Single Technology Appraisal: Bevacizumab in combination with a taxane for the first-line treatment of HER-2 negative breast cancer

Addendum to Evidence Review Group's Report following Roche's new evidence submission of 24th September 2010

Produced by	CRD and CHE Technology Assessment Group, University of York, Heslington, York, YO10 5DD
Authors	Mark Rodgers, Research Fellow, CRD Marta Soares, Research Fellow, CHE David Epstein, Honorary Visiting Research Fellow, CHE Huiqin Yang, Research Fellow, CRD Alison Eastwood, Senior Research Fellow, CRD
Date completed	4th October 2010

Contents

1. Introduction

The Evidence Review Group (ERG) was requested by NICE to provide commentary and validity checks on the additional evidence submitted by the manufacturer.[1] It should be recognised that the work undertaken by the ERG does not constitute a full critique of the manufacturer's additional submission and does not accord with the procedures and templates applied to the original submission due to the limited time available to review the additional submission. However, a number of detailed checks were undertaken to ensure the validity of the manufacturer's revised analyses based on the additional data and revised set of assumptions employed by the manufacturer.

2. Clinical effectiveness

Changes from original submission

In terms of clinical effectiveness data, the new submission:

1. Incorporates AVADO and RIBBON-1 phase III RCTs alongside the E2100 trial data presented in the original submission;
2. Focuses on two subgroups – patients with 'triple negative' disease and those who have received previous adjuvant taxane therapy.

1. Incorporates AVADO and RIBBON-1 phase III RCTs alongside the E2100 trial data presented in the original submission

RIBBON-1 was excluded from the original submission on the basis of having an insufficient sample size and AVADO was excluded for not representing routine clinical practice on the basis of the docetaxel dose. However, in the new submission, data from both trials were combined with data from the E2100 trial in an individual patient data (IPD) meta-analysis.

The IPD meta-analysis combining data from all three trials indicated a statistically significant improvement in progression free survival (PFS) for the addition of bevacizumab to chemotherapy (HR 0.64, 95% CI 0.58 to 0.71). Pooled 1-year survival rates favoured bevacizumab treated groups (76.5% vs. 81.6%, p=0.003) but no benefits were found in terms of overall survival (OS; HR 0.97, 95% CI 0.86 to 1.08). These findings were similar to those from the E2100 trial that were presented in the manufacturer's original submission (see Table 1 below).

2. Focuses on two subgroups – patients with 'triple negative' disease and those who have received previous adjuvant taxane therapy

The manufacturer stated that two subgroups of patients – those with 'triple negative' disease and those who have received previous adjuvant taxane therapy – are likely to have poorer prognosis and may gain greater benefit from bevacizumab therapy than the broader ITT[i]

[i] In the manufacturer's additional submission, 'intention to treat' (ITT) refers to the overall population of all randomised patients.

population. As the existing trials were not originally designed to examine these subgroups, the manufacturer presented a number of exploratory analyses (see Table 1).

A subgroup meta-analysis of 621 patients with 'triple-negative' disease was similar to the ITT meta-analysis in terms of both PFS (HR 0.63, 95% CI 0.52 to 0.76) and OS (HR 0.96; 95% CI 0.79 to 1.16).

An exploratory subgroup meta-analysis of 311 patients with 'prior-taxane treated' disease (excluding the capecitabine ± bevacizumab groups from RIBBON-1) was slightly more favourable to bevacizumab than the ITT analysis in terms of PFS (HR 0.47, 95% CI 0.35 to 0.62). This was the only meta-analysis which significantly favoured taxane plus bevacizumab over taxane alone in terms of OS (HR 0.73; 95% CI 0.55 to 0.97). Median OS among patients receiving taxanes alone was 21.3 months, compared with 26.9 months among those receiving taxane plus bevacizumab. The manufacturer did not report 1-year survival data for this subgroup.

Table 1: PFS, OS and 1-year survival outcomes presented in the manufacturer's additional submission

