

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Ramucirumab for previously treated locally advanced or metastatic non-small-cell lung cancer [ID838]

The following documents are made available to the consultees and commentators:

- 1. Response to consultee, commentator and public comments on the Appraisal Consultation Document (ACD)**
- 2. Consultee and commentator comments on the Appraisal Consultation Document** from:
 - Lilly
 - Pfizer
- 3. ERG's response to Lilly's ACD consultation comments**
- 4. Responses to questions posed to Dr Yvonne Summers, clinical expert**

'No comments' responses were received from British Thoracic Oncology Group (BTOG) and the Department of Health.

No responses were received from patient experts or through the NICE website consultation.

Any information supplied to NICE which has been marked as confidential, has been redacted. All personal information has also been redacted.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal

Ramucirumab for previously treated locally advanced or metastatic non-small-cell lung cancer [ID838]

Response to consultee, commentator and public comments on the Appraisal Consultation Document (ACD)

Definitions:

Consultees – Organisations that accept an invitation to participate in the appraisal including the companies, national professional organisations, national patient organisations, the Department of Health and the Welsh Government and relevant NHS organisations in England. Consultees can make a submission and participate in the consultation on the appraisal consultation document (ACD; if produced). All non-company consultees can nominate clinical experts and/or patient experts to verbally present their personal views to the Appraisal Committee. Company consultees can also nominate clinical experts. Representatives from NHS England and clinical commissioning groups invited to participate in the appraisal may also attend the Appraisal Committee as NHS commissioning experts. All consultees have the opportunity to consider an appeal against the final recommendations, or report any factual errors, within the final appraisal determination (FAD).

Clinical and patient experts and NHS commissioning experts – The Chair of the Appraisal Committee and the NICE project team select clinical experts and patient experts from nominations by consultees and commentators. They attend the Appraisal Committee meeting as individuals to answer questions to help clarify issues about the submitted evidence and to provide their views and experiences of the technology and/or condition. Before they attend the meeting, all experts must either submit a written statement (using a template) or indicate they agree with the submission made by their nominating organisation.

Commentators – Commentators can participate in the consultation on the ACD (if produced), but NICE does not ask them to make any submission for the appraisal. Non-company commentator organisations can nominate clinical experts and patient experts to verbally present their personal views to the Appraisal Committee. Commentator organisations representing relevant comparator technology companies can also nominate clinical experts. These organisations receive the FAD and have opportunity to report any factual errors. These organisations include comparator technology companies, Healthcare Improvement Scotland any relevant National Collaborating Centre (a group commissioned by NICE to develop clinical guidelines), other related research groups where appropriate (for example, the Medical Research Council and National Cancer Research Institute); other groups such as the NHS Confederation, the NHS Commercial Medicines Unit, the Scottish Medicines Consortium, the Medicines and Healthcare Products Regulatory Agency, the Department of Health, Social Services and Public Safety for Northern Ireland).

Public – Members of the public have the opportunity to comment on the ACD when it is posted on the Institute's web site 5 days after it is sent to consultees and commentators. These comments are usually presented to the appraisal committee in full, but NICE reserves the right to summarise and edit comments received during consultations, or not to publish them at all, where in the reasonable opinion of NICE, the comments are voluminous, publication would be unlawful or publication would be otherwise inappropriate.

Please note: Comments received in the course of consultations carried out by NICE are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the submissions that NICE has received, and are not endorsed by NICE, its officers or advisory committees.

Comments received from consultees

Consultee	Comment	Response
Lilly	The company provided comments on the approach to modelling survival using single model versus separate models. It considered that modelling treatment arms together (single model) with treatment as a covariate was more appropriate than separate models, which may result in less robust predictions. See section 1 of the company comments on the ACD for more details.	Comments noted. The committee considered the comments from the company regarding the models used in its own analyses and those of the ERG. It acknowledged that there were different approaches to the modelling but on balance preferred the ERG’s approach. The committee concluded that the ERG’s approach to modelling survival was more reasonable than the company’s approach and better reflected the data from the trial (see FAD section 4.9).
Lilly	The company provided comments on underestimation of observed mean survival for the docetaxel group using the log-logistic model. The company considered that in preferring the ERG’s model the committee agreed that ramucirumab plus docetaxel provides a small extension to overall survival compared to nintedanib plus docetaxel. It also commented that TA347 considered nintedanib plus docetaxel compared with docetaxel alone met end of life criteria and that the order of appraisals affects the end of life criteria decision. See section 2 of the company comments on the ACD for more details.	Comment noted. The committee heard from the ERG that the log-logistic curve was a good fit for the ramucirumab plus docetaxel data but not for the docetaxel-alone data. The committee considered the results of the linear trend model, when comparing ramucirumab plus docetaxel with docetaxel alone and the comparison with nintedanib plus docetaxel. The committee concluded that the extension to life criterion was met in the population with non-squamous disease when comparing with docetaxel alone but not met in the full population or the squamous population, when comparing with docetaxel alone or for the non-squamous population who would otherwise receive nintedanib plus docetaxel (see FAD section 4.9).

Confidential until publication

Consultee	Comment	Response
Lilly	The company provided comments on the model fitted by the ERG. The company considered that the ERG had not tested the point of extrapolation by smoothing the hazard rates and that the smoothed hazard rates did not support the assumption of a linear trend after 11 months. See section 3 of the company comments on the ACD for more details.	The committee was aware that the ERG presented exploratory analyses using a linear trend model to extrapolate survival from month 13 onwards because the trial data showed a constant hazard for death after 11 months but acknowledged that there was some arbitrariness in the selection of this time point. The committee concluded that the ERG's approach to modelling survival was more reasonable than the company's approach and better reflected the data from the trial (see FAD section 4.9).

