Evidence Review Group's Report

Title: Bevacizumab in combination with a taxane for the first-line treatment of HER2-negative metastatic breast cancer

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 Dave Fox, Information Specialist, CRD Alison Eastwood, Senior Research Fellow, CRD
Correspondence to	Mark Rodgers, Research Fellow, CRD
Date completed	17th May 2010

Source of funding: This report was commissioned by the NIHR HTA Programme as project number HTA 09/25.

Declared competing interests of the authors

None

Acknowledgements

The ERG would like to thank Prof Galina Velikova (Professor of Psychosocial and Medical Oncology, St James's Institute of Oncology, Leeds) for providing clinical advice and

commenting on drafts of the report, as well as Steve Palmer (Senior Research Fellow, CHE) and Mark Sculpher (Professor of Health Economics, CHE) for their advice and comments on this report.

Rider on responsibility for report

The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:

CRD and CHE Technology Assesment Group. Bevacizumab in combination with a taxane for the first-line treatment of HER2-negative metastatic breast cancer: a single technology appraisal. Centre for Reviews and Dissemination/ Centre for Health Economics, 2010.

Authors' contributions

Mark Rodgers and Huiqin Yang wrote the clinical effectiveness sections of the report. Marta Soares and David Epstein wrote the cost effectiveness sections of the report and conducted the economic analyses. Dave Fox wrote the sections on the search strategies. Alison Eastwood commented on and revised the report.

Table of Contents

List of Tables

List of Figures

Definitions of Terms

Adverse effect

An abnormal or harmful effect caused by and attributable to exposure to a chemical (e.g. a drug), which is indicated by some result such as death, a physical symptom or visible illness.

Functional Assessment of Cancer Therapy-Breast (Fact-B) Questionnaire

A 44-item self-report instrument designed to measure multidimensional quality of life in patients with breast cancer. The FACT-B consists of the FACT-General plus the Breast Cancer Subscale, which complements the general scale with items specific to quality of life in patients with breast cancer.

Human epidermal growth factor receptor 2 (HER2)

A type of receptor tyrosine kinase involved in regulating a diverse repertoire of cellular processes that control cell growth, survival, differentiation and migration. Cancer cells removed from the body can be tested for the presence of HER2 to help decide the best treatment.

Intention-to-treat

An intention-to-treat analysis is one in which all the participants in a trial are analysed according to the intervention to which they were allocated, whether they received it or not.

Mixed treatment comparison

Mixed treatment comparison is a form of meta-analysis used to strengthen inference concerning the relative efficacy of two treatments. It uses data based on direct comparisons (A vs. B and B vs. C trials) and indirect comparisons (A vs. C trials). It facilitates simultaneous inference regarding all treatments in order to select the best treatments.

Overall survival

Overall survival is a term that denotes the chances of staying alive for a group of individuals. It is often stated as the percentage of individuals in the group who are likely to be alive after a particular duration of time.

Progression free survival

Progression-free survival is a term that denotes the chances of staying free of disease progression for a group of individuals suffering from a cancer after a particular treatment. It is often stated as the percentage of individuals in the group whose disease is likely to remain stable (with no sign of progression) after a specified duration of time.

Quality Adjusted Life Year (QALY)

An index of health gain where survival duration is weighted or adjusted by the patient's quality of life during the survival period. QALYs have the advantage of incorporating changes in both quantity (mortality) and quality (morbidity) of life.

Quality of Life

A concept incorporating all the factors that might impact on an individual's life, including factors such as the absence of disease or infirmity as well as other factors which might affect their physical, mental and social well-being.

List of abbreviations

AE	Adverse events
CEA	Cost-effectiveness analysis
CI	Confidence interval
CNS	Central nervous system
ERG	Evidence Review Group
FACT-B	Functional Assessment of Cancer Therapy-Breast
HER2	Human epidermal growth factor receptor 2
HR	Hazard ratio
HRQoL	Health related quality of life
ICER	Incremental cost-effectiveness ratio
ITT	Intention-to-treat
mBC	Metastatic breast cancer
MS	Manufacturer's submission
MTC	Mixed treatment comparison
NHS	National Health Service
NICE	National Institute for Health and Clinical Excellence
N/A	Not applicable
NR	Not reported
OS	Overall survival
PFS	Progression-free survival
QALY	Quality-adjusted life years
RCT	Randomised controlled trial

SAE Serious adverse events

SPC Summary of product characteristics

TOI-B FACT-B trial outcome index

TOT-B FACT-B total score

TTP Time to progression

1 SUMMARY

1.1 Scope of the submission

This report presents the ERG's assessment of the manufacturer's (Roche) submission to NICE on the use of bevacizumab, in combination with a taxane, for the treatment of untreated metastatic HER-2 negative breast cancer in patients for whom anthracyclines are not appropriate. The report includes an assessment of both the clinical and cost effectiveness evidence submitted by the company.

The manufacturer's evaluation of clinical efficacy included only evidence relating to bevacizumab in combination with weekly paclitaxel. Evidence on bevacizumab in combination with docetaxel was excluded.

Cost effectiveness analysis focused on a comparison between weekly (qw) paclitaxel plus bevacizumab and paclitaxel qw based on the results of the E2100 trial. Although an indirect comparison was conducted, the results were not implemented in the manufacturer's model.

1.2 Summary of submitted clinical effectiveness evidence

The main clinical effectiveness data were derived from a single open-label randomised controlled trial (RCT), E2100, that evaluated the addition of bevacizumab to qw paclitaxel chemotherapy in patients with HER2-negative metastatic breast cancer (mBC) who had not previously received chemotherapy for advanced disease. This trial reported statistically significant increases in median progression-free survival (PFS), from 5.8 months to 11.3 months (Hazard Ratio (HR) 0.54, 95% CI 0.44 to 0.67) and an objective response rate from 22.2% to 49.8% ($p<0.0001$) for bevacizumab plus paclitaxel versus paclitaxel alone. Median overall survival was not significantly different between the two groups (26.5 vs. 24.8 months; HR 0.87, 95% CI 0.72 to 1.05). A *post-hoc* analysis indicated that overall survival at 1 year was significantly higher with paclitaxel plus bevacizumab than paclitaxel alone (81.4% vs. 74.0%, $p=0.017$). The manufacturer reported that the addition of bevacizumab to paclitaxel qw therapy was associated with a significant improvement in quality of life as measured by FACT-B trial outcome index (TOI-B) score at week 33 ($p=0.0042$), and in FACT-B total score (TOT-B) at week 17 ($p=0.0475$) and week 33 ($p=0.0046$) compared with paclitaxel alone.

In the absence of head-to-head comparisons for certain comparators, the manufacturer conducted an indirect comparison using the Bucher method.[1] This reported that bevacizumab plus qw paclitaxel was associated with a significant improvement in progression-free survival when compared with 3-weekly (q3w) docetaxel (HR 0.56, 95% CI 0.39 to 0.78), and with gemcitabine plus q3w paclitaxel (HR 0.46, 95% CI 0.34 to 0.64). No significant difference was found for progression-free survival between qw paclitaxel and q3w docetaxel (HR 1.15, 95% CI 0.89 to 1.48), or between qw paclitaxel and gemcitabine plus q3w paclitaxel (HR 0.96, 95% CI 0.76 to 1.21).

On the basis of E2100 and a large uncontrolled study of over 2000 patients, the manufacturer concluded that bevacizumab is not associated with the commonly recognised side-effects of cytotoxic anti-cancer therapies and that the most common adverse events associated with bevacizumab therapy are hypertension and proteinurea.

1.3 Summary of submitted cost effectiveness evidence

The submission identified six cost-effectiveness analyses, but stated that they were not relevant as they were all conducted outside the UK. The manufacturer, therefore, justified the development of a *de novo* economic model which considered patients with the same baseline characteristics as seen in women in the E2100 trial.

The model assessed:

- BEV+PAC: bevacizumab 10mg/kg (every 2 weeks) in combination with paclitaxel 90mg/m2 (weekly for 3 weeks followed by 1 week of rest).
- PAC qw: paclitaxel (monotherapy) 90mg/m² weekly for 3 weeks followed by 1 week of rest.
- DOC: docetaxel (monotherapy) 75 mg/m² on day 1 every 21 days (considered current UK NHS clinical practice in the submission).
- GEM+PAC: gemcitabine 1,250mg/m² days 1 and 8 plus paclitaxel 175 mg/m² on day 1 every 21 days

Pairwise comparisons were made between BEV+PAC and PAC (using the E2100 trial), BEV+PAC and DOC, and BEV+PAC and GEM+PAC.

The model was Markov model with three states: progression free, progressed and dead; and used a 10 year time horizon. Parametric survival functions were used to model the rate of progression based on data from the E2100 trial, and the rate of progression from the progression-free state was assumed to be the same after PAC qw as after DOC and after GEM-PAC. It was assumed that the hazard of death after progression was constant over time and the same across all treatments, meaning that any difference in progression free survival between treatments is mirrored in terms of overall survival. The costs and disutility associated with treatment-related adverse events were included based on the incidence of events in the E2100 trial. Utility estimates were derived from a non-systematic literature review of studies of patients with breast cancer. A number of cost categories were considered: drug acquisition costs, drug administration costs, duration of treatment, supportive care costs, adverse event costs and end of life costs. Two alternative base-case analyses were presented, one using product list prices and the other using PASA (Purchasing and Supply Agency, NHS) prices for paclitaxel. The latter included a capping scheme for BEV which the ERG understands has not been approved by the Department of Health.

Based on NHS list prices, the manufacturer's model estimated incremental cost effectiveness ratios (ICERs) for BEV+PAC of £117,803, £115,059, and £105,777 per QALY gained, respectively, relative to PAC, DOC and GEM+PAC regimens, respectively. If PASA prices for PAC with a 10g cap on the cost per patient of BEV are used instead, the ICERs for BEV+PAC are estimated at £77,314, £57,753 and £60,101 per QALY, respectively. The submission suggests that, based on the above results, the regimen of BEV+DOC is not cost effective because it is considered less effective than BEV+PAC and more costly.

1.4 Commentary on the robustness of submitted evidence

1.4.1 Strengths

The manufacturer's systematic review of the literature used appropriate search methods to identify the relevant comparative evidence on bevacizumab in mBC. The main findings were derived from an open-label RCT (E2100) which was conducted in a relevant population and steps were taken to mitigate against methodological limitations (e.g. intention-to-treat analyses of independently reviewed outcomes was undertaken). The safety evaluation included the most comprehensive and robust study available to assess this outcome.

The manufacturer's submission largely conforms to the NICE Reference Case for cost-effectiveness analysis and was reasonably clearly presented.

1.4.2 Weaknesses

The manufacturer's search for relevant evidence identified a second RCT (the AVADO trial) that evaluates bevacizumab plus a taxane. Though this trial provides information on the efficacy of adding bevacizumab to q3w docetaxel (the monotherapy currently recommended by NICE for first-line treatment of patients with advanced breast cancer who are not suitable for anthracyclines),[2] the manufacturer excluded AVADO because the dose was considered inappropriate compared to routine UK clinical practice. However, this decision was inconsistent with clinical advice received by the ERG.

Limitations in the collection and analysis of data in E2100 impact on the reliability of the trial's findings. Firstly, the treatment regimens received by patients after disease progression are unknown, as these data were not collected. Consequently, the influence of post-progression treatment on overall survival in this trial is unknown. Secondly, the relative improvements in quality of life reported in E2100 were based on analysis using extreme imputed data for missing data; without these imputed data, differences between groups are statistically insignificant. These data were not further used in the cost effectiveness model.

The ERG identified several limitations and inconsistencies relating to the indirect comparison of efficacy. One inclusion criterion (<60% of patients receiving second-line chemotherapy for mBC) may have been formulated to allow the inclusion of one specific trial. Though the AVADO trial was also excluded from the indirect comparison on the basis of docetaxel dose, another trial (Jones 2005[3]) that used the same dose was included, suggesting an inconsistent application of inclusion criteria. The manufacturer's alternative justification for excluding AVADO from the indirect comparison (that the Bucher method is only intended to compare two trials via a common comparator) highlights limitations in their existing network, which extends the Bucher method to indirectly link three trials via two comparators. Given these methodological limitations, combined with concerns about the validity and exchangeability of the included trials, the ERG does not consider the findings of the indirect comparison to be reliable.

The manufacturer's cost-effectiveness model did not formally consider all relevant comparators. Specifically, bevacizumab in combination with docetaxel, and q3w paclitaxel were not formally considered despite the latter being used in clinical practice in the UK.[a] The manufacturers assumed that the rate of death after progression is constant over time and the

[a] The manufacturer advises that bevacizumab in combination with docetaxel is not used in UK clinical practice

same for all initial treatments, with the implication that differences in mean progression free survival (PFS) between treatments are maintained in the mean overall survival (OS) estimates. This is a strong assumption and alternative model structures were not considered by the manufacturer. As a likely result, the model overestimated the difference in overall survival for BEV+PAC versus PAC compared with the E2100 trial.

The base-case model assumed that the regimens PAC, DOC and GEM+PAC are equally effective; no alternative scenarios presented.

Despite the use of a disease-specific health-related quality of life instrument in the E2100 trial (the FACT-B), no mapping algorithm was used to link this to a preference –based (utility) instrument such as the EQ5D. Instead, external utility estimates were used based on a literature search which was not systematic. No attempt was made to collate or synthesise the alternative estimates, and the selection of utilities for the model seemed arbitrary.

In an alternative base-case, paclitaxel (non-proprietary) was costed based on the average NHS Purchasing and Supply Agency (PASA) price, but the unit costs for all other (proprietary) treatments were based on list prices (BNF). The ERG understands that the price cap assumed for bevacizumab has not been agreed with the Department of Health and should not, therefore, have been assumed in the model. The patent for docetaxel is soon going to expire (November 2010)[b], but the manufacturer failed to explore the implications of a likely reduction in its acquisition cost. The analysis also ignored the possibility of dose reductions. The extent to which dose reductions occur may differ between alternative treatments, and the ERG expects this to impact on the results. The manufacturer undertook no sub-group analysis. The model results were presented as a series of pairwise ICERs comparing BEV+PAC individually with the alternative regimens. This is inappropriate and a fully incremental analysis should have been undertaken.

1.4.3 Areas of uncertainty

Efficacy outcomes for bevacizumab plus qw paclitaxel versus qw paclitaxel alone were based on an interim analysis of the E2100 trial – all progression free survival and response data were collected prior to February 2005 and overall survival data were collected prior to October 2006. Analysis of more complete follow-up data would be valuable, though the manufacturer stated that no such analyses are available.

---

[b] Amended from May 2010 by ERG on advice of Roche

The lack of overall survival benefit for combination therapy observed in the E2100 trial could be due to a number of reasons, including cross-over between treatment arms after disease progression. However, this cannot be established as data on post-progression treatment were not collected in E2100.

Multiple methodological limitations in the indirect comparison mean that the relative efficacy of bevacizumab plus qw paclitaxel versus comparators other than paclitaxel alone outlined in the decision model remains highly uncertain.

No evidence on the clinical efficacy or safety of bevacizumab in combination with docetaxel was presented in the manufacturer's submission.

The methodological weaknesses in the model described above give rise to a number of uncertainties; the ERG undertook a series of analyses to explore their implications:

- The use of PASA discount (without the cap on BEV) made very little difference to the incremental costs of BEV+PAC versus PAC, compared with using NHS list prices.

- The acquisition cost of docetaxel has very little effect on the ICER of DOC versus BEV+PAC. However, the price of DOC may be important in any comparison of the cost-effectiveness of taxane monotherapies (PAC versus DOC).

- The ERG evaluated BEV+DOC versus DOC alone based on the results of the AVADO RCT. This found that the ICER was more than £250,000 per QALY.

- The ERG constructed an alternative model that was calibrated to the E2100 results for OS. The ICER of BEV+PAC versus PAC was over £250,000 per QALY in the revised model.

1.5 Key issues

Despite some methodological limitations, the E2100 trial provides direct evidence to suggest that the addition of bevacizumab to qw paclitaxel can increase objective response and progression free survival in the first-line treatment of mBC. The same trial fails to show a benefit in terms of overall survival. Whether this is a true null finding or due to crossover between treatment arms cannot be established as relevant data were not collected.

The intervention specified in the scope issued by NICE required an evaluation of bevacizumab in combination with a taxane (including both paclitaxel and docetaxel). Since docetaxel is the taxane currently recommended for first-line treatment of patients with advanced breast cancer in existing NICE guidelines,[2] the ERG extracted the limited available data from published AVADO abstracts. In terms of response rate and PFS, the AVADO trial reported a markedly smaller benefit of adding bevaczumab to docetaxel than was reported for adding bevacizumab to qw paclitaxel in E2100. One explanation for this difference might be potentially greater benefits, in terms of disease progression, associated with docetaxel monotherapy relative to paclitaxel monotherapy, though, without any head-to-head comparison of the bevacizumab-taxane combinations, other confounding factors could be responsible. AVADO also reported no statistically significant effect of combination therapy versus docetaxel in terms of overall survival.

Given the important limitations around the evidence selected and methods used for the indirect comparison, the manufacturer's reporting of a statistically significant benefit of bevacuzimab plus qw paclitaxel over the currently recommended first-line treatment of docetaxel monotherapy cannot be considered reliable. Similar limitations apply to the indirect comparison of combined bevacuzimab and qw paclitaxel against combined gemcitabine and q3w paclitaxel.

The cost-effectiveness analysis presented by the manufacturers included judgments and assumptions which are subject to uncertainty. However, the manufacturers' own analysis suggested ICERs for BEV+PAC of £58,000 to £77,000 per QALY gained (based on PASA prices for PAC and a 10g cap on the cost per patient of BEV), and £106,000 to £118,000 based on NHS list prices. Further analysis by the ERG suggested that alternative assumptions can increase the ICERs yet further and, based on current prices, no plausible changes to the model assumptions will bring the ICERs for BEV+PAC down to NICE's £20,000 to £30,000 per QALY threshold.

2 BACKGROUND

2.1 Critique of manufacturer's description of underlying health problem

The manufacturer provides a brief but accurate summary of the key issues around metastatic breast cancer (mBC).

2.2 Critique of manufacturer's overview of current service provision

The manufacturer provides an accurate overview of the treatment pathway for mBC patients. The ERG asked for additional information regarding which regimens are currently used for the first line treatment of mBC in UK clinical practice. The manufacturer's response (see A1 of the manufacturer's response to clarifications document), based on IMS data, indicated that 36% of all mBC patients receive taxanes as first-line therapy, the majority of whom (29%) receive docetaxel. A similar proportion (34%) of patients receive either capecitabine or vinorelbine, having either had recent (<12 months) taxane therapy or being considered unable to tolerate taxane therapy. Among those patients who receive first-line taxane therapy for mBC, the manufacturer does not give details of the regimens used in routine practice. For example, it is not clear what proportion of patients receive standard q3w paclitaxel[4] versus the qw paclitaxel regimen advocated in the manufacturer's submission (MS).

3 Critique of manufacturer's definition of decision problem

3.1 Population

The decision problem specifies the relevant population as people with untreated HER2-negative mBC for whom anthracyclines are not appropriate. The key trial in the direct efficacy comparison (E2100) [5-13]was relevant to the decision problem, being limited to patients with previously untreated mBC, over 90% of whom were HER2-negative. An additional uncontrolled safety study (ATHENA)[14-20] also included patients with previously untreated mBC, 97.2% of whom were HER2-negative.

The indirect comparison included trials with populations other than those described in the decision problem: three trials in which the HER2 status of patients was unknown; [3, 21, 22] one trial predominately (55%) included patients who had been previously treated with chemotherapy for mBC;[3] and a second included a substantial minority (29%) of previously treated patients.[22]

3.2 Intervention

Bevacizumab is licenced in combination with paclitaxel or docetaxel for the first-line treatment of mBC,[23] and the final scope issued by NICE specifies bevacizumab in combination with a taxane to be the intervention of interest. However, the manufacturer's evaluation of clinical efficacy and cost-effectiveness includes only evidence relating to bevacizumab in combination with weekly paclitaxel. Addition of bevacizumab to the current licensed q3w paclitaxel regimen was not considered. Evidence on bevacizumab in combination with docetaxel was also excluded (see Section 4.1 of this report).

The uncontrolled study of safety (ATHENA)[14-20] included patients treated with bevacizumab in combination with paclitaxel monotherapy, docetaxel monotherapy, taxane-based combination regimens, or non-taxane therapies.

3.3 Comparators

The decision problem specifies that bevacizumab in combination with paclitaxel and bevacizumab in combination with docetaxel should be compared with each other. No head-to-head trials were available for this comparison, and the manufacturer did not attempt to address this in their indirect comparison and cost-effectiveness analysis.

The remaining comparators specified in the decision problem were: docetaxel monotherapy, paclitaxel monotherapy and paclitaxel in combination with gemcitabine. One included trial evaluating bevacizumab (E2100[5-13]) included a weekly paclitaxel comparator arm; the remaining comparators were addressed using indirect comparisons. One trial evaluating the addition of bevacizumab to docetaxel monotherapy was excluded from the manufacturer's submission (see Section 4.1).

In the manufacturer's cost-effectiveness analysis only a once weekly paclitaxel regimen was considered for paclitaxel monotherapy and the current licensed q3w paclitaxel regimen was excluded.

3.4 Outcomes

Each of the outcomes specified in the decision problem (overall survival, progression free survival, response rates, adverse events and health-related quality of life) were addressed to some extent for the evaluation of bevacizumab plus paclitaxel versus paclitaxel monotherapy, with a focus on progression free survival (PFS).

For all other comparisons, PFS was the only efficacy outcome reported.

Data on PFS were combined with assumptions on overall survival to estimate mean survival times in the cost-effectiveness analysis. Durations of PFS and post-progression survival for each individual regimen were quality-adjusted using utility weights to estimate quality-adjusted life years.

3.5 Time frame

Length of follow-up in the trial evaluating bevacizumab plus paclitaxel versus paclitaxel monotherapy (E2100[5-13]) appeared to be adequate for a mBC population; patients were followed until disease progression then death, or for 5 years after randomization. However, the median length of follow-up was not reported.

3.6 Other relevant factors

N/A

4 CLINICAL EFFECTIVENESS

4.1 Critique of manufacturer's approach

4.1.1 Description of manufacturer's search strategy and comment on whether the search strategy was appropriate

The submission was checked against the Single Technology Appraisal (STA) specification for manufacturer/ sponsor submission of evidence update (July 2008).

The manufacturer's submission described the search strategies used to identify relevant studies of bevacizumab and breast cancer, and full details of the search strategies used in each section were reported in the appendices or in the clarifications provided. Overall, the search strategies employed for each of the sections of the submission was appropriate.

4.1.1.1 Search strategy for clinical evidence

The submission gave detailed descriptions of the search strategies and met NICE requirements (MS, p50-57). It included the specific databases searched (MEDLINE, MEDLINE In-Process, EMBASE, BIOSIS Previews and the Cochrane Library); the service providers used; the dates when searches were conducted; the date spans of the searches; the complete strategies used; the number of records identified for each search set; and the final result number.

Conference abstracts from American Society of Clinical Oncology (ASCO), San Antonio Breast Cancer Symposium (SABCS), European Cancer Organisation (ECCO) and European Society for Medical Oncology (ESMO) were searched online.

The search strategies were structured using a combination of subject indexing and free text search terms; thesaurus terms were focussed where appropriate; and search facets were correctly combined using Boolean operators. Truncation and wildcards were not used. Animal studies were excluded where possible.

The searches did not include any search facet for taxanes, so aimed to retrieve all studies relating to bevacizumab and metastatic breast cancer. The terms used for each search facet were relatively narrow. Normal practice would be to include drug trade names in searches, but for these searches the term 'Avastin' was not included. However, the ERG conducted

some additional searches to check if this might have resulted in other relevant studies being retrieved, and in our opinion it would not have done so.

4.1.1.2 Search strategy for indirect treatment comparison

The submission gave detailed descriptions of the search strategies and met NICE requirements (MS, p.104-105). It included the specific databases searched (MEDLINE, EMBASE and BIOSIS); the service providers used; the dates when searches were conducted; the date spans of the searches; the complete strategies used; and the number of records identified for each search set. Conference abstracts from American Society of Clinical Oncology (ASCO) were also searched online.

The search strategies were structured using a combination of subject indexing and free text search terms; thesaurus terms were exploded when relevant; truncation and wildcards were appropriately used; and search facets were correctly combined using Boolean operators. Animal studies were excluded where possible.

The thesaurus terms used were appropriate to each database searched and the comparator terms searched for were comprehensive.

The filters used to identify study types in the searches were appropriate to each database searched.

The ERG, therefore, considers the search strategies for direct clinical evidence and the indirect treatment comparison to be appropriate.

4.1.2 Statement of the inclusion/exclusion criteria used in the study selection and comment on whether they were appropriate

The MS states two sets of inclusion/exclusion criteria: one for their direct comparison of efficacy and safety (MS, p.63), and a second for their indirect treatment comparison (MS, p.105).

Direct efficacy comparison: The direct comparison of efficacy included controlled clinical trials evaluating bevacizumab in combination with paclitaxel or docetaxel for the first-line treatment of mBC in predominantly (≥90%) HER2-negative patients, and reporting relevant efficacy (PFS, OS, response rates) and safety outcomes. These inclusion criteria appear

appropriate as they are in line with both the decision problem and the bevacizumab license. Those studies not meeting the inclusion criteria (e.g. bevacizumab not evaluated, non-first-line mBC setting, no available efficacy data) were excluded, as were any studies considered statistically underpowered or recruiting less than100 patients, and studies "in which the partner agent was not given according to routine UK clinical practice".

The ERG considered that excluding studies on the basis of inadequate statistical power might be inappropriate if there was the potential for several such studies to be combined in a meta-analysis. One potentially relevant trial was excluded on this basis (RIBBON-1[24-26]), though outcome data were not reported adequately to allow data extraction (see Section 4.1.4 of this report for details).

To ensure that small but relevant studies were not excluded, the ERG asked the manufacturer to provide details of the 12 studies excluded on the basis of recruiting <100 patients. The manufacturer provided abstracts for each of these and the ERG confirmed that the each study could be appropriately excluded from the review for reasons other than sample size.

The MS did not define *a priori* what was considered "routine UK clinical practice" for the administration of partner agents, allowing some subjectivity in the application of this exclusion criterion. One large double-blind RCT (AVADO[27-36]) was excluded on this basis, though the ERG considers the exclusion of this trial to be inappropriate (see Section 4.1.4 for details).

Indirect treatment comparison: RCTs in patients with mBC were eligible for inclusion in the indirect comparison, with inclusion criteria broadened to include comparisons of any of the agents described in the decision problem (i.e. bevacizumab + paclitaxel, paclitaxel monotherapy, docetaxel monotherapy, gemcitabine + paclitaxel, or bevacizumab + docetaxel). However, two selection criteria were inconsistent with the direct efficacy comparison.

Firstly, the threshold for the proportion of HER2 negative patients among eligible studies was lowered from >90% to >50%. However, three of the four trials included in the indirect comparison[3, 21, 22] did not report the HER2 status of participants and did not, therefore, strictly meet this criterion. In their response to the ERG's request for clarification (A3 and A6), the manufacturer stated that this was because many clinical studies recruited patients prior to widespread HER2 testing and - in the absence of enrichment of the population for HER2 positive patients - less than 50% would be expected to be HER2 positive. Clinical advice to

the ERG suggests that this is a reasonable assumption, and that no more than 30% of mBC patients would be expected to be HER2-positive. Since the precise proportion of HER2 positive/negative patients across these three studies remains unknown, any interaction between HER2 status and treatment effect could theoretically influence the findings of the indirect treatment comparison. However, the ERG does not consider this alone to be a major limitation.

Secondly, studies selected for the indirect comparison were allowed to include up to 60% of patients who were receiving second or later line treatment. In their response to the ERG's request for clarification (A4), the manufacturer stated that a threshold of 60% was chosen "to ensure that studies in which a majority of patients were not treated in the first-line setting were excluded", in line with the NICE scope. However, one of the key studies included in the indirect comparison (Jones 2005[3]) did include a majority of patients who were not treated in the first-line setting (55%). If "a majority of patients" was traditionally defined *a priori* as a proportion greater than 50% (as the manufacturer did for HER2 negative patients, see paragraph above), then the Jones 2005 trial[3] would be excluded and, therefore, could not be used to link docetaxel in the indirect comparison.

As in the direct efficacy comparison, the MS excluded the AVADO trial [27-36]from the indirect comparison on the basis of having a dosing schedule unrepresentative of routine clinical practice in the NHS (docetaxel 100mg/m2 q3w for up to 9 cycles). The manufacturer stated that, in routine clinical first-line mBC, patients receive docetaxel at "75mg/m2 for a maximum of 6, or in exceptional cases 8, cycles" (MS, p.110). However, it should be noted that the Jones 2005 trial[3], which was included in the indirect comparison, also assigned patients to a dose of docetaxel 100mg/m2 q3w, continued until tumor progression, unacceptable toxicity or withdrawal of consent. The mean dose of docetaxel received in the Jones 2005 trial[3] was 95mg/m2 per cycle, for a median of 6 cycles (See Table 3). Though mean dose and median cycles received were not available for AVADO, it was reported that 85% of AVADO patients completed six cycles and only 47% completed nine cycles. Therefore, the ERG considered the decisions to include Jones 2005 [3]and to exclude AVADO on the basis of docetaxel dose to be inconsistent with one another.