	Overall PFS HR (95% CI)	**Overall OS HR (95% CI)**	**1-year survival (placebo vs Bev)**
ITT population			
E2100 (n=722)	0.54 (0.44, 0.67)† 0.48 (0.39, 0.61)‡	0.87 (0.72, 1.05)	74% vs. 81.4%, p=0.017
AVADO (n=488)*	0.77 (0.64, 0.93)† 0.67 (0.54, 0.83)‡	1.03 (0.70, 1.33)	76% vs. 84%, p=0.02
RIBBON-1** Taxane/anthracycline (n=622) Capecitabine (n=615)	 0.64 (0.52, 0.80) 0.69 (0.56, 0.84)	 1.03 (0.77, 1.38) 0.85 (0.63, 1.14)	 NR NR
Meta-analysis (n=2447)	0.64 (0.58, 0.71)	0.97 (0.86, 1.08)	76.5% vs. 81.6%, p=0.003
Triple-negative subgroup			
E2100 (n=232)	0.49 (0.34, 0.70)†	0.89 (0.66, 1.19)	NR
AVADO (n=111)*	0.68 (0.46, 0.99)†	0.82 (0.51, 1.32)	NR
RIBBON-1 Taxane/anthracyclines (n=NR) Capecitabine (n=NR)	 NR NR	 NR NR	 NR NR
Meta-analysis (n=621)	0.63 (0.52, 0.76)	0.96 (0.79, 1.16)	NR
Prior taxane subgroup			
E2100 (n=140)	0.33 (0.20, 0.54)†	0.67 (0.45, 0.99)	NR
AVADO (n=78)*	0.53 (0.33, 0.85)†	0.58 (0.31, 1.08)	NR
RIBBON-1 Taxane/anthracyclines (n=NR) Capecitabine (n=245)	 NR [redacted]	 NR [redacted]	 NR NR
Meta-analysis (n=311)	0.47 (0.35, 0.62)† 0.53 (NR, NR)‡	0.73 (0.55, 0.97)	NR

† Unstratified
‡ Stratified
*Data refer to 15mg/kg bevacizumab versus placebo comparison
**Data taken from Robert et al[2]. Manufacturer's submission reports slightly different values for PFS and does not report OS values.
NR = not reported
Commercial in confidence data
Value inconsistently reported between different sources. Data from tables in the additional submission, unless otherwise stated

ERG commentary

It can be seen that the PFS benefit of bevacizumab in combination with a taxane was less pronounced in the AVADO and RIBBON-1 trials than in E2100, with any trend towards an OS benefit eliminated (Table 1). The IPD meta-analyses of PFS and OS findings are similar to those presented from E2100 in the original submission, though slightly (statistically non-significantly) less favourable to bevacizumab.

The authors stated that overall OS values may have been confounded due to placebo-group patients being allowed to cross over to bevacizumab treatment. However, as crossover data were not collected in E2100, were unreported for RIBBON-1 and were only partially reported for the AVADO trial, it was not possible to establish the potential impact of any such confounding.

It should also be noted that 25% (615/2447) of all patients included in the ITT meta-analysis received bevacizumab in combination with capecitabine, a combination that lies outside the scope of the current STA. The manufacturers did not present an ITT meta-analysis limited to those patient groups receiving taxanes ± bevacizumab.

The manufacturer stated that patients with triple-negative disease may gain greater benefit from bevacizumab therapy than the ITT population. However, PFS and OS hazard ratios for the ITT population and triple-negative subgroup were almost identical (Table 1). Unlike the ITT population, the manufacturer did not report median PFS, median OS or 1-year survival data for this subgroup. Neither were these data or the relevant subgroup hazard ratios reported for the RIBBON-1 study. Since the analysis used individual patient data, it is unclear why these values would not be available.

Hazard ratios for PFS and OS from the second subgroup meta-analysis of 'prior taxane treated' patients did show a trend towards being (statistically nonsignificantly) more favourable towards bevacizumab than the ITT meta-analysis. This is the only meta-analysis in which a statistically significant overall OS benefit was observed for bevacizumab. However, this was an exploratory meta-analysis that included only a small number of patients (n=311 in total; n=78 taxane-pretreated patients receiving standard-dose bevacizumab in AVADO) from trials with subgroups insufficiently powered to detect a difference in OS, of which only one (AVADO) appeared to stratify for taxane pretreatment. The author's conclusion that "There is no clear scientific rationale for the observed additional benefit of combining bevacizumab with a taxane in taxane-pretreated patients"[3] confirms the ERG's view that this analysis can only be considered adequate for hypothesis generation, and not as convincing evidence of a subgroup effect.

Despite these limitations, the manufacturers presented the subgroup meta-analyses as evidence of a subgroup effect for the addition of bevacizumab to taxanes in patients with triple-negative disease or prior taxane therapy. However, rather than use estimates from these subgroup meta-analyses in the subsequent economic model, only the subgroup data from the E2100 trial were used. The ERG considers this to be an inconsistency in the manufacturer's approach. If the manufacturer's assumption of equivalent effectiveness for 3-weekly docetaxel and weekly paclitaxel (as stated on p.27 of the additional submission) is valid, combined data on effectiveness from all trial taxane arms should be considered

eligible. The PFS and OS values for E2100 are more favourable towards bevacizumab than the IPD data combined across all the relevant comparisons, and may subsequently yield more favourable ICER estimates (see section 3, below).

In conclusion, the ERG considers the submitted meta-analyses incorporating two previously excluded trials to be adequate for hypothesis generation and may form the basis of future investigations. However, given the issues around selection of data, inconsistent reporting of outcomes and the exploratory nature of the subgroup analyses, the current submission fails to provide clear and convincing evidence of the proposed subgroup effects.

