Eribulin for the treatment of locally advanced or metastatic breast cancer

ADDENDUM

This report was commissioned by the NIHR HTA Programme as project number 10/19

Completed 13th September, 2011

CONTAINS IN CONFIDENCE DATA
Title: Eribulin for the treatment of locally advanced or metastatic breast cancer

Produced by: Liverpool Reviews and Implementation Group (LRiG)

Authors:
- Adrian Bagust, Professor of Modelling in Health, Liverpool Reviews and Implementation Group, University of Liverpool
- Angela Boland, Associate Director, Liverpool Reviews and Implementation Group, University of Liverpool
- Helen Davis, Assistant Director, North West Medicines Information Centre, Pharmacy Practice Unit, Liverpool
- Yenal Dundar, Research Fellow, Liverpool Reviews and Implementation Group, University of Liverpool
- Nigel Fleeman, Research Fellow (Clinical Effectiveness), Liverpool Reviews and Implementation Group, University of Liverpool
- Janette Greenhalgh, Research Fellow (Clinical Effectiveness), Liverpool Reviews and Implementation Group, University of Liverpool
- Susan O’Reilly, Consultant in Medical Oncology, Clatterbridge Centre for Oncology, Wirral
- James Oyee, Research Fellow (Medical Statistician), Liverpool Reviews and Implementation Group, University of Liverpool
- Nicola Trevor, Health Economist, BMJ Technology Assessment Group, London

Correspondence to: Dr Janette Greenhalgh, Research Fellow, Liverpool Reviews and Implementation Group, University of Liverpool, Room 2.2, Whelan Building, The Quadrangle, Brownlow Hill, Liverpool L69 3GB

Date completed: 24th May 2011

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 10/19

Declared competing interests of the authors: None
Table of contents

1 BACKGROUND ........................................................................................................................................ 4
2 SUMMARY OF SUPPLEMENTARY EVIDENCE .................................................................................. 5
  2.1 Submission overview ......................................................................................................................... 5
  2.2 New clinical evidence ......................................................................................................................... 5
  2.3 New economic evidence .................................................................................................................... 5
3 RE-DEFINING THE DECISION PROBLEM ....................................................................................... 7
  3.1 New basis for evidence synthesis ...................................................................................................... 7
  3.2 What base-case scenario should be used? ........................................................................................ 7
4 FURTHER ANALYSIS UNDERTAKEN BY ERG ........................................................................... 9
  4.1 Implementation of amendments and corrections previously identified ......................................... 9
  4.2 New survival analysis ....................................................................................................................... 10
  4.3 ERG revised cost-effectiveness results ............................................................................................ 15
5 SUMMARY ............................................................................................................................................ 17
6 REFERENCES ......................................................................................................................................... 18
7 Appendix ............................................................................................................................................... 19
1 BACKGROUND

On 23rd June 2011, the National Institute for Health and Clinical Excellence (NICE) Appraisal Committee (AC) considered the evidence for use of eribulin (Halaven®) as a treatment for patients with locally advanced or metastatic breast cancer (LABC/MBC) who have received two or more chemotherapy (CTX) regimens. On 12th July 2011 NICE issued its Appraisal Consultation Document (ACD) which stated that "eribulin is not recommended for the treatment of locally advanced or metastatic breast cancer in people whose disease has progressed after at least two chemotherapeutic regimens for advanced disease".¹

The manufacturer of eribulin (Eisai) made a successful request to NICE to submit new evidence at the ACD consultation stage for the appraisal of eribulin for the treatment of LABC/MBC, as permitted by the Single Technology Appraisal Process Guide.²

This document summarises an assessment by the Evidence Review Group (ERG) of the supplementary evidence submission provided by Eisai.
2 SUMMARY OF SUPPLEMENTARY EVIDENCE

2.1 Submission overview

The manufacturer’s supplementary evidence submission included an appendix to the initial submission document together with an updated version of the EXCEL cost-effectiveness model.