Safety evaluation: To identify data on the safety of bevacizumab in combination with paclitaxel/docetaxel, the inclusion criteria were broadened to include non-randomised studies reporting data from more than 1,000 patients. To ensure relevant studies were not missed, the ERG requested details of the 266 excluded non-RCTs noted in Figure 1 of the submission. Among these were three safety studies excluded on the basis of sample size.[37-39] Two small

phase II studies (n=27 and n=76) provided safety data on bevacizumab plus docetaxel and a larger study (n=307) provided data on bevacizumab plus paclitaxel. Though they address relevant interventions and populations, the ERG does not feel these three studies would substantially add to the safety data already included in the MS (the ATHENA study,[14-20] see Section 4.2).

4.1.3 Table of identified studies. What studies were included in the submission and what were excluded

Direct efficacy comparison: A single open-label phase III RCT comparing qw paclitaxel versus qw paclitaxel plus bevacizumab for first-line treatment of locally recurrent or metastatic breast cancer (the E2100 trial[5-13]) was included. The AVADO[27-36] and RIBBON-1 [24-26]RCTs were excluded (see section 4.1.4).

Indirect efficacy comparison: As well as the E2100 trial (qw paclitaxel versus qw paclitaxel plus bevacizumab),[5-13] three further studies were included in the indirect comparison: Seidman 2008/CALGB 9840 [22](qw paclitaxel versus q3w paclitaxel), Albain 2008 [21](q3w paclitaxel versus gemcitabine plus paclitaxel), and Jones 2005[3] (q3w paclitaxel versus q3w docetaxel). The ERG had some concerns about how these trials were selected for inclusion in the indirect comparison (see section 4.1.2). The AVADO trial[27-36] (q3w docetaxel versus q3w docetaxel plus bevacizumab) was excluded from this analysis (see section 4.1.4).

Safety evaluation: Alongside comparative safety results from the E2100 trial[5-13], the findings of a large multi-centre single-arm open label study of bevacizumab plus taxane-based chemotherapy for first-line treatment of locally recurrent or metastatic breast cancer (MO19391, the ATHENA study[14-20]) were included.

4.1.4 Details of any relevant studies that were not included in the submission?

Direct efficacy comparison: As discussed previously, one large (n=736) double-blind RCT comparing docetaxel versus docetaxel plus bevacizumab (AVADO[27-36]) was excluded from the MS on the basis that the docetaxel dose in AVADO (docetaxel 100mg/m2 q3w for up to 9 cycles) was unrepresentative of routine clinical practice (stated by the manufacturer to be"75mg/m2 for a maximum of 6, or in exceptional cases 8, cycles"). However, the comparison in this trial (bevacizumab plus docetaxel versus docetaxel) is highly relevant to

the decision problem, particularly as NICE currently recommends single-agent docetaxel for first-line treatment of patients with advanced breast cancer who are not suitable for anthracyclines.[2] In relation to docetaxel dose in routine clinical practice, advice from the ERG's clinical expert indicated that the majority of patients receive 100mg/m^2 (60-70%) with remainder receiving a lower dose of 80mg/m^2 (30-40%), typically for six cycles. Mean dose was not reported for AVADO, but 85% of patients in the trial completed six cycles and only 47% completed nine cycles, so the median number of cycles is likely to be lower than nine. Clinical advice received by the ERG indicates that around 30% of UK patients treated under standard conditions may receive up to nine cycles of chemotherapy. The ERG considered AVADO to be relevant to this evaluation and present the available characteristics and results of this trial in section 4.2.

A second potentially relevant placebo-controlled RCT, the RIBBON-1 trial[24-26] evaluated bevacizumab in combination with chemotherapy (either capecitabine, anthracycline or taxane therapy) in patients with HER2-negative previously untreated locally recurrent or metastatic breast cancer. The manufacturer justified the exclusion of the RIBBON-1 trial on the basis of the potentially relevant subgroup (bevacizumab plus docetaxel versus placebo plus docetaxel) not being powered to detect a significant difference in treatment effect. Though this trial was relatively small, given the limited available evidence for this comparison, the available outcome data from RIBBON-1 are briefly summarized in section 4.2.

Indirect efficacy comparison: As discussed in section 4.1.2 the exclusion of the AVADO trial[27-36] from the indirect comparison on the basis of dose was inconsistent with the inclusion of the Jones 2005 trial.[3] Both studies assigned patients to a dose of docetaxel 100mg/m^2, with the mean dose of docetaxel received in the Jones 2005 trial being 95mg/m^2 per cycle, for a median of 6 cycles. Therefore the ERG considered that if the Jones 2005 trial[3] was included, then the AVADO trial should also be included.

For the comparison of qw versus q3w paclitaxel, a UK trial by Verrill 2007 (the 'Will Weekly Win' trial)[40] was excluded on the basis of reporting insufficient information on the proportion of first-line metastatic breast cancer patients and because the number of treatment cycles was limited (12 cycles for qw, six cycles for q3w paclitaxel; p115). Therefore an alternative trial making the same comparison (Seidman 2008/CALGB9840[22]) was included as it allowed treatment until progression, as in the E2100 trial.[5-13] Though the available data for Verrill 2007 were limited, it should be noted that the Seidman 2008 trial[22] also had potential limitations in terms of relevance and internal validity (see section 4.2).

Safety evaluation: Given the relatively large numbers of patients involved, the AVADO trial[27-36] could potentially have been used to provide additional safety data on docetaxel plus bevacizumab in the MS, just as the E2100 trial provided additional safety data on paclitaxel plus bevacizumab. However, the ERG accepts the manufacturer's assertion that large-scale ATHENA [14-20]provides the most comprehensive and robust safety database for the use of bevacizumab in mBC (clarification response A14).

4.1.5 Description and critique of manufacturers approach to validity assessment

Direct efficacy comparison: Though no published validity assessment tool appeared to have been used, the MS provided a narrative critical appraisal of the validity of the E2100 trial (MS, p.88-89).[5-13] The issues addressed were appropriate and included justification of sample size, length of follow-up, statistical methods used, randomization method, approaches to minimize potential bias due to lack of blinding, and similarity of patients and interventions to routine clinical practice. For further discussion of study quality, see section 4.2.

Indirect efficacy comparison: The MS did not provide a systematic assessment of validity for all studies included in the indirect treatment comparison. Instead, it briefly acknowledges flaws in two of the four included studies (p.121), specifically: (1) Seidman 2008[22] allowed for an imbalance of trastuzumab treated patients in the two arms, which may have subsequently biased the results and (2) Jones 2005[3] used a docetaxel dose of 100mg/m2 which the manufacturer considered to be substantially greater than that used in routine UK clinical practice. However, other aspects relating to the internal validity of these studies were not addressed (see section 4.2).

Safety evaluation: A brief narrative critical appraisal of the ATHENA,[14-20] similar to that provided for the E2100 trial was presented in the MS (p.139). For further discussion of study quality, see section 4.2.

4.1.6 Description and critique of manufacturers outcome selection

To a greater or lesser extent, the MS addressed each of the outcomes specified in the final scope issued by NICE.

Direct efficacy comparison: The E2100 trial[5-13] was powered to detect a difference in the primary outcome of progression free survival (PFS). Overall survival (OS), overall response rate (ORR) and duration of objective response were included as secondary outcomes. An unplanned analysis of 1-year survival was also presented. As data were not collected on therapy after disease progression in E2100, the OS analyses are limited by the fact that the potential impact of any post-progression crossover between groups or use of additional treatments is unknown.

Health-related quality of life (HRQoL) was assessed in the E2100 trial[5-13] using the FACT-B questionnaire, which was administered at baseline (within 2 weeks prior to starting treatment), week 17 and week 33. The MS focused on a subset of the FACT-B, the Trial Outcome Index (TOI-B). This measure excludes the emotional and functional well-being subscales of the instrument, though the total FACT-B score (TOT-B) was also reported. Several issues around loss to follow-up and imputation of HRQoL data are presented in section 4.2.

Indirect efficacy comparison: The outcome of interest in the manufacturer's indirect comparison was PFS. Two of the included studies[3, 22] reported time to progression (TTP) rather than PFS. However, since these studies included death as an event in their TTP calculations, these outcomes could be considered synonymous with PFS and were therefore employed in the indirect comparison. No other outcomes relating to response or survival were assessed.

Safety evaluation: The incidence of Grade 3-5 non-haematologic and Grade 4-5 haematologic adverse events were reported for each arm of the E2100 trial on a per-protocol basis. [5-13]The primary outcome of the uncontrolled ATHENA study was the incidence of serious adverse events (SAEs) and specific adverse events (serious and non-serious), including hypertension, proteinuria, arterial and venous thromboembolism, congestive heart failure, CNS bleeding, other haemorrhages, wound healing complications and gastrointestinal problems.[14-20]

4.1.7 Describe and critique the statistical approach used

Direct efficacy comparison and safety evaluation: The manufacturer simply provided summary results of the E2100 trial[5-13] for the evaluation of efficacy and safety, with an additional summary of the ATHENA study[14-20] for the evaluation of safety.

Indirect efficacy comparison: The manufacturer used the method described by Bucher[1] and Song[41] to compare alternative therapies where no head-to-head RCT was available. Where standard errors were not reported, estimates were obtained using the indirect method described by Tuder.[42] Details of these calculations were not provided in the MS, but were provided at the request of the ERG (clarification response A18).

The Bucher indirect comparison method is intended to compare two randomized trials via a common comparator (i.e. the relative effect of intervention A vs. C derived from one trial comparing A vs. B, and one comparing B vs. C). This is acknowledged by the manufacturer in their response to the ERG's clarification letter (response B1). Bucher[1] details several potential limitations of the method and notes that the strength of inference is limited for such comparisons. Sources of uncertainty in the indirect comparison presented in the MS included issues around internal validity (the imbalance in proportion of patients receiving second-line treatment in the qw (16%) and q3w (41%) paclitaxel arms of Seidman et al[22]) and exchangeability (the majority of patients in Jones 2005[3] (55%) had received previous chemotherapy for mBC, compared with 0% in E2100[5-13] and Albain 2008[21]) of the included trials.

In addition, the approach as used in the MS extends the Bucher method beyond its intended application. The manufacturer's alternative justification for excluding AVADO[27-36] from the indirect comparison (that the Bucher method is only intended to compare two trials via a common comparator) highlights limitations in their existing network, which extends the Bucher method to indirectly link three trials via two comparators. In an attempt to link qw paclitaxel plus bevacizumab with docetaxel via q3w paclitaxel, the manufacturer has incorporated three trials into the indirect comparison (E2100, [5-13]Seidman 2008,[22] Jones 2005[3]). Similarly, three trials (E2100[5-13], Seidman 2008, [22]Albain 2008[21]) were used to connect qw paclitaxel plus bevacizumab with gemcitabine plus paclitaxel. Given the limitations of both the Bucher method and the available evidence to make these comparisons, the ERG considered this approach (in effect comparing A vs. D) to be inappropriate. Therefore the existing evidence base is inadequate for any useful indirect comparison of these comparators.

4.1.8 Summary statement

Though the manufacturer's search strategies were appropriate and likely to have identified all the evidence relevant to the decision problem, the ERG had several concerns about how this evidence was selected and presented in the MS.

Firstly, the direct efficacy comparison includes only the open-label E2100 trial, [5-13] which evaluated the addition of bevacizumab to qw paclitaxel. A randomized double-blind trial comparing the addition of bevacizumab to docetaxel (the AVADO trial[27-36]) is also available, but was excluded by the manufacturer because (a) the dose was considered inappropriate compared to routine clinical practice and (b) bevacizumab in combination with docetaxel would be far from cost-effective for the NHS. However, the decision problem specifies that this comparison (bevacizumab plus docetaxel versus docetaxel monotherapy) should be considered. Regardless of potential cost-effectiveness or otherwise, the AVADO trial would provide information on the efficacy of adding bevacizumab to the taxane monotherapy currently recommended by NICE for first-line treatment of patients with advanced breast cancer who are not suitable for anthracyclines.[2] In addition, clinical advice to the ERG indicated that the dose used in AVADO is often used in routine clinical practice, and the uncertainties introduced by this issue are likely to be less than uncertainties that appear to have been accepted elsewhere in the MS.

The ERG identified several limitations and inconsistencies relating to the indirect efficacy comparison. One inclusion criterion (<60% second-line treatment for mBC) may have been formulated to allow the inclusion of a specific trial required for a link in the comparison (Jones 2005[3]). Though the AVADO trial[27-36] was excluded from the indirect comparison on the basis of docetaxel dose, the Jones 2005 trial[3] that used the same dose was included, suggesting an inconsistent application of inclusion criteria. The manufacturer's alternative justification for excluding AVADO from the indirect comparison (that the Bucher method is only intended to compare two trials via a common comparator) highlights limitations in their existing network, which extends the Bucher method to indirectly link three trials via two comparisons. Given these methodological limitations, combined with concerns about the validity and exchangeability of the included evidence, the ERG does not consider the findings of the indirect comparison to be reliable.

4.2 Summary of submitted evidence

4.2.1 Summary of results

4.2.1.1 Direct efficacy comparison

On the basis of a single open-label RCT, study E2100,[5-13] the manufacturer concludes that the addition of bevacizumab to paclitaxel chemotherapy provides substantial benefit to patients with mBC who had not previously received chemotherapy for advanced disease (MS, chapter 6.4). This was based on statistically significant increases in median progression-free survival (PFS), from 5.8 months to 11.3 months (Hazard Ratio (HR) 0.54, 95% CI 0.44 to 0.67) and in objective response rate from 22.2% to 49.8% ($p<0.0001$) for bevacizumab plus paclitaxel versus paclitaxel alone. Median overall survival was not significantly different between the two groups (26.5 vs. 24.8 months; HR 0.87, 95% CI 0.72 to 1.05). A post-hoc analysis analysis indicated overall survival at 1 year was significantly higher with paclitaxel plus bevacizumab than paclitaxel alone (81.4% vs. 74.0%, $p=0.017$).

For quality of life in the E2100 trial,[5-13] the manufacturer reported that at both week 17 and 33, the FACT-B trial outcome index (TOI-B) and the FACT-B total score (TOT-B) were noticeably better in patients receiving bevacizumab plus paclitaxel therapy compared with those receiving paclitaxel alone. The addition of bevacizumab to paclitaxel therapy was associated with a significant improvement in TOI-B score at week 33 ($p=0.0042$), and in TOT-B score at week 17 ($p=0.0475$) and at week 33 ($p=0.0046$) compared with paclitaxel alone.

Data from two additional Phase III RCTs, the AVADO, [27-36]and the RIBBON-1 trial[24-26] are not presented in the MS because they were considered to have limited relevance (see section 4.1.4). However, the ERG noted that these trials fall within the final scope issued by NICE (see section 4.1.4) and have summarised them in section 4.2.2.2 below.

4.2.1.2 Indirect efficacy comparison

The MS states that bevacizumab plus qw paclitaxel was associated with a significant improvement in PFS when compared with q3w docetaxel (HR 0.56, 95% CI 0.39 to 0.78), and with gemcitabine plus q3w paclitaxel (HR 0.46, 95% CI 0.34 to 0.64). No significant difference was found for PFS between qw paclitaxel and q3w docetaxel (HR 1.15, 95% CI 0.89 to 1.48), or between qw paclitaxel and gemcitabine plus q3w paclitaxel (HR 0.96, 95% CI 0.76 to 1.21).

4.2.1.3 Safety evaluation

In E2100,[5-13] the addition of bevacizumab to paclitaxel therapy was associated with a 20% overall increase in the incidence of Grade 3–5 adverse events, with an increase in Grade 3 hypertension and sensory neuropathy being the most common. There appeared to be no increase in the incidence of Grade 3–5 vascular thromboembolic events with the addition of bevacizumab to paclitaxel. However, grade 3 and 4 adverse events that were increased by ≥5% in patients treated with paclitaxel plus bevacizumab compared with those treated with paclitaxel alone were sensory neuropathy (24.2% vs. 17.5%), hypertension (16.0% vs. 1.4%), and fatigue (10.7% vs. 5.2%). There was a higher incidence of neutropenia, and infection/febrile neutropenia events in patients receiving bevacizumab in combination of paclitaxel .

The large uncontrolled ATHENA study[14-20] showed that the most frequent serious adverse events were febrile neutropenia (5.1%), neutropenia (3.6%), and pyrexia (1.5%). Bevacizumab was discontinued in 18.9% of patients due to adverse events, most commonly due to hypertension, fatigue and proteinuria. There appeared to be a lower incidence of Grade 3-5 hypertension (4.4%), proteinuria (1.7%), arterial and venous thromboembolism (3.4%) and coronary heart failure (0.5%) in ATHENA than in E2100.

The MS concludes that bevacizumab is not associated with the commonly recognised side-effects of cytotoxic anti-cancer therapies. The most common adverse events associated with bevacizumab therapy included hypertension and proteinurea.

4.2.2 Critique of submitted evidence syntheses

Table 1: Results of quality assessment for RCTs identified in the manufacturer's submission

	Study				
	E2100[5-13]	**Jones 2005[3]**	**Seidman 2008[22]**	**Albain 2008[21]**	**AVADO[27-36]**
Included in direct comparison	●				
Included in indirect comparison	●	●	●	●	
Quality assessment criteria					
Eligibility criteria specified?	Y	Y	Y	Y	Y
Power calculation?	Y	Y	Y	Y	Y
Adequate sample size?	Y	Y	Y	Y	Y
Number randomised stated?	Y	Y	Y	Y	Y
True randomisation?	Y	Y	Y	Y	NR
Double-blind?	N	N	NR	N	Y
Allocation of treatment concealed?	N	N	NR	N	NR
Treatment administered blind?	N	N	NR	N	Y
Outcome assessment blind?	Y	N	NR	N	Y
Patients blind?	N	N	NR	N	Y
Blinding successful?	N	N	NR	N	Y
Adequate baseline details presented?	Y	Y	Y	Y	Y
Baseline comparability?	Y	Y	N	Y	Y
Similar co-interventions?	Y	Y	NR	Y	Y
Compliance with treatment adequate?	NR	NR	NR	NR	NR
All randomised patients accounted for?	Y	Y	NR	Y	Y
Valid ITT analysis?	Y	Y	NR	Y	Y
$\geq$ 80% patients in follow-up assessment?	Y	NR	NR	Y	Y

Y=yes; N=no; NR=not reported

4.2.2.1 Direct efficacy comparison (E2100)

4.2.2.1.1 Trial design and quality

The submission's findings for the direct comparison evidence draw exclusively on the E2100 trial,[5-13] an open-label, multicentre, randomised, phase III trial comparing paclitaxel plus bevacizumab with paclitaxel alone as first line chemotherapy for patients with HER2-negative metastatic or locally recurrent breast cancer. The primary outcome was PFS, defined as the time from randomisation until the first date that recurrent or disease progression was objectively documented by the independent review facility, or until death within 84 days of the last protocol treatment. The censoring methods for the survival analysis data appeared appropriate.

The E2100 trial [5-13]included patients eligible for first line treatment of mBC, more than 90% of whom were HER2-negative; this was representative of the target population specified by the NICE scope.

The ERG noted that the results reported in the MS were derived from interim analyses; PFS and objective response results were derived from data collected prior to 9 February 2005 and overall survival data were collected prior to 21 October 2006. More recent follow-up data from this trial are likely to valuable, particularly for survival outcomes. However, in response to the ERG's request for clarification, the manufacturer stated that more recent analyses were not available (response A10).

In the E2100 trial,[5-13] randomisation, intention-to-treat efficacy analyses, and comparability of patients' baseline characteristics were adequate (see Table 1 for quality assessment). However, like many trials in mBC, E2100 employed an open-label design, without blinding for patients or investigators. Such designs can lead to an overestimation of treatment effect.[43] In view of this potential limitation, the trialists conducted a retrospective, independent and blinded review of the response and progression outcomes, and compared these with investigator assessed outcomes using the same statistical methodology. The results from both assessment methods were similar, suggesting that any assessment bias did not substantially impact on the findings of the trial.

4.2.2.1.2 Dosing regimen

In the E2100 trial,[5-13] the estimated overall dose of bevacizumab was 17.9mg/kg/cycle. The estimated overall dose of paclitaxel was 216.8 mg/m2 in the combination therapy arm and 240.9mg/m2 in the arm with paclitaxel alone. The estimated overall doses for both bevacizumab and paclitaxel appeared to be representative of the routine clinical practice. [4, 23] In terms of treatment cycles, the ERG has been advised that routine UK practice is usually use six to nine treatment cycles, with about 30% of UK patients receiving up to nine cycles of chemotherapy. It should be noted that more than half of patients in the combination therapy arm of E2100 and more than 20% of patients in the control arm (paclitaxel alone) received more than ten cycles (clarification response A7). Therefore, the paclitaxel monotherapy dosing regimen used in the E2100 trial was generally reflective of clinical practice, though the decision to treat until progression means that patients receiving combination therapy received considerably more cycles of paclitaxel than typically seen in current practice.

4.2.2.1.3 Health-related quality of life outcomes

Evidence on health-related quality of life (HRQoL) in the clinical effectiveness Section of the MS was drawn solely from the E2100 trial. [5-13]Quality of life was appropriately measured using the FACT-B, a validated instrument specifically designed for use in breast cancer patients.

There was a substantial loss to follow-up from baseline for the HRQoL measures over time (46% of patients in the paclitaxel alone arm versus 35% of patients in the combination therapy arm at week 33), though the manufacturer was unable to provide summary characteristics of those patients who were lost to follow-up in their submission or in response to the ERG's request for this information (clarification response A16).

As the MS did not discuss the clinical significance of the observed changes in HRQoL scores, the ERG requested baseline data for each measure, though the manufacturer could only provide this for TOI-B (see clarification response A15, table 4). The number of patients providing baseline TOI-B scores in this response differ from those presented in the MS (p.100-101), though the reasons for this are unclear.

For patients who were missing at week 17 or 33, the manufacturer imputed a score of zero (i.e. the worst score) for all of those who had died or progressed by the time of follow-up, and excluded patients who had missing scores for other reasons. This resulted in a statistically significant difference in both FACT-B measures favouring combination therapy at week 33 and TOT-B at week 17; however, these differences were not statistically significant in the analyses without imputed data. Given the use of an extreme assumption for patients who progress and the greater observed progression in the paclitaxel monotherapy groups, this approach to imputation may have biased the results in favour of paclitaxel plus bevacizumab.

Given these limitations and uncertainties, the manufacturer's conclusion that the addition of bevacizumab to paclitaxel leads to a relative improvement in HRQoL (MS, p.101) may not be reliable.

4.2.2.2 Additional direct evidence not included in the manufacturer's submission

The manufacturer identified three further RCTs that the ERG considered to be potentially relevant to the direct comparison evidence: AVADO,[27-36] RIBBON-1[24-26] and Will Weekly Win[40] (see section 4.1.4). The available characteristics and results from AVADO and RIBBON-1 are presented in Appendix 1 of this report. The limited available data from Will Weekly Win is included in Table 3.

4.2.2.2.1 The AVADO trial

AVADO[27-36] is a randomized, placebo-controlled, double-blind trial evaluating docetaxel in combination with bevacizumab 7.5mg/kg or bevacizumab 15 mg/kg compared with docetaxel

monotherapy in 736 patients with HER2–negative inoperable locally recurrent or mBC. The AVADO trial reported that, with a median follow-up of 10.2 months, a significant improvement in PFS was observed in the bevacizumab (7.5mg/kg) plus docetaxel group (HR 0.79, 95% CI 0.63 to 0.98) and the bevacizumab (15mg/kg) plus docetaxel group (HR 0.72, 95% CI 0.57 to 0.90), when compared with docetaxel alone. However, there were no significant differences in overall survival between either of the combination therapy groups and the docetaxel group.

In addition, an 'updated analysis' of PFS from the AVADO trial is reported in the bevacizumab SPC. It is stated that this analysis incorporates an additional 18 months of follow-up, though no other details are available. The updated analysis is less favourable to combination therapy than the planned analysis, with a non-significant difference in PFS between bevacizumab (7.5mg/kg) plus docetaxel versus docetaxel alone (HR 0.86, 95% CI 0.72 to 1.04). Bevacizumab (15mg/kg) plus docetaxel versus docetaxel alone remained statistically significant, though the magnitude of the difference between the groups was slightly smaller in the updated analysis (HR 0.77, 95% CI 0.64 to 0.93).

A summary of the main characteristics and results of the AVADO and E2100 trials are presented in Table 2.

In terms of the study quality in the AVADO trial (see Table 1), blinding of patients and investigators, and comparability of patients' baseline characteristics between the treatment groups were adequate. Intention-to-treat analysis was appropriately performed. This well-designed trial was generally protected against bias.

Table 2: Key characteristics and efficacy data from direct comparison bevacizumab RCTs (E2100 and AVADO)

	E2100[5-13]		**AVADO**[27-36]		
Participants	HER2-negative metastatic breast cancer not previously treated with chemotherapy (n= 722).		HER2-negative previosouly untreated locally recurrent or metastatic breast cancer (n= 736).		
Intervention	Bevacizumab 10mg/kg + paclitaxel 90mg/m^2; weekly		Bevacizumab 7.5mg/kg + docetaxel 100mg/m^2; 3-weeky Bevacizumab 15mg/kg + docetaxel 100mg/m^2; 3-weekly		
Comparator	Paclitaxel 90 mg/m^2; weekly		Placebo + docetaxel: docetaxel 100mg/m2; 3-weekly		
Length of follow-up for the analysis	Patients were enrolled between December 2001 and May 2004. Progression-free survival and objective response: Data collected prior to 9 February 2005. Overall survival: Data collected prior to 21 October 2006.		Patients were enrolled between March 2006 and April 2007. Progression-free survival: Primary analysis: median follow-up 10.2 months Updated analysis**:** Conducted at time of final OS analysis (additional 18 months of follow-up) Overall survival: Median follow-up 25 months		
	Paclitaxel (n=354)	**Bevacizumab+ Paclitaxel (n=368)**	**Docetaxel + Placebo (n=241)**	**Bevacizumab 7.5mg/kg + Docetaxel (n=248)**	**Bevacizumab 15mg/kg + Docetaxel (n=247)**
Median progression-free survival (months)	5.8	11.3	8.0	8.7	8.8
Updated analysis*	_	_	8.2	9.0	10.1
Progression-free survival, Hazard ratio (HR)(95% confidence interval)	_	0.48 (95% CI 0.39 to 0.61)	_	0.79 (95%CI 0.63 to 0.98)	0.72 (95% CI 0.57 to 0.90)
Updated analysis*	_	_	_	0.86 (95% CI 0.72 to 1.04)	0.77 (95% CI 0.64 to 0.93)
Response rate	22.2%	49.8%	44.4%	55.2%	63.1%
Updated analysis*	_	_	46.4%	55.2%	64.1%
Overall survival (HR) (95% confidence interval)	_	0.87 (95% CI 0.72 to 1.05)	_	1.05 (95%CI 0.81 to 1.36)	1.03 (95% CI 0.79 to 1.33)

* Updated analysis only applies to the AVADO trial

4.2.2.2.2 The RIBBON-1 trial

RIBBON-1[24-26] is a double-blind, placebo-controlled RCT that randomised HER2-negative previously untreated locally recurrent or mBC patients to either bevacizumab 15mg/kg or placebo, in combination with a previously selected chemotherapy agent (capecitabine, anthracycline or taxane).

Very limited data were available for the subgroup of 307 patients receiving taxanes. However, RIBBON- 1[24-26] reported a significantly higher objective response rate in the taxane plus bevacizumab group (n=203) compared with the taxane plus placebo (n=104) group (50.3% versus 35.3%, p=0.03) and median PFS was greater in the taxane plus bevacizumab group (8.2 versus 9.2 months; HR 0.75, 95% CI 0.56 to 1.01, p=0.0547). Overall survival was not reported.

Some taxane-treated patients received docetaxel and some nab-paclitaxel. The docetaxel subgroup fits with the scope of this evaluation (being the same comparison as AVADO[27-36]), but the ERG could not identify separate outcome data for this subgroup.

A formal quality assessment of the RIBBON-1 trial was not possible due to limited reporting in the located abstracts.

4.2.2.2.3 Will Weekly Win trial

Verrill et al[40] briefly reported the findings of a UK-based RCT ('Will Weekly Win)' comparing qw paclitaxel with q3w paclitaxel for locally advanced or metastatic breast cancer. This trial was excluded from the MS, but for the sake of completeness, the limited available details of this trial are presented in Table 3. Briefly, this RCT reported a statistically non-significant reduction in time to progression in favour of qw paclitaxel (HR 0.92, p=0.06).

