknowledgements: The authors would like to thank Paula Scullin (Consultant Oncologist, Belfast HSC Trust) and Julian Nam (Health Economist, University of Glasgow) who provided feedback on draft versions of the report.

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

This report should be referenced as follows: Fleeman N, Bagust A, Boland A, Dwan K, Richardson M, Beale S, Kotas E, Banks L, Greystoke A. Nintedanib for previously treated locally advanced or metastatic non-small cell lung cancer: A Single Technology Appraisal. LR*i*G, University of Liverpool, 2014

Contributions of authors:

Fleeman N	Project lead, drafted clinical results section and supervised the final report
Bagust A	Checking and validation of the economic model and critique
Boland A	Critical appraisal of clinical and economic evidence
Dwan K	Critical appraisal of the statistical evidence
Richardson M	Critical appraisal of the statistical evidence
Beale S	Critical appraisal of clinical and economic evidence
Kotas E	Cross checking of the submission search strategy
Banks L	Critical appraisal of the submission
Greystoke A	Clinical advice and critical appraisal of the clinical sections of the manufacturer's submission

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

Table of contents

List of tables

List of figures

LIST OF ABBREVIATIONS

AE	adverse event
AESI	adverse events of special interest
AIC	Akaike information criteria
ALT	alanine aminotransferase
AST	aspartate transaminase
AUC	area under the curve
BSA	body surface area
BSC	best supportive care
BD	twice daily
CEAC	cost-effectiveness acceptability curve
CI	confidence interval
CS	company's submission
CTCAE	Common Terminology Criteria for Adverse Events
CTR	clinical trial report
D	Death
ECOG	Eastern Cooperative Oncology Group
EGFR	epidermal growth factor receptor
EMA	European Medicines Agency
ERG	Evidence Review Group
HR	hazard ratio
HRQoL	health related quality of life
ICER	incremental cost-effectiveness ratio
ITT	intention-to-treat
K-M	Kaplan-Meier
LUCADA	National Lung Cancer Audit database
MTC	mixed treatment comparison
NSCLC	non-small-cell lung cancer
ORR	overall response rate
OS	overall survival
PD	progressive disease
PF	progression-free
PFS	progression-free survival
PPS	post-progression survival
PS	performance status
PSA	probability sensitivity analysis
QALY	quality adjusted life year
QoL	quality of life
RCT	randomised controlled trial
RDI	Relative Dose Intensity
RECIST	Response Evaluation in Solid Tumours
SAEs	serious adverse events
SEER	Surveillance, Epidemiology and End Results
SmPC	summary of product characteristics
TKI	tyrosine-kinase inhibitor
ToT	time on treatment
TSAP	trial statistical analysis plan
TTD	time to deterioration
VEGF	vascular endothelial growth factor
VEGFR	vascular endothelial growth factor receptors

1 SUMMARY

1.1 Scope of the submission

The remit of the Evidence Review Group (ERG) is to comment on the clinical and cost-effectiveness evidence submitted to the National Institute for Health and Care Excellence (NICE) as part of the single technology appraisal (STA) process. Clinical and economic evidence have been submitted to NICE from Boehringer Ingelheim in support of the use of nintedanib (Vargatef) for previously treated locally advanced or metastatic non-small cell lung cancer (NSCLC) of adult patients with adenocarcinoma tumour histology.

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

The population specified in the scope is adults with locally advanced or metastatic NSCLC that has progressed following prior chemotherapy. The decision problem addressed by the company is patients with locally advanced, metastatic or locally recurrent NSCLC of adenocarcinoma tumour histology who had previously received first-line chemotherapy. This is in line with the anticipated full marketing authorisation for nintedanib which differs to that of the scope by including the term "locally recurrent" and restricting NSCLC to adenocarcinoma. The ERG notes that to be classified as locally recurrent, a patient would initially present with early stage disease (stage I, II or IIIA) and be treated with surgery or radical radiotherapy and then relapse in the same area without metastases. Since the anticipated license also stipulates patients must have previously received first-line chemotherapy, then all patients would have locally advanced or metastatic disease at the time of second-line treatment. Treatment for locally advanced (be it recurrent or present since diagnosis) or metastatic disease at this point in the disease course is identical.

The anticipated license also specifies that nintedanib should be administered in combination with docetaxel. Both docetaxel monotherapy and erlotinib monotherapy are considered as comparators in the company's submission (CS). However the company states that erlotinib is not a relevant comparator to nintedanib plus docetaxel and this is only considered a comparator by the company for secondary analyses. The ERG agrees with the company that erlotinib is not a relevant comparator. A preliminary recommendation by NICE in February and August 2014 is that erlotinib should not be recommended for treating locally advanced or metastatic NSCLC that has progressed after non-targeted chemotherapy in people with tumours that are EGFR-negative. Furthermore, in clinical practice the ERG notes that the majority of patients with EGFR-positive disease will already have received erlotinib (or another tyrosine-kinase inhibitor [TKI]) as first-line treatment so would not receive erlotinib as a second-line treatment. Finally, patients who would be considered fit enough (i.e. Eastern Cooperative Oncology Group [ECOG] performance status [PS] 0 to 1) to receive nintedanib

would also need to be considered fit enough to receive docetaxel (since docetaxel is administered in combination with nintedanib). Hence only docetaxel is considered to be a relevant comparator by the ERG.

Clinical evidence is presented for all outcomes specified in the scope and cost-effectiveness results are expressed in terms of incremental cost per quality adjusted life year (QALY) gained. No subgroups were specified in the decision problem and no equality issues were identified.

1.3 Summary of clinical effectiveness evidence submitted by the company

Direct evidence is presented for nintedanib plus docetaxel vs placebo plus docetaxel from one phase III double-blind randomised controlled trial (RCT) (LUME-Lung 1). The company states that as not all patients in LUME-Lung 1 had histology of adenocarcinoma but that as patients who did not have adenocarcinoma are expected to be outside the licensed population only data for patients with adenocarcinoma are presented. While some of these patients had locally recurrent, as opposed to locally advanced or metastatic disease at diagnosis, the vast majority (94.2%) had metastatic disease at screening.

The findings from LUME-Lung 1 suggested that nintedanib plus docetaxel significantly improve progression-free survival (PFS) and overall survival (OS) in comparison to placebo plus docetaxel. The gain in median PFS is 1.2 months (4.0 months vs 2.8 months; hazard ratio [HR] 0.77, 95% confidence interval [CI]: 0.62 to 0.96) based on the primary analysis with a median follow-up of 7.1 months. Based on the final analysis, after a median follow-up of 31.7 months, the gain in PFS is 1.4 months (4.2 months vs 2.8 months; HR 0.84, 95% CI: 0.71 to 1.00). The gain in median OS is 2.3 months (12.6 months vs 10.3 months; HR 0.83, 95% CI: 0.70 to 0.99). Pre-specified and post-hoc subgroup analyses for both PFS and OS support the findings for the population of patients with adenocarcinoma as a whole.

Specific adverse events (AEs) occurring more often in the nintedanib plus docetaxel arm than in the placebo plus docetaxel arm and considered to be AEs of special interest (AESIs) were diarrhoea (43.4% vs 24.6%), nausea (28.4% vs 17.7%) and vomiting (19.4% vs 12.3%). These AEs were successfully managed by dose reduction, dose interruption and/or symptomatic treatment and led to permanent nintedanib discontinuation in <1% of patients. Other reported AESIs associated with nintedanib treatment included increases in alanine aminotransferase (ALT) (37.8% vs 9.3%) and aspartate transaminase (AST) (30.3% vs 7.2%). These were reported to be generally reversible and led to permanent nintedanib discontinuation in <2% of patients. The incidence of Common Terminology Criteria for

Adverse Events (CTCAE) grade ≥3 AEs and CTCAE grade ≥3 SAEs were greater in the nintedanib plus docetaxel arm (75.9% and 31.3%) than the placebo plus docetaxel arm (68.5% and 26.6%). The AEs of greatest concern were fatal AEs and some imbalances were reported between treatment arms; fatal AEs being more common in the nintedanib plus docetaxel arm (6.3%) compared to the placebo plus docetaxel arm (2.4%). However, the company considers that these figures may be partially confounded by a longer median duration of treatment with nintedanib/placebo (4.2 months vs 3.0 months respectively) and docetaxel (median 5 and 4 cycles in the intervention and comparator arms respectively).

There was no significant difference over time, or between arms, in global health status/quality of life (QOL) or self-reported health related quality of life (HRQoL) assessments for cough, dyspnea or pain in LUME-Lung 1. Statistically significant improvements were observed for three individual pain items ('have pain', 'pain in chest' and 'pain in arm and shoulder') in favour of nintedanib plus docetaxel, while time to deterioration (TTD) for diarrhoea was significantly worsened in this arm.

Additional evidence is presented for nintedanib plus docetaxel compared to docetaxel and erlotinib by means of mixed treatment comparisons (MTCs) and, where possible, Bucher indirect comparisons. Compared to docetaxel, the base-case MTC analyses (which include four trials) report significant improvements in OS (HR 0.83, 95% CI: 0.70 to 0.99) and PFS (HR 0.77, 95% CI: 0.62 to 0.96) with the addition of nintedanib. The base-case MTC analyses also report significant improvements in OS (HR 0.64, 95% CI: 0.46 to 0.90) and PFS (HR 0.70, 95% CI: 0.50 to 0.998) for nintedanib plus docetaxel compared to erlotinib. The Bucher indirect comparisons (which includes two trials) support these findings (OS HR 0.56, 95% CI: 0.38 to 0.82; PFS 0.58, 95% CI: 0.39 to 0.87). Scenario analyses (including three of the trials from the base-case plus an additional trial) and sensitivity analyses of the base-case (including eight trials) and scenario analyses (including eight trials) were also conducted. These analyses all broadly support the base-case findings. For overall response rate (ORR), the base-case results suggest that there was no significant difference between nintedanib plus docetaxel in comparison with docetaxel or erlotinib.

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

The ERG is satisfied with the search strategy employed by the company to identify clinical effectiveness studies. It is not aware of any additional relevant ongoing or completed studies relevant to the decision problem.

The ERG is of the opinion that the LUME-Lung 1 study is well-designed and conducted, with low risk of bias. However, eligibility criteria mean that the patient population may not be representative of patients generally seen in clinical practice in England. Specifically, the trial excludes patients with any major pleural effusion or evidence of cavitary or necrotic tumours and therapeutic anticoagulation (except low dose heparin) or antiplatelet therapy (except for chronic low-dose therapy with acetylsalicylic acid ≤325mg/day). In addition, the proportion of patients aged ≥65 years is relatively small (28.3%) and such patients may have a poorer prognosis than younger patients. Given the focus of the decision problem on patients with adenocarcinoma, the ERG agrees it was appropriate for the company to only present data from LUME-Lung 1 for this patient population. Notwithstanding the exclusions of certain types of patients referred to above, the patient population is similar to the adenocarcinoma population likely to be treated for locally recurrent, locally advanced or metastatic disease in clinical practice in England. However, perhaps as a result of the eligibility criteria, it is noted that the rate of post-study therapy is relatively high (55.8%) which suggests this is an atypically fitter patient population than would be found in clinical practice in England. This is, however, not uncommon in clinical trials.

The ERG does not consider a comparison of nintedanib plus docetaxel to erlotinib is appropriate to decision problem. However, this was specified in the NICE scope and the company has therefore undertaken such a comparison via MTCs. The ERG has identified a number of methodological limitations related to the conduct of the MTCs (explored below in section 1.9.2) and advises that results from the MTCs should be treated with caution.

1.5 Summary of cost effectiveness evidence submitted by the company

The company developed a de novo partitioned survival Markov model that comprises three health states: progression-free (on or off treatment), progressive disease (PD) and death. All patients enter the model in the progression-free state. The model, when projecting PFS and OS data from LUME-Lung 1, fits a variety of standard parametric functions to the available trial data. Variants of this model structure have been used in the modelling of metastatic oncology for a number of previous NICE STAs. The model has been developed in Microsoft Excel using a 3-weekly cycle length. It includes a half-cycle correction and the time horizon is set at 15 years. As recommended by NICE, a discount rate of 3.5% has been used for both costs and outcomes; outcomes are measured in QALYs. The model perspective is that of the UK NHS. Resource use, costs and utilities were estimated based on information from LUME-Lung 1, published sources and clinical experts.

For the comparison of nintedanib plus docetaxel vs docetaxel, the company's incremental cost-effectiveness ratio (ICER) per QALY gained is £50,776. For the comparison of nintedanib plus docetaxel vs erlotinib, the company's ICER per QALY gained is £27,008. The company carried out a wide range of deterministic sensitivity analyses for these two comparisons. The results from the ten parameters that had the most influence on the ICER per QALY gained ranged from £44,034 to £59,711 for nintedanib plus docetaxel vs docetaxel and from £17,721 to £238,678 for nintedanib plus docetaxel vs erlotinib (in the latter comparison, the HR for OS was the single most influential variable). The results of the company's probabilistic sensitivity analysis (PSA) suggest that for nintedanib plus docetaxel vs docetaxel, there is a 2% and a 50% chance of nintedanib plus docetaxel being cost-effective at willingness to pay thresholds of £30,000 and £50,000 per QALY gained respectively; and a 65% and 94% chance of nintedanib plus docetaxel being cost-effective compared to erlotinib using the same thresholds.

1.6 Summary of the ERG's critique of cost effectiveness evidence submitted

The ERG is satisfied with the search strategy employed by the company to identify cost-effectiveness studies and is reasonably confident that no other relevant published articles exist.

Overall, the ERG found the company's model to be well structured. For most functions the assumptions and options are labelled and annotated where necessary; however, in some cases, the ERG has found it difficult to confirm details of the data sources employed (e.g. analyses related to Surveillance, Epidemiology and End Results program [SEER] and the National Lung Cancer Audit database [LUCADA]). The ERG identified eleven factors that limit confidence in the reliability of the company's model and/or results. These relate to: inappropriate methods used to project time-to-event outcomes (OS, PFS and time-on-treatment); mid-cycle adjustment error; inappropriate methods used to estimate cost of treatment doses; underestimate of true cost of febrile neutropenia; monitoring costs; non-UK standard approach to discounting; overall average disutility estimate for fatigue used for both regimens; error in stable disease costs and erroneous restriction of docetaxel to four cycles. The ERG is concerned by the number of implementation errors that have been identified, some of which have important consequences for the size of the estimated ICER per QALY gained for the comparison of nintedanib plus docetaxel vs docetaxel.

The ERG does not consider a comparison of nintedanib plus docetaxel to erlotinib is appropriate to decision problem. However, this was specified in the NICE scope and the company has therefore undertaken such a comparison. The ERG considers that this is

seriously flawed due to inconsistencies apparent in the available time-to-event data leading to conflicting results from the MTC. The ERG has applied other relevant amendments to the submitted model for this comparison, but the uncertainty in OS, PFS and time on treatment (ToT) probably far outweighs all other effects but cannot be quantified.

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

The company makes a case that nintedanib plus docetaxel meets the criteria set by NICE for end of life treatment. Namely:

- The life expectancy of the patient population was short (< 24 months). Patients with advanced NSCLC have a short life expectancy of less than 24 months on average. Using the extrapolated results from the LUME-Lung 1 trial data implemented in the cost effectiveness model, the median OS of patients on docetaxel monotherapy (current standard of care) is 10.23 months and the mean OS is 15.96 months.
- The number of patients who would be eligible for the treatment is small. The total eligible population in England for nintedanib plus docetaxel based on the anticipated marketing authorisation is estimated to be 703.
- The increase in OS is >3 months. Extension to life due to nintedanib plus docetaxel vs docetaxel monotherapy in the target population with the base-case assumptions within the model is a mean of 3.96 months. The extension in OS over erlotinib is a mean of 5.16 months.

1.8 ERG commentary on end of life criteria

The ERG agrees that patients with advanced NSCLC have a life expectancy of less than 24 months. It also agrees that only a small number of patients would be eligible for treatment with nintedanib plus docetaxel. By applying the Kaplan-Meier (K-M) trial results using the area under the curve (AUC) method until the long-term OS trends were established and then projecting remaining estimated survival using exponential trends, the ERG calculated the extension in mean OS to be 3.05 months for nintedanib plus docetaxel compared with docetaxel. It was not possible for the ERG to derive a mean estimate for OS gain for nintedanib plus docetaxel vs erlotinib.

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

1.9.1 Strengths

Clinical effectiveness

The ERG considers LUME-Lung 1 presents good quality evidence of clinical effectiveness which is directly relevant to the decision problem.

Cost-effectiveness

The company presented comprehensive and very detailed economic sections both within the CS and in the supplementary evidence. The company attempted to fully address the NICE scope. The ERG's requests to the company for additional economic analyses and further information were completed on time and to a high standard.

1.9.2 Weaknesses and areas of uncertainty

Clinical effectiveness

The ERG considers the MTCs are unnecessary because erlotinib is a comparator of no relevance to the vast majority of the patient population that would be considered for treatment with nintedanib plus docetaxel. The ERG further observes that LUME-Lung 1 is the only trial in which any patients (18.8%) received pemetrexed as a first-line treatment, as is now typically the case in clinical practice in England and so, arguably, all of the other trials included in the MTCs are of limited relevance to the decision problem. There are also other major methodological weaknesses and areas of uncertainty with the conduct of the MTCs, namely:

1. the proportional hazards assumption is not supported by the LUME-Lung 1 trial data for PFS or OS. Thus any estimation of the relative effectiveness of nintedanib plus docetaxel vs erlotinib (i.e. a calculated HR) will lack credibility and be effectively meaningless

2. differences in trial and patient characteristics mean that there is heterogeneity across trials which suggests that comparing data from these trials is inappropriate

Methodological issues also exist, namely: the use of both unadjusted and adjusted PFS and OS data, the use of PFS assessed by central independent review and local investigators and the use of primary PFS as opposed to updated PFS from LUME-Lung 1. However, these are not considered by the ERG to have major importance, particularly given the weaknesses and areas of uncertainty identified previously.

A greater number of fatal AEs have been observed in the nintedanib plus docetaxel arm than in the placebo plus docetaxel arm of the LUME-Lung 1 trial. However, the numbers are small and the company is using ongoing surveillance to monitor this issue.

Whilst LUME-Lung 1 is directly relevant to the decision problem, specific exclusion criteria employed in this trial may have excluded some patients who would ideally be considered for treatment in clinical practice in England. These are patients with major pleural effusion,

evidence of cavitary or necrotic tumours, or receiving therapeutic anticoagulation (except low dose heparin) or antiplatelet therapy (except for chronic low-dose therapy with acetylsalicylic acid ≤325mg/day). This may also partially explain why a higher proportion of patients in the trial than would be expected in clinical practice in England received third-line treatment.

Cost-effectiveness

The ERG identified a number of weaknesses and areas of uncertainty in the company's model for the comparison of nintedanib plus docetaxel vs docetaxel. The ERG considers that the high number of implementation errors is a major weakness of the model. These errors are present in estimates of both costs and benefits and therefore influence the size of the base-case ICER per QALY gained in a number of ways (mostly resulting in increasing the size of the ICER).

The most important area of uncertainty identified by the ERG is related to OS estimation. The company used a Log-Logistic survival model, whereas the ERG used the unadjusted trial data directly for the majority of patients, followed by projecting long-term survivors using trends evident in the data set. The company used data from the SEER and LUCADA to support the parametric survival modelling applied in the model. However, it was not possible for the ERG to assess whether this approach was valid; the analyses reported by the company did not provide references for the specific data sets used, nor did the company present sufficient explanation of the data employed. When the ERG replaced the company's preferred OS model with the ERG's preferred OS model, there was a major impact on the size of the ICER per QALY gained; it increased substantially as the size of the ERG's estimated OS incremental gain was reduced.

The ERG does not consider the company's comparison of nintedanib plus docetaxel vs erlotinib to be relevant to the decision problem. Furthermore, even if the comparison was considered to be relevant, the ERG has noted a number of flaws in the company's MTCs that render the clinical effectiveness results unreliable. The ERG considers that these problems are so fundamental that it is not possible to rectify them and modify the company's model to provide improved estimates of OS, PFS and the relative cost-effectiveness of nintedanib plus docetaxel and erlotinib.

1.10 Summary of exploratory and sensitivity analyses undertaken by the ERG

For the comparison of nintedanib plus docetaxel vs docetaxel, the company's base-case ICER (£50,776 per QALY gained) would increase to £85,292 per QALY gained if all 11 ERG

recommended revisions were applied and would increase to £82,995 per QALY gained if all but the limit on the number of cycles of docetaxel treatment were applied.

The ERG has been unable to estimate an ICER for the comparison of nintedanib plus docetaxel vs erlotinib for the reasons stated in the ERG's critique of the clinical effectiveness and cost-effectiveness evidence.

2 BACKGROUND

2.1 Critique of company's description of underlying health problems

section 2.1 of the CS[1] provides a brief overview of NSCLC.sections 2.1 and 2.2 of the CS[1] provide data on the number of patients with NSCLC and section 2.3 provides details about the life expectancy of people with NSCLC in England. These sections appropriately present the key issues relating to the underlying health problems of patients with NSCLC and are summarised as presented in the CS[1] in Box 1.

Box 1 Lung cancer disease course and epidemiology

Types of lung cancer

- [Non-small cell lung cancer (NSCLC) is] the most common type, accounting for 85% to 90% of cases
- Adenocarcinoma is the most common [40%] histological sub-type of NSCLC[2]
- Patients with NSCLC have a poor prognosis that has not changed significantly in the past decades

The disease course

- Lung cancer does not usually cause noticeable symptoms until it is locally advanced or has spread through much of the lungs or into other parts of the body (i.e. metastatic lung cancer)
- This means that the outlook for lung cancer is poor compared with other types of cancer[3]

Epidemiology

- Lung cancer is the second most common cancer in the UK; there are around 41,500 new cases diagnosed each year, with 35,406 new cases in England and Wales in 2010, and more than one in five cancer deaths (22%) in the UK are from lung cancer[4]
- Smoking causes more than 8 in 10 lung cancers in the UK[5]
- At diagnosis, 10 to 15% of patients have locally advanced cancer, i.e. stage IIIB and 40% of patients have metastatic cancer i.e. stage IV[6,7]
- Moreover, patients with stage IIIB and stage IV NSCLC have the lowest 5-year survival rate, at 5% and 1%, respectively[2,8-10]

In relation to epidemiology, the ERG adds that the LUCADA database audit published in 2012 reported approximately 57% of patients with NSCLC were stage IIIB or stage IV.[11] This figure is consistent with the estimates presented by the company in Box 1 (50% to 55%). A recent National Institute for Health Research Horizon Scanning Centre document[12] states that the incidence of stage III/IV NSCLC is 78%. This implies an incidence of stage IIIA disease of around 20% to 30% if the estimates for stage IIIB and IV cited by the company and ERG are subtracted.

2.2 Critique of company's overview of current service provision

As stated in section 2.1 of the CS,[1] the type of treatment that patients with locally advanced or metastatic NSCLC receive depends on several factors, including, but not limited to, tumour histology and EGFR mutation status. Patients with mutation free (i.e. EGFR-negative) locally advanced or metastatic lung cancer usually receive platinum doublet chemotherapy in the first-line setting, typically pemetrexed plus cisplatin for patients with adenocarcinoma.[3] As stated in section 2.5 of the CS,[1] TKIs - erlotinib, gefitinib or afatinib - are all NICE recommended options[13-15] for patients with EGFR-mutations. At present all three of these drugs have been made available to NHS patients at discount prices, as set out in patient access schemes.

According to the company, approximately 30%[16] to 50%[17] of patients with locally advanced or metastatic NSCLC receive second-line treatment. The current options for second and subsequent lines of treatment, as stated in sections 2.1 and 2.5 of the CS,[1] are summarised in Box 2. The company's advisory board, which comprised five clinical experts, estimated that [redacted] of all patients who had received second-line treatment would go on to receive third-line treatment, with approximately one third of these patients receiving this treatment as part of an ongoing clinical trial.[1] The company's own data on file[18] that reports on data from the final quarter of 2012 appears to support this view. These data show that 13.33% of patients who received second-line treatment also received third-line cytotoxic treatment.

Box 2 Current service provision for patients with NSCLC following first-line treatment

Second-line treatment

- The major goal of second-line treatment is to prolong life without worsening HRQ[o]L
- There are a number of new therapies that target patients with relatively rare mutations (e.g. EGFR), but patients with adenocarcinomas and without actionable mutations [e.g. EGFR] who progress following first-line chemotherapy have limited therapy options
- Following failure of first-line chemotherapy, treatment options are limited to docetaxel monotherapy or erlotinib [19,20]
- Docetaxel monotherapy can be considered for second-line treatment of locally advanced or metastatic NSCLC when cancer has relapsed after previous chemotherapy
- Erlotinib is recommended, within its licensed indication, as an alternative to docetaxel as a second-line treatment option for patients with NSCLC only on the basis that it is provided by the manufacturer at an overall treatment cost (including administration, AEs and monitoring costs) equal to that of docetaxel

Third-line treatment and subsequent lines of therapy

- Currently, there are no NICE-recommended technologies

In section 2.6 of their submission,[1] the company notes that the use of erlotinib as a second-line treatment is being reviewed by NICE and presents recommendations issued by NICE in February 2014. The ERG notes that this guidance is in the process of being replaced, with draft guidance published on the NICE website on 7 August 2014. One of the Appraisal Committee's preliminary recommendations[21] from both February and August 2014 is that

erlotinib should not be recommended for treating locally advanced or metastatic NSCLC that has progressed after non-targeted chemotherapy in people with tumours that are EGFR-negative. Furthermore, erlotinib is only recommended in second-line treatment for patients with tumours that are EGFR-positive, or of unknown status, in limited circumstances (Box 3).

In addition, the ERG notes that as the recommended first-line treatment for patients with tumours that are EGFR-positive is a TKI,[13-15] there are unlikely to be many patients with EGFR-positive tumours for whom erlotinib is considered an appropriate second-line treatment. Furthermore, as noted on page 35 of the CS,[1] the opinion of clinical experts is that patients who are sufficiently fit to allow them to tolerate treatment with docetaxel receive docetaxel rather than erlotinib. It is, therefore, unlikely that the same group of patients who would be eligible to receive erlotinib is the same as that who would be considered for docetaxel.

Box 3 Draft NICE guidance on the use of erlotinib as second-line treatment, 7th August 2014

- Erlotinib should not be recommended for treating locally advanced or metastatic NSCLC that has progressed after non-targeted chemotherapy in people with tumours that are EGFR-negative
- Erlotinib is recommended as an option for treating locally advanced or metastatic non-small-cell lung cancer that has progressed in people who have had non-targeted chemotherapy because of delayed confirmation that their tumour is epidermal growth factor receptor tyrosine kinase (EGFR-TK) mutation-positive, only if the manufacturer provides erlotinib with the discount agreed in the patient access scheme
- Erlotinib is recommended as an option for treating locally advanced or metastatic non-small-cell lung cancer that has progressed after chemotherapy in people with tumours of unknown epidermal growth factor receptor tyrosine kinase (EGFR-TK) mutation status, only if:
    - the result of an EGFR-TK mutation diagnostic test is unobtainable because of an inadequate tissue sample or poor quality DNA and
    - the treating clinician considers that the tumour is very likely to be EGFR-TK mutation-positive and
    - the person's disease responds to the first 2 cycles of treatment with erlotinib and
    - the manufacturer provides erlotinib with the discount agreed in the patient access scheme

According to the company: "Nintedanib fits well in the existing clinical pathway and can complement docetaxel treatment as an effective second-line option for patients with locally advanced/metastatic or recurrent NSCLC of adenocarcinoma tumour histology, previously treated with one line of chemotherapy." (page 34 of the CS[1]) As highlighted in the CS,[1] nintedanib is a potent, orally-administered small molecule triple angiokinase inhibitor targeting three receptor classes: vascular endothelial growth factor receptors (VEGFR), fibroblast growth factor receptors and platelet-derived growth factor receptors α and β.[22-24] These receptors have a key role in the formation and maintenance of new blood vessels (angiogenesis) and tumour growth.[25-27] Suppression of neo-angiogenesis via inhibition of VEGFR is considered a promising strategy for the treatment of human solid tumours, impacting tumour growth and spread.[25-27] The simultaneous targeting of all three pathways may be more effective than inhibition of angiogenesis via the VEGF pathway alone.

Largely based on the findings from the pivotal trial comparing nintedanib plus docetaxel to placebo plus docetaxel (LUME-Lung 1[24]), nintedanib is expected to be licensed in combination with docetaxel. Indeed, a positive opinion was received by the European Medicines Agency (EMA) on 25 September 2014 as follows: "Vargatef [nintedanib] is indicated in combination with docetaxel for the treatment of adult patients with locally advanced, metastatic or locally recurrent non-small cell lung cancer (NSCLC) of adenocarcinoma tumour histology after first-line chemotherapy."[28] As noted above, the group of patients who would be eligible to receive second-line docetaxel - and therefore nintedanib - is not likely to be the same as those who would be eligible to receive second-line erlotinib. Therefore the ERG considers with only very few exceptions, nintedanib plus docetaxel would fit into the existing treatment pathway as a comparator to docetaxel rather than erlotinib.

The ERG notes that the aforementioned positive opinion includes patients with locally recurrent NSCLC. In order to be classified as locally recurrent, a patient would initially present with early stage disease (stage I, II or IIIa). The company does not provide information on the service provision for these patients, presumably because the NICE scope is focussed on patients with locally advanced or metastatic disease. However, since the scope also focussed on second-line treatment following chemotherapy, the ERG considers these patients will have locally advanced or metastatic cancer by this stage. The ERG notes that patients with stage I, II or IIIa will initially be treated with surgery or radical radiotherapy and subsequently receive first-line chemotherapy when their disease has relapsed and/or spread.[29] The choice of chemotherapy will again depend on several factors, including, but not limited to, tumour histology and EGFR mutation status.

The estimated number of patients with locally advanced or metastatic adenocarcinoma potentially eligible for second-line treatment with nintedanib plus docetaxel in England is reported by the company to be 703. The ERG agrees with the company that a similar number of patients are likely to be eligible for treatment with nintedanib plus docetaxel. Based on data from the pivotal LUME-Lung 1[24] in which the median number of cycles with docetaxel was five (see also section 4.5) and given the norm in clinical practice in England is to provide a maximum of four cycles of docetaxel, the ERG considers the majority of patients would receive nintedanib in combination with four cycles of docetaxel.

3 CRITIQUE OF COMPANY'S DEFINITION OF DECISION PROBLEM

Table 1 displays the decision problem presented in the CS[1] and that addressed by the company. Each parameter is discussed in detail in the text following the table.

Table 1 Decision problem specified by NICE and addressed in the company's submission

Parameter	Final scope issued by NICE	Decision problem addressed in the company's submission
Population	Adults with locally advanced or metastatic non-small cell lung cancer (NSCLC) that has progressed following prior chemotherapy	Patients with locally advanced, metastatic or recurrent NSCLC of adenocarcinoma tumour histology after first-line chemotherapy
Intervention	Nintedanib in combination with docetaxel	As per final scope
Comparator(s)	Docetaxel monotherapy Erlotinib	Primary analysis: docetaxel monotherapy Secondary analysis: erlotinib monotherapy
Outcomes	The outcome measures to be considered include: • overall survival • progression-free survival • response rates • adverse effects of treatment • health-related quality of life	As per final scope
Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. The availability of any patient access schemes for the intervention or comparator technologies should be taken into account	As per final scope
Subgroups to be considered	None	Not applicable
Special considerations including equity or equality issues	None	Not applicable

Source: adapted from Table 5 of the CS[1]

3.1 Population

The population addressed in the CS[1] differs to the population specified in the scope. The scope states the population is adults with locally advanced or metastatic NSCLC that has progressed following prior chemotherapy. The decision problem addressed by the company is patients with locally advanced, metastatic or locally recurrent NSCLC of adenocarcinoma tumour histology who had previously received first-line chemotherapy. This is in line with the

anticipated full marketing authorisation for nintedanib which also specified nintedanib should be administered in combination with docetaxel (expected in December 2014). The ERG notes that to be classified as locally recurrent, a patient would initially present with early stage disease (stage I, II or IIIA) and be treated with surgery or radical radiotherapy and then relapse in the same area without metastases. Since the anticipated license also stipulates patients must have previously received first-line chemotherapy, then all patients would have locally advanced or metastatic disease at the time of second-line treatment regardless of their initial diagnosis. The ERG notes that while the scope makes no specification about the EGFR mutation status of tumours, in the UK the majority (85% to 90%) of patients have EGFR wild-type tumours (EGFR-negative).[30-32] The ERG further notes that as patients who receive nintedanib also receive docetaxel, the vast majority of eligible patients will be required to have ECOG PS 0 to 1.

3.2 Intervention

The intervention described in the CS[1] is nintedanib. Nintedanib does not currently have a full UK Marketing Authorisation. It does however have a positive opinion from the EMA and it is anticipated that it will be licensed in December 2014 in combination with docetaxel (the specified intervention in the final NICE scope). Nintedanib is provided orally at a dose of 200mg twice daily (BD) and dose adjustments are permitted in patients who experience AEs. The first dose reduction is to 150mg BD and, if required, the dose may be further reduced to 100mg BD. Docetaxel is administered intravenously alongside nintedanib on day 1 of a 21 day cycle at a dose of 75mg/m^2. If necessary, docetaxel doses may be reduced to 60mg/m^2 as per the docetaxel summary of product characteristics (SmPC)[33] and standard clinical practice. Nintedanib may be provided as monotherapy after discontinuation of docetaxel. In the pivotal LUME-Lung 1[24] trial, this was only permitted after four cycles of treatment with docetaxel. The ERG notes that in England, clinicians rarely administer more than four cycles of docetaxel due to the toxicity associated with this drug.

3.3 Comparators

Both docetaxel monotherapy and erlotinib monotherapy are considered as comparators for locally advanced or metastatic disease in the CS.[1] These are the same comparators that are specified in the scope. The company considers docetaxel monotherapy to be the comparator for the primary analysis and considers erlotinib to be the comparator for secondary analyses. This is because as stated on page 184 of the CS,[1] based on feedback from clinical experts, it does not believe that erlotinib is a relevant comparator. The ERG agrees with the company. As noted in section 2.2, the ERG notes that one of the NICE Appraisal Committee's preliminary recommendations[21] is that erlotinib should not be recommended for

treating locally advanced or metastatic NSCLC that has progressed after non-targeted chemotherapy in people with tumours that are EGFR-negative. Furthermore, the characteristics of patients who are considered suitable for second-line erlotinib treatment are different from those who are considered suitable for docetaxel treatment. Given that erlotinib is likely to be preferred when patients have a poorer ECOG PS and/or have EGFR-positive tumours, docetaxel is the most appropriate comparator to nintedanib plus docetaxel in the second-line setting. The company notes that no other agents are licenced or routinely used for this indication (pemetrexed is licensed but not NICE approved). Therefore, no other comparisons are presented (although as reported in section 4.4, there were other comparators employed in the MTCs). The ERG agrees that this is appropriate.

3.4 Outcomes

Clinical evidence is reported in the CS[1] for all outcomes specified in the scope: OS, PFS, response rate (reported as ORR]and disease control rate), AEs of treatment and HRQoL.

3.5 Economic analysis

Results are expressed in terms of incremental cost per QALY gained. Various time horizons are presented with lifetime (15 years) being that of the primary analysis (appropriate for a condition such as lung cancer, with low survival rates). Costs are considered from the perspective of the NHS. No patient access scheme has been submitted.

3.6 Subgroups

No subgroups were specified by NICE or identified by the company.

3.7 Other relevant factors

The company states on page 37 of the CS[1] that it does not consider there will be any equality issues if nintedanib is recommended by NICE.

4 CLINICAL EFFECTIVENESS

4.1 Introduction

This section provides a structured critique of the methods and clinical evidence submitted by Boehringer Ingelheim Ltd in support of the use of nintedanib in combination with docetaxel for the treatment of adult patients with locally advanced, metastatic or recurrent NSCLC of adenocarcinoma tumour histology after first-line chemotherapy. The key components of the clinical evidence presented in the CS[1] are (i) a report of the pivotal trial (LUME-Lung 1[24]) which compared nintedanib plus docetaxel to placebo plus docetaxel (ii) a report of the company's MTC which was conducted in order to compare nintedanib plus docetaxel to erlotinib.

4.2 Critique of the methods of review(s)

The company conducted a systematic literature review to identify RCTs of patients with previously treated second-line NSCLC. The review was designed to identify evidence for any drug, not limited to nintedanib plus docetaxel, erlotinib or docetaxel.

4.2.1 Searches

Sections 6.1.1 and Appendices 1 and 4 of the CS[1] describe the search strategies employed for the systematic review (direct evidence) and the multiple treatment comparison (MTC) (indirect evidence), respectively. While the ERG notes some potential minor limitations with the search strategy employed by the company (as outlined in Appendix 1), the ERG considers that the search strategies employed by the company were appropriate and sufficiently comprehensive to identify relevant studies.

In order to ascertain whether the company had missed any relevant studies or not, the ERG also conducted its own searches, as summarised in Appendix 1. However, the ERG did identify four additional conference presentations[34-37] for the pivotal LUME-Lung 1[24] trial not cited in the CS.[1]

4.2.2 Eligibility criteria

Although the same search strategy was employed to identify studies for inclusion in the systematic review (direct evidence) and MTC (indirect evidence), different eligibility criteria were appropriately employed for each. These are described in detail in Table 6 (pages 44 to 45) and Table 25 (pages 106 to 107) respectively of the CS[1] and summarised in Appendix 2. In general the ERG considers the criteria for both reviews were appropriate although notes that the eligibility of studies for inclusion into the MTC was limited to include only results with

abstracts, an unusual exclusion criterion which could potentially have removed relevant results. However, as noted in section 4.2.1, the ERG conducted its own searches and did not identify any additional eligible studies.

Although the same search was conducted to identify studies for both the systematic review and the MTC, it is unclear if the eligibility criteria for both reviews were simultaneously employed. The ERG notes from an examination of Figures 1 (page 46) and 19 (page 109) in the CS[1] that the number of records screened in the systematic review differed from the number screened in the MTC, suggesting this was not the case.

4.2.3 Quality assessment

The company conducted an assessment of the risk of bias of LUME-Lung 1,[24] the only study to meet the inclusion criteria for the systematic review, and all studies included in the MTC. This assessment included elements of the tool for assessing risk of bias, as recommended by the Cochrane Collaboration.[38] The ERG agrees this is an appropriate tool for assessing the quality of RCTs.

4.2.4 Evidence synthesis

One trial (LUME-Lung 1[24]) was identified by the searches for inclusion into the systematic review and hence the findings were appropriately presented narratively. This trial[24] compared nintedanib plus docetaxel to placebo plus docetaxel. In order to compare nintedanib plus docetaxel to erlotinib, the other comparator specified in the final NICE scope, the company conducted a MTC. The ERG's critique of the company's MTCs is presented in section 4.3.

4.3 Critique of the direct evidence

4.3.1 Identified studies

Only one RCT (LUME-Lung 1[24]) that presented direct evidence relevant to the decision problem was identified by the systematic review. The ERG is not aware of any additional relevant ongoing or completed studies. The company also referred to LUME-Lung 2[39] which compared nintedanib plus pemetrexed to placebo plus pemetrexed. However, data from LUME-Lung 2[39] were solely used to inform the pre-specified statistical analysis of LUME-Lung 1.[24]

As well as being published in a peer reviewed journal,[24] data from LUME-lung 1 were also provided by the company in two clinical trial reports (CTRs): primary PFS[40] and final OS[41] since analyses were conducted at both these time points (see section 4.3.4). Selected appendices to the CTR for final OS were also provided.[42] The company also provided the trial statistical analysis plan (TSAP),[43] the TSAP addendum[44] and the summary of clinical efficacy.[45] Three conference presentations were also cited, two poster presentations,[46,47] and an oral presentation, the slides of which were provided;[48] one of the poster presentations[46] also included data from LUME-Lung 2,[39] the focus of the presentation being to identify potential clinical biomarkers for second-line treatment. These findings are not presented by the company it the CS.[1] The ERG's search also identified four conference presentations not referred to by the company;[34-37] these do not appear to contain any additional data to that included in the CS.[1]

4.3.2 Trial characteristics

The key characteristics of LUME-Lung 1[24] are summarised in Table 2. The study was conducted internationally and randomised 1,314 patients in a 1:1 ratio to nintedanib plus docetaxel or placebo plus docetaxel. Randomisation was stratified by ECOG PS (0 vs 1), previous bevacizumab treatment (yes vs no), histology (squamous vs non-squamous) and presence of brain metastases (yes vs no). The ERG is of the opinion that the LUME-Lung 1[24] study is well-designed and conducted. A large number of patients were recruited to the study and the length of trial follow-up means that the data collected are mature and allow reasonable conclusions to be drawn from the data.

The ERG notes that some of the participating treatment centres were located in the UK although it is not known how many centres or numbers of patients were recruited (this was reported in Appendix 16.1.4 of the CTRs,[40,41] an appendix not included with the CS[1]). However, the ERG notes that the eligibility criteria for entry into this trial (see Appendix 3 for the full eligibility criteria as provided in the CS,[1] pages 58 to 59) do mean the patient

population was likely to be different to that of standard clinical practice in England in a number of different ways. Specifically the trial excludes, patients with clinically significant pleural effusion or evidence of cavitary or necrotic tumours and therapeutic anticoagulation (except low dose heparin) or antiplatelet therapy (except for chronic low-dose therapy with acetylsalicylic acid ≤325mg/day). In clinical practice these patients are likely to have a poorer prognosis than patients included in the trial[49-52] although it is recognised that cavitation may be less of a strong prognostic factor[49] than pleural effusions[50,51] or venous thromboembilsm.[52]

The ERG further notes that previous treatment with docetaxel is a specific exclusion criterion to entry in LUME-Lung 1.[24] Docetaxel is licensed for first-line treatment of NSCLC. However, this is rarely used in clinical practice in England, pemetrexed being the preferred choice for adenocarcinoma patients (see also section 2.2).

