Dacomitinib for untreated EGFR-positive non-small-cell lung cancer (ID1346). – ERG Erratum

Produced by: Warwick Evidence

Authors:
Peter Auguste, Research Fellow in Health Economics[1]
Emma Loveman, Senior Researcher [2]
Daniel Gallacher, Research Fellow in Statistics[1]
Mary Jordan, Research Associate in Health Economics[1]
Rachel Court, Information Specialist[1]
Jacoby Patterson, Honorary Clinical Research Fellow[1]
Jatinder Kaur, Academic Foundation Year Two Doctor[3]
John Green, Consultant in Medical Oncology[4]
Jill Colquitt, Senior Researcher[2]
Xavier Armoiry, Honorary Clinical Research Fellow[1] and Professor of Pharmacy[5]
James Mason, Professor of Health Economics[1]
Lazaros Andronis, Senior Research Fellow in Health Economics[1]

Correspondence to: Dr Lazaros Andronis,
Division of Health Sciences, Warwick Medical School,
University of Warwick, Coventry, CV4 7AL.
Tel: +44 (0) 24 765 74490.
Email: l.andronis@warwick.ac.uk

Date completed: 13th March 2019

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 17/156/04

Declared competing interests of the authors

None

gefitinib, and these changes were considered clinically meaningful (mean scores diarrhoea: 19.88 vs 7.32, p<0.0001; sore mouth: 15.09 vs 3.51, p<0.0001).

Statistically significant differences in the EQ-5D absolute VAS score and utility index were observed in favour of gefitinib.

Rates of any all-cause and treatment-related adverse events were similar between dacomitinib and gefitinib. There were slightly higher rates of any all-cause and any treatment-related grade 3 adverse event and serious adverse events with dacomitinib (based on observation of the proportions only), and dose reductions or temporary discontinuations were more frequently observed with dacomitinib.

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

The CS systematic review of clinical effectiveness was generally well executed. Two studies that the CS excluded from their broader review may have been eligible, however, these omissions would not have affected the results seen. Overall the ERG considered there to be a low chance of systematic error in the findings of the review.

The main clinical evidence for dacomitinib was drawn from the ARCHER 1050 trial which was a multi-centre study comparison with gefitinib. The trial had a high risk of performance bias (owing to the open-label design) but low risks of detection and attrition bias.

The ARCHER 1050 trial presents a number of potential issues in terms of representativeness to the population of England and Wales. There were no UK sites participating in the trial. The proportion of participants from European countries was approximately [redacted]. There was a high proportion of Asian participants, the population was limited to two epidermal growth factor receptor (EGFR) mutations only (albeit the most common ones), and the trial excluded people with brain metastases. In addition, there are imbalances in potential prognostic factors between arms.

The ERG have no concerns about the analysis sets used in the ARCHER 1050 trial or with the censorship and management of missing data used. The outcome measures appear appropriate. With regard to the trial statistics, the CS did not justify why a one-sided p-value was used for PFS and a two-sided p-value for OS and it is unclear to the ERG why there were different data time cut-off points for these two key analyses. The company did not provide significance thresholds alongside p-values presented throughout their submission, and it was unclear to the ERG whether formal

hypotheses were being tested or whether conclusions should be drawn, particularly for the sub-group analyses. The ERG considers that caution is required in the interpretation of the analysis of OS, as the proportional hazards assumption was violated. For patient reported outcomes, there was no adjustment for multiple comparisons.

The CS undertook a network meta-analysis (NMA) comparing dacomitinib with afatinib. The ERG agrees that other than the LUX-Lung 7 trial of afatinib versus gefitinib, there were no other relevant trials for the comparison. The CS adequately described the methods of their NMA approaches and provides a reasonable justification for using the fractional polynomial (FP) analysis. Despite this, the ERG has concerns over the use of the FP analysis with respect to the extrapolations for the survival outcomes but also because there are no detailed results or interpretation of the findings of the FP analysis.

In addition, the CS does not adequately assess the included study populations for transitivity and the ERG considers that the transitivity assumption may be violated. Finally, the CS does not present results of the indirect comparison between dacomitinib and afatinib. Although caution is recommended in the interpretation of the ERG analyses, these show no significant differences between the two respective treatments for PFS or OS.

