National Institute for Health and Care Excellence

Centre for Health Technology Evaluation

Pro-forma Response

ERG report

Afatinib for treating epidermal growth factor receptor mutation positive locally advanced or metastatic non-small cell lung cancer

You are asked to check the ERG report from Liverpool Reviews and Implementation Group to ensure there are no factual inaccuracies contained within it.

If you do identify any factual inaccuracies you must inform NICE by **5pm** on **Wednesday 11 December 2013** using the below proforma comments table. All factual errors will be highlighted in a report and presented to the Appraisal Committee and will subsequently be published on the NICE website with the Evaluation report.

The proforma document should act as a method of detailing any inaccuracies found and how and why they should be corrected.

Issue 1 ERG considers that there is insufficient evidence to answer the decision problem

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
The ERG has chosen not to provide updated ICERs as they believe that there is insufficient evidence presented in the submission to answer the decision problem.	The ERG report should reflect that sufficient information has been provided in order to answer the decision problem and that: • These issues have been raised in previous STAs but have not been considered serious enough to prevent the Committees issuing positive guidance • The perceived issues relating to the differences in Asian and non-Asian patients and the assumption regarding proportions hazards lead to afatinib becoming *more* cost-effective, as indicated by the ERG's comparison between afatinib and pemetrexed/cisplatin on the final two pages of the report. Therefore sufficient information exists for the committee to conclude that afatinib in its licensed indication is a cost-effective treatment option for these patients.	Previous NICE committees have been able to provide guidance with similar levels of evidence. In fact, the level of evidence in this submission has exceeded those of erlotinib and gefitinib for which there is guidance in this indication. Despite the ERG's comments, there are aspect to the report which enable NICE to make a decision with the available evidence: 1. [redacted] If the analysis undertaken by the ERG is correct, then afatinib would be expected to be *more* cost-effective than those presented by Boehringer Ingelheim in the submission. In fact, the ERG shows in it's comparisons of afatinib vs pemetrexed/cisplatin that afatinib is more cost-effective in the non-Asian population than the whole population (pages 120-121). 2. Despite the ERGs perceived issues around the difference between Asian and non-Asian patients and the assumptions over proportional hazards being raised during previous appraisal of TKIs in this indication (as stated by the ERG), the NICE committee has been able to come to a decision and issue positive guidance in previous STAs. In these cases, the Committee did not feel that these issues would substantially alter the Guidance. Indeed, given that that these	This is a matter of opinion. No amendment to the ERG report is required.

		issues have been highlighted again support the rationale that there are unlikely to be significant different between the TKIs and supports the case proposed by the ERG for cost-minimisation. 3. The ERG has shown that afatinib is cost-effective vs pemetrexed/cisplatin. Given that NICE has found both gefitinib and erlotinib to be cost-effective vs pemetrexed/cisplatin, afatinib can be considered a treatment option alongside these. Given these points and the additional justification below, we not only feel that there is sufficient information to make a decision, but also that additional committee meetings are unlikely to add to the level of evidence required to make a decision.	