Table 2 Trial characteristics of LUME-Lung 1

Characteristics of LUME-Lung 1[24]	
Location	211 locations in 27 countries (Austria, Belarus, Belgium, Bulgaria, China, Croatia, Czech Republic, Denmark, France, Georgian Republic, Germany, Greece, India, Israel, Italy, Lithuania, Poland, Portugal, Romania, Russia, Slovakia, South Korea, South Africa, Spain, Switzerland, Ukraine, United Kingdom)
Design	Phase III multi-centre, randomised, parallel-group, double-blind, placebo-controlled RCT
Population	Patients with locally advanced, metastatic (stage IIIB/IV) or recurrent NSCLC after failure of first-line chemotherapy
Duration of study	23 December 2008 to 15 February 2013 (data cut-off date)
Intervention and comparator	Nintedanib + docetaxel (n=655) Nintedanib 200mg twice daily, orally, on days 2 to 21 of a 21-day cycle in combination with docetaxel 75mg/m^2 IV on day 1 of a 21-day cycle Matched placebo + docetaxel (n=659) Matched placebo twice daily on days 2 to 21 of a 21-day cycle in combination with docetaxel 75mg/m^2 IV on day 1 of a 21-day cycle
Primary outcomes	PFS by central independent review
Secondary outcomes	OS (key secondary endpoint) PFS by local investigator review Tumour response by central independent review and local investigator assessment, including: confirmed objective response; disease control; time to confirmed objective response; duration of confirmed objective response; duration of disease control; change in tumour size; clinical improvement HRQoL Pharmacokinetics safety and tolerability
Duration of follow-up	Median follow-up at the primary PFS analysis (2 November 2010) was 7.1 months (interquartile range: 3.8 to 11.0) and 31.7 months (interquartile range: 27.8 to 36.1 months) at the time of the final OS analysis (15 February 2013)

AE=adverse event; HRQoL=health related quality of life; NSCLC=non-small-cell lung cancer; OS=overall survival; PFS=progression-free survival; RCT=randomised controlled trial
Source: adapted from Table 8 of the CS[1]

4.3.3 Participant characteristics

Not all patients in LUME-Lung 1[24] had histology of adenocarcinoma. As the expected marketing authorisation for nintedanib plus docetaxel is specifically for patients with adenocarcinoma, the company only presented data for the overall population of patients with NSCLC where the results were of relevance to statistical testing (see section 4.3.4). The ERG agrees that this is appropriate. The participant characteristics of 658 (50.1%) patients with adenocarcinoma in LUME-Lung 1[24] are summarised in Table 3. While some patients (15.8%) had early stage disease at diagnosis, at the time of treatment the ERG considers all would have locally advanced or metastatic disease since patients had all received first-line treatment and were now being treated second-line. Indeed, 94.2% of all patients had metastatic disease at screening. The mean time from diagnosis to randomisation into the trial reported in Table 15.1.8: 3 of the CTR[41] was 12.84 months (median 8.74 months).

The company comments that demographic and baseline disease characteristics are well balanced between the two arms of the trial, and that the population is largely representative of patients typically diagnosed with adenocarcinoma although it is noted by the ERG that the proportion of patients aged ≥65 years is relatively small (28.3%). The ERG agrees that the patient characteristics are well balanced.

Data on EGFR mutation status was not routinely collected in LUME-Lung 1[24] although in response to a query from the ERG during the clarification process, the company stated these data has been retrospectively collected for a sample of patients: [redacted]

Notwithstanding the exclusions of certain types of patients identified in section 4.3.2, the patient population is similar to the population who would be treated in clinical practice in England with the exception that a smaller proportion (18.8%) of patients than would be expected today had received prior pemetrexed. Additionally, perhaps as a result of the eligibility criteria, it is noted that post-study therapy is relatively high (55.8%) which suggests this is a fitter patient population than in clinical practice England in Wales. As noted in section 2.2, in England, in clinical practice around [redacted] of all patients who receive second-line treatment subsequently receive third-line treatment (including a third of patients who are enrolled into trials).

Table 3 Participant characteristics of patients with adenocarcinoma in LUME-Lung 1

Characteristic		Nintedanib + docetaxel (N=322)	Placebo + docetaxel (N=336)
Sex, n (%)	Male Female	203 (63.0) 119 (37.0)	208 (61.9) 128 (38.1)
Age, years	Mean (SD) Median (range)	58.5 (10.1) 60.0 (29 to 80)	58.6 (9.5) 59.0 (30 to 80)
Age ≥65 years, n(%)		90 (28.0)	96 (28.6)
Race, n (%)	Asian White Other	65 (20.2) 253 (78.6) 4 (1.2)	78 (23.2) 253 (75.3) 5 (1.5)
ECOG performance status, n (%)	0 1†	96 (29.8) 226 (70.2)	99 (29.5) 237 (70.5)
Stage of disease at diagnosis, n(%)	<IIIB IIIB IV	50 (15.6) 55 (17.2) 215 (67.2)	54 (16.1) 45 (13.4) 237 (70.5)
Local recurrence without metastases at screening		22 (6.8)	16 (4.8)
Smoking status, n (%)	Never smoked Ex-smoker Current smoker	115 (35.7) 151 (46.9) 56 (17.4)	115 (34.2) 162 (48.2) 59 (17.6)
Prior first-line therapy	Platinum-based therapy	308 (95.7)	323 (96.1)
	Non-platinum-based therapy	10 (3.1)	10 (3.0)
Prior pemetrexed, n (%)	As platinum therapy As non-platinum therapy	58 (18.0) 3 (0.9)	61 (18.2) 2 (0.6)
Prior bevacizumab, n (%)		24 (7.5)	21 (6.3)
Brain metastases at study entry, n (%)	Present Absent	26 (8.1) 296 (91.9)	23 (6.8) 313 (93.2)
Post study therapy	Any systemic therapy Any chemotherapy Pemetrexed Docetaxel Other chemotherapy EGFR-TK inhibitor Anti-angiogenesis agent Investigational agent	179 (55.6) 123 (38.2) 52 (16.1) 15 (4.7) 90 (28.0) 98 (30.4) 6 (1.9) 18 (5.6)	188 (56.0) 136 (40.5) 62 (18.5) 13 (3.9) 101 (30.1) 105 (31.3) 2 (0.6) 5 (1.5)

ECOG=Eastern Cooperative Oncology Group; EGFR-TK=epidermal growth factor receptor tyrosine kinase; SD=standard deviation
† Including one patient in the nintedanib arm who had an ECOG PS of 2 at screening and at randomisation (i.e. at baseline)
Source: adapted from Table 10 of the CS[1] with additional information taken from Table 15.1.8:2 of the CTR[41]

4.3.4 Description and critique of the statistical approach

Information relevant to the statistical approach taken by the company to analyse data from the pivotal study LUME-Lung 1[24] are taken from the TSAP,[43] trial protocol,[53] CTRs[40,41] and the CS.[1]

Sample size calculation

Details of the sample size calculation performed by the company are reported in the CS[1] (page 71). The study was powered (at the 90% level) to detect a HR for centrally independently assessed PFS for the comparison of nintedanib plus docetaxel vs placebo

plus docetaxel of 0.7843. This would require 713 PFS events. The ERG is satisfied that the company's pre-specified sample size calculation is correct. However as noted in section 3.1 above, only patients with adenocarcinoma were considered relevant to this STA. The company therefore only presents data for the adenocarcinoma population. The ERG notes that although around half of the patients in LUME-Lung 1[24] had adenocarcinoma (see section 4.3.3) this was not a stratification factor (see section 4.3.2) and so patients with adenocarcinoma were not strictly a randomised subgroup although they do constitute the majority of non-squamous patients which was a stratification factor. However, as noted in section 4.3.3, baseline characteristics were well balanced between the two groups suggesting the analyses were valid.

Protocol amendments

A list of changes implemented after a protocol amendment (dated 15 May 2009) is included in the CTR[40] (pages 120 to 121). The changes included slight adjustments to the exclusion criteria, clarification of ongoing safety evaluations, and timings of the screening period. All changes were made before analyses began, and so were not driven by the results of the trial. The ERG considers that it is very unlikely that any of the changes would influence the outcomes or analyses of LUME-Lung 1,[24] or would be a cause for concern.

Clinical endpoints and statistical analyses

The company provides a list of outcome measures used in LUME-Lung 1[24] in Table 13 (page 66) of the CS[1] (also summarised in Appendix 4 of the ERG report). The ERG is satisfied that all outcomes were pre-specified in the TSAP[43] and reported in full in the CTRs.[40,41]

The intention-to-treat (ITT) population was used in all efficacy analyses. The primary outcome of PFS by central review was analysed using the K-M method, and a stratified log-rank test. Cox regression analyses were also carried out to estimate treatment effect, including adjustment for stratification factors.

Secondary outcomes relevant to the decision problem included OS, PFS by local investigator review, best tumour response, HRQoL and AEs. OS, the key secondary outcome of the trial, was also analysed using a stratified log-rank test. Tumour response was reported for both central independent review and local investigator review according to modified Response Evaluation in Solid Tumours (RECIST) criteria and analysed using logistic regression.

Within clinical trials, time-to-event data like PFS and OS are commonly reported as HRs, derived from the Cox proportional hazards model. Such a model does not appear to be appropriate for the PFS and OS results of this trial since hazards are not independent of time (see Appendix 7) and the HR (and 95% CIs) presented for PFS and OS offer inaccurate estimates of relative efficacy. Instead of assuming proportional hazards, alternative approaches may be more appropriate to better reflect relative efficacy in the data..

The CS[1] (page 71) describes the stages of analyses in Table 14. These are summarised in Appendix 4 of the ERG report. The ERG is satisfied that each of these stages was pre-specified in the trial protocol.[53]

Following the hypothesis-generating trial LUME-Lung 2,[39] an amendment to the statistical plan of LUME-Lung 1[24] was implemented such that statistical testing of OS would only be conducted if a significant difference had been observed for the primary analysis of PFS and had been confirmed by the updated analysis of PFS. If this condition was satisfied, OS analyses would then be conducted in a sequential fashion, i.e. the null hypothesis was to be tested in each population only if a significant treatment effect had been shown in the previous population. This hierarchical method was utilised to control the type 1 error rate (detecting an effect when one is not present), which can be high when performing a large number of statistical tests. The sequence of populations was:

1. Adenocarcinoma patients who had progressed within 9 months of starting first-line therapy (i.e. the T<9m adenocarcinoma population)

2. Adenocarcinoma population

3. Overall trial population

The CS[1] clarifies that the amendment to the TSAP[43] was made before database lock and unblinding of data used in the final OS analysis; the ERG considers that this amendment is unlikely to bias the results from LUME-Lung 1.[24]

Subgroup analyses

A number of pre-specified analyses for the primary endpoint of PFS assessed by central review and for the secondary outcome OS were pre-specified in the protocol. The company also conducted post-hoc subgroup analyses and a number of baseline characteristics (CS,[1] page 77) were also investigated for subgroup effects. The subgroup types analysed are summarised in Appendix 5 of the ERG report. The ERG notes that there is a large number of

subgroup analyses but is satisfied that the results of all of the pre-specified and post-hoc subgroup analyses are provided in the CS.[1]

Sensitivity analyses

A number of sensitivity analyses for the primary endpoint of PFS assessed by central review and for the secondary outcome OS were pre-specified in the protocol.[53] These are summarised in Appendix 5 of the ERG report. However, no sensitivity analysis of PFS in the adenocarcinoma population was performed. The ERG is satisfied that the results of all of the pre-specified and post-hoc sensitivity analyses are provided in the CTR.[41]

4.3.5 Risk of bias

The company conducted an assessment of the risk of bias using the criteria recommended by NICE in the Guide to the Methods of Technology Appraisal.[54] The risk of bias assessment is presented in Table 4. The ERG is satisfied with the risk of bias assessment presented in the CS[1] and agrees that the study has an overall low risk of bias.

Table 4 Assessment of risk of bias conducted by company for LUME-Lung 1trial

Criteria	Response
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?	Yes
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

Source: Table 17 of the CS[1]

4.3.6 Results

The focus of the results section in both the CS[1] and this ERG report is the adenocarcinoma population from LUME-Lung 1,[24] as this is the population relevant to the decision problem.

However, results of the primary PFS analysis for the overall trial population and OS for the T<9m adenocarcinoma population have been presented wherever necessary (and are clearly labelled), due to the fact that these populations were part of the previously described hierarchical OS statistical analysis (section 4.3.4). By presenting these results, the justification for conducting the analysis of OS for the patients with adenocarcinoma has been demonstrated.

Progression-free survival

The CS[1] reports that median follow-up was 7.1 months at the time at the primary PFS analysis (2 November 2010), the results of which are presented in Table 5.

Table 5 Primary analysis of centrally independently assessed PFS in LUME-Lung 1 trial (November 2010)

Outcome	Nintedanib + docetaxel (median)*	Placebo + docetaxel (median)*	HR vs placebo arm (95% CI)†	p-value	Risk reduction
PFS in overall ITT population	3.4 months	2.7 months	0.79 (0.68 to 0.92)	0.0019	21%
PFS in adenocarcinoma population§	4.0 months	2.8 months	0.77 (0.62 to 0.96)	0.0193	23%

CI=confidence interval; HR=hazard ratio; ITT=intention-to-treat; PFS=progression-free survival
* Based on unadjusted Kaplan-Meier estimates for each treatment arm
† A proportional hazards model stratified by three factors (ECOG PS at baseline, presence of brain metastases at baseline, prior bevacizumab therapy) was used to derive the HR, 95% CI, and p-value (corresponding to the stratified log-rank test p-value)
§ Analysis conducted retrospectively
Source: Table 18 of the CS[1]

The results suggest that the use of nintedanib plus docetaxel significantly improved PFS in comparison to placebo plus docetaxel in both the overall trial population and in the subgroup of patients with adenocarcinoma. However, the ERG suggests that these results should be interpreted with caution, due to the violation of the proportional hazards assumption (see section 4.3.4). In particular, it is evident that the trial survival arms converge within the duration of the trial indicating that PFS gain from use of nintedanib is restricted to the first 12 months of treatment (see Figure 12 in section 5.5.3) and that the hazard ratio is not time-invariant (Figure 19 in Appendix 7).

The CS[1] reports that median follow-up was 31.7 months at the time of the updated PFS analysis (15 February 2013), the results of which are summarised in Table 6.

Table 6 Updated analysis of centrally independently assessed PFS in in LUME-Lung 1 trial (February 2013)

Outcome	Nintedanib + docetaxel (median)*	Placebo + docetaxel (median)*	HR vs placebo arm (95% CI)†	p-value	Risk reduction
PFS in the overall trial population	3.5 months	2.7 months	0.85 (0.75 to 0.96)	0.0070	15%
PFS in adenocarcinoma population§	4.2 months	2.8 months	0.84 (0.71 to 1.00)	0.0485	16%

CI=confidence interval; HR=hazard ratio; PFS=progression-free survival
* Based on unadjusted Kaplan-Meier estimates for each treatment arm
† A proportional hazards model stratified by three factors (ECOG PS at baseline, presence of brain metastases at baseline, prior bevacizumab therapy) was used to derive the HR, 95% CI, and p-value (corresponding to the stratified log-rank test p-value)
§ Analysis conducted retrospectively
Source: Table 19 (page 89) of the CS[1]

The CS[1] states that the results obtained in the updated analysis support the findings from the primary PFS analysis. The ERG agrees that the results are consistent across both analyses as nintedanib plus docetaxel is shown to significantly improve PFS in comparison to placebo plus docetaxel in both the overall trial population and the adenocarcinoma population at the updated analysis.

Progression-free survival by local investigator review

The ERG notes that the PFS results as assessed by local investigator review were very similar to those obtained by central review. The treatment effect for the adenocarcinoma population significantly favoured nintedanib plus docetaxel over placebo plus docetaxel (HR 0.78, 95% CI: 0.62 to 0.97, p=0.0246).

Progression-free survival subgroup analyses

Subgroup analyses were performed at the time of the final OS analysis (15 February 2013). Results from the PFS (central review) subgroup analyses of baseline characteristics for adenocarcinoma patients are provided by the company in Figure 17 of the CS[1] (page101). The majority of pre-specified and post-hoc subgroup analyses show the effect of nintedanib plus docetaxel to be consistent with the treatment benefit observed in the primary analysis. The only exceptions to this are two subgroups (i) more than 9 months since start of first-line treatment and (ii) Asian region where there was a trend in favour of placebo plus docetaxel.

The results of tests for interaction were also provided to identify whether any subgroup of patients experienced a significantly greater treatment benefit than the remaining population. Significant interactions were observed for 'time since start of first-line therapy' (p=0.0032) and metastases in 'adrenal glands' (p=0.0336); these results suggest that patients who progressed within 9 months of first-line therapy, and those with metastases in the adrenal glands, experience a greater treatment effect than the remaining population.

Progression-free survival sensitivity analyses

Sensitivity analyses were only performed for PFS in the whole trial population, not only for those with adenocarcinoma.

Overall survival

Nintedanib plus docetaxel significantly improved median OS in comparison to placebo plus docetaxel in the population of adenocarcinoma patients who progressed within 9 months of first-line therapy (10.9 months vs 7.9 months respectively; HR 0.75, 95% CI: 0.60 to 0.92, p=0.0073). Therefore, analysis of OS in the population of interest, all adenocarcinoma patients, was permitted and the results are summarised in Table 7. Median OS was significantly longer with nintedanib plus docetaxel than with placebo plus docetaxel in the

adenocarcinoma population. However, the ERG is concerned that survival hazards appear not to be time invariant (see Figure 20, Appendix 7) and therefore may be misleading.

Table 7 OS in the adenocarcinoma population in LUME-Lung 1 trial (February 2013)

Outcome	Nintedanib + docetaxel (median)*	Placebo + docetaxel (median)*	HR vs placebo arm (95% CI)†	p-value
Overall survival	12.6 months	10.3 months	0.83 (0.70 to 0.99)	0.0359

CI=confidence interval; HR=hazard ratio; OS=overall survival
* Based on unadjusted Kaplan-Meier estimates for each treatment arm
† A proportional hazards model stratified by three factors (ECOG PS at baseline, presence of brain metastases at baseline, prior bevacizumab therapy) was used to derive the HR, 95% CI, and p-value (corresponding to the stratified log-rank test p-value)
Source: Table 20 of the CS[1]

Overall survival subgroup analyses

Subgroup analyses were performed at the time of the final OS analysis (15 February 2013) in the adenocarcinoma population.

Results from the pre-specified and post-hoc OS subgroup analyses of baseline characteristics for adenocarcinoma patients are provided by the company in Figure 18 of the CS[1] (page102). The subgroup analyses also show treatment effects in favour of nintedanib plus docetaxel, supporting the findings of the primary analysis. The only exceptions to this are two baseline characteristics: (i) presence of brain metastases and (ii) below stage IIIB disease at diagnosis. The company notes that a significant interaction was observed for 'best response to first-line treatment' (p=0.0766), indicating that patients whose best response to first-line therapy was PD would benefit more in terms of OS than the rest of the population. The ERG agrees with the company that this subgroup has a relatively small sample size (n=117) and the results should therefore be interpreted with caution.

The ERG is satisfied that all pre-specified subgroups were reported and show a consistent effect for OS across the majority of baseline characteristics.

Overall survival sensitivity analyses

The results of the two sensitivity analyses performed for OS in the adenocarcinoma population are presented in the text of the CS[1] and summarised here in Table 8.

Table 8 Sensitivity analyses of OS in the adenocarcinoma population in LUME-Lung 1 (February 2013)

Analysis	HR (95% CI)	p-value
Main OS analysis	0.83 (0.70 to 0.99)	0.0359
Sensitivity analysis 1 - Cox proportional hazards model with three of the stratification factors used at randomisation as covariates (ECOG PS at baseline, prior bevacizumab treatment, presence of brain metastases at baseline)	0.83 (0.70 to 0.98)	0.0295
Sensitivity analysis 2 - Model included the stratification factors and the baseline sum of the longest diameters (SLD) of the target lesions (mm) as covariates	0.81 (0.69 to 0.97)	0.0186

CI=confidence interval; HR=hazard ratio; OS=overall survival
Source: Text (page 99) of the CS[1] and Table 11.4.1.2.1.7: 2 of CTR[41]

The sensitivity analyses show that the results of the OS analysis remain very similar when including three of the stratification factors (ECOG PS at baseline, prior bevacizumab treatment and presence of brain metastases at baseline), or the stratification factors and baseline sum of the longest diameters of the target lesions as covariates in the model..

Tumour response based on central independent review

The results from the tumour response assessment (central independent review) are summarised in Table 9. No significant difference in ORR between nintedanib plus docetaxel patients and placebo plus docetaxel patients (4.7% vs 3.6%, odds ratio 1.32 [95% CI 0.61 to 2.93], p=0.4770) was observed. The ERG considers the ORRs in both arms to be lower than would be anticipated in typical clinical trials (see also section 4.4.5).

Table 9 Tumour response and disease control in the adenocarcinoma population in LUME-Lung 1 (February 2013)

Type of response (according to modified RECIST version 1.0 by central independent review)	Nintedanib + docetaxel (n=322)	Placebo + docetaxel (n=336)	Odds ratio* (95% CI)
Patients with objective tumour response, ORR [n (%)]	15 (4.7)	12 (3.6)	1.32 (0.61 to 2.93) p=0.4770
Complete response, n (%)	0	0	-
Partial response, n (%)	15 (4.7)	12 (3.6)	-
Unconfirmed complete/partial response n (%)	10 (3.1)	7 (2.1)	-
Median duration of confirmed objective response (months)	4.9	4.3	-
Median time to confirmed objective response (months)	1.6	5.1	-
Stable disease† n (%)	179 (55.6)	136 (40.5)	-
Patients with disease control§ n (%)	194 (60.2)	148 (44.0)	1.93 (1.42 to 2.64) p<0.0001
Median duration of disease control (months)	5.7	6.3	-
Progressive disease‡ n (%)	87 (27.0)	147 (43.8)	-
Other¥ n (%)	41 (12.7)	41 (12.2)	-

CI=confidence interval; ORR=overall response rate
* Odds ratios were obtained from logistic regression model adjusted for baseline ECOG PS
† stable disease was assumed if a follow-up imaging indicated stable disease at least once and at least 6 weeks after randomisation (i.e. at or after Day 43).
§ A patient was considered to have disease control if he/she had a best objective response of stable disease or better.
‡ Including patients with stable disease from a radiological imaging earlier than Day 43 followed by progressive disease
¥ Including patients with stable disease from a radiological imaging earlier than Day 43 followed by a non-evaluable response
Source: Table 21 of the CS[1]

4.4 Critique of the indirect evidence

4.4.1 Included studies in the MTC and statistical approach

Nine trials[24,55-62] were included in the review of the indirect evidence. The ERG did not identify any additional studies that met the company's eligibility criteria. However, not all nine studies were incorporated in any single MTC analysis. Four studies[48,56,59,62] were included in the base-case analyses, three[24,59,62] of which were also included in scenario analyses alongside a fifth study.[60] The remaining four studies[55,57,58,61] were only included in sensitivity analyses alongside those included in the base-case (sensitivity analyses i) or scenario analyses (sensitivity analyses ii); hence there were only ever a maximum of eight studies included in any given analysis.

The features of the types of analyses are as follows:

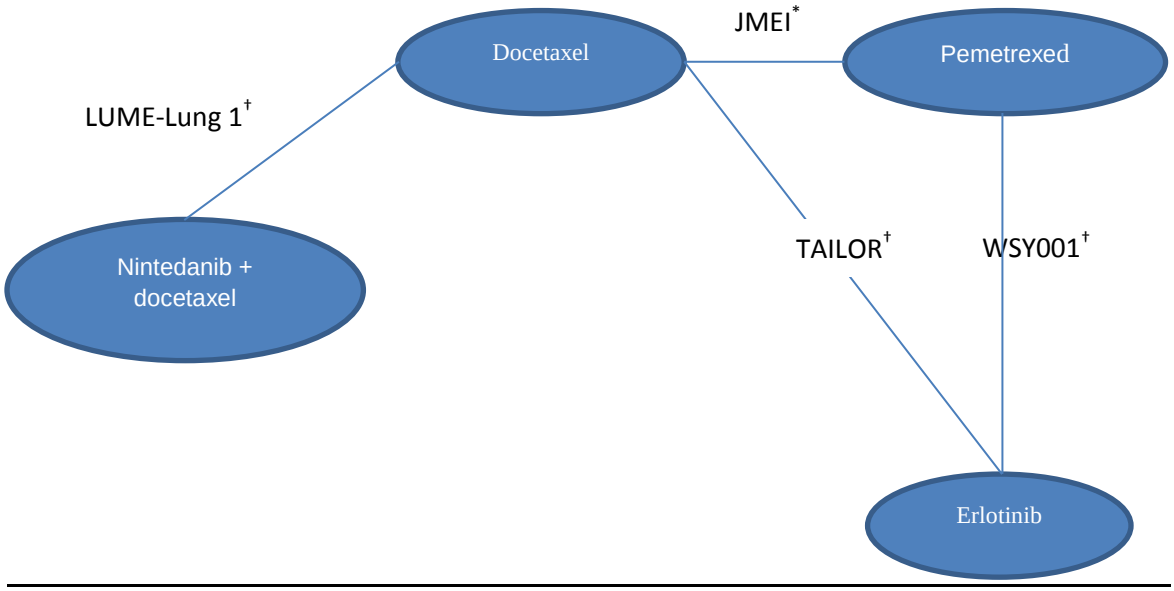
1. Base-case: includes all trials that meet eligibility criteria but excludes studies in which a high proportion (>20%) of patients have EGFR-positive adenocarcinoma and studies which include 'chemotherapy' as a single comparator where chemotherapy could be one or more possible regimens i.e. it must be possible to compare the intervention to all included comparators separately. The base-case analysis network diagram is reported in Figure 1.

2. Scenario analysis: assumes docetaxel and pemetrexed are of equal efficacy. The CS[1] states that this assumption was used to allow as many treatments to be compared with nintedanib plus docetaxel as possible. Hence the TITAN[60] study, excluded from the base-case, could be included in the scenario analysis because chemotherapy (docetaxel or pemetrexed) was the comparator. However the JMEI[56] study could not be included since this trial compared docetaxel to pemetrexed.

3. Sensitivity analyses: studies in which >20% of patients had EGFR-positive adenocarcinoma were also included in a MTC alongside

    i. the trials included in the base-case or

    ii. the trials included in the scenario analyses.



* Trial included only patients with adenocarcinoma
† Subgroup of patients with adenocarcinoma

Figure 1. Network diagram for MTC base-case analyses
Source: adapted from Figure 20 of the CS[1]

The company explains that the rationale for excluding patients with EGFR-positive adenocarcinoma from all but the sensitivity analyses was to enable a comparison between nintedanib plus docetaxel and other TKIs in a population similar to the patient population in LUME-Lung 1.[24] The majority of patients in LUME-Lung 1[24] would be expected to have EGFR-mutation negative adenocarcinoma [redacted] (see also section 4.3.3 and Table 14, section 4.4.3).

For each analysis, the company attempted to compare efficacy and safety. Efficacy outcomes were OS, PFS and ORR and safety outcomes were AEs for the following: fatigue, nausea and diarrhoea. However, for AEs, it was not possible to conduct a MTC for the base-case because none of the AE outcomes were reported in a sufficient number of trials in the base-case in order to be able to conduct a MTC.

The studies, comparators and analyses are summarised in Table 10 and Table 11.

Table 10 Studies included in the review of indirect evidence identified by the company

Trial name	Intervention	Comparator	Analyses included in
LUME-Lung 1[24]	Nintedanib + docetaxel	Placebo + docetaxel	Base-case, scenario and sensitivity
TAILOR[59]	Erlotinib	Docetaxel	Base-case, scenario and sensitivity
WSY001[62]	Erlotinib	Pemetrexed	Base-case, scenario and sensitivity
JMEI[56]	Pemetrexed	Docetaxel	Base-case and sensitivity
TITAN[60]	Erlotinib	Chemotherapy (docetaxel or pemetrexed)	Scenario and sensitivity
GEF-ERL[55]	Gefitinib	Erlotinib	Sensitivity
KCSG-LU08-01[57]	Gefitinib	Pemetrexed	Sensitivity
V-15-32[61]	Gefitinib	Docetaxel	Sensitivity
S103[58]	Pemetrexed + erlotinib	Pemetrexed or erlotinib	Sensitivity

Source: adapted from Figure 20 and Figure 21 of CS[1]

Table 11 Comparisons with nintedanib plus docetaxel in the MTCs undertaken by the company

Analyses	Comparators	Outcomes
Base-case	Docetaxel Erlotinib Pemetrexed	Overall survival Progression-free survival Overall response rate
Scenario	Chemotherapy (docetaxel and/or pemetrexed) Erlotinib	Overall survival Progression-free survival Overall response rate Safety
Sensitivity i	Docetaxel Pemetrexed Erlotinib Gefitinib Pemetrexed + erlotinib	Overall survival Progression-free survival Overall response rate
Sensitivity ii	Chemotherapy (docetaxel and/or pemetrexed) Erlotinib Gefitinib Pemetrexed + erlotinib	Overall survival Progression-free survival Overall response rate Safety

Sensitivity i: sensitivity analyses for base-case; sensitivity II; sensitivity analyses for scenario analyses
Source: adapted from Table 36 and Table 37 of CS[1]

For efficacy and safety outcomes, the company conducted MTCs and, where possible, Bucher indirect comparison results using the methods described in Appendix 6. The ERG is satisfied that the modelling approach was suitable. The ERG considers that conducting Bucher indirect comparisons is an effective method of assessing consistency within the network and therefore the reliability of the MTC results. If results from the MTC for any given comparison are considerably different to those obtained by the Bucher indirect comparison, it is likely that the MTC is not measuring the treatment effect accurately.

However, the ERG does not consider conducting any MTC was appropriate. There are multiple reasons for this, the primary reasons relating to the appropriateness of the MTC to the decision problem:

1. Erlotinib is not an appropriate comparator for the population of patients who would potentially be eligible to receive nintedanib plus docetaxel. As noted earlier in sections 0 and 3.3, the NICE Appraisal Committee's preliminary recommendations[21] are that erlotinib should not be recommended for treating locally advanced or metastatic NSCLC that has progressed after non-targeted chemotherapy in people with tumours that are EGFR-negative. Furthermore the characteristics of the vast majority of patients who are considered suitable for second-line erlotinib treatment are different from those who are considered suitable for second-line docetaxel treatment, most notably in terms of ECOG PS, EGFR mutation status and previous treatment received. Therefore the comparison with docetaxel is most appropriate and direct evidence for this is available from LUME-Lung 1.[24]

2. The ERG further observes that LUME-Lung 1[24] is the only trial in which any patients received pemetrexed as a first-line treatment as is now typically the case in clinical practice in England. In this trial, 19.1% of patients were previously treated with pemetrexed, more than in any other included trial.

In addition, there are methodological issues:

3. Although this is still an issue of some academic debate, the ERG considers that the proportional hazards assumption is not supported by LUME-Lung 1[24] trial data for PFS or OS. As LUME-Lung 1[24] is the only trial providing evidence for nintedanib plus docetaxel, any comparison also including evidence from this trial will incorporate this HR, which affects the robustness of these other comparisons. Thus any estimation of the relative effectiveness of nintedanib plus docetaxel vs erlotinib (i.e. a calculated HR) will lack credibility and be effectively meaningless. A full assessment of this issue is provided by the ERG in Appendix 7.

4. Differences in trial and patient characteristics (as described in detail in sections 4.3.2 and 4.3.3) suggest there is heterogeneity across trials which may mean MTCs are inappropriate:

    a. In the base-case analyses, while LUME-Lung 1[24] and TAILOR[59] both report similar median follow-up times, JMEI[56] and WSY001[62] report much shorter follow-up times. This heterogeneity may mean that the trials are too dissimilar

to allow a valid comparison of outcomes in an MTC. Additional sources of heterogeneity have also been identified in terms of differences in eligibility criteria across trials (see section 4.4.2) and participant characteristics (see section 4.3.3).

b. There also appears to be heterogeneity across the trials included in the scenario analyses.[24,59,60,62] TITAN[60] includes many more patients with ECOG PS 2 (20%) than would be expected in a patient population considered for treatment with nintedanib plus docetaxel. Furthermore, unlike any of the other trials in the base-case, TITAN[60] also permitted treatment crossover following disease progression. Hence the median OS for the chemotherapy arm may be inflated. It also compares erlotinib to chemotherapy in which chemotherapy consists of docetaxel or pemetrexed, thereby assuming the two treatments to be of equal efficacy. The ERG is not aware of any evidence that supports this assumption specifically in an adenocarcinoma population. Finally, by including this trial, the MTC is no longer making comparisons to docetaxel but to chemotherapy. However the chemotherapy arm includes pemetrexed which is not a second-line treatment option in England. Taking these factors into account, the ERG considers that the efficacy and safety results generated by the scenario analyses are neither relevant nor robust.

c. Trials included only in the sensitivity analyses[55,57,58,61] appear to be different to those in the base-case and scenario analyses, in particular these trials have high proportions of patients with EGFR-positive mutations and are based in Asia. Combining data from these trials with data from trials in the base-case and scenario analyses appears to be inappropriate as patients from Asia may have different tumour biology and comorbidities to those in the UK and EGFR mutation status is known to be related to the efficacy of some drugs. The ERG considers that the efficacy and safety results generated by the sensitivity analyses are not robust.

d. For the MTCs of safety outcomes, the company explains that due to low event rates, and the fact that only a small number of trials reported these outcomes, a network could only be formed when assuming equal tolerability of docetaxel and pemetrexed (scenario analysis). However, the findings from JMEI,[56] which compared these two drugs, albeit in a broader NSCLC population (52.9% had adenocarcinoma), reported differences between the two drugs, with a more favourable safety profile for pemetrexed. Therefore

this assumption does not hold. Furthermore, as identified above, the ERG considers there are differences in trial and patient characteristics between the trials included in the base-case, scenario and sensitivity analyses. The ERG considers that none of safety results generated by the MTC analyses are robust.

For all of the reasons outlined above, the ERG does not consider the comparison of nintedanib plus docetaxel with erlotinib is relevant to this STA.

4.4.2 Trial characteristics of included studies

The characteristics of trials included in the base-case and scenario analyses are summarised in Table 12 as well as the characteristics of those trials included in the sensitivity analyses only. The ERG notes that only TAILOR[59] was conducted solely in Europe whereas four trials[55,57,61,62] were conducted solely in Asia; three[55,57,61] of the Asian studies were included only in the sensitivity analyses. The location of trials is likely to be important because patients may have different tumour biology and comorbidities depending on their ethnic origin and where they live.

The company argues that all of the included trials had similar eligibility criteria. However, the ERG notes that there were some differences.

The ERG considers that two eligibility criteria (ECOG PS and complications such as brain metastases and pleural effusions) may be the main drivers of outcome in patients with adenocarcinoma. In the base-case and scenario analyses, only LUME-Lung 1[24] restricted trial entry to ECOG PS ≤1. Six trials explicitly stated they excluded patients with brain metastases: three in the base-case: LUME-Lung 1,[24] WSY001[62] and JMEI;[56] TITAN[60] in the scenario analyses and KCSG-LU08-01[57] and S103[58] in the sensitivity analyses. It is however noted by the ERG that LUME-Lung 1[24] excluded patients with active brain metastases and so this exclusion criterion may have enabled patients with a poorer prognosis to have been included than in the other trials in this respect. Two trials (LUME-Lung 1[24] and JMEI[56]), both in the base-case, excluded clinically significant or uncontrolled pleural effusions. Therefore patients in LUME-Lung 1[24] and JMEI[56] in particular may be expected to have slightly better prognoses than patients in the other trials although similar exclusion criteria may have been employed in the other trials but were not reported; for example TAILOR,[59] GEF-ERL[55] and V-15-32[61] reported only limited eligibility criteria. The ERG acknowledges that existence of brain metastases is a relatively common exclusion criteria for entry into trials. Nevertheless, such exclusions do result in a patient population different to those who would be treated in clinical practice. In this instance, because patients who receive nintedanib do so in

combination with docetaxel, the exclusion of patients with ECOG PS≥2 is however appropriate.

Additional eligibility criteria which could also impact on patient outcomes include EGFR mutation status, previous treatment and smoking status. WSY001[62] included only patients with EGFR wild type disease (EGFR-negative) whereas GEF-ERL[55] included only patients with EGFR activating mutations (EGFR-positive). Patients in the latter study would be expected to perform better when treated with a TKI or chemotherapy than patients in the former study. Furthermore, it should be noted that the majority of patients treated in clinical practice in England would be EGFR-negative. KCSG-LU08-01[57] and V-15-32[61] only permitted entry to never-smokers whereas the majority of patients with NSCLC treated in England are current or ex-smokers. Prior pemetrexed (or drugs directed at pemetrexed molecular targets) or TKIs were explicitly not permitted in three trials in the base-case (WSY001[62] and JMEI[56] and S103[58]), TITAN[60] in the scenario analyses and three trials (TAILOR,[59] GEF-ERL[55] and KCSG-LU08-01[57]) in the sensitivity analyses. These are potentially important exclusion criteria as not only may these affect outcomes but in clinical practice in England today, as noted in section 2.2, these are the first-line treatments of choice: pemetrexed for EGFR-negative disease and a TKI for EGFR-positive disease.

Alongside differences in eligibility criteria, the ERG also observes that in V-15-32,[61] docetaxel was administered every 3 weeks as a one-hour intravenous infusion of 60 mg/m^2 (the approved dosage in Japan). This trial,[61] alongside KCSG-LU08-01[57] and TITAN,[60] also permitted treatment crossover, unlike any of the other trials. This is an important consideration because treatment crossover could confound OS in these trials. Finally, it should also be noted that the median follow-up times varied considerably in the trials (range 7.5 to 33 months). This is important because if follow-up is not similar across trials, bias may be introduced into studies with shorter follow-up and less mature data as a result of increased censoring.

Table 12 Trial characteristics of trials included in only the MTC base-case and scenario analyses

Trial	Location	Inclusion criteria	Exclusion criteria	Median follow-up
LUME-Lung 1[24]	Europe, Asia, South Africa	• Histologically or cytologically confirmed stage IIIB-IV or recurrent NSCLC of any histology, following relapse or failure of one previous first-line chemotherapy (in the case of recurrent disease one additional previous regimen was allowed for adjuvant, neoadjuvant, or neoadjuvant + adjuvant therapy) • Life expectancy of ≥3 months • At least one target lesion measurable according to RECIST criteria • ECOG PS 0 to 1	• Prior docetaxel or VEGF/VEGFR inhibitor (other than bevacizumab) usage • Radiographic evidence of cavitary or necrotic tumours, centrally located tumours with radiographic evidence (CT or MRI) of local invasion of major blood vessels, or a recent history (<3 months) of clinically significant haemoptysis or a major thrombotic or clinically relevant major bleeding event in the past 6 months • Active brain metastases or leptomeningeal disease • Pre-existing ascites and/or clinically significant pleural effusion	31.7 months
TAILOR[59]	Italy	• Patients with wild-type EGFR advanced NSCLC, who had recurrence or progression after failing platinum-based chemotherapy • Adequate vital functions • ECOG PS ≤2	• Previous treatment with taxanes or anti-EGFR drugs or drugs directed at pemetrexed molecular targets (i.e., thymidylate synthase and dihydrofolate reductase inhibitors)	33 months
WSY001[62]	China	• Aged 18 to 75 years • Pathologically or cytologically confirmed stage IIIB or IV lung adenocarcinoma or postoperative recurrent lung adenocarcinoma incurable by surgery or radiotherapy within 6 months of neoadjuvant or adjuvant chemotherapy • EGFR wild-type and EGFR FISH-positive disease • Received 1 prior platinum-based chemotherapy (including neoadjuvant or adjuvant chemotherapy) • Adequate bone marrow function • Adequate liver function • Adequate renal function • Presence of 2-dimensional measurable disease • Life expectancy of ≥3 months • ECOG PS 0 to 2	• Prior treatment with TKI or pemetrexed • Symptomatic brain metastases • Prior malignant disease (except for basal cell carcinomas) • Pregnancy	14.7 months

Trial	Location	Inclusion criteria	Exclusion criteria	Median follow-up
JMEI[56]	Not reported	• Histologically or cytologically confirmed stage III or IV NSCLC not amendable to curative therapy • Received treatment with only one prior chemotherapy for advanced disease (one prior additional therapy allowed for neoadjuvant, adjuvant, or neoadjuvant + adjuvant therapy) • Adequate bone marrow function • Adequate hepatic function • Adequate renal function • ECOG PS 0 to 2	• Patients with prior docetaxel or pemetrexed treatment • CTCAE ≥grade 3 peripheral neuropathy • An inability to interrupt nonsteroidal anti-inflammatory drugs • Uncontrolled pleural effusions, symptomatic or uncontrolled brain metastases, or significant weight loss (≥ 10% body weight in the preceding 6 weeks) were ineligible.	7.5 months
TITAN[60]	International	• Histologically documented locally advanced, recurrent, or metastatic NSCLC • Disease progression while receiving four cycles of a standard first-line platinum-based chemotherapy doublet (representing a population with poor prognosis); patients who had disease progression during the four cycles of a standard platinum-based chemotherapy doublet could enrol once they had recovered from any toxic effects of the chemotherapy treatment • Adequate haematologica function • Adequate hepatic function • Adequate renal function • Ability to comply with study and follow-up procedures • ECOG PS 0 to 2	• Previous exposure to anti-human-EGFR-directed drugs or drugs directed at pemetrexed molecular targets (i.e., thymidylate synthase and dihydrofolate reductase inhibitors) • Prior chemotherapy or systemic anti-neoplastic therapy other than the permitted platinum-based regimens • Uncontrolled or untreated brain metastasis • Spinal cord compression or other malignancies within the past 5 years (except carcinoma in situ)	24.8 months (chemotherapy arm) 27.9 months (erlotinib arm)

CTCAE= Common Terminology Criteria for Adverse Events; ECOG=Eastern Cooperative Oncology Group; EGFR=epidermal growth factor receptor; FISH=fluorescence in situ hybridisation**;** NSCLC=non-small-cell lung cancer; PS=performance status; RECIST=Response Evaluation Criteria in Solid Tumors; TKI=tyrosine-kinase inhibitor; VEGFR=vascular endothelial growth factor receptor

Source: adapted from Tables 9, 26 and 36 of the CS[1] with additional criteria added from cited source publications (For JMEI[56] eligibility criteria were reported in Hanna *et al*[63])

Table 13 Trial characteristics of trials included in only the MTC sensitivity analyses

Trial	Location	Inclusion criteria	Exclusion criteria	Median follow-up
GEF-ERL[55]	South Korea	• Histologically confirmed stage IIIB or IV NSCLC including recurrent or metastatic disease following failure of first-line chemotherapy • WHO performance status of 0 to 2 • Presence of either an activating EGFR mutation, or two of three clinical factors associated with higher incidence of EGFR mutations. • Brain metastasis permitted if treated at least 4 weeks before entry and clinically stable without steroid treatment for 1 week	• Previous treatment with EGFR signalling inhibitors and radiation therapy within the preceding 4 weeks	16.3 months
KCSG-LU08-01[57]	Korea	• Histologically or cytologically confirmed pulmonary adenocarcinoma that progressed after just 1 previous platinum-based chemotherapy regimen for advanced disease (stage not reported) • Never-smoked (a total of ≤100 cigarettes in their lifetime) • ECOG PS 0 to 2	• Patients with prior TKI or pemetrexed treatment • Symptomatic or uncontrolled brain metastases were ineligible.	15.9 months
V-15-32[61]	Japan	• Histologically or cytologically confirmed stage IIIB or IV NSCLC not amenable to curative surgery or radiotherapy, or postoperative recurrent NSCLC • Failure of prior treatment with one or two chemotherapy regimens (≥1 platinum-based regimen) • WHO PS 0 to 2 • Protocol amendment allowed recruitment of patients without measurable lesions	• Not reported	21 months
S103[58]	Not reported	• Histologically or cytologically confirmed, locally advanced or metastatic non-squamous NSCLC following failure of first-line chemotherapy regimen • ECOG PS 0 to 2 • Only never-smoking patients (<100 lifetime cigarettes) were eligible.	• Prior exposure to agents directed at the human EGFR axis or at pemetrexed molecular targets (e.g. TS or DHFR inhibitors) • Brain metastasis (unless treated and stable after radiotherapy ≥2 weeks) • Concurrent administration of any other antitumour therapy.	14.7 months

DHFR= dihydrofolate reductase; ECOG=Eastern Cooperative Oncology Group; EGFR=epidermal growth factor receptor;; NSCLC=non-small-cell lung cancer; PS=performance status; TKI=tyrosine-kinase inhibitor; TS= thymidylate synthase; VEGFR=vascular endothelial growth factor receptor; WHO=World Health Organisation
Source: adapted from Tables 27 and 36 of the CS[1]

4.4.3 Participant characteristics of included studies

Baseline characteristics of patients summarised in the CS[1] are reported in Table 14 (for trials included in the base-case and scenario analyses) and Table 15 (for trials included in the sensitivity analyses). The ERG considers that the baseline characteristics that are the main drivers of outcomes in patients with adenocarcinoma are ECOG PS, response to prior therapy and EGFR mutation status.