1.4 Summary of cost-effectiveness evidence submitted by the company

The CS included a systematic review of economic evidence, a review of evidence on resource use and costs, a separate review to identify studies that measure health-related quality of life (HRQoL) in people with non-small-cell lung cancer (NSCLC) patients, and an electronic partitioned survival model built in a widely available spreadsheet application (Microsoft Excel ®).

The search for cost-effectiveness studies comparing the use of dacomitinib against other treatments did not identify any relevant references. The majority of the studies identified evaluated the cost-effectiveness of other treatments. Few relevant studies reporting resource use and costs were identified.

The company constructed a partitioned survival model to trace a cohort of treatment naïve patients with locally advanced or metastatic EGFR-positive NSCLC who may undergo treatment with dacomitinib compared to gefitinib, erlotinib or afatinib. Partitioned survival modelling considers the PFS and the OS curve directly, with the time in post-progression calculated using the difference

in area between the two curves. The company's model comprised three health states: progression-free, post-progression (progressed disease) (PD) and dead.

The model started from a hypothetical cohort of people reflective of the participants in the ARCHER 1050 trial,[1] all of whom began in the progression-free (PF) health state. Over time, people were at risk of progression or death. Transitions between health states was unidirectional and occurred at the end of each 28-day cycle, where people remained in the same health state or progressed. In each cycle, people incurred costs and accrued benefits depending on the health state they occupied. A half-cycle correction was applied in the base-case and the model concluded at a 15-year time horizon.

The company modelled PFS for gefitinib and erlotinib using a generalised gamma curve fitted to the gefitinib arm of ARCHER 1050. They then performed a FP NMA to obtain time-varying hazard ratios for afatinib and dacomitinib and apply these to the gefitinib extrapolation. The ERG found the company's predictions to be pessimistic and preferred a log-normal extrapolation and alternative adjustments for the comparators.

Similarly, for OS the company used a generalised gamma curve for gefitinib and applied HR obtained from FP NMA for the dacomitinib and afatinib. The ERG argues for a log-logistic extrapolation for gefitinib and suggests assuming a HR of 1 from 3 years onwards for the comparators.

Health-related quality of life values for the pre-progression health states were derived from the EQ-5D collected from the ARCHER 1050 study,[1] while utility values for the post-progression health state were obtained from the literature.[2] On clarification, the company provided utility values collected from participants in ARCHER 1050 trial[1] who were post-progression state at follow-up visit (28-35 days following end of treatment visit).. The ERG preferred the use of these utility values; hence they were included in the ERG's base-case. The impact of treatment related adverse events was not accounted for directly in the company's base-case analysis, as it was assumed that these would have been captured by EQ-5D data collected in the trial. However, the ERG argues that it is unlikely that quality of life decrements associated with treatment related AEs are captured by the EQ-5D, unless it is arranged for the instrument to be administered at the same time of these events. Utility decrements (disutilities) were included in the ERG's base-case for treatment related adverse events.

3 Critique of company's definition of decision problem

3.1 Decision problem

The company's decision problem is as follows:

- **Population**: People with untreated locally advanced or metastatic NSCLC with EGFR activating mutation(s).
- **Intervention**: dacomitinib
- **Comparators**: afatinib, erlotinib, gefitinib
- **Outcomes**: overall survival (OS); progression-free survival (PFS); overall response rate (ORR); duration of response (DoR); adverse events (AE); health-related quality-of-life (HRQoL).

There are no subgroups in the NICE scope or in the company decision problem, and there are no special considerations. The company's decision problem is consistent with the NICE scope. The evidence presented from the ARCHER 1050 trial has some deviations from the decision problem as summarised in Table 1.

Table 1: Differences between the decision problem and the evidence provided in the CS.