The main focus of the new evidence submitted involves the restriction of the population considered for treatment to those patients previously treated with capecitabine, who constitute 74% of all patients included in the EMBRACE\(^3\) clinical trial. The manufacturer argues that this is consistent with current NICE guidelines,\(^4\) especially if vinorelbine is used as the primary comparator.

The main body of the supplementary evidence submission is taken up with discussions of different approaches to estimating survival outcome benefits within the economic model, and the differing results obtained with alternative assumptions. The objectives of these analyses are two-fold: to indicate the range of incremental cost-effectiveness ratios (ICERs) which can be generated under different scenarios, and also to establish the likely range of estimated gain in overall survival (OS) which can be attributed to use of eribulin, for use in the consideration of the NICE 'end of life' criteria.\(^5\)

2.2 New clinical evidence

The manufacturer provides information for the post-capecitabine subgroup of EMBRACE\(^3\) for median OS (the primary outcome of the trial), showing a significant improvement of 2.9 months (hazard ratio\([HR]=0.787\))

In addition, the manufacturer includes details of a meta-analysis of other clinical studies in which the median OS attributable to eribulin (4.33 months) is comparable to that indicated for eribulin vs vinorelbine in the EMBRACE\(^3\) intention-to-treat (ITT) population (4.2 months), despite the small number of patients receiving vinorelbine in the EMBRACE\(^3\) trial.

2.3 New economic evidence

Data inclusion

The manufacturer presents arguments against the employment of data from the whole EMBRACE\(^3\) trial, preferring to use only Region 1 data (North America and Europe) arguing that data from Regions 2 (Eastern Europe, Russia and Turkey) and 3 (Latin America and South Africa) are less mature due to patient enrolment starting later, and that this leads to a bias which may underestimate the survival benefit from use of eribulin.
Projective modelling

The manufacturer presents arguments for earlier use of projective modelling (from 35% of patients remaining alive), rather than the ERG’s approach which minimises the influence of projective modelling and maximises use of the unadjusted trial data. The manufacturer presents arguments to support alternative forms of projective modelling, and also argues against the simple exponential form employed by the ERG. In addition, the manufacturer indicates their preference for a proportional hazard approach to survival modelling, suggesting that this provides more stable results.

Model results

The manufacturer provides summary results for:

- 24 different OS modelling scenarios involving different populations, different comparators and different survival projection methods
- 21 different cost-effectiveness analysis model scenarios
- 9 different cost-effectiveness analysis model scenarios adjusted for increased utility values to represent ‘end of life’ thresholds
- a sensitivity analysis using the costs of intravenous (rather than oral) vinorelbine
- two probabilistic sensitivity analyses.

The ICERs reported range from £63,761 - £23,790 per QALY gained. The manufacturer’s preferred scenarios reflect three proportional hazard projection models using different functional forms (log-logistic, exponential and log-normal) and are restricted to patients from Region 1 who had been planned to receive vinorelbine, with ICERs between £23,790 and £26,475 per QALY gained (Tables 28-30 in the Eisai document "Appendix: Additional evidence in response to the Appraisal Consultation Document (ACD)").
3 RE-DEFINING THE DECISION PROBLEM

3.1 New basis for evidence synthesis

Although the manufacturer’s supplementary evidence submission does not explicitly redefine the decision problem for the appraisal it is clear from the first paragraph of the document, that a new basis is being proposed.

Whereas the initial manufacturer’s submission imposed no limitation related to prior therapy, it is clear that the AC is being asked to consider only the use of eribulin in patients who have relapsed after previous treatment with capecitabine. In addition, the manufacturer has modified the comparator technology from "treatment of physician's choice" (TPC) to vinorelbine, on the basis that this more closely matches the stage within the current NICE guidelines at which it is envisaged that eribulin may be used.

The ERG considers that both these alterations are reasonable and realistic in the UK context where capecitabine is widely used for LABC/MBC patients. However, there are likely to be practical implications for the analyses required to reflect these changes in order to rework estimates of clinical benefit and cost effectiveness. The reduction of the overall volume of admissible trial data by a quarter is likely to increase uncertainty in analytical results. In addition, the restriction of the comparator to less than 25% of the overall trial population suggests that it may prove impossible to obtain meaningful results for comparisons of clinical effectiveness, or as a basis for projective modelling of outcomes.