3. Cost-effectiveness

Changes from original submission

In terms of cost effectiveness data, the additional submission:

1. Reports cost-effectiveness estimates for two new subgroups – patients with 'triple negative' disease and those who have received previous adjuvant taxane therapy
2. Uses an alternative approach to modeling overall survival;
3. Excludes the 10g capping scheme for the cost of bevacizumab from the base-case analysis in the revised model.

1. Reports cost-effectiveness estimates for the two new subgroups

Tables 2 and 3 summarise the main cost-effectiveness results presented by the manufacturer using PASA and list prices for paclitaxel.

Table 2: Cost-effectiveness results for prior taxane treated patients

	BEV+PAC	PAC	DOC
Mean Life Years (yrs)	2.624	1.969	1.969
Mean QALYs	1.559	1.058	1.057
Mean Total Cost (PASA)	£51,449	£17,557	£22,625
Mean Total Cost (NHS list prices)	£59,576	£22,218	£22,625
Incremental Life Years	-	0.654	0.654
Incremental QALYs	-	0.501	0.502
Incremental Cost (PASA prices)	-	£33,892	£28,824
Incremental Cost (NHS list prices)	-	£37,358	£36,951
Cost per QALY (PASA)	-	£67,714	£57,416
Cost per QALY (NHS list prices)	-	£74,640	£73,605

Table 3: Cost-effectiveness results for patients with triple negative disease

	BEV+PAC	PAC	DOC
Mean Life Years (yrs)	2.179	1.815	1.815
Mean QALYs	1.308	0.996	0.995
Mean Total Costs (PASA)	£42,342	£16,637	£22,269
Mean Total Costs (NHS list prices)	£48,809	£21,422	£22,269
Incremental Life Years	-	0.364	0.364
Incremental QALYs	-	0.312	0.313
Incremental Cost (PASA prices)	-	£25,705	£20,073
Incremental Cost (NHS list prices)	-	£27,387	£26,540
Cost per QALY (PASA)	-	£82,469	£64,092
Cost per QALY (NHS list prices)	-	£87,865	£84,740

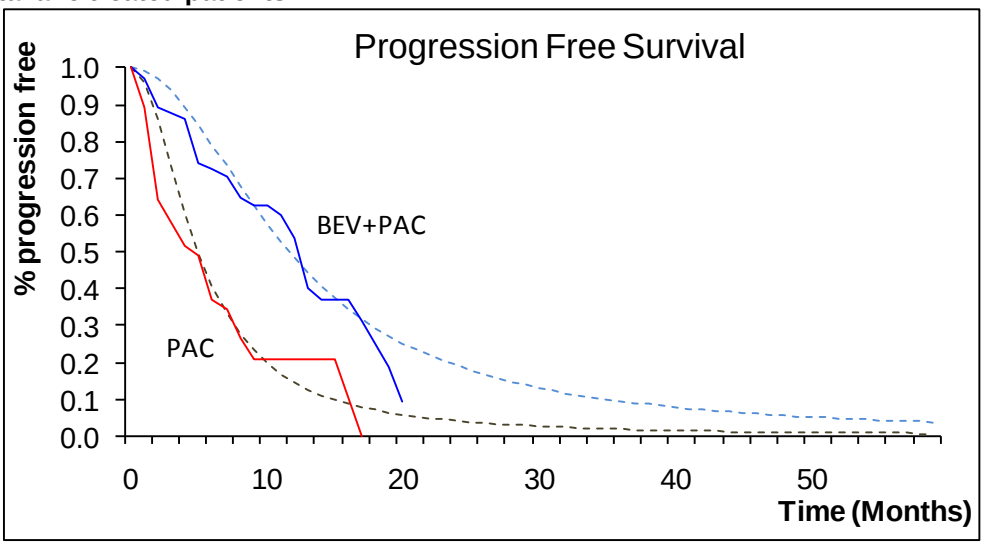
2. Uses an alternative approach to modeling overall survival

The previous model (presented in the manufacturer's original submission) assumed that the benefits on PFS between treatments would be maintained after progression of metastatic disease, leading to equivalent differences in mean OS between treatments over patients' lifetimes. The revised model uses an alternative approach based on fitting parametric extrapolations to both the PFS and OS data from the E2100 study.

Goodness of fit statistics for a range of alternative parametric functions were reported by the manufacturer for OS/PFS for the prior taxane group (p.8 of the additional submission); however, equivalent statistics were not reported for the triple negative group. The log-logistic model was selected by the manufacturer as the best fit for both OS and PFS.

Figures 1 and 2 provide a graphical comparison, undertaken by the ERG, of the Kaplan-Meier (KM) curves based on the observed data from the E2100 study and the parametric log-logistic curves used by the manufacturer in the cost-effectiveness analysis for the two subgroups. The areas between the log-logistic curves provide the basis for the manufacturer's incremental estimates of PFS and OS applied in the cost-effectiveness analysis.