Issue	ERG comments
Population	
The ARCHER 1050 trial population have either exon 19 deletion (del19) or exon 21 L858R (L858R) substitutions.	This is a narrower population than all EGFR mutations as covered in the scope. Clarification A10 confirms that these were the established activating mutations at the time of design of the ARCHER 1050 trial. The ERG clinical advisor states this has been a common eligibility criteria in clinical trials. These two mutations make up approximately 90% of EGFR mutations (clarification response A4) but it is possible that the other mutations have less favourable responses to treatment. The European Medicines Agency (EMA) have published a positive opinion in January 2018 for dacomitinib monotherapy for the first-line treatment of adult patients with locally advanced or metastatic NSCLC with EGFR activating mutations.[9] The ERG notes that the FDA approval for dacomitinib is in EGFR del19 or L858R substitutions specifically.[10]

However, of these potential comparator studies only one study was actually used in the comparison because there were no links between the other studies in the network (discussed in Section 4.3). This study, the LUX-Lung 7 trial[27, 28] included participants who would match the decision problem of the current appraisal (previously untreated adults with stage IIIB or IV NSCLC and EGFR mutation positive). Appendix D.1.7 lists reference details for included studies and excluded studies with reasons.

A PRISMA style flow-diagram with numbers is presented. Not all excluded studies were available in the original CS, but these were subsequently provided in response to clarification request A2. A two-stage study selection process was undertaken (titles and abstract screening, full paper screening) by two independent reviewers with arbitration from a third reviewer if necessary, for the main SLR. The CS does not state how studies were screened out of the network but the exclusions appear appropriate.

4.1.3 Critique of data extraction

The approach to the data extraction is appropriate (data extraction was by two independent, blinded reviewers and after reconciling differences a third reviewer could be included to reach consensus for any remaining discrepancies, data were extracted in to a pre-specified extraction form).

4.1.4 Quality assessment

The company assessed the quality of the ARCHER 1050 RCT using NICE recommended criteria (CS Table 15) and the Cochrane risk of bias tool (CS Appendix Table 75). There were some differences in the company's responses between these tools, which are summarised in Table 2. The ERG generally agrees with the company's judgements, and notes the potential performance bias (systematic differences in care or in exposure to other factors) that may arise from the open-label design. The risk of detection bias was considered to be low due to blinded IRC review of PFS and ORR (details of how blinding was achieved was provided in clarification response A12).

There was a higher proportion of women (64.3% vs 55.6%) and people with ECOG PS 0 (33% vs 28%) in the dacomitinib arm compared with the gefitinib arm (Section 4.2). The reason for these imbalances is unclear and could be due to selection bias (despite appropriate procedures in place) or could be due to chance. The CS refers to the difference in proportions of females and males in the trial (which is expected given higher frequency of EGFR mutations in females) but makes no

reference to the difference in the rates of females across the arms of the trials. Gender is a potential prognostic factor (see Section 4.2.1). The CS states that generally there is no difference in outcomes between ECOG PS 0 and 1, citing ARCHER 1050 as evidence. The ERG's clinical advisor noted that PS 0 and 1 are usually grouped together in trials, however there is evidence overall that ECOG PS may be an independent prognostic factor (see Section 4.2). The risk of selection bias is therefore uncertain.

The company gives the trial an overall judgement of high risk of bias due to the open-label design. The ERG agrees with this as the trial has a high risk of performance bias (differences between groups in care provided or in exposure to other factors), but notes that the risk of detection bias and attrition bias is low.

Table 2: Risk of bias assessment of ARCHER 1050.

Assessment criteria	**Company response**		**ERG response** (Cochrane tool)
	NICE criteria (CS Table 15)	**Cochrane tool (CS Appendix Table 75)**	
Method used to generate random allocations adequate?	Yes	Low risk	Low risk
Allocation adequately concealed?	Not applicable[a] Open label study	Low risk	Low risk
Groups similar at the outset of the study in terms of prognostic factors?	Yes	Low risk	Imbalance in gender and PS[b]
Care providers and participants blind to treatment allocation?	Not applicable Open label study	High risk	High risk
Outcome assessors blind to treatment allocation?		Low risk	Low risk
Unexpected imbalances in drop-outs between groups?	No	Low risk [d]	Low risk [d] Differences explained
Were the statistical analyses undertaken appropriate? [c]	Yes		
Evidence to suggest authors measured more outcomes than they reported?	No	Low risk	Low risk
Other bias	NR	Unclear Sponsored by pharmaceutical company	Low risk No other bias apparent
Overall judgement	NR	High risk Open-label	High risk