3.2 What base-case scenario should be used?

In line with the manufacturer's revised submission, it is appropriate that patients who had not received prior treatment with capecitabine should be excluded from consideration. In addition, the model calibration should initially be based on patients intended for treatment with vinorelbine prior to randomisation to eribulin or vinorelbine.

The manufacturer argues that data from Region 1 only should be used in the analysis, on the basis that centres in Regions 2 and 3 began recruiting at a later date, and that the greater degree of censoring would bias survival estimates and HRs. This is a curious supposition since techniques such as Kaplan-Meier and Cox regression analysis are specifically designed to take account of differing proportions of censoring within data sets. However, to test this hypothesis the ERG have applied Cox regression analysis to the OS data from the EMBRACE study, restricted to patients with previous experience of capecitabine therapy. The explanatory variables included in the analysis are the trial arm (eribulin or TPC), the recorded “Best Response to therapy”, as well as the two remaining randomisation variables (HER2 status and Region). Using a step-wise procedure, the first variable
entered into the model was “Best Response to therapy” (p < 0.001). No more variables achieved the 5% significance level required, indicating that neither treatment, HER2 status nor geographical region could contribute significant additional explanatory power. When the analysis was repeated omitting the “Best Response to therapy” variable, only the trial arm (eribulin vs TPC) proved to be significant. Thus it appears that these two factors are sufficient to account for the observed results, and there are no grounds for distinguishing between subgroups of patients on the basis of either HER2 status or geographic region. The ERG therefore considers it is appropriate to use suitable trial data from all regions in calibrating the economic model.

On this basis, the newly submitted cost-effectiveness scenario results were examined to identify which mostly match these criteria. Amongst the scenarios using Kaplan-Meier values combined with long-term projective modelling, the closest equivalent yields the results shown in Table 16 of the re-submission with an ICER of £38,005. There are no equivalent scenarios based on proportional hazards (PH) modelling, since it was noted by the manufacturer that:

"Unfortunately, no post capecitabine analysis (vs. vinorelbine all regions) was ready in time for this submission. However as mentioned 95% of patients in Region 1 had received prior capecitabine and the committee have confidence that they are unlikely to be greatly different from the Region 1 outcomes."

It should be noted that the PH scenarios using data from all regions (Tables 25-27) yield greater ICERs than those using only Region 1 data (Tables 28-30): £35,242-£41,480 per QALY gained compared to £23,790-£26,475 per QALY gained. This indicates that using the data from all regions is likely to result in higher ICERs irrespective of whether patients not previously treated with capecitabine are excluded. In order to assess the relative impact of various changes made by the ERG to the economic model, the scenario represented by the manufacturer's Table 16 is used below as the starting point for comparison.
4 FURTHER ANALYSIS UNDERTAKEN BY ERG

4.1 Implementation of amendments and corrections previously identified

The original ERG report identified a series of logic errors and amendments and recommended that these should be applied to the manufacturer’s model (Table 33). In each case the ERG has examined the re-submitted model to determine whether these changes have been correctly implemented by the manufacturer. In several cases it was found that changes had not been incorporated; in response, the ERG has applied the necessary alterations to the model to ensure that the impact of these factors can be properly understood. The results are shown in Table 1.