Figure 1: ERG comparison of Kaplan-Meier and log-logistic curves (PFS and OS) for prior taxane treated patients



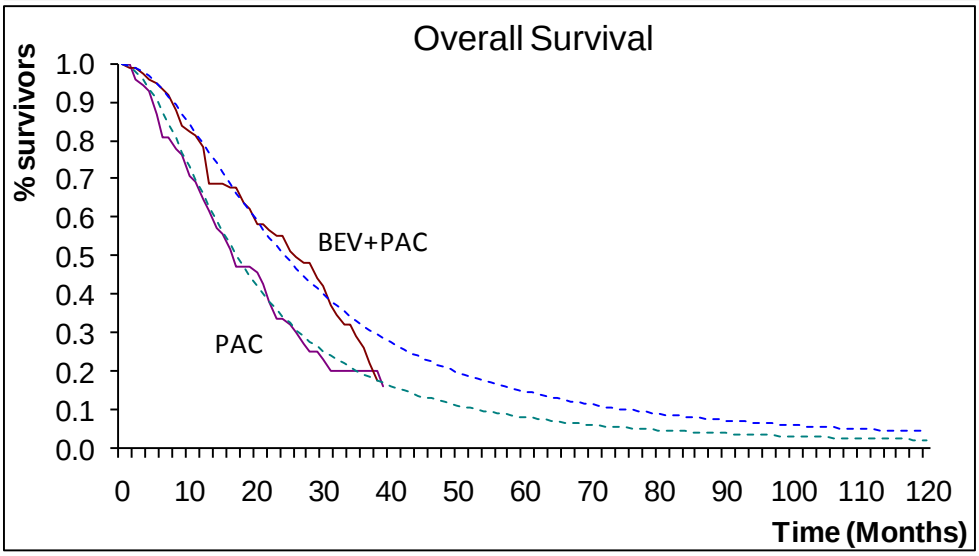
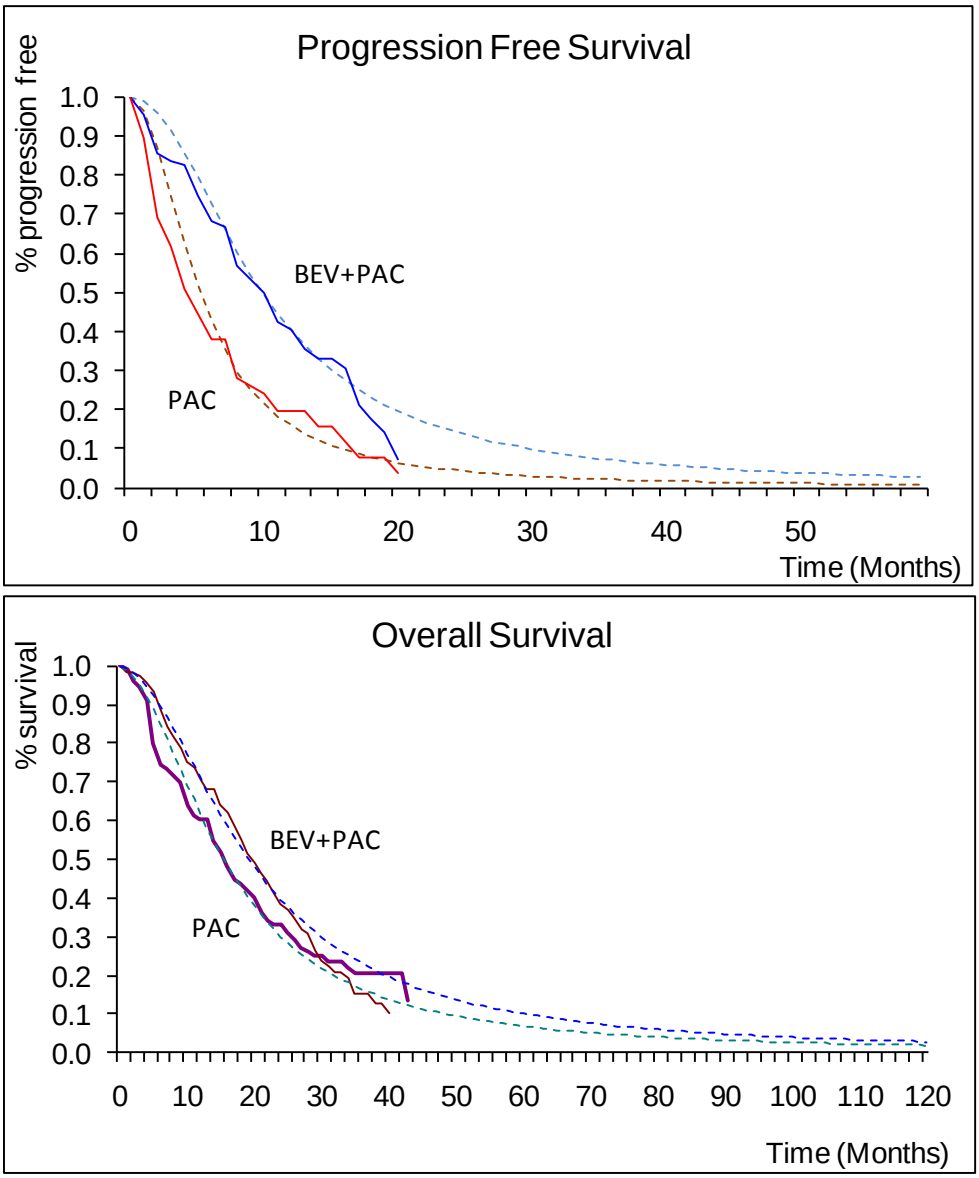


