STA REPORT

LIVERPOOL REVIEWS AND IMPLEMENTATION GROUP (LRiG)

Bevacizumab in combination with capecitabine for the first-line treatment of metastatic breast cancer

This report was commissioned by the NIHR HTA Programme as project number 08/206/02

Completed 23rd February 2012

DOES NOT CONTAIN IN CONFIDENCE DATA

UNIVERSITY OF LIVERPOOL

LIVERPOOL REVIEWS AND IMPLEMENTATION GROUP

A MEMBER OF THE RUSSELL GROUP

Produced by: Liverpool Reviews & Implementation Group (LR*i*G)

Authors: Nigel Fleeman, Research Fellow (Clinical Effectiveness), Liverpool Reviews and Implementation Group, University of Liverpool

Sophie Beale, Research Associate (Decision Analysis), Liverpool Reviews and Implementation Group, University of Liverpool

Helen Davis, Assistant Director, North West Medicines Information Centre, Pharmacy Practice Unit, Liverpool

Rumona Dickson, Director, Liverpool Reviews and Implementation Group, University of Liverpool

Juliette Loncaster, Consultant Clinical Oncologist, The Christie NHS Foundation Trust, Manchester

Susan O'Reilly, Consultant in Medical Oncology, Clatterbridge Centre for Oncology, Wirral

James Oyee, Research Fellow (Medical Statistician), Liverpool Reviews and Implementation Group, University of Liverpool

Correspondence to: Nigel Fleeman, Research Fellow, Liverpool Reviews and Implementation Group, University of Liverpool, Room 2.10, Whelan Building, The Quadrangle, Brownlow Hill, Liverpool L69 3GB

Date completed: 23rd February 2012

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 08/206/02

Declared competing interests of the authors: None

Acknowledgements: Thanks to James Mahon and Adrian Bagust who provided feedback on the economic sections of the 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:

Fleeman N, Beale S, Davis H, Dickson R, Loncaster J, O'Reilly S, Oyee J. Bevacizumab in combination with capecitabine for the first-line treatment of metastatic breast cancer : A Single Technology Appraisal. LR*i*G, University of Liverpool, 2011

Contributions of authors:

Nigel Fleeman: Project lead, cross checking of manufacturer's search strategies, drafted the clinical results section and supervised the production of the final report

Sophie Beale: Critical appraisal of the economic evidence and economic model

Helen Davis: Critical appraisal of the manufacturer's submission

Juliette Loncaster: Clinical advice and critical appraisal of the clinical sections of the manufacturer's submission

Rumona Dickson: Summary and critical appraisal of the clinical evidence

Susan O'Reilly: Clinical advice and critical appraisal of the clinical sections of the manufacturer's submission

James Oyee: Critical appraisal of clinical statistical approach

All authors read and commented on draft versions of the ERG report.

Table of contents

Abbreviations

AE (s)	Adverse event (s)
AUC	Area under the curve
AVADO	Avastin Plus Docetaxel
BEV	Bevacizumab
BNF	British National Formulary
BSC	Best supportive care
CAPE	Capecitabine
CHD	Coronary heart disease
CI	Confidence interval
CSR	Clinical study report
CVD	Cardiovascular disease
DFI	Disease-free interval
ECOG	Eastern Cooperative Oncology Group
EMA	European Medicines Agency
EQ-5D	EuroQol 5D (a standardised instrument used as a measure of health outcome)
ER+ve	Oestrogen receptor-positive
ER-ve	Oestrogen receptor-negative
ERG	Evidence Review Group
EORTC	European Organisation for Research and Treatment of Cancer
FACT-B	Functional Assessment of Cancer Therapy-Breast
FDA	Food and Drug Administration
HER2	Human epidermal growth factor receptor 2
HER2+ve	Human epidermal growth factor receptor 2 positive
HER2-ve	Human epidermal growth factor receptor 2 negative
HR	Hazard ratio
HRQoL	Health-related quality of life
ICER	Incremental cost-effectiveness ratio
IPCW	Inverse Probability of Censoring Weights
IRC	Independent review committee
ITT	Intention to treat
KM	Kaplan-Meier
LYG	Life year gained
mBC	Metastatic breast cancer
MS	Manufacturer's submission
NCI-CTC	National Cancer Institute Common Terminology Criteria
NICE	National Institute for Health and Clinical Excellence
NPTs	Non-protocol specified antineoplastic therapies
ORR	Objective response rate
OS	Overall survival
PD	Progressive disease
PFS	Progression-free survival
PgR+ve	Progesterone receptor positive
PgR-ve	Progesterone receptor negative
PSA	Probabilistic sensitivity analysis
PSS	Personal Social Services
QALY	Quality adjusted life year
RCT	Randomised controlled trial
RECIST	Response Evaluation Criteria in Solid Tumors
RIBBON	Regimens in Bevacizumab for Breast Oncology
RPSFT	Rank Preserving Structural Failure Time
SA	Sensitivity analysis
SD	Stable disease
SPC	Summary of Product Characteristics
STA	Single Technology Appraisal
VEGF	Vascular endothelial growth factor
VIN	Vinorelbine
vs	Versus