Table 1 ERG revisions to re-submitted cost-effectiveness model results

<table>
<thead>
<tr>
<th>Scenario / changes</th>
<th>Eribulin</th>
<th>Vinorelbine</th>
<th>Incremental</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cost per patient</td>
<td>QALYs per patient</td>
<td>Cost per patient</td>
</tr>
<tr>
<td>Table 16*</td>
<td>£20,673</td>
<td>0.7530</td>
<td>£13,253</td>
</tr>
<tr>
<td>+ discounting logic</td>
<td>£20,920</td>
<td>0.7640</td>
<td>£13,405</td>
</tr>
<tr>
<td>+ terminal period logic</td>
<td>£22,585</td>
<td>0.7302</td>
<td>£15,190</td>
</tr>
<tr>
<td>+ mid-cycle logic</td>
<td>£20,274</td>
<td>0.7320</td>
<td>£13,165</td>
</tr>
<tr>
<td>+ IV vinorelbine cost</td>
<td>£20,673</td>
<td>0.7530</td>
<td>£12,769</td>
</tr>
<tr>
<td>+ febrile neutropenia</td>
<td>£20,744</td>
<td>0.7529</td>
<td>£13,279</td>
</tr>
<tr>
<td>Table 16 revised*</td>
<td>£22,522</td>
<td>0.7311</td>
<td>£14,812</td>
</tr>
</tbody>
</table>

*Table 16 refers to manufacturer’s revised submission

The net effect of applying all these modifications is very small, indicating that none of these concerns is likely to prove influential for decision-making.
4.2 New survival analysis

The major alterations made to the decision problem have the potential to lead to different methods in the estimation of patient outcomes as the relevant data subset is likely to exhibit reduced heterogeneity. In particular, omission of patients not previously treated with capecitabine and restriction of patients to those initially identified for treatment with vinorelbine result in reduced patient numbers, but may give important insights into the manner by which patient benefit accrues over time.

Examination of the OS Kaplan-Meier plot for this patient subgroup (Figure 1) is suggestive that the survival experience of patients receiving vinorelbine and eribulin may converge after about 2 years. However, the small number of patients in the comparison means that this pattern could easily have arisen by chance. This is an important matter to resolve, since if the convergence can be confirmed from the trial evidence this would imply that an accurate estimate of survival gain could be obtained directly from the Kaplan-Meier analysis without need for any parametric projective modelling. Moreover, the magnitude of the estimated gain is likely to be considerably smaller than that obtained by accumulating additional benefit indefinitely (equivalent to the area of the gap between the two projective curves shown in Error! Reference source not found.).

Figure 1
In the ERG report it can be observed (Table 31 and Figure 3) that the vinorelbine, gemcitabine and capecitabine subgroups exhibited similar mean OS and statistically significant survival gain. Although capecitabine is now excluded from consideration in this analysis, since all patients have received capecitabine previously, there remains the possibility of augmenting the vinorelbine data with additional gemcitabine patients if it can be shown that patient outcomes (OS and progression-free survival [PFS]) are sufficiently similar.

This hypothesis was tested by the log-rank test, and examination of the Kaplan-Meier plots. Error! Reference source not found. and Error! Reference source not found. illustrate the close correspondence between the vinorelbine and gemcitabine planned treatment cohorts for both arms of the clinical trial for OS.

Figure 2
Similar results were obtained for investigator PFS. As a consequence it was deemed appropriate by the ERG to pool the vinorelbine and gemcitabine cohorts to obtain more robust survival results with an 80% increase in patient numbers available for analysis.

Error! Reference source not found. and Error! Reference source not found. confirm the suspected convergence of eribulin and TPC trial arms when the larger pooled data set is analysed.
Figure 4

Figure 5
Table 2 shows the results of estimating the survival gain attributable to eribulin compared to TPC for both the planned vinorelbine subgroup, and the pooled vinorelbine and gemcitabine subgroups. The latter shows slightly larger estimated gains in OS (99 vs 85 days), and narrower confidence intervals as expected. In both cases the additional survival is split evenly between the pre- and post-progression periods.

These estimates of extended OS contrast strongly with modelled estimates from the manufacturer’s Table 16 scenario (4.5 months) and preferred proportional hazards scenarios in Tables 28-30 (over 7 months).