Comments received from commentators

Commentator	Comment	Response
Pfizer	Any recommendation for ramucirumab should exclude its use in ALK-positive NSCLC	Comment noted. The committee did not recommend ramucirumab in combination with docetaxel for treating locally advanced or metastatic non-small-cell lung cancer in adults whose disease has progressed after platinum-based chemotherapy. Therefore no action required.

No comment received from:

British Thoracic Oncology Group
Department of Health



16th May 2016

Helen Knight
Associate Director – Appraisals
Centre for Health Technology Evaluation, NICE
Level 1A, City Tower
Manchester
M1 4BT

Eli Lilly and Company Limited

Lilly House
Priestley Road
Basingstoke
Hants
RG24 9NL
+44 (0)1256 315000

RE: Lilly response to appraisal consultation document (ACD): ramucirumab for previously treated locally advanced or metastatic non-small-cell lung cancer [ID838]

Dear Helen

Lilly are naturally disappointed that NICE has not recommended ramucirumab, in combination with docetaxel, within its marketing authorisation for treating locally advanced or metastatic non-small-cell lung cancer in adults whose disease has progressed after platinum-based chemotherapy.¹ However we are pleased that the committee acknowledged that the REVEL trial was of good quality and concluded that current evidence suggests that RAM+DOC has an acceptable safety profile compared with DOC alone.

Furthermore Lilly welcomes that there is a great deal of agreement between Lilly and the committee on the general approach to economic analysis and that it was primarily the plausibility of the survival extrapolation that was questioned by the committee. We explain in technical detail why we believe our approach to extrapolation is appropriate in Appendix 1.

Whilst welcoming the level of agreement on the economic approach at this stage, Lilly requests that the end-of-life criteria be further considered by the committee at the next meeting. With respect to paragraph 4.16 of the ACD, which considers the extension of life criterion, Lilly understand how the committee reached the conclusion stated in the ACD, given the choice of ERG model they preferred. However, Lilly notes that in preferring this model the committee therefore agree that RAM+DOC provides a small extension to overall survival compared to NIN+DOC, and furthermore notes that in TA347 NICE considered that the extension of life criterion applied to the consideration of NIN+DOC vs DOC.² Lilly therefore requests that the committee reconsider whether it would be more consistent with TA347 to allow the end-of-life criteria to be applied in consideration of the non-squamous subgroup.

In patients with non-squamous NSCLC for whom NIN+DOC is a treatment option, i.e. adenocarcinoma, Lilly notes that the extension to life criterion cannot be met if the only comparator considered relevant is NIN+DOC. Given Lilly's estimates of current market share (table 96 of the submission) and that erlotinib is no longer a relevant comparator in this (EGFR negative) population, the remaining market shares for DOC and NIN+DOC represent 70% and 30% of the non-squamous, NSCLC market respectively. Lilly therefore considers that it would be appropriate for the committee to consider a comparison of RAM+DOC to DOC in such patients and that it would be appropriate for the end-of-life criteria to be applied to such a comparison.

The ERG report,³ Figure 9, shows that in this subgroup the ERG's linear trend model (preferred by the committee) delivers an estimated 3.9 months mean overall survival gain, which clearly meets the standard extension of life criterion for the comparison of RAM+DOC to DOC. As it stands, the ACD represents a situation where the order in which the NIN and RAM appraisals have fallen is the only factor which affects whether the end-of-life criteria are considered to apply to the appraisal of RAM. Given that DOC remains the standard of care in this patient population the end of life criteria should be considered for this comparator in addition to NIN+DOC. If such a comparison from RAM+DOC to DOC with the end-of-life criteria were not considered by the committee this would be inconsistent with TA347.

We have provided some additional points of clarification in appendix 2

Please contact me if you have any further queries.

Yours sincerely,

██████████
████████████████████

References

1. NICE. Appraisal consultation document: Ramucirumab for previously treated locally advanced or metastatic non-small-cell lung cancer, 2016.
2. NICE. Nintedanib for previously treated, locally advanced, metastatic, or locally recurrent non-small-cell lung cancer (TA347). Available at: <http://www.nice.org.uk/guidance/ta347> (accessed October 20, 2015). 2015.
3. Loveman E, Cummins E, Connock M, et al. Ramucirumab for previously treated locally advanced or metastatic non-small-cell lung cancer: A Single Technology Appraisal: Warwick Evidence, 2016.

Appendix 1 – Lilly response to Appraisal Committee comments on the approach to modelling survival (ACD Section 4.9, page 15)

“The committee heard from the ERG that the log-logistic curve was a good fit for the ramucirumab plus docetaxel data but not for the docetaxel-alone data so separate models should have been fitted to the different groups. The committee noted the ERG’s comment that the company’s modelling approach underestimated survival for the docetaxel group compared with the observed data in the Kaplan–Meier curve.” (ACD page 15)

“The committee noted that the company’s modelling approach underestimated survival for the docetaxel group compared with the observed data and this would continue in the extrapolation. The committee was also concerned that the company’s approach assumed that the probability of death reduced over time.” (ACD page 24)