NR, not reported. [a]The company's response is referring to masking of treatment, rather than concealment of the allocation sequence, which the ERG considered appropriate as a central interactive web response system was used. [b] Potential prognostic factors (although not an item on the Cochrane tool). [c]Question as worded in CS Table 15; the full question should be 'Did the analysis include an intention to treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?' [d]Cochrane risk of bias criterion: Attrition bias due to amount, nature or handling of incomplete outcome data.

Table 4: Baseline characteristics in ARCHER 1050

Baseline characteristic	Dacomitinib N=227	Gefitinib N=225
Male, n (%)	81 (35.7)	100 (44.4)
Female, n (%)	146 (64.3)	125 (55.6)
Age, years, median (range)	62 (28-87)	61 (33-86)
Age, years, mean (SD)	61.2 (11.26)	60.9 (10.17)
White, n (%)	56 (24.7)	49 (21.8)
Black, n (%)	1 (0.4)	0
Asian, n (%)	170 (74.9)	176 (78.2)
Japanese, n (%)	40 (17.6)	41 (18.2)
Mainland Chinese, n (%)	114 (50.2)	117 (52.0)
Other East Asian, n (%)	16 (7.0)	18 (8.0)
Never smoked, n (%)	147 (64.8)	144 (64.0)
Ex-smoker, n (%)	65 (28.6)	62 (27.6)
Smoker, n (%)	15 (6.6)	19 (8.4)
ECOG performance status 0, n (%)	75 (33)	62 (28)
ECOG performance status 1, n (%)	152 (67)	163 (72)
Stage IIIB at screening, n (%)	18 (8)	16 (7)
Stage IV at screening, n (%)	184 (81)	183 (81)
Unknown at screening [a], n (%)	25 (11)	26 (12)
del19, n (%)	134 (59)	133 (59)
L858R, n (%)	93 (41)	92 (41)
[a] Newly diagnosed with stage IV at time of study entry.		

There are more female compared with male participants in both treatment groups. The company has referred to this in their summary and have stated that this was to be expected, given the higher proportion of EGFR mutations in NSCLC occurring in females. Nonetheless, 64.3% are female in the dacomitinib group compared with 55.6% in the gefitinib group. The CS states in section B.2.13.2.2 that gender is not a prognostic factor for PFS in EGFR+ NSCLC, citing the ARCHER 1050 trial as evidence. The ERG has identified evidence which suggests that female gender may be an independent prognostic factor for NSCLC, including in those with EGFR-mutation positive NSCLC, summarised below.[30-33] In addition, evidence suggests that females respond better to

dacomitinib were acneiform ([redacted]), diarrhoea ([redacted]) and paronychia ([redacted]), while those related to gefitinib were and increase in ALT ([redacted]) and AST ([redacted]). All-causality adverse events are presented in CS Table 24 and follow a similar pattern.

Specific adverse events resulting in dose reductions, temporary discontinuation and permanent discontinuation are presented in CS Tables 26, 27 and 28, respectively. The most common all-cause adverse events leading to dose reduction were dermatitis acneiform (20.3%), paronchyia (16.7%) and diarrhoea (8.4%) for dacomitinib and increased ALT (2.7%) and AST (2.2%) for gefitinib. The median time to dose reduction was 2.8 months (Inter-quartile range, IQR, 1.3–4.2 months) for dacomitinib, with a median duration of 11.3 months (IQR 4.8–18.9 months). The median time to dose reduction for gefitinib was 3.3 months (IQR 2.4–4.2 months); median duration 5.2 months (IQR 2.5–7.9 months). Temporary discontinuation of dacomitinib resulted most commonly from dermatitis acneiform ([redacted]), paronchyia ([redacted]) and diarrhoea ([redacted]), and of gefitinib due to increased ALT ([redacted]) and AST ([redacted]). Treatment-related adverse events leading to permanent discontinuation of dacomitinib included dermatitis acneiform ([redacted]) and diarrhoea ([redacted]), and those leading to permanent discontinuation of gefitinib included increased ALT ([redacted]).