Narratively, the company only focuses on ECOG PS, noting that in TITAN[60] and GEF-ERL,[55] which were included only in the scenario and/or sensitivity analyses, there were a higher proportion of patients (20.0% and 14.6% respectively) with ECOG PS 2 than any of the trials in the base-case. Because patients receive docetaxel with nintedanib, then the ERG considers that the vast majority of patients included in studies should all have ECOG PS≤1. While it is difficult to quantify the proportion, a minimum of 85% would seem reasonable. TITAN[60] (included only in the scenario analyses) does not meet this criterion.

The ERG notes that the proportion of patients with adenocarcinoma ranged from 50% to 100% in the base-case. The three studies (LUME-Lung 1,[24] JMEI[56] and TAILOR[59]) with <75% are appropriately included because they do report subgroup analyses for patients with adenocarcinoma. The ERG further notes that while LUME-Lung 1[24] included some patients with early stage disease at diagnosis, the majority (91.2%) of patients with adenocarcinoma had stage III/IV disease at diagnosis and even more (94.2%) had metastatic disease at screening. In WSY001[62] 71.5% were reported to have stage III/IV disease, the remainder (28.5%) described as having recurrent disease. In JMEI[56] all adenocarcinoma patients were reported to have stage III (18%) or IV (82%) disease at baseline. No information about staging is provided in TAILOR,[59] it being stated patients with metastatic disease were enrolled who “had recurrence or progression after failing platinum-based chemotherapy.”

Response to prior therapy differed across the trials. The proportion of patients with a complete or partial response or stable disease to previous treatment varied from 56.1% in WSY001[62] to 70.7% in JMEI;[56] in LUME-Lung 1[24] it was 70.7% and in TAILOR[59] was 63.9%. As noted in section 4.3.6, in LUME-Lung 1[24] a significant interaction was observed for ‘best response to first-line treatment’ indicating that patients whose best response to first-line therapy was PD would benefit more in terms of OS than the rest of the population. However, as noted by both the company and ERG, this subgroup has a relatively small sample size (n=117) and the results should be interpreted with caution.

With regard to other baseline characteristics, in studies included in the base-case, the ERG observes that median age varied from 54.3 to 60 years, proportion of females from 27.3% to 39.2%, patients with wild-type mutations (EGFR-negative) ranged from [redacted] to 100% and the proportion of never smokers from 17.4% to 35.7%. However, data were not presented for mutation status or never smokers for JMEI[56] and data were incomplete for mutation status for LUME-Lung 1;[24] it is assumed the majority of patients in both trials would be EGFR-negative, an assumption apparently supported by the limited data available from LUME-Lung 1.[24]

In some respects, the characteristics of TITAN,[60] which is included only in the scenario analyses, is like those included in the base-case. There were again a high proportion of patients with unknown EGFR status but it appears from the data available, if it is assumed the ratio of EGFR-positive to EGFR-negative patients in the patients with unknown mutation status is the same as that in the known mutation status, that the majority were EGFR-negative. Arguably what makes this trial most unlike those in the base-case, however, is the aforementioned higher proportion of patients with ECOG PS 2 suggesting a greater proportion of patients with more severe disease in this trial. The ERG further notes that all patients in this trial had stage IIIB (21.7%) or stage IV (88.3%) disease at baseline.

With regard to the trials included only in the sensitivity analyses, it is apparent from Table 15 that three trials appear similar to each other in most respects (GEF-ERL,[55] KCSG- LU08-01[57] and S103[58]) whereas the fourth (V-15-32[61]) appears to be different as it has fewer numbers of female patients, never smokers and patients with adenocarcinoma. The ERG notes that in all trials, EGFR-mutation status is only available from a minority of patients. If it is assumed the ratio of EGFR-positive to EGFR-negative patients in the patients with unknown mutation status is the same as that in the known mutation status patients, then the data appear to support the company's assertion that the proportion of patients with EGFR-positive disease ≥20%; indeed, in each trial there would be a majority of patients with EGFR-positive disease. All patients had stage IIIB and IV disease at baseline in S103[58] and KCSG-LU08-01.[57] In GEF-ERL[55] the proportion was 84.4% with the majority of other patients described as having recurrent disease (13.5%). In V-15-32[61] all patents had stage III/IV disease or were described as being recurrent (83.0% and 17.0% respectively).

Table 14 Patient characteristics of trials included in only the MTC base-case and scenario analyses

Trial and arm	Number at baseline	Adenocarcinoma		Age (years)	Wild-type mutations (EGFR-negative) (%)	ECOG PS 0 to 1 (%)	Female (%)	Never smokers (%)
		%	N					
LUME-Lung 1[24]	1314	50.1	658			100.0		
Nintedanib + docetaxel	655	49.2	322	Median: 60 Range: 53 to 67	[redacted]	100.0	27.3	35.7*
Placebo + docetaxel	659	51.0	336	Median: 60 Range: 54 to 66	[redacted]	100.0	27.3	34.2*
TAILOR[59]	219	69.4	152			92.7		
Erlotinib	109	63.3	69	Median: 66 Range: 40 to 81	100	93.6	29.4	17.4
Docetaxel	110	75.5	83	Median: 67 Range: 35 to 83	100	91.7	33.6	27.2
WSY001[62]	123	100	123			94.3		
Erlotinib	61	100	61	Median: 54.3 Range: 30 to 74	100	93.4	34.4	24.6
Pemetrexed	62	100	62	Median: 55.1 Range: 33 to 75	100	95.2	37.1	27.4
JMEI[56]	571	52.9	302			86.8*		
Pemetrexed	283	55.8	158	Median: 57.4* Range not reported	Not reported	84.8*	39.2*	Not reported
Docetaxel	288	50.0	144	Median: 56.7* Range not reported	Not reported	88.9*	34.0*	Not reported
TITAN[60]	424	49.5	201			80.0		
Erlotinib	203	47.3	96	Median: 59 years Range: 36 to 80 years	36.9 Indeterminate: 15.8 Missing: 43.3	80.8	20.7	14.8
Chemotherapy	221	51.6	114	Median: 59 years Range: 22 to 79 years	33.5 Indeterminate: 16.3 Missing: 45.7	79.2	27.6	19.9

* Subgroup of patients with adenocarcinoma only. [redacted]
Source: adapted from Table 26 with additional data on EGFR mutations and ECOG PS taken from the cited source publications

Table 15 Patient characteristics of trials included in only the MTC sensitivity analyses

Trial and arm	Number at baseline	Adenocarcinoma		Age (years)	Wild-type mutations (EGFR-negative) (%)	ECOG PS 0 to 1 (%)	Female (%)	Never smokers (%)
		%	N					
GEF-ERL[55]	96	90.6	87			85.4		
Gefitinib	48	91.7	44	Median: 60 Range: 37 to 83	25.0 Missing: 56.3	85.4	85.4	91.7
Erlotinib	48	89.6	43	Median: 56 Range: 32 to 81	41.7 Missing: 41.7	85.4	85.4	95.8
KCSG-LU08-01[57]	135†	100.0	135			91.1		
Gefitinib	68†	100.0	68	Median: 58 Range: 40 to 77	22.1 Missing: 50.0	91.2	85.3	100.0
Pemetrexed	67†	100.0	67	Median: 64 Range: 30 to 78	23.9 Missing: 44.8	91.0	85.1	100.0
V-15-32[61]	489¥	77.7	380		5.3 Missing: 88.3	95.7		
Gefitinib	244¥	78.4	191	≤64 years: 56.3		95.5	38.4	29.0
Docetaxel	239¥	77.0	184	≤64 years: 55.3		95.9	38.1	35.7
S103[58]	240	93.8	225		7.9 Missing: 82.1	92.9		
Erlotinib + pemetrexed	78	92.3	72	Median: 55.8 Range not reported		91.0	74.4	100.0
Erlotinib	82	92.7	76	Median: 53.9 Range not reported		92.7	65.9	100.0
Pemetrexed	80	96.3	77	Median: 55.9 Range not reported		95.0	56.3	100.0

Source: adapted from Table 27 with additional data on EGFR mutations and ECOG PS taken from the cited source publications
† Population analysed for safety and efficacy analyses
¥ Population evaluated for safety (described as intention-to-treat population in source paper)

4.4.4 Risk of bias

The company conducted an assessment of the risk of bias of the studies included in the base-case MTC, the results are presented in the CS[1] and shown in Table 16. The ERG considers that the conclusions drawn by the company are valid and that the included studies have an overall low risk of bias.

Table 16 Company's assessment of risk of bias for trials included only in the MTC base-case analyses

Criteria	LUME-Lung 1[24]	TAILOR[59]	WSY001[62]	JMEI[56]
1. Was randomisation carried out appropriately?	Yes	Yes	Yes	Yes
2. Was the concealment of treatment allocation adequate?	Yes	Yes	Not clear	Not clear
3. Were the groups similar at the outset of the study in terms of prognostic factors?	Yes	Yes	Yes	Yes
4. Were the care providers, participants and outcome assessors blind to treatment allocation?	Yes	Yes	No	Not clear
5. Were there any unexpected imbalances in drop-outs between groups?	Not clear	No	No	No
6. Is there any evidence to suggest that the authors measured more outcomes than they reported?	No	No	No	No
7. 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	Yes	Yes	Yes
Overall quality (" + + ", " + ", "-")	**+ +**	**+ +**	**+ +**	**+ +**

Source: Appendix 5 (Table 155) of the CS[1]
Information for LUME-Lung 1[24] taken from trial protocol[53]

Trials in the scenario and sensitivity analyses were also assessed for risk of bias (Appendix 5 of the CS[1]). These were considered to be of similarly low risk of bias with the exception of KCSG-LU08-01[57] which was deemed to be at higher risk of bias because of unexpected imbalances in drop-outs between treatment arms.

4.4.5 Individual study findings

Efficacy results from the studies included in the base-case analyses are provided in Table 17. The findings from the studies included in the scenario and sensitivity analyses are not presented here because the patient characteristics of these trials are considered by the ERG to be too different to those of the patient population relevant to the decision problem.

Significant improvements in OS were reported for nintedanib plus docetaxel compared to placebo plus docetaxel in LUME-Lung 1[24] and erlotinib vs pemetrexed in TAILOR.[59] Significant improvements in PFS were only reported in LUME-Lung 1[24] for nintedanib plus docetaxel vs placebo plus docetaxel. In patients treated with adenocarcinoma, median OS varied from 9.2 months (adjusted OS in JMEI[56]) to 13.4 months (pemetrexed arm of WSY001[62]); the OS for nintedanib plus docetaxel therefore appears to compare favourably in LUME-Lung 1[24] (12.6 months). Median PFS ranged from 2.8 months (placebo plus docetaxel arm in LUME-Lung 1[24] to 4.2 months (nintedanib plus docetaxel arm in LUME-Lung 1[24]).

Although median OS was not presented for the adenocarcinoma population in TAILOR,[59] the ERG notes that for the overall population median OS was 8.2 months in the erlotinib arm as compared to 11.7 months for erlotinib in WSY001[62] in which all patients had adenocarcinoma. The median OS for docetaxel in the overall population of TAILOR,[59] was 5.4 months and was lower than the adjusted median OS reported for the adenocarcinoma subgroup of patients treated with docetaxel in JMEI[56] (9.2 months) and OS for the placebo plus docetaxel arm in the adenocarcinoma subgroup of LUME-Lung 1[24] (10.3 months). The median PFS for the erlotinib arm in the overall population in TAILOR[59] (2.9 months) was also slightly lower than for the erlotinib arm in WSY001[62] (4.1 months). Median PFS for the docetaxel arm in the overall population in TAILOR,[59] (2.4 months) was however similar to that of the placebo plus docetaxel arm of adenocarcinoma patients in LUME-Lung 1[24] (2.8 months) and slightly less than the adjusted PFS in the docetaxel arm of JMEI[56] (3.5 months). These findings may be indicative that the trials included different patient populations, as suggested by the ERG in 4.4.1.

Response rates were only reported for three of the trials.[24,56,62] The ERG notes that the ORR for patients treated with docetaxel in JMEI[56] (9.9%) was much greater than reported for placebo plus docetaxel in LUME-Lung 1[24] (3.6%). For patients treated with pemetrexed it was also higher (12.8%) in JMEI[56] than in WSY001[62] (8.1%). The highest ORR was reported for erlotinib (19.7%) in WSY001.[62] The findings for ORR were lowest for either arm in LUME-Lung 1.[24]

Table 17 Individual study findings (inputted into the MTC base-case analyses by the company)

Outcomes		LUME-Lung 1[24]‡		TAILOR[59]		WSY001[62]		JMEI[56]	
		Nintedanib + docetaxel	Placebo + docetaxel	Erlotinib	Docetaxel	Erlotinib	Pemetrexed	Pemetrexed	Docetaxel
N efficacy		322	336	69	83	61	62	158	144
Unadjusted OS	Months	12.6	10.3	NR¥	NR¥	11.7	13.4	NR	NR
	HR (95% CI) p-value	0.83 (0.7 to 0.99) p=0.0359		0.67 (0.48 to 0.95); reported as significant		1.01 (0.66 to 1.54) p= 0.97		NR	
Adjusted OS†	Months	NR	NR	NR	NR	NR	NR	9.0	9.2
	HR (95% CI) p-value	0.81 (0.69 to 0.97) p= 0.0186 (two-sided)		NR		NR		0.92 (0.69 to 1.22) p=0.551	
Unadjusted PFS	Months	4.0	2.8	NR¥	NR¥	4.1	3.9	NR	NR
	HR (95% CI) p-value	0.77(0.62 to 0.96) p= 0.0193		0.76 (0.54 to 1.05)		0.92 (0.62 to 1.37) p= 0.683		NR	
Adjusted PFS†	Months	4.2	2.8	NR	NR	NR	NR	3.5	3.5
	HR (95% CI) p-value	0.84 (0.71 to 1) p= 0.0485 (two-sided)		NR		NR		0.83 (0.65 to 1.06) p= 0.135	
Response	Criteria	RECIST		NR		RECIST		Southwest Oncology Group Criteria	
Objective response	Definition	Objective tumour response (CR + PR)		NR		PR + CR		CR, PR*	
ORR	N evaluated	322	336	NR	NR	61	62	158	144
	N	15	12	NR	NR	12	5		
	%	4.7	3.6	NR	NR	19.7	8.1	12.8	9.9

CI=confidence interval; CR=complete response; HR=hazard ratio; NR=not reported; ORR=overall response rate; OS=overall survival; PFS=progression-free survival; PR=partial response
† No study reported what variables were adjusted for except LUME-Lung 1[24] for OS: [redacted]
‡ For the LUME-Lung 1[24] trial adjusted OS, PFS, and ORR data for the adenocarcinoma subgroup are available from CTR[41]
¥ For TAILOR, median OS and median PFS are only reported for overall population, not adenocarcinoma subgroup
* Complete response: complete disappearance of all measurable and evaluable disease; Partial response: ≥50% decrease in the sum of products of perpendicular diameters of all measurable lesions
Source: adapted from Table 32 of the CS[1] with additional data taken from the source papers

4.4.6 Results from mixed treatment comparisons

As noted in sections 2.2, 3.3 and 4.4.1 above, the ERG does not consider a comparison of nintedanib plus docetaxel to erlotinib is appropriate to decision problem. Furthermore, the ERG also considers there are a number of methodological issues with the MTC and taken together, the ERG does not therefore consider a comparison of nintedanib plus docetaxel with erlotinib is relevant to this STA. Nevertheless, for completeness, a brief description of the results and critique follows.

The following analyses were conducted by the company:

- Base-case analyses
- Sensitivity analyses for base-case (sensitivity analysis i)
- Scenario analyses
- Sensitivity analyses for scenario analyses (sensitivity analysis ii)

While only comparisons of nintedanib plus docetaxel to docetaxel and erlotinib are considered relevant to the NICE scope, some results are presented relative to other comparators included in the MTCs for completeness.

Summary of company's results: overall survival

The results from the base-case analysis for OS are presented in Table 18 and the probabilities of each treatment being the best at improving OS are presented in Table 19. The results from the base-case analysis suggest that nintedanib plus docetaxel is significantly more effective than either docetaxel alone or erlotinib alone. Results from the Bucher indirect comparisons support the findings from the MTC. Nintedanib plus docetaxel is most likely to be the best treatment, suggesting superiority over docetaxel and erlotinib.

Table 18 Summary of OS findings from MTC base-case analysis

Treatment	HR (95% CI) to fixed-effects
Nintedanib + docetaxel vs docetaxel	
Result from MTC	0.83 (0.70 to 0.99)
Result from Bucher indirect comparison	Not applicable
Nintedanib + docetaxel vs pemetrexed	
Result from MTC	0.82 (0.60 to 1.11)
Result from Bucher indirect comparison	0.90 (0.65 to 1.26)
Nintedanib + docetaxel vs erlotinib	
Result from MTC	0.64 (0.46 to 0.90)
Result from Bucher indirect comparison	0.56 (0.38 to 0.82)
Deviance information criterion	0.4095

OS=overall survival; HR=hazard ratio; CI=confidence interval
Notes: The results from the base-case analysis do not feature the random-effect model as there were no instances of two trials with the same comparison
Source: adapted from Table 38 of the CS[1]

Table 19 Probabilities of each treatment being the best at improving OS in base-case analysis

Treatment	Probability of being best
Nintedanib + docetaxel	70.44%
Docetaxel	9.81%
Pemetrexed	16.42%
Erlotinib	3.33%

Source: adapted from Table 39 of the CS[1]

The findings from the scenario and sensitivity analyses broadly support those of the base-case analyses. Nintedanib plus docetaxel also had the highest probability of being the best treatment in the sensitivity analysis (i) of the base-case (49.2%), followed by erlotinib plus pemetrexed (37.17%), a comparator that was not included in the original base-case. In the scenario analysis that assumes equivalent efficacy of docetaxel and pemetrexed, nintedanib plus docetaxel also had the highest probability of being the most effective treatment (78.95%). In the sensitivity analysis (ii) for the scenario analysis, erlotinib plus pemetrexed (54.39%), a comparator that was not included in the base-case analysis, had the highest probability of being the most effective, followed by nintedanib plus docetaxel (34.21%).

Summary of company's results: progression-free analyses

The results from the base-case analysis for PFS are presented in Table 20. The probabilities of each treatment being the best at improving PFS are presented in Table 21. The results suggest that nintedanib plus docetaxel significantly improves PFS in comparison to docetaxel and erlotinib. Results from the Bucher indirect comparisons support the findings from the MTC. Nintedanib plus docetaxel was most likely to be the best treatment, suggesting superiority over docetaxel and erlotinib.

Table 20 Summary of PFS findings from MTC base-case analysis

Treatment	HR (95% CI) to fixed-effects
Nintedanib + docetaxel vs docetaxel	
Result from MTC	0.77 (0.62 to 0.96)
Result from Bucher indirect comparison	Not applicable
Nintedanib + docetaxel vs pemetrexed	
Result from MTC	0.84 (0.61 to 1.15)
Result from Bucher indirect comparison	0.93 (0.67 to 1.29)
Nintedanib + docetaxel vs erlotinib	
Result from MTC	0.70 (0.50 to 1.00)¥
Result from Bucher indirect comparison	0.58 (0.39 to 0.87)
Deviance information criterion	1.568

PFS=progression-free survival; HR=hazard ratio; CI=confidence interval
The results from the base-case analysis do not feature the random-effect model as there were no instances of two trials with the same comparison
¥ The estimate for the upper bound of the 95% credible interval was 0.9958, making the result statistically significant
Source: adapted from Table 40 of the CS[1]

Table 21 Probabilities of each treatment being the best at improving PFS in base-case analysis

Treatment	Probability of being best
Nintedanib + docetaxel	69.69%
Docetaxel	5.01%
Pemetrexed	18.53%
Erlotinib	6.77%

Source: adapted from Table 41 of the CS[1]

The findings from the scenario and sensitivity analyses broadly support those of the base-case analyses although not to the same extent as for the OS analyses. Erlotinib plus pemetrexed, a comparator not in the original base-case, had the highest probability of being the best treatment in the sensitivity analysis (i) of the base-case (61.99%), followed by nintedanib plus docetaxel (25.01%). In the scenario analysis that assumes equivalent efficacy of docetaxel and pemetrexed, nintedanib plus docetaxel had the highest probability of being the most effective treatment (83.57%). In the sensitivity analysis (ii) for the scenario analysis, erlotinib plus pemetrexed (72.23%), a comparator that was not included in the base-case analysis, had the highest probability of being the most effective, followed by nintedanib plus docetaxel (16.42%).

Summary of company's results: overall response rate

Table 22 shows the results of the base-case analysis for ORR. The results suggest that there was no significant difference in ORR between nintedanib plus docetaxel in comparison with docetaxel or erlotinib.

Table 22 Summary of ORR findings from MTC base-case analysis

Treatment	HR (95% CI) to fixed-effects
Nintedanib + docetaxel vs docetaxel	
Result from MTC	1.33 (0.61 to 2.95)
Result from Bucher indirect comparison	Not applicable
Nintedanib + docetaxel vs pemetrexed	
Result from MTC	0.98 (0.33 to 2.84)
Result from Bucher indirect comparison	0.98 (0.34 to 2.83)
Nintedanib + docetaxel vs erlotinib	
Result from MTC	0.33 (0.07 to 1.56)
Result from Bucher indirect comparison	Not applicable
Deviance information criterion	37.47

CI=confidence interval; ORR=overall response rate
The results from the base-case analysis do not feature the random-effect model as there were no instances of two trials with the same comparison.
Source: adapted from Table 42 of the CS[1]

In the sensitivity analysis (i) for the base-case, nintedanib plus docetaxel was statistically inferior to erlotinib, gefitinib and erlotinib plus pemetrexed using a fixed-effects model. The findings from the random-effects model also suggest nintedanib plus docetaxel to be inferior although the wider confidence intervals mean that the difference is no longer statistically significant. The scenario analysis found nintedanib plus docetaxel shows no significant difference in ORR compared with chemotherapy (docetaxel or pemetrexed) or erlotinib. The sensitivity analysis (ii) for the scenario analysis found nintedanib plus docetaxel was not significantly different from chemotherapy (docetaxel or pemetrexed) or erlotinib but was significantly inferior to gefitinib and erlotinib plus pemetrexed.

Summary of company's results: adverse events

The safety outcomes of any grade fatigue, nausea and diarrhoea were only able to be analysed as part of the sensitivity analysis where docetaxel and pemetrexed were assumed to be of comparable efficacy. Although LUME-Lung 1[24] reported additional AEs, including CTCAE grade ≥3 fatigue and nausea, these outcomes could not be compared as either no other linked trial reported equivalent data, or the event rates in one or more of the treatment arms were zero.

ERG critique of the company's results from the mixed treatment comparisons

If the problems with the appropriateness and conduct of the MTCs highlighted in section 4.4.1 are ignored, the ERG makes a number of further observations in relation to the findings reported from the MTCs:

1. Results from the Bucher indirect comparisons support the findings from the MTCs suggesting that inconsistency in the network is not a concern, as additional evidence

from the wider treatment networks corroborate the evidence from simple indirect comparisons.

2. It is stated that unadjusted data were used wherever possible, although only adjusted data were available for JMEI.[56] This trial did not specify the variables which were adjusted for and this lack of information makes it difficult to assess the impact that these adjustments may have had on the data, and therefore the results of the MTCs.

3. Data used to derive results for PFS was PFS assessed by central independent review for LUME-Lung 1,[24] whereas for JMEI[56] and TAILOR[59] local investigator assessed PFS data were used; it is unclear whether the results used in the MTC from WSY001[62] were the results from central independent review or local investigator assessment. However, considering the similarities in the findings from central independent review (HR 0.77, 95% CI: 0.62 to 0.96) and local investigator assessment (HR 0.78, 95% CI: 0.62 to 0.97) for LUME-Lung 1,[24] it seems unlikely that this would greatly impact the results of the MTC.

4. The ERG observes that the company inputted data from the primary PFS analysis for LUME-Lung 1[24] into the MTC; it would have been more informative to use the data from the updated analysis (HR 0.84, 95% CI: 0.71 to 1.00).

5. The company states that trials which provided the active treatment arm with placebo versions of the comparator were not distinguished from trials which did not provide a placebo. The ERG does not consider this to be of major concern, as although the one trial (LUME-Lung 1[24]) which provided a placebo is less likely to be at risk of bias (see also section 4.4.4 [risk of bias]), it is unlikely that this difference would introduce a significant amount of heterogeneity between trials. The ERG notes that due to the small number of studies, a comparison between the fixed and random-effects models to test for heterogeneity could only be conducted for the sensitivity analyses for both OS and PFS. The base-case analysis showed some inconsistency for both OS and PFS effect sizes when direct and indirect evidence was compared. The company suggests that this may be due to differences in EGFR mutation status across studies. The ERG agrees with this assessment and believes the inconsistency may also be caused by differences in patient populations as discussed in sections 4.3.2 and 4.3.3.

The ERG's critique of AEs, including consideration of the evidence input into and derived from the MTC, is presented in section 4.5.

4.5 Critique of the adverse events data

Comparison of adverse events from the direct evidence

In LUME-Lung 1[24] AEs were collected for the full trial population and the subgroup of patients with a histology of adenocarcinoma. In the CS[1] AEs are appropriately only presented for the adenocarcinoma subgroup since this is the population that is relevant to the decision problem.

The company reports that treatment with nintedanib plus docetaxel resulted in additional AEs compared with docetaxel treatment alone. Indeed, drug-related AEs reported in Table 59 of the CS[1] were 81.3% in the nintedanib plus docetaxel arm compared to 72.4% in the placebo plus docetaxel arm. However, the company argues that these data must be considered in the context of there being longer median treatment duration in the nintedanib plus docetaxel arm (Table 23). The ERG notes that in clinical practice in England, the maximum number of docetaxel cycles is likely to be four but notes the median number in the nintedanib plus docetaxel arm was five. It further notes that in both arms, the maximum number of cycles exceeded 40.

Table 23 Treatment exposure in the adenocarcinoma population in LUME-Lung 1

Length of treatment and dose intensity	Nintedanib + docetaxel	Placebo + docetaxel
Median duration of nintedanib/placebo treatment (range)	4.2 months (0.10 to 41.53)	3.0 months (0.07 to 31.10)
Mean dose intensity of nintedanib/placebo (%, SD)	91.2 (15.0)	93.8 (13.3)
Number of docetaxel courses (median, range)	5.0 (1 to 45)	4.0 (1 to 41)
Mean overall dose intensity of docetaxel (%, SD)	98.1 (4.5)	98.7 (3.7)

Source: Table 53 of the CS[1]

The most common specific types of AEs reported by adenocarcinoma patients in LUME-Lung 1[24] are summarised in Table 24. Types of AEs reported by patients in the nintedinab plus docetaxel arm included diarrhoea (43.4%), nausea (28.4%) and vomiting (19.4%) which the company states were successfully managed by dose reduction, dose interruption and/or symptomatic treatment and led to permanent nintedanib discontinuation in <1% of patients (Table 25). These were identified as AESIs relating to nintedanib by the company. Other reported AESIs associated with nintedanib treatment included ALT/AST increase (37.8% vs 9.3% and 30.3% vs 7.2% respectively, Table 24) which were reported to be generally reversible and led to permanent nintedanib discontinuation in <2% of patients (Table 25). For the majority of patients with adenocarcinoma requiring a dose reduction to manage AEs, a single dose reduction of nintedanib or placebo was sufficient (Table 25).

Table 24 Proportion of types of AEs in the adenocarcinoma population in LUME-Lung 1

AEs	Nintedanib + docetaxel		Placebo + docetaxel	
	% any CTCAE grade (% CTCAE grade ≥3)		% any CTCAE grade (% CTCAE grade ≥3)	
All AEs	96.3	(75.9)	94.3	(68.5)
Occurring in ≥5% in either arm:				
• Diarrhoea	43.4	(6.3)	24.6	(3.6)
• Neutrophil count decrease	40.9	(36.3)	40.5	(34.8)
• ALT increased	37.8	(11.6)	9.3	(0.9)
• Fatigue	30.9	(4.7)	29.4	(4.2)
• AST increased	30.3	(4.1)	7.2	(0.6)
• Nausea	28.4	(0.9)	17.7	(0.6)
• WBC decreased	27.8	(19.7)	28.2	(18.3)
• Decreased appetite	23.4	(1.3)	15.6	(1.5)
• Vomiting	19.4	(1.3)	12.3	(0.6)
• Alopecia	17.5	(0.3)	20.4	(0)
• Dyspnoea	16.9	(4.7)	15.6	(6.0)
• Neutropenia	13.8	(11.9)	15.3	(13.5)
• Cough	13.1	(0.9)	18.9	(0.6)
• Pyrexia	12.2	(0.6)	14.1	(0.3)
• Stomatitis	11.3	(1.3)	7.8	(0.3)
• Haemoglobin decreased	10.9	(0.9)	13.8	(2.1)
• Constipation	6.9	(0)	11.7	(0.3)
SAEs	34.7	(31.3)	32.1	(26.6)
Occurring in ≥5% in either arm:				
• Febrile neutropenia	5.6	(5.6)	1.8	(1.8)
• Malignant neoplasm progression	3.8	(3.8)	2.4	(2.1)
• Dyspnoea	2.8	(2.5)	5.4	(4.8)
• Pneumonia	2.8	(2.2)	3.6	(1.8)
• Diarrhoea	1.9	(1.6)	2.1	(1.8)
• General physical health deterioration	1.9	(1.9)	1.5	(1.2)
• Neutropenia	1.9	(1.6)	3.3	(3.3)
• Asthenia	1.6	(1.3)	0.6	(0.3)
• Respiratory failure	1.6	(1.6)	0.3	(0.3)
• Vomiting	1.6	(0.6)	1.2	(0.6)
• Atrial fibrillation	1.3	(0.9)	0	(0)
• Chest pain	1.3	(0.9)	1.8	(1.5)
• Pleural effusion	1.3	(1.3)	1.8	(1.2)
• Sepsis	1.3	(1.3)	0.3	(0.3)
• Pyrexia	0.6	(0)	1.2	(0)

AEs=adverse events; ALT=Alanine aminotransferase; AST=Aspartate transaminase; CTCAE= Common Terminology Criteria for Adverse Events; WBC=white blood cell
* As judged by the local investigator
Source: adapted from Tables 62 and 63 of the CS[1]

Table 25 AEs leading to dose interruptions, reductions or discontinuations in LUME-Lung 1

AEs	Nintedanib + docetaxel	Placebo + docetaxel
	%	%
At least 1 temporary interruption of nintedanib/placebo	52.2	41.4
At least 1 temporary interruption of nintedanib/placebo >14 consecutive days	10.0	6.6
1 dose reduction of nintedanib/placebo	17.2	6.6
2 dose reductions of nintedanib/placebo	4.7	0
AEs leading to dose reduction of nintedanib or placebo	21.6	6.6
AEs leading to dose reduction of nintedanib or placebo occurring in ≥1% in either arm:		
• Diarrhoea	8.1	3.3
• ALT increased	7.8	0.6
• AST increased	3.8	0
• Vomiting	2.2	0.6
• Nausea	1.3	0.3
AEs leading to dose reduction of docetaxel	16.6	12.3
AEs leading to permanent discontinuation of last study medication	20.9	17.7
AEs leading to permanent discontinuation of last study treatment occurring in ≥1% in either arm:		
• ALT increased	1.6	0
• Malignant neoplasm progression	1.6	1.5
• AST increased	1.3	0.3
• Dyspnoea	1.3	3.3

AEs=adverse events; ALT=Alanine aminotransferase; AST=Aspartate transaminase
Source: adapted from Tables 54, 55, 56, 58 and 59 of the CS[1]

The ERG notes from Table 24 that the incidence of AEs and SAEs was similar between treatment arms but the incidence of grade ≥3 AEs and SAEs was greater in the nintedanib plus docetaxel arm. The ERG notes grade 3 AEs tend to be particularly significant and can lead to drug discontinuation and hospitalisation but grade 2 AEs may also be clinically relevant by also impacting negatively on HRQoL. It is further noted that dose reduction schemes for nintedanib/placebo specified in Table 11 (pages 63 to 64) of the CS[1] included grade 2 AEs, namely vomiting of CTCAE grade ≥2 within 3 days after docetaxel therapy, diarrhoea of CTCAE grade 2 for >7 consecutive days and AST or ALT elevations of CTCAE grade 2 in conjunction with bilirubin elevations of CTCAE grade ≥1, or AST or ALT elevations of CTCAE grade ≥3. CTCAE grade 2 diarrhoea was included as an AE in its economic model (see Table 35 in section 5.4.7). From the CTR[41] (page 332, Table 12.2.2.4.1.2: 2) the ERG observes 17.8% of patients in the nintedanib plus docetaxel arm reported CTCAE grade 2 diarrhoea compared to 7.2% in the placebo plus docetaxel arm; CTCAE grade 2 + was 24.0% and 10.8% respectively.

Aside from AESIs related to nintedanib, a number of other AESIs were also identified and reported in the CS.[1] These were generally balanced across treatment arms. Exceptions identified by the company were:

- AESIs related to VEGFR inhibitor class effects: a higher frequency of any CTCAE grade hypertension in the nintedanib plus docetaxel arm (3.4% vs 0.6%). However, the incidence of CTCAE grade ≥3 hypertension was balanced across arms (0.9% vs 0.6%)
- AESIs based on potential associations/complications of AEs: any CTCAE grade dehydration only occurred in the nintedanib plus docetaxel arm (1.9% any CTCAE grade and 0.6 % CTCAE grade ≥3)
- AESIs related to potential interaction with concomitant chemotherapy: mucositis was more common in the nintedanib plus docetaxel arm (16.6%) than the placeboplus docetaxel arm (11.4%); however, the incidence of grade ≥3 mucositis was balanced across arms (1.3 % vs 0.6% respectively)
- AESIs selected based on competitor labelling: any CTCAE grade cutaneous skin reactions and any CTCAE grade rash were more common in the nintedanib plus docetaxel arm than placebo plus docetaxel arm (15.6% vs 10.5% and 12.5% vs 8.7% respectively; the incidence of both grade ≥3 cutanous skin reactions and grade ≥3 rash was however balanced across arms (1.3 % vs 0.6% for cutaneous skin reactions and 0.3% and 0% for rash)
- AESIs related to cardiac events: any CTCAE grade cardiac arrhythmias occurred at a slightly higher incidence in the nintedanib plus docetaxel arm (11.6%) compared with the placebo plus docetaxel arm (7.5%); however, the incidence of grade ≥3 cardiac arrhythmias was balanced across arms (2.2 % vs 1.5% respectively).

Other AEs identified as AESIs by the company were interstitial lung disease, photosensitivity conditions and anaphylactic reaction. Frequencies of these AESIs were uncommon (1.3%, 0.3% and 0 respectively for nintedanib plus docetaxel compared to 0.3%, 0.6% and 0.3% respectively in the placebo plus docetaxel arm). All were CTCAE grade <3 except for interstitial lung disease (0.3%) and anaphylactic reaction (0.3%) in the placebo plus docetaxel arm.

The AEs reported in LUME-Lung 1[24] which are of greatest concern, are fatal AEs where some imbalances were reported between treatment arms, fatal AEs being more common in the nintedanib plus docetaxel arm (Table 26). The only exception was fatal AEs occurring within 6 weeks of treatment which the company argues were well-balanced "indicating that the combination therapy with nintedanib and docetaxel had no acute toxicity[(42)]" (page 154 of the CS[1])

Table 26 Summary of fatal AEs in LUME-Lung 1 in the adenocarcinoma population

Fatal adverse events (AEs)	Nintedanib + docetaxel	Placebo + docetaxel
	n (%)	n (%)
All fatal AEs	56 (17.5)	32 (9.6)
• Fatal AEs occurring within 6 weeks	13 (4.0)	12 (3.6)
• Fatal AEs not attributed to progressive disease	20 (6.3)	8 (2.4)
• Fatal AEs attributed to progressive disease*	36 (11.3)	24 (7.2)
Drug-related fatal AEs	6 (1.9)	1 (0.3)
• Sepsis	2 (0.6)	0
• Dehydration	1 (0.3)	0
• Diverticulum intestinale†	1 (0.3)	0
• Ischaemic stroke	1 (0.3)	0
• Large intestine perforation†	1 (0.3)	0
• Neutropenic infection	1 (0.3)	0
• Dyspnoea	0	1 (0.3)

* Attribution to progressive disease by the local investigator, as documented on the Case Report Form
† One patient experienced more than 1 fatal AE considered drug-related (patient with large intestine perforation and diverticulum intestinale)
Source: adapted from Tables 60 and 61 of the CS[1]

The company argues that data on fatal AEs are confounded in two ways. Firstly, the company argues the extent of exposure was longer on nintedanib plus docetaxel compared to docetaxel alone. As noted in Table 23, the median number of cycles of docetaxel that patients received was greater in the nintedanib plus docetaxel arm than in the docetaxel arm (5 vs 4 respectively). Therefore it is argued that the higher exposure to docetaxel may have contributed, at least in part, to the higher incidence of fatal AEs of sepsis caused by neutropenia in the nintedanib arm through the known myelotoxic effect of docetaxel. Consequently neutropenia and sepsis are considered possible side effects of nintedanib therapy in combination with docetaxel and are regarded as important identified risks for future monitoring and ongoing safety surveillance. Secondly, the analysis focusing on the on-treatment fatal AEs resulted in a skewed view of the deaths that occurred during the study. The company states that further review of PD and non-PD deaths occurring during the entire observation period revealed no other safety pattern suggestive of nintedanib associated toxicities.[42]

The ERG considers that the number of deaths related to AEs is relatively small but agrees with the company that AE related deaths need to be monitored in future. The ERG considers that the greater number of PD related deaths in the nintedanib plus docetaxel arm could be related to the fact that PFS was longer in this arm and so patients were on treatment longer; this may also account for differences in non- PD deaths. However, the ERG does not consider that the greater number of cycles of docetaxel received by patients treated with nintedanib is likely to have been a confounder since, as reported by the National Confidential

Enquiry into Patient Outcome and Death, most patients with life threatening toxicity tend to experience fatal AEs during the first cycle of treatment;[64] it is reported by the company that dose intensity was similar between arms (98.1% vs 98.7%) .

Comparison of adverse events from the indirect evidence

As noted in section 4.4.1, it was only possible to conduct MTCs for safety outcomes if it was assumed pemetrexed and docetaxel had equal tolerability, an assumption which the ERG reiterates is not supported by the evidence (e.g. see JMEI[56]). The ERG has however presented the data input into the MTC as this shows AEs across two trials: LUME-Lung 1[24] and WSY001.[62] However, as WSY001[62] it should be noted that WSY001[62] is a trial of Asian patients conducted in China and AEs in a population in England may differ as a result of differences in co-morbidities, smoking history and pharmacokinetics between these populations. The ERG notes that the data from these trials support the generally held view that erlotinib is generally better tolerated than nintedanib plus docetaxel or docetaxel alone.