We understand the need to investigate different modeling approaches, as finding the best fitting model to clinical data is rarely straightforward. There are multiple criteria for selecting the most appropriate modeling approach. In the case of REVEL overall survival data the test of proportionality of hazard suggested that modeling treatment arms together with treatment as covariate would be appropriate – as suggested by the NICE technical support document (TSD) 14¹: “Generally, when patient-level data are available, it is unnecessary to rely upon the proportional hazards assumption and apply a proportional hazards modelling approach – the assumption should be tested which will indicate whether it may be preferable to separately fit parametric models to each treatment arm, or to allow for time-varying hazard ratios.”¹

Under this approach, with individual patient level data, either proportional hazards models or accelerated failure time models (e.g. log-logistic) can be used, as indicated in the TSD 14; “...but log-logistic and log normal models are accelerated failure time models and do not produce a single hazard ratio (HR), and thus the proportional hazards assumption does not hold with these models. However, modelling using treatment group as a covariate can still be undertaken with these models, with the treatment effect measured as an ‘acceleration factor’ rather than a HR.”¹

1) Single model versus separate models

“In the opinion of the ERG, if log log models are adopted, the preferred option would be to fit separate models for each arm.”

While visual inspection indicated “non-perfect fit” of the log-logistic model at the end of the overall survival curve for the control arm this may not describe fit comprehensively and the AIC/BIC statistics provide additional information.

The ERG preferred the use of models fitted to separate study arms. However, only the visual fit was used to justify this statement. Further model fit tests could have been conducted to support this statement.

Model fit for these types of models can be tested for by allowing both the scale and shape parameters to differ by treatment using what is often referred to as a stratified model in statistical software packages. A number of models have been fitted using this technique and also flexible spline-based models with time varying effects and piecewise constant hazard models to further investigate the best choice of model.

Figure 1 presents the AIC scores for the different models and Figure 2 presents the BIC scores.

Figure 1 AIC fit statistics for models tested

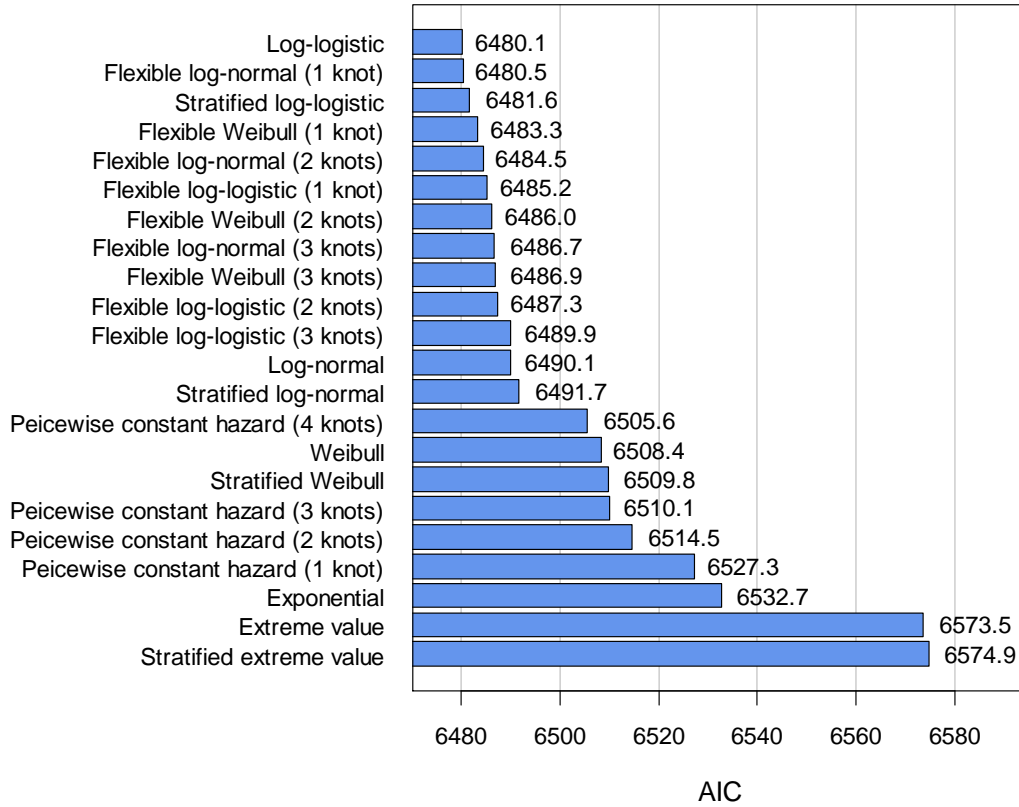
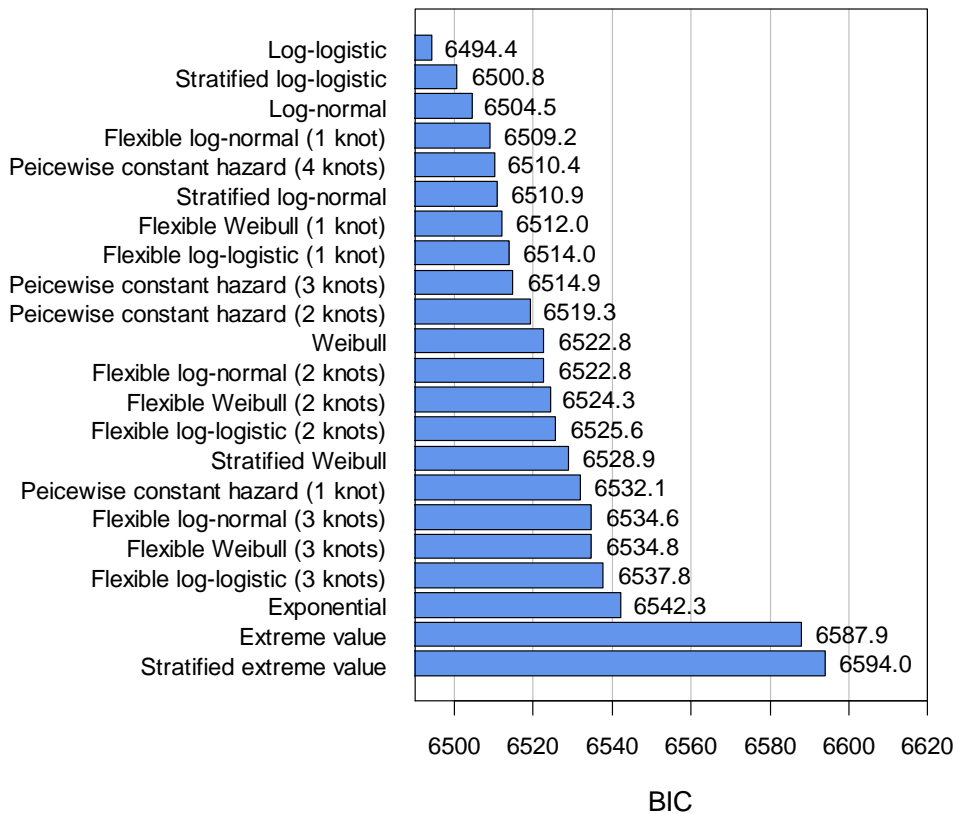