1 SUMMARY

1.1 Critique of the decision problem in the manufacturer's submission

The remit of the Evidence Review Group (ERG) is to comment on the clinical and cost-effectiveness evidence submitted to the National Institute for Health and Clinical Excellence (NICE) as part of the single technology appraisal (STA) process. Clinical and economic evidence has been submitted to NICE from Roche Ltd in support of the use of bevacizumab (BEV) (Avastin) in combination with capecitabine (CAPE) (Xeloda) as a first-line treatment for patients with metastatic breast cancer (mBC).

BEV+CAPE has a marketing authorisation in Europe. It is licensed for the first-line treatment of patients with mBC in whom treatment with other chemotherapy options including taxanes or anthracyclines is not considered appropriate. Patients who have received taxane and anthracycline-containing regimens in the adjuvant setting within the last 12 months should be excluded from treatment with BEV+CAPE for mBC.

The ERG believes that the patient population for whom the manufacturer presents its clinical evidence is the same patient population that is stipulated in the decision problem in the scope issued by the National Institute for Health and Clinical Excellence (NICE) and for whom BEV+CAPE is licensed. The manufacturer has based the economic evidence on a subgroup of patients that have previously been treated with a taxane in the adjuvant setting. Based on current clinical practice, it is assumed that all patients within this subgroup have also received prior adjuvant treatment with an anthracycline. According to the manufacturer, this more stringent patient population is representative of the population of patients for whom CAPE is licensed, i.e. patients requiring treatment for mBC after failure of taxanes and an anthracycline-containing chemotherapy regimen or for whom further anthracycline therapy is not indicated.

Other than the population utilised in the economic model, the only other deviation from the decision problem as issued in the scope by NICE is the exclusion of vinorelbine (VIN) as a comparator in the clinical section; VIN is only used as a comparator for a scenario analysis in the economic model. Given CAPE is usually preferred to VIN in clinical practice and in the absence of any studies comparing BEV+CAPE to VIN and of evidence to suggest that VIN is superior to CAPE, the ERG is satisfied that CAPE is considered the main comparator.

1.2 Summary of clinical effectiveness evidence submitted by the manufacturer

The clinical effectiveness section of the manufacturer's submission (MS) is derived from a single manufacturer supported randomised controlled trial (RCT) known as RIBBON-1. RIBBON-1 was a superiority trial using parallel groups of patients in which patients were considered suitable for treatment with CAPE (or in the other cohort of the trial, a taxane/anthracycline) and then randomised in a 2:1 ratio to receive BEV+CAPE or CAPE (or BEV+taxane/anthracycline or a taxane/anthracycline in the other cohort of the trial). The primary endpoint was investigator assessed progression-free survival (PFS) according to Response Evaluation Criteria in Solid Tumors (RECIST) criteria. Secondary outcomes included independent review committee (IRC) assessed PFS and overall survival (OS). An additional PFS and OS benefit of around 3 months for patients in the BEV+CAPE arm over the CAPE arm was reported (investigator assessed median PFS: 8.6 vs 5.7 months; IRC assessed median PFS: 9.8 vs 6.2 months; median OS: 25.7 vs 22.8 months). However, despite significant improvements also in overall response rate (ORR) for the BEV+CAPE arm (35.4% compared to 23.6%), only the PFS and not the OS findings were statistically significant. The lack of a statistically significant difference in OS between the groups may be explained by differences in the nature and frequency of subsequent treatments received in both arms of the trial following disease progression.