ERG summary: overall there were similar rates of all-cause and treatment-related adverse events between dacomitinib and gefitinib. However, there were higher rates of any all-cause and any treatment related Grade 3 adverse event and serious adverse events with dacomitinib (based on observation of the proportions only), and dose reductions or temporary discontinuations were more frequently observed with dacomitinib.

4.2.4 Other dacomitinib trials

The ERG identified five additional trials of dacomitinib 45 mg once daily in NSCLC to inform the evidence base on adverse events (dose escalation studies were not considered). One of these[36] was undertaken in treatment naive patients, the rest were undertaken in previously treated patients therefore a different patient population to the NICE scope (Table 9).

- Phase 2 single-arm study (NCT00818441) of dacomitinib in advanced NSCSC (adenocarcinoma subtype) as first-line treatment.[36]

5.2.6 Treatment effectiveness and extrapolation

Two clinical outcomes from ARCHER 1050 were used to inform the transitions between health states in the model:

- Progression-free survival
- Overall survival

5.2.6.1 Time-to-event extrapolation

The company chose to extrapolate both PFS and OS using a combination of parametric and fractional polynomial (FP) models. A generalised gamma model was fitted separately to both the observed PFS and OS data of the gefitinib arm of ARCHER 1050 and extrapolated across the model time horizon. The data used for PFS extrapolation was obtained from Wu et al. (data cut July 2016),[1] whereas the OS data came from Mok et al. (data cut February 2017).[29]

Equivalent efficacy was assumed between gefitinib and erlotinib, whereas for afatinib and dacomitinib, time-varying hazard ratios of their relative effects to gefitinib were estimated using a FP network meta-analysis (NMA). The hazard ratios were then applied to the generalised gamma extrapolation of gefitinib to predict PFS and OS for dacomitinib and afatinib. The ERG have concerns over this approach, as the resulting extrapolations from FP models can be unstable and extremely implausible. The company encountered this during their model selection process, and excluded many FP models based on their implausibility. FP models offer benefits to traditional parametric curves when fitting to survival data with unusual hazard profiles, but it is unclear whether they offer a benefit when extrapolating.

5.2.6.2 Progression-free survival

The company selected a parametric curve for gefitinib through consideration of visual fit, goodness-of-fit statistics and clinical plausibility. They also compared predicted quantiles to those observed in ARCHER 1050. The parametric models had a similar visual fit, and the goodness-of-fit statistics suggested that the Weibull, log-logistic and generalised gamma were the best fitting to the data. The company's clinical experts stated that whilst the 3-year estimates for the log-logistic and log-normal were too optimistic, the 5-year estimates were realistic. The log-logistic and log-normal

It is apparent that the effect of afatinib relative to gefitinib decreases and appears to tend towards 1 across the time horizon of the model, which could be interpreted as a gradual waning effect. However, the efficacy of dacomitinib appears to increase across the time horizon, with the hazard ratio reducing. This results in a clear difference between the long term OS predictions for dacomitinib and gefitinib (Figure 13), of which the ERG's clinical advisor was unconvinced. When applied to the log-logistic extrapolation, the company's FP predicts [redacted]% alive on dacomitinib compared to [redacted]% on gefitinib at 5 years.

The ERG are not aware of any evidence or clinical rationale to support this optimistic prediction for dacomitinib, nor the contrast between the behaviours of the dacomitinib PFS and OS ratios. Recall that for PFS, under the company's FP NMA the HR for dacomitinib grew seemingly exponentially, yet for OS the HR for dacomitinib improves constantly over time. Whilst for afatinib, the PFS improved over time whilst the OS worsened. This contrast between afatinib and dacomitinib is not supported by any clinical rationale, and sheds further doubt over the reliability of the FP analysis extrapolations. The ERG did not consider the other single order FP model to be an improvement in terms of plausibility, with the survival curves for dacomitinib appearing to be almost identical.