Figure 2 BIC fit statistics for models tested



Both the AIC and BIC scores support the choice of the log-logistic model with a common shape parameter. This suggests that although fitting separate models may look visually appealing there is a risk of over-fitting models to the data which will result in less robust predictions being made. The choice of splitting the data by subgroups such as histology may also cause further problems of over-fitting.

2) Underestimation of observed mean survival for docetaxel group by log-logistic model

“The committee noted that the company’s modelling approach underestimated survival for the docetaxel group compared with the observed data and...” (ACD page 24)

When the judgement of goodness-of-fit is based on visual inspection only it may appear that the fit of selected model does not “adequately” follow the Kaplan–Meier data from clinical trial. However, as summarised above, a strategy of fitting models that “may look visually appealing” as opposed to models indicated by statistical tests may carry a risk of “over-fitting”. This may misguide the conclusion regarding the model predictions.

We wish to demonstrate this point by comparing model predictions directly with REVEL overall survival data. In Table 1 below the mean survival times for both treatment arms based on different modeling approaches are presented using REVEL trial based restricted means as reference. The model horizon is set for “trial length” which has been set at 33 months.

Table 1 Mean overall survival time estimates using different modelling approaches.

REVEL ITT population	Mean OS estimated for trial length (33 months)		
	RAM+DOC	PBO+DOC	Difference
REVEL restricted mean	13.98	11.97	2.02
Lilly log-log model	13.92	12.11	1.81
ERG KM + linear Trend	13.83	12.40	1.43
ERG log-log (by treatment-arm)	13.86	12.50	1.36

ITT: intention-to-treat; OS: overall survival; RAM: ramucirumab; DOC; Docetaxel; PBO: placebo; ERG: evidence review group; KM: Kaplan–Meier

The table demonstrates that while all models slightly underestimate the trial based mean overall survival time for the RAM+DOC arm, they *all* (including the manufacturer’s model) *overestimate* the mean overall survival in the control arm. Therefore it appears that all models underestimate the mean overall survival benefit of RAM+DOC and both of the ERG’s models produce very conservative estimates of mean survival – keeping in mind that the *median* overall survival difference in REVEL was 1.4 months. Of note is that the relative overall survival benefit for RAM+DOC vs DOC during the trial period (i.e. the restricted mean) is 2.02 months. The ERG’s modelling approach significantly underestimates this survival benefit and seems to be at odds with the relative efficacy observed in the REVEL double blind randomised controlled trial.

In conclusion, based on these data, Lilly’s approach to modelling overall survival is justified.

3) The hybrid model fitted by the ERG

“It was aware that the ERG presented exploratory analyses using a linear trend model to extrapolate survival from month 13 onwards because the trial data showed a constant hazard for death after 11 months. The committee preferred the ERG’s approach because the linear trend model provided a better fit to the trial data than the company’s log-logistic model.” (ACD page 15)