Figure 2: ERG comparison of Kaplan-Meier and log-logistic curves (PFS and OS) for patients with triple negative disease



3. Excludes the 10g capping scheme for the cost of bevacizumab from the base-case analysis in the revised model

The previous submission included a 10 gram cap on the cost of bevacizumab per patient. The base-case of the revised model does not include the capping scheme. A sensitivity analysis presented by the manufacturer shows that the capping scheme would reduce the ICER of BEV+PAC vs. PAC in the prior taxane subgroup from £67,714 to £46,447 per QALY. All the numbers quoted in this ERG addendum exclude the capping scheme because it has not yet been approved by the Department of Health.

ERG commentary

In this additional submission, the manufacturer has assumed that docetaxel has an equivalent effectiveness to weekly paclitaxel (p.27 of the additional submission), based on conclusions drawn from the original submission regarding the ITT population. The ERG considers that, for this assumption to be sustained, its validity in the subgroups themselves should have been demonstrated. The ERG will thus not present or discuss results of the comparison against docetaxel.

The ERG commentary focuses on the validity of the alternative approach to modeling OS and the robustness of the revised ICER estimates.

The original submission from the manufacturer was based on an assumption that gains in PFS would be translated into equivalent gains in OS. The ERG considered this a potentially optimistic assumption for OS, and noted the low external validity of the model predictions compared to the trial results themselves. The estimates of OS in the revised model from the manufacturer are now informed directly by OS data from the E2100 trial. The use of the log-logistic survival function for OS is based on the goodness of fit statistics reported for the prior taxane group. The ERG has identified two areas of potential concerns regarding this revised approach:

a. The use of goodness of fit statistics (and the lack of comparable statistics for the triple negative disease subgroup)
b. The exclusion of potentially relevant evidence from the AVADO and RIBBON-1 studies.

a. The use of goodness of fit statistics (and the lack of comparable statistics for the triple negative disease subgroup)

The log-logistic function was selected by the manufacturer as the best fit for both OS and PFS based on goodness of fit statistics (BIC and AIC) for the prior taxane subgroup. However, the difference in goodness of fit statistics is small, indicating that several of the competing functions appear to have a similar fit with the observed data. This suggests that the choice between the different functions appears marginal which is an important consideration given the subsequent variation in the ICER estimates based on the different survival functions.

A visual comparison of the alternative parametric functions is presented by the ERG in the Appendix. The Exponential and Gompertz models were not considered by the ERG due to the relatively poor goodness of fit statistics compared to the other functions. The statistical models differ substantially in their ‘tails’, that is, the predictions of OS beyond the trial period. The log-logistic has the heaviest tail and consequently predicts the largest difference in OS between the treatments. The Weibull has the shortest tail and hence predicts the smallest difference in OS compared to the other alternative parametric fits. The differences in the assumptions concerning the tail of the distributions (and the subsequent impact on the differences in OS) have important implications for the cost-effectiveness estimates. For example, the base-case ICER for BEV+PAC vs. PAC in the prior taxane subgroup is

£67,714, whereas using a Weibull function the ICER is £86,854 (pg 43 of the manufacturer's additional submission).

It should also be recognized that the goodness of fit statistics only provide an indication of how well the curves fit to the observed trial data. Consequently, they do not provide any indication of how valid the curves are beyond the observed data (i.e. to the 'extrapolation period' required to estimate mean PFS and OS for the economic analysis). The validity of extrapolations beyond the observed data can only be adequately determined based on comparisons with external data (i.e. longer term observational data) and/or clinical judgment.