The manufacturer also presents both *a priori* and *post-hoc* exploratory subgroup analyses for PFS and OS. All subgroup analyses suggested improvements in terms of PFS for the BEV+CAPE arm compared with the CAPE arm. A similar pattern was reported for OS. While the majority of subgroups reported statistically significant differences in PFS (all in favour of BEV+CAPE), the only subgroups that reported significant differences in OS (all in favour of BEV+CAPE) were those aged <50 years and those previously treated with a taxane, anthracycline or neoadjuvant/adjuvant chemotherapy.

The subgroup of patients previously treated with a taxane is the population of patients used by the manufacturer in the economic model. For this subgroup, the differences in PFS (4.5 months) and OS (7.9 months) between the BEV+CAPE and CAPE arms appeared to be greater (median PFS: 8.7 vs 4.2 months; median OS: 28.4 vs 20.5 months) than in the Intention To Treat (ITT) population.

A greater proportion of BEV+CAPE patients than CAPE patients in RIBBON-1 reported any adverse event (AE) (40% vs 27%), serious adverse events (SAEs) (25% vs 20%) and National Cancer Institute Common Terminology Criteria (NCI-CTC) grade 3-5 AEs (37% vs 23%). While a greater number of patients in the BEV+CAPE arm reported AEs than in the CAPE arm, no new safety concerns were identified.

1.3 Summary of the ERG's critique of clinical effectiveness evidence submitted

Overall, the range of databases selected and search strategies employed to identify RIBBON-1 appear to be appropriate. The ERG does not believe there are any relevant RCTs omitted. With regard to the clinical evidence submitted, the main issues are:

- RIBBON-1 appears to be a well conducted trial, the results of which are likely to be generaliseable to patients in the UK. Generally, baseline characteristics within RIBBON-1 appeared to be balanced across the treatment groups.
- The CAPE dose was not given to patients at the licensed dose of 1250mg/m^2 but rather at a dose of 1000mg/m^2. However, this is a dose that is commonly used in clinical practice.
- For CAPE vs BEV+CAPE, since the hazard ratios (HRs) for investigator and IRC assessed PFS were almost identical (HR=0.69 [95% CI: 0.56 to 0.84] and HR=0.68 [95% CI: 0.54 to 0.86] respectively), the evidence suggesting a benefit in terms of PFS does appear to be robust.
- Interpreting differences in OS is difficult because patients were able to 'cross-over' from the CAPE arm to receive subsequent BEV and those in the BEV+CAPE arm were also able to receive subsequent BEV. Other anti-cancer therapies were also available on progression and, in a minority of instances, prior to progression.
- The ERG urges caution in interpreting subgroup results for all outcomes (adjusted and unadjusted) because no statistical adjustments were performed to control for multiple testing in all subgroups and of all outcomes, thus increasing the likelihood of significant results emerging by chance.
- The ERG agrees that there were a greater proportion of AEs in the BEV+CAPE arm, including AEs 'of special interest' but that no new safety concerns were identified. The ERG believes the difference between the two arms can largely be attributed to differences in grade 3 AEs (27% vs 14%).

Because, as noted above in section 1.1, the economic model is based on the subgroup of patients who received a prior taxane, and as this was not the population for whom evidence was presented by the manufacturer in the clinical section of the MS (with the exception of the PFS and OS findings), the ERG attempted to extract as much data as possible on this subgroup from relevant Microsoft Excel worksheets submitted as part of the economic model. There appear to be baseline differences between this subgroup and the entire ITT population, in particular, the subgroup appears to be younger and healthier than the ITT population.