Table 27 Safety results for adenocarcinoma populations of trials included in MTC base-case analysis

Outcomes		LUME-Lung 1[24]		WSY001[62]	
		Nintedanib + docetaxel (n=320)	Placebo + docetaxel (n=333)	Erlotinib (n=61)	Pemetrexed (n=62)
Any CTCAE grade AE: fatigue	N	99	98	12	16
	%	30.9	29.4	19.7	25.8
Any CTCAE grade AE: nausea	N	91	59	1	15
	%	28.4	17.7	1.6	24.2
Any CTCAE grade AE: diarrhoea	N	139	82	10	2
	%	43.4	24.6	16.4	3.2
CTCAE grade ≥3 fatigue	N	15	14	0	0
	%	4.7	4.2	0	0
CTCAE grade ≥3 nausea	N	3	2	0	2
	%	0.9	0.6	0	3.2

AE=adverse event; CTCAE= Common Terminology Criteria for Adverse Events
Source: adapted from Table 34 of the CS[1]

4.6 Critique of the health related quality of life data

The company reports data on HRQoL data for LUME-Lung 1[24] that appears to have been reported in a poster presentation at the World Conference on Lung Cancer, Sydney, Australia, October 2013.[47] It is stated that data are reported for patients with adenocarcinoma only although baseline data were only available for all patients, regardless of histology. The ERG also notes that the company states that longitudinal analysis reported that nintedanib plus docetaxel did not result in a change in global health status/QOL in patients with adenocarcinoma. Self-reported HRQoL assessments by EORTC

questionnaires also revealed that there were no significant differences in cough, dyspnea or pain in patients over time or between those receiving nintedanib plus docetaxel and those receiving placebo plus docetaxel. Nintedanib-treated patients did however achieve numerically better cough and pain scores than placebo-treated patients, suggesting an improvement in HRQoL for these domains. Furthermore, statistically significant differences were observed between groups for three individual pain items ('have pain', 'pain in chest' and 'pain in arm and shoulder'). On the other hand, the TTD for diarrhoea was significantly worsened in the nintedanib plus docetaxel arm; there was no significant difference between arms for nausea and vomiting, or appetite loss (Table 28).

Table 28 Time to deterioration of nausea and vomiting, appetite loss and diarrhoea in patients with adenocarcinoma in LUME-Lung 1

Symptom	HR (95% CI)
Nausea and vomiting	1.23 (1.00 to 1.51)
Appetite loss	1.13 (0.92 to 1.38)
Diarrhoea	1.86 (1.51 to 2.30)*

*p<0.05
HR=hazard ratio
Source: Table 24 of the CS[1]

The ERG notes that the response to the HRQoL questionnaire appears to be very good; the company states that over 80% of patients completed HRQoL questionnaires over the first 20 cycles of treatment and approximately 70% of patients completed the questionnaire at the end of the treatment visit. It is noted that the main drivers of HRQoL in this population tend to be cancer related symptoms. Taking into account the findings for ORR and PFS (see section 4.3.6) in which it was observed that the addition of nintedanib did not make a major difference to response rates but did lead to increased rates of tumour control and slower progression on average, it is perhaps unsurprising that dramatic differences in HRQoL were not seen on initiation of therapy. It is interesting to observe significant differences in pain symptoms as fatigue, dyspnoea and cough are often reported to be more troublesome to patients and their families.[65] The worsened TTD for diarrhoea for patients treated with nintedanib plus docetaxel is unsurprising given the greater proportion of diarrhoea AEs in this arm (see section 4.5). The increased rates of diarrhoea did not seem have any major impact on global health status/QoL.

No attempt was made by the company to compare HRQoL between nintedanib plus docetaxel and erlotinib. For reasons stated above (sections 2.2, 3.3 and 4.4.1), the ERG does not consider such a comparison is relevant to the decision problem, even if such a comparison were possible.

4.7 Conclusions of the clinical effectiveness section

Clinical evidence has been submitted to NICE from the company in support of the use of nintedanib for previously treated locally advanced, metastatic or recurrent NSCLC of adult patients with adenocarcinoma tumour histology. The NICE scope did not specify adenocarcinoma nor did it refer to locally recurrent disease. This population is however in line with the anticipated marketing authorisation. While none of the scope, decision problem or anticipated marketing authorisation refer to the EGFR mutation status of NSCLC tumours, in England, the majority of patients (85 to 90%) are likely to be EGFR-negative. The ERG further notes that because patients who receive nintedanib also receive docetaxel, then patients who are likely to be eligible for treatment with nintedanib will also be ECOG PS 0 to 1.

Direct evidence is presented for nintedanib plus docetaxel vs placebo plus docetaxel from one RCT (LUME-Lung 1[24]). Indirect evidence for nintedanib + docetaxel vs erlotinib is presented from MTCs. While both docetaxel and erlotinib are specified as comparators in the NICE scope, given that erlotinib is likely to be preferred when patients have a poorer performance status and/or have EGFR-positive tumours, or be treatment naïve for a TKI, the ERG agrees with the company that erlotinib is not a relevant comparator and that docetaxel is the only appropriate comparator for this STA.

LUME-Lung 1[24] is a phase III double-blind RCT which compares nintedanib plus docetaxel vs placebo plus docetaxel. It is considered to have a low risk of bias. As a result of the exclusions of certain types of patients, the patient population appears to be fitter than would be found in clinical practice in England. This could partially explain why the post-study therapy rate is relatively high (55.8%).

The findings from this trial suggest nintedanib plus docetaxel significantly improved PFS and OS in comparison to placebo plus docetaxel in the subgroup of patients with adenocarcinoma. After a median follow-up of 31.7 months the gain in median PFS was 1.4 months (4.2 months vs 2.8 months) and gain in median OS was 2.3 months (12.6 months vs 10.3 months). However, the ERG does not consider that the assumption of proportional hazards is consistent with the trial data, and therefore use of these results in cost-effectiveness modelling should not be based implicitly or explicitly on this assumption. Nintedanib plus docetaxel also resulted in an increase in some types of AEs, particularly diarrhoea, nausea, vomiting and increases in ALT/AST. The majority of these AEs can be managed by dose reductions of nintedanib. The ERG is in agreement with the company that apparent improvements seen in terms of PFS and OS in the adenocarcinoma patients were achieved without substantial alterations in self-reported HRQoL.

5 COST EFFECTIVENESS

5.1 Introduction

This section provides a structured critique of the economic evidence submitted by Boehringer Ingelheim Ltd in support of the use of nintedanib in combination with docetaxel for the treatment of adult patients with locally advanced, metastatic or recurrent NSCLC of adenocarcinoma tumour histology after first-line chemotherapy. The two key components of the economic evidence presented in the CS[1] are (i) a systematic review of the relevant literature and (ii) a report of the company's de novo economic evaluation. The company also provided an electronic copy of their economic model that was developed in Microsoft Excel.

5.2 ERG comment on company's review of cost-effectiveness evidence

5.2.1 Objective of the company's cost-effectiveness review

On page 175 of the CS,[1] the company explains that "The scope of the systematic review is to review all available published data on economic evaluations of second-line therapies for locally advanced or metastatic NSCLC that could inform a HTA submission based on Boehringer Ingelheim's second-line comparative trials of nintedanib". This single systematic review was performed to identify clinical, cost-effectiveness, resource use and cost data as well as studies reporting utility scores for health states within the model.

Details of the cost-effectiveness search strategies employed are included in Appendix 10 of the CS. Medline (via PubMed), Medline R-In Process (via PubMed), EMBASE, and The Cochrane Library (via NHS EED) were searched for data on economic models, costs, resource use associated with NSCLC, HRQoL and utilities. HEED and EconLit were searched for data on HRQoL and utilities. The time horizon for the search for full economic studies was 2000 to February 2014 and for cost analyses was 2012 to 2013.

The search of the literature yielded no relevant studies. The ERG is satisfied with the company's search strategy and is confident that the company did not miss any relevant published articles.

5.2.2 Eligibility criteria used in the study selection

The inclusion/exclusion criteria used in the study selection are presented in Table 29. The ERG is satisfied that these criteria are relevant to the decision problem.

Table 29 Economic evaluation search inclusion/exclusion criteria

Parameter	Inclusion criteria	Exclusion criteria
Population	Relapsed or refractory NSCLC (RR NSCLC) (receiving second-line chemotherapy or relapsed/refractory to first-line chemotherapy)	Any patient population other than RR NSCLC
Interventions	Any second-line chemotherapy for RR NSCLC: • Monotherapy • Combination therapy with other chemotherapy Other interventions that are considered standard care in the patient population that will be relevant to the economic model	Patients who were treatment-naïve or had received more than first-line therapy
Outcomes	Economic models: • Cost-utility analyses • Cost-effectiveness analyses • Cost-benefit analyses • Cost-minimisation analyses	No outcomes of interest included
Study design	Economic models: Economic studies	Not an economic model
Language restrictions	English language	Non-English language
Date	Economic models: 2002 onwards	Prior to the year 2002*
Country	Any	None

*Abstracts published prior to 2011 and systematic reviews published prior to 2009 were excluded
Source: Table 72 of CS[1]

5.2.3 Included and excluded studies

No relevant studies were identified by the company.

5.2.4 Conclusions of the cost effectiveness review

The ERG notes that since nintedanib in combination with docetaxel has not yet received a full marketing authorisation from the EMA for the second-line treatment of adult patients with adenocarcinoma, the lack of economic evaluations of relevance to the decision problem is not unexpected.

5.3 ERG critique of the company's literature review

The ERG is satisfied with the company's search strategy and stated eligibility criteria for inclusion/exclusion. The ERG is confident that the company did not miss any relevant published papers.

The ERG acknowledges that the company reported the methods and results of a series of literature reviews at key points throughout the cost-effectiveness section in the CS;[1] the ERG considered the results of these additional reviews to be very helpful.

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

5.4.1 NICE reference case checklist

Table 30 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	Partial - the population was limited to patients with adenocarcinoma
Comparator(s)	Alternative therapies routinely used in the NHS	Yes
Perspective costs	NHS and Personal Social Services	Partial - only NHS costs were included in the model
Perspective benefits	All health effects on individuals	Yes, health effects to the individual are captured via QALYs
Form of economic evaluation	Cost-effectiveness analysis	Cost-effectiveness analysis
Time horizon	Sufficient to capture differences in costs and outcomes	Lifetime horizon was used (15 years)
Synthesis of evidence on outcomes	Systematic review	Nintedanib + docetaxel vs docetaxel: direct trial data from LUME-Lung 1[24] was used. Nintedanib + docetaxel vs erlotinib: hazard ratios were taken from the results of the company's network meta-analysis (fixed-effects model)
Outcome measure	Quality adjusted life years (QALYs)	QALYs were used which is appropriate
Health states for QALY	Described using a standardised and validated instrument	EQ-5D was used, with data collected mainly from participants in LUME-Lung 1.[24] Data from published sources[66,67] were used in the sensitivity analysis
Benefit valuation	Time-trade off or standard gamble	Time-trade off
Source of preference data for valuation of changes in HRQoL	Representative sample of the public	UK preference tariff based on public sample. Data for assigning valuation health states were collected directly from trial participants
Discount rate	An annual rate of 3.5% on both costs and health effects	Benefits and costs were 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	All QALYs estimated by the model have the same weight
Sensitivity analysis	Probabilistic sensitivity analysis	Deterministic, scenario and probabilistic sensitivity analyses were undertaken by the company

Table 31 Critical appraisal checklist for the economic analysis completed by ERG

Question	Critical appraisal	ERG comment
Was a well-defined question posed in answerable form?	Yes	
Was a comprehensive description of the competing alternatives given?	Yes	
Was the effectiveness of the programme or services established?	Partial	For the direct comparison of nintedanib + docetaxel vs docetaxel, effectiveness was established using data from LUME-Lung 1[24] For the indirect comparison of nintedanib + docetaxel vs erlotinib, an MTC was undertaken. The ERG does not consider the results of the MTC to be valid or reliable, nor does it consider the comparison to be relevant to the decision problem
Were all the important and relevant costs and consequences for each alternative identified?	Mostly	Specific issues are discussed insection 5.5 and the impact on the ICER is presented insection 6 of the ERG report
Were costs and consequences measured accurately in appropriate physical units?	No	Specific issues are discussed insection 5.5 and the impact on the ICER is presented insection 6 of the ERG report
Were the cost and consequences valued credibly?	No	Specific issues are discussed insection 5.5 and the impact on the ICER is presented insection 6 of the ERG report
Were costs and consequences adjusted for differential timing?	Yes (with errors)	Specific issues are discussed insection 5.5 and the impact on the ICER is presented insection 6 of the ERG report
Was an incremental analysis of costs and consequences of alternatives performed?	Yes	ICERs were 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	

5.4.2 Description of the company's economic model

A schematic of the company's submitted economic model is provided in the CS[1] and is reproduced in Figure 2.The company's cost-effectiveness model is a partitioned survival Markov model which comprises three health states: progression-free (on or off treatment) (PF); PD and death (D). All patients enter the model in the PF state. At the beginning of each time period patients can either remain in the same health state or progress to a worse health state, i.e. from PF to PD or death; or from PD to death. The model uses the partitioned survival (also known as area under the curve or AUC) method to determine the proportion of patients in each of the three health states during each model cycle. The proportion of patients in the PD state is estimated as the difference between OS and PFS. Estimates of OS and PF are based on PF and OS survival data from LUME-Lung 1[24] and the corresponding parametric survival models. The model assumes that patients receive best supportive care (BSC) following the discontinuation of active second-line treatment. The model also allows some patients in the progressed state to have subsequent treatments. The costs of subsequent treatment are included in the economic evaluation; however, the impact of subsequent therapy is not included in the model. Variants of this model structure have been used in the modelling of metastatic oncology for numerous STAs including two recent NICE STAs that considered locally advanced or metastatic NSCLC (NICE TA192[14] and NICE TA258[13]).

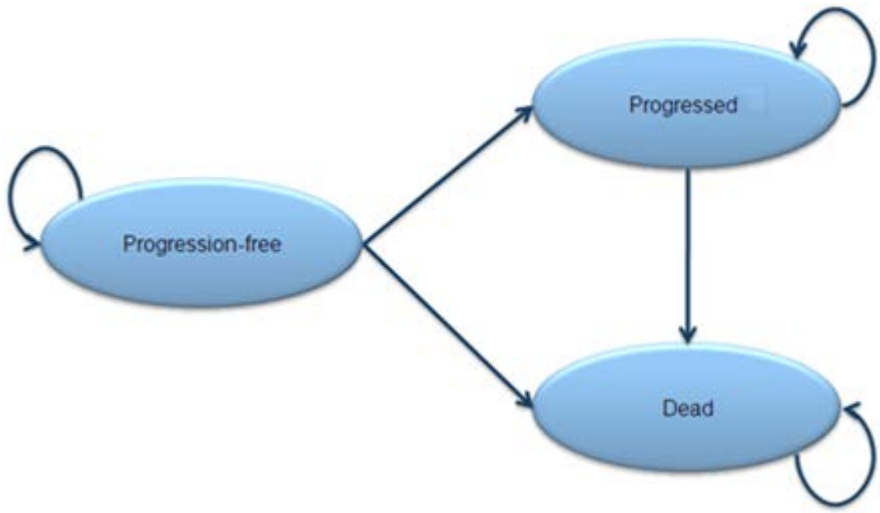


Figure 2 Schematic of company's model
Source: Figure 24 of the CS[1]

The model has been developed in Microsoft Excel using a 3-weekly cycle length. It includes a half-cycle correction and the time horizon is set at 15 years. Health effects are measured in QALYs. A summary of all of the variables applied in the economic model is shown in

Table 79 of the CS;[1] details displayed in the table include the values used, range (distribution) and source.

5.4.3 Population

The company states on page 180 of the CS[1] that the model population was based on the findings of LUME-Lung 1[24] and included patients with locally advanced and/or metastatic, stage IIIB/IV or recurrent NSCLC with adenocarcinoma histology who failed after first-line chemotherapy.

5.4.4 Interventions and comparators

The company's base-case economic evaluation compares nintedanib plus docetaxel with docetaxel. The interventions are implemented in the model in accordance with their current, or anticipated, full marketing authorisations and doses.

Patients receiving nintedanib plus docetaxel are assumed to take two 100mg capsules of nintedanib twice daily; there are 120 capsules in each 100mg pack. The assumed NHS list price per 30-day pack is £2151.10. The ERG notes that there is also a 150mg capsule available; there are 60 capsules in each 150mg pack. In response to a clarification question raised by the ERG, the company indicated that the price of both packs is likely to be the same. Nintedanib plus docetaxel therapy needs to be given for a minimum of four cycles before nintedanib can be administered as monotherapy. There is no administration cost associated with nintedanib. Patients receiving intravenous docetaxel are assumed to receive 75mg/m^2 on day 1 of a 21-day cycle. The monthly cost of docetaxel is estimated to be £49 (using electronic Marketing Information Tool [eMIT] prices[68]) and has a monthly administration cost of £221.43 (NHS Reference Cost 2012/13).[69]

The submitted economic model also permits the comparison of nintedanib plus docetaxel with erlotinib. In the model the dose of erlotinib is assumed to be 150mg per day and the MIMS 2013 price for a pack of 30 tablets is £1631.53.[70] It is noted that erlotinib has an associated patient access scheme, which the company took into account by undertaking a number of sensitivity analyses in which a range of discounts were applied to the list price. However, the company emphasises on page 184 of the CS[1] that erlotinib is not a relevant comparator and considers that patients treated with erlotinib are a different patient population.

Some patients in the model go on to receive subsequent therapy after progression: the company's external expert stipulated that 5% would receive erlotinib, 25% would receive a platinum doublet and 70% would receive BSC. The cost per month of BSC (£406.63 per

cycle [3 weeks]) is as per TA310 (Afatinib NICE submission) [15] as recommended by the company's external expert.

5.4.5 Perspective, time horizon and discounting

The company states that the economic evaluation is undertaken from the perspective of the NHS and Personal Social Services. However, it should be noted that the model does not include all likely Personal Social Services costs. The time horizon is set at 15 years and both costs and outcomes are discounted at 3.5% in line with the NICE Methods Guide to Technology Appraisal.[54]

5.4.6 Treatment effectiveness and extrapolation

Modelling treatment effectiveness (nintedanib plus docetaxel vs docetaxel)

Kaplan-Meier survival curves for OS and PFS for nintedanib plus docetaxel and docetaxel monotherapy were available from LUME-Lung 1[24] and show the proportion of patients in the model's three health states at each time point. These data were incorporated into the cost-effectiveness model by using full parametric approximation of the raw data in the base-case. In the sensitivity analyses, K-M data from LUME-Lung 1[24] were used to model OS (until at least 5% of trial patients are still at risk) and were extrapolated using parametric function as a tail to the Kaplan-Meier data to provide a lifetime time horizon. Survival data from LUME-Lung 1[24] were mature and the proportion of censored patients in both treatment arms were similar. However, in order to facilitate extrapolation of trial data beyond the time horizon, OS and PFS data were analysed using parametric survival models. Parametric survival curves were fitted to PFS and OS K-M curves using two approaches: 1) joint models, statistical models including data for both treatment groups with a term for treatment and 2) separate models, statistical models fitted to each randomised treatment arm separately.

Choice of statistical model

The "goodness of fit" based on Akaike information criteria (AIC) indicated that joint models were appropriate. However, the intercept and scale parameters of the separately fitted curves indicated that the curves should not be forced into the same model, thus separate curves were selected for OS and PFS. The log-logistic model had the lowest AIC among the separately fitted OS models, and the Weibull model had the lowest AIC among the separate proportional hazard models for OS; therefore, these were selected to model the OS data. The log-normal model had the lowest AIC among the separate PFS fits, and the Weibull had the lowest AIC among the separate proportional hazard models for PFS; therefore, these were selected to model PFS.

The company states that the long-term extrapolation of trial data was validated with a group of UK clinicians and against data from SEER using the SEER*Stat software, as well as against data from LUCADA. As noted in section 5.5, no references were provided to identify the specific DEER and LUCADA data sets employed.

Survivals implemented in the model

Survival modelling options programmed into the cost-effectiveness model are displayed in Table 32. In the base-case, the analysis used separate models for PFS and OS, with log-normal distribution for the PFS and log-logistic distribution for the OS.

Table 32 Survival estimation models employed in the company's model

Progression-free survival	Overall survival (OS)
Separate model - Log-normal (base-case)	Separate model - Log-logistic (base-case)
Separate model - Weibull	Separate model - Weibull
Kaplan-Meier curve*	Kaplan-Meier curve*
	Kaplan-Meier curve and SEER lognormal†
	Kaplan-Meier curve and Separate Log-logistic†
	Kaplan-Meier curve and Separate Weibull†
	Kaplan-Meier curve and LUCADA lognormal†

LUCADA= National Lung Cancer Audit database; OS=overall survival; PFS=progression-free survival; SEER=Surveillance, Epidemiology and End Results
* With this option, the model does not extrapolate the PFS/OS with the use of parametric models but it uses the K-M curves for PFS/OS obtained directly from the LUME-Lung 1[24] trial. Note that this option only applies for nintedanib + docetaxel and docetaxel monotherapy
† With this option, the Kaplan-Meier curves from the LUME-Lung 1[24] trial are used for the estimation of OS until patient number at risk drops down to 5% of original patients, afterwards parametric models are used
Source: Table 75 of the CS[1]

Modelling treatment effectiveness of erlotinib

As OS and PFS K-M curves for erlotinib were not available, model OS and PFS inputs for erlotinib were derived by applying HRs (i.e., vs nintedanib plus docetaxel) obtained from the mixed treatment comparisons to the OS and PFS of nintedanib plus docetaxel. The company considers that HRs can only be used if the survival distribution is a proportional hazard model such as exponential, Weibull, or Gompertz. Thus, in the model, erlotinib can be evaluated only if a Weibull distribution is selected for both OS and PFS. The model base-case analysis utilised HRs from the MTC base-case. The HR for OS was 0.64 (95% CI: 0.46 to 0.90) and the HR for PFS was 0.7 (95% CI: 0.5 to 1.0). The company used results from the fixed-effects model because there was one trial per comparison.

5.4.7 Health related quality of life

Utility

Health related quality of life data were collected during LUME-Lung 1[24] using the EQ-5D instrument, in line with the NICE Methods Guide to Technology Appraisal.[54] Data from the LUME-Lung 1[24] were analysed using a longitudinal model adjusted for baseline ECOG score, prior treatment with bevacizumab, presence of brain metastases, controlling for health status and key adverse events. Key model utility values for PF and PD are displayed in Table 33.

Table 33 Utilities for progression-free and post-progression states

Nintedanib + docetaxel and docetaxel arms - Pooled	Progression free without adverse events	
	Mean	Standard error
Week 0	0.710	0.01
Week 3	0.721	0.01
Week 6	0.707	0.01
Week 9	0.699	0.01
Week 12	0.692	0.01
Week 15	0.687	0.01
Week 18	0.682	0.01
Week 21	0.677	0.02
Week 24	0.671	0.02
Week 27	0.666	0.02
Week 30	0.661	0.02
Treatment arm	**Progressive disease**	
	Mean	**Standard error**
Nintedanib + docetaxel	0.64	0.01
Docetaxel	0.64	0.01

Source: Table 80 of the CS[1]

Progression free utility estimates

The analysis estimated utility values over time for PF patients from week 0 to week 30 in 3-week intervals - without a treatment term. An assumption of the linear extrapolation of trend observed until week 30 for the PF health state is employed in the base-case to allow modelling of continuing change in utility in the PF state beyond the trial data.

Progressed disease utility estimates

In contrast to the estimation of PF utilities over time, mean PD utilities were used in the base-case model to accommodate the memory-less feature of the Markov approach.

Utility values used in the model

The company's model uses the utility values derived from LUME-Lung 1[24] in the base-case. Utility values from the literature are also tested within the model. The company used utility values from a recently published paper by Chouaid et al[66] in a sensitivity analysis. This paper reports utilities recorded from relevant patients in Europe, Canada, Australia and Turkey as well as the UK and uses the EQ-5D to obtain utilities for the health states that were used in the company's model.

Table 34 Utilities used in the sensitivity analysis (Chouaid et al[66] 2013)

Health state	Mean (Standard error)
Progression free survival (PFS)	0.74 (0.03)
Post-progression	0.46 (0.08)

Source: Table 83 of the CS[1]

Disutility

The company's model also incorporated the impact of AEs on HRQoL; utility decrements associated with each AE were applied for a period of one model cycle. The company notes that the model may have double counted disutilities as some patients may experience multiple AEs simultaneously. Disutilities due to AEs are presented in Table 35.

Table 35 Disutilities associated with AEs

Adverse event	Disutility	Sources
ALT increased	-0.05	Assumption
Anaemia	-0.07	Nafees et al[67]
Diarrhoea - grade 2	-0.02	Assumption: half of the disutility for grade 3/4 diarrhoea
Diarrhoea - grade 3/4	-0.04	Data on file, Table 18.1[71]
Fatigue	-0.21	Data on file, Table 18.1[71]
Febrile neutropenia	-0.09	NICE TA192,[14] Nafees et al. 2008[67]
Infection	-0.05	Assumption
Liver-related investigations	-0.05	Assumption
Nausea and vomiting	-0.05	Nafees et al[67]
Neutropenia	-0.09	Nafees et al[67]
Neutrophil count decreased	-0.09	Assumption: same as disutility of neutropenia
Rash	-0.033	Nafees et al.[67]
Thrombocytopenia	-0.05	NICE TA181[72]
WBC count decreased	-0.05	Assumption

WBC=white blood cell
Source: Table 84 of the CS[1]

5.4.8 Resources and costs

Drug acquisition and administration costs

Table 36 presents a summary of the drug and IV administration costs per cycle for each comparator for the active second-line treatment phase, the BSC phase and, where relevant, the third-line treatment phase. Adjustments in drug costs due to change in dose intensity and treatment discontinuation as observed in LUME-Lung 1[24] were included in the company's model for second-line treatments. Changes in dose intensity or treatment discontinuation inputs only affected drug costs outcomes; they did not affect clinical outcomes (e.g. OS, PFS and AEs). Wastage was taken into account when calculating the cost of IV treatments.

As nintedanib is taken orally, it is not associated with any additional administration costs.

Table 36 Drug costs used in the company's model

Drug	Units per administration	Price per unit	Route	Administrations per cycle	Administration cost	Costs per cycle*
Nintedanib	400 mg	£0.18	Oral	21	-	£1,354
Docetaxel in combination with nintedanib	75 mg/m^2 400 mg	£5.68 £0.18	IV Oral	1 21	£155 -	£196 + £1,354= £1550
Docetaxel	75 mg/m^2	£5.68	IV	1	£155	£196
Erlotinib	150 mg	£0.36	Oral	21	-	£1,051
Carboplatin†	750 mg	£0.33	IV	1	£155	£250
Vinorelbine†	30 mg/m^2	£2.78	IV	3		£465

IV=intravenous
* Mean dose intensity taken into account: (nintedanib + docetaxel=98.1%, nintedanib=91.2%, docetaxel=98.7% and erlotinib=92%)
† third-line treatment
Source: Table 96 of the CS[1]

Health state costs

The company considered that there was little published literature exploring the detailed resource use commonly associated with NSCLC or other metastatic cancer. To estimate the treatment patterns in NSCLC a resource use questionnaire was constructed. This formed the basis of an interview with an oncologist who specialised in the treatment of patients with lung cancer and who had experience of working on NICE health technology assessment reports. A series of questions was posed separately for each different health state (stable on nintedanib plus docetaxel, stable on docetaxel, stable on erlotinib, stable on BSC; progressed on active treatment, progressed on BSC; and a one-off estimate of resource use at the time of progression) under the umbrella term 'monitoring'. Three main areas of resource use were considered: routine follow up (type and frequency of physician visit,

laboratory tests, radiological scans); treatment at time of progression (hospitalisations, physician visits, laboratory tests, radiological scans, procedures use; and resources used during BSC/palliative care (initial tests, procedures, hospitalisations, physician visits, laboratory tests, radiological scans and procedures). Detailed descriptions of resource use are displayed in Tables 98 to 105 in the CS;[1] in addition a full range of the unit costs employed is also presented in Table 106 of the CS.[1] The unit costs of visit procedures and laboratory tests were mainly derived from the National Schedule of Reference costs (2012/3),[69] whilst some visit costs were taken from the Personal Social Services Research Unit (PSSRU).[73]

Adverse events costs

A single UK consultant provided AE management costs. Estimates were generated via survey and face-to-face discussion. Costs for inpatient hospitalisations were taken from the NHS National Schedule of Reference Costs (2012/13).[69] Outpatient costs were taken from the same source[69] or from the PSSRU.[73] The cost of each AE is summarised in Table 37.

Table 37 Adverse events costs

Type of adverse event	Cost of adverse events
ALT increased	£587
Anaemia	£978
AST increased	£336
Diarrhoea - CTCAE grade 1 and 2	£250
Diarrhoea - CTCAE grade 3 and 4	£1796
Fatigue	£370
Febrile neutropenia	£2012
Infection	£2181
Nausea and vomiting	£1919
Neutropenia	£346
Rash	£639
Thrombocytopenia	£422
WBC count decreased	£423

ALT=alanine aminotransferase; AST=aspartate aminotransferase; CTCAE= Common Terminology Criteria for Adverse Events; WBC=white blood cell
Source: Table 107 of the CS[1]

5.4.9 Model validation

The company reports that a number of steps were taken to ensure that the analysis was validated, including:

- External review by a leading UK clinical expert to ensure that the model adheres to the clinical course of the disease and is reflective of current clinical practice
- Sensitivity analyses
- A senior modeller within the model developers' organisation (with no involvement in the model's development) performed a detailed quality assurance check on the model
- The company performed validation checks (varying parameter values and assumptions). This involved increasing and decreasing various parameters or changing assumptions in the model and then monitoring the impact on outputs. If the outputs were unexpected, further checks were made to determine whether this was the result of an error in the model.

5.4.10 Results included in the company's model

The incremental cost-effectiveness results generated by the company's economic model are presented in Table 38 and Table 39. The ICER for nintedanib plus docetaxel vs docetaxel is estimated by the company to be £50,677 per QALY gained. The ICER for nintedanib plus docetaxel vs erlotinib is estimated by the company to be £27,008.

Table 38 Company's base-case cost-effectiveness results: nintedanib plus docetaxel vs docetaxel

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£) vs baseline (QALYs)	ICER (£) incremental (QALYs)
Nintedanib + docetaxel	[redacted]	[redacted]	[redacted]	-	-	-	-	-
Docetaxel	[redacted]	[redacted]	[redacted]	£11,051	0.33	0.22	£50,776	£50,776

ICER=incremental cost-effectiveness ratio; LYG=life years gained; QALY=quality adjusted life year
N.B. Distributions used - OS: Log-logistic; PFS: Log-normal
Source: Table 129 of the CS[1]

Table 39 Company's secondary cost-effectiveness results: nintedanib plus docetaxel vs erlotinib

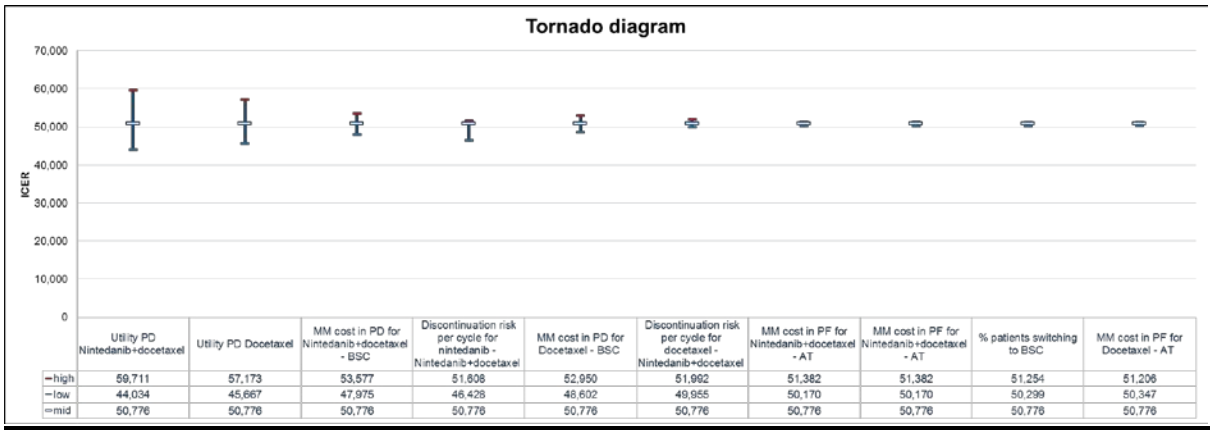
Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£) vs baseline (QALYs)	ICER (£) incremental (QALYs)
Nintedanib + docetaxel	[redacted]	[redacted]	[redacted]	-	-	-	-	-
Erlotinib	[redacted]	[redacted]	[redacted]	£7,571	0.43	0.28	£27,008	£27,008

ICER=incremental cost-effectiveness ratio; LYG=life years gained; QALY=quality adjusted life year
N.B. Distributions used - OS: Weibull distributions; PFS: Weibull survival
Source: Table 130 of the CS[1]

5.4.11 Sensitivity analyses

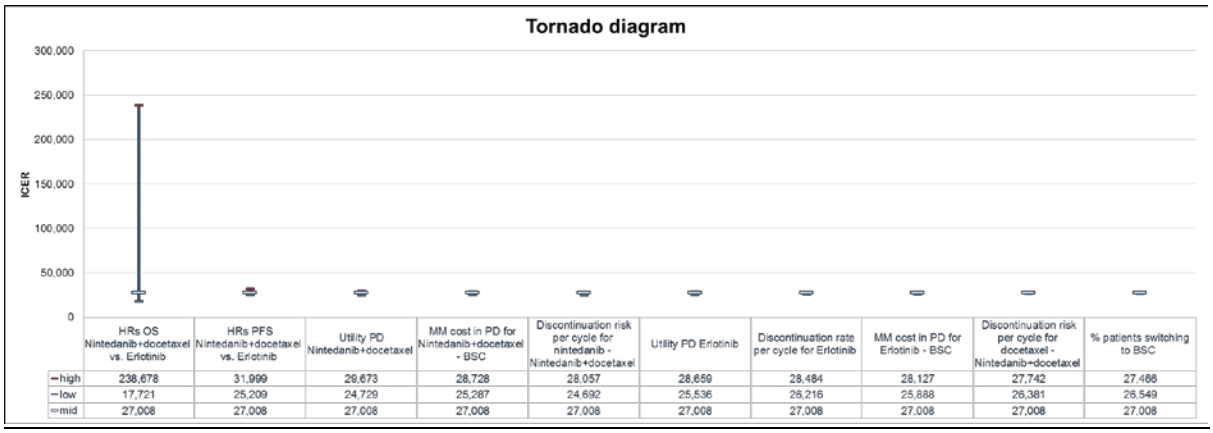
Deterministic sensitivity analyses

The company carried out a wide range of deterministic sensitivity analyses. Results for the ten parameters showing the greatest variability for the comparisons of nintedanib plus docetaxel vs docetaxel and vs erlotinib are shown in Figure 3 and Figure 4 respectively. For the comparison of nintedanib plus docetaxel vs docetaxel, the two most influential variables were univariate changes in utility values after progression for both intervention and comparator. For the comparison of nintedanib plus docetaxel vs erlotinib, the single most influential variable was the HR used for OS.



	Utility PD Nintedanib+docetaxel	Utility PD Docetaxel	MM cost in PD for Nintedanib+docetaxel - BSC	Discontinuation risk per cycle for nintedanib - Nintedanib+docetaxel	MM cost in PD for Docetaxel - BSC	Discontinuation risk per cycle for docetaxel - Nintedanib+docetaxel	MM cost in PF for Nintedanib+docetaxel - AT	MM cost in PF for Nintedanib+docetaxel - AT	% patients switching to BSC	MM cost in PF for Docetaxel - AT
high	59,711	57,173	53,577	51,606	52,950	51,992	51,382	51,382	51,254	51,206
low	44,034	45,667	47,975	46,428	48,602	49,955	50,170	50,170	50,299	50,347
mid	50,776	50,776	50,776	50,776	50,776	50,776	50,776	50,776	50,776	50,776

Figure 3 One-way sensitivity analysis: nintedanib plus docetaxel vs docetaxel
Source: Figure 33 of the CS[1]



	HRs OS Nintedanib+docetaxel vs. Erlotinib	HRs PFS Nintedanib+docetaxel vs. Erlotinib	Utility PD Nintedanib+docetaxel	MM cost in PD for Nintedanib+docetaxel - BSC	Discontinuation risk per cycle for nintedanib - Nintedanib+docetaxel	Utility PD Erlotinib	Discontinuation rate per cycle for Erlotinib	MM cost in PD for Erlotinib - BSC	Discontinuation risk per cycle for docetaxel - Nintedanib+docetaxel	% patients switching to BSC
high	238,678	31,999	29,673	28,728	28,057	28,659	28,484	28,127	27,742	27,466
low	17,721	25,209	24,729	25,287	24,692	25,536	26,216	25,888	26,381	26,549
mid	27,008	27,008	27,008	27,008	27,008	27,008	27,008	27,008	27,008	27,008

Figure 4 One-way sensitivity tornado diagram: nintedanib plus docetaxel vs erlotinib
Source: Figure 34 of the CS[1]

Probabilistic sensitivity analysis

The company undertook probabilistic sensitivity analysis (PSA) to derive the mean ICERs per QALY gained for nintedanib plus docetaxel vs docetaxel and vs erlotinib. PSA was carried out using 5000 iterations of the cost-effectiveness model.

The PSA result for nintedanib plus docetaxel vs docetaxel shows that nintedanib plus docetaxel has a 2% probability of being cost-effective at the £30,000 per QALY gained threshold and a 50% chance of being cost-effective at the £50,000 per QALY gained threshold. The cost-effectiveness plane and cost-effectiveness acceptability curve (CEAC) for this comparison are displayed in Figure 5 and Figure 6.

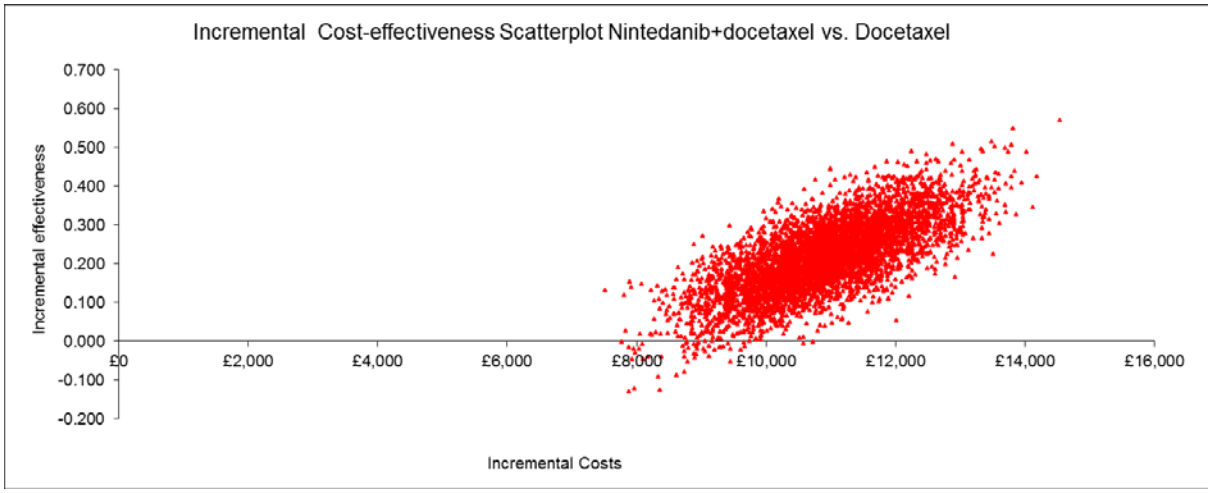


Figure 5 Incremental cost-effectiveness plane for nintedanib plus docetaxel vs docetaxel
Source: Figure 35 of the CS[1]

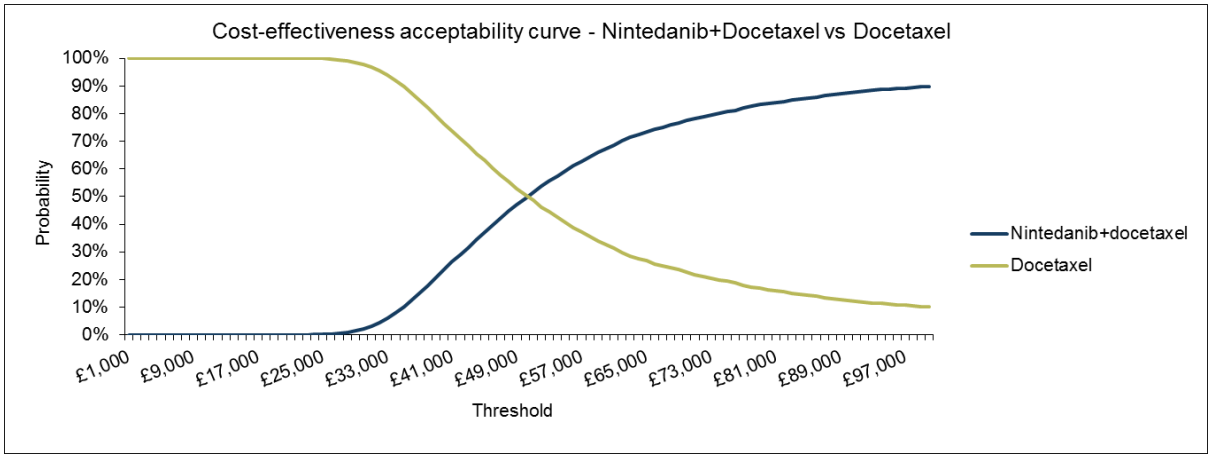


Figure 6 Cost-effectiveness acceptability curve for nintedanib plus docetaxel vs docetaxel
Source: Figure 36 of the CS[1]

The PSA result for nintedanib plus docetaxel vs erlotinib shows that nintedanib plus docetaxel has a 65% probability of being cost-effective at the £30,000 per QALY gained threshold and a 94% chance of being cost-effective at the £50,000 per QALY gained threshold. The cost-effectiveness plane and CEAC for this comparison are displayed in Figure 7 and Figure 8.