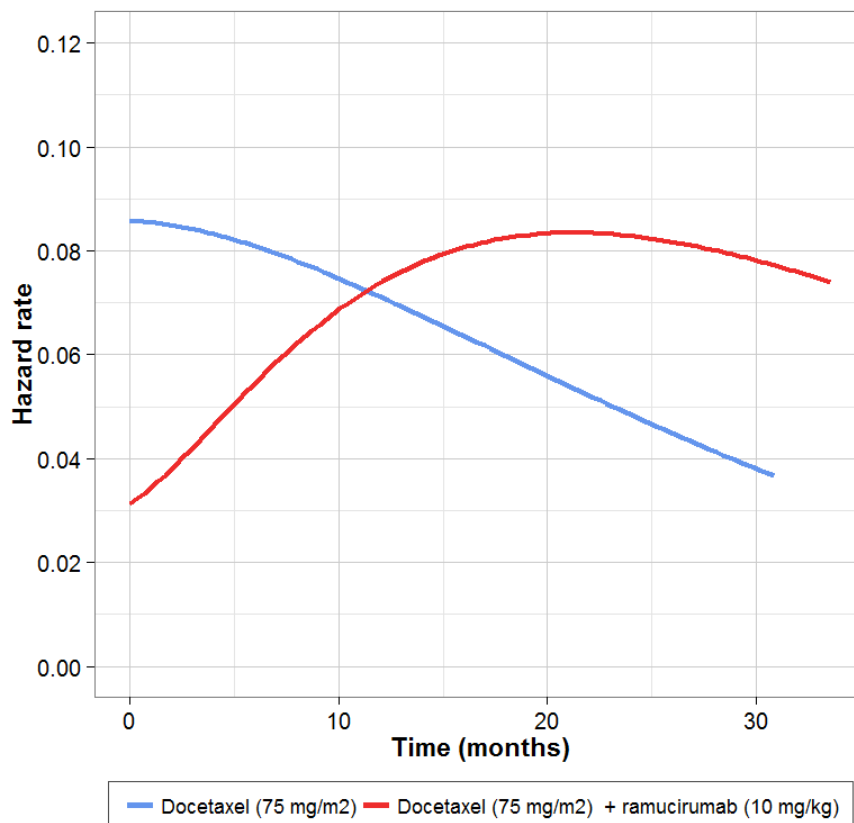
We think it would also be worth highlighting that there may be uncertainty related to alternative modeling approaches.

The ERG fitted linear trend models to each arm after 11 months.

“When trial cumulative hazard plots are examined it is evident that beyond about 11 months there is linear trend for each arm, suggesting that from this time a constant hazard fits the observed data.”

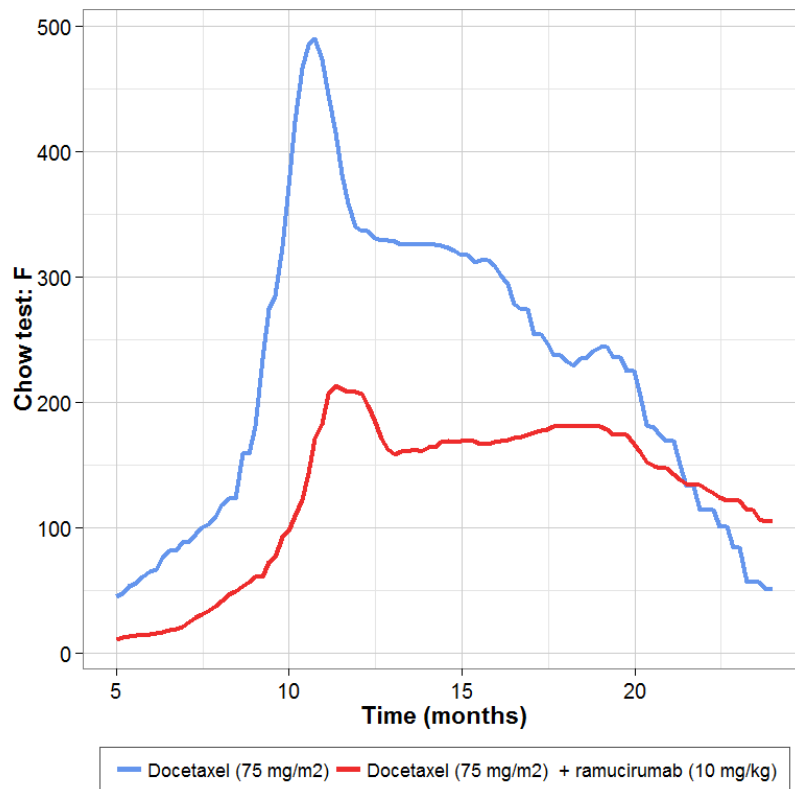
However, the ERG did not appear to have tested this by plotting smoothed hazard rates. The plot based on the total sample in the REVEL trial is shown in Figure 3.

Figure 3 Smoothed hazard rates estimated from the REVEL overall survival data



The smoothed hazard rates do not appear to support the assumption of a linear trend after 11 months. The hybrid model used by the ERG appeared to make a number of assumptions and no justification for the cut-off point was made other than a visual assessment. This could have been more objectively tested through the use of the Chow test.² The results of this technique are presented in Figure 4.