Table 2 Kaplan-Meier survival estimates using the Area Under Curve (AUC) difference until survival curves converge

<table>
<thead>
<tr>
<th>Upper limit (days)</th>
<th>Mean survival (days)</th>
<th>Survival gain (days)</th>
<th>Survival gain (months)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TPC</td>
<td>Eribulin</td>
<td>Mean</td>
</tr>
<tr>
<td>Vinorelbine subset</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PFS</td>
<td>230</td>
<td>76</td>
<td>118</td>
</tr>
<tr>
<td>OS</td>
<td>766</td>
<td>333</td>
<td>418</td>
</tr>
<tr>
<td>PPS</td>
<td>N/A</td>
<td>257</td>
<td>299</td>
</tr>
<tr>
<td>Vinorelbine+gemcitabine pooled subsets</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PFS</td>
<td>381</td>
<td>94</td>
<td>141</td>
</tr>
<tr>
<td>OS</td>
<td>827</td>
<td>341</td>
<td>440</td>
</tr>
<tr>
<td>PPS</td>
<td>N/A</td>
<td>247</td>
<td>299</td>
</tr>
</tbody>
</table>

N/A=not applicable; PFS=investigator assessment PFS; TPC=treatment of physician choice; CI=confidence interval; PPS=post progression survival (estimated as the difference between OS and PFS)
4.3 ERG revised cost-effectiveness results

Reference case utility values

It was noted in the first AC meeting that the utility values used by the manufacturer did not match the requirements of the NICE reference case. Subsequent to the meeting, Simon Dixon, a committee member unable to be present at the meeting, wrote to NICE staff drawing attention to the evidence previously presented at an earlier appraisal (see copy of his letter in Appendix 1). This used a value of 0.69 obtained from EQ-5D data collected in the EFG100151 clinical trial for pre-progression survival. Since this type of data would normally be considered more reliable than either figures obtained by standard gamble responses by non-patients, or inferred by mapping from a quality of life instrument, it is suggested that it should be used in place of the manufacturer's parameter value in the submitted model. The impact of this change is relatively minor. When applied to the scenario shown in Table 1, the ICER increases from £38,005 to £38,408 per QALY gained. When combined with all the other modifications in Table 1, the overall revised ICER increases from £38,737 to £39,137 per QALY gained. This additional amendment is included in the results shown below.

ERG survival estimates

In order to recalculate the impact of the Kaplan-Meier survival gains estimated by the ERG (Table 2), two sets of Kaplan-Meier have been added to the manufacturer's resubmitted model as alternatives to the original OS and PFS survival data for the vinorelbine subgroup and the pooled vinorelbine and gemcitabine subgroups.

Since there are insufficient patient data beyond the convergence point in each of the survival curves to allow any meaningful projection beyond this point, the Kaplan-Meier are terminated at that point (equivalent to all remaining patients dying at that time). This underestimates survival in both arms of the comparison by exactly the same amount so has no effect on either incremental costs or patient outcomes and so does not bias the calculation of the ICER in any way.

ERG revised cost-effectiveness estimates

Table 3 summarises the effect of these changes to the cost-effectiveness results generated by the modified decision model. The comparator in all cases is vinorelbine with its associated treatment costs and adverse event profile. The choice between using Kaplan-Meier survival estimates based on the small subgroup of patients with vinorelbine as the intended management or using the extended subgroup including also those intended for gemcitabine treatment has only a minor effect on the size of the estimated ICER (£53,538 vs £53,446 respectively). On the basis of minimising uncertainty (i.e. maximising the data set used for analysis) the ERG is inclined to prefer the use of the pooled subgroups.
Table 3 ERG revised cost-effectiveness model results including EQ-5D utility value and convergent Kaplan-Meier survival estimates

<table>
<thead>
<tr>
<th>Scenario/changes</th>
<th>Eribulin</th>
<th>Vinorelbine</th>
<th>Incremental</th>
<th>ICER</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cost per patient</td>
<td>QALYs per patient</td>
<td>Cost per patient</td>
<td>QALYs per patient</td>
</tr>
<tr>
<td>MS Table 16*revised</td>
<td>£22,522</td>
<td>0.7311</td>
<td>£14,812</td>
<td>0.5321</td>
</tr>
<tr>
<td>+ EQ-5D utility value</td>
<td>£22,522</td>
<td>0.7257</td>
<td>£14,812</td>
<td>0.5287</td>
</tr>
<tr>
<td>Vinorelbine subgroup survival</td>
<td>£20,633</td>
<td>0.6050</td>
<td>£13,623</td>
<td>0.4740</td>
</tr>
<tr>
<td>Vinorelbine/gemcitabine subgroup survival</td>
<td>£22,902</td>
<td>0.6455</td>
<td>£14,719</td>
<td>0.4924</td>
</tr>
</tbody>
</table>