Issue 2 Asian Ethnicity may be a treatment effect modifier

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
On page 10 of the assessment report the ERG (Evidence Review Group) comments, "results of subgroup analyses of LUX-Lung 3 data demonstrate support for the hypothesis that being of Asian ethnicity may be a treatment effect modifier in patients who have activating EGFR mutations". In addition to this on Page 11 of the Assessment Report the ERG reports states, "The comparison of Asian vs non-Asian survival outcomes in LUX-Lung 3 demonstrate that ethnicity may be a treatment effect modifier in patients with activating EGFR mutations".	We do not believe that the ERG has demonstrated that there is a meaningful effect or that that this will change the final guidance from the Committee. We therefore propose that the following be added: • The interaction term between the Asian and non-Asian patients should be included in the report. This has been supplied to the ERG who has chosen to omit it from their report. Race as a stratification factor (Asian vs non-Asian) interaction between treatment and subgroup has a P-value of **0.6541.** • An Asain vs non-Asian subgroup analysis was pre-specified and was not found by any regulatory authorities to be a treatment modifier. Specifically, the EMA did not identify race as a treatment modifier. The EMA stated "The treatment effect of afatinib was similar in relevant subgroups defined by gender, age, race, geographical region, and the ECOG performance status at baseline." • The rationale for looking only at the Asian vs non-Asian subgroup should be provided as this appears to be inconsistent with the way other subgroups have been analysed in the ERG report. • Appropriate weight should be given to the fact that no statistical differences exist	Whilst we welcome debate over the applicability of data in the decision making process, we feel that in this case the data has been misrepresented. As indicated in the proposed amendment, crucial parts of the evidence, such as the interaction term, statistical analysis and goodness-of-fit statistics have been omitted from the report or placed in the Appendix. More concerning is why the Asian vs non-Asian subgroup has been selected given that there are a number of different subgroups with greater differences in the hazard ratios (see Fig. 15). These are discussed in the report with the conclusion that the findings were consistent with the primary analysis. No rationale is provided as to why the Asian vs non-Asian patients were selected for further analysis despite these differences also being consistent with the primary analysis. The EMA did not consider race to be a treatment effect modifier. The	There has been a longstanding debate as to whether outcomes for EGFR-positive Asian patients with NSCLC who are treated with a TKI differ from those for non-Asians. This debate arose from the observation that Asian patients may be considered genetically, physically, socially and epidemiologically different from non-Asian patients. This is the first time that a trial has used Asian vs non-Asian as a randomisation factor and the ERG therefore took the opportunity to carry out exploratory analyses. The ERG considers that the use of Cox proportional hazards comparisons is of questionable value as some Kaplan-Meier survival plots show evidence of non-proportionality.

	between these Asian and non-Asian patients in LUX-Lung 3. • The report should emphasise that the bias identified by the ERG was in favour of non-Asian patients whereas the effect from the other TKIs was in favour of the Asian patients, indicating there is no consistent treatment effect across TKIs. • The forest plots in the Appendix on page 104 should be given due prominence as these clearly show the there is no difference in the treatment effect between Asian and non-Asian patients. • Figures 1 to 4 selectively aggregate and disaggregate the data for the Asian and non-Asian patients leading to the appearance of a treatment modifying effect. A consistent approach should be taken when presenting this type of data, as well as the statistics around the goodness-of-fit, in order to allow the reader to make an informed decision on whether these differences are real or an artefact of the statistical analysis. **Finally, it should be made clear that the use of a non-Asian population *improves* cost-effectiveness; this is shown by the analysis undertaken by the ERG at the end of the Appendix.**	EMA's assessment report concludes "The treatment effect of afatinib was similar in relevant subgroups defined by gender, age, race, geographical region, and the ECOG performance status at baseline." (European Medicines Agency. Assessment Report for Giotrif. Procedure number EMEA/H/C/002280 2013). We do not think that the case has been presented to differentiate between the Asian and non-Asian patients and that if the Committee had a balanced presentation of the data then we do not believe that this analysis would be considered sufficient justification to change the decision making process. Given that the ERG have estimated that afatinib is more-cost-effective in the non-Asian patients, then this distinction will not change the decision problem. At best the analysis stands as it is, and at worse afatinib is more cost-effective. We would like this to be considered when the Committee are making their decsion.	