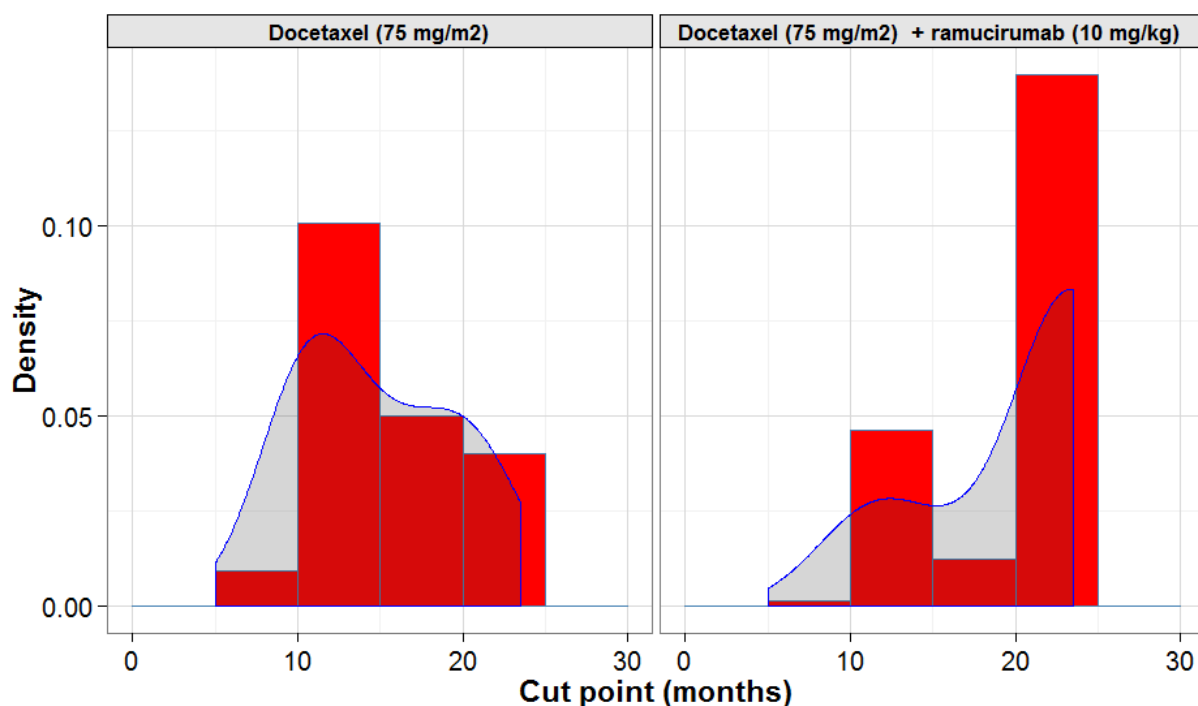
Figure 4 Results from the Chow break test



The optimum cut-off point is where the F statistic from the Chow test reaches its maximum. The optimum cut-off point for the docetaxel arm appears to be close to 11 months. However, there appears to be uncertainty as to where the cut-off point for the docetaxel plus ramucirumab arm should be.

The ERG also did not appear to address the overall uncertainty of fitting a hybrid model. In the original publication of this type of approach by Gelber³ bootstrapping was used to take account of the error. This model can also be extended so that the Chow break test is performed for each bootstrap sample so that the uncertainty of where the cut-off point should be applied is taken into account. Figure 5 presents the distributions of the cut-off points from 1000 bootstrap samples.

Figure 5 Distributions of the cut-off points found from the Chow break test for different bootstrap samples



The choice of cut-off for the two arms in the study appears to differ. It is not clear if applying the same cut-off point for both treatment arms is the correct approach. The more general conclusion would be that it is unclear if the novel method applied by the ERG actually reduces the uncertainty compared to the manufacturer's submitted base-case model.

We would also argue that since hybrid models have not yet been validated for over-fitting data and currently there are no methods to compare the fit of these types of models with parametric models the results from hybrid models do need to be treated with caution.

References

1. Latimer, N. NICE DSU Technical Support Document 14: Undertaking survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data. 2011. Available from <http://www.nicedsu.org.uk>
2. Chow GC. Tests of Equality Between Sets of Coefficients in Two Linear Regressions. *Econometrica* 1960;28:591.
3. Gelber RD, Goldhirsch A, Cole BF. Parametric extrapolation of survival estimates with applications to quality of life evaluation of treatments. *International Breast Cancer Study Group. Control Clin Trials* 1993;14:485-99.

Appendix 2 - Points of clarification for the ACD

Lilly considers that a number of points of clarification warrant consideration for the next draft of the appraisal document, i.e. in the FAD, as follows:

Issue 1

Description of problem	Description of proposed amendment	Justification for amendment
<p>Paragraph 4.4 states “For the subgroup with squamous disease, the committee noted that the difference of 1.5 months in progression-free survival was statistically significant between the 2 treatment groups but that the overall survival difference was not statistically significant.”</p>	<p>For the subgroup with squamous disease, the committee noted that the difference of 1.5 months in progression-free survival was statistically significant between the 2 treatment groups but that the overall survival difference was not statistically significant. <i>However the committee acknowledges that the REVEL trial was not powered for subgroup histology and that the lack of statistical significance noted in the overall survival result given for the squamous subgroup is likely to be due to small patient numbers in this subgroup</i></p>	<p>This acknowledgment would avoid allowing the reader to develop any impression that RAM+DOC had been found ineffective in a fully powered analysis and would reinforce the point made in the ACD elsewhere in paragraph 4.4, viz. “REVEL was not powered for subgroup histology</p>

Issue 2

Description of problem	Description of proposed amendment	Justification for amendment
<p>Paragraph 3.17 states “The ERG noted that the company had not used the actual EQ-5D data collected in REVEL. The company had instead assumed that quality of life was the same in each group while on treatment (that is, the company pooled the EQ-5D values from the trial) but made small allowances for different side effects.</p>	<p>The ERG noted that the company had not used the actual EQ-5D data collected in REVEL. The company had instead assumed that quality of life was the same in each group while on treatment (that is, the company pooled the EQ-5D values from the trial) but made small allowances for different side effects.</p>	<p>The first sentence in the statement is not factually accurate as the EQ-5D data from REVEL were in fact used, as whole-trial data rather than per trial-arm data. These changes will avoid confusion due to the contradicting statements currently in the ACD.</p>