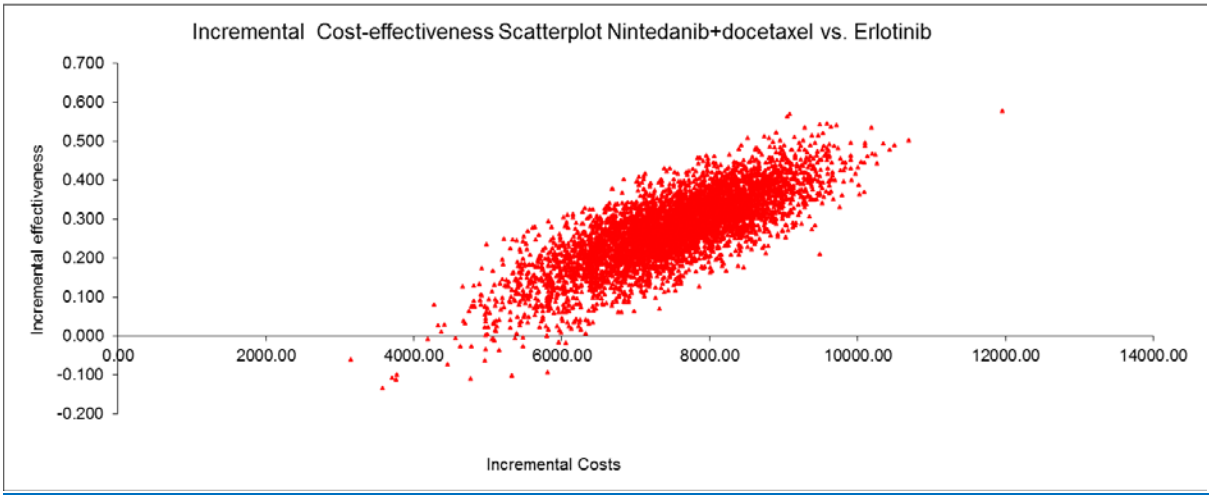


Figure 7 Cost-effectiveness plane for nintedanib plus docetaxel vs erlotinib
Source: Figure 37 of the CS[1]

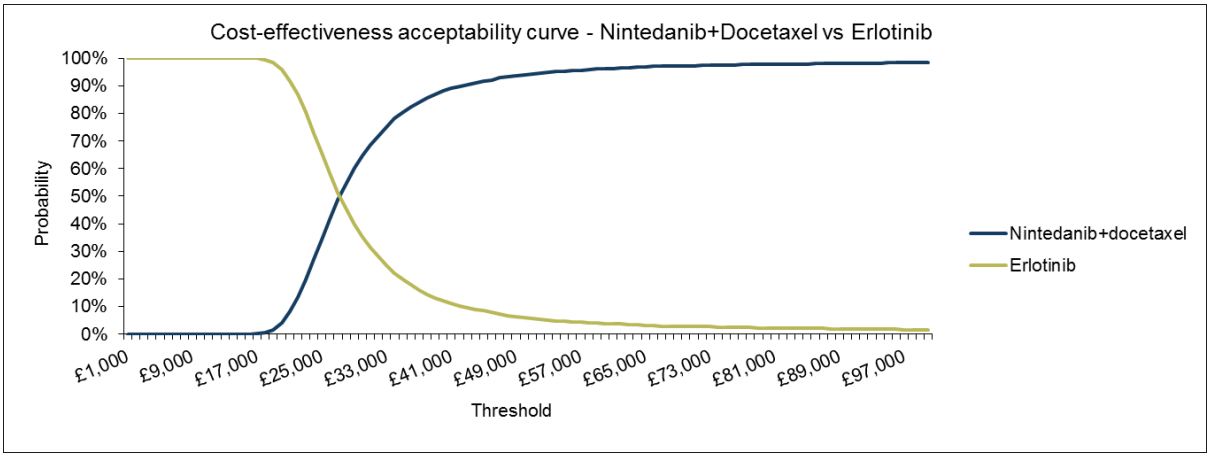


Figure 8 Cost-effectiveness acceptability curve for nintedanib plus docetaxel vs erlotinib
Source: Figure 38 of the CS[1]

Scenario analyses

The company also undertook a series of scenario analyses and explored how varying scenarios relating to survival modelling, indirect comparisons, resource use, utility, time horizon and discount rate might affect the results of the economic evaluation. The results of these scenario analyses are displayed in Tables 135 to 140 in the CS.[1] The company concluded that the base-case ICERs are mainly sensitive to changes in the PFS and OS HRs as well as the costs and utilities associated with the post-progression states.

5.5 Detailed critique of company's economic model

The model submitted by the company for this appraisal is structured as a partitioned survival spreadsheet model following a structure broadly similar to those used in similar appraisals.[13,14] For most functions the assumptions and options are labelled and annotated where necessary. However, in some cases, the ERG has found it difficult to confirm details of the data sources employed.

In line with the issues previously discussed (section 2.2 and 3.3) concerning the relevance of erlotinib as a comparator (largely due to the challenge of identifying a meaningful population for such a comparison), and the unreliability of indirect evidence of relative efficacy (section 4.4.1 and 4.4.6), this critique is primarily focussed on the direct comparison between nintedanib plus docetaxel and docetaxel monotherapy in the adenocarcinoma subgroup of LUME-Lung 1.[24]

A particular concern of the ERG relates to the analyses reported by the company of OS data from the SEER and LUCADA registers (Appendix 13 of CS[1]). No references were provided which identify the specific data sets employed and relevant details (such as date of extraction, selection criteria, duration of follow-up) are missing. The ERG has had to infer from the text that the SEER results appear to relate to all-cause mortality from the date of Stage 4 diagnosis and that the LUCADA data relate to second-line chemotherapy, but without any specific indication of prior treatments, PS and/or other relevant characteristics. The value of these analyses to support the company's chosen parametric survival modelling is therefore difficult to assess, and in particular the relevance of the SEER dataset to the population recruited to LUME-Lung 1[24] must be considered weak.

The following sections detail eleven specific issues identified by the ERG involving errors in data analysis, parameter values or methodology which have been identified in the submitted model, together with estimates made by the ERG of the impact of correcting these problems on the estimated ICER for nintedanib plus docetaxel compared with docetaxel. Within the time available to the ERG, it has not been possible to be certain that other problems do not remain undetected in the company's model.

5.5.1 Methods used to project time-to-event outcomes

In seeking to project OS and PFS data from LUME-Lung 1[24] to represent expected lifetime experience, the company has followed a convention of seeking to fit a variety of standard parametric functions to the available data, and employed the derived functions in place of the trial data throughout their decision model.

The ERG considers that this approach to model calibration to be flawed on several counts:

- The primary purpose of curve-fitting is to anticipate what is likely to happen to the minority of patients who remain at risk (i.e. alive with or without disease progression or remaining on treatment) at the time of data cut. However, the great majority of data events used for this purpose are drawn from patients who are unlike those remaining at risk at the time of data cut, since that majority were at greater propensity to fail (i.e. die, progress or cease treatment) than those still remaining. This is an example of bias against survivors and frequently results in the fitting of inappropriate functions and misleading projection estimates.

- The methods used for fitting parametric functions to a survival data set are essentially descriptive and lack any external validity based on the appropriateness of an underlying disease/treatment process governing them. Therefore, the analyst may be content in having achieved a reasonable correspondence to the available data, but lacks any basis for confidence in any future projection based thereon.

- When a single clinical trial is the primary source for cost-effectiveness assessment, it is important to make the maximum direct use of the available evidence. Replacing a large part of the trial results with a fitted model adds additional uncertainty from imposed modelling assumptions to the unavoidable data sampling uncertainty, so that rather than clarifying the underlying disease dynamic, it only serves to obscure it.

- Most of the 'standard' statistical functions used by the company to model survival lack any logical or empirical basis for representing a biological phenomenon, being only selected for their analytical convenience.

As part of the clarification process, the ERG requested detailed K-M survival analysis results for all of the time-to-event trial data employed in the company's model. The ERG has, for each of the K-M survival analysis results, identified a projective model, using only those data in the period towards the end of the survival curve in which it is apparent that a long-term trend has become established. The early K-M data are used directly in the company's

model, giving way to the projective model only to represent the segment of patient experience which cannot be reliably estimated otherwise.

In projecting ToT the company's model considers only a single parametric function (exponential model with fixed hazard per cycle calibrated over the whole trial period). Here, the same methodology flaws are present, except that no attempt has been made to assess the comparative validity of the exponential hazard function against possible alternatives.

5.5.2 Overall survival estimation

The company's model base-case comparison of nintedanib plus docetaxel vs docetaxel indicates a gain in (undiscounted) overall survival of 4.7 months; only 15% (0.7 months) of this gain is attributed to the pre-progression phase. This is unusual in locally advanced and metastatic cancers where treatment benefit is largely confined to the active treatment period (i.e. PFS). In order to validate this claim, the ERG has carried out its own analysis of the OS and post-progression survival (PPS) trial data, based on K-M results provided by the company in response to a clarification request.

Figure 9 shows a cumulative hazard chart for OS. After about 300 days, a simple linear trend is established in both trial arms and continues indefinitely. This indicates that in both arms OS can be estimated by use of a simple exponential projective model (i.e. there is a constant hazard irrespective of time). Comparing the slopes of the trend lines allows a long-term HR of 0.83 in favour of nintedanib plus docetaxel to be estimated. To verify this finding a similar cumulative hazard chart was prepared for PPS (shown as Figure 10). This confirms that patients in LUME-Lung 1[24] who survived a disease progression event continued to gain survival benefit from treatment with nintedanib plus docetaxel compared with those receiving only docetaxel. Long-term linear trends are apparent in both trial arms beyond 200 days in PPS, and the trends continue to diverge with an estimated long-term HR of 0.79 in favour of nintedanib plus docetaxel.

Estimates of lifetime OS were obtained by the ERG by applying the K-M trial results directly using the area under curve (AUC) method until the long-term OS trends were established and then projecting remaining estimated survival using the exponential trends (as shown in Figure 11). Mean OS in the docetaxel arm is estimated as 453.0 days (14.9 months) compared with 545.7 days (17.9 months) in the nintedanib plus docetaxel arm, a net survival gain of 92.7 days (3.05 months) attributable to the addition of nintedanib to docetaxel. This result is considerably lower than the OS gain obtained from the company's model (4.7 months), and indicates the effect of replacing the company's preferred Log-Logistic survival model to represent the whole trial data set with the ERG's approach (direct use of

unadjusted trial data for the majority of patients, followed by projecting long-term survivors using trends evident in the trial data set).

Replacing the company's preferred OS model with the ERG's approach has a major impact on the cost-effectiveness of nintedanib plus docetaxel compared with docetaxel alone. The incremental discounted cost per patient is reduced by [redacted] while the incremental discounted QALY gain is reduced by [redacted], resulting in the estimated ICER increasing from £50,776 per QALY gained to £68,587 per QALY gained.

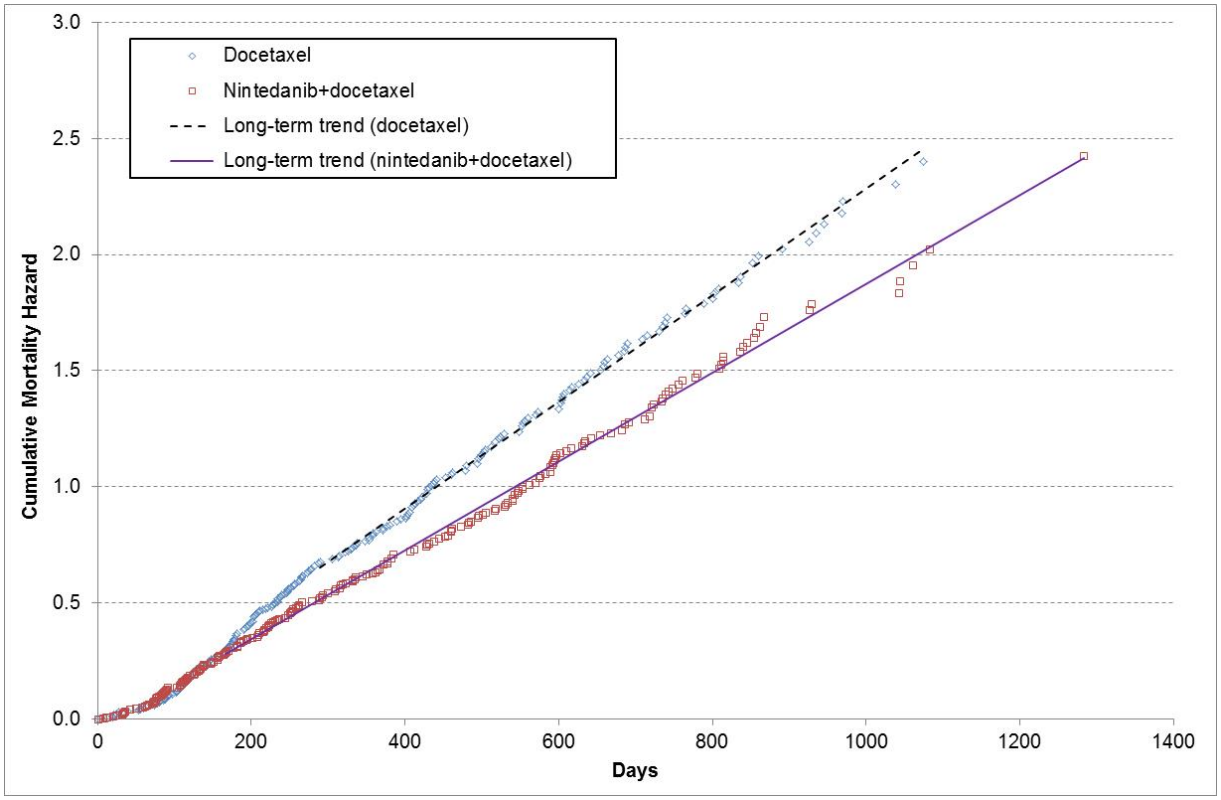


Figure 9 Cumulative OS hazard plot for nintedanib plus docetaxel vs docetaxel

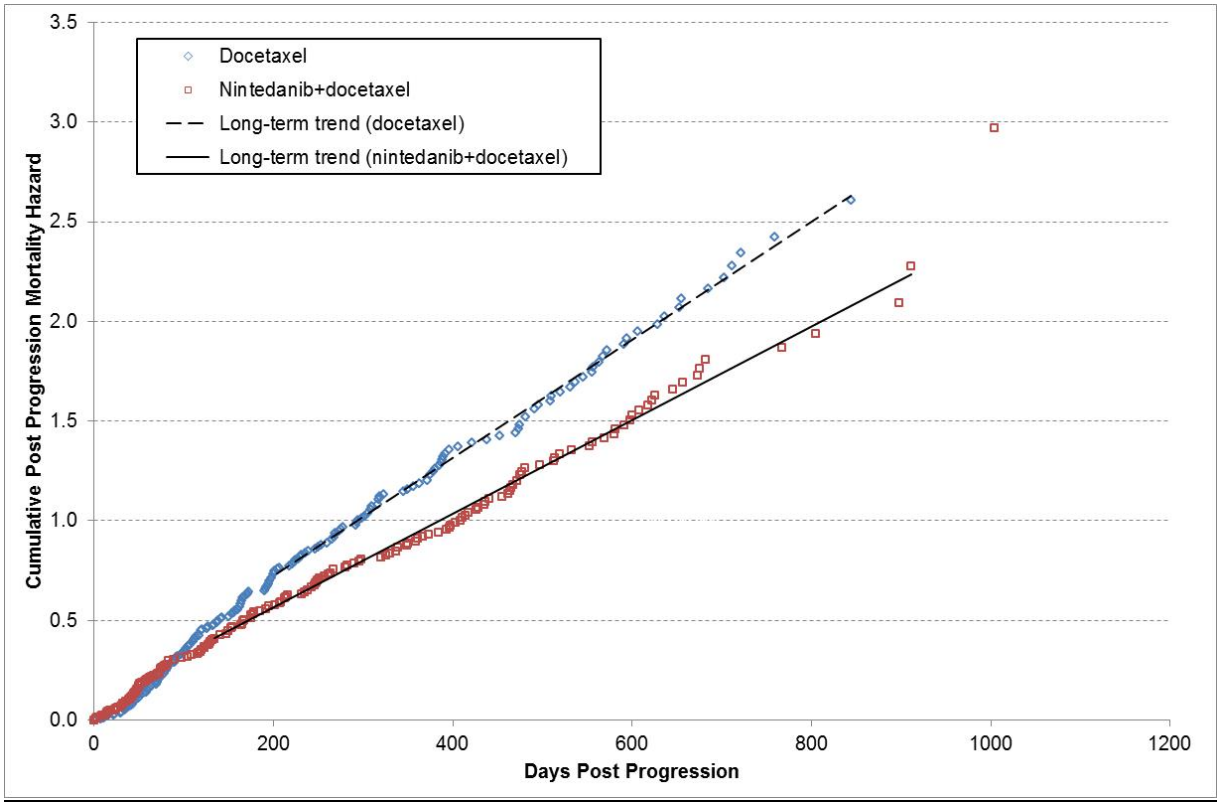


Figure 10 Cumulative PPS hazard plot for nintedanib plus docetaxel vs docetaxel

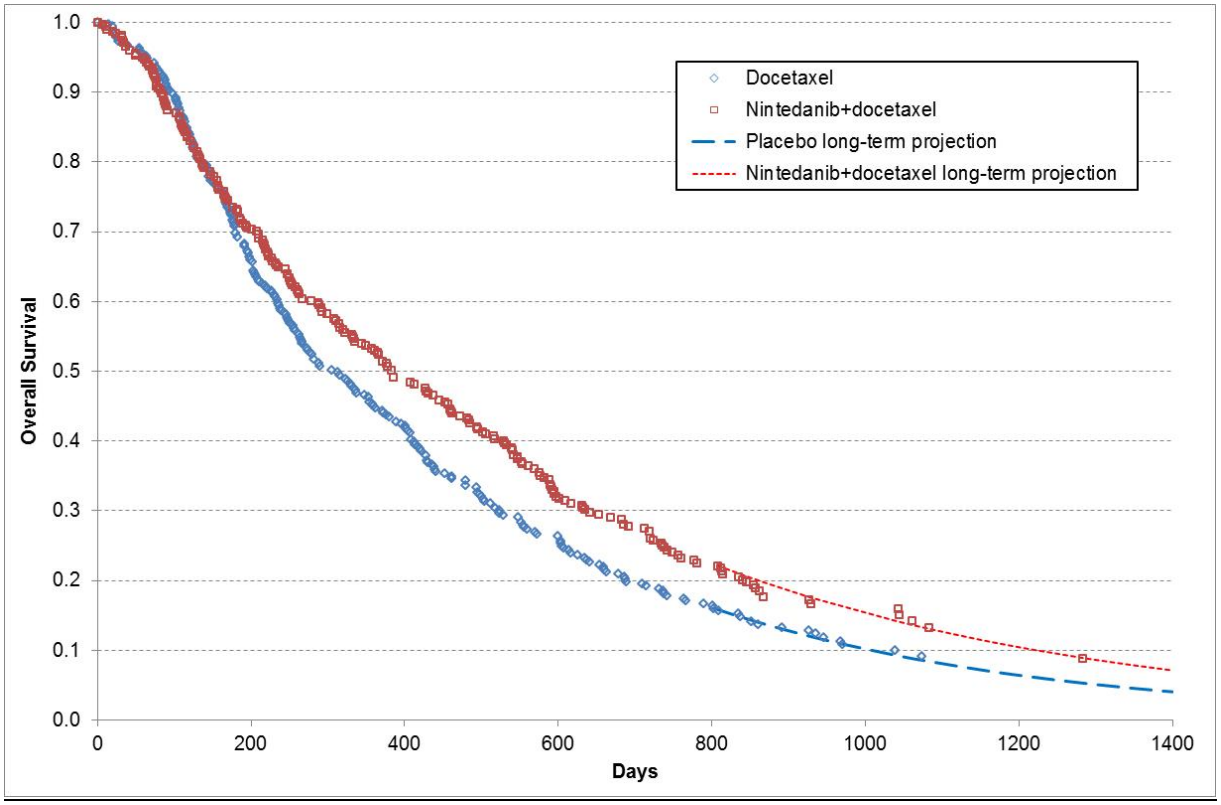


Figure 11 OS plot with ERG long-term projections for nintedanib plus docetaxel vs docetaxel

5.5.3 Progression-free survival estimation

The company's model base-case comparison of nintedanib plus docetaxel vs docetaxel indicates a gain in (undiscounted) PFS of 28.6 days, based on calibrating a Log-Normal hazard distribution to each trial arm and applying these to represent patient experience until all patients have died or suffered disease progression.

Examination of the PFS temporal profile (Figure 12) indicates that although the addition of nintedanib to docetaxel therapy generates a short-term delay in disease progression for some patients (i.e. the PFS curves begin to separate), subsequently this advantage progressively dissipates until the PFS experience of patients in the two trial arms is indistinguishable. Here, the extent of advantage in mean PFS can be readily estimated directly from the K-M analysis results by comparing the AUC estimates up to the point when the curves converge. The ERG identified that convergence occurred at day 375, and the difference in AUC at this time is 36.4 days. This suggests a small additional PFS benefit compared with the gain obtained in the company's model (28.6 days).

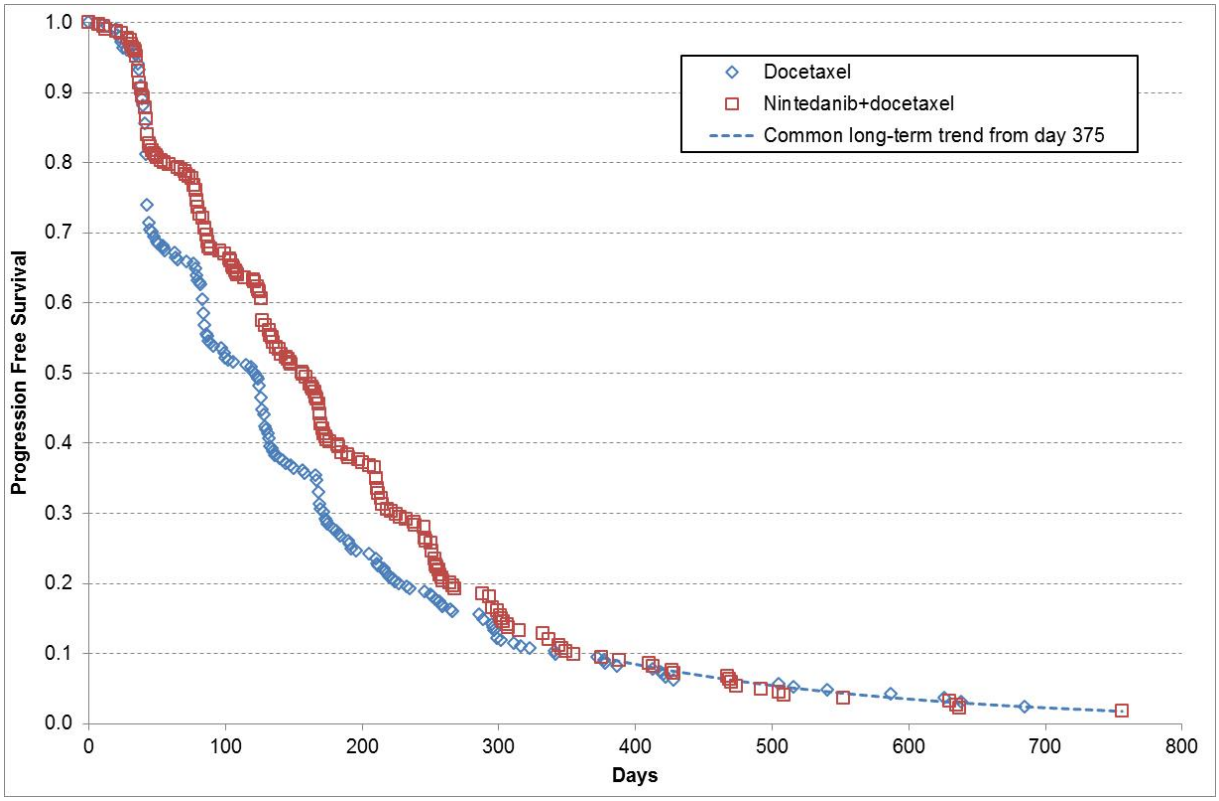


Figure 12 PFS plot with ERG common long-term projection for nintedanib plus docetaxel vs docetaxel

In order to apply the results of this re-analysis to the company's model, the ERG carried out a K-M landmark analysis for patients who were still progression-free at day 375. This indicated that a common long-term exponential model is appropriate for use in both

treatment arms from day 375 onwards, and this is shown in Figure 12. However, it should be noted that any projective model could be employed to both arms of the trial without any effect on the cost-effectiveness analysis as the incremental gain in PFS is unaffected.

Replacing the company's preferred PFS model with the ERG's approach has a modest impact on the cost-effectiveness of nintedanib plus docetaxel compared with docetaxel. The incremental discounted cost per patient is increased by [redacted] while the incremental discounted QALY gain is increased by [redacted], resulting in the estimated ICER increasing from £50,776 per QALY gained to £52,445 per QALY gained.

5.5.4 Time on treatment estimation

The ERG has used the same approach to obtain an accurate representation of the duration of treatments in the arms of LUME-Lung 1.[24] This approach uses the K-M results directly until a long-term exponential trend is established for projection until all patients have died (shown in Figure 13 to Figure 15).

Replacing the company's preferred exponential model with the ERG's approach has a modest impact on the cost-effectiveness of nintedanib plus docetaxel compared with docetaxel. The discounted cost per patient is increased in both treatment arms, so that the incremental cost per patient rises by [redacted], resulting in the estimated ICER increasing from £50,776 per QALY gained to £51,930 per QALY gained.

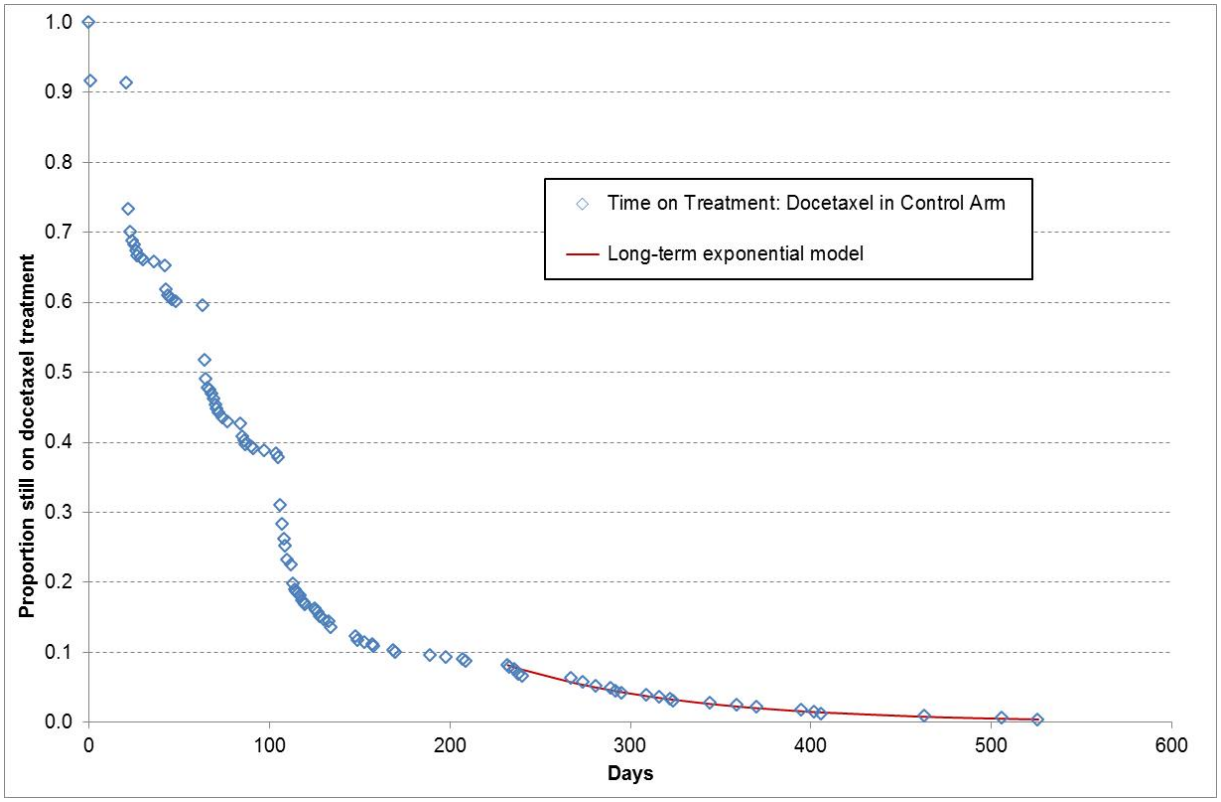


Figure 13 Time on Treatment: docetaxel in control arm with ERG long-term projection

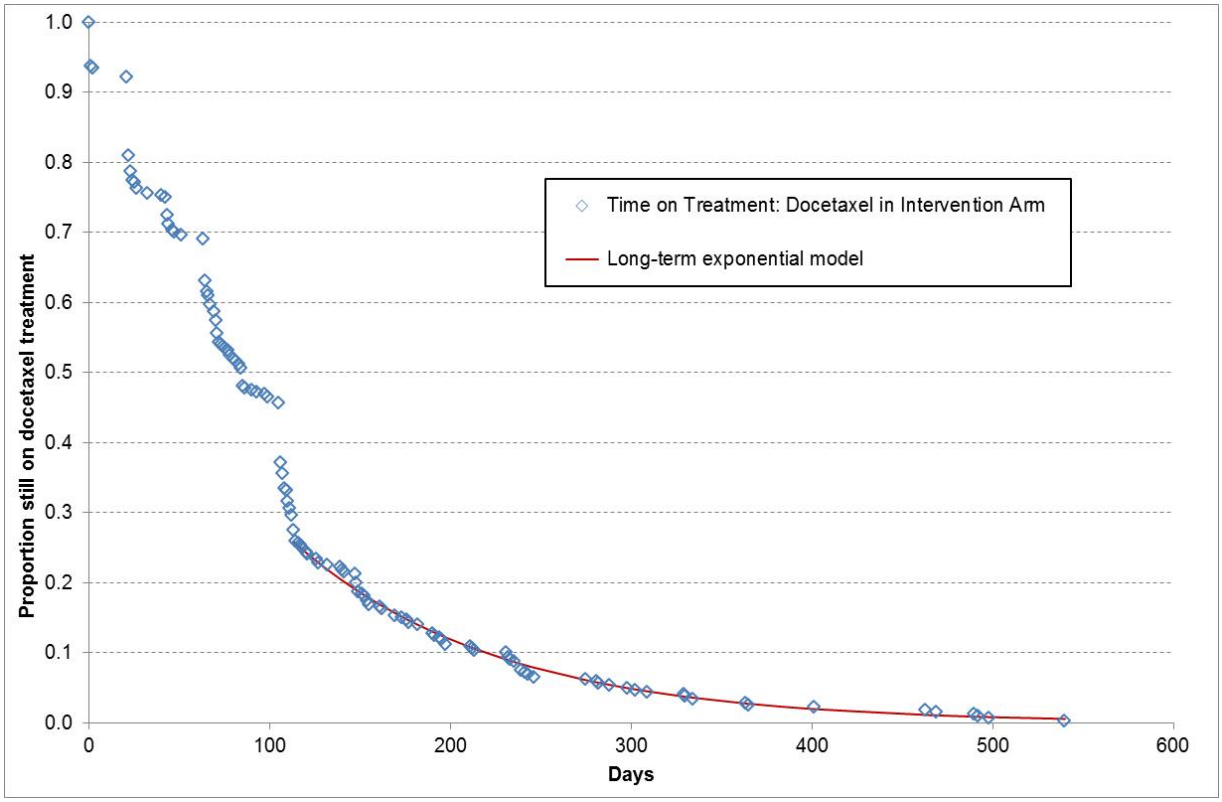


Figure 14 Time on Treatment: docetaxel in intervention arm with ERG long-term projection

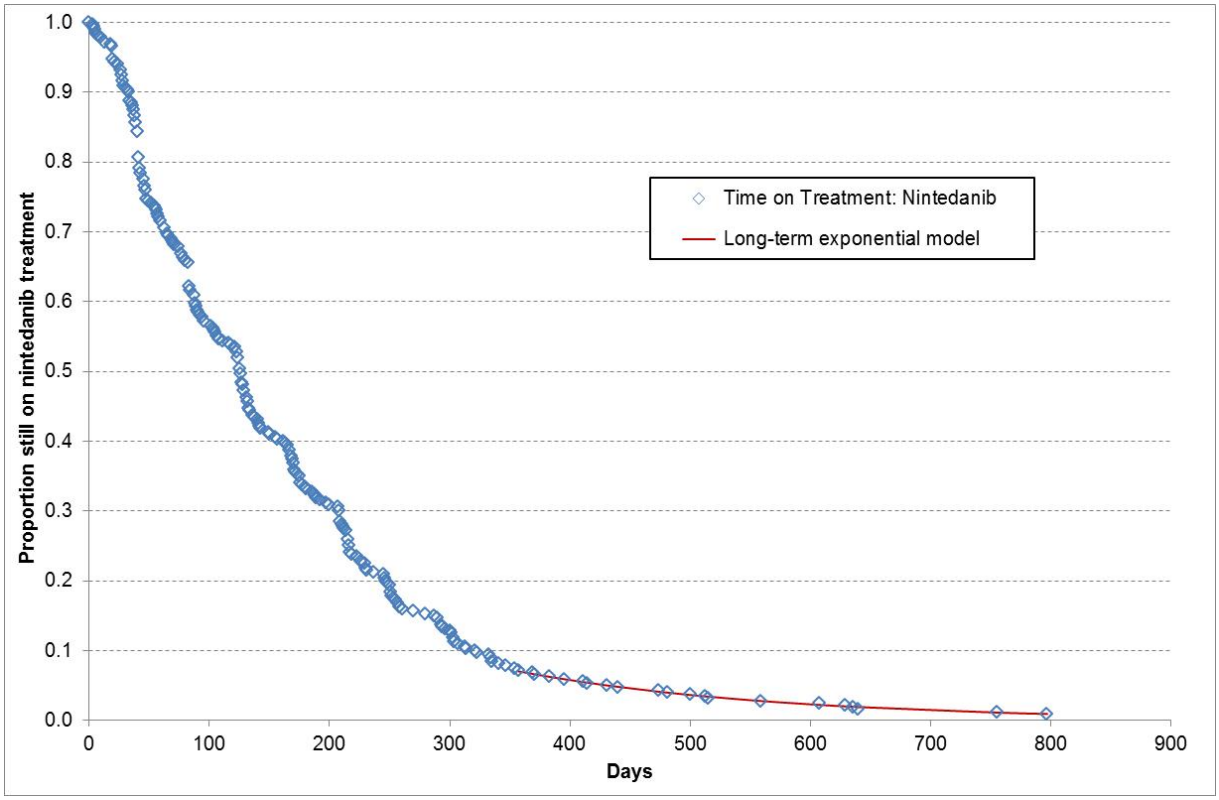


Figure 15 Time on Treatment: nintedanib in intervention arm with ERG long-term projection

5.5.5 Incorrect mid-cycle adjustment for drug costs

In the company's model the costs of both docetaxel and nintedanib are calculated for the average number of patients on treatment across each cycle. This mid-cycle adjustment for docetaxel is not accurate since three-weekly docetaxel is delivered on the first day of each cycle. Clinical advice also indicates that nintedanib doses are also dispensed on the first day of each cycle. The effect of this error is to under-estimate the quantity and cost of drugs used throughout the trial and in both arms of the comparison.

When this error is remedied the incremental discounted cost per patient increases by [redacted], and the estimated ICER increases from £50,776 per QALY gained to £53,839 per QALY gained.

5.5.6 Calculations for drug costs per dose

The average cost per dose of docetaxel delivered has been calculated by the company according to the distribution of body surface area (BSA) within the relevant UK population as a whole, though neglecting the important distinction between males and females whose mean BSA differs sufficiently to affect the overall estimated cost per dose. In addition, only the cost of the full 75mg/m^2 dose is estimated and adjusted using a relative dose intensity (RDI) index from trial data. It is more accurate to estimate the cost of a reduced dose (60mg/m^2) and then create a weighted average cost based on the balance between full and reduced doses recorded in the trial. The ERG has therefore re-estimated the overall average cost per dose of docetaxel using separate male and female subgroups, and also re-estimated the RDI multiplier to match the balance of full and reduced doses.

In addition, the ERG received clinical advice from a centre currently using nintedanib indicating that in practice nintedanib tablets are dispensed to patients at the time of docetaxel administration in blister packs sufficient to self-treat until the date of the next docetaxel dose (i.e. for days 2 to 21 of each cycle). Any missed doses are unlikely to alter the dispensing pattern, and thus missed doses will not alter the amount and cost of product dispensed. Therefore a reduction in cost through a RDI index is inappropriate. The company's method of calculating the cost per dose of nintedanib does not take account of the effect of three separate doses used (full dose, and two reduced doses) when part packs are dispensed as required at each cycle visit. Using data from LUME-Lung 1[24] of the differing balance between dose levels at each cycle, it has been possible to estimate an overall mean cost of treatment with nintedanib per cycle.

Applying these revised ERG parameter values to the company's base-case model, results in a [redacted] increase in the incremental cost per patient, and raises the estimated ICER from £50,776 per QALY gained to £52,587 per QALY gained.

5.5.7 Cost of treating febrile neutropenia

The company's model includes an estimated cost of treatment for grade 3/4 febrile neutropenia of £2,012.10 per patient affected, based on clinical advice. This figure is substantially lower than the average cost estimated by the NICE Decision Support Unit in 2007[74] which was revised for the recent MTA of second-line chemotherapy in NSCLC.[75] The ERG further updated the DSU estimate using National Reference costs[69] for 2012/13, to a mean cost per episode of £5,240.40 and mean cost per patient of £7,352.54 (assuming 1.4 episodes per patient).

Using these revised cost estimates in the company's model increases the incremental cost of nintedanib plus docetaxel vs docetaxel by £130 per patient, and raises the base-case ICER from £50,776 per QALY gained to £51,372 per QALY gained.

5.5.8 Monitoring cost

In the company's model the ERG has observed that there is a discrepancy between the cost of disease monitoring in patients who are on active treatment but who have not yet suffered disease progression (i.e. patients with stable disease). The model assigns a cost of £188 per cycle to patients in the nintedanib plus docetaxel arm and assigns a value of £205 per cycle to patients in the docetaxel arm, when the only difference in treatment relates to self-administered nintedanib tablets. On examination, it appears that the advice given by the company's clinical expert, concerning additional physician monitoring every 2 to 3 months for patients who have completed active treatment but who have not yet suffered disease progression, has been wrongly applied to patients still on active treatment with docetaxel. Moreover the unit cost employed is erroneously that of a GP consultation not an oncology out-patient visit.

When this misallocation is corrected, the incremental cost per QALY gained for nintedanib plus docetaxel vs docetaxel increases by £364, and the base-case ICER increases from £50,776 per QALY gained to £51,140 per QALY gained.

5.5.9 Discounting method

The submitted model applies discounting at a different rate for every 3-week model cycle based on the time elapsed. By convention in the UK, in line with the use of annual public sector budgets, discounting is applied annually considering the first 12 month period as involving current costs and each subsequent 12 month period requiring discounting for an additional year's delay. In some models with extended survival and multiple future events the choice of discounting method may have a large impact on the modelled ICER. However, using annual discounting in the company's model for this appraisal has only a minor effect, reducing the estimated base-case ICER from £50,776 per QALY gained to £50,532 per QALY gained.

5.5.10 Disutility of fatigue related adverse events

The key AEs identified from LUME-Lung1[24] were CTCAE grade 3 or 4 diarrhoea and fatigue. The company's analysis of EQ-5D utility data indicates that the estimated disutility for diarrhoea is low (-0.04). By contrast CTCAE grade 3 or 4 fatigue appears to have the largest effect in terms of patient disutility, amounting to an average of -0.21 across both treatment arms. However, Table 24 of the company's submitted Health Economics report[71] indicates a

large statistically significant difference between effect sizes in the two treatment arms: -0.326 for nintedanib plus docetaxel vs -0.101 for docetaxel, suggesting that patients experiencing serious fatigue on treatment are more seriously affected by the combination therapy. The company's model uses the overall average disutility estimate for both regimens. The ERG has applied a model amendment to apply the separate disutility values, resulting in a small reduction in the incremental QALY gain for nintedanib plus docetaxel vs docetaxel, and a corresponding increase of £54 per QALY gained in the base-case estimated ICER (from £50,776 per QALY gained to £50,830 per QALY gained).

5.5.11 Specification of second-line stable disease costs

Details of health care costs incurred by patients in various health states were derived from evidence provided by a panel of clinical advisors. A summary of this evidence is included in the appendices document accompanying the CS[1] (pages 70 to 77). A comparison between the details shown in the advisors evidence and the calculations used in the model to estimate average costs reveal important differences with respect to the cost of care for patients who have ceased active treatment and remain in a stable condition without evidence of further disease progression. The submitted model includes an assumption that these patients will require an hour of palliative nursing care every week and a bone scan every 3 weeks. This is in addition to a chest X-ray every 2 to 3 months and a physician visit once a year. The evidence of the clinical advisors only refers to the latter two items, and it appears that the palliative care and bone scans are included in error. Correcting this error substantially reduces the care costs per patient for any patient in a stable condition after second-line treatment. This has the effect of increasing the incremental cost per patient by [redacted] and increasing the estimated ICER for nintedanib plus docetaxel vs docetaxel from £50,776 per QALY gained to £53,470 per QALY gained.

5.5.12 Duration of docetaxel treatment

The company's base-case model follows the protocol of the LUME-Lung1[24] trial in permitting unlimited continuation of docetaxel treatment in either trial arm. One patient in the nintedanib plus docetaxel arm received 45 cycles of docetaxel, and one patient in the docetaxel monotherapy arm received 42 cycles. In the UK, standard clinical practice is to limit docetaxel to a maximum of four cycles per patient to avoid unacceptable AEs and associated poor QoL. The company's model includes an option to restrict docetaxel therapy to a maximum of four cycles. However, a formula error has been detected in the company's model which implements a limit of five rather than four cycles. The ERG has applied its own model adjustment which limits treatment to four cycles. It should be noted that this feature only affects the cost of drug acquisition and administration; it does not address the issue of

whether limiting exposure to docetaxel will impact on the prognosis of patients, nor does it attempt to adjust for consequent changes in AEs and the resulting cost and QoL effects.