4.2.2.3 Indirect efficacy comparison

A detailed critique of the selection and analysis of studies in the indirect efficacy comparison are presented in Sections 4.1.2 and 4.1.7 of this report. The quality of studies included in the indirect comparison was variable (see Table 1). Though the NICE scope specifies the population of interest to be HER2-negative mBC receiving first-line treatment, only the E2100[5-13] and Albain 2008[21] trials exclusively included first-line treatment metastatic breast cancer patients. The Jones 2005 trial[3] had 55% of patients receiving second line treatment and the Seidman 2008 trial[22] had 29% of patients receiving second line treatment. The ERG has been advised that patients receiving second line chemotherapy usually have worse

response rates and survival than those receiving first line treatment. Therefore, in the presence of any interaction between the different patient samples and the treatment effect, these differences in trial characteristics could potentially bias the estimates of relative treatment efficacy.[1]

The ERG further noted that the Seidman 2008 trial[22] had an imbalance in the proportion of patients receiving second line treatment between the treatment and control arm: 41% of patients in the q3w paclitaxel arm versus 16% of patients in the qw paclitaxel arm (see Table 3). This was due to a large proportion of the q3w patients having been 'imported' from an earlier study, 75% of whom had previous treatment for mBC. Since response rates to second-line chemotherapy are likely to be lower than for first-line, this imbalance between the arms may have compromised the internal validity of the trial. A biased estimate from this trial could, therefore, lead to biased estimates of relative efficacy in the indirect comparison analysis.

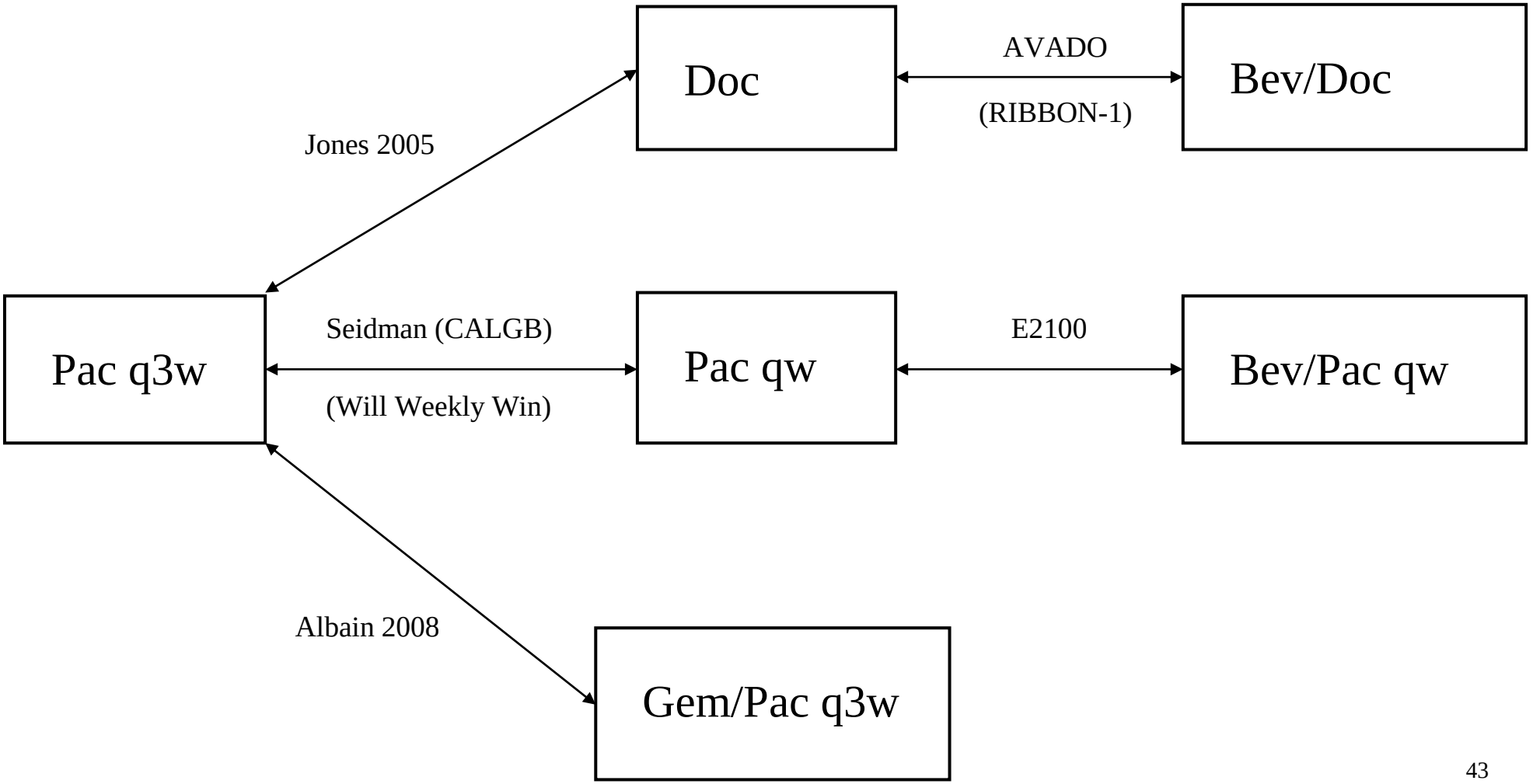
As in the direct comparison, two additional manufacturer-funded trials (AVADO[27-36] and RIBBON-1[24-26]) could theoretically have been incorporated in the indirect comparison network as conceived by the manufacturer in order to establish the relative treatment efficacy. Table 3 summarises all the identified trials potentially eligible for inclusion in the indirect comparison. Figure 1 is an adaptation of the network presented in Figure 8 (p.125) of the MS, including all the available comparisons. It should be noted that this is an illustration only, and that the ERG has major concerns about using this evidence network in any formal indirect comparison (see Section 4.1.7).

Table 3: Trials potentially eligible for indirect treatment comparison

Study	Comparison	Participants	Intervention	Comparator	Median PFS/TTP/TTF (months)	Study design	Included in MS ITC?	ERG Comments
E2100[5-13]	Bev/Pac qw vs. Pac qw	HER2-negative mBC not previously treated* with chemotherapy N (total)= 722 Bev+Pac: n= 368 Pac: n=354	**Bevacizumab + paclitaxel:** 10 mg/kg bevacizumab following paclitaxel treatment on weeks 1 and 3 of every cycle	**Paclitaxel:** 90 mg/m2 IV infusion over 1 hour every week for 3 weeks followed by 1 week of rest	**PFS** Bev/Pac: 5.8 Pac: 11.3	Open label RCT	Yes	
AVADO (BO17708)[27-36]	Bev/Doc vs. Doc	HER2-negative previously untreated locally recurrent (LR) or mBC N (total)= 736 Bev 7.5mg/kg + Doc: n= 248 Bev 15mg/kg + Doc: n= 247 Placebo + Doc: n= 241	**Bevacizumab 15mg/kgq3w + docetaxel** OR **Bevacizumab 7.5mg/kg q3w + docetaxel**	**Placebo + docetaxel:** 100mg/m2 q3w,for a maximum of nine cycles	**TTP** BEV7.5/DOC: 9.0 BEV 15/DOC: 10.1 DOC/Placebo: 8.2 **HR for PFS vs placebo** BEV7.5/DOC 0.86 (0.72 to 1.04) BEV15/DOC 0.77 (0.64 to 0.93)	Double blind RCT	No, wrong DOC dose	Clinical advice suggested that this dose is used in practice Patients receiving 9 cycles of Doc: BEV15/DOC: 51% Placebo/DOC: 42%
RIBBON-1[24-26]	Bev/Doc vs. Doc Bev/nab-Pac q3w vs. nab-Pac q3w	HER2-negative previously untreated LR or mBC N (total)= 1237 Within taxane/anthracycline cohort: Bev+ Doc: n=122 Placebo+Doc: n= 58	**Chemotherapy** (capecitabine, taxane or anthracycline selected prior to randomisation) **+ bevacizumab** 15mg/kg q3w	**Chemotherapy** (capecitabine, taxane or anthracycline selected prior to randomisation) **+placebo**	N/A	Double blind RCT	No, too few bev/taxane vs taxane pts	Published outcome data are not available separately for different taxanes
Seidman (CALGB 9840)[22]	Pac qw vs. Pac q3w	mBC; up to one line of previous chemotherapy allowed (29% (213/735) of patients had previous treatment for mBC) 40% (349/577) of HER-2 tested	**Paclitaxel 175mg/m2 q3w** until progression or unacceptable toxicity	**Paclitaxel 80mg/m2 qw** until progression or unacceptable toxicity	**TTP** Pac qw:9.0 Pac q3w: 5.0	Open label(?) RCT plus additional	Yes	41% (157/385) of 3-weekly patients 2nd line, compared with 16% (56/350) weekly, potentially biasing against 3-

		patients were HER-2 positive. N (total)=735 Weekly: n=350 3-weekly: n=385				control patients (q3w)		weekly treatment 40% (349/577) of HER-2 tested patients were HER-2 positive.
Verrill et al 2007 (Anglo Celtic IV/Will Weekly Win)[40]	Pac qw vs. Pac q3w	LR or mBC 569 patients	**Paclitaxel q3w 175 mg/m2** for 6 cycles (18 weeks)	**Paclitaxel qw 90 mg/m2** for 12 weeks	**TTP** Pac qw: 23.9 weeks Pac q3w: 22.0 weeks HR=0.92, p=0.06	RCT	No, insufficient details.	% 2nd line and HER-2 positive unknown. Longer treatment duration in q3w arm
Jones et al 2005[3]	Pac q3w vs. Doc q3w	Locally advanced or mBC that had progressed ofter previous chemotherapy or within 12 months of adjuvant/neoadjuvant chemotherapy.	**Paclitaxel** q3w 175 mg/m2, until disease progression or unacceptable toxicity	**Docetaxel** 100mg/m2 q3w, until disease progression or unacceptable toxicity	**TTP** Pac q3w: 3.6 Doc q3w: 5.7 HR 1.64 (1.33, 2.02) P<0.0001	Open label RCT	Yes	58% of Doc and 53% of Pac patients 2nd line for mBC. Doc dose may be similar to AVADO (mean dose 95mg/m2 per cycle; median 6 cycles)
Albain 2008[21]	Pac q3w vs. Gem/Pac	Patients with recurrent or mBC who relapsed after adjuvant anthracyclines	**Paclitaxel** q3w 175 mg/m2	**Paclitaxel** q3w 175 mg/m2 + **Gemcitabine** 1,250 mg/m2	**TTP** Pac q3w: 3.98 Gem/Pac: 6.14 HR 0.70 (0.59, 0.85) **PFS** Pac q3w: 5.9 Gem/Pac: 3.9 HR 0.73 (0.61, 0.87)	Open label RCT	Yes	

Figure 1: Revised network of the indirect treatment comparison showing all relevant comparisons identified in the manufacturer's submission



4.2.2.4 Safety comparison

The evidence in the MS for the safety of bevacizumab was drawn from two studies: a RCT (E2100)[5-13] and a non-randomised study (ATHENA).[14-20]

4.2.2.4.1 Study design and quality

The study quality of E2100[5-13] is discussed in section 4.2.2.1 of this report. Safety data were analysed on the basis of the treatment received and included all patients who received any amount of protocol therapy. The ERG noted some inconsistencies in the reported number of patients withdrawing and/or analysed between the CONSORT flow charts (MS p.82, and clarification response A2) and the safety outcome table (MS Table 22, p.132), though it appears these are unlikely to have substantially changed the findings of the trial.

The ATHENA study is a multicentre, non-randomised, single-arm, open-label study.[14-20] It evaluated the safety of bevacizumab in combination with taxane-based chemotherapy as first-line treatment in 2,251 patients with locally recurrent or metastatic breast cancer, with 97.2% of patients being HER2-negative. The MS states that this sample size allowed the study to evaluate the occurrence of rare adverse events with a rate of less than 1%. Though all uncontrolled studies are susceptible to confounding,[44] the adverse event rates reported in ATHENA are likely to be representative of those observed in clinical practice. Median follow-up was 12.7 months, so data on any potential longer-term adverse events of bevacizumab/taxane combination treatment are not available.

4.2.2.4.2 Dosing regimen

The dosing regimens of both bevacizumab and paclitaxel used in the E2100[5-13] were considered to be representative of the UK current practice (see section 4.2.2.1). In the ATHENA study,[14-20] patients received bevacizumab 10mg/kg every 2 weeks, or 15mg/kg every 3 weeks, on the basis of the recommendations in the SPC. The dosing of the taxane-based chemotherapy was based on the standard of care in the participating institutions. Therefore, the dosing regimen used in the ATHENA study appeared to be representative of UK clinical practice.

4.2.3 Summary

The MS evaluation of clinical efficacy of bevacizumab was primarily based on a single paclitaxel RCT (E2100) in which the patient sample was considered to be representative of the target population defined by the NICE scope, and the dosing regime was generally reflective of UK clinical practice.

The E2100 trial demonstrated that, compared with paclitaxel alone, the addition of bevacizumab to qw paclitaxel roughly halved the risk of disease progression. This represents a significant absolute improvement in median time to progression from 5.8 months to 11.3 months. Though there were limitations (e.g. a lack of blinding) in the methodological quality of E2100, reasonable attempts were made to minimise the potential for bias in data collection and analysis.

In terms of quality of life, a significant improvement in FACT-B score was observed in the bevacizumab plus paclitaxel group in the E2100 trial. However, due to methodological issues associated with loss to follow-up and imputation methods for missing data, this finding may not be reliable.

The manufacturer excluded a large high-quality RCT (AVADO) evaluating q3w docetaxel in combination with two alternative doses of bevacizumab or placebo; a trial in which the docetaxel dosing regimen may be applicable to routine UK current practice. Therefore, the ERG examined clinical efficacy data from both E2100 and AVADO trials. Relative to E2100, a smaller effect size was observed for the addition of bevacizumab to a taxane in AVADO. The AVADO trial revealed that, compared with docetaxel and placebo, the addition of bevacizumab to docetaxel was associated with a significant absolute improvement of median PFS from 8.0 months to 8.7 (bevacizumab 7.5mg/kg) or 8.8 months (bevacizumab 15mg/kg). In a poorly reported ‘updated analysis’, this PFS benefit for bevacizumab 7.5mg/kg became statistically insignificant.

A potentially relevant RCT, the RIBBON-1 trial, was identified but did not report separate outcomes for patients receiving docetaxel.

The available trial evidence indicates that the addition of bevacizumab to taxane treatment can be an efficacious first-line treatment for patients with HER2- negative mBC, in terms of prolonged PFS. No trial reported a benefit of combination therapy on overall survival. The

magnitude of PFS benefit observed for combining bevacizumab with qw paclitaxel (E2100) was larger than that seen for bevacizumab plus q3w docetaxel (AVADO), though the results these studies were not formally compared.

In the absence of a head-to-head comparison on the relative efficacy between different therapies, an indirect comparison was performed to estimate the relative efficacy between these therapies in the MS. This stated that bevacizumab plus qw paclitaxel was associated with a significant improvement in PFS compared to both q3w docetaxel (HR 0.56, 95% CI 0.39 to 0.78), and gemcitabine plus q3w paclitaxel (HR 0.46, 95% CI 0.34 to 0.64). However, due to differences between patient populations both within and between selected studies and the use of inappropriate statistical methods, the estimates derived from the indirect comparison may not be reliable.

The limited available evidence on the evaluation of safety suggests that bevacizumab is not associated with the commonly recognised side-effects of cytotoxic anti-cancer therapies. Serious adverse events were rare, but the most common adverse events associated with bevacizumab treatment include hypertension, proteinurea and sensory neuropathy. The longer-term safety profile of bevacizumab remains unknown.

5 ECONOMIC EVALUATION

This section focuses on the economic evidence submitted by the manufacturer in their initial report. The submission is subject to a critical review on the basis of the manufacturer's report and by direct examination of the electronic version of the economic model. The critical appraisal is conducted with the aid of a checklist to assess the quality of economic evaluations and a narrative review to highlight key assumptions and possible limitations. Section 6 presents a description of the additional information provided by the manufacturer following ERG points of clarification and a critique of this by the ERG, alongside additional work undertaken by the ERG to address any remaining uncertainties.

5.1 Overview of manufacturer's economic evaluation

This section provides a structured critique of the economic evaluation reported in the MS to NICE. The manufacturer's initial economic submission included:

1. A description of the databases and websites searched in the literature review (MS, p.149-151). A description of the systematic search strategy used to identify existing cost-effectiveness studies for bevacizumab in the first line treatment of mBC with full details in a separate Appendix (MS, p.298, Appendix 3).

2. A report on the *de novo* economic evaluation conducted by the manufacturer. The report described the technology; comparators and patient population; the categories of resource use costed; the resource use and unit cost assumptions and sources; the base-case results; and sensitivity analysis (MS, p.151-221).

3. An Excel-based model comprising the manufacturer's electronic economic model.

An overall summary of the manufacturer's approach and signposts to the relevant sections in the MS are reported in Table 4.

Table 4: Summary of the manufacturer's economic evaluation (and signposts to MS)

	Approach	**Source / Justification**	**Signpost (location in MS)**
Model	Cost effectiveness and cost utility analysis using a decision model based on a Markov model	MS justifies that mBC is a chronic disease and that progression in the disease can be described by movement between few discrete health states.	*Sections 7.2.6.1 to 7.2.6.8 p.162*
States and events	The model included three health states: progression-free survival, progressed, and dead.	The states defined mirrored the endpoints of the E2100 trial. MS states that the structure reflects a very	*Sections 7.2.6.1 to 7.2.6.8 p.166*

		common structure used in oncology.	
Comparators	The intervention evaluated is BEV+PAC. The model makes pairwise comparisons with: PAC qw, DOC and GEM+PAC.	Several interventions in the NICE scope were not included (PAC q3w, BEV+DOC). BEV+DOC was stated a-priori by the manufacturer not to be cost-effective, and PAC q3w was stated to be less effective than PAC q1w.[c]	*Sections 7.1.2.1 to 7.1.2.2 p.152*
Natural History	Based on the three state Markov model.	See justification provided under states and events. Data derived from extrapolating patient level data from the E2100 trial.	*Sections 7.2.6.1 to 7.2.6.8 p.166*
Treatment effectiveness	The model assumes treatments to impact only on the rate of occurrence of disease progression. Remaining transitions, survival for progressed patients and death in progression free survival, do not depend on treatment. It is assumed that all comparators are equally effective.	Progression free survival was informed by the E2100 trial only and the indirect treatment comparison was only used to justify the assumption of equality of treatment effects. Remaining transitions were also mainly derived from E2100 trial.	*Sections 7.2.6.9 to 7.1.2.2 p.168*
Adverse events	Adverse events were included based on incidence observed in the E2100 trial. For DOC, an increased incidence of febrile neutropenia was assumed, based on TA162 [45].	Events with fewer than 3% frequency in the E2100 trial were excluded, with the exception of febrile neutropenia.	*Section 7.2.8.5 p. 185*
Health related QoL	Derived from two studies identified through a non-systematic literature review. Within these studies the standard gamble technique was used to derive utility weights.	Although health related QoL had been collected in the E2100 trial, the measure used was disease specific and thus inadequate to inform utility scores	*Section 7.2.8.3 p. 180*
Resource utilisation and costs	The following cost categories were considered in the manufacturer analyses: drug acquisition costs, drug administration costs, duration of treatment, supportive care costs, adverse event costs and end of life costs.	The data sources used included UK reference costs, published literature, clinical expert opinion and the E2100 trial. The unit cost of drugs was based on NHS list prices. However, a second base case used PASA prices to cost paclitaxel and assumed a 10g cost cap for BEV	*Section 7.2.9 p. 185*
Discount rates	A 3.5% discount rate was employed for both costs and health benefits	In accordance with the NICE reference case approach.	*Section 7.2.10 p. 197*
Sub groups	Subgroup analysis not undertaken	The MS states that the E2100 trial was underpowered to study outcomes on patient subgroups.	*Section 7.2.2.2 p. 159*
Sensitivity analysis	Detailed scenario analysis and probabilistic sensitivity analysis (PSA) undertaken	Not all uncertain parameters are defined. Cost effectiveness plane and CEACs provided.	*Section 7.2.11 p. 197*

BEV+PAC - bevacizumab in combination with paclitaxel; PAC qw – one weekly paclitaxel monotherapy regimen; DOC - docetaxel monotherapy; GEM+PAC - gemcitabine plus paclitaxel; PASA - Purchasing and Supply Agency, NHS

The manufacturer conducted a literature search to identify published cost-effectiveness studies (CEA) for bevacizumab in the treatment of mBC. No attempt was made to search for CEA that only considered the comparator treatments. An overview of the search was described in the MS (p.149) and detailed search strategies are presented separately in Appendix 3 of the MS (p.298).

[c] Wording amended by ERG on advice of Roche

The databases searched for the cost effectiveness section included all of those defined by NICE in the specification for manufacturer/sponsor submission of evidence: MEDLINE, MEDLINE In-Process, EMBASE, HEED and NHS EED.

The submission gave detailed descriptions of the search strategies and met NICE requirements. It included the specific databases searched; the service providers used; the dates when searches were conducted; the date spans of the searches; and the complete strategies used. The number of records identified for each search set were included for the MEDLINE, MEDLINE In-Process and EMBASE searches, but not for the HEED and NHS EED searches.

The strategies and thesaurus terms used in each database varied but were appropriate to each database searched. The filter used to identify study types in the searches were appropriate to each database searched. However, the use of economic search terms in the HEED and NHS EED search strategies appears unnecessary, given that these sources contain only economic evaluation/cost study references. The ERG considered the search strategy for Section 7 (cost-effectiveness) to be appropriate.

The submission (MS p151, Table 27) identified six cost-effectiveness analyses, but stated that they were not relevant as they were all conducted outside the UK. However, the ERG considered that, in the absence of any published UK studies, these may still be of interest to the Appraisal Committee, and thus asked for clarifications on the main methods and results of any full economic evaluations (i.e., that compare both costs and outcomes of two or more relevant interventions). Only one publication fulfilled these criteria (Dedes 2009).[46] The analysis is based on the E2100 trial and evaluates the cost effectiveness of bevacizumab plus paclitaxel compared to paclitaxel alone. This was conducted from the Swiss healthcare perspective. A Markov model was defined where relevant health states were: progression-free, disease progression and death. The base-case analysis reported an ICER of €189,427 per QALY gained (where prices relate to 2008).

Following the literature search, the manufacturer developed a *de novo* economic model. A description of the assumptions and structure of this model is detailed next.

5.1.1 Interventions and comparators

The intervention evaluated in the manufacturer's model is bevacizumab 10mg/kg (every 2 weeks) in combination with paclitaxel 90mg/m^2 (weekly for 3 weeks followed by 1 week of

rest), abbreviated throughout as BEV+PAC. The model makes a series of pairwise comparisons with this intervention, with the following comparators:

- PAC qw - paclitaxel (monotherapy) 90mg/m² weekly for 3 weeks followed by 1 week of rest (Study E2100[5-13]).
- DOC - docetaxel (monotherapy) 75 mg/m² on day 1 and every 21 days. The manufacturer considers this regimen to be current UK NHS clinical practice.
- GEM+PAC - gemcitabine 1,250mg/m² days 1 and 8 plus paclitaxel 175 mg/m² on day 1 every 21 days (Albain et al. 2008[21]).

With BEV+PAC and PAC qw therapies, the treatment is assumed to be continued according to the protocol of the E2100 trial. That is, patients are assumed to remain on treatment until the first of one of the following events:

- Disease progression,
- Unacceptable toxicities, or
- Refusing further treatment

The docetaxel SPC does not state that the therapy should be administered for a fixed number of cycles, but the manufacturer reports that UK expert opinion and market research data show that it is usually administered for six cycles and no more than 9 cycles. Hospital sales data from IMS show that the average docetaxel treatment for metastatic breast cancer in UK is 6.13 cycles q3w, at an average dose of 150mg (or 88mg/m^2 for an average 1.7m^2 patient). An assumption was made in the manufacturer's model of treatment until disease progression or a maximum of 6 months (or approximately 8.7 cycles) of treatment. When accounting for the rate of disease progression, the average time on treatment in the model was 4.86 months, equating to 7.0 cycles of treatment. Using a dose of 75 mg/m^2 every 3 weeks, the manufacturer states this was considered a reasonable representation of UK clinical practice.

For gemcitabine + paclitaxel dosed as per SPC, it was assumed that time on treatment would be similar to that of qw paclitaxel monotherapy and, therefore, the comparator arm curve for 'duration of treatment' generated from E2100 was used as a proxy for this comparator's time on treatment.

Other comparators and interventions that are potentially relevant were excluded from the evidence synthesis and model. These were:

- Bevacizumab in combination with docetaxel (BEV + DOC). The manufacturer stated (without any explicit analysis) that this was not cost-effective.
- Paclitaxel monotherapy q3w. The manufacturer stated weekly paclitaxel is more effective and less toxic than three-weekly paclitaxel

5.1.2 Natural history

The patient cohort within the economic evaluation is assumed to have the same baseline characteristics as those observed in the E2100 trial (Section 4.2.1.1). The mean age was 56 years, and patient weight was assumed to be 70kgs and Body Surface Area to be 1.7m^2.

The model has three health states: progression-free survival (PFS), progressed and dead (MS Sections 7.2.6.1 to 7.2.6.8). The cycle length in the model was one month and the overall time horizon considered was 10 years. Patients start in the progression-free state. During each discrete monthly cycle of the model patients who are progression-free can remain progression-free, die of any cause, or the metastatic disease can progress. For patients whose metastatic disease has progressed, at each cycle they can either remain in the progressed state or die of any cause. The "progressed" health state represents the time period from first treatment relapse until death and. Therefore. includes the possible sequence of remission and relapse of second and following lines of treatments common to metastatic breast cancer. The transition probabilities in the model were estimated from the E2100 trial.

5.1.3 Treatment effectiveness within the submission

The model incorporates estimates 5 sets of parameters related to the effectiveness of the treatments:

1. The probability of disease progression during each month for patients who are progression-free on BEV+PAC
2. The probability of disease progression during each month for patients who are progression-free on PAC qw
3. The probability of death from any cause during each month for patients who are progression-free on BEV+PAC
4. The probability of death from any cause during each month for patients who are progression-free on PAC qw
5. The probability of death from any cause during each month after progression of metastatic disease, for all initial therapies.

In each case, parametric survival analysis was used to estimate these probabilities. The parametric approach requires an assumption to be made about the shape of the hazard function in each case. The advantage of parametric survival analysis is that it allows the hazard function to be extrapolated beyond the follow-up period of the RCT. A non-parametric survival function would not be able to estimate mean life expectancy if patients survive beyond the end of the RCT.

The probability of disease progression during each month for patients who are progression-free at the start of the month (BEV+ PAC and PAC qw)

This was estimated from a parametric survival function for 'progression-free survival' using the data from the E2100 trial.

In response to a request from the ERG for clarifications, the manufacturer defined PFS as follows (MS clarifications A17):

"PFS is defined as the time from randomization until the first date that recurrent or progressive disease was objectively documented by the Independent Review Facility (IRF) or death within 84 days of the last study treatment. For patients who did not have disease progression or death by 9 February 2005, PFS was censored at the date of their last tumor assessment in the IRF reviewed database (or if no tumor assessments were performed after the baseline visit, at the time of randomization plus 1 day). Data for patients who died after the data cutoff date of 9 February 2005 without progressive disease (PD) were censored at the last tumor evaluation date before the cutoff date. Data for patients who died before the cutoff date but after 84 days following the last treatment date were censored at the last tumor evaluation date. Data for patients who receive non–protocol-specified cancer therapy prior to experiencing documented disease progression were also censored at the time of the last tumor assessment prior to receiving non–protocol-specified cancer therapy. Data for patients with no scans or pertinent medical information submitted to the IRF were censored at the randomization date".

The manufacturer compared several parametric functions or the rate of progression, and the Gompertz was found to have the best fit with the observed data, as measured by Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). The hazard of progression, $h_p(t)$, at month t for patients on BEV+PAC and PAC qw monotherapy was estimated by the manufacturer from the E2100 data (Table 5).

Table 5: Parameters of Gompertz function for progression-free survival (Source: Excel spreadsheet of MS)

	Estimate	*StdErr*
Intercept (μ)	-2.519	0.09
BEV	-0.617	0.11
Shape (γ)	0.053	0.01

Note: $h_p(t) = exp(\mu+BEV)exp(\gamma t)$

The exponential of the intercept parameter is the rate of events at the start of the study, which are exp(-2.519) = 0.080 per month in the PAC qw arm and exp(-2.519-0.617) = 0.043 per month in the BEV+PAC arm. A characteristic of the Gompertz function is that the rates of events are assumed to increase exponentially with time. This is seen by the positive value of the shape parameter. The rate of events (in both arms) doubles every 13 months, seen by the expression exp(0.053x13) = 2.

The predicted hazards from the Gompertz hazard function are shown in Figure 2. The analysis assumes proportional hazards - that is, at all time points, the rate of events in the BEV+PAC arm is 46% lower than in the PAC qw arm, seen by 1-exp(-0.617) = 0.46. The assumption of proportional hazards implies that the difference in the absolute rate of events between the treatments is also increasing exponentially.