Issue 3

Description of problem	Description of proposed amendment	Justification for amendment
<p>Paragraph 3.20 states “When comparing ramucirumab plus docetaxel with docetaxel alone the linear trend models showed a mean extension in overall survival of 2.20 months for the full population and 1.10 months for the population with squamous disease.”</p> <p>Paragraph 4.16 states “It also considered the results of the linear trend model, when comparing ramucirumab plus docetaxel with docetaxel alone (mean extension in overall survival of 2.20 months for the full population and 1.10 months for the population with squamous disease)”</p>	<p>When comparing ramucirumab plus docetaxel with docetaxel alone the linear trend models showed a mean extension in overall survival of 2.20 months for the full population, 3.9 months for the population with non-squamous disease and 1.10 months for the population with squamous disease.</p> <p>It also considered the results of the linear trend model, when comparing ramucirumab plus docetaxel with docetaxel alone (mean extension in overall survival of 2.20 months for the full population, 3.9 months for the population with non-squamous disease and 1.10 months for the population with squamous disease)</p>	<p>Paragraph 3.20 and Paragraph 4.16 both relate to the extension of life modelled in the overall population and the squamous and non-squamous subgroups; however the overall survival gains are not uniformly reported in these three populations. Lilly request that both paragraphs should have added to them explicit mention of the committee’s preferred overall survival gain of 3.9 months in the comparison of RAM+DOC to DOC in non-squamous patients; as presently written, these paragraphs do not reflect a balanced report of the modelled survival outcomes in these three populations.</p>

ACD - Consultees & Commentators: Lung cancer (non-small cell, metastatic) - ramucirumab [ID838]

Response from Pfizer Ltd | 16th May 2016

Patients with anaplastic lymphoma kinase (*ALK*)-positive NSCLC account for around 3% of non-squamous NSCLC. Ramucirumab is an anti-angiogenic agent which has no biomarker and hence is unlikely to be a preferred option in *ALK*-positive disease. In the REVEL trial, data were not presented by *ALK* status. As such, the ACD for this appraisal indicates that the manufacturer did not submit a comparison versus crizotinib in *ALK*-positive patients, and that the NICE Committee agreed that a comparison versus crizotinib was not appropriate.

Any recommendation for ramucirumab should therefore explicitly exclude its use in patients with *ALK*-positive NSCLC or state that NSCLC patients with *ALK* mutations should have first progressed on an approved *ALK*-targeted therapy, prior to being considered for ramucirumab.

ERG's response to Lilly's ACD consultation comments

Ramucirumab for previously treated locally advanced or metastatic non-small-cell lung cancer

Produced by ERG: Warwick Evidence

Authors: Emma Loveman, Senior Researcher, Effective Evidence LLP
Ewen Cummins, Health Economist, McMDC Ltd
Martin Connock, Senior Research Fellow, Warwick Evidence
Aileen Clarke, Professor of Public Health Research, Warwick Evidence

Correspondence to: Aileen Clarke
Warwick Medical School
University of Warwick
Coventry
CV4 7AL
02476 574505

Date completed: 23/05/2016

Warwick Evidence response to Lilly letter and Appendices 16/05/16

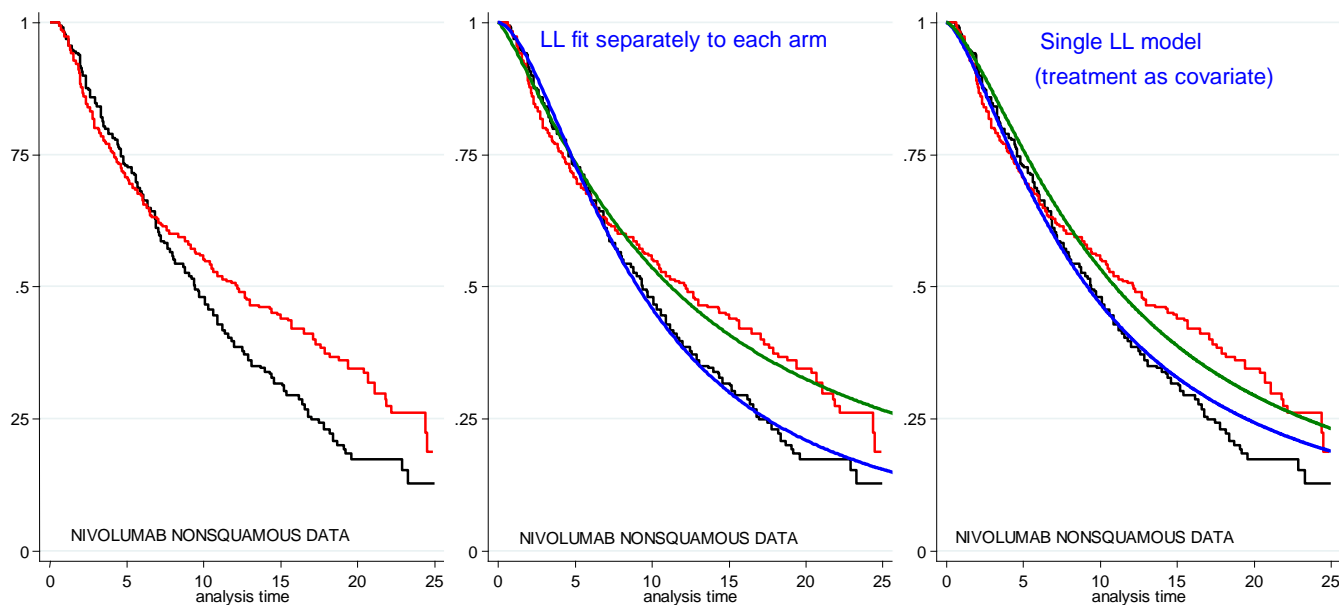
Response to letter clarification points:

Appendix 1

1a] Single versus separate models.