The ERG considers that there is significant uncertainty surrounding the choice of statistical function used in the cost-effectiveness analysis. It should also be noted that all of the parametric functions investigated by the manufacturer continue to assume a long-term sustained treatment effect with BEV+PAC for OS. However, the KM curves from E2100 trial data appear to demonstrate that the difference in OS may not be sustained over a longer time horizon. Therefore, the ERG conducted an additional sensitivity analysis using the MS model directly employing the KM survival estimates from the E2100 RCT up to approximately 3.2 years (20% of patients in the PAC arm and 18% of patients in the BEV arm were still alive at 38 months), and assumed that there was no further difference in incremental survival after this time.

The cost-effectiveness results based on the different distributions (including the additional KM analysis undertaken by the ERG) are reported in Table 4. The ICER for BEV+PAC vs. PAC in this sensitivity analysis ranged from £67,714 to £117,587 per QALY based on PASA prices for paclitaxel and £74,640 and £129,794 per QALY based on NHS list prices.

Table 4: ERG comparison of cost-effectiveness results using different survival distributions (prior taxane group)

Cost effectiveness outcomes for the comparison BEV+PAC vs. PAC – *prior taxane group*	Distribution function used for PFS and OS				
	Log logistic	Weibull	Gamma	Log Normal	KM
Incr QALYs	0.501	0.387	0.463	0.486	0.284
Incr Costs (PASA prices)	£33,892	£33,628	£34,024	£34,126	£33,390
Incr Costs (NHS list prices)	£37,358	£37,094	£37,491	£37,592	£36,857
Cost per QALY (PASA prices)	£67,714	£86,854	£73,524	£70,180	£117,587
Cost per QALY (NHS list prices)	£74,640	£95,807	£81,014	£77,309	£129,794

It is important to note that the equivalent goodness of fit statistics was not reported by the manufacturer for the triple negative disease subgroup. Nor were the equivalent parameter estimates supplied in the Excel model. This precluded the ERG from undertaking any assessment of the goodness of fit and the robustness of the ICER estimates to alternative parametric survival functions. However, the ERG was able to undertake an additional sensitivity analysis using the KM data. A comparison of the cost-effectiveness results from the log-logistic and KM analysis for the triple negative group is reported in Table 5. The results appear highly sensitive to the choice of survival distribution and sustaining potential OS benefits during the extrapolation period.

Table 5: ERG comparison of cost-effectiveness results using different survival distributions (triple negative group)

Cost effectiveness outcomes for the comparison BEV+PACvs. PAC – *triple negative group*	Distribution function used to describe time to progression and time to death	
	Log logistic (base case)	KM
Incr QALYs	0.312	0.143
Incr Costs (PASA prices)	£25,705	£25,107
Incr Costs (NHS list prices)	£27,387	£26,789
Cost per QALY (PASA prices)	£82,469	£175,575
Cost per QALY (NHS list prices)	£87,865	£187,339

b. The exclusion of potentially relevant evidence from the AVADO and RIBBON-1 studies.

The approach to estimate the relative effectiveness data in the model is informed exclusively from the E2100 trial. This approach ignores potentially relevant evidence from the AVADO and RIBBON-1 trials which were included with the E2100 trial in the IPD meta-analysis. Details of both these trials and the meta-analysis are presented by the manufacturer (see section 2 above). The ERG considers that alternative scenarios related to the treatment effect should have been explored in the model based on the additional RCT evidence presented. The ERG has undertaken an additional exploratory analysis using the hazard ratio estimated from the IPD meta analysis of the prior taxane subgroup presented by the manufacturer (HR=0.738). The parameterisation of this approach was only possible using the Weibull function. Since the Weibull regression coefficients were not reported for the triple negative group, results are only presented for the prior taxane subgroup. The results of this analysis are presented in Table 6 and demonstrate that applying an estimate of the relative effectiveness of BEV+PAC vs. PAC from the IPD meta analysis (as opposed to the E2100 trial only), increases the ICER from £86,844 to £98,834 (PASA prices) and from £95,807 to £109,242 (NHS list prices).

Table 6: ERG exploratory analysis using the treatment effect estimated by the IPD meta-analysis (prior taxane group) and assuming that OS and PFS are estimated by a Weibull distribution

Cost effectiveness outcomes for the comparison BEV+PACvs. PAC – *prior taxane group*	Cost effectiveness estimates
Incr QALYs	0.333
Incr Costs (PASA prices)	£32,917
Incr Costs (NHS list prices)	£36,383
Cost per QALY (PASA prices)	£98,834
Cost per QALY (NHS list prices)	£109,242

In conclusion, the ERG considers that the removal of the capping scheme from the model was appropriate because the scheme has not yet been approved by the Department of Health. The additional submission presented a revised model to estimate cost-effectiveness of BEV combination therapy vs. a taxane in two subgroups. The ERG considers that the manufacturer has not demonstrated the equivalence of efficacy of DOC and PAC in the subgroups and therefore results for BEV+DOC vs. DOC are not robust in either subgroup.