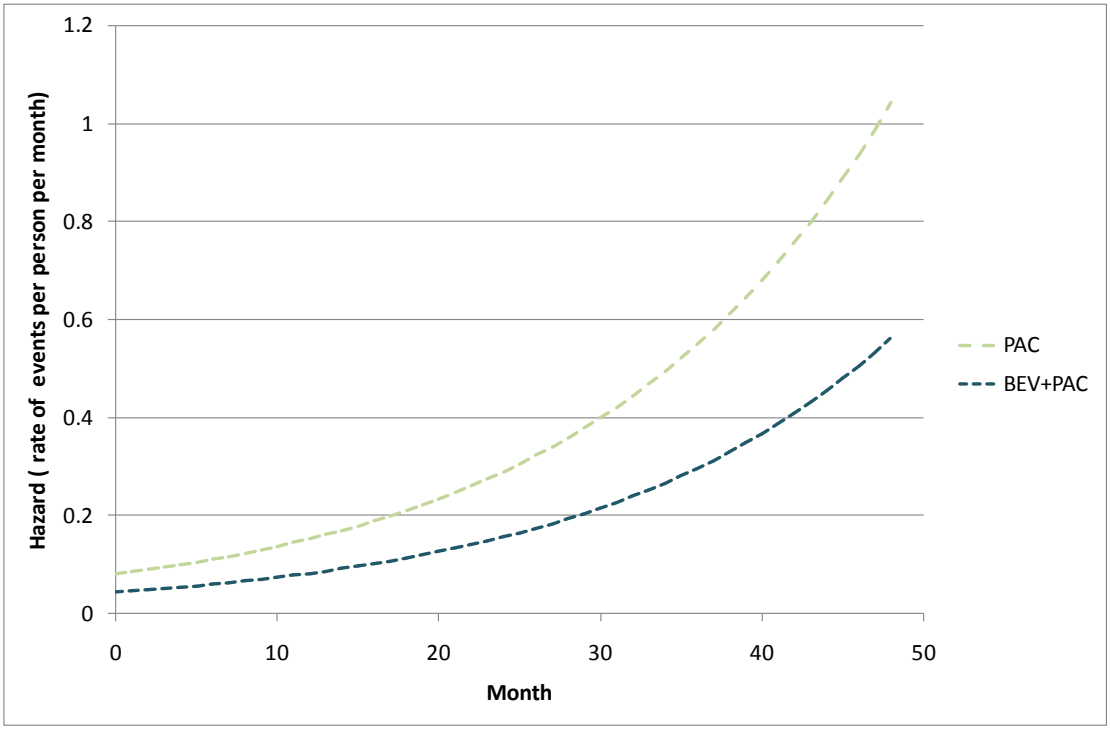


Figure 2: Predicted rate of death or progression of disease during each month for patients who were progression-free at the start of the month

The model estimates the proportion of the original cohort which is alive and progression free at time PFS(t) using the Gompertz formula

$$PFS(t) = \exp(\exp(\mu+BEV)/\gamma)(1-\exp(\gamma t)))$$

Follow up in the E2100 trial was up to 4.5 years. The model estimated that all patients receiving PAC qw would have progressed or died by 4 years (that is, PFS(month=48)≈0), which is consistent with the trial.

The rate of progression from the PFS state was assumed to be the same after PAC qw as after DOC and after GEM-PAC. The overlapping confidence intervals estimated by the indirect treatment comparison and the Sparano 2008[47] study were used to support the assumption that the E2100 paclitaxel qw arm was a reasonable proxy for the docetaxel monotherapy and for gemcitabine in combination with paclitaxel comparisons.

The probability of death from any cause during each month for patients who are progression-free on BEV+PAC and PAC qw

The definition of a 'failure' event in the survival function PFS(t) includes either death within 84 days of the end of treatment or progression of disease. Therefore, in order to estimate the proportion of the cohort which is in the 'progressed' state at time t, an adjustment is made in the Excel model to take account of the proportion which dies in the progression-free state without registering progression of disease.

This rate of mortality (for any cause) without progression of disease was estimated by the manufacturer from the E2100 data in each arm (Table 6).

Table 6: Monthly rate of deaths without progression of disease (Source: Excel spreadsheet of MS)

	BEV+PAC	PAC
Nr of PFS Deaths	19	21
PFS Person-Months	2764	1769
Monthly Rate of PFS Deaths	0.007	0.012

In the model, the rate of mortality from the progression-free state was assumed to be at least as great as the underlying sex and age-related mortality in the general population. The rate of mortality from the PFS state was assumed to be the same after PAC qw as after DOC and after GEM+PAC.

The probability of death from any cause during each month after progression of metastatic disease, for all initial therapies.

An important element of the model was the assumption that the hazard of death after progression is constant over time and the same across all treatments. This hazard rate was estimated by the manufacturer from the E2100 trial data to be 0.043 deaths per patient-month (SE 0.00033). The number of events and the number of patient-years at risk used to determine this estimate were not stated. It was not reported if alternative assumptions were considered or if other survival functions (with non-constant hazards) were fitted and compared.

The manufacturer justified the assumption that the hazard of death after progression was constant across treatments because a log-rank test for differences between BEV+PAC and PAC qw in the E2100 trial data was non-significant at the 5% level (p=0.2441).

5.1.4 Adverse events

The costs and disutility associated with treatment-related adverse events were included in the model. Adverse events were included based on incidence in the E2100 trial (see Section 4.2.1.3). According to the Excel model spreadsheet, in the PAC qw arm, the adverse events assumed to affect costs in the trial were: peripheral sensory neuropathy (affected about 6% of patients) and hypersensitivity (2%). The adverse events in the BEV+PAC arm that affected costs were: hypersensitivity (3% of patients), hypertension (4%), infection (6%), peripheral sensory neuropathy (2%) and febrile neutropenia (1%). It is not clear why these proportions differ from those reported for E2100 in the MS. Only peripheral sensory neuropathy and febrile neutropenia were assumed to have an impact on utility. It was assumed that the docetaxel adverse events would be equivalent to those for paclitaxel with the exception of an increased incidence of febrile neutropenia (60% of patients). The rate of febrile neutropenia for docetaxel in lung cancer (in TA162[45]) was assumed to apply to docetaxel in breast cancer.

5.1.5 Health related quality of life

The manufacturer reported that, although health related quality of life had been collected in the E2100 trial (see Section 4.2.1.1), the measure used was disease specific and thus inadequate to inform utility scores (MS, p.184). The estimates applied in the model were, therefore, derived from external data sources from a non-systematic literature review of studies of patients with breast cancer. A summary of the results of six main studies found was

presented in the MS (Table 37, p.181). The manufacturer points out that all the studies found used the standard gamble technique to derive utility weights.

In the base case analysis, the manufacturer uses one of the identified studies (Winstanley 2009[2]/ Cooper 2003[48]), and does not justify this choice or make any attempt to combine the estimates derived from the different studies. This study used data extracted from other published studies where quality of life was valued by oncology doctors and nurses, rather than by the cancer patients themselves, using the standard gamble.[48] The manufacturer provides no further details on the methods used and the ERG considers these to be unclear. These data consider the disease progression status of the patients (progression free or progressed disease), and not the treatment received. Table 7 shows the utility estimates assumed in the base case. Note that the manufacturer assumed the utility associated with progression free survival time to be an average of the utility associated with the response and stable disease health states.

Table 7: Base-case analysis utility scores

Health state	***Base case utility score***	***Source***
Response	0.81	Winstanley 2009 [2]
Stable disease	0.65	Winstanley 2009[2]
Progression-free survival	0.73	Winstanley 2009[2]
Progressive disease	0.45	Winstanley 2009[2]
Disutility from febrile neutropenia	-0.21	Winstanley 2009[2]
Disutility from peripheral sensory neuropathy	-0.21	Brown 1998 [49]

The manufacturer's analysis also applied a decrement to reflect the disutility of patients suffering adverse events. However, only febrile neutropenia and peripheral sensory neuropathy were considered (see Section 5.1.4). The disutility associated with these events was also derived from the same non-systematic literature search described above, and the base case estimates used are also detailed in Table 7.

The estimates obtained were applied to time spent in the corresponding health states of the decision model.

5.1.6 Resources and costs

The cost analysis was conducted from an NHS and PSS perspective and considers only resources relevant to the management of the disease. The year of pricing was 2008 (MS p. 196).

The manufacturer states that resource use was not collected within the E2100 trial (with the exception of adverse event rates and duration of treatment) and thus these data were identified from published sources (MS p.186, p.194). The following broad areas were considered in the manufacturer's analyses: drug acquisition costs, drug administration costs, duration of treatment, supportive care costs, adverse event costs and end of life costs. Expert opinion was also used in establishing assumptions regarding costing procedures, for which no data were available. The different areas are described below.

5.1.6.1 Drug acquisition costs

Costs associated with drug acquisition were calculated from assumed drug use and not from patient level data of the available RCTs.

The doses and frequency of administration of the different treatments constituting the alternative regimens assumed are described in Section 5.1.1. An average adult dose was calculated from these by assuming the characteristics of an average patient (weight of 70kg and a body surface area of $1.7m^2$). Wastage was incorporated in the evaluation by rounding up this dose to the nearest vial size. A monthly cost was calculated by using the frequency of administration of infusions (as described in Section 5.1.1) and the unit cost of vials of the drugs.

The unit costs of drugs was based on NHS list prices obtained from the British National Formulary. However, a second base case considered PASA (Purchasing and Supply Agency, NHS) prices to cost paclitaxel (Table 8). The PASA price represents the average (weighted arithmetic mean) price paid for a specific generic product over the last four months in the NHS hospital-sector (English trusts), and was derived from the Generic Pharmaceuticals Electronic Market Information Tool (eMIT, http://www.pasa.nhs.uk/PASAWeb/Productsandservices/Pharmaceuticals/Medicines/Generic/eMIT.htm).

This second base case also implements a capping scheme for BEV. Under this scheme, if a patient continues treatment after receiving a total dose of 10g of bevacizumab, the NHS does not pay any further acquisition costs beyond the cap. However, following clarification with NICE, the ERG understands that the capping scheme has not been approved by the Department of Health and, therefore, should not form part of any base case analysis. Since the submitted model considers a constant dose per unit of time (the monthly total dose of BEV for an average patient as considered in the model was 1400mg), the capping scheme was

implemented by assuming that no costs were incurred for the acquisition of BEV from the 7th month of treatment onwards.

Table 8: Unit costs for the treatments

Treatment	*Description of vial*	*Costs (£)*	*Source*
BEV	100mg vial	242.66	BNF 58 [50]
	400mg vial	924.40	
PAC	150mg vial (list price)	300.52	BNF 58 [50] (non-propriety)
	150mg vial (PASA price)	25.28	average PASA price *
DOC	80mg	534.75	BNF 58 [50]
GEM	1 gram for gemcitabine	162.76	BNF 58 [50]

* derived from the Pharmaceuticals Electronic Market Information Tool; average price paid for paclitaxel weighted by use over the last four months of the period ending April 2009.

The monthly acquisition cost assumed for each of the regimens is shown in Table 9. Monthly costs were applied to the duration of treatment, estimated from the E2100 trial as described below (and p. 153-159 of the MS).

Table 9: Monthly acquisition cost for the recommended dose of the treatments composing the regimens of interest

Regimen	*Treatment*	*Monthly Cost*
BEV	Bevacizumab	£3,592
BEV + PAC	Bevacizumab	£3,592
	Paclitaxel	£1 176 (BNF price used); £98.93 (PASA price used)
GEM + PAC	Gemcitabine	£1 038
	Paclitaxel	£871 (BNF price used); £73 (PASA price used)
DOC	Docetaxel	£1 150.14

PASA: Pharmaceuticals Electronic Market Information Tool; average price paid for paclitaxel weighted by use over the last four months of the period ending April 2009.

In practice, although patients start treatment with a chemotherapy regimen on a certain dose of its constituents, this dose is commonly reduced if, for example, patients experience toxicity. In most of the trials evaluated in this submission (e.g. E2100 and Jones 2005) dose reductions were defined in the protocol for patients experiencing adverse events. In the submission, the manufacturer did not apply dose reduction in calculating acquisition costs, and justified this with the lack of comparable data across the different trials.

5.1.6.2 Drug administration costs

Drug administration costs for the alternative regimens include the location of delivery (e.g. day-case), hospital pharmacist time for drug preparation, pre-medication costs and response assessment costs.

The MS assumes that all treatments are administered in a day-case setting. The drug administration costs were assumed to correspond to the HRG reference cost category described as 'deliver more complex parenteral chemotherapy at first attendance' [£237, National Schedule of Reference Costs (NSRC) 2007/08]. Each infusion of each treatment within a regimen was assumed to take 15 minutes of a pharmacist time to prepare (derived from expert opinion). The cost of a hospital pharmacist (salary based costs disregarding costs incurred with further qualifications gained) per hour was assumed to be £28 [Personal Social Services Research Unit (PSSRU) 2008].[51]

Patients were also assumed to require premedication whilst on treatment with docetaxel and paclitaxel (as recommended in the SPC). For treatment with docetaxel and paclitaxel, premedication was estimated to cost £3.34 and £4.08, respectively (a detailed description is provided in the MS, p.189 and 190).

Patients on each treatment regimen were also assumed to have an assessment of response to therapy every three months costing £71.88. In calculating this cost a consultation with a clinician (Reference costs – Clinical Oncology, consultant led: follow-up attendance non admitted face to face visit) and a CT scan (Reference costs – RA12Z – Computerised Tomography Scan, two areas, with contrast, outpatient) were assumed.

5.1.6.3 Duration of treatment

Drug acquisition and administration costs were applied to the duration of treatment by multiplying these costs by the proportion of patients on treatment in each cycle of the model.

For the regimens BEV+ PAC and PAC alone, the manufacturer evaluated the duration of treatment by using data from the E2100 trial (MS p.153). In this trial, participants were recommended treatment until unacceptable toxicity, progressive disease or loss to follow up. Duration of treatment was defined as the time from randomisation until censoring or experiencing an event. An event was assumed when participants: did not complete the protocol therapy due to disease progression; died due to the disease; had been taken off drug prior to disease progression due to unacceptable toxicities; or refused further treatment whilst not yet experiencing disease progression. Patients were censored if they were still considered progression free and on the protocol specified study drug at the time of the data cutoff (21 OCT 2006), or they died for other than disease related reasons.

Parametric survival functions (see MS Section 7.2.6.9, p.156 and 157) were used to model duration of treatment, analogously to the modelling of progression free survival (see Section

5.1.3). When fitting parametric distributions, duration of treatment with bevacizumab was modelled separately to duration of treatment with paclitaxel. The model evaluating paclitaxel assumed proportional hazards to distinguish duration of treatment with paclitaxel when this treatment was used alone or in conjunction with bevacizumab. Alternative parametric distributions were used (Gompertz, Weibull, Exponential, Log Logistic, and Log Normal), and the Weibull was chosen as the best fitting one (goodness of fit was assessed by AIC and BIC) in the two models implemented (results in MS Table 31, p.158). The manufacturer estimated the proportion of patients on treatment [survival $S_p(t)$] at time *t* (in months) for patients on bevacizumab to be:

$$S_p(t) = \exp(-\lambda t^{\gamma})$$

where the values of the parameters λ and γ were treatment specific (results shown in MS Table 31, p.158). The assumption of duration of treatment being described by a Weibull distribution implies that the mean and median duration of drug use are as described in Table 10.

Table 10: Mean and median time on drug assumed by describing duration of treatment using Weibull distributions.

Regimen	***Treatment***	***Duration of treatment,*** *months*	
		mean	median
BEV+PAC	Bevacizumab	7.323	6.098
	Paclitaxel	6.651	5.465
PAC	Paclitaxel	4.846	3.982

For the regimen GEM+PAC, time on treatment was assumed equal that of PAC. To describe duration of treatment with DOC the manufacturer assumed patients receive a conservative dose of 75 mg/m2 every 3 weeks until disease progression or a maximum of 6 months of treatment. This implies an average time on treatment of 4.86 months (equating to 7.0 cycles of treatment). For further details see the MS (p.158 and 166).

5.1.6.4 Supportive care costs

Supportive care costs are described in the MS (p. 191-192) and summarised in Table 11.

Table 11: **Supportive care costs**

Description	***Costs***	**Unit cost source**
PFS on treatment [2]		
2 nurse home visits per month,	£23 per visit	Community nurse per home visit without qualifications [51]
1 GP visit (including direct care staff) per month	£46 per visit	General practitioner unit cost including direct care staff costs without qualifications [51]

1 clinical nurse visit per month	£73 per visit	Clinical nurse specialist per hour with patient [51]
Total costs per month	*£165*	
PFS off treatment [2]		
1 consultation with a specialist every 2 months	£86 per visit	Clinical Oncology - consultant led: follow-up attendance non admitted face to face visit (National Reference Costs 2007/2008)
Total costs per month	*£42.81*	
Progressed [2]		
4 nurse home visit per month	£23 per visit	Community nurse per home visit without qualifications [51]
4 visits with a clinical nurse per month,	£73 per visit	Clinical nurse specialist per hour with patient [51]
2 GP home visit per month,	£50 per visit	General practitioner unit cost including direct care staff costs without qualifications home visit [51]
and 2 therapist home visits per month	£40 per visit	NHS therapist 1 hour home visit without qualifications [51]
Total costs per month	*£564*	

The manufacturer assumed that the costs incurred per unit of time whilst in the progressive state are independent from the first line treatment. Total costs incurred after progression do, however, depend on the time patients remain in the progressive state.

5.1.6.5 Adverse events (AE) costs

The management costs associated with AEs are given on p.193 of the MS. These costs were derived from national reference costs and expert opinion, and are summarised in Table 12 (Table 44 in the MS).

Table 12: Unit costs of adverse event assumed in the MS and source of unit costs (Table 44, MS)

Adverse event	*Cost per event*	*Source*
Febrile neutropenia	£3803	NHS reference costs 2008/2009 – PA45Z
Hypersensitivity	£274	NHS reference costs 2008/2009 – WA17X
Hypertension	£367	Coon 2008 [52]
Infection	£243	NHS reference costs 2008/2009 – WA09W
Peripheral Neuropathy	£0	Expert Opinion

In the submission, these unit costs were multiplied by the probability of events (described in Section 5.1.4), to generate an overall cost associated with AEs. Within the Markov model, the manufacturer incorporated AE costs as a one-off cost incurred at time zero (i.e. discounting was not applied).

For the regimens PAC and GEM+PAC, treatment related AEs increased mean total costs by £9. In the BEV+PAC regimens, adverse events increased the mean costs by £108. The model also assumed an overall increase in mean costs of £332 for the DOC monotherapy regimen.

5.1.6.6 End of life costs

The cost associated with the management of a patient at the end of their life (£3,805), was derived from the literature (MS p.193). This cost represents the average cost weighted for the setting (assuming a proportion of 0.4 die in hospital costing £4,706, 0.1 in a Marie Curie hospice costing £5,867 and 0.5 die at home costing £2,428), updated for inflation. In the model, the manufacturer multiplied this cost by the proportion of deaths occurring in each cycle of the model.

5.1.7 Discounting

The manufacturer's model applied a discount rate of 3.5% per annum to expected costs and health effects (MS p.197), in line with the NICE reference case.

5.1.8 Subgroup analyses

The manufacturer did not conduct subgroup analysis as part of the cost-effectiveness modelling (MS p. 159). The manufacturer justifies this by stating that the E2100 trial, the main source of clinical information used to evaluate cost effectiveness, was underpowered to study outcomes on patient subgroups.

5.1.9 Sensitivity analyses

Scenario analysis and probabilistic sensitivity analyses (PSA) were undertaken by the manufacturer. The terms scenario analyses and one way (or univariate) sensitivity analyses were used interchangeably. In practice, the MS evaluated the impact of using alternative assumptions or of using alternative values for a specific parameter (or for a set of parameters) in these analyses. Throughout we will refer to these as scenario analyses.

5.1.9.1 Scenario analyses

A range of alternative scenarios were used to explore the implications of distinct model assumptions and of the use of alternative sources of data. These are summarised in Table 13 and are described in full in p.197-204 of the MS.

Table 13: Specification of the scenario analyses performed by the MS

Brief identification (signpost to MS)		***Description of assumptions in base case***	***Description of assumptions in scenario analysis***	***Main impact on***
A	Parametric function used to describe time to disease progression (MS p.197)	The Gompertz distribution was used,	Alternative distributions were evaluated: Weibull, Exponential, log Logistic, log Normal and generalized Gamma distributions	LYG, QALY
B	Costs associated to continuing treatment until	Treatment was assumed to stop as	Treatments were assumed to be administered until progression.	Costs

	progression (MS p.200)	observed in the E2100 trial		
C	Body surface area (BSA) and weight (MS p.201)	Patients assumed to weight 70kg and have 1.7 m^2 of BSA	Two scenarios were tested: weight of 60 kg and BSA of 1.6 m^2 and weight of 80 kg and BSA of 1.8m^2	Costs
D	Utility of progression free patients (MS p.201 and 202, Tables 46-49)	Unweighted average of the utility of patients with stable disease and of patients that are responsive to therapy	Weighted average of the utility of patients with stable disease and of patients that are responsive to therapy, where the weights were given by the proportion of patients in these categories observed in the E2100 trial (for PAC and BEV+PAC). For DOC and GEM+PAC, the weights were derived by using the relative proportions observed in trials included in the indirect comparison.	Utilities
E	Source of utility values (MS p.202 and 203, Tables 50-51)	Winstanley (2009) [2]	Lloyd (2006) [no reference provided by the manufacturer]	Utilities
F	Supportive care costs (MS p.203)	-	Scenarios evaluated the impact of considering supportive care costs 50% higher and 50% lower	Costs
G	PFS costs, Progressive disease costs, end of life costs (MS p.203 and 204)	See description in Section 5.1.6	Disease management before disease progression was assumed to involve only one CT scan and one consultant visit every three months. The cost of progressive disease was assumed to be £771 (from the published literature) and the cost at end of life was assumed to be £1,503 (Remak et al. 2004 [53], inflated to 2008 prices).	Costs

5.1.9.2 Probabilistic sensitivity analysis

The main sources of parameter uncertainty considered in the PSA were: (i) the estimates of the parameters of the distributions characterising time to event; (ii) utilities for PFS and progression; (iii) monthly supportive care costs; (iv) adverse event costs; (v) drug administration costs; and (vi) end of life costs. This is described in the MS, p. 204 and 205.

When describing the uncertainty over some multi-parametric distributions, such as the Gompertz fitted to PFS, the correlation between parameters was considered (Electronic model submitted by the manufacturer). For other distributions fitted to the data, such as duration of treatment, uncertainty was not considered in PSA.

Table 14: Model parameter values assumed in the base case and assumptions used in the PSA, in the MS

Model Variable	*Base case value*	*PSA (MS p.204)*
Transition Probabilities (tp)		
PFS to Progressed	Gom(α_i,β), i= treat	The Choleski matrix was used to maintain the correlation between the parameter estimates
PFS to death	Max(age-specific mortality, monthly death rate in PFS)	Methods are unclear
Progression to death	Cte hazard of dying	Monthly death rate was assumed exponential. Rate parameter calculated from restricted mean. Unclear method of characterising uncertainty.
Patient characteristics		

Age	55.5	Not considered uncertain
Weight	70 kg	Not considered uncertain
Body Surface Area	1.7 m^2	Not considered uncertain
Costs		
Supportive-care costs		
Monthly PFS - Background care	£165	*Distribution*: BetaPert; *Range*: 115.5 to 214.5
Monthly PFS - Assessment of response	£72	Not considered uncertain
Monthly PFS - After therapy	£42.81	*Distribution*: BetaPert; *Range*: 29.97 to 55.56
Monthly Progressed - Background care	£564	*Distribution*: BetaPert; *Range*: 394.8 to 733.2
Monthly Progressed - Last 14 days of life	£3,804.59	*Distribution*: BetaPert; *Range*: 3043.67 to 4565
Monthly Drug costs		
bevacizumab	£3,592	Not considered uncertain
docetaxel	£1550	Not considered uncertain
paclitaxel weekly	£1176 / £99	Not considered uncertain
paclitaxel every 3 weeks	£871 / £73	Not considered uncertain
gemcitabine	£1038	Not considered uncertain
Duration of treatment		
PAC	Weib(α_i,β), i= regimen (PAC alone or BEV+PAC)	Considered uncertain[d]
BEV	Weib(α_{BEV},β_{BEV})	Considered uncertain[e]
Monthly administration costs		
bevacizumab + paclitaxel weekly	£896	Methods unclear
paclitaxel weekly	£881	Methods unclear
docetaxel	£430	Methods unclear
gemcitabine + paclitaxel every 3 weeks	£795	Methods unclear
Adverse event costs		
Febrile neutropenia	£3803	*Distribution*: BetaPert; *Range*: 3042 to 4564
Hypersensitivity	£274	*Distribution*: BetaPert; *Range*: 219 to 329
Hypertension	£367	*Distribution*: BetaPert; *Range*: 184 to 276
Infection	£243	*Distribution*: BetaPert; *Range*: 194 to 292
Peripheral Neuropathy	£0	
Utilities – values		
PFS	0.73	*Distribution*: Beta(0.73*1000, (1-0.73)*1000)
Progressed	0.45	*Distribution*: Beta(0.45*1000, (1-0.45)*1000)
Febrile Neutropenia	-0.21	Not considered uncertain
Peripheral Sensory Neuropathy	-0.21	Not considered uncertain
Discount rates		
Costs	3.5%	Not considered uncertain
QALYs	3.5%	Not considered uncertain

5.1.10 Model validation

The MS reports that internal validation and debugging of the model was performed by an independent consultant company. The manufacturer reported to have conducted tests involving extreme values of parameters (for assessment of consistency) and checks on the

[d] Amended by ERG following response from Roche

[e] Amended by ERG following response from Roche

completeness of reported results. It is unclear to the ERG, however, how these latter checks were conducted.

The MS did not include a validation of the model results against the values observed in the clinical trials that informed the model.

5.2 Critique of approach used

The manufacturer conducted a systematic search of the economic literature and identified no relevant prior studies on the cost-effectiveness of treatments for mBC in a UK setting. As such, the submission of a *de novo* economic evaluation was appropriate. However, one economic evaluation identified was conducted from the perspective of the Swiss Health Service. In the absence of any existing UK studies, the ERG considered the exclusion of this study to be a potentially important omission from the MS given it was based on the results of the E2100 trial and used a similar model structure to the manufacturer in their submission to NICE.

The MS presented a cohort Markov model constructed in Excel. The ERG has assessed the manufacturer's economic evaluation using Philips *et al.'s* checklist for quality assessing decision analytic models. [54] This is shown in Appendix 1 and is used to assist the narrative critique in the following sections. In Table 15, the methods used in the MS are also compared to those detailed in the NICE reference case.

Table 15: **A consideration of the MS using a checklist based on NICE's reference case and other methodological recommendations, together with an indication of the inclusion of each of the elements in the MS and ERG's comments on whether the *de-novo* evaluation meets the requirements of NICE reference case**

Elements of the economic evaluation	***Reference Case***	***Included in submission***	***Comment on whether de-novo evaluation meets requirements of NICE reference case***
Comparator(s)	Therapies routinely used in the NHS, including technologies regarded as current best practice	No	Not all relevant interventions (as defined by the scope) have been compared in the model. BEV+PAC, DOC, PAC qw and GEM+PAC were considered, whereas PAC q3w and BEV+DOC were not.
Type of economic evaluation	Cost-effectiveness analysis	Yes	
Perspective on costs	NHS and PSS	Yes	NHS and PSS costs have been taken into account.
Perspective on outcomes	All health effects on individuals	Yes	QALY benefits to treated individuals were considered.
Time horizon	Sufficient to capture differences in costs and outcomes	Yes	The economic model has a time horizon of 10 years.

Synthesis of evidence on outcomes	Systematic review	??	Although a systematic review and indirect comparisons were performed to evaluate PFS benefits, results were not used when populating the decision model used to inform cost effectiveness. Other input parameters were seldom based on a systematic search and synthesis of evidence.
Measure of health effects	QALYs	Yes	
Source of data for measurement of HRQL	Reported directly by patients and/or carers	No	The model uses utility values from the literature. The base case uses a study where the standard gamble was applied to oncology nurses in order to directly value changes in HRQL.
Source of preference data for valuation of changes in HRQL	Representative sample of the public	No	The model uses utility values from the literature. The base case uses a study where oncology nurses directly valued health states (not possible to assess representativeness due to incomplete reporting).
Discount rate	Annual rate of 3.5% on both costs and health effects	Yes	
Equity weighting	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	Yes	No special weighting was undertaken.
Sensitivity analysis	Probabilistic sensitivity analysis	Yes	Probabilistic sensitivity analysis and scenario analysis were undertaken. Results are presented graphically using cost-effectiveness planes and acceptability curves.

Abbreviations: HRQL, health related QoL; NHS, National Health Service; PSS, personal social services; QALYs, quality-adjusted life years

The ERG identified a number of shortcomings with the manufacturer's model.

5.2.1 Interventions and comparators

Not all relevant comparators and interventions have been compared in the model. Omitting relevant comparators from the analysis can lead to biased ICERs if the intervention is not being compared to the next best alternative, and also to an incorrect characterisation of the uncertainty surrounding the decision.