This modification to the company's model reduces the base-case incremental cost per patient by [redacted], and reduces the estimated ICER for nintedanib plus docetaxel vs docetaxel from £50,776 per QALY gained to £48,060 per QALY gained.

5.5.13 Comparison with erlotinib

As noted in sections 2.2, 3.3 and 4.4.1 above, the ERG does not consider a comparison of nintedanib plus docetaxel to erlotinib is appropriate to the decision problem, a view also shared by the company. Nevertheless the company has attempted to incorporate into their model a facility to compare the relative cost-effectiveness of erlotinib and nintedanib plus docetaxel, as indicated in the NICE scope. In the absence of a trial directly comparing these regimes, it was necessary to attempt an MTC to generate estimated outcomes for patients treated with erlotinib, consistent with all relevant information in related studies. The base-case MTC includes three RCTs[56,59,62] in addition to the LUME-Lung1[24] trial: JMEI[56] which compared docetaxel with pemetrexed, WSY001[62] which compared pemetrexed with erlotinib and TAILOR[59] which compared docetaxel with erlotinib (see also Figure 1, page 39). This provides two connection pathways linking nintedanib plus docetaxel to erlotinib:

1) LUME-Lung1[24] ⇒ JMEI[56] ⇒ WSY001[62]
2) LUME-Lung1[24] ⇒ TAILOR[59]

In principle, it is desirable to employ this network to generate HRs for each time-to-event outcome as a basis for estimating the corresponding survival profiles for erlotinib, consistent with that obtained for nintedanib plus docetaxel in the LUME-Lung1[24] trial.

Time on Treatment

Employing a network may be possible for OS and PFS, but is not feasible for ToT of erlotinib, since none of the connecting trial reports (for JMEI,[56] WSY001[62] and TAILOR[59]) report results for this outcome. Instead the company has assumed that a simple exponential function is appropriate for ToT in all treatments and have calibrated this function for each trial based on an estimated mean number of treatments per patient. It has already been demonstrated in section 5.5.4 that such an assumption is not correct in the case of the LUME-Lung1[24] trial and there is no reason to believe that it would be any more successful in the other trials in the network. In particular, the company modellers have assigned a parameter value for erlotinib consistent with a mean number of erlotinib cycles (i.e. 28 days) taken from the ERG report for NICE assessment TA162,[19] without recognising that this

figure was obtained indirectly from PFS trial data (which may overstate ToT) and that the ERG on that occasion employed a 2-phase exponential model with a high risk of discontinuation in the first 11 weeks, and a lower risk thereafter. Without access to detailed patient-level ToT data for each of the studies in the MTC, it is not possible to rectify the substantial uncertainty associated with the estimation of drug acquisition costs in the company model.

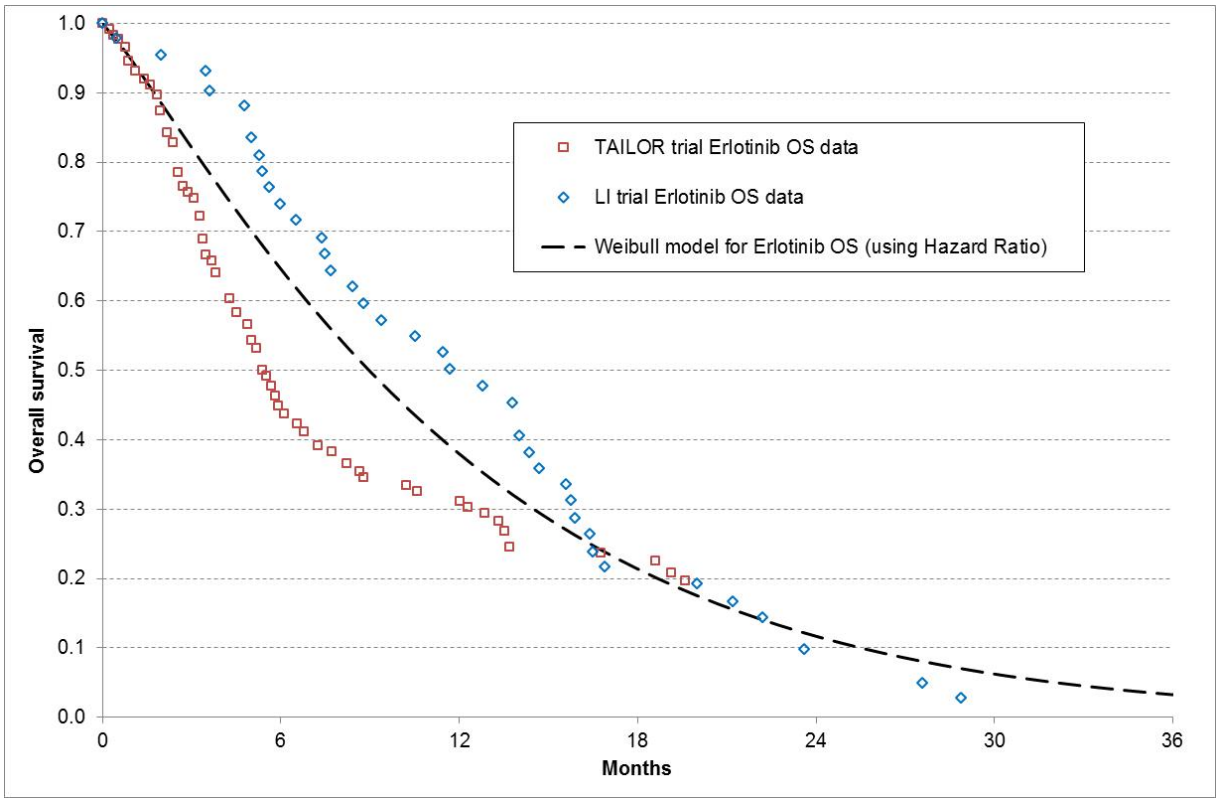
Overall survival and progression-free survival

Meta-analysis of time-to-event data in a network relies on a number of conditions being met:

- Within each trial the assumption of proportional hazards should apply
- Between trials featuring the same treatment at nodes in the MTC, treatment outcomes should be equivalent (i.e. both proportional hazards and very similar outcomes at all time points)
- If a parametric survival function is to be propagated through the network then it should be inherently proportional hazard compliant (i.e. Weibull or Exponential)

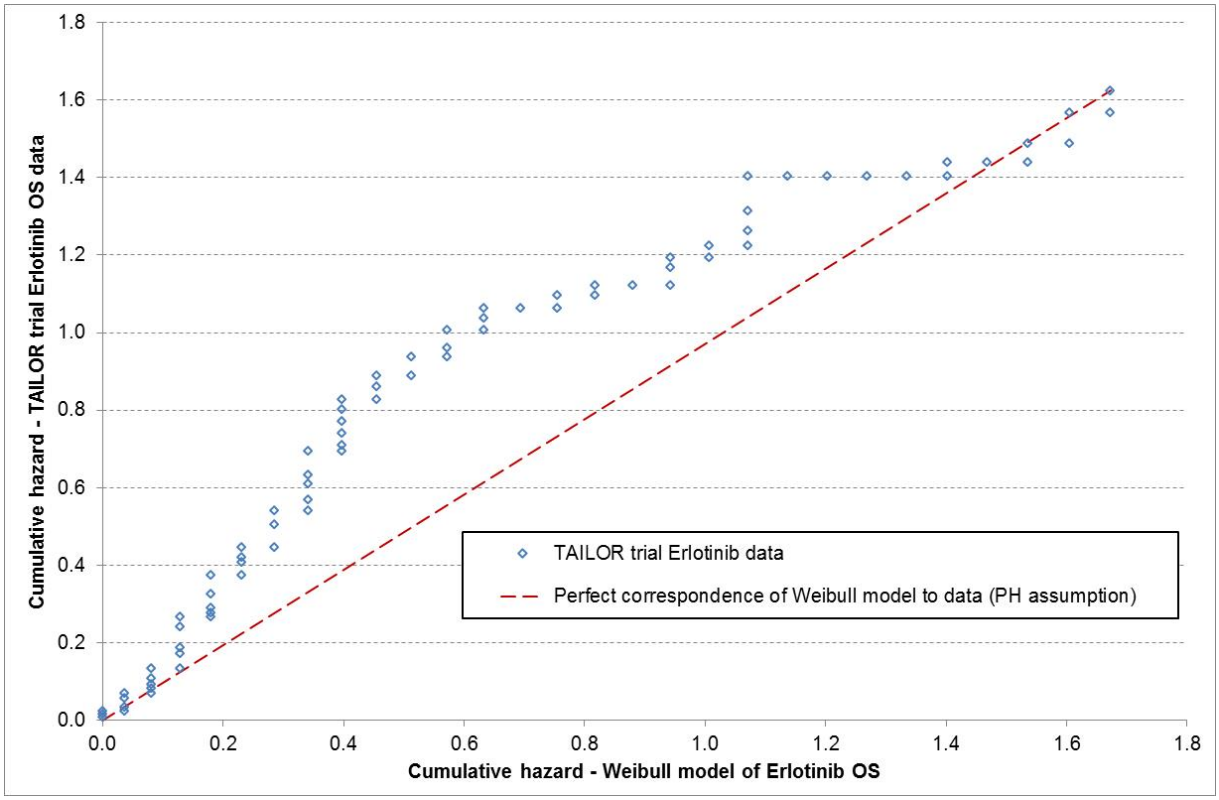
For the company's MTC of OS, a Weibull formulation was therefore used, despite this not appearing to give the best match to the nintedanib plus docetaxel LUME-Lung1[24] trial OS data. If all the above criteria are met, the resulting time-to-death profile should be a Weibull curve adjusted by an overall HR (0.64 for nintedanib plus docetaxel vs erlotinib in OS) so that it is consistent with the corresponding profiles for erlotinib in both the TAILOR[59] and WSY001[62] trials.

Figure 16 compares the fitted Weibull model for erlotinib with the erlotinib Kaplan-Meier data from the TAILOR[59] and WSY001[62] trials. It is apparent that during the first 18 months there are large differences between the three profiles. It is also possible to test whether the proportional hazards assumption is violated in both arms of the network. Figures 17 and 18 show plots of cumulative hazard data from each erlotinib trial arm against the cumulative hazard at the same time points from the Weibull OS model. The proportional hazards assumption is confirmed if the data points (corresponding to trial events) all lie close to and evenly spaced around the diagonal 'proportionality' line. It is clear that for both the erlotinib trials (TAILOR[59] and WSY001[62]) the proportional hazards assumption is seriously violated. This is likely to have been caused by multiple problems, including non-proportional hazards results in LUME-Lung1[24] trial OS data (as discussed in Appendix 7), proportional hazards violations in one or more of the other three trials in the MTC and non-equivalence of trial arms at network nodes.



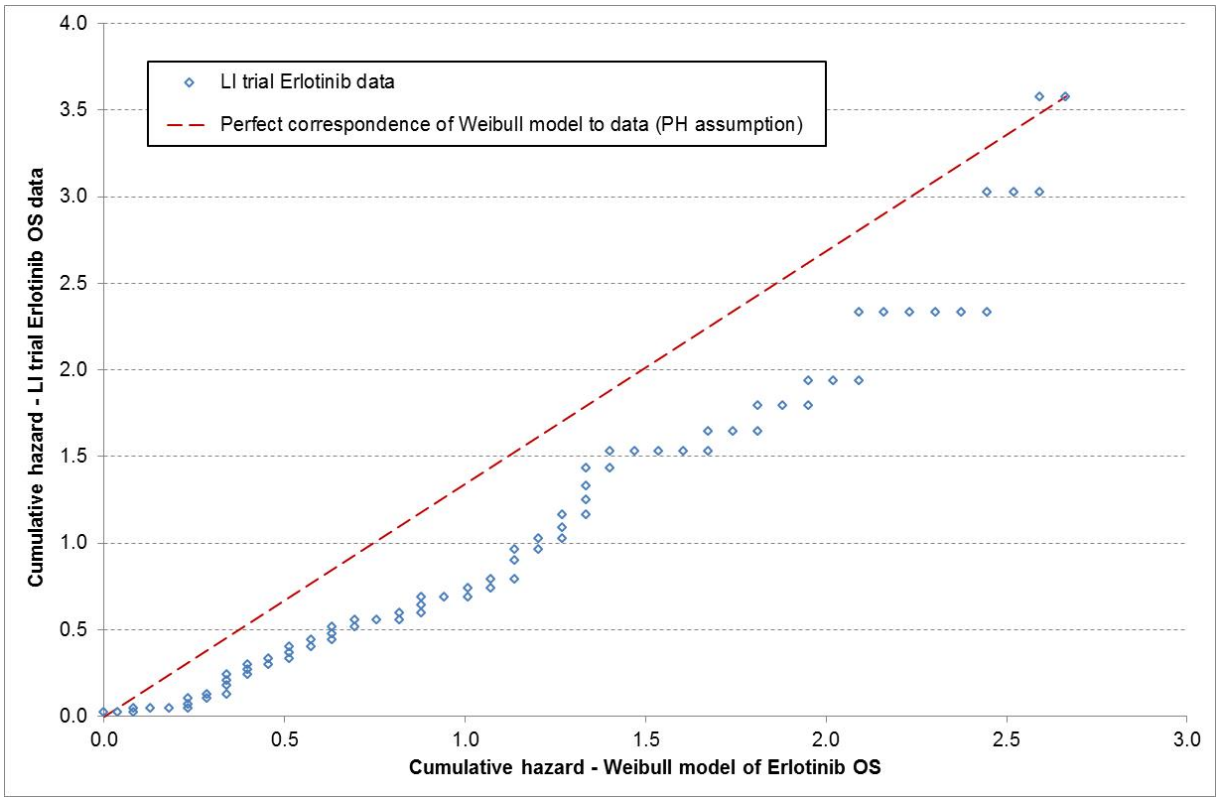
LI=WSY001; OS=overall survival

Figure 16 Weibull OS model for erlotinib-treated patients compared with original trials data



OS=overall survival; PH=proportional hazards

Figure 17 Proportionality check of Weibull erlotinib OS model vs erlotinib data from TAILOR trial



LI=WSY001; OS=overall survival; PH=proportional hazards

Figure 18 Proportionality check of Weibull erlotinib OS model vs erlotinib data from WSY001 trial

These diagnostic checks indicate not only that the estimated OS model estimates are inconsistent within the evidence network, but that the Weibull functional form calibrate from LUME-Lung1[24] trial data when transmitted through the network does not accord with the outcome patterns seen in other network trials. This calls into question the use of both Weibull parametric form and the HR for erlotinib vs nintedanib plus docetaxel estimated from the network.

The potential impact of alterations in OS far outweigh all other aspects of the model (see 5.5.2 above and Table 40 below) and therefore the importance of this finding cannot be over-estimated. The ERG has not been able to complete a full assessment of the PFS network in a similar manner due to time limitations, but early indications are that similar inconsistencies are present. However, PFS data are more complete and have less influence on cost-effectiveness results than OS.

Unfortunately, these problems with the evidence networks are so fundamental that it is not possible to rectify them and modify the company's model to provide improved estimates of OS, PFS and the relative cost-effectiveness of nintedanib plus docetaxel and erlotinib.

5.6 Conclusions of the cost-effectiveness section

Although the structure of the economic model submitted by the company is generally appropriate, the ERG is concerned by the number of implementation errors that have come to light with important consequences for the economic results generated. The ERG has identified eleven specific aspects of the submitted base-case model that are subject to challenge, or involve implementation errors. In each case an appropriate amendment has been introduced into the company's model with results ranging from minor changes to important and substantial changes to the estimated ICER per QALY gained.

Neither the company nor the ERG considers a comparison of nintedanib plus docetaxel to erlotinib to be appropriate to the decision problem. Nevertheless, this was specified in the NICE scope and the company has therefore undertaken such a comparison. However, the ERG considers that this is seriously flawed due to inconsistencies apparent in the available time-to-event data leading to conflicting results from the MTC. The ERG has applied other relevant amendments to the submitted model for this comparison, but the uncertainty in OS, PFS and ToT probably far outweighs all other effects but cannot be quantified.

6 ADDITIONAL WORK UNDERTAKEN BY THE ERG

Table 40 summarises the effects of the various ERG amendments made to the company's decision model (see also Appendix 8). The consequence of applying each proposed amendment is shown separately for comparison with the company's base-case analysis. The joint effect of applying all ERG changes to the model simultaneously is included. In addition, a second summary result is provided excluding the limitation of docetaxel treatment to four cycles because this change reflects an issue of principle (clinical evidence vs UK practice), and because the impact of applying a model revision is necessarily incomplete (the ERG cannot estimate what the effect might be on outcomes of restricting treatment).

Generally these amendments result in increased costs (both absolute and incremental) and/or reduced outcomes (survival and QALYs) and hence lead to increases in the estimated ICER per QALY gained. The company's base-case ICER (£50,776 per QALY gained) is increased to either £85,292 per QALY gained with all revisions applied, or to £82,995 per QALY gained if no limit is placed on the number of cycles of docetaxel treatment allowed.

The most influential change is the application of the ERG OS estimates. If this revision is not accepted, the revised ICER using the other ten revisions becomes £62,719 per QALY gained. The ERG's estimate of the gain in undiscounted mean OS is 3.05 months.

Cost-effectiveness results of applying the non-Time To Event ERG amendments are detailed in Table 41, with a full sensitivity analysis for a range of possible patient access scheme discounts on the list price of erlotinib in Table 42. It should be borne in mind that were it possible to estimate the mean OS for patients treated with erlotinib rather than docetaxel monotherapy in second-line chemotherapy, it is quite likely that the estimated incremental gain in life-years would diminish and the estimated ICER rise substantially.

Table 40 Cost-effectiveness results for nintedanib plus docetaxel vs docetaxel with ERG revisions to company's base-case comparison in the adenocarcinoma population

Model scenario & ERG revisions	Nintedanib + docetaxel			Docetaxel			Incremental			ICER	ICER
	Cost	QALYs	Life years	Cost	QALYs	Life years	Cost	QALYs	Life years	£/QALY	Change
Company's base-case	[redacted]	[redacted]	**1.810**	[redacted]	[redacted]	**1.419**	**+ £11,051**	**+ 0.218**	**+ 0.391**	**£50,776**	-
1) ERG OS estimates	[redacted]	[redacted]	1.493	[redacted]	[redacted]	1.238	+ £10,497	+ 0.153	+ 0.255	£68,587	**+ £17,811**
2) ERG PFS estimates	[redacted]	[redacted]	1.810	[redacted]	[redacted]	1.419	+ £11,527	+ 0.220	+ 0.391	£52,445	**+ £1,669**
3) ERG ToT estimates	[redacted]	[redacted]	1.810	[redacted]	[redacted]	1.419	+ £11,298	+ 0.218	+ 0.391	£51,930	**+ £1,154**
4) Mid-cycle adjustment	[redacted]	[redacted]	1.810	[redacted]	[redacted]	1.419	+ £11,717	+ 0.218	+ 0.391	£53,839	**+ £3,062**
5) Cost of treatment doses	[redacted]	[redacted]	1.810	[redacted]	[redacted]	1.419	+ £11,445	+ 0.218	+ 0.391	£52,587	**+ 1,811**
6) Febrile neutropenia cost	[redacted]	[redacted]	1.810	[redacted]	[redacted]	1.419	+ £11,180	+ 0.218	+ 0.391	£51,372	**+ £595**
7) Monitoring cost	[redacted]	[redacted]	1.810	[redacted]	[redacted]	1.419	+ £11,130	+ 0.218	+ 0.391	£51,140	**+ £364**
8) Discounting method	[redacted]	[redacted]	1.810	[redacted]	[redacted]	1.419	+ £11,189	+ 0.221	+ 0.391	£50,532	**-£244**
9) Disutility of fatigue	[redacted]	[redacted]	1.810	[redacted]	[redacted]	1.419	+ £11,051	+ 0.217	+ 0.391	£50,830	**+ £54**
10) Stable disease costs	[redacted]	[redacted]	1.810	[redacted]	[redacted]	1.419	+ £11,637	+ 0.218	+ 0.391	£53,470	**+ £2,693**
11) Docetaxel ≤4 cycles	[redacted]	[redacted]	1.810	[redacted]	[redacted]	1.419	+ £10,452	+ 0.217	+ 0.391	£48,060	**-£2,716**
Base-case + revisions 1-10	[redacted]	[redacted]	**1.493**	[redacted]	[redacted]	**1.238**	**+ £13,087**	**+ 0.158**	+ 0.255	**£82,995**	**+ £32,219**
Base-case + all revisions	[redacted]	[redacted]	**1.493**	[redacted]	[redacted]	**1.238**	**+ £13,437**	**+ 0.158**	**+ 0.255**	**£85,292**	**+ £34,516**

Costs and QALYs discounted; Life years undiscounted
OS=overall survival; PFS=progression-free survival; QALYs=quality adjusted life years; ToT=time on treatment

Table 41 Cost-effectiveness results for nintedanib plus docetaxel vs erlotinib with ERG revisions to company's base-case comparison in the adenocarcinoma population

Model scenario & ERG revisions	Nintedanib + docetaxel			Erlotinib			Incremental			ICER	ICER
	Cost	QALYs	Life years	Cost	QALYs	Life years	Cost	QALYs	Life years	£/QALY	Change
Company's base-case	[redacted]	[redacted]	**1.445**	[redacted]	[redacted]	**0.979**	**£7,571**	**0.280**	**0.465**	**£27,008**	-
1) ERG OS estimates	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**
2) ERG PFS estimates	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**
3) ERG ToT estimates	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**
4) Mid-cycle adjustment	[redacted]	[redacted]	1.445	[redacted]	[redacted]	0.979	£7,815	0.280	0.465	£27,878	**+ £870**
5) Cost of treatment doses	[redacted]	[redacted]	1.445	[redacted]	[redacted]	0.979	£7,926	0.280	0.465	£28,275	**+ £1,267**
6) Febrile neutropenia cost	[redacted]	[redacted]	1.445	[redacted]	[redacted]	0.979	£7,897	0.280	0.465	£28,173	**+ £165**
7) Monitoring cost	**NA**	**NA**	**NA**	**NA**	**NA**	**NA**	**NA**	**NA**	**NA**	**NA**	**NA**
8) Discounting method	[redacted]	[redacted]	1.445	[redacted]	[redacted]	0.979	£7,679	0.285	0.465	£26,927	**-£81**
9) Disutility of fatigue	[redacted]	[redacted]	1.445	[redacted]	[redacted]	0.979	£7,571	0.280	0.465	£27,020	**+ £12**
10) Stable disease costs	[redacted]	[redacted]	1.445	[redacted]	[redacted]	0.979	£7,576	0.280	0.465	£27,027	**+ £19**
11) Docetaxel ≤4 cycles	[redacted]	[redacted]	1.445	[redacted]	[redacted]	0.979	£7,069	0.283	0.465	£24,975	**-£2,033**
Base-case + revisions 4-11	[redacted]	[redacted]	**1.445**	[redacted]	[redacted]	**0.979**	**£8,147**	**0.288**	**0.465**	**£28,307**	**+ £1,299**

Costs and QALYs discounted; Life years undiscounted
NA = not applicable; OS=overall survival; PFS=progression-free survival; QALYs=quality adjusted life years; ToT=time on treatment; X = meaningful amendments for time-to-event estimates are not possible due to unreliable data network or absence of data

Table 42 Cost-effectiveness results for nintedanib plus docetaxel vs erlotinib with ERG revisions to company's base-case comparison in the adenocarcinoma population: sensitivity of ICER to different patient access scheme discount levels for erlotinib.

Model scenario & ERG revisions	**Patient access scheme discount for erlotinib**										
	0%	**5%**	**10%**	**15%**	**20%**	**25%**	**30%**	**35%**	**40%**	**45%**	**50%**
Company's base-case	**£27,008**	**£27,939**	**£28,870**	**£29,802**	**£30,733**	**£31,664**	**£32,596**	**£33,527**	**£34,458**	**£35,390**	**£36,321**
1) ERG OS estimates	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**
2) ERG PFS estimates	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**
3) ERG ToT estimates	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**
4) Mid-cycle adjustment	£27,878	£28,902	£29,926	£30,950	£31,975	£32,999	£34,023	£35,047	£36,071	£37,095	**£38,119**
5) Cost of treatment doses	£28,275	£29,206	£30,138	£31,069	£32,000	£32,932	£33,863	£34,794	£35,726	£36,657	**£37,588**
6) Febrile neutropenia cost	£28,173	£29,104	£30,035	£30,967	£31,898	£32,830	£33,761	£34,692	£35,624	£36,555	**£37,486**
7) Monitoring cost	**NA**	**NA**	**NA**	**NA**	**NA**	**NA**	**NA**	**NA**	**NA**	**NA**	**NA**
8) Discounting method	£26,927	£27,851	£28,775	£29,699	£30,623	£31,547	£32,471	£33,395	£34,319	£35,243	**£36,167**
9) Disutility of fatigue	£27,020	£27,951	£28,883	£29,815	£30,747	£31,678	£32,610	£33,542	£34,474	£35,405	**£36,337**
10) Stable disease costs	£27,027	£27,958	£28,890	£29,821	£30,752	£31,684	£32,615	£33,546	£34,478	£35,409	**£36,340**
11) Docetaxel ≤4 cycles	£24,975	£25,897	£26,820	£27,742	£28,664	£29,587	£30,509	£31,431	£32,354	£33,276	**£34,198**
Base-case + revisions 4-11	**£28,307**	**£29,314**	**£30,320**	**£31,327**	**£32,334**	**£33,341**	**£34,348**	**£35,354**	**£36,361**	**£37,368**	**£38,375**

Costs and QALYs discounted; Life years undiscounted
NA = not applicable; OS=overall survival; PFS=progression-free survival; QALYs=quality adjusted life years; ToT=time on treatment; X = meaningful amendments for time-to-event estimates are not possible due to unreliable data network or absence of data

7 END OF LIFE

The company makes a case that nintedanib plus docetaxel meets the criteria set by NICE for end of life treatment. Namely:

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

The company states on page 288 of the CS:[1]

- Patients with advanced NSCLC have a short life expectancy of less than 24 months on average. Using the extrapolated results from the LUME-Lung 1[24] trial data implemented in the cost-effectiveness model, the median OS of patients on docetaxel monotherapy (current standard of care) is 10.23 months and the mean OS is 15.96 months.
- The total eligible population for nintedanib plus docetaxel in England is 703.
- Extension to life due to nintedanib plus docetaxel vs docetaxel monotherapy in the target population with the base-case assumptions within the model is a mean of 3.96 months. The extension in OS over erlotinib is a mean of 5.16 months.

The ERG agrees that patients with advanced NSCLC have a short life expectancy of less than 24 months. It also agrees that the patients who would be eligible for the treatment is small. As noted in section 5.5.2, by applying the K-M trial results using the AUC method until the long-term OS trends were established and then projecting remaining estimated survival using the exponential trends, the ERG calculated the mean extension in OS to be 3.05 months for the base-case analysis of nintedanib plus docetaxel vs docetaxel. The ERG were only able to carry out a partial comparison of nintedanib plus docetaxel to erlotinib for reasons outlined in section 5.5.13 (excluding the time-to-event outcomes known to be subject to the most uncertainty) and were therefore unable to derive a mean estimate for OS for nintedanib plus docetaxel vs erlotinib.

8 DISCUSSION

The NICE scope for this STA stipulates the population should be adults with locally advanced or metastatic NSCLC that has progressed following prior chemotherapy. The decision problem differs in that it is restricted to NSCLC with adenocarcinoma histology. It also includes patients with locally recurrent disease. The ERG considers both differences to be appropriate since they reflect the relevant population stipulated by the anticipated licensed indication for nintedanib plus docetaxel.[28] Based on the LUME-Lung 1[24] trial, the majority (94.2%) of these patients will have metastatic disease at the time of second-line treatment. The majority (85% to 90%) of such patients in England would be expected to have EGFR-negative disease,[30-32] [redacted].

The NICE scope also states that docetaxel and erlotinib are relevant comparators. The company notes the preliminary recommendation issued by NICE in February 2014 is that erlotinib should not be recommended for treating locally advanced or metastatic NSCLC that has progressed after non-targeted chemotherapy in people with tumours that are EGFR-negative. The ERG also notes the same recommendation from August 2014.[21] Furthermore, in current clinical practice in England, the majority of patients with EGFR-positive NSCLC receive erlotinib (or another TKI) as first-line treatment.[13-15] These patients would, therefore, be unlikely to receive erlotinib as a second-line therapy. Because nintedanib is administered in combination with docetaxel, patients in receipt of nintedanib must be fit enough to receive docetaxel. Such patients are, therefore, likely to be assessed as ECOG PS 0 to 1. The general opinion of clinical advisors to both the company and ERG is that patients who are sufficiently fit to tolerate treatment with docetaxel will receive docetaxel rather than erlotinib. In view of these factors, while the decision problem does include erlotinib as a comparator for secondary analyses, this is nevertheless considered by the company to be an irrelevant comparator to nintedanib plus docetaxel. The ERG agrees with the company that erlotinib is not a relevant comparator for the same reasons.

Evidence for the relative effectiveness of nintedanib plus docetaxel is derived from the LUME-Lung 1[24] trial which compares nintedanib plus docetaxel to placebo plus docetaxel. This, therefore, provides direct evidence for the clinical effectiveness of nintedanib plus docetaxel vs docetaxel alone. The trial appears to be of good quality and low risk of bias and reports that nintedanib plus docetaxel is superior to placebo plus docetaxel in terms of OS (median improvement of 2.3 months) and PFS (median improvement of 1.4 months). However, the ERG does not consider that the assumption of proportional hazards is consistent with the trial data, and therefore use of these results in cost-effectiveness

modelling should not be based implicitly or explicitly on this assumption. The reported gain in efficacy is accompanied by an increase in CTCAE grade ≥3 AEs and SAEs but these AEs are reported to be generally manageable. Some differences in HRQoL between treatment arms have been reported but none result in differences between arms in terms of overall global health status/QoL. The AEs of greatest concern are fatal AEs. More fatal AEs have been reported in the nintedanib plus docetaxel arm than in the placebo plus docetaxel arm. However, the numbers are small and the company is using ongoing surveillance to monitor this issue. Neutropenia and sepsis have also been identified as important risks.

One potential limitation with regard to the generalisability of the findings from LUME-Lung 1[24] to clinical practice in England relates to three of the exclusion criteria that the trial employed. First, patients with major pleural effusion were excluded. Second, patients with evidence of cavitary or necrotic tumours were excluded. Third, patients receiving therapeutic anticoagulation (except low dose heparin) or antiplatelet therapy (except for chronic low-dose therapy with acetylsalicylic acid ≤325mg/day) were excluded. Pleural effusions[50,51] and venous thromboembilsm [52] appear to predict poor prognosis; evidence of cavitary or necrotic tumours may also result in a worse prognosis, although cavitation may be a less strong prognostic factor.[49] These exclusion criteria may partially explain why, in LUME-Lung 1,[24] a higher proportion of patients than would be expected in clinical practice also received third-line treatment on disease progression. This may in turn also be an indicator that patients included in this trial were fitter than those generally seen in NHS clinical practice.

In order to derive an estimate for cost-effectiveness of nintedanib plus docetaxel to docetaxel alone, the company have developed a de novo partitioned survival Markov model, which incorporates data from LUME-Lung 1[24] alongside other published sources. The company's estimate of cost-effectiveness for nintedanib plus docetaxel vs docetaxel is £50,766 per QALY gained. However, the ERG identified a number of weaknesses in the company's model and is concerned about the number of implementation issues that it identified. The most important area in terms of its impact on the ICER related to OS estimation. Here inadequate information was provided about specific data sources used for SEER and LUCADA used to validate the long-term extrapolation of OS. Furthermore OS projection was based on the flawed assumption that there is constant hazard over time.

In total the ERG made 11 changes to the company's model. These related to: inappropriate methods used to project time-to-event outcomes (OS, PFS and time-on-treatment); mid-cycle adjustment error; inappropriate methods used to estimate cost of treatment doses; underestimate of true cost of febrile neutropenia; monitoring costs; non-UK standard approach to discounting; overall average disutility estimate for fatigue used for both

regimens; error in stable disease costs and erroneous restriction of docetaxel to four cycles. When all of the ERG's alterations are implemented, the ERG's revised estimate of cost-effectiveness for the comparison of nintedanib plus docetaxel with docetaxel is £85,292 per QALY gained. Independently, implementing each of the ERG's changes in the model results in ICERs ranging from £50,532 to £68,587 per QALY gained. The change which has the largest impact on the size of the ICER is the method used to estimate OS. If all of the other changes in the model are implemented, except replacement of the company's OS model, the ICER increases to £62,719 per QALY gained.

There is no direct evidence for the relative effectiveness of nintedanib plus docetaxel compared with erlotinib. In order to compare the relative clinical effectiveness for these two regimens, the company conducted a number of MTCs. However, the ERG has identified a number of uncertainties and weaknesses in relation to these MTCs. Relating to the generalisability of the trials to clinical practice, the ERG notes that only patients in one trial, LUME-Lung 1,[24] had received pemetrexed as a first-line treatment and even then, this was only a minority (18.8%). Pemetrexed is now the treatment of choice for adenocarcinoma patients with EGFR-negative disease who, as noted above, constitute the majority of adenocarcinoma patients in England.

There are also a number of methodological weaknesses with the conduct of the MTCs, the most important being that proportional hazards are presumed to hold throughout the MTC networks for both PFS and OS. As discussed above, the ERG has found that this is not the case within the LUME-Lung 1[24] and, as a consequence, any results generated comparing nintedanib plus docetaxel with erlotinib cannot be considered reliable. Important differences in trial and patient characteristics of trials included in the MTCs have also been observed which question the validity of the base-case, scenario and sensitivity analyses.

To compare nintedanib plus docetaxel to erlotinib, the results from the MTCs have been incorporated into the company's model. The company's estimate of cost-effectiveness for nintedanib plus docetaxel vs erlotinib is £27,008 per QALY gained. However, as discussed above, there are a number of methodological issues with the conduct of the MTCs which undermine any confidence in this estimate. Furthermore, in addition to those discussed above, additional problems have been identified in relation to ToT where again the assumption for proportional hazards is assumed. It is impossible to ascertain whether this is true for any trial other than LUME-Lung 1[24] as these data were not available for any other trial. However this assumption did not hold for LUME-Lung 1.[24] Furthermore, the ERG also established that not only is the assumption of proportional hazards for OS violated for LUME-Lung 1[24] but this is also violated for OS reported in two other trials (WSY001[62] and

TAILOR[59]) included in the MTC. Thus because of concerns about the relevance of erlotinib as a comparator and the appropriateness of the analyses conducted, the ERG only considers it feasible to estimate a reliable ICER per QALY gained using the direct trial data from LUME-Lung 1[24] for patients with locally advanced, metastatic or locally recurrent NSCLC of adenocarcinoma tumour histology after first-line chemotherapy.

9 OVERALL CONCLUSIONS

The ERG agrees that LUME-Lung 1[24] is a high quality trial that demonstrates the efficacy of nintedanib plus docetaxel over docetaxel for patients with adenocarcinoma after first-line chemotherapy. Based on the clinical and cost-effectiveness evidence available, the ERG only considers it feasible to estimate an ICER per QALY gained using the direct trial data from LUME-Lung 1[24] for this population. The ERG concludes that the comparison of nintedanib plus docetaxel vs docetaxel yields an ICER that is higher than £50,000 per QALY gained.

10 REFERENCES

1. Boehringer Ingelheim. Single technology appraisal (STA) submitted by Boehringer Ingelheim Ltd: Nintedanib in combination with docetaxel for the treatment of adult patients with locally advanced, metastatic or recurrent non-small cell lung cancer (NSCLC) of adenocarcinoma tumour histology after first-line chemotherapy. 2014.
2. American Cancer Society. Lung Cancer (Non-Small Cell). 2013; Available from: http://www.cancer.org/cancer/lungcancer-non-smallcell/detailedguide/non-small-cell-lung-cancer-what-is-non-small-cell-lung-cancer (Accessed 19/06/2013).
3. NHS Choices. Lung cancer. 2013 [updated 02/08/2013]; Available from: http://www.nhs.uk/Conditions/Cancer-of-the-lung/Pages/Introduction.aspx (Accessed 16/09/2014).
4. Roy Castle Lung Cancer Foundation. Lung Cancer Facts and Figures. 2014; Available from: http://www.roycastle.org/lung-cancer/Lung-Cancer-Facts-and-Figures (Accessed 16/09/2014).
5. Cancer Research UK. Lung Cancer Key Facts. 2014 [updated 22/02/13 ('What causes lung cancer?' section)]; Available from: http://www.cancerresearchuk.org/cancer-info/cancerstats/keyfacts/lung-cancer/ (Accessed 16/09/2014).
6. Jett JR, Schild SE, Keith RL, Kesler KA. Treatment of non-small cell lung cancer, stage IIIB: ACCP evidence-based clinical practice guidelines (2nd edition). Chest. 2007; 132:266s-76s.
7. Socinski MA, Crowell R, Hensing TE, Langer CJ, Lilenbaum R, Sandler AB*, et al.* Treatment of non-small cell lung cancer, stage IV: ACCP evidence-based clinical practice guidelines (2nd edition). Chest. 2007; 132:277S-89S.
8. Bottomley A, Efficace F, Thomas R, Vanvoorden V, Ahmedzai SH. Health-related quality of life in non-small-cell lung cancer: methodologic issues in randomized controlled trials. J Clin Oncol. 2003; 21:2982-92.
9. Degner LF, Sloan JA. Symptom distress in newly diagnosed ambulatory cancer patients and as a predictor of survival in lung cancer. J Pain Symptom Manage. 1995; 10:423-31.
10. Maione P, Rossi A, Bareschino MA, Sacco PC, Schettino C, Falanga M*, et al.* Factors driving the choice of the best second-line treatment of advanced NSCLC. Rev Recent Clin Trials. 2011; 6:44-51.
11. The Healthcare Quality Improvement Partnership (HQIP) The Health and Social Care Information Centre (HSCIC) The Royal College of Physicians (RCP). National Lung Cancer Audit Report. Report for the audit period 2011. 2012; Available from: http://www.hqip.org.uk/assets/NCAPOP-Library/NCAPOP-2012-13/Lung-Cancer-National-Audit-Report-pub-2012.pdf (Accessed 16/09/2014).
12. National Institute for Health Research. MK-3475 for advanced or recurrent PD-L1 positive non-small cell lung cancer – second line. 2013; Available from: http://www.hsc.nihr.ac.uk/files/downloads/2210/2508.10efb920.MK3475_nscl_Nov13.pdf (Accessed 10 October 2014).
13. National Institute for Health and Care Excellence (NICE). Erlotinib for the first-line treatment of locally advanced or metastatic EGFR-TK mutation-positive non-small cell lung cancer. NICE technology appraisal guidance 258. 2012; Available from: http://guidance.nice.org.uk/TA258 (Accessed 16/09/2014).
14. National Institute for Health and Care Excellence (NICE). Gefitinib for the first-line treatment of locally advanced or metastatic non-small-cell lung cancer. NICE

technology appraisal guidance 192. 2010; Available from: http://www.nice.org.uk/guidance/ta192 (Accessed 16/09/2014).

15. National Institute for Health and Care Excellence (NICE). Afatinib for treating epidermal growth factor receptor mutation-positive locally advanced or metastatic non-small-cell lung cancer. NICE technology appraisal guidance 310. 2014; Available from: http://www.nice.org.uk/guidance/ta310 (Accessed 16/09/2014).

16. Leighl NB. Treatment paradigms for patients with metastatic non-small-cell lung cancer: first-, second-, and third-line. Curr Oncol. 2012; 19:S52-8.

17. Stinchcombe TE, Socinski MA. Considerations for second-line therapy of non-small cell lung cancer. Oncologist. 2008; 13 Suppl 1:28-36.

18. Boehringer Ingelheim. Cytotoxic regimens by line of treatment for the UK. Data On File. ONC 13-01. 2013.

19. National Institute for Health and Care Excellence (NICE). Erlotinib for the treatment of non-small-cell lung cancer. NICE technology appraisal guidance 162. 2008; Available from: http://www.nice.org.uk/guidance/ta162 (Accessed 16/09/2014).

20. National Institute for Health and Clinical Excellence (NICE). The diagnosis and treatment of lung cancer (update of Clinical Guideline 24). NICE clinical guideline 121. 2011; Available from: http://www.nice.org.uk/guidance/CG121 (Accessed 16/09/2014).

21. National Institute for Health and Care Excellence (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 16/09/2014).

22. Hilberg F, Roth GJ, Krssak M, Kautschitsch S, Sommergruber W, Tontsch-Grunt U*, et al.* BIBF 1120: triple angiokinase inhibitor with sustained receptor blockade and good antitumor efficacy. Cancer Res. 2008; 68:4774-82.

23. Reck M, Kaiser R, Eschbach C, Stefanic M, Love J, Gatzemeier U*, et al.* A phase II double-blind study to investigate efficacy and safety of two doses of the triple angiokinase inhibitor BIBF 1120 in patients with relapsed advanced non-small-cell lung cancer. Ann Oncol. 2011; 22:1374-81.

24. Reck M, Kaiser R, Mellemgaard A, Douillard JY, Orlov S, Krzakowski M*, et al.* Docetaxel plus nintedanib versus docetaxel plus placebo in patients with previously treated non-small-cell lung cancer (LUME-Lung 1): a phase 3, double-blind, randomised controlled trial. Lancet Oncol. 2014; 15:143-55.

25. Ellis LM, Hicklin DJ. VEGF-targeted therapy: mechanisms of anti-tumour activity. Nat Rev Cancer. 2008; 8:579-91.

26. Folkman J. Seminars in Medicine of the Beth Israel Hospital, Boston. Clinical applications of research on angiogenesis. N Engl J Med. 1995; 333:1757-63.

27. Bergers G, Benjamin LE. Tumorigenesis and the angiogenic switch. Nat Rev Cancer. 2003; 3:401-10.

28. European medicines Agency (EMA). Vargatef. 2014; Available from: http://www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/human/medicines/002569/smops/Positive/human_smop_000727.jsp&mid=WC0b01ac058001d127 (Accessed 07/10/2014).