The results of the model for BEV+PAC vs. PAC are highly sensitive to assumptions made about extrapolation of treatment effects on overall survival and PFS beyond the time horizon of the E2100 trial. The ERG considers that the manufacturer has not convincingly demonstrated that the base-case (using the log-logistic function) is the most appropriate method of estimating all cause and progression-free life expectancy in either subgroup. All alternative methods of extrapolation estimate higher ICERs for BEV+PAC vs. PAC in the prior taxane subgroup. Furthermore, the model has not made use of evidence from other RCTs about the treatment effect in either subgroup. Incorporating the IPD meta-analysis into the model would give higher estimates of the ICER of BEV+PAC vs. PAC in the prior taxane subgroup.

4. References

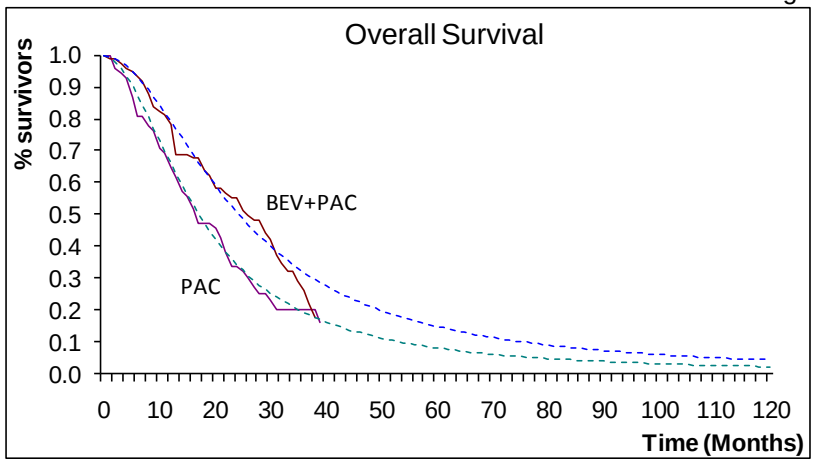
1. Bevacizumab in combination with taxanes for the treatment of HER-2 negative 1st line metastatic breast cancer. New evidence based on two subgroups: prior-taxane treated patients and triple negative patients. Roche new evidence submission to the National Institute for Health and Clinical Excellence, submitted 24th September 2010.

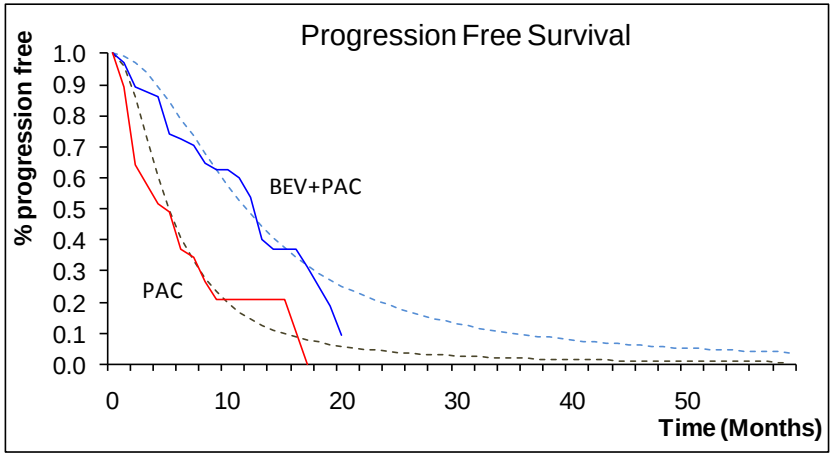
2. Robert NJ, Dieras V, Glaspy J, Brufsky A, Bondarenko I, Lipatov O, et al. RIBBON-1: Randomized, double-blind, placebo-controlled, phase III trial of chemotherapy with or without bevacizumab (B) for first-line treatment of HER2-negative locally recurrent or metastatic breast cancer (MBC). *J Clin Oncol* 2009;27:(suppl; abstr 1005).

3. Miles DW, et al. Meta-analysis of patients previously treated with taxanes from three randomized trials of bevacizumab (BV) and first-line chemotherapy as treatment for metastatic breast cancer (MBC). ESMO Congress, Milan, Poster Discussion October 11, 2010; Roche data on file.