Table 16: Summary of Product Characteristics and comparison with the drug dose and continuation rule as used in the model

	Indication	*Dose*	*Treatment continuation rule*
Bevacizumab in combination with paclitaxel (SPC)	Bevacizumab in combination with paclitaxel or docetaxel is indicated for first-line treatment of patients with metastatic breast cancer.	The recommended dose of bevacizumab is 10 mg/kg of body weight given once every 2 weeks or 15 mg/kg of body weight given once every 3 weeks as an intravenous infusion.	It is recommended that treatment be continued until progression of the underlying disease. Dose reduction for adverse events is not recommended. If indicated, therapy should either be permanently discontinued or

			temporarily suspended
Model	Bevacizumab in combination with paclitaxel Bevacizumab in combination with docetaxel is not included	BEV: 10 mg/kg of body weight given once every 2 weeks PAC: 175mg/kg once every 3 weeks	Mean number of cycles observed in BEV+PAC arm of E2100 trial
Docetaxel monotherapy (SPC)	Docetaxel monotherapy is indicated for the treatment of patients with locally advanced or metastatic breast cancer after failure of cytotoxic therapy. Previous chemotherapy should have included an anthracycline or an alkylating agent.	Docetaxel is administered as a one-hour infusion every three weeks. For the treatment of patients with locally advanced or metastatic breast cancer, the recommended dose of docetaxel is 100 mg/m2 in monotherapy. In first-line treatment, docetaxel 75 mg/m2 is given in combination therapy with doxorubicin (50 mg/m2).	In patients who experienced either febrile neutropenia, neutrophil < 500 cells/mm3 for more than one week, severe or cumulative cutaneous reactions or severe peripheral neuropathy during docetaxel therapy, the dose of docetaxel should be reduced from 100 mg/m2 to 75 mg/m2 and/or from 75 to 60 mg/m². If the patient continues to experience these reactions at 60 mg/m², the treatment should be discontinued.
Model		75 mg/m2 once every 3 weeks	Until disease progression or a maximum of six months
Paclitaxel monotherapy (SPC)	As a single agent, treatment of metastatic carcinoma of the breast in patients who have failed to respond adequately to standard treatment with anthracyclines or in whom anthracycline therapy has not been appropriate.	Second-line chemotherapy of breast carcinoma: The recommended dose of paclitaxel is 175 mg/m2 administered over a period of 3 hours, with a 3-week interval between courses.	Patients who experience severe neutropenia (neutrophil count <0.5 x 109/l for a minimum of 7 days) or severe peripheral neuropathy, should receive a dose reduction of 20% for subsequent courses
Model		90mg/m² weekly for 3 weeks followed by 1 week of rest	Mean number of cycles observed in weekly PAC arm of E2100 trial
Gemcitabine, in combination with paclitaxel (SPC)	Gemcitabine, in combination with paclitaxel, is indicated for the treatment of patients with unresectable, locally recurrent or metastatic breast cancer who have relapsed following adjuvant/neoadjuvant chemotherapy. Prior chemotherapy should have included an anthracycline unless clinically contraindicated.	Gemcitabine in combination with paclitaxel is recommended using paclitaxel (175 mg/m2) administered on Day 1 over approximately 3-hours as an intravenous infusion, followed by gemcitabine (1250 mg/m2) as a 30-minute intravenous infusion on Days 1 and 8 of each 21-day cycle.	Dose reduction with each cycle or within a cycle may be applied based upon the grade of toxicity experienced by the patient.
Model		As SPC	Mean number of cycles observed in weekly PAC arm of E2100 trial

Bevacizumab in combination with docetaxel was not considered as a comparator in the manufacturer's model. The manufacturer stated, without formal analysis, that bevacizumab with docetaxel is more costly than bevacizumab with paclitaxel and has a similar health outcome, and therefore would be dominated. However, this approach does not permit

sensitivity analyses using alternative estimates of relative treatment effects (or costs). Furthermore, even if interventions are dominated in terms of the mean ICER estimates, they should still be incorporated within a PSA to correctly characterise the overall decision uncertainty.

Table 16 compares the SPC with the dosing in the model for each therapy. Docetaxel monotherapy is included in the model at a dose of 75mg/m^2. The manufacturer has excluded studies (principally AVADO Chan, 2009 #195; Cortés, 2009 #194; Dirix, 2009 #192; Fumoleau, 2009 #196; Greil, 2009 #190; Harbeck, 2009 #198; Miles, 2009 #199; Miles, 2008 #197; Pivot, 2009 #191; Wardley, 2009 #193} comparing BEV+DOC to DOC) that used a regimen of docetaxel 100 mg/m2, stating that this dose is inconsistent with UK clinical practice (see Section 4.1.2). However, The SPC recommends 100 mg/m², and the clinical advisor to the ERG stated that docetaxel 100mg/m^2 is used in UK clinical practice particularly in younger and fitter patients. Therefore, the ERG believes that studies using this dose should be included in the evidence synthesis and analysis.

The SPC for paclitaxel monotherapy recommends a dose of 175mg/m² every 3 weeks. The model dose is 90mg/m^2 weekly for 3 weeks followed by 1 week of rest. The manufacturer justifies the decision not to include q3w paclitaxel as a comparator in the model, stating that recent studies (Seidman 2008 [22]; Sparano 2008 [47]) have indicated that qw paclitaxel is more effective than q3w paclitaxel and the E2100 trial used this more effective administration regimen. The clinical advisor for the ERG has confirmed that qw paclitaxel is less toxic and would generally be used in older or frailer patients, but q3w paclitaxel remains a relevant comparator.

5.2.2 Natural history

The model has good internal validity in the prediction of PFS compared with the E2100 trial (see Section 5.2.4). However, the model has poor internal validity in the prediction of overall survival (see Section 5.2.4). This may be because some strong assumptions are being made in the model about mortality following progression.

The model structure assumes that the rate of death after progression of metastatic disease is constant over time and is the same for all initial treatments. That is, the differences between initial treatments are only modelled with respect to PFS and, once patients progress, they are assumed to face a common mortality rate irrespective of the initial treatment. The implication of this is that any differences in mean PFS between treatments are assumed to be maintained

in the mean OS estimates. Alternative model structures were not considered by the manufacturer.

One possible alternative structure is an Area Under the Curve (AUC) approach. This takes as inputs the estimated parametric survival functions for both overall survival and progression free survival. The proportion of the initial cohort which is in the progressed state at any point in time is then the difference between the proportion that is alive and the proportion that is progression-free. The mean life-years of the cohort with progression of metastatic disease is the area between these two curves. This approach allows the model to be accurately calibrated with the trial results for both PFS and OS. The model can be used to compare treatments that were not evaluated in the E2100 trial if a proportional hazards assumption is valid (or if there is no difference in the effectiveness of the comparators). If the proportional hazards assumption is not valid, then the separate survival functions would need to be fitted to each therapy. The information required to do this would not be available from summary data of other trials and, therefore, the model could not easily evaluate other treatments. The ERG has undertaken further analyses using the AUC approach, using the direct evidence comparing BEV+PAC with PAC qw, and incorporating the indirect evidence linking to the other possible comparators and interventions (PAC q3w, DOC and BEV+DOC) assuming proportional hazards. The approach employed and results are discussed in detail in Section 6.2.

Another alternative modelling approach would be to set up a series of tunnel states after progression to allow the rate of death after progression to be time-dependent.

5.2.3 Treatment effectiveness within the submission

The main assumptions made in the model about treatment effectiveness are:

- The relative treatment effects of progression free survival are proportional hazards (the hazard ratios are constant over time)
- The rate of death following progression is constant over time and the same for all initial treatments
- The effectiveness of all the comparators (docetaxel monotherapy and gemcitabine, in combination with paclitaxel) is assumed the same as qw paclitaxel monotherapy

The assumption of proportional hazards, together with the Gompertz survival function, implies that the difference in the absolute rate of progression between the treatments is increasing exponentially over time in the model (see Figure 2). This may be a strong

assumption and should be tested. The MS did not state if alternative assumptions about the hazard ratio were explored. The manufacturer presented data justifying the proportional hazards assumption in clarifications for the ERG (MS clarifications 16 April 2010).

The internal validity of the model can be checked by comparing the median survival time for PFS and overall survival found by the E2100 trial with the model predictions. Although the mean survival time is more relevant than the median for cost-effectiveness analysis, the trial did not provide estimates of mean survival time as extrapolation would have been required.

Table 17: Comparison of model predictions with E2100 trial

		Mean PFS *(months)*	***Median PFS*** *(months)*	***Mean overall survival*** *(months)*	***Median overall survival*** *(months)*
Model prediction	PAC	8.2	6.5	28.0	23
E2100 trial estimate	PAC	N/A	5.8	N/A	24.8
Model prediction	BEV+PAC	12.5	11	32.2	28
E2100 trial estimate	BEV+PAC	N/A	11.3	N/A	26.5
Model prediction	Difference	4.3	4.5	4.2	5
E2100 trial estimate	Difference	N/A	5.5	N/A	1.7

N/A: E2100 did not estimate mean survival. PFS: progression free survival

The estimates of PFS correspond closely for the model and the trial (Table 17). The median PFS is predicted by the model to be 6.5 months with PAC and 11 months with BEV+PAC, a difference of about 4.5 months. These results are similar to the E2100 trial results for PFS, which found median PFS was 5.8 months with PAC and 11.3 months with BEV+PAC, a difference of 5.5 months.

However, the model overestimates the difference in *overall survival* for BEV+PAC versus PAC compared with the E2100 trial. The model predicts a mean difference in overall survival between PAC and BEV+PAC of 0.35 years, or 4.2 months. The median overall survival is predicted by the model to be 23 months with PAC and 28 months with BEV+PAC, a difference of about 5 months. This difference is considerably greater than the overall survival gain estimated by the E2100 trial, in which median survival was only improved by 1.7 months, from 24.8 months with paclitaxel alone to 26.5 months with paclitaxel + bevacizumab. The reason for the discrepancy in overall survival between the model and trial is likely to be because the model assumes the gain in progression-free survival from the trial is maintained and translated into an equivalent gain in overall survival. As previously noted, this assumption was justified by manufacturer by the lack of a statistical treatment difference reported between BEV+PAC or PAC alone based on a log-rank test (p=0.2441) from the stratified comparison of the Kaplan-Meier curves from E2100 of patients in the progressive health state. However, the lack of a statistically significant difference does not imply that the

hazard rates are equivalent. Importantly, no sensitivity analysis was presented of alternative assumptions about mortality after progression of disease by the manufacturer. Given the importance of this assumption and the poor internal validity of the subsequent model results, the ERG considers that the alternative structures discussed in the previous section should have been explored by the manufacturer.

The base-case model assumes that the regimens paclitaxel, docetaxel and gemcitabine+paclitaxel are equally effective. No alternative scenario was presented about the relative effectiveness of these comparator treatments. The ERG considers that an alternative exploratory approach to have considered within the scenarios presented by the manufacturer would have been to conduct the cost-effectiveness analysis using the results of the indirect comparison evidence synthesis in the model (and propagating uncertainty through the model using probabilistic sensitivity analysis).

5.2.4 Adverse events

The MS spreadsheet shows the costs and QALY decrements associated with adverse drug-related events (AEs). In the PAC arm, it was estimated that AEs diminished total QALYs by an average of -0.0186 and increased costs on average by £9. In the BEV+PAC arm, AEs diminished QALYs by -0.0083 and increased costs by £108.

It was assumed that the costs and disutility of AEs following GEM+PAC would be the same as PAC. It was estimated that all AEs in DOC monotherapy (at dose of 75mg/m2) would diminish total QALYs by an average of -0.1985 and increased costs on average by £332. On average, these effects of AEs on costs and HRQoL are relatively modest. These assumptions seem reasonable given that patients will either withdraw or reduce dose following a severe adverse event, and the effect of these changes of therapy on effectiveness are reflected in the mean hazard ratios estimated by the E2100 trial.

5.2.5 Health related quality of life

Evidence on HRQoL from the E2100 trial (described in Section 4.2.2.1.3, and MS p. 100) was not used in the economic evaluation. Given that a disease specific instrument was used within this trial, the FACT-B QoL instrument, the ERG considers the use of a mapping algorithm an option in estimating EQ-5D scores. This would allow estimating utility at baseline and each follow-up in each treatment group. The manufacturer was asked if they considered using such methodology and if they searched the literature for such an algorithm.

Whilst no search has apparently been conducted, the manufacturer states that they are not aware of the existence of a mapping algorithm (see clarification response B16).

The approach undertaken by the manufacturer was to derive utility estimates from the literature. However, the search that was conducted was not systematic and no attempt was made to collate or synthesise the alternative estimates found, to create a comprehensive evidence base. Instead, the manufacturer used estimates from only one of the studies found in their base case, although as part of the sensitivity analysis the manufacturer explored the use of an alternative study. In Table 18 , the variation in the utility values reported across the separate studies identified by the manufacturer (MS, Table 37, p.181) is shown for selected health states. Notably, the utility value for progressive disease for individuals not suffering toxicity is especially variable, ranging from 0.33 to 0.65. The ERG feels that the choice of utility estimates employed in the manufacturer's base case is somewhat arbitrary, and that the impact on results of using alternative utility values is not sufficiently well explored.

Table 18: Utility scores assumed in the base case analysis and variation in utility scores reported in the studies found through a non-systematic literature review

Health state	***Base case utility score***	***Between study differences,*** *mean (min to max), number of studies*
Response	0.81	0.82 (0.79 to 0.84), N=8
Stable disease	0.65	0.67 (0.62 to 0.75), N=8
Progression-free survival	0.73	-
Progressive disease	0.45	0.44 (0.33 to 0.65), N=8
Disutility from febrile neutropenia	-0.21	-0.21, N=1 - stable disease -0.47 (-0.51 to -0.42), N=2 - response
Disutility from peripheral sensory neuropathy	-0.21	-0.25, N=1 - stable disease -0.25 (-0.28 to -0.22), N=2 - response

Additionally, as mentioned above, only the more frequent AEs in E2100 were selected for inclusion in the model. However, rarer adverse events may impact on HRQoL and thus could be relevant for explicit inclusion in the model. As a minor point, the disutility value attributed to neuropathy, -0.21, is not coherent with the data presented for the reference which suggests a disutility of -0.25 (as seen in Table 18).

5.2.6 Resource utilisation and costs

In general, the ERG considers that the manufacturer has identified all the relevant cost categories. However, several potential shortcomings were identified regarding methods of costing and sources of unit. These will be described next, for each of the cost categories used.

The ERG would like to note that, in the manufacturer model, drug costs are applied to duration of treatment. This quantity is, in the MS, independent of the effectiveness parameters, although the length of treatment used, and thus costs, would be expected to relate to the health outcomes attained. This is not expected to impact on deterministic estimates; however, it may impact on the characterisation of decision uncertainty when undertaking PSA (which propagates explicit correlation between the inputs through to cost effectiveness).

5.2.6.1 Drug acquisition costs

The manufacturer submits two alternative scenarios as part of their base case analysis. In the first, list prices (from the BNF) were used to cost all treatments. However, in the second, paclitaxel (non-proprietary) was costed based on the average NHS Purchasing and Supply Agency (PASA) price, whilst the unit costs for all other (proprietary) treatments were based on list prices (BNF). Moreover, in this second base case, the manufacturer also implements a capping scheme for BEV.

NICE's guidance [55] states that the public list price should be used in the reference case analysis, but that the implications of variations from this price can be assessed using sensitivity analysis. Consequently, rather than presenting two alternative base-case analyses, it would have been more appropriate to present the results using the PASA price for paclitaxel as part of a separate sensitivity analysis in order to comply with existing NICE guidance.

Moreover, the PASA price is significantly different from the listed price (approximately a tenth) and, as the manufacturer suggests (MS p.196), a nationally agreed discounted price may not be uniformly available throughout the country. NICE is clear, however, that analyses based on price reductions for the NHS will only be considered when the reduced prices are transparent and can be consistently available across the NHS, and if the period for which the specified price is available is guaranteed. The variation of average prices across the NHS may be explored by using information supplied by the eMIT tool on the PASA prices – this interface reports not only the average price but also the standard deviation associated with this average price. According to this source, consulted by the ERG on 27/04/2010, the average price of a 150mg/25ml vial of paclitaxel is £26.10 (excluding VAT) and the standard deviation over this price is £9.36, suggesting a relatively wide variation in the average price throughout the trusts. The manufacturer did not explore the impact of the existence of price variations for paclitaxel across the NHS.

Furthermore, the ERG considers that it is likely that the other chemotherapy treatments are acquired locally with a discount, and this was not considered in the analysis provided by the

manufacturer. By doing so, the estimates incremental analysis of the regimen BEV+PAC may not be conservative.

The capping scheme considered in costing bevacizumab is not approved by the Department of Health, and consequently is not implemented in the NHS. The ERG thus considers analyses based on this scheme should not have been presented as part of a base case analysis.

Still related to the unit costs of chemotherapy regimens, another shortcoming was identified regarding the acquisition costs of docetaxel. The ERG identified that the patent over this treatment is soon to expire (November 2010).[f] It is thus likely that non-proprietary products will be available soon, and that their list price will be lower than the price used in the current evaluation [56]. The manufacturer failed to identify or explore a possible reduction in acquisition costs for docetaxel, which may impact significantly the cost effectiveness of this treatment and consequently the overall conclusions of this study.

The current analysis also ignores the possibility of dose reductions. The extent to which dose reductions occur may differ between alternative treatments, and the ERG expects this to impact on the results. As an example, more toxic drugs may lead to higher rates of dose reduction, and consequently the acquisition costs would be lower. In the current analysis, the ERG expects costs to be overestimated in more toxic drugs, such as PAC and DOC. It is not clear how this will affect incremental cost-effectiveness of BEV combination therapy, but is unlikely have a large effect on the ICER.

5.2.6.2 Drug administration costs

The ERG has not identified any issue with the manufacturer's approach to estimating drug administration costs.

5.2.6.3 Duration of treatment

Duration of treatment was mainly estimated from the E2100 trial (MS p.156-157), and was used to calculate the average cost of treatment. Because the estimation was conducted independently of the rates of recurrence with treatments this has no bearing on effectiveness/natural history – if lower or higher time on treatment is assumed in the current model, effectiveness estimates (e.g. PFS) will not be affected. The effectiveness of the interventions and comparators described in Section 5.1.1 and critiqued in 5.2.1 is, in the current submission, drawn directly from data of the E2100 trial. This means that distinct

[f] Amended from May 2010 by ERG on advice of Roche

treatments characterised, for example, by a distinct frequency of administration or distinct stopping rules cannot be evaluated within the submitted model. Thus, altering the way in which the drugs are administered or altering time off treatment does not impact on the model outcomes on effectiveness, but only on costs.

Since almost all patients had stopped treatment by the end of the E2100 trial, the ERG does not expect the application of alternative extrapolation methods to duration of treatment to impact significantly on cost effectiveness results. The manufacturer has provided model based estimates (derived from probabilistic sensitivity analysis) and trial based estimates [e.g. by using the area under the curve (AUC) method applied to the Kaplan Meier curves] of mean duration of treatment. These are shown in Table 19.

Table 19: **Trial based and model based estimates of mean duration of treatment (from clarification B4 p.45 Table 12)**

Duration of treatment (months)	**Trial based estimate,** mean (SE)*	**Model based estimate,** mean (95%CI)
Bevacizumab	7.38 (SE 0.29)	7.83 (95% CI: 7.20 to 8.42)
Paclitaxel (Bev/Pac arm)	6.72 (SE 0.26)	7.16 (95% CI: 6.57 to 7.71)
Paclitaxel (Pac arm)	4.84 (SE 0.22)	5.35 (95% CI: 4.98 to 5.77)

* derived from Kaplan Meier based on last observed time
** from PSA with 10,000 runs using Weibull best fit

The expected durations of treatment estimated by the model are higher than observed within the E2100 trial, although this bias is fairly constant across the treatments. The manufacturer has not commented on whether the time on treatment in the E2100 trial is representative of clinical practice in the UK.

With regards to the regimens GEM+PAC and DOC, assumptions were used to define duration of treatment. The ERG considers that there are few data to support these assumptions, and an assessment of their impact should have been conducted by the manufacturer.

5.2.6.4 Supportive care costs

The manufacturer did not include the costs of second line treatment for patients in the model with progressed disease. This is because, firstly, the same second line treatment is expected to be used in both arms. Secondly, model assumptions establish that patients spend the same time in the progressed disease state for all treatments, and thus the use of distinct first line treatments is assumed not to impact on this. The ERG considers the latter assumption, critiqued in Section 5.2.2, to be a shortcoming of the analysis. As a consequence, the manufacturer was asked to provide further justification on the costs of second line therapies

following progression and also to detail relevant protocols followed in UK clinical practice. The manufacturer did not provide additional clarification on this (see clarification response B15).

5.2.6.5 Adverse events costs

The unit costs assumed for AEs were derived from national tables and the ERG deemed these to be reasonable.

5.2.6.6 End of life costs

The ERG considers the evidence on end of life costs not to be robust, and that the impact on cost effectiveness of varying these costs should have been explored within the MS.

5.2.7 Discounting

Discounting was appropriately conducted.

5.2.8 Subgroup analysis

Subgroup analyses of cost effectiveness estimates are relevant when the relative effect of a treatment is expected to differ (and/or when the uncertainty over this differs) between subgroups of individuals. The evidence produced within the E2100 trial for there being an important subgroup effect in terms of the relative effect of bevacizumab is weak (MS p.92). The ERG thus accepts the assumption that the relative capacity to benefit clinically from this treatment may not differ for patients with different characteristics. However, the manufacturer has not shown this assumption necessarily to hold for the comparator regimens.

Even if the relative effectiveness of treatments is equivalent between subgroups, it may still be relevant to explore variation in cost effectiveness within these same subgroups. This happens when the subgroups differ in prognosis or in the overall costs incurred, which affects the incremental costs or the incremental effects. In the current assessment, the ERG considers there may be important variation across subgroups in the baseline PFS and OS estimates, and hence in absolute gains in mean PFS and OS. In the submission, the manufacturer did not explore these subgroup effects. As a consequence, the ERG requested time to progression data for the alternative subgroups explored in the E2100 trial (see clarification response A11). These data show that different patient characteristics (e.g. severity of illness or frailty of patients) affect median PFS. However, the impact of these differences on cost effectiveness is unexplored.

Additionally, the ERG considers that particular treatment regimens may not be considered to be relevant comparators for individuals with certain characteristics (e.g. intolerant, frailer patients). After a request for clarification, the manufacturer confirmed that this may be the case (clarification response A1), by suggesting that taxanes are more toxic than, for example, capecitabine and thus frailer patients may receive the later as first line treatment. The manufacturer further explored the potential impact for the current evaluation. The ERG feels that, for this subgroup of patients, a comparison with capecitabine would have been more appropriate, although this was not specified in the scope for this STA.

5.2.9 Sensitivity analysis

The manufacturer undertook a detailed set of scenario analyses and PSA. However, the ERG considers that parameter uncertainty was not fully explored. This is because not all relevant parameters seem to have been considered uncertain in PSA (see Table 14). The ERG considers this to preclude a correct characterisation of uncertainty.

5.3 Results included in manufacturer's submission

The results of the model are presented in the manufacturer's submission from p.207 to 219.

Base case

In analysing the cost-effectiveness of bevacizumab in combination with paclitaxel two base case analyses are provided, the first evaluating paclitaxel at NHS list prices (case 1) and the second at PASA prices (case 2). Case 2 also includes a 10g cap for bevacizumab (as described in MS p.196).

A breakdown of the costs incurred is shown in the MS (p. 207 and 208). Time spent in PFS and progressive states is also detailed in the MS (p.209). For case 1, based on NHS list prices, the manufacturer's model estimated ICERs for BEV+PAC of £117,803, £115,059 and £105,777 per QALY gained, respectively, relative to PAC, DOC and GEM+PAC regimens (Table 20). If PASA prices are used instead, the ICERs for BEV+PAC is estimated at £77,314, £57,753 and £60,101 per QALY, respectively.

Table 20: Deterministic cost effectiveness results for BEV+PAC over a time horizon of 10 years – case 1 (NHS list price for paclitaxel) and case 2 (PASA price for paclitaxel and 10 grams capping scheme for bevacizumab). Adapted from Tables 56 and 57 in the MS.

	Results	***BEV+PAC***	***PAC***	***DOC***	***GEM+PAC***

Case 1.— NHS list prices	Mean Life Years (yrs)	2.682	2.330	2.330	2.330
	Mean QALYs	1.498	1.239	1.225	1.239
	Mean Total Cost	£56,473	£26,004	£25,057	£29,115
	Incremental Life Years		*0.352*	*0.352*	*0.352*
	Incremental QALYs		*0.259*	*0.273*	*0.259*
	Incremental Cost		*£30,469*	*£31,416*	*£27,358*
	Cost per Life Year Gained		**£86,572**	**£89,263**	**£77,734**
	Cost per QALY Gained		**£117,803**	**£115,059**	**£105,777**
Case 2 – PASA prices and BEV cap	Mean Life Years (yrs)	2.682	2.330	2.330	2.330
	Mean QALYs	1.498	1.239	1.225	1.239
	Mean Total Cost	£40,826	£20,829	£25,057	£25,281
	Incremental Life Years		*0.352*	*0.352*	*0.352*
	Incremental QALYs		*0.259*	*0.273*	*0.259*
	Incremental Cost		*£19,997*	*£15,769*	*£15,545*
	Cost per Life Year Gained		**£56,818**	**£44,805**	**£44,168**
	Cost per QALY Gained		**£77,314**	**£57,753**	**£60,101**

The manufacturer qualitatively concluded that, based on the above results, the regimen BEV+DOC is not cost effective. This is because the manufacturer assumed that, according to the AVADO trial,[27-36] BEV+DOC was no more effective than BEV+PAC. Also, higher costs are expected to be incurred, since DOC is more costly than PAC. The MS concluded that BEV+DOC would be dominated by BEV+PAC, but a formal analysis was not carried out to demonstrate this.

Sensitivity analyses

Scenario analyses

In Table 21 the results of the alternative scenarios examining the impact of alternative assumptions and parameter estimates are presented. Importantly, these analyses all use the PASA prices for PAC and the capping scheme for BEV cap. The ERG notes that, because costing of the treatments is independent from the effectiveness, careful interpretation of the following scenarios is needed. This is especially relevant for scenario B.

Table 21: Scenario analyses results for pairwise comparison of the regimen BEV+PAC with identified regimens (adapted from Table 58 in the MS). See Table 10 for further details on specification of these analyses.

Scenario		***ICER** (costs per QALY gained) **of BEV+PAC compared to***		
ID	**Description**	**PAC qw**	**DOC**	**GEM+PAC**
	Base case	£77,314	£57,753	£60,101
A	Weibull function	£70,662	£52,128	£54,951
A	Exponential function	£57,838	£44,766	£45,055
A	Log logistic function	£53,492	£40,448	£41,660
A	Log normal function	£58,969	£44,363	£45,919
A	Generalized Gamma function	£62,591	£46,743	£48,716

B	First line treatment administered until progression	£97,308	£60,832	£67,833
??	Utilities: Weighted by response rates (Cooper 2003)	£68,343	£50,655	£53,746
E	Utilities: Weighted by response rates (Lloyd 2006)	£65,977	£50,066	£51,500
G	Monthly supportive care cost: alternative values (Remak 2004)	£74,728	£55,376	£57,515
F	Monthly supportive care cost decrease by 50%	£75,844	£56,397	£58,631
F	Monthly supportive care cost increase by 50%	£78,784	£59,109	£61,571
C	Patient weight = 60kg; BSA = 1.6 m2	£67,350	£48,023	£49,921
C	Patient weight = 80kg; BSA = 1.8 m2	£85,289	£65,307	£66,233

The scenario analyses performed by the manufacturer resulted in ICERs ranging from (i) £53,492 and £97,308 per QALY for BEV+PAC when compared to PAC qw; (ii) £40,448 and £65,307 per QALY for BEV+PAC when compared to DOC; and (iii) £41,660 and £67,833 per QALY for BEV+PAC when compared to GEM+PAC. The equivalent range of ICERs based on the list prices and removing the BEV capping scheme was not presented by the manufacturer.

Probabilistic sensitivity analysis

The model produced probabilistic cost effectiveness results comparable to those obtained from the deterministic analysis (Table 22).

Table 22: Cost effectiveness results obtained from deterministic and probabilistic analysis

Description	***ICER of BEV+PAC compared to***		
	PAC	***DOC***	***GEM+PAC***
Deterministic Results (Case 2, base case)			
Cost per life year gained (£)	£56,818	£44,805	£44,168
Cost per QALY gained (£)	£77,314	£57,753	£60,101
PSA results			
Cost per life year gained (£)	£56,248	£45,323	£38,628
Cost per QALY gained (£)	£76,571	£58,645	£51,450
	Probability of BEV+PAC being cost-effective compared to		
	PAC	***DOC***	***GEM+PAC***
PSA results			
at a threshold of £20k	0	0	0
at a threshold of £30k	0	0	0

Incremental cost effectiveness planes and cost effectiveness acceptability curves are shown in the MS (p.216 to 219), for pairwise comparisons of BEV+PAC in relation to each of the comparator treatments considered. An example is presented in Figure 3.