29. Zimmermann FB, Molls M, Jeremic B. Treatment of recurrent disease in lung cancer. Semin Surg Oncol. 2003; 21:122-7.

30. Rosell R, Carcereny E, Gervais R, Vergnenegre A, Massuti B, Felip E*, et al.* Erlotinib versus standard chemotherapy as first-line treatment for European patients with advanced EGFR mutation-positive non-small-cell lung cancer (EURTAC): a multicentre, open-label, randomised phase 3 trial. Lancet Oncol. 2012; 13:239-46.

31. Peters S, Adjei AA, Gridelli C, Reck M, Kerr K, Felip E. Metastatic non-small-cell lung cancer (NSCLC): ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2012; 23 Suppl 7:vii56-64.

32. Cancer Research UK. Biological therapy for lung cancer. Available from: http://www.cancerresearchuk.org/about-cancer/type/lung-cancer/treatment/biological-therapy-for-lung-cancer (Accessed 07/10/2014).

33. European medicines Agency (EMA). Taxotere 20 mg/0.5 ml concentrate and solvent for solution for infusion. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Product_Information/human/000073/WC500035264.pdf (Accessed 30/09/2014).

34. Mellemgaard A, Kaiser R, Douillard JY, Orlov SV, Krzakowski MJ, Von Pawel J*, et al.* Analysis of overall survival in adenocarcinoma NSCLC patients receiving 2nd line combination treatment with nintedanib (BIBF 1120) + docetaxel in the LUME-Lung 1 trial: A randomized, double-blind, placebo-controlled phase 3 study. Eur J Cancer. 2013; 49:S798.

35. Reck M, Mellemgaard A, Douillard J, Orlov S, Krzakowski M, Von Pawel J*, et al.* Nintedanib (BIBF 1120) + docetaxel as 2nd-line therapy in patients with stage IIIB/IV or recurrent NSCLC: Results of the Phase III, randomised, double-blind LUME-Lung 1 trial. (For the LUME-Lung 1 Study Group). J Thorac Oncol. 2014; 9:S39-S40.

36. Reck M, Mellemgaard A, Orlov SV, Krzakowski MJ, Von Pawel J, Gottfried M*, et al.* Antiangiogenic-specific adverse events (AEs) in patients with non-small cell lung cancer (NSCLC) treated with nintedanib (N) and docetaxel (D). J Clin Oncol. 2014; 32:15_suppl:8100.

37. Reck M, Novello S, Mellemgaard A, Orlov S, Kaiser R, Barrueco J*, et al.* Impact of tumor burden on the overall survival analysis of the lume-lung 1 study: A randomized, double-blind phase 3 trial of nintedanib (BIBF 1120) + docetaxel in NSCLC patients progressing after first-line chemotherapy. J Thorac Oncol. 2013; 8:S196.

38. The Cochrane Collaboration. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. [updated 21/03/2011]; Available from: http://www.cochrane.org/training/cochrane-handbook (Accessed 16/09/2014).

39. Hanna NH, Kaiser R, Sullivan RN, Aren OR, Ahn MJ, Tiangco B*, et al.* Lume-lung 2: A multicenter, randomized, double-blind, phase III study of nintedanib plus pemetrexed versus placebo plus pemetrexed in patients with advanced nonsquamous non-small cell lung cancer (NSCLC) after failure of first-line chemotherapy. J Clin Oncol. 2013; 31:15_suppl:8034.

40. Boehringer Ingelheim. Multicentre, randomised, double-blind, Phase III trial to investigate the efficacy and safety of oral BIBF 1120 plus standard docetaxel therapy compared to placebo plus standard docetaxel therapy in patients with stage IIIB/IV or recurrent non small cell lung cancer after failure of first line chemotherapy (LUME Lung 1). Clinical Trial Report (Primary PFS analysis). Data on file. ONC 14-19. 2012.

41. Boehringer Ingelheim. Multicentre, randomised, double-blind, Phase III trial to investigate the efficacy and safety of oral BIBF 1120 plus standard docetaxel therapy compared to placebo plus standard docetaxel therapy in patients with stage IIIB/IV or

recurrent non small cell lung cancer after failure of first line chemotherapy (LUME Lung 1). Clinical Trial Report (Final OS Analysis). Data on file. ONC 14-18. 2013.

42. Boehringer Ingelheim. LUME-Lung 1 CTR (1199.13) (final OS analysis). Appendix. Data on file. ONC 14-24. 2013.

43. Boehringer Ingelheim. Multicentre, randomised, double-blind, Phase III trial to investigate the efficacy and safety of oral BIBF 1120 plus standard docetaxel therapy compared to placebo plus standard docetaxel therapy in patients with stage IIIB/IV or recurrent non small cell lung cancer after failure of first line chemotherapy. Trial Statistical Analysis Plan. ONC 14-20. 2011.

44. Boehringer Ingelheim. Multicentre, randomised, double-blind, Phase III trial to investigate the efficacy and safety of oral BIBF 1120 plus standard docetaxel therapy compared to placebo plus standard docetaxel therapy in patients with stage IIIB/IV or recurrent non small cell lung cancer after failure of first line chemotherapy. Trial Statistical Analysis Plan addendum. Data on file. ONC 14-21. 2011.

45. Boehringer Ingelheim. Summary of Clinical Efficacy. Data on file. ONC 14-22. 2013.

46. Kaiser R, Barrueco JR, Reck M, Hanna N, Gann C, Glomb P*, et al.* Identification of a clinical biomarker for 2nd line combination with nintedanib in adenocarcinoma non-small cell lung cancer (NSCLC) patients in two phase III trials. Eur J Cancer. 2013; 49:S822.

47. Novello S, Mellemgaard A, Kaiser R, Douillard JY, Orlov S, Krzakowski M*, et al.* Analysis of patient-reported outcomes from the lume-lung 1 trial: A randomized, double-blind, placebo-controlled phase 3 study in second-line advanced non-small cell lung cancer (NSCLC) patients. J Thorac Oncol. 2013; 8:S1207.

48. Reck M, Kaiser R, Mellemgaard A, Douillard JY, Orlov S, Krzakowski MJ*, et al.* Nintedanib (BIBF 1120) plus docetaxel in NSCLC patients progressing after first-line chemotherapy: LUME Lung 1, a randomized, double-blind phase III trial. J Clin Oncol. 2013; 31:abstract LBA8011.

49. Phernambucq EC, Hartemink KJ, Smit EF, Paul MA, Postmus PE, Comans EF*, et al.* Tumor cavitation in patients with stage III non-small-cell lung cancer undergoing concurrent chemoradiotherapy: incidence and outcomes. J Thorac Oncol. 2012; 7:1271-5.

50. Morgensztern D, Waqar S, Subramanian J, Trinkaus K, Govindan R. Prognostic impact of malignant pleural effusion at presentation in patients with metastatic non-small-cell lung cancer. J Thorac Oncol. 2012; 7:1485-9.

51. Ryu JS, Ryu HJ, Lee SN, Memon A, Lee SK, Nam HS*, et al.* Prognostic impact of minimal pleural effusion in non-small-cell lung cancer. J Clin Oncol. 2014; 32:960-7.

52. Hicks LK, Cheung MC, Ding K, Hasan B, Seymour L, Le Maitre A*, et al.* Venous thromboembolism and nonsmall cell lung cancer: a pooled analysis of National Cancer Institute of Canada Clinical Trials Group trials. Cancer. 2009; 115:5516-25.

53. Boehringer Ingelheim. Multicentre, randomised, double-blind, Phase III trial to investigate the efficacy and safety of oral BIBF 1120 plus standard docetaxel therapy compared to placebo plus standard docetaxel therapy in patients with stage IIIB/IV or recurrent non small cell lung cancer after failure of first line chemotherapy. Clinical Trial Protocol. Data on file. ONC 14-30. 2008.

54. National Institute for Health and Clinical Excellence (NICE). Guide to the methods of technology appraisal 2013. 2013; Available from: http://publications.nice.org.uk/pmg9 (Accessed 03/10/2014).

55. Kim ST, Uhm JE, Lee J, Sun JM, Sohn I, Kim SW*, et al.* Randomized phase II study of gefitinib versus erlotinib in patients with advanced non-small cell lung cancer who failed previous chemotherapy. Lung Cancer. 2012; 75:82-8.

56. Scagliotti G, Hanna N, Fossella F, Sugarman K, Blatter J, Peterson P*, et al.* The differential efficacy of pemetrexed according to NSCLC histology: a review of two Phase III studies. Oncologist. 2009; 14:253-63.

57. Sun JM, Lee KH, Kim SW, Lee DH, Min YJ, Yun HJ*, et al.* Gefitinib versus pemetrexed as second-line treatment in patients with nonsmall cell lung cancer previously treated with platinum-based chemotherapy (KCSG-LU08-01): an open-label, phase 3 trial. Cancer. 2012; 118:6234-42.

58. Lee DH, Lee JS, Kim SW, Rodrigues-Pereira J, Han B, Song XQ*, et al.* Three-arm randomised controlled phase 2 study comparing pemetrexed and erlotinib to either pemetrexed or erlotinib alone as second-line treatment for never-smokers with non-squamous non-small cell lung cancer. Eur J Cancer. 2013; 49:3111-21.

59. 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.

60. Ciuleanu T, Stelmakh L, Cicenas S, Miliauskas S, Grigorescu AC, Hillenbach C*, et al.* Efficacy and safety of erlotinib versus chemotherapy in second-line treatment of patients with advanced, non-small-cell lung cancer with poor prognosis (TITAN): a randomised multicentre, open-label, phase 3 study. Lancet Oncol. 2012; 13:300-8.

61. 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.

62. Li N, Ou W, Yang H, Liu QW, Zhang SL, Wang BX*, et al.* A randomized phase 2 trial of erlotinib versus pemetrexed as second-line therapy in the treatment of patients with advanced EGFR wild-type and EGFR FISH-positive lung adenocarcinoma. Cancer. 2014; 120:1379-86.

63. Hanna N, Shepherd FA, Fossella FV, Pereira JR, De Marinis F, von Pawel J*, et al.* Randomized phase III trial of pemetrexed versus docetaxel in patients with non-small-cell lung cancer previously treated with chemotherapy. J Clin Oncol. 2004; 22:1589-97.

64. National Confidential Enquiry into Patient Outcome and Death. For better, for worse? A review of the care of patients who died within 30 days of receiving systemic anti-cancer therapy. London: National Confidential Enquiry into Patient Outcome and Death 2008.

65. Hirsh V. Is the Evaluation of Quality of Life in NSCLC Trials Important? Are the Results to be Trusted? Front Oncol. 2014; 4:173.

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

67. Nafees B, Stafford M, Gavriel S, Bhalla S, Watkins J. Health state utilities for non small cell lung cancer. Health Qual Life Out. 2008; 6. doi: 10.1186/1477-7525-6-84

68. Department of Health. Electronic Market Information Tool. Commercial Medicines Unit; 2014; Available from: http://cmu.dh.gov.uk/electronic-market-information-tool-emit (Accessed 7 October 2014).

69. Department of Health. NHS Reference costs guidance 2012-13. London 2013; Available from: https://www.gov.uk/government/publications/reference-costs-guidance-for-2012-13--2 (Accessed 15/10/2014).

70. MIMS. Tarceva (erlotinib). 2013; Available from: http://www.mims.co.uk/Drugs/cancer/antineoplastics/tarceva/ (Accessed 7 October 2014).

71. Boehringer Ingelheim. Health Economic Report for LUME-Lung 1. Data on file. ONC 14-29. 2014.

72. National Institute for Health and Care Excellence (NICE). Pemetrexed for the first-line treatment of non-small-cell lung cancer. NICE technology appraisal guidance 181. London 2009; Available from: http://guidance.nice.org.uk/TA181 (Accessed 16/09/2014).

73. Curtis L. Unit Costs of Health and Social Care (PSSRU) 2012. 2012; Available from: http://www.pssru.ac.uk/project-pages/unit-costs/2012/ (Accessed 14/10/2014).

74. Morgan A, Sutton A, Wailoo A. The risk and costs of febrile neutropenia in patients with non small cell lung cancer treated with docetaxel. Sheffield: NICE Decision Support Unit, ScHARR; 2007; Available from: http://www.nicedsu.org.uk/PDFs%20of%20reports/Erlotinib%20DSU%20final%20report1.pdf (Accessed 14/10/2014).

75. 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). Health Technol Assess. 2014 (in press).

11 APPENDICES

Appendix 1: Detailed critique of the company's search strategy

The ERG's critique of the company's search strategy was undertaken in two parts: (i) An examination of the sources searched and the terms used to make a judgement whether the strategy appeared to be sufficient; (ii) The conduct of its own search strategy to determine if any additional relevant studies were identified. The sources searched by the company and the ERG are summarised in Table 43.

Table 43 Databases searched

Databases searched by company	Databases searched by ERG
Bibliographic databases: • MEDLINE and MEDLINE In-Process (PubMed) • EMBASE (Interface not stated) • Cochrane Library (Wiley Interscience): o Cochrane Database of Systematic Reviews o Cochrane Central Register of Controlled Trials	Bibliogrphic databases: • MEDLINE and MEDLINE In-Process (OvidSP) • EMBASE (OvidSP) • Cochrane Library (Wiley Interscience): o Cochrane Database of Systematic Reviews o Cochrane Central Register of Controlled Trials o Database of Abstracts of Reviews of Effects (DARE) o Health Technology Assessment Database (HTA) o NHS Economic Evaluation Database (NHS EED)
The following sources were searched for grey literature: • ClinicalTrials.gov (www.clinicaltrials.gov) • American Society for Clinical Oncology (ASCO) annual meeting (www.asco.org) • European Society for Medical Oncology (ESMO) annual meeting (www.esmo.org/) • National Guidelines Clearinghouse In addition, reference lists of identified systematic reviews were assessed for additional relevant studies	The following online sources were searched for grey literature: • ClinicalTrials.gov (www.clinicaltrials.gov) • American Society for Clinical Oncology (ASCO) annual meeting (www.asco.org) • European Society for Medical Oncology (ESMO) annual meeting (www.esmo.org/) • National Institute for Health and Care Excellence (www.nice.org.uk) • metaRegister of Controlled Trials (http://www.controlled-trials.com/mrct/) • US Food and Drug Administration (www.fda.gov/) • European Medicines Agency (www.ema.europa.eu/) • National Institute for Health and Clinical Excellence (http://www.nice.org.uk/) • International Society for Pharmacoeconomics and Outcomes Research (www.ispor.org) • Scottish Medicines Consortium (https://www.scottishmedicines.org.uk/) • Summary of Product Characteristics (www.medicines.org.uk/emc/medicine/20929/SPC/tyverb) • Medicines and Healthcare products Regulatory Agency (http://www.mhra.gov.uk/) • The European Union Clinical Trials Register (https://www.clinicaltrialsregister.eu/)

Direct evidence

Five databases were searched by the company on 28 February 2014. These are the minimum specified by NICE and the ERG considers would be sufficient to identify relevant studies. The same search strategy was run across all databases and included free text and MeSH terms of lung cancer, relapsed and second line search terms and randomised controlled trial. The search was limited to humans. The company limited online grey literature searching to the past four years (from January 2011 to February 2014) as they stated conference proceedings older than four years of high quality can be expected to be published in peer viewed journals and therefore picked up in the search results. In addition to the databases searched, the citation lists of relevant systematic reviews published since 2009 were also examined to identify other relevant studies. The ERG considers this search to be adequate although some cancer synonyms have been missed and combining search terms with 'AND' as opposed to 'adjacency' reduces the precision of the search.

The ERG conducted its own searches on 8th August 2014. The ERG search strategy also included free text and MeSH terms, drug search terms and a search term filter to identify RCTs. It did not identify any additional studies.

Indirect evidence

The company completed MTC searches on the same date as the systematic review searches using the same search terms and the same databases. The ERG conducted searches on 21st August 2014 and searched the same databases as its previous search. The search terms included free text and MeSH search terms. An RCT filter was used. The strategy also included a drug comparison concept combined as follows:

- Nintedinab + docetaxel vs docetaxel
- Docetaxel vs gefitinib
- Docetaxel vs erlotinib
- Docetaxel vs pemetrexed
- Pemetrexed vs gefitinib
- Pemetrexed vs erlotinib
- Pemetrexed vs pemetrexed + erlotinib
- Pemetrexed + erlotinib vs erlotinib
- Erlotinib vs gefitinib

No additional studies were identified by the ERG that met the company's eligibility criteria for inclusion into the MTC.

Appendix 2: Eligibility criteria for study inclusion into the company's systematic review and MTC

Table 44 describes the eligibility criteria employed by the company for inclusion into its systematic review. In addition, all non-nintedanib studies were subsequently excluded from the results of the search.

Table 44 Eligibility criteria for inclusion into the company's systematic review

Parameter	Inclusion criteria	Exclusion criteria
Population	**Relapsed or refractory NSCLC** Adults with histologically or cytologically confirmed, locally advanced and/or metastatic NSCLC of stage IIIB or IV (according to American Joint Committee on Cancers) or recurrent NSCLC (all histologies): • Squamous-cell carcinoma • Adenocarcinoma • Large cell carcinoma	Any patient population other than relapsed or refractory NSCLC
Interventions	Any second-line pharmacological treatment for relapsed or refractory NSCLC • Monotherapy • Combination chemotherapy	Patients who were treatment-naïve, had received more than first-line therapy, or had received only non-pharmacological interventions
Outcomes	Relevant outcomes for full-text inclusion: • Overall survival and progression-free survival • Time to relapse • Time to death • Adverse events (all CTCAE grades and CTCAE grade 3 to 4) • Withdrawals • Mean dose and number of cycles of therapy received	No outcomes of interest
Study design	Randomised controlled trials (RCTs) only	Not an RCT (e.g. observational)
Language restrictions	Any language‡	
Date	2000 onwards*	Prior to 2000*
Country	Any	None

NSCLC=non-small cell lung cancer
‡ Non-English-language publications were identified for the efficacy review but none met the inclusion criteria.
*Abstracts published prior to the year 2011 and systematic reviews published prior to the year 2009 were excluded.
Source: Table 6 of the CS[1]

Table 45 describes the eligibility criteria employed by the company for inclusion, with rationale, into its MTC. The search was also limited to include only results with abstracts.

For both the systematic review and MTC, all abstracts obtained from the database search were each examined manually by two researchers applying the predefined eligibility criteria. Following this, a random sample of excluded abstracts was checked for accuracy by a third researcher to confirm the exclusion decisions. Any discrepancy in the decision to include or exclude a study was reviewed by and resolved between researchers. The full-text articles for

abstracts deemed potentially relevant during this first level of screening were retrieved in order to confirm their inclusion in the review. All full-text publications were independently reviewed by two researchers, with all disagreements being resolved by consensus.

Table 45 Eligibility criteria for inclusion in the company's MTC

Parameter	Inclusion criteria	Exclusion criteria	Rationale
Population	**Relapsed or refractory NSCLC (RR NSCLC)** Adults with histologically or cytologically confirmed, locally advanced and/or metastatic NSCLC of stage IIIB or IV (according to American Joint Committee on Cancers) or recurrent NSCLC (all histologies, including patients with mixed histology): • Squamous-cell carcinoma • Adenocarcinoma • Large cell carcinoma Additional inclusion criteria applied during feasibility assessment: • Study must report data for adenocarcinoma subgroup, or 75% or more of participants should have adenocarcinoma	Studies not assessing patients with locally advanced or metastatic, stage IIIB, or IV/recurrent NSCLC Additional exclusion criteria applied during feasibility assessment: • Study does not report data for an adenocarcinoma subgroup • Fewer than 75% of participants overall had adenocarcinoma	The patient population evaluated in our MTC matches the population for which nintedanib is being considered for approval.
Interventions	Any second-line pharmacological treatment for RR NSCLC: • Monotherapy • Combination therapy with other pharmacological agents Additional inclusion criteria applied during feasibility assessment: • Intervention should be licensed for use as second-line treatment for NSCLC	• Trials evaluating non-second-line treatment (e.g., first-, third- or subsequent-line therapy) without subgroup data provided for second-line treatment only • Dose comparison studies without a placebo or control arm • Studies evaluating maintenance treatment	To evaluate nintedanib vs currently available licensed interventions for the second-line treatment of RR NSCLC.
Comparators	Any pharmacotherapy or no treatment: • Other second-line pharmacological treatment • Usual care/no additional intervention • Placebo	None in addition to the above criteria	To compare included interventions with common comparators currently available for the second-line treatment of RR NSCLC, as well as usual care/no intervention and placebo.

Parameter	Inclusion criteria	Exclusion criteria	Rationale
Outcomes	Outcomes relevant to clinical efficacy and safety which were reported in the LUME-Lung 1 study, including: • OS • PFS • ORR • AEs Additional inclusion criteria applied during feasibility assessment: • Study must report relevant data from at least one outcome that has been reported for other studies, thus enabling a comparison across treatments	• Study protocols without outcome data presented • Studies with only patient baseline characteristics reported	We considered outcomes for which an MTC comparing nintedanib + docetaxel with other second-line treatments was feasible, and only included studies with published results for these outcomes.
Study design	Randomised controlled trials (RCTs) only	Non-RCTs Pooled analyses of RCTs	RCTs provide the highest quality clinical trial data.
Language restrictions	Any language		To minimise bias, RCTs published in languages other than English were included in the search, but no relevant non-English language papers were identified
Date	2000 onwards If a study is an abstract only (for example, from a conference), it was only included if it was published in 2011 or onwards	Primary studies published prior to 2000, systematic literature reviews published before 2010 and conference abstracts published prior to 2011 were also excluded	Limiting the review to studies published from 2000 enabled us to focus on the latest trials evaluating the second-line treatment of NSCLC that reflect current clinical practice and patient populations. Conference abstracts were limited to those presented in 2011 onwards, as full text publications of earlier abstracts reporting on studies of a high quality would be expected to have been published. Systematic reviews were limited to those published in the previous 4 years, as these were used only to identify additional relevant primary research papers and therefore needed to be as up-to-date as possible.

AE=adverse event; NSCLC= non-small-cell lung cancer; ORR=overall response rate; OS=overall survival; PFS=progression-free survival; RCT=randomised controlled trial
Source: Table 25 of the CS[1]

Appendix 3: Eligibility criteria for patient inclusion into LUME-Lung 1

Table 46 Inclusion and exclusion criteria for selection of the trial population in LUME-Lung 1

Eligibility criteria for LUME-Lung 1	
Inclusion criteria	• Male or female patient aged 18 years or older • Histologically or cytologically confirmed, locally advanced and/or metastatic NSCLC of stage IIIB or IV or recurrent NSCLC • Relapse or failure of one first-line prior chemotherapy • At least one target tumour lesion that has not been irradiated within the past 3 months and that can accurately be measured • Life expectancy of at least 3 months • ECOG PS of 0 or 1 • Patient has given written informed consent
Exclusion criteria	• More than one prior chemotherapy regimen for advanced and/or metastatic or recurrent NSCLC • More than one chemotherapy treatment regimen (either neoadjuvant or adjuvant or neoadjuvant + adjuvant) prior to first-line chemotherapy • Previous therapy with other VEGFR inhibitors (other than bevacizumab) or docetaxel for treatment of NSCLC • Persistence of clinically relevant therapy related toxicities from previous chemotherapy and/or radiotherapy • Treatment with other investigational drugs or other anti-cancer therapy, or treatment in another clinical trial within the past 4 weeks before start of therapy or concomitantly with this trial • Radiotherapy (except extremities and brain) within the past 3 months prior to baseline imaging • Active brain metastases or leptomeningeal disease • Radiographical evidence of cavitary or necrotic tumours • Centrally located tumours with radiographical evidence (CT or MRI) of local invasion of major blood vessels • History of clinically significant haemoptysis within the past 3 months • Therapeutic anticoagulation (except low dose heparin) or antiplatelet therapy (except for chronic low-dose therapy with acetylsalicylic acid ≤325mg/day) • History of major thrombotic or clinically relevant major bleeding event in the past 6 months • Known inherited predisposition to bleeding or thrombosis • Significant cardiovascular diseases • Inadequate safety laboratory parameters • Significant weight loss (>10 %) within the past 6 weeks • Current peripheral neuropathy greater than CTCAE grade 2 except due to trauma • Pre-existing ascites and/or clinically significant pleural effusion • Major injuries and/or surgery within the past 10 days prior to randomisation with incomplete wound healing • Serious infections requiring systemic antibiotic therapy • Decompensated diabetes mellitus or other contraindication to high-dose corticosteroid therapy • Gastrointestinal disorders or abnormalities that would interfere with absorption of the study drug • Active or chronic hepatitis C and/or B infection • Serious illness or concomitant non-oncological disease or laboratory abnormality that may increase the risk associated with study participation or study drug administration • Patients who are sexually active and unwilling to use a medically acceptable method of contraception during the trial and for at least 12 months after end of active therapy • Pregnancy or breast feeding • Psychological, familial, sociological, or geographical factors potentially hampering compliance with the study protocol and follow-up schedule • Patients unable to comply with the protocol

Eligibility criteria for LUME-Lung 1	
	• Active alcohol or drug abuse • Other malignancy within the past 3 years other than basal cell skin cancer, or carcinoma in situ of the cervix • Any contraindications for therapy with docetaxel • History of severe hypersensitivity reactions to docetaxel or other drugs formulated with polysorbate 80 (Tween 80) • Hypersensitivity to nintedanib and/or the excipients of the trial drugs • Hypersensitivity to contrast media

CT=computerised (or computed) tomography, CTCAE=Common Toxicity Criteria for Adverse Events. ECOG PS=Eastern Cooperative Oncology Group Performance Status; MRI=magnetic resonance imaging, NSCLC=non-small-cell lung cancer, VEGFR=vascular endothelial growth factor receptor
Source: adapted from Table 9 of the CS[1]

Appendix 4: Clinical endpoints and statistical analyses plan in LUME-Lung 1

Outcomes measured are summarised in Table 47. The TSAP[43] is summarised in Table 48.

Table 47 LUME-Lung 1 Outcomes measured

Endpoint/ assessment	Details
Primary outcome	
PFS	• PFS by central review, using modified RECIST (version 1.0) criteria. Tumour assessments performed at baseline (within 4 weeks of randomisation), and every 6 weeks after first docetaxel administration • PFS was defined as time from date of randomisation to date of disease progression, or to date of death, whichever occurred earlier • Disease progression was defined as: o new lesions, including new lesions in a previously irradiated field o an unequivocal increase in a tumour within a previously irradiated field o an increase in sum of longest diameter (SLD) of the target lesions of 20% from nadir (lowest value measured since treatment started) • Patients who experienced a 30% reduction from baseline in SLD of target lesions and a single instance of a 20% increase in SLD from nadir were considered as having progressed • The primary PFS analysis considered all data collected until the cut-off date for the efficacy analysis, which was the date of the 713th PFS event • The stratified log-rank test was used to test for the effect of nintedanib at the 2-sided alpha-level of 0.05. The log-rank test included the four stratification factors used at randomisation.
Secondary outcomes	
OS	• OS was the key secondary endpoint • OS was defined as the time from date of randomisation to date of death (irrespective of cause of death). Patients who stopped active trial treatment were followed until death or lost to follow-up • Stratified log-rank test and a two-look Lan-DeMets group sequential design with an O'Brien-Fleming-type boundary at a two-sided cumulative 5% level of significance.
PFS by local investigator review	PFS by local investigator review
Tumour response evaluation	Tumour response by central independent review and local investigator assessment, according to modified RECIST (version 1.0) criteria was assessed at baseline (within 4 weeks of randomisation) and every 6 weeks after first docetaxel administration, and categorised into one of the following categories: • complete response (CR) - disappearance of all target lesions and non-target lesions • partial response (PR) - at least a 30% decrease in the SLD of target lesions, taking as reference the baseline SLD • stable disease (SD) - neither sufficient shrinkage of target lesions to qualify for PR nor sufficient increase to qualify as PD; persistence of one or more non-target lesions • progressive disease (PD): o new lesions, including new lesions in a previously irradiated field o an unequivocal increase in a tumour within a previously irradiated field o an increase in SLD of the target lesions of 20% from nadir (lowest value measured since treatment started) • unknown (UNK)

Endpoint/ assessment	Details
	Evaluation of tumour response was based on radiological tumour assessments (CT or MRI) • Tumour images were centrally reviewed by a panel of central independent radiologists. Following radiological review, all patient information was presented to an oncologist. The radiologists and the oncologist were blinded to treatment • Best overall response: o represents the best response a patient has had during their time in the study up until progression, last evaluable assessment in the absence of progression or the start of subsequent anti-cancer therapy. o for patients whose progression event is death, best objective response will be calculated based on data up until the last evaluable RECIST assessment prior to death. • Confirmed objective response o A patient was considered to have a confirmed objective response if a CR or PR was confirmed by imaging no earlier than 28 days after the first occurrence of the response • Disease control o Disease control was defined as a best overall response of CR, PR, or SD recorded at least 6 weeks after the date of randomisation • Time to confirmed objective response o Time from randomisation to first documented confirmed response (CR or PR) recorded at least 6 weeks after the date of randomisation • Duration of confirmed objective response o Time from first documented confirmed response (CR or PR) to progression, or death in the absence of progression • Duration of disease control o Time from randomisation to progression, or death in the absence of progression (whichever occurs earlier) amongst patients with disease control • Change in tumour size o The best change in size (i.e. SLD) of target lesions from baseline was analysed. The maximum SLD decrease from baseline (or the minimum increase in SLD for patients with no reduction in target lesion size) was considered as the best change of the target lesion size in a patient
Clinical improvement	• Clinical improvement quantified the maintenance of body weight and ECOG PS, by measuring the time from randomisation to deterioration in body weight of more than 10% from baseline, and/or increase in ECOG performance score of at least 1 category from baseline, whichever occurred earlier. Patients who died without prior deterioration were considered as having deteriorated at the time of death. • Clinical improvement was analysed until end-of treatment only
HRQoL	• HRQL was measured at the screening visit, at 21-day intervals during treatment, at the end of active treatment, and at the first follow-up visit by the following standardised self-assessment questionnaires: o EQ-5D health status self-assessment questionnaire o EORTC Quality of Life Questionnaire (EORTC QLQ-C30) o EORTC lung cancer specific supplementary module (EORTC QLQ-LC13) • The EQ-5D includes the following two questionnaires, which were analysed descriptively: o Five dimensions of health (mobility, self-care, usual activities, pain/discomfort and anxiety/depression), which are analysed descriptively. Each dimension comprised three levels (no problems, some problems, severe problems)

Endpoint/ assessment	Details
	○ A visual analogue scale (VAS) recorded the respondents self-rated health status on a vertical graduated (0 to 100) scale • The EORTC QLQ-C30 questionnaire includes a global health status/HRQL scale, 5 functional scales, 3 symptom scales, and 6 single items to assess dyspnoea, insomnia, appetite loss, constipation, diarrhoea and financial difficulties. The QLQ-LC13 supplementary module was designed to be used by patients receiving chemotherapy or radiotherapy. It incorporates a multi-item scale to assess dyspnoea, and a series of single items to assess pain, coughing sore mouth, dysphagia, peripheral neuropathy, alopecia and haemoptysis. • The main HRQL endpoints were the time to deterioration for cough (QLQ-LC13, question 1), dyspnoea (QLQ-LC13, questions 3 to 5) and pain (QLQ-C30, Questions 9 and 19) and were evaluated as follows: ○ Distribution of patients with improved, stable, or worsened scores. Improvement was defined as scores that improve by ≥10 points (0 to 100 point scale) at any time during study. Worsening was defined as a worsening in EORTC scores of ≥ 10 points at any time in patients with no improvement. Otherwise, a patient was considered stable. ○ Time to deterioration: defined as time from randomisation to the first 10-point increase (i.e. worsening) from baseline score
Pharmacokinetics	• Pharmacokinetics of nintedanib and of its clinical relevant metabolites BIBF1202 and BIBF1202 glucuronide were determined from blood samples taken at Visit 2 of Treatment Course 2 and 3; both prior to and after the administration of nintedanib.
Safety	• Incidence and intensity of AEs according to the CTCAE version 3.0 • Changes in safety laboratory parameters • The safety analysis included data collected until the safety cut-off date

CR=complete response; CT=computed tomography; CTCAE=Common Terminology Criteria for Adverse Events; ECOG=Eastern Cooperative Oncology Group; EORTC QLQ LC=European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire (Lung Cancer Module); EMA=European Medicines Agency; EORTC= European Organisation for Research and Treatment of Cancer; EQ-5D=European Quality of Life-5 Dimensions; HRQL=health related quality of life; MRI=Magnetic resonance imaging; OS=overall survival; PD=progressive disease; PFS=progression-free survival; PR=partial response; QLQ=quality of life questionnaire; PRO=patient reported outcome; RECIST=Response Evaluation Criteria in Solid Tumours; SD=stable disease; SLD=sum of longest diameters; VAS=visual analogue scale
Source: Table 13 of the CS[1]

Table 48 Trial statistical analysis plan for LUME-Lung 1

Stage of analysis	Description
Futility analysis	A pre-planned futility analysis was to be performed by the central independent DMC after approximately 50% of the PFS events needed for the primary PFS analysis had occurred (~356 events), for the purpose of advising the sponsor as to whether or not the study should continue as planned. The sponsor was blinded to the results of this analysis. Although PFS by central independent review was the primary endpoint, PFS as assessed by the local investigator was used for the futility analysis because of the logistical complexity and the time it took to complete the central independent review of patients' imaging data.
Primary PFS analysis	The primary PFS analysis was to be performed when 713 patients had experienced a centrally independently assessed PFS event (cut-off date 2 November 2010). At this time, a protocol-defined interim analysis of OS was also to be performed. The primary analysis was based on the ITT population.
Final OS analysis	The final analysis of OS was performed when 1,151 patients had died (cut-off date 15 February 2013). At the time of the final OS analysis an updated analysis of all available PFS events was also performed.

DMC=data monitoring committee; ITT=intention-to-treat; OS=overall survival; PFS=progression-free survival
Source: TSAP[43]

Appendix 5: Subgroup and sensitivity analyses in LUME-Lung 1

Subgroup analyses

A number of subgroup analyses for the primary endpoint of PFS assessed by central review and for the secondary outcome OS were pre-specified in the protocol:

- tumour histology (squamous vs non-squamous)
- baseline ECOG PS (0 vs 1)
- presence of brain metastases at baseline (yes vs no)
- prior treatment with bevacizumab (yes vs no)
- sex (male, female)
- age (<65years, ≥65 years)
- race (Asian vs non-Asian patients; information was derived from the race categories as documented on the CRF)
- smoking status (never smoked vs currently smokes/ex-smoker)

The following subgroup analyses were added post-hoc:

- geographical region (Asia, Europe, South Africa; based on country of enrolment)
- best response to first-line therapy (CR/PR/SD, PD, unknown/missing/NA)
- sum of longest diameters at baseline (<7.5cm vs ≥7.5cm)
- time since first-line therapy (<9 months vs ≥9 months)

The company lists a number of baseline characteristics (CS[1], p.77), which were also investigated for subgroup effects. However, neither the protocol[53] or CTR[40] specified whether these were pre-specified or post-hoc analyses. Therefore, the ERG asked for clarification on this issue and the company responded stating that three variables were pre-specified in the protocol[53] and three were included in an amendment to the TSAP[44] for the final OS analysis of LUME-Lung 1.[24]

- presence of liver metastases (yes vs no) (a priori).
- disease stage at diagnosis (<IIIB/IV, IIIB, IV) (a priori).
- concomitant therapy with biphosphonates at baseline (yes vs no) (a priori).
- presence of adrenal metastases (yes vs no) (included in amendment)
- number of metastatic organs at baseline (≤2 metastatic organs, >2 metastatic organs, not centrally reviewed) (included in amendment)
- lactate dehydrogenase (LDH) level at baseline (LDH ≤1, LDH >1) (included in amendment)

Sensitivity analyses

The following sensitivity analyses were pre-specified in the protocol for PFS:

- Analysis using a Cox proportional hazards model fitting the four stratification factors as covariates
- Analysis using a stepwise variable selection method to identify covariates that might be relevant to efficacy
- Analysis replacing actual tumour imaging dates with the originally scheduled dates of radiological assessments
- Analysis using an interval-censoring approach

However, the ERG found that the list of covariates included in the model for the second sensitivity analysis were listed in the CS[1] (pages 74 to 75), but were not all pre-specified in the protocol. The ERG asked for clarification on whether these factors were pre-specified, and the company responded with the following information, stating that only four out of twelve were pre-specified:

- Brain metastases at baseline: predefined strata and also for this analysis in interim TSAP LUME-Lung 1 before unblinding of primary PFS data (a priori).
- Prior treatment with bevacizumab: predefined strata and also for this analysis in interim TSAP LUME-Lung 1 (a priori).
- Body- surface area: (post-hoc).
- Age: Predefined in interim TSAP LUME-Lung 1 in subgroup section (a priori).
- Duration of first-line chemotherapy: (post-hoc).
- Time to first progression: specified in TSAP amendment (post-hoc).
- Time since first histological diagnosis: (post-hoc).
- Presence of ipsilateral metastases in the lung at baseline: (post-hoc).
- Presence of contralateral metastases in the lung at baseline: (post-hoc).
- Bone metastases at baseline: (post-hoc).
- Adrenal metastases at baseline: specified in TSAP amendment (post-hoc).
- Sum of target lesions at baseline: predefined in interim TSAP LUME-Lung 1 (a priori).

The following sensitivity analyses were pre-specified in the protocol for OS:

- Analysis using a Cox proportional hazards model with three of the stratification factors used at randomisation as covariates(ECOG PS at baseline, prior bevacizumab treatment, presence of brain metastases at baseline)
- Analysis using a model which included the stratification factors and the baseline sum of the longest diameters (SLD) of the target lesions (mm) as covariates.

Appendix 6: Methods utilised by the company for making indirect comparisons and mixed treatment comparisons

Mixed treatment comparisons

MTCs were performed using the Markov chain Monte Carlo software package OpenBUGs. The company ran all analyses using fixed-effects models, which assume there is no heterogeneity in relative effects. Random-effects models were also performed if sufficient data was available to estimate a random-effects coefficient, i.e. there were comparisons in the network with evidence from more than one trial. The company chose not to fit random-effects models in situations where the data was sparse, as the estimate of random-effects variation would be too reliant on the choice of prior. The company chose to use vague (non-informative) priors for study and treatment effects, in order to enable a moderate amount of random-effects variation.

Three chains were used to run the analyses, and in all cases, the first 50,000 burn-in simulations were discarded to allow for convergence. Estimates were then obtained from a further 50,000 iterations. The company performed several validation checks to ensure that the models had converged sufficiently and that the estimates produced were reliable. These included examining the Brooks-Gelman-Rubin (BGR) plots and inspection of the values of the Monte Carlo error (Monte Carlo standard error of the mean) to assess validity.

Bucher indirect comparisons

A Bucher indirect comparison is a simple method of comparing two treatments for which there is no direct evidence. In order to obtain an estimate of the treatment effect of A vs C, it is possible to look at two trials which have a common comparator, i.e. Trial 1 considering A vs B, and Trial 2 considering B vs C. The Bucher method does not incorporate random-effects variance from trials elsewhere in the evidence network, i.e. trials which consider C vs D.

Wherever possible, the company conducted Bucher indirect comparisons.

Appendix 7: Assessment of proportional hazards assumption in LUME-Lung 1 Trial

Both indirect comparisons and MTCs require the trials included in the analysis to conform to the assumption of proportional hazards for meaningful and robust results to be generated. This means that the hazard (i.e. the risk of an event occurring at a particular time) is in a constant ratio between the patterns of events observed in the two treatment arms, independent of the time since randomisation. This is a strong assumption which is frequently violated, and it is important that its validity is confirmed prior to carrying out any meta-analysis of outcomes from multiple clinical trials.

In this appraisal a single trial (LUME-Lung 1[24]) compares nintedanib plus docetaxel treatment with erlotinib through a network of trials in which the only links are trials which feature docetaxel monotherapy as a treatment arm. If the proportional hazards assumption is not supported by the LUME-Lung 1[24] trial data, any estimation of the relative effectiveness of nintedanib plus docetaxel vs erlotinib (i.e. a calculated HR) will lack credibility and be effectively meaningless. In this appendix the validity of the proportional hazards assumption in LUME-Lung 1[24] is considered for two key outcomes (PFS and OS) critical to the modelling of cost-effectiveness.

PFS

Figure 19 shows clearly that the PFS survival curve LUME-Lung 1[24] trial arms diverge after about six weeks and then converge and cross after about one year, indicating that the patient PFS advantage from nintedanib plus docetaxel treatment is limited to the first year after treatment. To test the proportional hazards assumption in this data set the HR has been calculated at each event time in either arm of the trial and are shown in Figure 19. If the proportional hazards assumption is supportable the HR values should vary randomly about a horizontal line corresponding to the conventional estimated HR for the trial. Clearly this is not the case as a strong upward trend is apparent following the initial fluctuations (which are due to the small numbers of events recorded in the first few weeks of the trial).

On this basis it must be concluded that any HR estimated from a meta-analysis aimed at comparing PFS outcomes between nintedanib plus docetaxel and any treatment other than docetaxel does not satisfy the essential requirement for validity and reliability, and cannot be considered appropriate for populating a cost-effectiveness model.