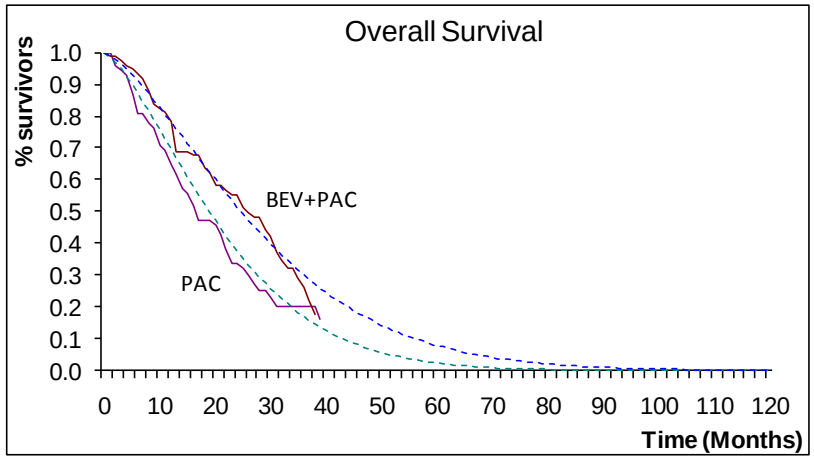
5. **Appendix -** ERG comparison of Kaplan-Meier and alternative parametric distributions (PFS and OS) for prior taxane treated patients

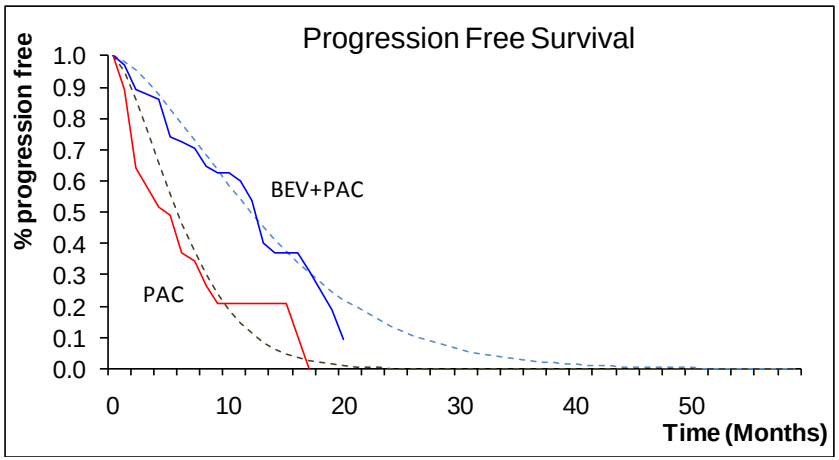
Log logistic



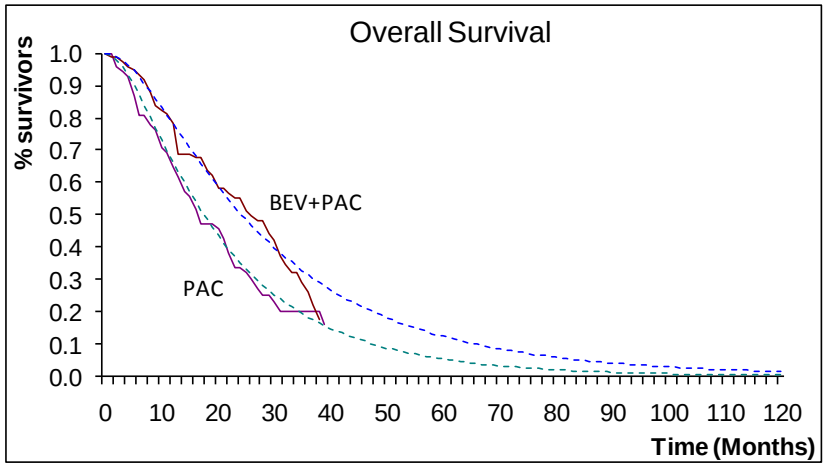


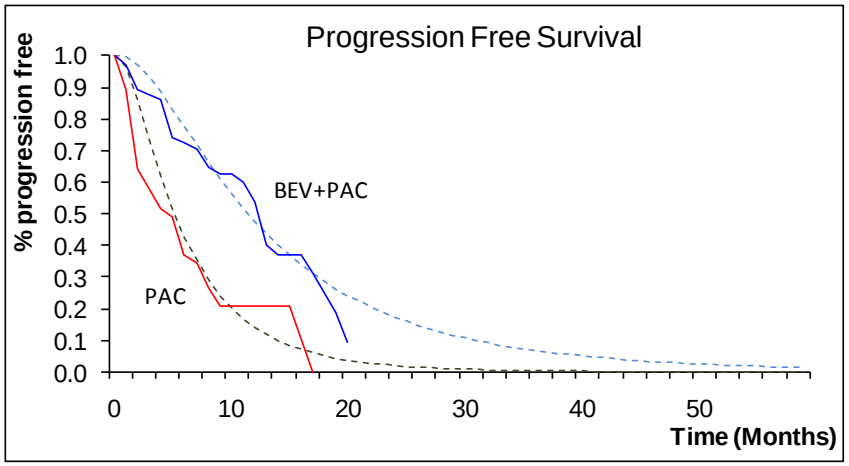
Weibull





Generalised Gamma





Log normal