Selecting a single model is unnecessary when data is available for both arms. The single model approach uses treatment as a covariate and arguably will produce models less faithful to the observed data than separate models. An extreme example below (NB. far more extreme than in REVEL) illustrates the point:

The observed data (represented by the KM plots left) departs from proportional hazards (the plots cross over). With separate loglog models for each arm (middle) the fits cross over (faithful to observed) and visually fit each arm better than the single loglog model (treatment as covariate) in which arm fits do not cross over or fit as well.



1b) AIC and BIC values.

As Lilly point out that many models are possible and may provide useful fits.

The company have produced an interesting set of information criteria values for a variety of models. Most of these, including flexible parametric models, were explored by the ERG but not reported because of the bulk of the ERG report. AIC and BIC values are a widely used and often useful guide in selecting models, however it should be emphasised that they are only a guide and that no statistical tests have yet been developed to differentiate between different scores; often models that generate very different extrapolations beyond the data exhibit trivial differences in AIC BIC scores. The choice of models often rests on the reasonableness of extrapolation and other considerations than AIC BIC scores. The ERG's opinion remains that the single loglog model approach of the company is less acceptable than their own approach because:

- i. in the company modelling almost half of the gain from ramucirumab accrues well after patients have progressed and ceased ramucirumab treatment. Such benefit gain is rather implausible.
- ii. The loglog model predicts decreasing probability of death over the whole of the model time beyond progression; this is not supported by examination of cumulative hazard beyond ~12 months which clearly indicates that hazard (probability of death) becomes constant after about 12 months; it is reasonable to expect that this would continue in extrapolation and less reasonable that as time extends it would decrease. (The ERG concede that the linear trend is less clear for certain histological subgroups in REVEL, but think decreasing probability of death is implausible for these also)
- iii. The company adopts proportional hazards assumptions for both network meta-analysis and overall survival in the REVEL trial, but then generates loglog models which do not conform to proportional hazards. There is an obvious inconsistency of approach in this.

- iv. The company models survival for the comparator nintedanib in a different way to survival for the intervention (ramucirumab). For nintedanib, a hazard ratio from network meta-analysis is applied to the loglog model for the control arm of REVEL (derived using a single model for both REVEL arms) but for ramucirumab a meta-analysis HR is not applied to the control arm instead the loglog model for ramucirumab (with treatment as covariate) is used directly. The ERG do not think this is an even-handed approach.

2] Smoothed hazard plots

“The ERG did not appear to have tested this by plotting smoothed hazard rates. The plot based on the total sample in the REVEL trial is shown in Figure 3”. Smoothed hazard plots are available in statistical packages such as Stata. Multiple and different smoothed plots can be obtained from a given data depending on the settings set for smoothing. The ERG’s preferred plot in the circumstances was a the cumulative hazard plot of the observed data; no *arbitrary* selection of smoothing method is made and the hazard change through time can be seen from the shape of the plot, which in the case of REVEL became obviously *linear because hazard became constant after about 12 months*. The ERG linear trend models were not “visual fits” but *least squares linear regression fits* to the observed cumulative hazard data. The ERG acknowledge there is some arbitrariness in the selection of the time point from which the linear trend is fitted

Further considerations by NICE DSU regarding modelling approaches, including that adopted by the ERG, can be found in the following two publications which are additional to those referenced by Lilly:

Bagust A and Beale S. Survival Analysis and Extrapolation: Modeling of Time-to-Event Clinical Trial Data for Economic Evaluation. An Alternative Approach. Med Decis Making 2014;34: 343–351

Latimer NR. .Response to “Survival Analysis and Extrapolation Modeling of Time-to- Event Clinical Trial Data for Economic Evaluation: An Alternative Approach” by Bagust and Beale. Med Decis Making 2014;34: 279–282

Ramucirumab for previously treated locally advanced or metastatic non-small-cell lung cancer [ID838]

Yvonne Summers, one of the clinical experts who attended the first committee meeting, was invited back to the second committee meeting, to respond to some questions raised during the ACD consultation period, but was unable to attend. Therefore some clinical questions were sent to her ahead of today's meeting. These questions and Yvonne's responses are shown below.

Question	Response [sic]
<p>Would there ever be a circumstance where people with non-squamous NSCLC could not receive nintedanib plus docetaxel but could receive ramucirumab plus docetaxel? If so could these patients receive docetaxel alone even though they are unsuitable for nintedanib plus docetaxel?</p>	<p>I can't think of a good example except perhaps if a patient was unable to tolerate the large capsule (eg I once had a patient with previous treatment for head and neck cancer who had such a difficulty) The exclusion criteria for both studies were very similar (eg tumours invading blood vessels, uncontrolled hypertension).</p>
<p>During the consultation period for this appraisal we received a comment suggesting that a proportion of non-squamous patients would still be receiving docetaxel alone rather than nintedanib plus docetaxel. Is this a true reflection of current clinical practice?</p>	<p>Only a few patients are non squamous (NOS) , less than 10%. There are some patients who have contraindications to nintedanib, but these contraindications would apply to ramucirumab too and so these patients would receive docetaxel alone.</p>
<p>As nintedanib plus docetaxel is recommended for people with adenocarcinoma could there be a group of patients with unspecified non-squamous NSCLC who would receive docetaxel alone?</p>	<p>Yes but the proportion of patients in the non-squamous group would be small (less than 10%)</p>