Recall also that the company failed to provide evidence of a significant difference between dacomitinib and afatinib for OS in their clinical section. The ERG considered alternative approaches to the extrapolation of dacomitinib OS, including the assumption of equivalency of dacomitinib and afatinib by assuming the FP OS HR from afatinib for both interventions. A summary of the predictions made by the models explored by the ERG is presented in Table 24.

progression survival, and less common for the PFS benefits to be extended in post-progression survival. Under the models fitted with ERG's preferred PFS and OS assumptions, dacomitinib provides an OS and PFS benefit over the comparators, but has a shorter post-progression survival time, consistent with the scenario of the degree of the pre-progression benefit not being repeated in the OS benefit as has been observed in ARCHER 1050.

Table 25: Comparison of pre-progression and post-progression survival gains

Scenario		**Pre Progression Incremental Life Years (Dacomitinib difference)**	**Post Progression Incremental Life Years (Dacomitinib difference)**	**Total Incremental Life Years (Dacomitinib difference)**
Company Base-case	Dacomitinib Gefitinib/Erlotinib Afatinib	[redacted]	[redacted]	[redacted]
ERG PFS and OS log-logistic	Dacomitinib Gefitinib/Erlotinib Afatinib	[redacted]	[redacted]	[redacted]
ERG PFS and OS matched to Afatinib	Dacomitinib Gefitinib/Erlotinib Afatinib	[redacted]	[redacted]	[redacted]
ERG PFS and OS HR=1 at 3 years	Dacomitinib Gefitinib/Erlotinib Afatinib	[redacted]	[redacted]	[redacted]
ERG PFS and OS HR=1 at 4 years	Dacomitinib Gefitinib/Erlotinib Afatinib	[redacted]	[redacted]	[redacted]
ERG PFS and OS HR=1 at 5 years	Dacomitinib Gefitinib/Erlotinib Afatinib	[redacted]	[redacted]	[redacted]

5.2.7 ERG's exploratory survival analysis

5.2.7.1 Progression-free survival

The ERG investigated modelling separately the PFS of dacomitinib from ARCHER 1050 using parametric models, and KM data followed by parametric models fitted to data beyond 8 month, chosen because it was at this point that the KM curves separated. However, none of the parametric curves produced using these models produced a better model than those discussed above. Since the

Table 26: List of adverse events included in the model

Adverse event	Dacomitinib (n=227)	Gefitinib (n=224)	Afatinib (n=160)	Erlotinib*
ALT increased	2 (0.9%)	18 (8.0%)	0 (0.0%)	18 (8.0%)
Diarrhoea	18 (7.9%)	1 (0.4%)	21 (13.1%)	1 (0.4%)
Fatigue	0 (0.0%)	0 (0.0%)	9 (5.6%)	0 (0.0%)
Paronychia	17 (7.5%)	3 (1.3%)	3 (1.9%)	3 (1.3%)
Rash (grouped term)	55 (24.2%)	1 (0.4%)	15 (9.4%)	1 (0.4%)
*Erlotinib assumed equivalent to Gefitinib (see Section B.2.9.1)				

The ERG finds error in the company's use of 7.9% as the proportion of patients experiencing diarrhoea at Grade 3 or above. This is the cumulative value for Grades 3-4 but 0.4% patients experienced grade 5 diarrhoea in the dacomitinib group which has not been counted. Therefore the model input should have been 8.3% rather than the 7.9% used.

Additional data supplied by the company during the clarification process provided AE data per cycle by frequency reported for grades 3 and above at a >2% threshold. Table 27 summarises this new data showing a discernible difference between AE frequencies in patients from the >5% to >2% level. (N.B. rash was not calculated due to lack of clarity as to how the company had defined rash as a grouped term.)