*Table 16 refers to manufacturer's revised submission

N.B. Cost and QALY totals using Kaplan-Meier sub-group survival estimates are not directly comparable with other scenarios due to truncation of model at point of convergence. However, incremental cost and QALY estimates and ICERs are directly comparable with all other scenarios.
5 SUMMARY

The ERG welcomes the manufacturer's reorientation of their submission to restrict use of eribulin to those patients previously treated with capecitabine, rather than an unspecific comparison with any other treatment that a physician may choose. This appears to locate eribulin more clearly within the context of current NICE guidelines for the treatment of breast cancer. However, the ERG considers that this restriction is likely to alter the characteristics of the data set selected for analysis from the pivotal clinical trial, with the prospect of eliminating at least some of the evident heterogeneity which makes projective modelling uncertain and contentious.

The ERG's own analysis of the data led to two important conclusions:

- that there is no basis for excluding any records from the data set on the basis of the location of trial centre (no regional bias);

- that it is likely that the net outcome benefits attributable to use of eribulin in terms of PFS and OS are limited to a specific time period from randomisation, and do not extend indefinitely.

The main consequence of these observations is that the most reliable estimates of benefit are obtained directly from the non-parametric Kaplan-Meier analysis of the trial data, obviating the need for any parametric projective modelling (as discussed at length in the manufacturer's resubmission) and avoiding the need for any regional exclusions of trial records.

On the basis of this much simpler approach to estimating cost effectiveness, the ERG has concluded that the most reliable estimated ICER for eribulin compared to vinorelbine in treating patients previously treated with capecitabine is £53,446 per QALY gained.
6 REFERENCES


Robert

I was unable to attend the Eribulin meeting, but wish to make a comment on the ACD. Is this best done through the public consultation process or at the next meeting? If use the consultation process, you could ask the manufacturer (or LRiG) to do the necessary analyses. My concern is that the manufacturer is now doing further analyses based on (what I suggest to be) the wrong estimates of QoL - it would be better if they did the analyses on the correct estimates. FYI, my comment is below:

Within the appraisal, the utility values used are based on those of Lloyd et al., which are non-reference case. However, in the absence of more appropriate evidence these have necessarily been accepted. The utility values used by the manufacturer are 0.715 for stable disease and 0.79 for response, which were revised upward by the ERG to 0.756 and 0.823, respectively.

However, reference case utility values for a similar patient population are available. In the suspended NICE appraisal of lapatinib for HER2 over-expressing breast cancer, the manufacturer provided EQ-5D utilities from the pivotal trial on which the appraisal was based (EFG100151). The pre-progression utility used by the manufacturer was 0.69, which is in essence an weighted average of stable and responding disease. This figure is considerably lower than that used in the eribulin appraisal and suggests a possible bias within the estimates of cost-effectiveness.

The key question then becomes, how similar are the patient populations in the respective appraisals?

The patient population for EFG100151 is advanced or metastatic HER2 over-expressing breast cancer who have received two prior therapies, whilst for EMBRACE it is locally advanced or metastatic breast cancer who have received at least two prior therapies.

The median age in EFG100151 is 52. The median age in EMBRACE is 56.

ECOG 0 at baseline in the lapatinib arm of EFG100151 is 58% and 43% in the eribulin arm of EMBRACE.

Median overall survival is 67.7 weeks for lapatinib (=474 days) in EFG100151 and 399 days for eribulin in EMBRACE.

All figures are taken from the publicly available ERG reports and/or manufacturer submissions.

This suggests that the reference case estimate of pre-progression utility from EFG100151 is suitable for the eribulin appraisal, and as such, is preferred to the figures employed by the manufacturer and the ERG.

Thanks, Simon.