Issue 3 Assumption over proportional hazards

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
The ERG believes that the assumption over proportional hazards has been violated and therefore the results from the mixed treatment comparison (MTC) are invalid.	The ERG claims that the assumption for proportional hazard ratios has been violated. However, it should be made clear that: • The ERG has not provided any details of the goodness-of-fit for it's 2-stage proportional hazard model meaning that it is not possible to compare the single proportional hazard model to the ERG two-stage model. This omission means that the reader cannot make an informed choice over which is the best model. • This 2-stage proportional hazard model is inconsistently fitted to the trial data as it is variously aggregated or disaggregated by Asian or non-Asian patients. This means that the reader is unable to compare these models to the proportional hazard model. • Confidence intervals for the models are not provided which again makes comparison between models difficult. • Despite this issue arising in previous HTAs (as stated by the ERG), the Committee was able to produce guidance, and is also possible in this case. • Despite these issues, the ERG has managed to estimate the ICER for afatinib vs pemetrexed/cisplatin meaning that this	[redacted] We note that the goodness-of-fit of almost any model to data can be improved by adding more parameters but the ERG has omitted to supply the information crucial for judging which model is best. They have also avoided discussing the fact that the MTC fits the clinical data well. Therefore, the ERG have alluded to an issue but not demonstrated that this leads to misleading results and conclusions. Whilst we recognise that adding additional parameters to models can always improve the fit to the data, this does not always provide additional insight. The assumption of proportional hazard has been recognised by the regulatory authorities who have not identified this as an issue, and we also believe that in this case this is not an issue.	There are two separate issues here. Firstly, the construction of a MTC using hazard ratios depends on the assumption of proportional hazards being valid across all parts of the network. If this proves not to be the case with one or more component trials, the network may degenerate, fragment or become wholly disconnected depending on the number and location of invalid segments. On this basis the proposed MTCs appear to be vulnerable due to the inclusion of several trials with obvious violation of the proportional hazard assumptions evidenced by crossing Kaplan-Meier plots. Secondly, fitting of parametric functions to individual trial data is not a mechanistic process, and must take account of patterns of response and effect which are present in the trial data. Clinical trials of advanced cancer do not normally follow smooth trends (as seen in all 'standard' functions) due to the

	is possible. • A wide variety of sensitivity analysis has been undertaken to show that the MTC is robust, and replicating the analysis with a more complex model is unlikely to show that the TKIs differ in a meaningful way from each other.		nature of the trial which involves the use of new intervention(s) at a fixed time point designed to substantially alter outcome hazard for a period before losing effectiveness in the long term. The resulting temporal hazard changes are protocol driven and require more sophisticated analytical methods than generally used in epidemiological studies.

Issue 4 ERG conclusion: Adenocarcinoma populations with EGFR mutations

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
On page 11 of the assessment report the ERG concludes “Direct data from LUX-Lung 3 and LUX-Lung 6 can only be used to inform treatment decisions for patients who have adenocarcinoma and activating EGFR mutations in the first-line setting”.	Boehringer Ingelheim notes the conclusion of the ERG and that the LUX Lung 3 and LUX Lung 6 studies enrolled patients with EGFR mutation positive adenocarcinoma of the lung. However, the conclusion of the ERG, in the opinion of Boehringer Ingelheim is not consistent with the medically accepted importance of EGFR mutational status rather than histology status in informing treatment decisions with an EGFR TKI in first line. Boehringer Ingelheim would advise that the LUX Lung 3 and LUX Lung 6 data can be used to inform treatment decisions for patients who have EGFR mutation positive NSCLC regardless of histology.	The rationale for adenocarcinoma histology inclusion within these studies is due to the higher rate of EGFR mutations in patients with adenocarcinoma histology (Billah et al 2011). The EMA have not restricted the afatinib label to adenocarcinoma histology (European Medicines Agency. Assessment Report for Giotrif. Procedure number EMEA/H/C/002280 2013). Boehringer Ingelheim highlights the registration studies for gefitinib (Mok et al. 2009) and erlotinib (Rosell et al 2012) which had 95.4% and 95% adenocarcinoma histology respectively in the intervention treatment arms. Furthermore, Rosell et al 2012 report progression free survival subgroup analysis by histology with a non-significant interaction p value (p=0.8721)	This is a matter of opinion rather than a factual error. The extent to which trial results may be generalised to a population that is wider than that recruited to the trial is likely to be an area that the AC will wish to discuss. No amendment to the ERG report is required.