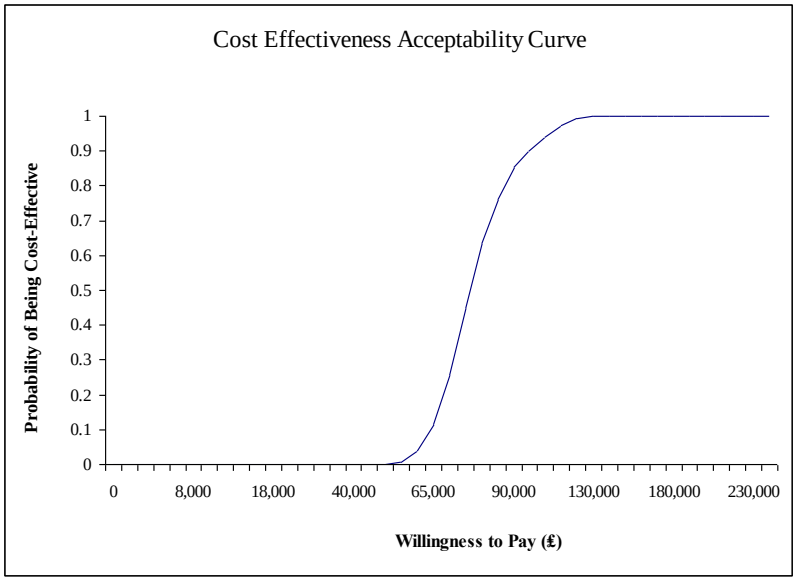


Figure 3: **Cost-effectiveness acceptability curve for the adoption decision of BEV+PAC compared to PAC qw**

Table 22 details the probability of BEV+PAC being cost-effective when compared to the remaining regimens (as pairwise comparisons). The cost-effectiveness acceptability curves shown in the MS (Figures 29-31 p 218-219, MS) demonstrate that there is no uncertainty that BEV+PAC is not cost-effective in relation to the comparator regimens based on a cost-effectiveness threshold of £20,000-£30,000 per QALY. Figure 3 shows the cost-effectiveness acceptability curve reported by the manufacturer for the comparison of BEV+PAC vs PAC qw.

5.4 Comment on validity of results presented with reference to methodology used

The manufacturer's results indicate that bevacizumab used in combination with paclitaxel is unlikely to be cost effective at a conventional threshold of £20 000 to £30 000 per QALY. The ICERs for BEV+PAC, based on NHS list prices, in the manufacturer's model were estimated at £117,803, £115,059, and £105,777 relative to PAC, DOC, and GEM+PAC regimens, respectively (Table 17). If PASA prices with a 10g per patient cap for the cost of BEV are used instead, the ICERs for BEV+PAC are estimated at £77,314, £57,753 and £60,101 per QALY compared to PAC, DOC and GEM+PAC, respectively.

It should be noted that the ICER estimates presented by the manufacturer are based on a series of pairwise comparisons for BEV+PAC relative to each separate comparator regimen. However, calculating a series of pairwise ratios is not appropriate when considering more

than two regimens and, in particular circumstances, can be misleading. To establish the correct estimate of the ICER for BEV+PAC this should be undertaken using a fully incremental analysis comparing all the regimens simultaneously. This is a central tenet of cost-effectiveness analysis and involves assessing the incremental cost of generating additional health effects when moving from one option to a more effective one, and assessing this against a relevant measure of opportunity cost (e.g. the NICE threshold). Regimens that are dominated (i.e. those which are more expensive and less effective than one or more alternatives) are removed from further consideration. So too are options which are extendedly dominated – that is, more costly and less effective than a combination of two alternatives. The ICERs of each of the remaining regimens are then calculated as the additional costs divided by the additional effects by comparing on option with the next least costly/effective.

The validity of the estimated ICERs are subject to a number of remaining uncertainties and issues in relation to the modelling undertaken by the manufacturer. These are outlined by the ERG in the section below and are summarised in Table 16. The ERG considers that the assumptions in the base case are not necessarily inappropriate individually, but that alternative assumptions have not been adequately explored and that, taken together, the assumptions in the base-case tend to be optimistic towards the estimated cost-effectiveness of bevacizumab. The main assumptions are:

- Mortality after disease progression is independent of initial treatment
- All comparators are equally effective, measured by both PFS and OS
- The cost of BEV to the NHS is limited to 10g per patient
- Paclitaxel is available to the NHS at discounted PASA prices
- Docetaxel is available to the NHS at its proprietary price

Detailed inspection of assumptions made throughout, however, revealed a number of additional issues that may impact on the validity of the cost effectiveness results.

5.5 Summary of uncertainties and issues

A number of potential uncertainties are identified and described in Section 5.2, and summarised in Table 23. Several of these issues were subject to additional analyses by both the manufacturer, as part of their response to the ERG's points for clarification, and the ERG. The results of these additional analyses are presented in Section 6.

Table 23: **Summary of uncertainties and issues identified in Section 5.2**

Topic, uncertainty or issue	*Likely consequences for the results and conclusions*	*Additional analysis by manufacturer*	*Additional analysis by ERG*
5.2.1 Interventions and comparators			
Not all relevant interventions (as defined by the scope and as used in clinical practice) have been compared in the model.	Major, impact unknown	Yes (Section 6.1.1)	Yes (Section 6.2.2)
5.2.2 Natural history			
-			
5.2.3 Treatment effectiveness within the submission			
The relative treatment effects of progression free survival are proportional hazards (the hazard ratios are constant over time)	Minor, impact unknown	No	No
The rate of death following progression is constant over time	Minor, impact unknown	No	No
The rate of death following progression is the same for all initial treatments	Major, expected to benefit BEV+PAC	No	Yes (Section 6.2.3)
The effectiveness of the comparators (DOC and GEM+PAC) is the same as PAC qw	Minor, expected to benefit BEV+PAC	Yes (Section 6.1.2)	Yes (Section 6.2.4)
5.2.4 Adverse events			
Not all available evidence was used to derive rates of adverse events	Minor, impact unknown	No	No
Adverse events occurring in less than 3% of the E2100 trial were excluded	Minor, impact unknown	No	No
5.2.5 Health related quality of life			
The literature search conducted to identify utility values was not systematic and no attempt was made to collate the alternative estimates found	Unknown	No	Yes (Section 6.2.1.2)
5.2.6 Resource utilisation and costs			
Drug acquisition costs: In a second base case scenario, non listed prices were used to cost paclitaxel.	Major	No	This case should be disregarded as a base case
10 g cap on costs of BEV. The base case comprises two alternative analysis one assuming list prices and another assuming PASA prices and a 10g capon the expenditure for BEV. The cap has not been approved by the Department of Health and no analysis was presented by the manufacturer that considered the effect of PASA prices separately from the effect of the 10g cap.	Major	No	Yes (Sections 6.1 and 6.2)
Drug acquisition costs: Patent over docetaxel expires in November 2010.[g] Generic products with lower list price are likely to be available soon.	Major, impact unknown	No	Yes (Section 6.2.1.1)
Drug acquisition costs: Assessment ignores the possibility of dose reduction	Minor, expected to benefit BEV+PAC	No	No
Duration of treatment: The impact of considering alternative distributions to describe this quantity was not evaluated	Minor, expected to benefit BEV+PAC	No	No
Duration of treatment: The way in which treatments are used in E2100 differ from UK's clinical practice which may impact on costs	Expected to benefit comparators	No	No
Duration of treatment: Impact of assumption on duration of treatment of GEM+PAC and DOC	Unknown	No	No
5.2.7 Discounting			
No issue was identified			

[g] Amended from May 2010 by ERG on advice of Roche

5.2.8 Subgroup analysis			
There may be important variation across subgroups in the baseline PFS and OS estimates, and hence in absolute gains in mean PFS and OS.	Minor, impact unknown	No	No
Particular treatments (or regimens) may not be considered to be relevant comparators for individuals with certain characteristics (e.g. intolerant, frailer patients).	Minor, impact unknown	No	No
5.2.9 Sensitivity analysis			
Univariate sensitivity analysis was not undertaken	Minor	No	No
Not all relevant parameters seem to have been considered uncertain in PSA	Minor, impact unknown	No	No

6 Additional 'exploratory' or other work undertaken by the manufacturer and ERG

6.1 Additional work undertaken by the manufacturer

Following a number of points of clarification raised by the ERG, the results from two additional analyses were presented by the manufacturer:

1. Incorporating additional comparators into the existing economic analysis to reflect the NICE scope and current licensing for paclitaxel (clarification response B1).
2. Using the results of the evidence synthesis to evaluate cost effectiveness as opposed to assuming that all comparators were equally effective (clarification response B2).

The first of these analyses reflects the concerns raised by the ERG that not all regimens identified in the NICE scope were subsequently included in the manufacturer's analysis. Of particular relevance is the exclusion of the regimen BEV+DOC, as this excludes potentially relevant evidence from the AVADO trial [27-36] (see Section 4.2.2.2.1). Furthermore, the paclitaxel regimen included in the MS does not conform to the SPC and current licensing for paclitaxel, which is based on a q3w regimen. While the inclusion of a non-licensed comparator is permitted within current NICE guidance when it is used in current clinical practice (a point which was confirmed by our clinical advisor), the ERG considers that both paclitaxel regimens should have been explicitly included as comparators in this evaluation, particularly since the cost-effectiveness of a qw paclitaxel regimen has not been previously demonstrated compared to the licensed regimen.

The second analysis uses the available evidence on relative treatment effects (derived from the indirect comparison) instead of using the assumption of equality of effects made in the MS.

The results of the analyses presented by the manufacturer are discussed by the ERG below. However, it should be noted that all the revised analyses submitted by the manufacturer were based on the PASA prices for paclitaxel and the capping scheme for bevacizumab (i.e. Case 2). The corresponding results based on NHS list prices were not reported by the manufacturer. To assist in the interpretation of the revised results presented by the manufacturer, the ERG has also re-run the same analyses using NHS list prices (i.e. Case 1). In addition, the ERG has undertaken a third approach (Case 3) which includes the PASA prices for paclitaxel and

excludes the capping scheme for bevacizumab. In summary, the 3 cases referred to in the subsequent tables are:

- Case 1 (ERG re-analysis) – NHS list prices *excluding* capping scheme for bevacizumab
- Case 2 (manufacturer re-analysis) – PASA prices for paclitaxel *including* capping scheme for bevacizumab
- Case 3 (ERG re-analysis) – PASA prices for paclitaxel *excluding* capping scheme for bevacizumab

6.1.1 Inclusion of relevant comparators (PAC q3w and BEV+DOC)

The manufacturer did not attempt to incorporate BEV+DOC as an additional comparator as part of any re-analysis presented as part of their response to the ERG points of clarification (see clarification response B1). Instead the manufacturer re-iterated that they did not consider BEV+DOC to be a relevant comparator for the following reasons: (i) it was unlikely to be cost-effective; and (ii) it is not recommended by NICE and (iii) it is not used in standard UK practice. Nevertheless, since NICE included this regimen as part of the scope and the use of BEV+DOC forms part of the current license for bevacizumab, the ERG considers that this regimen should have been formally incorporated to provide the Appraisal Committee with an explicit and quantitative basis to support their decision and to appropriately characterise decision uncertainty in relation to the full range of relevant alternatives.

Although the manufacturer did not incorporate a BEV+DOC regimen in their revised analyses, they did incorporate the q3w paclitaxel monotherapy strategy (PAC q3w). The revised economic model incorporating this regimen was based on the original model using the gemcitabine + q3w paclitaxel regimen. The manufacturer simply changed the existing parameterisation of the gemicitabine + q3w paclitaxel regimen to model the PAC q3w regimen. The treatment benefit (in terms of PFS) of BEV+PAC relative to PAC q3w was derived from the manufacturer's indirect treatment comparison (ITC). The calculated PFS hazard ratio for BEV+PAC compared to PAC q3w was estimated to be 0.338 (95% CI 0.26 to 0.44) by the manufacturer. To model the acquisition and monitoring costs of the PAC qw3 regimen, the manufacturer simply removed the drug and administration cost associated with gemcitabine from the model.

The revised cost effectiveness results for the comparison between BEV+PAC versus PAC q3w are presented in Table 24 (column identified as PAC q3w ITC) alongside the results of the original base case analysis. The ICER of BEV+PAC compared to PAC 3qw was £90,761 per QALY based on NHS list prices (Case 1) and £59,339 per QALY using PASA prices and incorporating the capping scheme for bevacizumab. The equivalent ICERs of BEV+PAC versus PAC 3qw for Cases 2 and 3 were £59,339 and £82,151 per QALY, respectively.

6.1.2 Incorporation of results from the evidence synthesis

The manufacturer also presented the results of an additional analysis incorporating t the indirect comparisons for the effectiveness data (i.e. not assuming that the comparator regimens were equally effective). The manufacturer replaced the assumption of equality in the effectiveness of the regimens DOC and GEM+PAC in relation to PAC qw with the results of the ITC. Further details on the analysis of relative effectiveness are reported in the clarification response, B2. The cost effectiveness results based on the indirect comparisons are also summarised in Table 24 for each of the 3 separate cases considered. The use of the ITC did not appear to alter significantly the ICER estimates, although the ICER estimates for BEV+PAC were marginally higher (i.e. less favourable) versus DOC than when assuming equal effects with PAC qw and marginally lower (i.e. more favourable) versus GEM+PAC.

6.1.3 ERG's commentary on the additional analyses

In addition to presenting the equivalent ICER results for the separate cases not considered by the manufacturer, the ERG has also undertaken a fully incremental analysis based on the manufacturer's revised analyses for each of the 3 cases. The results from the fully incremental analysis are reported separately in Table 25. These results show that GEM+PAC is dominated throughout the 3 sets of analysis by PAC qw. The DOC regimen is also dominated by the PAC qw regimen in Cases 2 and 3. In these two cases PAC q3w is the cheapest and least effective regimen. The ICER of PAC qw in both these cases is below conventional threshold of cost-effectiveness (£19,769 per QALY). The ICER of BEV+PAC presented in these two cases is versus PAC qw (£77,314 per QALY including the capping scheme and £110,475 per QALY excluding the capping scheme). Hence, in cases 2 and 3 the pairwise ICER presented by the manufacturer between BEV+PAC versus PAC qw is the relevant ICER comparison for BEV+PAC derived from a fully incremental analysis. However, when NHS list prices are used throughout (Case 1), PAC qw appears to be extendedly dominated by DOC. Accordingly, the relevant ICER comparison for BEV+PAC is now versus DOC (£118,362 per QALY).

For the sake of completeness, the ERG has also conducted a second fully incremental analysis based on the original approach employed by the manufacturer to estimate the effectiveness inputs of the comparator regimens i.e. where the effects on PFS of DOC and GEM+PAC were assumed equal to those of PAC. These results are presented in Table 23.The overall results are similar to the revised analyses, except for the case where list prices are used (Case 1). In this analysis, DOC is no longer dominated. The ICER for PAC qw is now estimated versus DOC (£67,643 per QALY) and the ICER for BEV+PAC (£117,641 per QALY) is estimated versus PAC qw. Hence, in each of the 3 cases, the pairwise ICER comparison presented by the manufacturer of BEV+PAC versus PAC qw is the relevant ICER comparison derived from a fully incremental analysis.

Table 24: Deterministic cost effectiveness results for pairwise comparisons of BEV+PAC over a time horizon of 10 years – case 1 (NHS list price for paclitaxel), case 2 (PASA price for paclitaxel and capping scheme for bevacizumab) and case 3 (PASA price for paclitaxel). Adapted from Tables 56 and 57 in the MS, Tables 7 and 8 of clarifications document, and accompanying Excel files.

	Results	***BEV+PAC***	***PAC qw***	***PAC q3w*** *(ITC)*	***DOC*** *(effects equal to PAC qw)*	***DOC*** *(ITC)*	***GEM+PAC*** *(effects equal to PAC qw)*	***GEM+PAC*** *(ITC)*
		Base case	*Base case*	*Revised*	*Base case*	*Revised*	*Base case*	*Revised*
Case 1— NHS list prices	Mean Life Years (yrs)	2.682	2.330	2.195	2.330	2.340	2.330	2.282
	Mean QALYs	1.498	1.239	1.122	1.225	1.233	1.239	1.197
	Mean Total Cost	£56,473	£26,004	£22,350	£25,057	£25,111	£29,115	£29,104
	Incremental Life Years		*0.352*	*0.487*	*0.352*	*0.343*	*0.352*	*0.400*
	Incremental QALYs		*0.259*	*0.376*	*0.273*	*0.265*	*0.259*	*0.300*
	Incremental Cost		*£30,469*	*£34,124*	*£31,416*	*£31,403*	*£27,358*	*£27,611*
	Cost per Life Year Gained *		**£86,572**	**£70,071**	**£89,263**	**£91,530**	**£77,734**	**£68,442**
	Cost per QALY Gained *		**£117,803**	**£90,761**	**£115,059**	**£118,362**	**£105,777**	**£91,133**
Case 2 — PASA price and BEV cap	Mean Life Years (yrs)	2.682	2.330	2.195	2.330	2.340	2.330	2.282
	Mean QALYs	1.498	1.239	1.122	1.225	1.233	1.239	1.197
	Mean Total Cost	£40,826	£20,829	£18,516	£25,057	£25,111	£25,281	£25,271
	Incremental Life Years		*0.352*	*0.487*	*0.352*	*0.343*	*0.352*	*0.400*
	Incremental QALYs		*0.259*	*0.376*	*0.273*	*0.265*	*0.259*	*0.300*
	Incremental Cost		*£19,997*	*£22,310*	*£15,769*	*£15,715*	*£15,545*	*£15,555*
	Cost per Life Year Gained *		**£56,818**	**£45,812**	**£44,805**	**£45,865**	**£44,168**	**£38,899**
	Cost per QALY Gained *		**£77,314**	**£59,339**	**£57,753**	**£59,310**	**£60,101**	**£51,795**
Case 3 — PASA prices	Mean Life Years (yrs)	2.682	2.330	2.195	2.330	2.340	2.330	2.282
	Mean QALYs	1.498	1.239	1.122	1.225	1.233	1.239	1.197
	Mean Total Cost	£49,403	£20,829	£18,516	£25,057	£25,111	£25,281	£25,271
	Incremental Life Years		*0.352*	*0.487*	*0.352*	*0.343*	*0.352*	*0.400*
	Incremental QALYs		*0.259*	*0.376*	*0.273*	*0.265*	*0.259*	*0.300*
	Incremental Cost		*£28,573*	*£30,886*	*£24,346*	*£24,292*	*£24,121*	*£24,132*
	Cost per Life Year Gained *		**£81,187**	**£63,424**	**£69,174**	**£70,679**	**£68,537**	**£60,347**
	Cost per QALY Gained *		**£110,475**	**£82,151**	**£89,164**	**£91,679**	**£93,262**	**£80,345**

ITC – Indirect treatment comparison

* ICER estimates regard the pairwise comparisons between BEV+PAC and each comparator regimen

Table 25: **Full incremental analysis of the revised results regarding the cost effectiveness of alternative chemotherapy regimens for mBC – case 1 (NHS list price for all treatments), case 2 (PASA price for paclitaxel and capping scheme for bevacizumab) and case 3 (PASA price for paclitaxel).**

			Mean costs	*Mean QALYs*	*Incremental cost,* next best	*Incremental QALYs,* next best	*ICER (£/QALY)* next best
Case 1— NHS list prices	PAC q3w (ITC)	Revised	£22,350	1.122	-	-	-
	DOC (ITC)	Revised	£25,111	1.233	£2,761	0.111	£24,874
	PAC qw (E2100)	Base case	£26,004	1.239			Extendedly dominated
	GEM+PAC (ITC)	Revised	£29,104	1.197			Dominated
	BEV+PAC (E2100)	Base case	£56,473	1.498	£31,362	0.265	£118,362
Case 2 — PASA price and BEV cap							
	PAC q3w (ITC)	Revised	£18,516	1.122	-	-	-
	PAC qw (E2100)	Base case	£20,829	1.239	£2,313	0.117	£19,769
	DOC (ITC)	Revised	£25,111	1.233			Dominated
	GEM+PAC (ITC)	Revised	£25,271	1.197			Dominated
	BEV+PAC (E2100)	Base case	£40,826	1.498	£19,997	0.259	£77,314
Case 3 — PASA price	PAC q3w (ITC)	Revised	£18,516	1.122	-	-	-
	PAC qw (E2100)	Base case	£20,829	1.239	£2,313	0.117	£19,769
	DOC (ITC)	Revised	£25,111	1.233			Dominated
	GEM+PAC (ITC)	Revised	£25,271	1.197			Dominated
	BEV+PAC (E2100)	Base case	£49,403	1.498	£28,574	0.259	£110,475

ITC – Indirect treatment comparison

Table 26: **Full incremental analysis of the non-revised (original MS) results regarding the cost effectiveness of alternative chemotherapy regimens for mBC – case 1 (NHS list price for all treatments), case 2 (PASA price for paclitaxel and capping scheme for bevacizumab) and case 3 (PASA price for paclitaxel).**

		Mean costs	*Mean QALYs*	*Incremental cost,* next best	*Incremental QALYs,* next best	*ICER (£/QALY)* next best
Case 1— NHS list prices	DOC (PFS equal to PAC qw)	£25,057	1.225	-	-	-
	PAC qw (E2100)	£26,004	1.239	£947	0.014	£67,643
	GEM+PAC (PFS equal to PAC qw)	£29,115	1.239			Dominated
	BEV+PAC (E2100)	£56,473	1.498	£30,469	0.259	£117,641
Case 2 — PASA price and BEV cap						
	PAC qw (E2100)	£20,829	1.239	-	-	-
	DOC (PFS equal to PAC qw)	£25,057	1.225			Dominated
	GEM+PAC (PFS equal to PAC qw)	£25,281	1.239			Dominated
	BEV+PAC (E2100)	£40,826	1.498	£19,997	0.259	£77,314
Case 3 — PASA price	PAC qw (E2100)	£20,829	1.239	-	-	-
	DOC (PFS equal to PAC qw)	£25,057	1.225			Dominated
	GEM+PAC (PFS equal to PAC qw)	£25,281	1.239			Dominated
	BEV+PAC (E2100)	£49,403	1.498	£28,574	0.259	£110,475

ITC – Indirect treatment comparison; PFS – Progression free survival

6.2 Additional work undertaken by the ERG

In addition to presenting the results using a separate costing approach (Case 3) and undertaking a fully incremental analysis, the ERG has undertaken several analyses to address some of the other limitations and uncertainties identified in Section 5. These are:

1. Docetaxel is soon to be off-patent (from November 2010)[h], and it is likely that the price will fall in the near future. Although the price reduction is likely to depend on market conditions and is difficult to predict, the ERG has undertaken sensitivity analyses around a range of potential prices for generic docetaxel.
2. The ERG considered that the selection of utility estimates employed in the manufacturer's base case appeared relatively arbitrary and that the impact on results of using alternative utility values had not been sufficiently well investigated. The ERG has undertaken additional analyses using alternative utility (HRQOL) values that could have been used.
3. The MS did not incorporate a BEV+DOC regimen in the model. The ERG has, therefore, explored alternative approaches to incorporating this regimen to consider its cost-effectiveness.
4. The ERG is aware of the limitations of the evidence base used to undertaken the indirect treatment comparison (Section 4.2.2.3), and that the MS evidence synthesis has made a reasonable attempt to compare the effectiveness of the different therapies using the available data. Nevertheless, the ERG has noted a number of potential weaknesses in the ITC which are also addressed.
5. The ERG was concerned about the internal validity of the results for OS from the model and the strong assumptions used by the manufacturer which effectively resulted in a comparable gain in OS to that for PFS and noted that other approaches could have been explored using alternative assumptions and model structures. The ERG has, therefore, undertaken an additional analysis using an area under the curve (AUC) model.

6.2.1 Further exploration of the manufacturer's revised model by the ERG

6.2.1.1 Impact of a reduction of the list price of docetaxel

The ERG has evaluated the impact of a reduction of the list price of docetaxel on cost effectiveness. For this, we evaluated the change in total costs incurred by docetaxel regimen

[h] Amended from May 2010 by ERG on advice of Roche

for discounts between 0 and 100 %.We used the revised model submitted by the manufacturer to produce the results shown in Table 27. The results show the impact on total costs of docetaxel and on the ICER of BEV+PAC in relation to DOC for case 1 (NHS list price for all treatments) and case 3 (PASA price for paclitaxel).

Table 27: Impact of potential reductions on the list price of docetaxel.

% price reduction for docetaxel	*Total costs of docetaxel regimen (revised model)*	*ICER of BEV+PAC vs. DOC* revised model Case 1 – NHS list prices	*ICER of BEV+PAC vs. DOC* (revised model, Case 3 – PASA prices)
0	£25,111	£118,362	£91,679
10%	£24,434	£120,915	£94,232
20%	£23,758	£123,468	£96,785
30%	£23,082	£126,021	£99,337
40%	£22,405	£128,574	£101,890
50%	£21,729	£131,126	£104,443
60%	£21,052	£133,679	£106,996
70%	£20,376	£136,232	£109,549
80%	£19,699	£138,785	£112,101
90%	£19,023	£141,338	£114,654
100%	£18,347	£143,890	£117,207

As expected the overall costs of the DOC regimen decrease as the reduction in price increases. Using list prices for costing the other drugs (Case 1) and considering a price reduction compared to current patent prices for a generic formulation of docetaxel of between 20 to 30%, the ICER of BEV+PAC versus DOC increases from £118,362 per QALY(i.e. assuming current prices) to £123,468 to £126,021 per QALY, respectively. A similar increase in the ICER is also evident for case 3 where (discounted) PASA prices are used in costing paclitaxel.

6.2.1.2 Impact of using alternative utility values

To evaluate the impact on results of using alternative utility values the ERG varied, in turn, each of the utility parameters included in the model. The alternative values considered were based on the minimum and maximum values reported for each health state derived from the full set of studies identified by the manufacturer. The range of values and the impact on the ICER estimates for the comparison between BEV+PAC versus PAQ qw are reported in Table 28. The results demonstrate that the ICER estimates only alter marginally across the range of utility values considered in the literature.

Table 28: Impact of alternative utility estimates

	Utility score	***BEV+PAC vs. PAC qw***	
Health state	***Between study differences,*** *mean (min to max)*	***Incremental QALY's***	***ICER***

Response	0.82 (0.79 to 0.84)	0.260 (0.255 to 0.264)	£117,000 (£119,443 to £115,425)
Stable disease	0.67 (0.62 to 0.75)	0.262 (0.253 to 0.276)	£116,207 (120,281 to 110,234)
Progressive disease	0.44 (0.33 to 0.65)	0.259 (0.259 to 0.258)	£117,788 (£117,625 to £118,100)
Disutility from peripheral sensory neuropathy	-0.25 (-0.28 to -0.22)	0.259 (0.259 to 0.259)	£117,714 (£117,648 to £117,781)

6.2.2 Other comparators and interventions

Using the manufacturer's revised model and the effectiveness results from the AVADO trial (Section 4.2.2.2.1), we explored approaches to formally incorporating a BEV+ DOC regimen as part of the analysis. The AVADO trial [27-36] evaluated the use of two alternative BEV+DOC regimens, differing in the dose of bevacizumab administered – 15 and 7.5 mg/kg of bevacizumab in addition to 100mg/m^2 docetaxel, administered every 3 weeks - by comparing outcomes attained with a regimen where 100mg/m^2 of docetaxel is administered every 3 weeks. From the alternative bevacizumab dosing schedules only the first dosing approach (15mg/kg bevacizumab administered) is recommended in the SPC and hence our additional analyses applied this dosing schedule.

In the absence of patient level data from the AVADO trial it was not possible to undertake a comparison of BEV+DOC versus DOC using an equivalent approach to that employed for the E2100 trial (i.e. using statistical extrapolation of the PFS and OS data). Consequently, it was necessary to link to the existing model and comparators using assumptions. As this was primarily an exploratory analysis, we made the assumption that DOC regimen in the AVADO trial was equivalent to the DOC regimen already in the model. In addition, given the limitations of the available information on the results of the AVADO trial, in modelling the relative effectiveness of the combination of bevacizumab and docetaxel, we used only the hazard ratio reported for PFS for BEV+DOC compared to DOC alone of 0.77 (0.64 to 0.93, see Section 4.2.2.2).

Other model parameters (such as mortality after progression and adverse events) were unchanged from the revised analyses provided by the manufacturer. These assumptions are clearly a shortcoming of the additional analyses and hence, while exploratory in nature, a certain amount of caution should be applied in subsequent interpretation of the ICER results.

It should be noted that no information is available on the mean or median number of chemotherapy cycles received in the AVADO trial. Thus, for the current analyses these regimens were costed by assuming the same duration of treatment as reported in the E2100

trial for bevacizumab and paclitaxel. Unit costs were changed for the respective drugs and the dosing considered for costing was based on the protocol dose reported in the AVADO trial.

The cost effectiveness results based on this analysis are shown in Table 29 for the comparison of BEV+DOC vs DOC.

Table 29: **Exploratory results on the cost effectiveness of BEV+DOC in comparison to DOC, as given by the relative measure of increased time to progression observed in the AVADO trial.**

	Results	***BEV+DOC*** *(relative effects form AVADO)*	***DOC*** *(effects equal to PAC qw)*
		ERG	*Base case*
Case 1— NHS list prices	Mean Life Years (yrs)	2.514	2.330
	Mean QALYs	1.361	1.225
	Mean Total Cost	£59,769	£25,057
	Incremental Life Years		*0.183*
	Incremental QALYs		*0.136*
	Incremental Cost		*£34,712*
	Cost per Life Year Gained		**£189,220**
	Cost per QALY Gained		**£254,530**

The results show that the ICER of BEV+DOC versus DOC is £254,530 per QALY. Sensitivity analyses to the list price of docetaxel show that a reduction of 20% or 30% does not affect overall conclusions (ICERs change to £249,467 and £246,936, respectively), mainly because docetaxel is being used in both arms.