Table 27: Treatment-related adverse events occurring in >2% of patients

Type of AE	Dacomitinib n=227 Frequency (percentage)	Gefitinib n=224 Frequency (percentage)
Stomatitis	[redacted]	[redacted]
ALT increase	[redacted]	[redacted]
Diarrhoea	[redacted]	[redacted]
Paronychia	[redacted]	[redacted]

The mix of AEs represented in table 27 shows the inclusion of stomatitis. Clarification data from the company reported [redacted]. However, stomatitis is a dermatological AE commonly associated with irreversible EGFR TKI inhibitors[71] and showed clinically meaningful worsening in HRQOL in dacomitinib patients assessed using the EORTC QOL questionnaire in ARCHER 1050. Therefore, inclusion of stomatitis in AE data may be relevant and warranted. The further data provided could not definitively show a low rate of recurrence as suggested by the company and to which they attribute as the ability of clinicians to manage these events through dose reduction and discontinuation. It is noted that dose reduction due to AEs was required in 64.8% of dacomitinib patients versus 8.0% of gefitinib patients (see Table 7 and further discussion in clinical effectiveness section).

"Patient reported outcomes were assessed at days 1 (baseline), 8, and 15 of cycle one, on day 1 of subsequent cycles, at the end-of treatment visit, and at the post-treatment follow up visit." [1] (p. 1456).

The company were able to provide these values during the clarification process which are presented in Table 29.

Table 29: Summary of mean EQ-5D health index score.

Time point	Dacomitinib				Gefitinib			
	Median	**Mean**	**95% CI**	**n**	**Median**	**Mean**	**95% CI**	**n**
End of treatment	[redacted]	[redacted]	[redacted]	121	[redacted]	[redacted]	[redacted]	145
Post-treatment	[redacted]	[redacted]	[redacted]	75	[redacted]	[redacted]	[redacted]	107
CI : confidence interval								

The company did not recognize the post-treatment utility values from ARCHER 1050 as a PD progression value, therefore it used a PD utility of 0.64 from the literature.[2] (see Table 30). The use of health utility scores resulting from the Labbé et al, 2017 study[2] was well justified as an alternative source of utility data as it represents real world data obtained from a longitudinal study of EGFR-NSCLC patients directly from EQ-5D-3L. An alternative value from Nafees et al. (2008) sourced in the systematic review was rejected by the company as this study did not meet the NICE reference case.[56] EQ-5D was not used and values were not derived from patients, therefore it was not considered a robust source. In addition, a recent repeat of this study by Nafees et al. in 2017 that was identified in the systematic review, reported a PF value of 0.883 and PD of 0.166,[65] which further demonstrates the unreliability of the original Nafees et al. (2008) study.[50]

The health state utility values obtained from the Labbé, 2017 study[2] would be the ERGs preferred source of utility values from the literature. The ERG finds the company's reasoning for using these values over the Nafees, 2008[50] compelling and appropriate. Similarly, the ERG finds the health state utility values used by Huang (2017)[52] equally less convincing for use in the model as the KEYNOTE-024 trial was directed at patients specifically without *EGFR*-activating mutations or *ALK* translocations.

However, whilst the ERG considers the Labbé et al, 2017[2] values appropriate for use in scenario analysis (when both the PF value of 0.77 and PD value 0.64 are used simultaneously) it prefers the use of PD scores obtained from the ARCHER 1050 trial to be used in the base-case. This continues

Table 46: Results of base-case scenario analysis for the comparison between dacomitinib and afatinib.

Scenario	Dacomitinib			Afatanib			ICER	% change
	Total LYs	Total QALYs	Total costs	Total LYs	Total QALYs	Total costs		
Base-case	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gefitinib survival projection (PFS)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gefitinib survival projection (OS)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
FP model (PFS)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
FP model (OS)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
NMA methodology (PFS and OS)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Utility (PF - ARCHER) with AEs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Utility (PF - Labbé) with AEs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment beyond progression	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

The company presented results for the PSA undertaken around the outcome cost per QALY gained. In general, the company used appropriate distributions around the model input parameters varied. However, the ERG noted that distributions could have been placed around other inputs to reflect the uncertainty, instead of keeping these inputs fixed. A range of sensitivity and scenario analyses were undertaken. These results showed that the discount rate applied to benefits, using the log-logistic parametric curve to model OS for gefitinib and including treatment beyond progression were the key drivers to the ICERs and including treatment beyond progression.