Issue 5 ERG conclusion; comparators not first line treatment in NSCLC with EGFR mutations

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
On page 11 of the assessment report the ERG concludes "Main comparators in the presented afatinib trials are not specifically recommended by NICE for the first-line treatment of patients with activating EGFR mutations in England and Wales"	Boehringer Ingelheim would highlight that the choice of comparator chemotherapy in LUX Lung 3 and LUX Lung 6 are appropriate given the licensed and available treatment options available at the time of study design. Therefore, despite these platinum doublets not being recommended first line treatment options in patients with EGFR mutation positive NSCLC (non-small cell lung cancer), they are appropriate comparators given their demonstrated efficacy in NSCLC. Cisplatin/gemcitabine the comparator in LUX-Lung 6, was due to the label restrictions of pemetrexed in China and hence differs from LUX Lung 3.	At the time of trial design for LUX Lung 3 and LUX Lung 6 oral EGFR TKIs were not the standard of care. Boehringer Ingelheim comments that the comparator was chosen as cisplatin/pemetrexed in LUX Lung 3 as this has been demonstrated to be an effective treatment in adenocarcinoma patients (Scagliotti et al. 2009) and is currently recommended in adenocarcinoma patients without EGFR mutations in the first line setting (NICE Technology appraisal guidance 181)	This is not a factual error, but an accurate statement based on the wording of marketing authorisations. No amendment to the ERG report is required.

Issue 6 Post progression treatment

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
Page 28 of the ERG report states; "with most patients in LUX-Lung 3 and a large minority in LUX-Lung 6 receiving chemotherapy following disease progression on afatinib (%% and %% respectively); for patients starting on chemotherapy, the majority of patients in both trials received a TKI following progression on chemotherapy (%% and %% respectively)."	This is incorrect and should read as follows: "The proportion of patients in LUX-Lung 3 and in LUX-Lung 6 receiving chemotherapy following disease progression on afatinib was %% and **%%** respectively; for patients starting on chemotherapy, the proportion of patients in both trials receiving a TKI following progression on chemotherapy was %% and **%%** respectively." (LUX Lung 3 Clinical Trial Report, 2012; LUX Lung 6 Clinical Trial Report, 2013).	The post progression cross-over has been incorrectly reported in the ERG assessment report. Overall survival in considered to be influenced by post progression treatment. (McCain JA, 2010)	The ERG acknowledges this transcription error and has corrected the text accordingly. Please note, in accordance with the manufacturer's proposed amendment, the data has been marked as commercial in confidence in the ERG's erratum. However, the same data were not marked as CiC in the MS.

Issue 7 Treatment related adverse event data

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
On page 114 of the assessment report the treatment related adverse event data for LUX Lung 3, 6 and 2 has been summarised in table 64. The adverse event data supplied for LUX Lung 6 within the manufacturer's submission is adverse event data regardless of relatedness to treatment.	The treatment related adverse event data with an incidence >10% by treatment arm has been supplied in Appendix A (LUX Lung 6 Clinical Trial Report, 2013)	The adverse event data supplied within the manufacturer's submission for LUX Lung 6 was not treatment related adverse events. The data supplied within the manufacturers submission was adverse events with an incidence >20%. Treatment related adverse event data for LUX Lung 6 is contained within the clinical trial report and has been supplied within this response.	The ERG did not receive the appendices to the LUX-Lung 6 trial report. Nevertheless, the correct data have been extracted from Table 12.2.2.2:1 of the LUX-Lung 6 CTR into Table 61 of the ERG report (page 114). The ERG also notes that the LUX-Lung 2 AE data presented in the MS and reported in the ERG report are also for all AEs, not drug-related AEs, and therefore these figures have also been revised (using data from the published paper by Yang et al 2012) Finally, the text that relates to Table 61 has also been revised (page 110 of the ERG report).