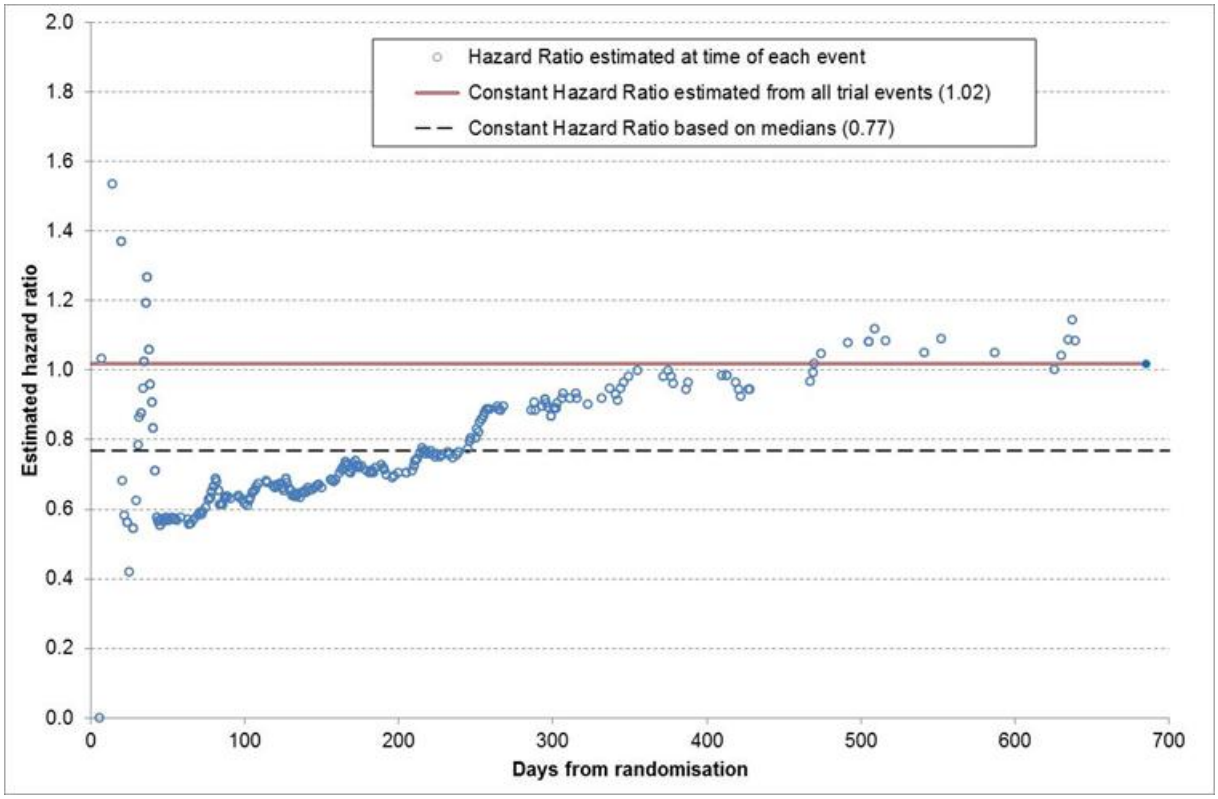


Figure 19 Variation in estimated PFS HR with time in the LUME-Lung 1 clinical trial

OS

Similarly the trend of OS HR estimates also show systematic variations over time (Figure 20): from a peak of 1.1 at four months, falling to less than 0.75 at 400 to 500 days, and increasing thereafter. This pattern is not consistent with the presumption of a steady common HR independent of time, and therefore indicates that the proportional hazards assumption cannot be applied to the LUME-Lung 1[24] OS data set with confidence.

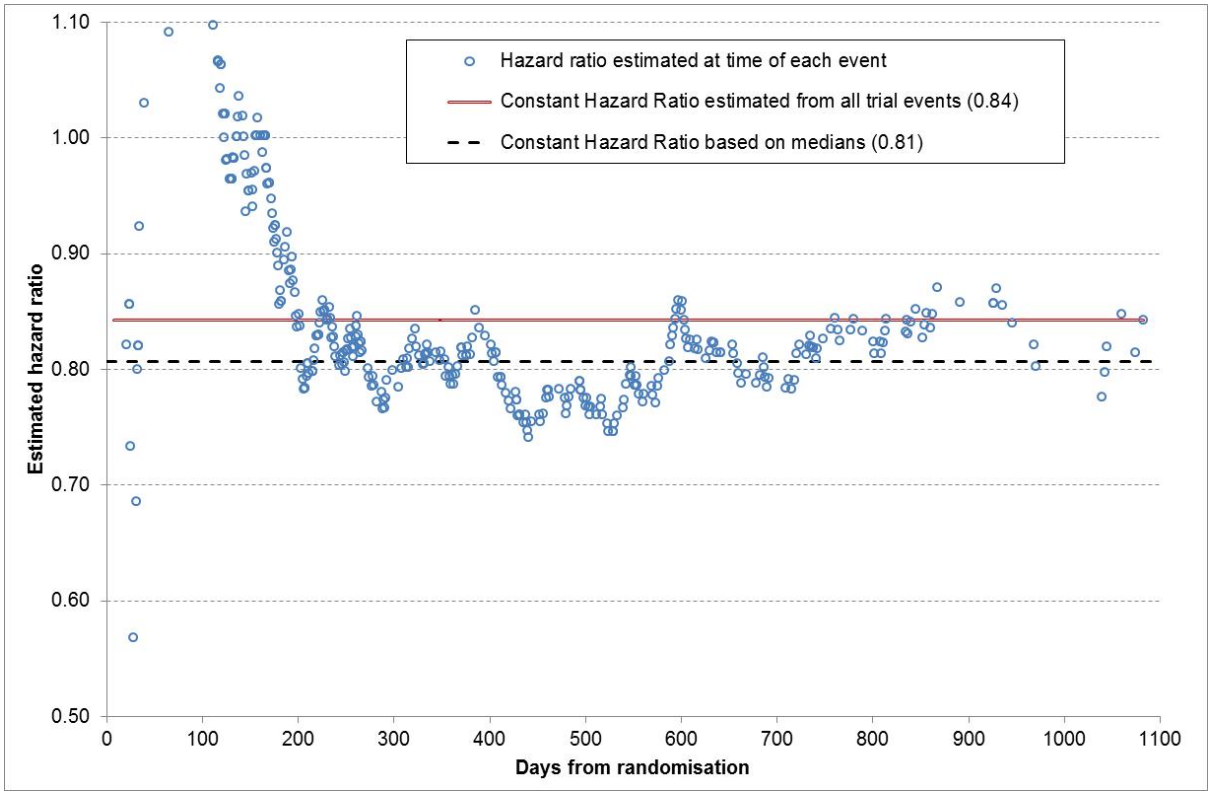


Figure 20 Variation in estimated OS HR with time in the LUME-Lung 1 clinical trial

CONCLUSION

Without a single robust time-invariant HR for either PFS or OS it is not possible to use conventional methods to link and compare the outcomes of patients treated with nintedanib plus docetaxel to patients treated with erlotinib in either the TAILOR[59] or the WSY001[62] trials, regardless of the characteristics of the other trials in the network. Without such comparison meaningful cost-effectiveness analysis involving erlotinib is not possible.

Appendix 8: ERG Revisions to company's model: Nintedanib STA

All revisions are activated by a binary logic switch with 0 = unchanged, 1 (or any non-zero number) = apply ERG modification.

Logic switches are indicated by range variables Mod_*n* where n = 1 - 12. The Mod numbers do not directly match the Table Row numbers, and one Table Row involves applying 2 similar Mod revisions simultaneously.

A menu of revisions/Mod numbers appears on the 'Results' worksheet together with summary results as used to transfer to the ERG report.

ERG Table 14 Row	Binary switch	Associated detail	Implementation instructions
		ERG_Survival_Tables.xlsx	Copy this worksheet as an additional sheet in the model. Ensure that the named ranges ERG_OS, ERG_PFS, ERG_TOT are correctly named in the model.
1. ERG OS estimates	Mod_7	LUME1_OS40-1-3.xlsx	In Sheet 'Survival', Replace formula in cell AW119 by =IF(Mod_7=0,OFFSET(AI119,0,2*ch_OS-2),VLOOKUP(B119,ERG_OS,2)) Copy formula in cell AW119 to range AW120:AW405 Replace formula in cell AX119 by =IF(Mod_7=0,OFFSET(AJ119,0,2*ch_OS-2),VLOOKUP(B119,ERG_OS,3)) Copy formula in cell AX119 to range AX120:AX405

ERG Table 14 Row	Binary switch	Associated detail	Implementation instructions
2. ERG PFS estimates	Mod_6	LUME1_PFS-1-1.xlsx	In Sheet 'Survival', Replace formula in cell M119 by =IF(Mod_6=0,IF(ch_PFS=1,G119,IF(ch_PFS=2,I119,K119)),VLOOKUP(B119,ERG_PFS,2)) Copy formula in cell M119 to range M120:M405 Replace formula in cell N119 by =IF(Mod_6=0,IF(ch_PFS=1,H119,IF(ch_PFS=2,J119,L119)),VLOOKUP(B119,ERG_PFS,3)) Copy formula in cell N119 to range N120:N405

ERG Table 14 Row	Binary switch	Associated detail	Implementation instructions
3. ERG TOT estimates	Mod_8	LUME1_TOT_DocArm_40-1-6.xlsx LUME1_TOT_NinArm_DocTx_40-1-7.xlsx LUME1_TOT_NinArm_NinTx_40-1-8.xlsx	In Sheet 'comp1Model', Replace formula in cell L15 by =IF(Mod_8=0,100%,VLOOKUP(E15,ERG_TOT,4)) Replace formula in cell L16 by =IF(Mod_8=0,L15*(1-rDiscontinuation_nine_Comp1),VLOOKUP(E16,ERG_TOT,4)) Copy formula in L16 to range L17:L301 Replace formula in cell M15 by =IF(Mod_8=0,100%,VLOOKUP(E15,ERG_TOT,3)) Replace formula in cell M16 by =IF(Mod_8=0,IF(OR(Efficacy!$F$43="no",E15<4),M15*(1-rDiscontinuation_doce_Comp1),0),VLOOKUP(E16,ERG_TOT,3))*IF(AND(Mod_4=1,E16>3),0,1) Copy formula in M16 to range M17:M301 In Sheet 'comp2Model', replace formula in cell M15 by =IF(Mod_8=0,100%,VLOOKUP(E15,ERG_TOT,2)) Replace formula in cell M16 by =IF(Mod_8=0,IF(OR(Efficacy!$F$43="no",E15<4),M15*(1-rDiscontinuation_Comp2),0), VLOOKUP(E16,ERG_TOT,2))*IF(AND(Mod_4=1,E16>3),0,1) Copy formula in M16 to range M17:M301

ERG Table 14 Row	Binary switch	Associated detail	Implementation instructions
4. Mid-cycle adjustment	Mod_10	None	In Sheet 'comp1Model', replace formula in cell BE16 by =IF(Mod_10=0,AVERAGE(S15:S16),S15)*cDrugAdmin_doxa_Comp1 Copy formula in BE16 to ranges BE17:BE19, BE21:BE301 Replace formula in cell BE20 by =IF(Efficacy!$F$43="no",IF(Mod_10=0,AVERAGE(S19:S20),S19)*cDrugAdmin_doxa_Comp1,0) In Sheet 'comp2Model', Replace formula in cell AV16 by =IF(Mod_10=0,AVERAGE(S15:S16),S15)*cDrugAdmin_Comp2 Copy formula in AV16 to ranges AV17:AV19, AV21:AV301 Replace formula in cell AV20 by =IF(Efficacy!$F$43="no",IF(Mod_10=0,AVERAGE(S19:S20),S19)*cDrugAdmin_Comp2,0)
5. Cost of treatment doses	Mod_1	DrugCalcs.xlsx Sheet 'Calcs_75mg' LUME1_MeanDoseCostEstimates(adjusted fordose reductions).xlsx Sheet: 'LUME1_DoseLevels_40_3_1'	In Sheet 'UnitCosts', Replace formula in cell H35 by =IF(Mod_1=0,Y37,98.480134%) Replace formula in cell H36 by =IF(Mod_1=0,Y38,99.08405%) Replace formula in cell H37 by =IF(Mod_1=0,Y39,99.08405%) Replace formula in cell I35 by =IF(Mod_1=0,DrugCostCalc!$S$71,37.5) Replace formula in cell I36 by =IF(Mod_1=0,G36*F36*BSA,37.5) Replace formula in cell I37 by =IF(Mod_1=0,DrugCostCalc!$S$88,37.5)
	Mod_11	LUME1_MeanDoseCostEstimates(adjusted fordose reductions).xlsx Sheet: 'LUME1_DoseLevels_40_2_1'	In Sheet 'UnitCosts', Replace formula in cell K33 by =IF(Mod_11=0,I33*J33,1409.920164)
6. Febrile neutropenia cost	Mod_2	FNcost.xlsx	In Sheet 'AdverseEvents', Replace formula in cell I195 by =IF(Mod_2=0,SUM(H196:H209),7352.543797)

ERG Table 14 Row	Binary switch	Associated detail	Implementation instructions
7. Monitoring cost	Mod_9	None	In Sheet 'ResourceUse', Replace formula in cell E91 by =100%*IF(Mod_9=0,1,0) Replace formula in cell F91 by =timeDaysInCycle/(2.5*timeDaysInMonth)*IF(Mod_9=0,1,0) Replace formula in cell M67 by =IF(Mod_9=0,0,100%) Replace formula in cell N67 by =timeDaysInCycle/(2.5*timeDaysInMonth)*IF(Mod_9=0,0,1) Replace formula in cell P67 by =IF(Mod_9=0,UnitCosts!$E$73,0) Replace formula in cell Q67 by =IF(Mod_9=0,M67*N67*P67,0)
8. Discounting method	Mod_3	None	In Sheet 'comp1Model', Replace formula in cell H16 by =IF(Mod_3=0,H15/(1 + discCc),1/(1 + iDiscCost)^INT(F16)) Replace formula in cell I16 by =IF(Mod_3=0,I15/(1 + discHc),1/(1 + iDiscHealth)^INT(F16)) Copy range H16:I16 to rows 17-301 In Sheet 'comp2Model', Replace formula in cell H16 by =IF(Mod_3=0,H15/(1 + discCc),1/(1 + iDiscCost)^INT(F16)) Replace formula in cell I16 by =IF(Mod_3=0,I15/(1 + discHc),1/(1 + iDiscHealth)^INT(F16)) Copy range H16:I16 to rows 17-301
9. Disutility of fatigue	Mod_5	None	In Sheet 'Utilities' Replace formula in cell E66 by =(SUMPRODUCT(Utilities!$E$50:$E$62,AdverseEvents!$E$34:$E$46) + IF(Mod_5=0,0,(-0.326-E55)*AdverseEvents!E39))/AdverseEvents!$E$48 Replace formula in cell E66 by =(SUMPRODUCT(Utilities!$E$50:$E$62,AdverseEvents!$F$34:$F$46) + IF(Mod_5=0,0,(-0.101-E55)*AdverseEvents!F39))/AdverseEvents!$F$48

ERG Table 14 Row	Binary switch	Associated detail	Implementation instructions
10. Stable disease costs	Mod_12	None	In Sheet 'Resource Use' Replace formula in cell M65 by =IF(Mod_12=0,100%,0%) Replace formula in cell N65 by =IF(Mod_12=0,timeWeeksInCycle,0) Replace formula in cell M78 by =IF(Mod_12=0,100%,0%) Replace formula in cell M78 by =IF(Mod_12=0,1,0)
11. Docetaxel ≤4 cycles	Mod_4	None	See details for #3

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Patient/carer expert statement (STA)

Nintedanib for previously treated locally advanced or metastatic non-small cell lung cancer

Thank you for agreeing to give us your views on this treatment that is being appraised by NICE and how it could be used in the NHS. Patients, carers and patient organisations can provide a unique perspective on conditions and their treatment that is not typically available from other sources. We are interested in hearing about:

- the experience of having the condition or caring for someone with the condition
- the experience of receiving NHS care for the condition
- the experience of having specific treatments for the condition
- the outcomes of treatment that are important to patients or carers (which might differ from those measured in clinical studies, including health-related quality of life)
- preferences for different treatments and how they are given
- expectations about the risks and benefits of the treatment.

We have already asked your nominating organisation to provide an organisation's view. We are asking you to give your views as an individual whether you are:

- a patient
- a carer (who may be voicing views for a patient who is unable to) or
- somebody who works or volunteers for a patient organisation.

To help you give your views, we have provided a questionnaire. You do not have to answer every question — the questions are there as prompts to guide you. The response area will expand as you type. The length of your response should not normally exceed 10 pages.

1. About you

Your name: JESME FOX
Name of your nominating organisation: ROY CASTLE LUNG CANCER FOUNDATION
Do you know if your nominating organisation has submitted a statement?

Yes

Do you wish to agree with your nominating organisation's statement?

yes

(We would encourage you to complete this form even if you agree with your nominating organisation's statement.)

Are you:

- a patient with the condition?

No

- a carer of a patient with the condition?

No

- a patient organisation employee or volunteer?
-

Yes

Do you have experience of the treatment being appraised?

Yes

If you wrote the organisation submission and do not have anything to add, tick here YES☐ (If you tick this box, the rest of this form will be deleted after submission.)

2. Living with the condition

What is your experience of living with the condition as a patient or carer?

3. Current practice in treating the condition

Which treatment outcomes are important to you? (That is, what would you like treatment to achieve?) Which of these are most important? If possible, please explain why.

What is your experience of currently available NHS care and of specific treatments? How acceptable are these treatments – which did you prefer and why?

4. What do you consider to be the advantages of the treatment being appraised?

Benefits of a treatment might include its effect on:

- the course and/or outcome of the condition
- physical symptoms
- pain
- level of disability
- mental health
- quality of life (such as lifestyle and work)
- other people (for example, family, friends and employers)
- ease of use (for example, tablets rather than injection)
- where the treatment has to be used (for example, at home rather than in hospital)
- any other issues not listed above

Please list the benefits that you expect to gain from using the treatment being appraised.

Please explain any advantages that you think this treatment has over other NHS treatments in England.

If you know of any differences in opinion between you and other patients or carers about the benefits of the treatment being appraised, please tell us about them.

5. What do you consider to be the disadvantages of the treatment being appraised?

Disadvantages of a treatment might include:

- aspects of the condition that the treatment cannot help with or might make worse
- difficulties in taking or using the treatment (for example, injection rather than tablets)
- side effects (for example, type or number of problems, how often, for how long, how severe. Please describe which side effects patients might be willing to accept or tolerate and which would be difficult to accept or tolerate)
- where the treatment has to be used (for example, in hospital rather than at home)
- impact on others (for example, family, friends and employers)
- financial impact on the patient and/or their family (for example, the cost of travel to hospital or paying a carer)
- any other issues not listed above

Please list any concerns you have about current NHS treatments in England.

Please list any concerns you have about the treatment being appraised.

If you know of any differences in opinion between you and other patients or carers about the disadvantages of the treatment being appraised, please tell us about them.

6. Patient population

Do you think some patients might benefit more from the treatment than others? If so, please describe them and explain why.

Do you think some patients might benefit less from the treatment than

others? If so, please describe them and explain why.

7. Research evidence on patient or carer views of the treatment

Are you familiar with the published research literature for the treatment?

☐ Yes ☐ No

If you answered 'no', please skip the rest of section 7 and move on to section 8.

Please comment on whether your experience of using the treatment as part of routine NHS care reflects the experience of patients in the clinical trials.

Do you think the clinical trials have captured outcomes that are important to patients? Are you aware of any limitations in how the treatment has been assessed in clinical trials?

If the treatment being appraised is already available in the NHS, are there any side effects that were not apparent in the clinical trials but have emerged during routine NHS care?

Are you aware of any relevant research on patient or carer views of the condition or existing treatments?

☐ Yes ☐ No

If yes, please provide references to the relevant studies.

8. Equality

NICE is committed to promoting equality of opportunity and eliminating discrimination. Please let us know if you think that recommendations from this appraisal could have an adverse impact on any particular groups of people, who they are and why.

9. Other issues

Do you consider the treatment to be innovative?

☐ Yes ☐ No

If yes, please explain what makes it significantly different from other treatments for the condition.

Is there anything else that you would like the Appraisal Committee to consider?

10. Key messages

In no more than 5 bullet points, please summarise the key messages of your submission.

-
-
-
-
-

STA ADDENDUM

LIVERPOOL REVIEWS AND IMPLEMENTATION GROUP (LRiG)

Nintedanib for previously treated locally advanced or metastatic non-small cell lung cancer

ADDENDUM

This report was commissioned by the NIHR HTA Programme as project number 13/106/01

Completed 13th November 2014

UNIVERSITY OF LIVERPOOL

LIVERPOOL REVIEWS AND IMPLEMENTATION GROUP

A MEMBER OF THE RUSSELL GROUP

1 INTRODUCTION

This document provides additional information requested of the ERG by NICE in advance of the first Appraisal Committee Meeting, to be held 19 November 2014.

The following information was requested:

1. Conduct two further exploratory analyses, based around the 11 amendments the ERG reported:

    1.1. Calculate the ICER for nintedanib plus docetaxel compared with docetaxel alone excluding the ERG's OS modification but including the other 10 amendments

    1.2. Calculate this ICER excluding the ERG's OS or PFS modifications but including the other 9 amendments

2. Provide information concerning the comparative accuracy of the company OS modelling and the ERG's alternative OS method, including graphs showing:

    2.1. Model residuals (difference between modelled estimates and clinical trial data).

    2.2. Restricted mean OS estimated by the Area Under Curve (AUC) method, comparing the company base case model and the ERG's alternative method.

2 ERG RESPONSE

2.1 Additional Scenarios

The requested new scenarios have been added as Scenarios D & E in the modified version of Table 40 from the ERG report shown below. The eleven ERG revisions are now labelled R1-R11 for clarity. As noted in the ERG report, the most influential change is the application of the ERG OS estimates. If this revision (R1) is not accepted, the revised ICER using the other ten revisions becomes £62,719 per QALY gained, or reduces to £61,311 per QALY gained if the ERG's PFS modification (R2) is also excluded.

Table 40: Cost-effectiveness results for nintedanib plus docetaxel vs docetaxel with ERG revisions to company's base-case comparison in the adenocarcinoma population (with additional scenarios)

Model scenario & ERG revisions	**Nintedanib + docetaxel**			**Docetaxel**			**Incremental**			**ICER**	**ICER**
	Cost	**QALYs**	**Life**	**Cost**	**QALYs**	**Life**	**Cost**	**QALYs**	**Life**	**£/QALY**	**Change**
A. Company's base-case	[redacted]	[redacted]	**1.810**	[redacted]	[redacted]	**1.419**	**+ £11,051**	**+ 0.218**	**+ 0.391**	**£50,776**	-
R1) ERG OS estimates	[redacted]	[redacted]	1.493	[redacted]	[redacted]	1.238	+ £10,497	+ 0.153	+ 0.255	£68,587	**+ £17,811**
R2) ERG PFS estimates	[redacted]	[redacted]	1.810	[redacted]	[redacted]	1.419	+ £11,527	+ 0.220	+ 0.391	£52,445	**+ £1,669**
R3) ERG ToT estimates	[redacted]	[redacted]	1.810	[redacted]	[redacted]	1.419	+ £11,298	+ 0.218	+ 0.391	£51,930	**+ £1,154**
R4) Mid-cycle adjustment	[redacted]	[redacted]	1.810	[redacted]	[redacted]	1.419	+ £11,717	+ 0.218	+ 0.391	£53,839	**+ £3,062**
R5) Cost of treatment doses	[redacted]	[redacted]	1.810	[redacted]	[redacted]	1.419	+ £11,445	+ 0.218	+ 0.391	£52,587	**+ 1,811**
R6) Febrile neutropenia cost	[redacted]	[redacted]	1.810	[redacted]	[redacted]	1.419	+ £11,180	+ 0.218	+ 0.391	£51,372	**+ £595**
R7) Monitoring cost	[redacted]	[redacted]	1.810	[redacted]	[redacted]	1.419	+ £11,130	+ 0.218	+ 0.391	£51,140	**+ £364**
R8) Discounting method	[redacted]	[redacted]	1.810	[redacted]	[redacted]	1.419	+ £11,189	+ 0.221	+ 0.391	£50,532	**-£244**
R9) Disutility of fatigue	[redacted]	[redacted]	1.810	[redacted]	[redacted]	1.419	+ £11,051	+ 0.217	+ 0.391	£50,830	**+ £54**
R10) Stable disease costs	[redacted]	[redacted]	1.810	[redacted]	[redacted]	1.419	+ £11,637	+ 0.218	+ 0.391	£53,470	**+ £2,693**
R11) Docetaxel ≤4 cycles	[redacted]	[redacted]	1.810	[redacted]	[redacted]	1.419	+ £10,452	+ 0.217	+ 0.391	£48,060	**-£2,716**
B. Base-case + R1 to R10	[redacted]	[redacted]	**1.493**	[redacted]	[redacted]	**1.238**	**+ £13,087**	**+ 0.158**	**+ 0.255**	**£82,995**	**+ £32,219**
C. Base-case + R1 to R11	[redacted]	[redacted]	**1.493**	[redacted]	[redacted]	**1.238**	**+ £13,437**	**+ 0.158**	**+ 0.255**	**£85,292**	**+ £34,516**
D. Base-case + R2 to R11	[redacted]	[redacted]	**1.810**	[redacted]	[redacted]	**1.419**	**+ £14,000**	**+ 0.223**	**+ 0.391**	**£62,719**	**+ £11,943**
E. Base-case + R3 to R11	[redacted]	[redacted]	**1.810**	[redacted]	[redacted]	**1.419**	**+ £13,549**	**+ 0.221**	**+ 0.391**	**£61,311**	**+ £10,535**

Costs and QALYs discounted; Life years undiscounted
OS=overall survival; PFS=progression-free survival; QALYs=quality adjusted life years; ToT=time on treatment

2.2 Comparing OS models and trial data

The company's preferred base-case log-logistical OS model and the alternative method proposed by the ERG are directly compared with the LUME-Lung 1 Kaplan-Meier (K-M) trial data in Figure 1. This suggests that the ERG's projected OS more closely fits the K-M data in both arms of the trial. It is also apparent that the main difference in OS estimation between the estimation methods takes effect only after the end of trial follow-up.

However the nature of the deviations from the trial data is visually difficult to appreciate. A better understanding of the differences in model estimates is to calculate and display the deviations of each estimate from the trial data (model residuals) for each trial arm to assess whether there are systematic patterns of deviation, rather than a random scattering of deviations above and below the trial data within a narrow band (i.e. a 'good fit').

Figures 2 and 3 present these differences (residuals) for the company's log-logistic OS model and the method preferred by the ERG (i.e. the OS estimates minus the trial data). The ERG uses K-M data directly from the trial for most of the trial period (i.e. there is no difference between the ERG estimate and K-M data) and only applies a projective model at the end of the OS curve. For both trial arms the log-logistic model over-estimates survival for the first period (up to about 300 days) and then under-estimates OS from 300 to 800 days. This consistent pattern of over- then under-estimation strongly suggests that the log-logistic function is unable to reflect accurately the survival experience of the trial population. By contrast, the ERG's approach relies directly on the trial K-M data for the first 800 to 900 days before employing a simpler projective model calibrated specifically from the patients with longer survival within the trial.

2.3 Area Under Curve (AUC) OS Trends

Figure 4 presents a comparison of the AUC estimated mean OS from the company's log-logistic model and the modelling of OS preferred by the ERG throughout the trial follow-up period and then when projected for the duration of the company's decision model. This shows that there is little difference apparent during the trial. Subsequently the difference between trial arms (i.e. the mean gain in OS per patient) reaches a stable maximum after 9 to 10 years when using ERG projective modelling, whereas the company's log-logistic model is still generating additional OS gain after 16 years. This contrast may be relevant when considering the inherent plausibility of long OS projections in a population acknowledged to have a poor prognosis.

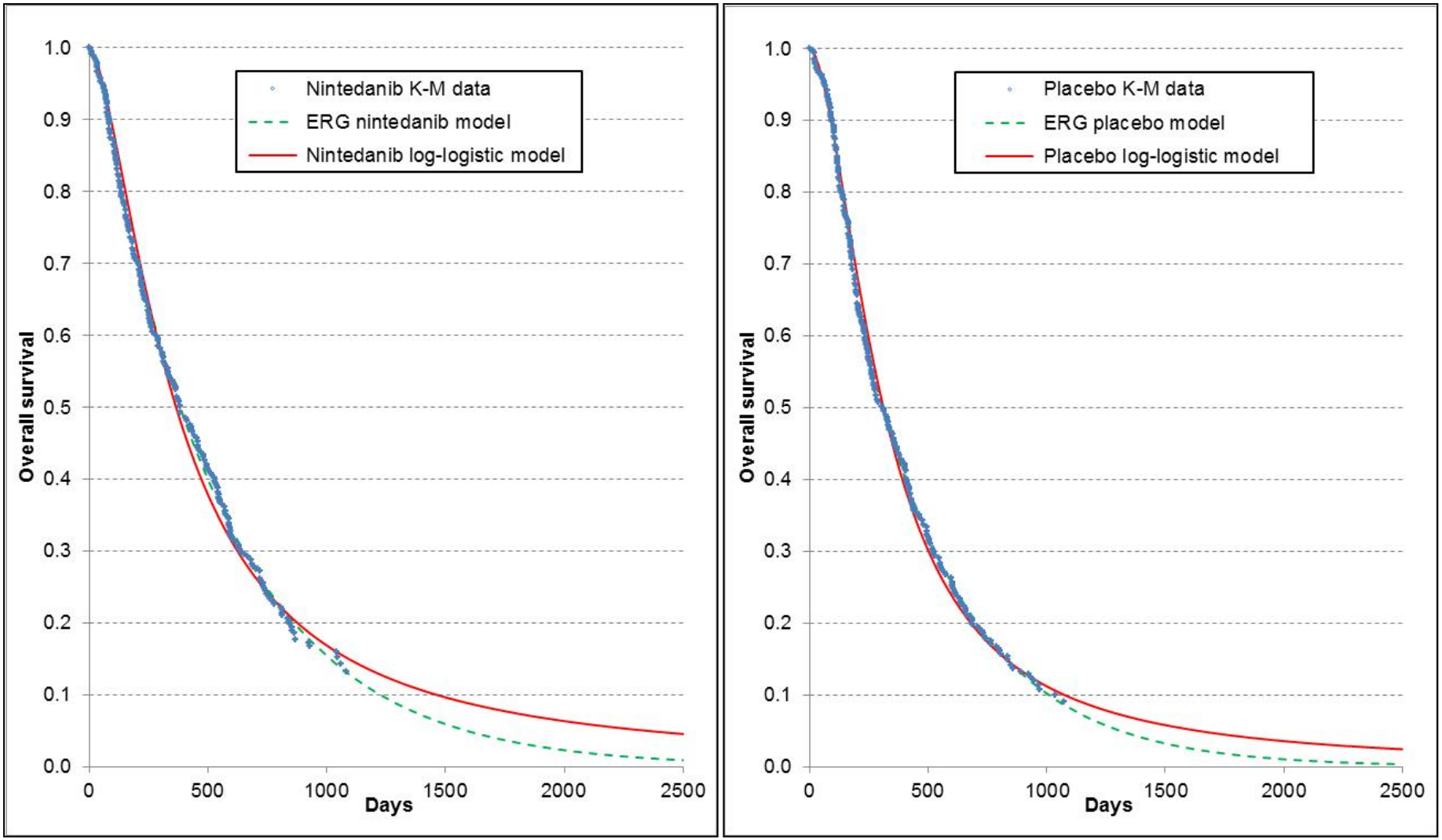


Figure 1 Comparison of company and ERG OS models to the LUME-Lung 1 trial data

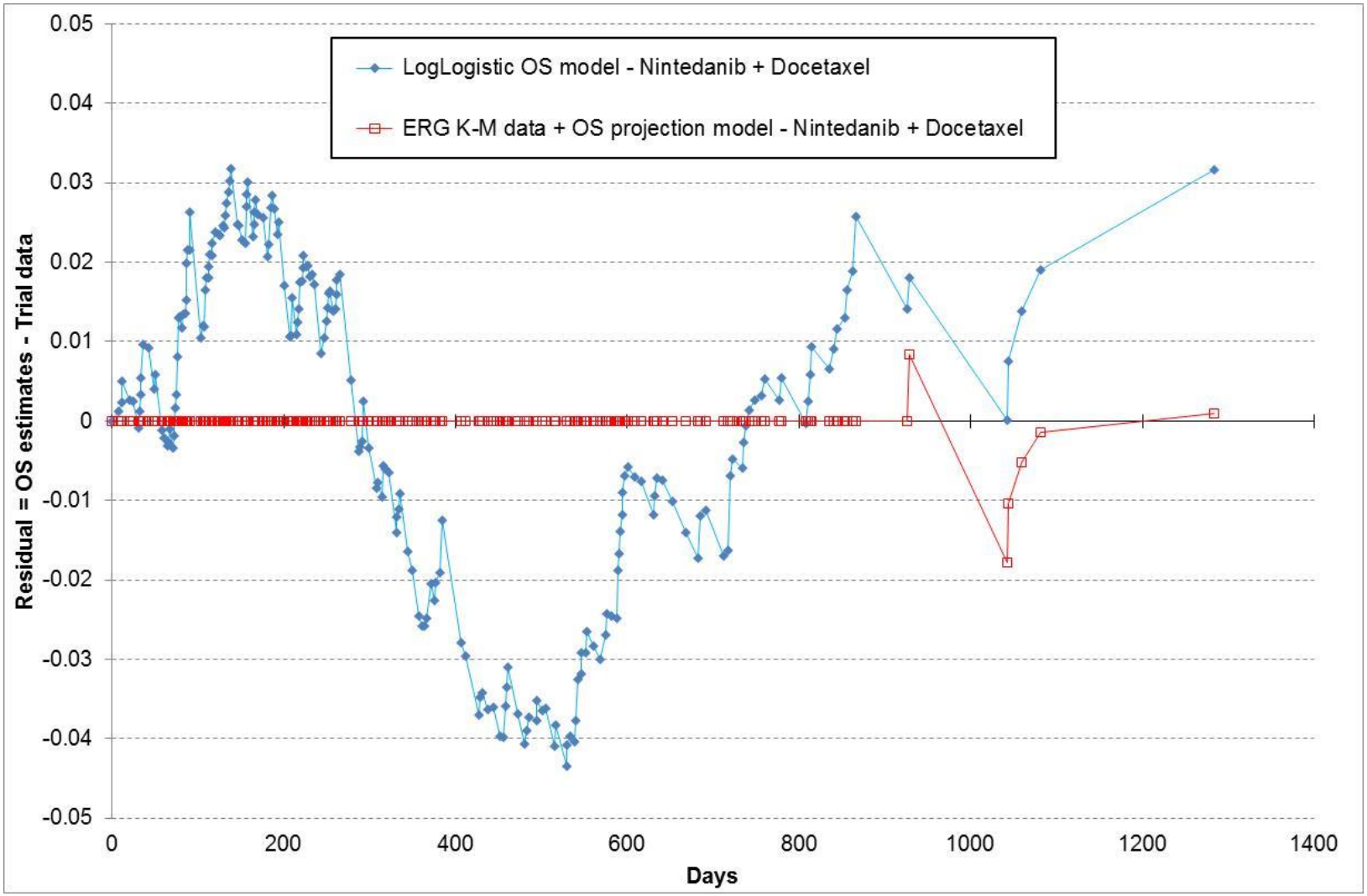


Figure 2 Differences between model OS estimates and LUME-Lung 1 trial data: nintedanib arm

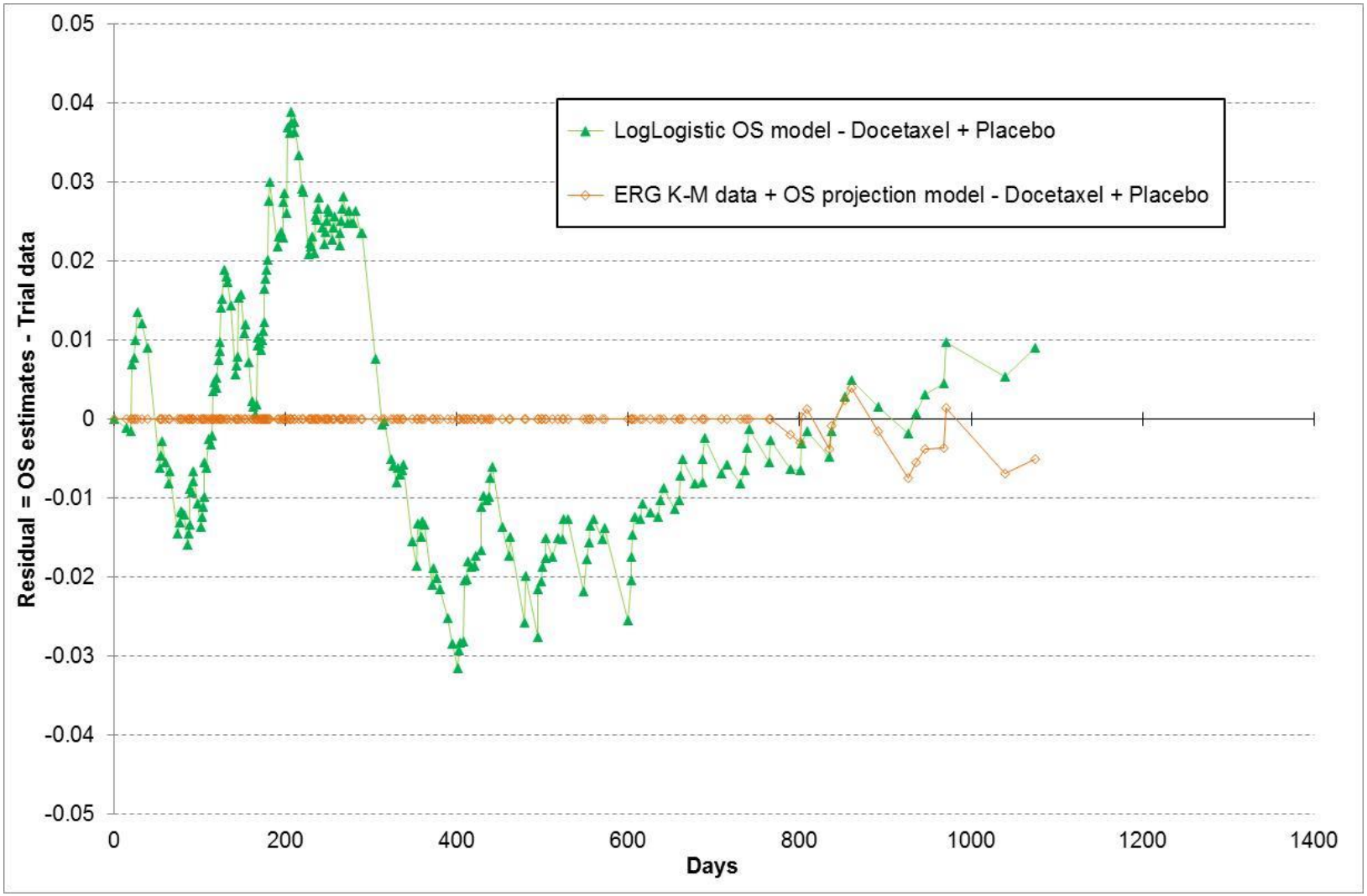


Figure 3 Differences between model OS estimates and LUME-Lung 1 trial data: docetaxel arm

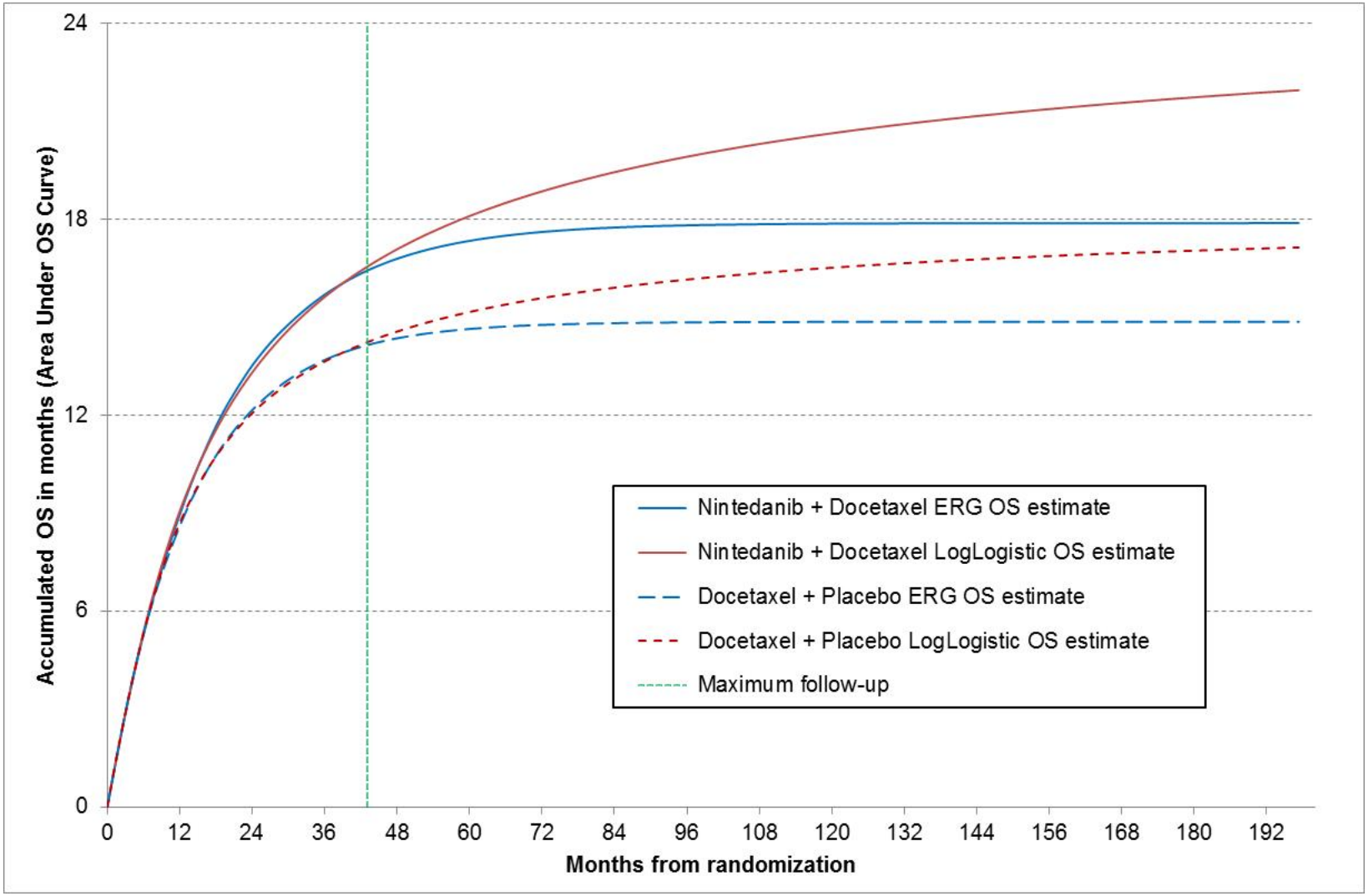


Figure 4 Comparison of Area Under Curve OS estimates over Time