6.2.3 Alternative ERG model structure

As previously identified in Section 5.2.2, the model submitted by the manufacturer has limited internal validity for OS when compared with corresponding estimates reported in the E2100 trial. The MS justified the approach applied in the model for OS (i.e. assuming that the rate of mortality post-progression was the same for all regimens) based on a number of considerations. Firstly, the lack of a statistical treatment difference reported between BEV+PAC or PAC alone based on a log-rank test (p=0.2441) from the stratified comparison of the Kaplan-Meier curves from E2100 of patients in the progressive health state. In addition, the manufacturer also argued that since the primary outcome of the RCT was PFS, the study was not powered to evaluate OS, and that the results of the RCT were diluted because of crossover to the intervention therapy after progression of disease. The ERG was advised by the clinical expert that these could be reasonable explanations of the trial results. Crossover and differing treatments after progression are problems with most current cancer trials and that gains in PFS rarely translate into differences in OS. However, since no data were collected on treatments received post-progression in the E2100 trial, it is not possible to

confirm or refute these potential explanations. Consequently, the ERG considered it was important to conduct sensitivity analyses where the model parameters are more closely calibrated with the trial results. Differences in OS are likely to be an important influence on estimates of cost-effectiveness and there may be other equally plausible explanations for the lack of differences in OS. For example, patients who survive longer without progression may be frailer at the time they do progress.

To explore these scenarios, the ERG has constructed an alternative decision model, based on an AUC approach that corresponds more closely to the observed trial results.

6.2.3.1 Methods of ERG alternative model

Given that the AUC approach is based on estimating the respective areas under the PFS and OS curves, the ERG model is restricted to a comparison of BEV+PAC vs. PAC qw, based on the survival curves from the E2100 trial.

PFS was modelled using the Gompertz survival function with proportional hazards, which was found by the MS to be the function with the best fit to the observed data. The ERG modelled OS in a similar way to the MS using the individual patient data, by exploring a range of parametric survival functions, either fitted independently to each arm of the trial, or fitted to both arms assuming proportional hazards.

The alternative model developed by the ERG is a Markov model built in Excel. PFS is estimated over 10 years in the same way as the MS model. However, in contrast to the manufacturer's assumption that the rate of mortality following progression is the same for all regimens, in the ERG model the OS for BEV+PAC and PAC qw is estimated separately and extrapolated over the 10 years based on the observed survival data from the E2100 trial using the survival function described above. The proportion of the cohort alive and in the 'progressed' state is then calculated as the difference between OS and PFS at each time point.

In the base-case ERG model, the costs of treatments are BNF list prices (i.e. Case 1). In a sensitivity analysis, PASA prices from the MS are used for PAC (i.e. Case 3). All other costs associated with the separate health states and the associated utilities are the same as in the MS model. The model is deterministic.

6.2.3.2 Results of ERG alternative model

Overall survival of BEV+PAC versus PAC

Alternative parametric distributions were fitted to the deaths for any cause observed in the E2100 trial: Exponential, Weibull, Lognormal (accelerated failure time metric) and Gompertz. Table 30 shows estimated coefficients of the functions, fitted independently to each arm or assuming proportional hazards. Table 30 also shows, for each function, the estimated mean survival time, for each arm (the area under the curve where a closed form solution exists).

Table 30: **Parameters of parametric survival functions fitted to overall survival**.

Function, coefficients (log scale) and predicted life expectancy (LE)	*PAC arm* Mean	SE	*BEV+PAC arm* Mean	SE
Exponential				
Intercept	-3.50	0.065	-3.57	0.034
Predicted LE	33.0 months	-	35.6 months	-
Weibull				
Intercept	-4.66	0.26	-5.39	0.3
Ln_p	0.300	0.056	0.434	0.056
Predicted LE	29.0 months		29.5 months	
Lognormal				
Intercept	3.12	0.059	3.22	0.056
Ln_sigma	0.037	0.049	-0.012	0.048
Predicted LE	38.8 months		40.6months	
Gompertz				
Intercept	-3.91	0.12	-4.21	0.038
Gamma	0.025	0.006	0.038	0.006
Predicted LE	No closed form solution		No closed form solution	
Gompertz	**Proportional hazards model**			
Intercept	-3.91	0.163		
Slope (BEV+PAC v PAC)	-0.068	0.091		
Gamma	0.032	0.004		

Figure 4 illustrates the survival curve for each parametric function, together with the Kaplan Meier estimates of survival. The Gompertz and Weibull functions have similar fit to the observed data but have differ in the extrapolation of survival beyond the follow-up in the trial. The Gompertz predicts a higher rate of mortality for the tail of the distribution than the Weibull model.

Various tests for the proportional hazards assumption for OS were carried out, but did not find consistently for or against this assumption. First, a plot of log (cumulative hazards) versus log(time) found that the survival curves cross early on, which violates the proportional hazards assumption. Second, the data were split at 6 months and 12 months and a test was carried out to see if there was an interaction between analysis time and the randomised treatment group. This found a significant interaction (p= 0.04), also indicating the proportional hazards assumption was incorrect. Third, a test based on Schoenfeld residuals did not find evidence against the proportional hazards assumption.

The base-case model uses the Gompertz function, fitted independently to each arm. A sensitivity analysis uses the Gompertz function with proportional hazards.

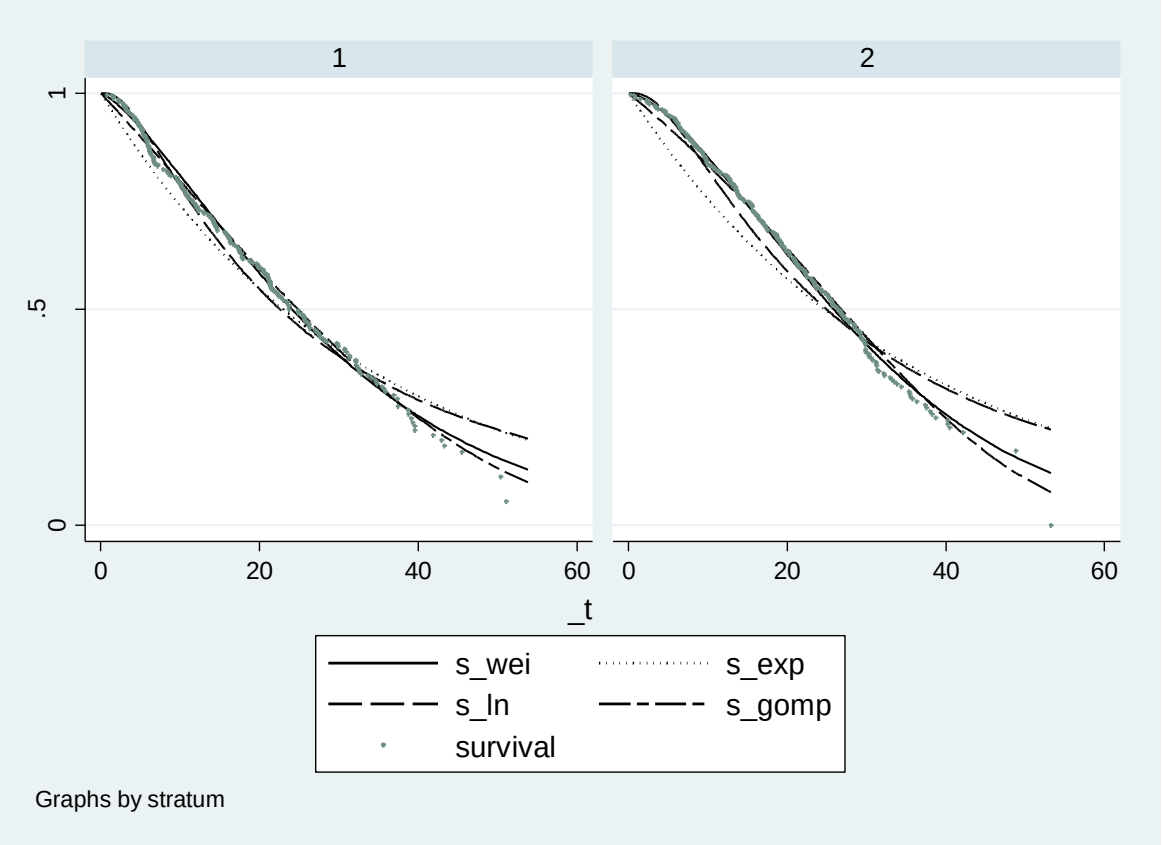


Figure 4: **Predicted parametric survival curves for OS for various functions and Kaplan Meier estimates from E2100 trial, by randomised treatment group**

Key (1= PAC, 2 = BEV+PAC). Time is measured in months. S_wei = Weibull, S:exp = exponential, S_ln = lognormal, s_gomp = Gompertz, Survival = Kaplan-Meier

Cost-effectiveness of BEV+PAC versus PAC

Table 31: Costs and QALYs of BEV+PAC versus PAC (10 years) with ERG model compared with MS model Table 31 shows the costs and QALYs of the treatments over 10 years in the ERG model compared with the MS model, using list prices for drugs. The estimates of incremental mean PFS are identical in the MS and ERG models, because they are based on the same parameters. The MS model, based on actual OS from the E2100 trial, estimates that the mean difference in OS is 0.03 years, while the MS model, based on extrapolating differences in PFS to OS, estimates a difference of 0.35 years. The MS model predicts that there is no difference in the time spent with progressed disease between the arms, while the ERG predicts that patients will spend 0.32 fewer years in the progressed state in the BEV+PAC arm than the PAC arm.

Sensitivity analyses with the ERG model using (i) PASA prices for PAC and (ii) assuming proportional hazards for OS did not reduce the ICER below £200,000 per QALY.

Table 31: Costs and QALYs of BEV+PAC versus PAC (10 years) with ERG model compared with MS model

	ERG model results			*MS model results*		
	BEV+PAC	*PAC*	*Incremental*	*BEV+PAC*	*PAC*	*Incremental*
Mean Life Years (yrs)	2.165	2.133	0.033	2.682	2.330	0.352
Mean Life Years in PFS (yrs)	1.000	0.644	0.356	1.041	0.686	0.355
Mean Life Years in Progression (yrs)	1.165	1.489	-0.323	1.641	1.645	-0.003
Mean QALYs	1.315	1.201	0.114	1.498	1.239	0.259
Mean QALY in PFS	0.791	0.531	0.260	0.759	0.499	0.260
Mean QALY in Progression	0.524	0.670	-0.145	0.739	0.740	-0.001
Mean Total Cost	£48,566	£18,891	£29,675	£56,473	£26,004	£30,469
Cost per QALY Gained (£)		£259,267			£117,803	

6.2.3.3 Conclusion of ERG alternative model

The revised model was based on OS estimates from the E2100 trial. In all scenarios tested the ICER of BEV+PAC versus PAC was over £200,000 per QALY.

6.2.4 Indirect treatment comparison

The ERG is aware of the limitations of the evidence base used to carry out the indirect treatment comparison, and that the MS evidence synthesis has made a reasonable attempt to compare indirectly the effectiveness of the different therapies using the available data. Nevertheless, the ERG has noted a number of potential weaknesses in the ITC (Section 4.2), and alternative analyses to address these issues are explored in this section. First, the MS ITC included a trial which did not report an ITT comparison of PFS outcomes [22]. Instead, non-randomised patients from another study were included in one of the arms. This study is excluded from the ERG revised analysis. Second, the MS ITC did not include relevant RCTs ('Will Weekly Win' trial [40] and AVADO [27-36]). Third,[i] the Bucher indirect comparison method is intended to compare two randomized trials via a common comparator. This is handled in

[i] Erratum. The following text was deleted: the hazard ratios are incorrectly calculated in the MS analysis. The MS calculated the mean hazard ratio (HR) according to the formula E(HR) = exp(μ), where HR is assumed to take a lognormal distribution log(HR)~Normal(μ,σ^2). This formula in fact estimates the median HR, not the mean. The correct formula is E(HR) = exp(μ+σ^2/2). The difference can be substantial when the standard error is large. Fourth,

the revised analysis by limiting the indirect comparisons to BEV+PAC versus PAC q3w (via a common comparator or PAC qw) and BEV+DOC versus PAC q3w (via a common comparator of DOC q3w).

The revised ITC remains exploratory, because of the weak evidence base. Many studies are only available as abstracts. Jones et al [3] has a high proportion of patients who had already received chemotherapy for mBC (Section 4.2.2). There are no direct comparisons of PAC qw versus DOC or PAC qw versus BEV+DOC. These weaknesses remain as limitations in the revised analysis.

Results of revised ITC

Figure 5 shows the network of trials used in the revised indirect comparison for hazard ratios of PFS, and Table 32 shows the direct and indirect estimates of the hazard ratios.

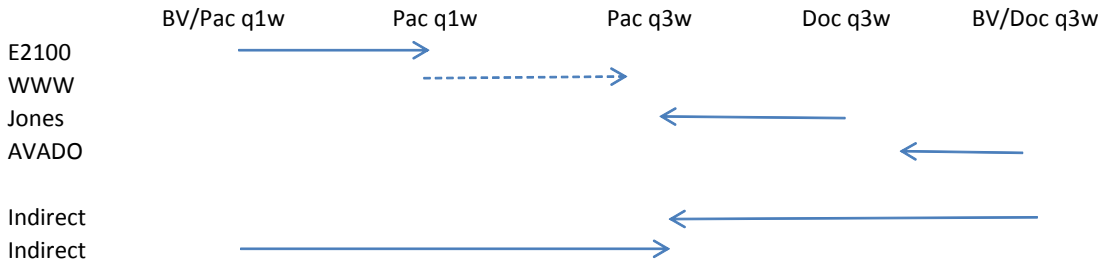


Figure 5: Network diagram for PFS hazard ratios

An arrow from therapy A to B indicates A is more effective than B. A solid line represents a significant hazard ratio, a broken line represents a non-significant effect at the 5% level.

The revised exploratory ITC finds that PAC 3w is significantly less effective than either BEV+PAC or BEV+DOC for PFS, with a HR of 0.40 (95% CI 0.31 to 0.51) and 0.48 (95% CI 0.37 to 0.63), respectively. Figure 6 and Table 33 shows that PAC 3w is also significantly less effective than either BEV+PAC or BEV+DOC for OS, with a HR of 0.76 (95% CI 0.57-0.99) and 0.72 (95% CI 0.51-0.99) respectively.

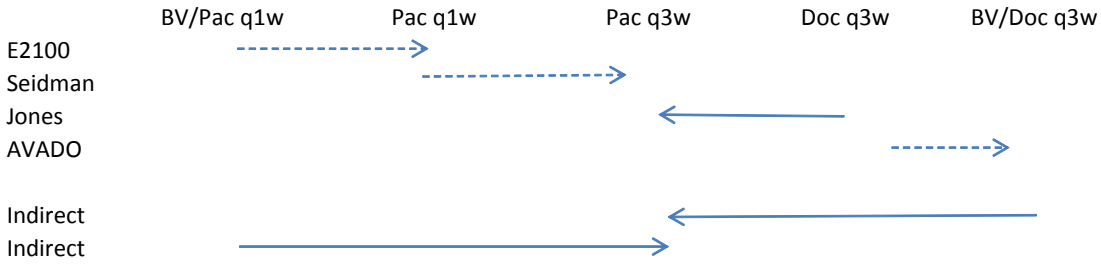


Figure 6: Network diagram for OS hazard ratios

An arrow from therapy A to B indicates A is more effective than B. A solid line represents a significant hazard ratio, a broken line represents a non-significant effect at the 5% level.

Table 32: Direct and indirect comparisons of hazard ratios of PFS

		LN(HR)	*SE(LN HR)*	*HR*	*LCL*	*UCL*		Comments
A v B	HR(BEV/Pac q1w vs Pac q1w)	-0.732	0.113	0.48	0.38	0.60	Study E2100	
C v B	HR(Pac q3w vs Pac q1w)			na			Seidman et al. JCO (2008)	Not randomised comparison
C v B	HR(Pac q3w vs Pac q1w)	0.083	0.048	1.09	0.99	1.20	WillWeekly Win	No SE reported, p-value 0.04
C v D	HR(Pac q3w vs Doc q3w)	0.489	0.104	1.64	1.33	2.02	Jones et al JCO (2005)	High % of 'second line' patients
E v D	HR(BV/Doc q3w vs Doc q3w)	-0.265	0.085	0.77	0.64	0.9	AVADO	
A v C	HR(BV/Pac q1w vs Pac q3w)	-0.815	0.123	0.45	0.35	0.57	Indirect	
E v C	HR(BV/Doc q3w vs Pac q3w)	-0.754	0.135	0.47	0.36	0.62	Indirect	

Table 33: Direct and indirect comparisons of hazard ratios of OS

		LN(HR)	*SE(LN HR)*	*HR*	*LCL*	*UCL*		*Comments*
A v B	HR(BV/Pac q1w vs Pac q1w)	-0.144	0.094	0.87	0.72	1.05	Study E2100	
C v B	HR(Pac q3w vs Pac q1w)	0.152	0.104	1.17	0.95	1.44	Seidman et al. JCO (2008)	ITT comparison
C v B	HR(Pac q3w vs Pac q1w)			na			WillWeekly Win	No OS results reported. Unknown % 2nd line patients
C v D	HR(Pac q3w vs Doc q3w)	0.338	0.102	1.41	1.15	1.73	Jones et al JCO (2005)	High % of 'second line' patients
E v D	HR(BV/Doc q3w vs Doc q3w)	0.021	0.130	1.03	0.79	1.33	AVADO	

A v C	HR(BV/Pac q1w vs Pac q3w)	-0.295	0.140	0.75	0.56	0.99	Indirect
E v C	HR(BV/Doc q3w vs Pac q3w)	-0.317	0.165	0.74	0.52	1.00	Indirect

Conclusion of revised ITC

The revised ITC undertaken by the ERG found that the hazard ratios of BEV+PAC and BEV+DOC, relative to PAC q3w, are similar. This implies that, given the available indirect evidence, BEV+PAC and BEV+DOC would be expected to be of similar effectiveness. The most cost-effective strategy in the choice between these two combination therapies would then depend on the acquisition and administration cost of the taxanes and the profile of adverse drug-related events. The MS reached the same conclusion, and stated that docetaxel was more expensive than paclitaxel. However, the ERG notes that, as docetaxel is soon to come off patent, the price of docetaxel is likely to fall in the near future.

7 Discussion

7.1 Summary of clinical effectiveness issues

Despite some methodological limitations, the E2100 trial provides direct evidence to suggest that the addition of bevacizumab to qw paclitaxel can increase objective response and progression free survival in the first-line treatment of metastatic breast cancer. The same trial fails to show a benefit in terms of overall survival. Whether this is a true null finding or due to crossover between treatment arms cannot be established as relevant data were not collected.

The intervention specified in the scope issued by NICE required an evaluation of bevacizumab in combination with a taxane (including both paclitaxel and docetaxel). However, the manufacturer provided several different reasons for excluding the well-conducted AVADO trial that evaluated the addition of bevacizumab to q3w docetaxel. Since this is the taxane currently recommended for first-line treatment of patients with advance breast cancer in existing NICE guidelines,[2] the ERG extracted the limited available data from published AVADO abstracts. In terms of response rate and PFS, the AVADO trial reported a markedly smaller benefit of adding bevaczumab to docetaxel than was reported for adding

bevacizumab to qw paclitaxel in the E2100 trial. One explanation for this difference might be potentially greater progression benefits associated with docetaxel monotherapy relative to paclitaxel monotherapy,[2, 3] though without any head-to-head comparison of the bevacizumab-taxane combinations, any number of confounding factors could also be responsible. The AVADO trial also reported no significant effect of combination therapy versus docetaxel in terms of overall survival.

Though the manufacturer did not look at the direct effects of adding bevacizumab to docetaxel as investigated in the AVADO trial, they did attempt to indirectly compare docetaxel monotherapy against combined bevacuzimab and qw paclitaxel. This indirect analysis indicated a statistically significant benefit of the bevacuzimab and qw paclitaxel combination over the currently recommended first-line treatment of docetaxel monotherapy. However, given important limitations around the evidence selected and the methods used, this finding cannot be considered reliable. Similar limitations apply to the indirect comparison of combined bevacuzimab and qw paclitaxel against combined gemcitabine and q3w paclitaxel.

7.2 Summary of cost effectiveness issues

Based on the E2100 trial, the manufacturer compared BEV+PAC versus PAC qw. This analysis concluded that the ICER was £118,000 per QALY, with NHS list prices, or £77,000 per QALY, with PASA (discounted) prices and a 10g (7 cycles) cap on the cost to the NHS of BEV per patient. Sensitivity analyses conducted by the manufacturer did not greatly change these conclusions.

Based on an indirect treatment comparison, the manufacturer concluded that the effectiveness of other comparators they considered (GEM+PAC and DOC q3w) was similar to PAC qw. On this basis, the manufacturer found that PAC qw was less costly and with more QALYs than either of those comparators. In a clarification, PAC q3w was found to be less costly and with fewer expected QALYs than PAC qw. The manufacturer stated that BEV+DOC would be more costly than BEV+PAC with no greater effectiveness, and so would be dominated although this was not formally shown in the modelling.

While the manufacturer conducted a number of sensitivity analyses, the ERG noted that several uncertainties remained. The ERG conducted further analyses to explore these issues.

The ERG understands that the Department of Health has not accepted the arrangement to cap the costs of BEV, in which case this arrangement should not be represented in any of the analyses. The ERG carried out a further analysis assuming no cap on the cost of BEV. This

found that the PASA discount (without the cap) made very little difference to the incremental costs of BEV+PAC versus PAC, compared with using NHS list prices.

Docetaxel is soon to come off-patent (November 2010).[j] The acquisition cost would be expected to fall but the magnitude is as yet unknown. Further analyses by the ERG found that the acquisition cost of docetaxel had very little effect on the ICER of DOC versus BEV+PAC. However, the price of DOC may be important in any comparison of the cost-effectiveness of taxane monotherapies (PAC versus DOC).

Alternative assumptions about utility values for the health states did not markedly affect the results.

The ERG evaluated BEV+DOC versus DOC based on the results of the AVADO trial. This found that the ICER was more than £250,000 per QALY.

The manufacturer's model was calibrated to the E2100 trial results for PFS, assuming survival time from the date of progression was the same in both initial treatments. While this assumption may be reasonable, no sensitivity analysis was undertaken to test its importance to the results. The ERG constructed an alternative model that was calibrated to the E2100 results for OS. The ICER of BEV+PAC versus PAC was over £250,000 per QALY in the revised model.

The manufacturer acknowledged many of the limitations of the indirect treatment comparison given the weak evidence base. The ERG also noted a number of errors in this analysis, in terms of selection of trials and the methods. The revised ERG analysis found that PAC q3w seemed to be the least effective treatment, and found that the hazard ratio of BEV+PAC versus PAC q3w was similar to the hazard ratio of BEV+DOC versus PAC q3, for both PFS and OS. This implies that a choice between BEV+PAC and BEV+DOC strategies might be based on the acquisition and administration cost of the drugs and their expected profile of adverse events. However, this simple analysis does not inform whether either PAC or DOC monotherapy might be more cost-effective than BEV in combination with either taxane.

7.3 Implications for research

As noted earlier in this report, available evidence on the clinical efficacy of bevacizumab plus taxane therapy is currently limited to two randomized controlled trials: E2100 (BEV+PAC qw

[j] Amended from May 2010 by ERG on advice of Roche

vs. PAC qw) and AVADO (BEV+DOC q3w vs. DOC q3w+placebo). A third trial (RIBBON-1) includes a BEV+taxane subgroup but has yet to fully report its findings. There are no known RCTs evaluating the addition of bevacizumab to routine 3-weekly paclitaxel treatment in mBC.

Because of limitations in the existing evidence base, the reliability of indirect estimates of efficacy between bevacizumab plus taxane regimens and other relevant comparators remains uncertain. While further direct trial comparisons (such as BEV+PAC qw vs DOC q3w) might provide more accurate estimates of efficacy of bevacizumab plus taxanes in mBC, the value of the additional information they would provide to the overall evidence base may be limited. Therefore, rather than undertaking new evaluations, resources might be better used to investigate other uncertainties, such as the impact of crossover effects on overall survival among currently available trials.

8 References

1. Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin Epidemiol* 1997;50:683-91.

2. *Clinical Guidelines CG81: Advanced breast cancer:diagnosis and treatment*. London: National Institute for Health and Clinical Excellence (NICE); 2009.

3. Jones SE, Erban J, Overmoyer B, Budd GT, Hutchins L, Lower E, et al. Randomized phase III study of docetaxel compared with paclitaxel in metastatic breast cancer. *J Clin Oncol* 2005;23:5542-51.

4. *Summary of Product Characteristics- Paclitaxel 6 mg/ml concentrate for solution for infusion (Hospira UK Ltd)*. 2009. [cited 2009 October 27th]. Available from: http://www.medicines.org.uk/emc/printfriendlydocument.aspx?documentid=15842&compa.

5. Cameron D. Bevacizumab in the first-line treatment of metastatic breast cancer. *Eur J Cancer Suppl* 2008;6:21-28.

6. Gray R, Bhattacharya S, Bowden C, Miller K, Comis RL. Independent review of E2100: a phase III trial of bevacizumab plus paclitaxel versus paclitaxel in women with metastatic breast cancer. *J Clin Oncol* 2009;27:4966-72.

7. Klencke BJ, Bhattacharya S, Samant MK, Gralow JR, Dickler MN, Cobleigh MA, et al. Independent review of E2100 progression-free survival (PFS) with the addition of bevacizumab (B) to paclitaxel (P) as initial chemotherapy for metastatic breast cancer (MBC). *J Clin Oncol* 2008;ASCO annual meeting proceedings:50, Abstract 1036.

8. Miller K, Wang M, Gralow J, Dickler M, Cobleigh M, Perez EA, et al. Paclitaxel plus bevacizumab versus paclitaxel alone for metastatic breast cancer. *N Engl J Med* 2007:2666-76.

9. Miller K, Wang M, Gralow J, Dickler M, Cobleigh MA, Perez EA, et al. First-line bevacizumab and paclitaxel in patients with locally recurrent or metastatic breast cancer: a randomized, phase III trial coordinated by the Eastern Cooperative Oncology Group (E2100). *Eur J Cancer* 2005:77-78.

10. Miller KD. E2100: a phase III trial of paclitaxel versus paclitaxel/bevacizumab for metastatic breast cancer. *Clinical Breast Cancer* 2003;3:421-2.

11. Miller KD, Wang W, Gralow J. A randomized phase III trial of paclitaxel versus paclitaxel plus bevacizumab as first-line therapy for locally recurrent or metastatic breast cancer: a trial coordinated by the Eastern Cooperative Oncology Group (E2100) *Breast Cancer Res Treat* 2005;94 (Suppl 1):S6 Abstract 3.

12. Wagner L, Wang M, Miller K. Health-related quality of life among patients with metastatic breast cancer receiving paclitaxel versus paclitaxel plus bevacizumab: results from the eastern cooperative oncology group (ECOG) study E2100. *Breast Cancer Res Treat* 2006;S239:Abstract 5078.

13. Zon R, Miller KD, Wang M, Gralow J, Dickler M, Cobleigh M, et al. A randomized phase III trial of paclitaxel with or without bevacizumab as first-line therapy for locally recurrent or metastatic breast cancer: Eastern Cooperative Oncology Group trial E2100. *Eur J Cancer Suppl* 2006;4:47.

14. Biganzoli L, Cortes-Funes H, Thomssen C, Pritchard KI, Pierga J, Kupp A, et al. Tolerability and efficacy of First-line bevacizumab (B) plus chemotherapy (CT) in elderly patients with advanced breast cancer (aBC): subpopulation analysis of the MO19391 study. *J Clin Oncol* 2009;27 (Suppl):Abstract 1032.

15. Cortes-Funes H, Pritchard KI, Biganzoli L, Thomssen C, Pierga J, Koza I, et al. Multinational study (n=2041) of first-line bevacizumab (Bev) plus taxane-based chemotherapy (CT) for locally recurrent or metastatic breast cancer (LR/mBC): updated results of MO19391. *Eur J Cancer Suppl* 2009;7:265-66.

16. Pierga J, Pritchard KI, Cortes-Funes H, Biganzoli L, Thomssen C, Franke FA, et al. MO19391: an open-label safety study of bevacizumab plus taxane-based therapy as first-line treatment of patients with locally recurrent (LR) or metastatic breast cancer (MBC). *J Clin Oncol* 2008;26 (Suppl):Abstract 1140.

17. Pierga J, Pritchard KI, Thomssen C, Cortes-Funes H, Biganzoli L, Padrik P, et al. Safety and efficacy of 1st-line bevacizumab (B) plus chemotherapy (CT) for locally recurrent or metastatic breast cancer (LR/mBC): analysis of MO19391 according to CT. *J Clin Oncol* 2009;27 (Suppl):Abstract 1033.

18. Smith I, Biganzoli L, Cortes-Funes H, Franke F, Mostafa Kamel Y, Pierga JY, et al. MO19391: an open-label safety study of bevacizumab plus taxane monotherapy or in combination as first-line treatment of patients with locally recurrent or metastatic breast cancer (LR or MBC). *Eur J Cancer Suppl* 2007;5:221.