Issue 8 ERG states that the model does not reproduce PFS results

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
Page 9 states that the ERG did not find that the model could reproduce PFS results	Analyses to support this claim are not provided by the ERG. However, Table 124 in the main submission clearly shows that not only is the PFS observed in both LUX-Lung 3 and LUX-Lung 6 reproduced by the submitted economic model, but so is overall survival.	The ERG mistakenly made this claim in the summary but has not justified it in their report.	Data from the manufacturer's model, comprising monthly PFS Kaplan-Meier data and associated Weibull models, have been plotted by the ERG against detailed time of event data for the latest LUX-Lung 3 data cut analysed using an unbiased (non-informative) censoring rule. These data are displayed in Figure 14 of the ERG report (p80) and demonstrate large differences between the data used in the manufacturer's model and trial data. No amendment is required to the ERG report.

Issue 9 Derivation of utility values

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
Pg81 third paragraph "The reliability of the results presented by Chouaid et al is unclear"	This section needs to include details around the differences in methodology between the two studies. This should include details of the method of elicitation and details around the populations. This should be accompanied by a detailed explanation of which values are most appropriate. In addition, there should be a discussion around the impact of these values on the decision as currently this comment is not qualified.	What is absent is which set of utilities is more in line with those outlined in the NICE Methods of technology appraisals, in particular section 5.3. The Chouaid study uses the EQ-5D questionnaires in patients with the condition of interest and receiving treatment and then used the UK preference weights to calculate the utilities, in-line with the NICE reference case. This is in comparison to Nafees et al which used standard gamble techniques in members of the general public. These values are applicable when alternatives are not available. While arguments can be made for either of these studies there needs to be a detailed explanation of why the EQ-5D values were considered inappropriate.	The ERG maintains their view that the reliability of the Chouaid et al results are unclear. This is primarily due to the fact that the only available information is presented in a poster that has not been peer reviewed. As detailed in the ERG report, utility values are based on the views of small numbers of patients, particularly with regard to BSC where the authors of the poster point out that the number of patients was too small to warrant meaningful analysis. The manufacturer's sensitivity analysis (p22 of the PAS submission) shows that ICERs per QALY gained for afatinib vs gefitinib and afatinib vs erlotinib rise from £17,933 to £20,256 and from £10,079 to £10,588 respectively when the manufacturer's Nafees *et al* derived utiity values, rather than Chouaid *et al*, utility values are employed. However, the ERG notes that if different survival models were employed the use of different utility values might

			lead to substantial differences in resultant ICERs per QALY gained. No amendment is required to the ERG report.

Issue 10 Adverse event costs – grade 3/4 fatigue

Description of problem	Description of proposed amendment	Justification for amendment	ERG comment
Pg 84 Adverse event costs in relation to the cost used for grade 3/4 fatigue	This discussion requires additional detail on the significance of this issue on the Committee's decision making. Currently it does not provide any information on the magnitude of effect this could have on the cost effectiveness estimates.	The statement as written reflects a point of view rather than providing evidence that the value utilised is incorrect. The ERG is correct in stating that the definition of a grade 3/4 adverse event is one that may require hospitalisation. However, what occurs in clinical practice is often very different. The truth is probably somewhere in between the two estimates. However, it should also be stated what the rates of grade 3/4 fatigue are so that the Committee could evaluate the potential impact of this.	The ERG notes the manufacturer's point of view and appreciates that the manufacturer recognises that the ERG's understanding of the definition of a grade 3/4 adverse event is accurate. The ERG is of the view that any change to the cost associated with grade 3/4 fatigue will only have a small effect on any ICER per QALY gained estimates. No amendment to the ERG report is required.

Additional errors identified by the ERG and corrected in the erratum

Pages 41 and 42, Tables 14 and 15, should read ‘difference in median PFS’ and not ‘difference in mean PFS’

Page 57 – ‘**Error! Reference source not found.Error! Reference source not found.**’ Should read ‘Table 21’