Figure 25: Tornado diagram for the comparison of dacomitinib versus afatinib, using the list prices

Company's scenario analysis results (all treatments at list prices)

The results for each change made and the impact to the base-case results are presented in Table 50 through to Table 52 for the comparison between dacomitinib and the comparators, using the list prices. As seen in these tables, under the scenario of using log-logistic parametric curves for OS and including treatment beyond progression had the greatest impact to the base-case ICER for dacomitinib against all comparators.

64. *National life tables: UK*. Office for National Statistics; 2018. URL: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/nationallifetablesunitedkingdomreferencetables (Accessed 10 Feb 2019).

65. Nafees B, Lloyd AJ, Dewilde S, Rajan N, Lorenzo M. Health state utilities in non-small cell lung cancer: An international study. *Asia Pac J Clin Oncol* 2017;13(5):e195-e203. http://dx.doi.org/10.1111/ajco.12477

66. Brown J, Cook K, Adamski K, Lau J, Bargo D, Breen S, *et al*. Utility values associated with advanced or metastatic non-small cell lung cancer: data needs for economic modeling. *Expert Rev Pharmacoecon Outcomes Res* 2017;17(2):153-64. http://dx.doi.org/10.1080/14737167.2017.1311210

67. *EQ-5D-3L: Health Questionnaire: English version for the UK*. Rotterdam, The Netherlands: EuroQol URL: https://euroqol.org/wp-content/uploads/2016/10/Sample_UK__English__EQ-5D-3L_Paper_Self_complete_v1.0__ID_23963.pdf (Accessed 7 February 2019).

68. Curtis LA, Burns A. *Unit Costs of Health and Social Care 2017*. Canterbury, UK: Personal Social Services Research Unit, University of Kent; 2017. URL: https://doi.org/10.22024/UniKent/01.02/65559 (Accessed 3 February 2019).

69. NHS Improvement. *2016/17 reference cost data*. 2017. URL: https://improvement.nhs.uk/documents/3479/201617_ReferenceCostData.zip (Accessed 25 October 2018).

70. Brown T, Pilkington G, Bagust A, Boland A, Oyee J, Tudur-Smith C, et al. Clinical effectiveness and cost-effectiveness of first-line chemotherapy for adult patients with locally advanced or metastatic non-small cell lung cancer: a systematic review and economic evaluation. Health Technol Assess 2013;17(31):1-278. http://dx.doi.org/10.3310/hta17310

71. Melosky B, Leighl NB, Rothenstein J, Sangha R, Stewart D, Papp K. Management of EGFR TKI-induced dermatologic adverse events. Curr Oncol. 2015;22(2):123-32

Appendix Table 5. Exploratory results, fitting the log-logistic parametric curve to OS data for gefitinib (dacomitinib PAS vs. comparators assumed PAS)

Treatment	Expected mean costs (£)	Incremental costs (£)	Expected mean QALY	Incremental QALY	ICER (£)
Gefitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Erlotinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Afatinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Dacomitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life years gained

Appendix Table 6. Exploratory results, fitting the log-logistic parametric curve to OS data for gefitinib (all treatments at list prices)

Treatment	Expected mean costs (£)	Incremental costs (£)	Expected mean QALY	Incremental QALY	ICER (£)
Erlotinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gefitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Afatinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Dacomitinib	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life years gained

OS, log-logistic curve to OS data for gefitinib and assumed equal efficacy, on the hazard scale, from 36 months onwards

Here, we fitted the log-logistic parametric curve to the OS data for gefitinib and assumed a hazard ratio of 1 is applied from 36 months onwards. Applying a HR=1 results in reduction to the post-progression survival and thus post-progression QALYs (results not shown). For comparison (i) afatinib is associated with an approximate ICER of [redacted] per QALY as compared to gefitinib. The ICER for dacomitinib compared to afatinib is approximately [redacted] per QALY. Under these assumptions resulted in an increase to the ICERs. (see Appendix Table 7). For comparison (ii), the ICER for afatinib (versus gefitinib) and dacomitinib (versus afatinib) are over [redacted] and [redacted] per QALY, respectively (Appendix Table 8).