19. Smith I, Biganzoli L, Cortés-Funes H, Stroyakovskiy D, Franke FA, Chlistalla A, et al. Primary analysis of study MO19391, an open-label safety study of bevacizumab plus taxane-based therapy as first-line treatment of patients with locally recurrent (LR) or metastatic breast cancer (mBC). *Cancer Res* 2009;69 (Suppl):Abstract 4118.

20. Thomssen C, Pierga J, Pritchard K, Biganzoli L, Cortes-Funes H, Petrakova K, et al. First-Line bevacizumab (Bev) combination therapy in triple-negative (TN) locally recurrent/metastatic breast cancer (LR/MBC): subpopulation analysis of study MO19391 in >2000 patients (Pts). *Cancer Res* 2009;69 (Suppl):Abstract 6093.

21. Albain KS, Nag SM, Calderillo-Ruiz G, Jordaan JP, Llombart AC, Pluzanska A, et al. Gemcitabine plus Paclitaxel versus Paclitaxel monotherapy in patients with metastatic breast cancer and prior anthracycline treatment. *J Clin Oncol* 2008;26:3950-7.

22. Seidman AD, Berry D, Cirrincione C, Harris L, Muss H, Marcom PK, et al. Randomized phase III trial of weekly compared with every-3-weeks paclitaxel for metastatic breast cancer, with trastuzumab for all HER-2 overexpressors and random assignment to trastuzumab or not in HER-2 nonoverexpressors: final results of Cancer and Leukemia Group B protocol 9840. *J Clin Oncol* 2008;26:1642-9.

23. *Summary of Product Characteristics- Avastin 25 mg/ml concentrate for solution for infusion (Roche Products Ltd)*. 2009. [cited 2010 May 10th]. Available from: http://emc.medicines.org.uk/medicine/15748/SPC/Avastin+25mg+ml+concentrate+for+solution+for+infusion/.

24. Dieras V, Glaspy J, Brufsky A, Bondarenko I, Lipatov O, Perez EA, et al. Efficacy in patient subgroups in RIBBON-1, a randomized, double-blind, Phase III trial of chemotherapy with or without bevacizumab (B) for first-line treatment of HER2-negative locally recurrent or metastatic breast cancer (MBC). *Eur J Cancer Suppl* 2009;7:265.

25. Robert N, Dieras V, Glaspy J, Brufsky A, Bondarenko I, Lipatov O, et al. Clinical benefit rate and time to response in RIBBON-1, a randomized, double-blind, phase III trial of chemotherapy with or without bevacizumab (B) for the first-line treatment of HER2-negative

locally recurrent or metastatic breast cancer (MBC). *Cancer Res* 2009;69 (Suppl):Abstract 6084.

26. Robert N, 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 1st-line treatment of HER2-negative locally-recurrent or metastatic breast cancer (MBC). *J Clin Oncol* 2009;27 (Suppl):Abstract 1005.

27. Chan A, Vanlemmens L, Conte PF, Beith J, Samonigg H, Verma S, et al. Efficacy of bevacizumab (BV) plus docetaxel (D) does not correlate with hypertension (HTN), or G-CSF use in patients (pts) with locally recurrent (LR) or metastatic breast cancer (mBC) in the AVADO phase III study. *Cancer Res* 2009;69 (Suppl):Abstract 1027.

28. Cortés J, Pivot X, Schneeweiss A, Wilcken N, Ellis P, Linderholm BK, et al. Safety of surgery in patients (pts) with locally recurrent (LR) or metastatic breast cancer (mBC) treated with docetaxel (D) plus bevacizumab (BV) or placebo (PL) in the AVADO phase III study *Cancer Res* 2009;69 (Suppl):Abstract 1030.

29. Dirix LY, Romieu G, Provencher L, Grimes D, de Souza Viana L, Paterson A, et al. Safety of bevacizumab (BV) plus docetaxel (D) in patients (pts) with locally recurrent (LR) or metastatic breast cancer (mBC) who developed brain metastases during the AVADO phase III study. *Cancer Res* 2009;69 (Suppl):Abstract 4116.

30. Fumoleau P, Greil R, Rayson D, Müller V, Barni S, Aleknavicius E, et al. Maintenance therapy of bevacizumab (BV) results in superior PFS compared with placebo (PL) in the AVADO trial (BV + docetaxel [D] vs D + PL in first-line HER2-negative locally recurrent [LR] or metastatic breast cancer [mBC]). *Cancer Res* 2009;69 (Suppl):Abstract 903.

31. Greil R, Im YH, Pienkowski T, Wardley A, Awada A, Ciruelos E, et al. Quality of life (QoL) in patients (pts) treated with bevacizumab (BV) and taxane therapy for locally recurrent (LR) or metastatic breast cancer (mBC). *Eur J Cancer Suppl* 2009;7:266.

32. Harbeck N, Chan A, ten Bokkel Huinink D, Chollet P, Gimenes D, Fabiani C, et al. No clinical evidence for increase in tumour aggressiveness or metastatic spread in patients with metastatic breast cancer (mBC) treated with bevacizumab (BV) and docetaxel (D) in the phase III AVADO study. *Cancer Res* 2009;69 (Suppl):Abstract 6086.

33. Miles D, Chan A, Romieu G, Dirix L, Cortés J, Pivot X, et al. Final overall survival (OS) results from the randomised, double-blind, placebo-controlled, phase III AVADO study of

bevacizumab (BV) plus docetaxel (D) compared with placebo (PL) plus D for the first-line treatment of locally recurrent (LR) or metastatic breast cancer (mBC) *Cancer Res* 2009;69 (Suppl):Abstract 41.

34. Miles D, Chan A, Romieu G, Dirix LY, Cortes J, Pivot X, et al. Randomised, double-blind, placebo-controlled, phase III study of bevacizumab with docetaxel or docetaxel with placebo as first-line therapy for patients with locally recurrent or metastatic breast cancer (mBC): AVADO. *J Clin Oncol* 2008;26 (Suppl):Abstract LBA1011.

35. Pivot X, Verma S, Thomssen C, Passos-Coelho JL, Latini L, Ciruelos E, et al. Clinical benefit of bevacizumab (BV) + first-line docetaxel (D) in elderly patients (pts) with locally-recurrent or metastatic breast cancer (mBC): AVADO study. *J Clin Oncol* 2009;27 (Suppl):Abstract 1094.

36. Wardley A, Lohrisch C, Joy AA, Young S, Yi FF, Marme A, et al. Effect of anticoagulation therapy on bleeding and thromboembolic events (TEs) in the AVADO phase III study of bevacizumab (BV) plus docetaxel (D) in locally recurrent (LR) or metastatic breast cancer (mBC) *Cancer Res* 2009;69 (Suppl):Abstract 1035.

37. Geberth M, Foerster F, Klare P, Schneeweiss A, Schumacher C, Hahn L, et al. Efficacy and safety of first-line bevacizumab (BEV) combined with paclitaxel (PAC) according to age: subpopulation analysis of a large, multicenter, non-interventional study in patients (pts) with HER2-negative metastatic breast cancer (MBC). *Cancer Res* 2009;69:Abstract nr 6085.

38. Hurvitz SA, Kabbinavar FF, Allen HJ, Moroose RL, D. C, Hagenstad C, et al. A phase II trial of docetaxel with bevacizumab as first line therapy for Her2/neu negative metastatic breast carcinoma. In: *30th Annual San Antonio Breast Cancer Symposium*; 2007 December 13-16; San Antonio, TX. 2007. p. Poster 4062.

39. Ramaswamy B, Elias AD, Kelbick NT, Dodley A, Morrow M, Hauger M, et al. Phase II trial of bevacizumab in combination with weekly docetaxel in metastatic breast cancer patients. *Clinical Cancer Research* 2006;12:3124-9.

40. Verrill M, Lee J, Cameron DA, Agrawal R, Coleman RE, McAdam K. Anglo-Celtic IV: First results of a UK National Cancer Research Network randomised phase 3 pharmacogenetic trial of weekly versus 3 weekly paclitaxel in patients with locally advanced or metastatic breast cancer (ABC). *J Clin Oncol* 2007;25:LBA1005.

41. Song F, Altman DG, Glenny A-M, Deeks JJ. Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published meta-analyses. *BMJ* 2003;326:472.

42. Tudur C, Williamson PR, Khan S, Best LY. The value of the aggregate data approach in meta-analysis with time-to-event outcomes. *Journal of the Royal Statistical Society Series A* 2001;164:357-70.

43. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;273:408-12.

44. Grimshaw J, Wilson B, Campbell M, Eccles M, Ramsay C. Epidemiological methods. In: Fulop N, Allen P, Clarke A, Black N, editors. *Studying the organisation and delivery of health services: research methods*. London: Routledge; 2001. p. 56-72.

45. *Technology Appraisal (TA)162: Erlotinib for the treatment of non-small-cell lung cancer*. London: National Institute for Health and Clinical Excellence (NICE); 2008.

46. Dedes KJ, Matter-Walstra K, Schwenkglenks M, Pestalozzi BC, Fink D, Brauchli P, et al. Bevacizumab in combination with paclitaxel for HER-2 negative metastatic breast cancer: an economic evaluation. *Eur J Cancer* 2009;45:1397-406.

47. Sparano JA, Wang M, Martino S, Jones V, Perez EA, Saphner T, et al. Weekly paclitaxel in the adjuvant treatment of breast cancer. *N Engl J Med* 2008;358:1663-71.

48. Cooper NJ, Abrams KR, Sutton AJ, Turner D, Lambert PC. A Bayesian approach to Markov modelling in cost-effectiveness analyses: application to taxane use in advanced breast cancer. *J Roy Stat Soc a Sta* 2003;166:389-405.

49. Brown RE, Hutton J. Cost-utility model comparing docetaxel and paclitaxel in advanced breast cancer patients. *Anticancer Drugs* 1998;9:899-907.

50. Joint Formulary Committee. *British National Formulary*. London: : British Medical Association and Royal Pharmaceutical Society of Great Britain; 2009.

51. Curtis L. *Unit Costs of Health and Social Care 2008*. Canterbury: Personal Social Services Research Unit (PSSRU); 2008.

52. Coon JT, Hoyle M, Green C, Liu Z, Welch K, Moxham T. *Bevacizumab, sorafenib tosylate, sunitinib and temsirolimus for renal cell carcinoma: a systematic review and*

economic evaluation. London: National Institute for Health and Clinical Excellence (NICE); 2008. Available from: http://www.nice.org.uk/nicemedia/pdf/RenalCellCarcinomaACDAR.pdf

53. Remak E, Brazil L. Cost of managing women presenting with stage IV breast cancer in the United Kingdom. *Br J Cancer* 2004;91:77-83.

54. Philips Z, Bojke L, Sculpher M, Claxton K, Golder S. Good practice guidelines for decision-analytic modelling in health technology assessment: a review and consolidation of quality assessment. *Pharmacoeconomics* 2006;24:355-71.

55. *Guide to the methods of technology appraisal*. London: National Institute for Clinical Excellence; 2008. Available from: http://www.nice.org.uk/media/B52/A7/TAMethodsGuideUpdatedJune2008.pdf

56. Magazzini L, Pammolli F, Riccaboni M. Dynamic competition in pharmaceuticals. Patent expiry, generic penetration, and industry structure. *Eur J Health Econ* 2004;5:175-82.

9 Appendices

Appendix 1: Additional data extraction of RCTs: AVADO and RIBBON-1

Study	Participants	Intervention/compactor	Results
AVADO [27-36] **Type of publication** Abstract **Funding** Roche **Study design** Double- blind RCT	HER2-negative inoperable locally recurrent (LR) or metastatic breast cancer (mBC). No previous chemotherapy for advanced disease. **Number randomised** 736 **Age (mean)** ~54 years **Gender, n(%)** 736 (100%) **HER 2 (+), n(%)** 3 (0.41%) **Triple negative (ER-/PR-/HER2-), n (%)** NR **Previous adjuvant chemotherapy; n (%)** No chemotherapy 6 months prior to randomisation (≥ 12 months if taxane-based) **Previous anthracycline therapy, n (%)** NR	**Intervention 1 (Bevacizumab + Docetaxel); n = 248** Bevacizumab 7.5mg/kg + Docetaxel 100mg/m2; 3-weekly **Intervention 2 (Bevacizumab + Docetaxel); n= 247** Bevacizumab 15mg/kg + Docetaxel 100mg/m2; 3-weekly **Comparator (Docetaxel + placebo); n =241** Docetaxel 100mg/m2 + placebo; 3-weekly Docetaxel was given up to 9 cycles. Bevacizumab/placebo given until disease progression or unacceptable toxicity.	***EFFICACY*** ***Progression-free survival*** **Median time to progression (median follow-up 10.2 months)** BEV 7.5mg + DOC: 8.7 months BEV 15mg + DOC: 8.8 months DOC + Placebo: 8.0 months **Number of patients with events (median follow-up 10.2 months)** BEV 7.5mg + DOC: 149 BEV 15mg + DOC: 142 DOC + Placebo: 162 **Progression free survival (median follow-up 10.2 months)** BEV 7.5mg + DOC vs. Placebo + DOC: Hazard ratio (HR) = 0.79 (95% CI 0.63 to 0.98); p=0.03. BEV 15mg + DOC vs. Placebo + DOC: HR=0.72 (95% CI 0.57 to 0.90); p=0.01. **Updated analysis: median time to progression (with additional 18 months of follow-up)** BEV 7.5mg + DOC: 9.0 months BEV 15mg + DOC: 10.1 months DOC + Placebo: 8.2 months **Updated analysis: number of patients with events (with additional 18 months of follow-up)** BEV 7.5mg + DOC: 218 BEV 15mg + DOC:208 DOC + Placebo: 214 **Updated analysis: progression free survival (PFS) (with additional 18 months of follow-up)** BEV 7.5mg + DOC vs. Placebo + DOC: HR = 0.86 (95% CI 0.72 to 1.04); p =0.12. BEV 15mg + DOC vs. Placebo + DOC: HR= 0.77 (95% CI 0.64 to 0.93); p = 0.006

	Patients receiving 9 cycles of docetaxel BEV 15mg + DOC: 51% Placebo + DOC: 42%		***Overall response rate*** **Overall response rate, n (%) (median follow-up 10.2 months)** BEV 7.5mg + DOC: 111/201 (55%) BEV 15mg + DOC: 130/206 (63%) Placebo + DOC: 92/207 (44%) BEV 7.5mg + DOC vs. Placebo + DOC: 55% vs. 44%; p=0.03 BEV 15mg + DOC vs. Placebo + DOC : 63% vs. 44%; p=0.0001 **Median duration of response(median follow-up 10.2 months)** BEV 7.5mg + DOC:7.2 months BEV 15mg + DOC:7.0 months DOC + Placebo:6.4 months **Number of patients with events (median follow-up 10.2 months)** BEV 7.5mg + DOC:58 BEV 15mg + DOC:68 DOC + Placebo:53 BEV 7.5mg + DOC vs. Placebo + DOC: HR=0.74 (95% CI 0.51 to 1.08) BEV 15mg + DOC vs. Placebo + DOC : HR = 0.80 (95% CI 0.56 to 1.14) **Updated analysis: overall response rate, n (%) (with additional 18 months of follow-up)** BEV 7.5mg + DOC: 111/201 (55.2%) BEV 15mg + DOC: 132/206 (64.1%) Placebo + DOC: 96/207 (46.4%) BEV 7.5mg + DOC vs. Placebo + DOC: 55.2% vs. 46.4%; p=0.07 BEV 15mg + DOC vs. Placebo + DOC : 64.1% vs. 46.4%; p=0.0003 **Updated analysis: median duration of response(with additional 18 months of follow-up)** BEV 7.5mg + DOC: 8.5months BEV 15mg + DOC: 8.5 months DOC + Placebo: 6.9 months **Updated analysis: number of patients with events (with additional 18 months of follow-up)** BEV 7.5mg + DOC:72 BEV 15mg + DOC:81 DOC + Placebo: 65 BEV 7.5mg + DOC vs. Placebo + DOC: HR 0.88 (95% CI 0.63 to 1.24) BEV 15mg + DOC vs. Placebo + DOC : HR 0.80 (95% CI 0.58 to 1.11)

Time to treatment failure

Median time to treatment failure (median follow-up 10.2 months)
BEV 7.5mg + DOC:7.0months
BEV 15mg + DOC: 7.7 months
DOC + Placebo: 6.1months
Number of patients with events (median follow-up 10.2 months)
BEV 7.5mg + DOC:185
BEV 15mg + DOC: 186
DOC + Placebo: 188

BEV 7.5mg + DOC vs. Placebo + DOC: HR= 0.85 (95% CI 0.69 to 1.04)
BEV 15mg + DOC vs. Placebo + DOC : HR=0.79 (95% CI 0.65 to 0.97)

Updated analysis: median time to treatment failure (with additional 18 months of follow-up)
BEV 7.5mg + DOC: 7.7 months
BEV 15mg + DOC: 7.9 months
DOC + Placebo: 6.3 months

Updated analysis : Number of patients with events (with additional 18 months of follow-up)
BEV 7.5mg + DOC: 233
BEV 15mg + DOC:224
DOC + Placebo: 226

BEV 7.5mg + DOC vs. Placebo + DOC: HR =0.86 (95% CI 0.71 to 1.03)
BEV 15mg + DOC vs. Placebo + DOC : HR=0.79 (95CI: 0.66 to 0.95)

Overall survival
Overall survival (median) (median follow-up 25 months)
BEV 7.5mg + DOC vs. Placebo + DOC: HR = 1.05 (95% CI 0.81 to 1.36); p=0.72.
BEV 15mg + DOC vs. Placebo + DOC: HR= 1.03 (95% CI 0.79 to 1.33); p=0.85

One year survival, n(%)
BEV 7.5mg + DOC vs. Placebo + DOC: 81% vs. 76%; p=0.198.
BEV 15mg + DOC vs. Placebo + DOC: 84% vs. 76%.; p= 0.02.

ADVERSE EVENTS

Table 1: Adverse events (AEs), n (%) (Length of follow-up: NR)

	Placebo+ DOC	BEV 7.5mg +	BEV 15mg +

	(n=233)	DOC (n=250)	DOC (n=247)
Adverse event*	232 (99.6%)	250 (100%)	246 (99.6%)
NCI -CTC grade 3.4.5 adverse event*	156 (67.0%)	187 (74.8%)	183 (74.1%)
Serious adverse event*	76 (32.6%)	92 (36.8%)	104 (42.1%)
Adverse event leading to discontinuation (any study drug)*	62 (26.6%)	58 (23.2%)	69 (27.9%)
Deaths due to adverse event §	6 (2.6%)	9 (3.6%)	5 (2.0%)
Bleeding (all events)	66 (28.3%)	131 (52.4%)	135 (54.7%)
Neutropenia	45 (19.3%)	54 (21.6%)	53 (21.5%)
Febrile neutropenia	28(12.0%)	39 (15.6%)	45 (18.2%)
Venous Thromboembolism	18 (7.7%)	15 (6.0%)	18 (7.3%)
Hypertension	21(9.0%)	34 (13.6%)	44 (17.8%)
Wound healing complication	3 (1.3%)	8 (3.2%)	12 (4.9 %)
Proteinuria	4 (1.7%)	3 (1.2%)	8 (3.2%)
Abscess and fistula	1 (0.4%)	6 (2.4%)	6 (2.4%)
Congestive heart failure	1 (0.4%)	3 (1.2%)	2 (0.8%)
Gastrointestinal perforation	2 (0.9%)	1 (0.4%)	2 (0.8%)
Arterial thrombombolic events	2 (0.9%)	0	1 (0.4%)

*Patient with at least one event
§also including patients who died more than 21 days after last drug administration

Withdrawals due to toxicity:

Number of patients discontinuing treatment (median follow-up ~ 25 months), n (%)
BEV 7.5mg + DOC: 27 (10.9%)
BEV 15mg + DOC: 35 (14.2%)
DOC + Placebo: 29 (12.0%)

Median time to progression after discontinuation (Median follow-up ~25 months)
BEV 7.5mg + DOC: 6.4 months
BEV 15mg + DOC: 6.8 months
DOC + Placebo: 3.3 months

Progression free survival (PFS) after discontinuation (median follow-up ~ 25 months)
BEV 7.5mg + DOC vs. Placebo + DOC: HR =0.71 (95% CI 0.40 to1.27)
BEV 15mg + DOC vs. Placebo + DOC: HR= 0.73 (95% CI0.42 to1.24)

QUALITY OF LIFE

Table 2: Total FACT-B Score: Mean score (changes from baseline); (n. of patients). (Length of follow-up: NR)

	Placebo+ DOC (n=241)	BEV 7.5mg + DOC (n=248)	BEV 15mg + DOC (n=247)
Baseline	21.69; (n=215)	21.98; (n=219)	22.11; (n=222)
Cycle 3	21.08 (-0.61); (n=181)	22.92 (0.94); (n=191)	22.77 (0.66); (n=198)
Cycle 5	20.95 (-0.74); (n=155)	21.96 (-0.02); (n=172)	22.05 (-0.06); (n=183)
Cycle 11	20.95 (-0.74) (n=81)	22.39 (0.41); (n=94)	22.44 (0.33); (n=114)

Study	Participants	Intervention/compactor	Results
RIBBON-1[24-26] **Type of publication** Abstract **Funding** Genentech **Study design** Double- blind RCT	HER2-negative previously untreated LR or mBC **Number randomised** n (total)= 1237 The taxane cohort (n=307) **Age (mean)** NR **Gender, n(%)** NR **HER 2 (+), n(%)** NR **Triple negative (ER-/PR-/HER2-), n (%)** NR **Previous adjuvant chemotherapy; n (%)** NR **Previous anthracycline therapy, n (%)** NR	**Intervention (Bevcizumab + Taxane), n= 203** Bevcizumab 15mg/kg + Docetaxel 75-100mg/m2 or protein-bound paclitaxel 260mg/m2 ; 3-weekly Number of cycles: NR **Comparator (Taxane+ Placebo), n = 104** Docetaxel 75-100mg/m2 or protein-bound paclitaxel 260mg/m2 + placebo; 3-weekly Number of cycles: NR Treatment could be continued up to 2 years in the absence of disease progression or unacceptable toxicity.	***EFFICACY*** **Patients with measurable disease at baseline** BEV + Taxane: 161 (79.3%) Taxane + placebo: 85 (81.7%) **Progression free survival (PFS)** BEV + Taxane: 9.2 months Taxane + placebo: 8.0 months HR= 0.64, p<0.0001 **Objective response rate, n (%)** BEV + Taxane: 81/161 (50.3%) Taxane + placebo: 30/85 (35.4%) BEV + Taxane vs. Taxane + placebo: 50.3% vs. 35.3%; p = 0.03 **Clinical benefit rate (complete or partial response or stable disease) at week 24, n (%)** BEV + Taxane: 146/203 (71.9%) Taxane + placebo: 68/104 (65.4%) BEV + Taxane vs. Taxane + placebo: 71.9% vs. 65.4%; p=0.24 **Duration of objective response (median)** BEV + Taxane: 8.4months Taxane + placebo: 8.6months BEV + Taxane vs. Taxane + placebo: HR= 0.75 (95% CI 0.45 to 1.27); p=0.28 **Overall survival** NR ***ADVERSE EVENTS*** Adverse events (no. patients (%)) NR **Withdrawals due to adverse events** NR

NR: Not reported

Appendix 2: Quality Assessment of the economic model

Quality criterion	*Question(s)*	*Response* (Y, N, or NS)	*Comments*
S1	Is there a clear statement of the decision problem?	*Y*	
	Is the objective of the evaluation and model specified and consistent with the stated decision problem?	*Y*	
	Is the primary decision-maker specified?	*Y*	The report is written for NICE.
S2	Is the perspective of the model stated clearly?	*Y*	NHS and PSS perspective were stated.
	Are the model inputs consistent with the stated perspective?	*Y*	
	Has the scope of the model been stated and justified?	*Y*	
	Are the outcomes of the model consistent with the perspective, scope and overall objective of the model?	*Y*	
S3	Is the structure of the model consistent with a coherent theory of the health condition under evaluation?	*Y*	Structure of the model is developed as per the existing UK guidelines.
	Are the sources of data used to develop the structure of the model specified?	*N*	The manufacturer has not referenced previous cost effectiveness studies, although the model structure used is coherent with these.
	Are the causal relationships described by the model structure justified appropriately?	*Y*	
S4	Are the structural assumptions transparent and justified?	*Y*	
	Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	*N*	The assumption of equal mortality after progression for the alternative treatments being evaluated was not explored by the manufacturer in the submission.
S5	Is there a clear definition of the options under evaluation?	*Y*	
	Have all feasible and practical options been evaluated?	*N*	The intervention evaluated is BEV+PAC. The model makes pairwise comparisons with: PAC qw, DOC and GEM+PAC. However, several interventions in the NICE scope were not included (PAC q3w, DOC 100mg/m2, BEV+DOC).
	Is there justification for the exclusion of feasible options?	*N*	The manufacturer states that BEV+DOC is not a cost-effective option, and that DOC 100mg/m^2 is not relevant to NHS practice. However, the ERG considers that

			these options should not have been excluded.
S6	Is the chosen model type appropriate given the decision problem and specified causal relationships within the model?	*Y*	A Markov model is deemed appropriate to represent the decision problem. However, the ERG considers there is only limited evidence to support the assumption of equal mortality after progression and that alternative assumptions are also plausible. A more conservative approach (i.e. that the gains in PFS may not be wholly translated into OS gains) was not explored by the manufacturer in the submission.
S7	Is the time horizon of the model sufficient to reflect all important differences between options?	*Y*	
	Are the time horizon of the model, the duration of treatment and the duration of treatment effect described and justified?	*Y*	
S8	Do the disease states (state transition model) or the pathways (decision tree model) reflect the underlying biological process of the disease in question and the impact of interventions?	*Y*	
S9	Is the cycle length defined and justified in terms of the natural history of disease?	*Y*	The cycle length of the Markov model is monthly.
D1	Are the data identification methods transparent and appropriate given the objectives of the model?	*Y*	With a few limitations, most of the data sources are well described and justified.
	Where choices have been made between data sources, are these justified appropriately?	*N*	Most of the choices between data sources were justified. However, data on some parameters were identified through non-systematic reviews of the literature and the selection of estimates appeared relatively ad-hoc, e.g. utilities.
	Has particular attention been paid to identifying data for the important parameters in the model?	*N*	The submission does not refer explicitly to whether special attention has been given to the identification of data to inform important parameters.
	Has the quality of the data been assessed appropriately?	*N*	The quality of data used to inform most model parameters has not been assessed explicitly.
	Where expert opinion has been used, are the methods described and justified?	*N*	
D2	Is the data modelling methodology based on justifiable statistical and epidemiological techniques?	*Y*	
D2a	Is the choice of baseline data described and justified?	*Y*	
	Are transition probabilities calculated appropriately?		

	Has a half-cycle correction been applied to both cost and outcome?	*Y*	
	If not, has this omission been justified?	*NA*	
D2b	If relative treatment effects have been derived from trial data, have they been synthesised using appropriate techniques?	*Y*	
	Have the methods and assumptions used to extrapolate short-term results to final outcomes been documented and justified?	*Y*	
	Have alternative extrapolation assumptions been explored through sensitivity analysis?	*N*	Alternative parametric distributions for time to progression were considered , although alternative assumptions for overall survival were not explored.
	Have assumptions regarding the continuing effect of treatment once treatment is complete been documented and justified?	*Y*	
	Have alternative assumptions regarding the continuing effect of treatment been explored through sensitivity analysis?	*N*	Although alternative stopping rules of paclitaxel were evaluated by the manufacturer, due to the model structure implemented, these impact only on costs.
D2c	Are the costs incorporated into the model justified?	*Y*	
	Has the source for all costs been described?	*Y*	
	Have discount rates been described and justified given the target decision-maker?	*Y*	
D2d	Are the utilities incorporated into the model appropriate?	*Y*	
	Is the source for the utility weights referenced?	*Y*	
	Are the methods of derivation for the utility weights justified?	*Y*	
D3	Have all data incorporated into the model been described and referenced in sufficient detail?	*Y*	
	Has the use of mutually inconsistent data been justified (i.e. are assumptions and choices appropriate)?	*NA*	
	Is the process of data incorporation transparent?	*Y*	
	If data have been incorporated as distributions, has the choice of distribution for each parameter been described and justified?	*Y*	
	If data have been incorporated as distributions, is it clear that second order	*Y*	

	uncertainty is reflected?		
D4	Have the four principal types of uncertainty been addressed?	*N*	Structural uncertainty has only been superficially addressed in relation to the use of alternative distributions describing time to progression.
	If not, has the omission of particular forms of uncertainty been justified?		
D4a	Have methodological uncertainties been addressed by running alternative versions of the model with different methodological assumptions?	*Y*	
D4b	Is there evidence that structural uncertainties have been addressed via sensitivity analysis?	*N*	
D4c	Has heterogeneity been dealt with by running the model separately for different subgroups?	*N*	
D4d	Are the methods of assessment of parameter uncertainty appropriate?	*Y*	
	If data are incorporated as point estimates, are the ranges used for sensitivity analysis stated clearly and justified?	*Y*	
C1	Is there evidence that the mathematical logic of the model has been tested thoroughly before use?	*Y*	
C2	Are any counterintuitive results from the model explained and justified?	*NA*	
	If the model has been calibrated against independent data, have any differences been explained and justified?	*N*	
	Have the results of the model been compared with those of previous models and any differences in results explained?	Y	Yes, after request for clarifications.