1.4 Summary of cost effectiveness submitted evidence by the manufacturer

In the absence of any relevant published economic evaluations the manufacturer developed a *de novo* economic model. The model, which has been constructed in Microsoft Excel, is made up of three health states (PFS, progressive disease (PD) and death). The modelled population comprises a subgroup of patients treated in the RIBBON-1 trial, namely those who have previously been treated

with a taxane, rather than the whole population licensed to receive BEV+CAPE. The economic evaluation adopts a time horizon of 15 years, and the perspective is that of the UK NHS. Resource use, costs and utilities have been estimated based on information from trial data and published sources.

The manufacturer's reported base case incremental cost-effectiveness ratio (ICER) is £77,318 per quality adjusted life year (QALY) gained. The manufacturer showed this ICER to be generally robust when subjected to deterministic and probabilistic sensitivity analysis (PSA) (ICERs ranging from £71,662 to £110,092 per QALY gained).

1.5 Summary of the ERG's critique of cost effectiveness evidence submitted

Overall, the ERG found the manufacturer's model to be clearly set out with adequate labelling of tables and parameters. The main areas that give cause for concern are:

- The modelled population is a subgroup of the RIBBON-1 Trial. Baseline characteristics indicate that this subgroup is younger and healthier than the licensed (ITT) population. The model results are therefore unlikely to be generalisable to the licensed population.

- The manufacturer has used the Rank Preserving Structural Failure Time (RPSFT) model *post-hoc* to ameliorate any effect that might arise from patients in both treatment arms receiving the study drug (i.e. BEV) after progression. This approach is recognised as having serious limitations when the proportions of patients receiving the study drug are high and when other therapies are permitted. Model subgroup data from the RIBBON-1 trial show that 44.7% of patients in the BEV+CAPE arm and 52.4% of patients in the CAPE arm received BEV post progression. The ERG is unable to ascertain whether the RPSFT model results in bias towards any particular treatment arm.

- The base case model generates a total incremental life-year gain of 0.863 for the BEV+CAPE arm, most of which (60.87%) accrues during PD, a period during which patients were receiving a variety of different therapies.

1.6 ERG commentary on the robustness of evidence submitted by the manufacturer

1.6.1 Strengths

The clinical evidence is derived from a well conducted RCT (RIBBON-1) that compares the intervention of interest (BEV+CAPE) to one of the comparators of interest (CAPE). This comparator is considered by the manufacturer and the ERG as the most significant of the comparators (CAPE and VIN) listed in the decision problem. The population of patients included in RIBBON-1 is the same group of patients who are specified in the decision problem and for whom BEV+CAPE has received a marketing licence from the European Union.

1.6.2 Weaknesses and areas of uncertainty

Evidence for cost effectiveness is derived from a subgroup of patients in the RIBBON-1 trial. Because no statistical adjustments were performed to control for multiple testing in all subgroups and of all outcomes, the results must be treated with caution. Furthermore, this is a more stringent patient population than that which is licensed to receive BEV+CAPE. From data extracted by the ERG from relevant spreadsheets of the economic model, it would appear this population differs to the overall population of RIBBON-1 in that it is a younger and healthier population.

The ERG has concerns about the modelling of PD, namely:

- The reliance on the RPSFT model to ameliorate any effect arising from patients in both treatment arms receiving the study drug (BEV); and
- In the model 60.87% of the incremental life years gained for BEV+CAPE accrue during PD, a period during which patients receive multiple therapies.

Three other, relatively minor, areas of uncertainty relating to the model are the estimation of drug costs, the absence of any cost of terminal care and the calculation of utility values.

1.7 Summary of exploratory and sensitivity analyses undertaken by the ERG

The manufacturer's ICER is £77,318 per QALY gained. The ERG made three relatively minor alterations/corrections to the model, namely:

- Recalculation of drug costs based on the distribution of patient body weight and body surface area of a UK specific cohort of patients, rather than a simple average based on trial data.
- Addition of the cost of terminal care.
- Correction to the calculation of utility values.

Implementing these three changes increased the ICER to £91,607 per QALY gained.

Because of concerns regarding the use of the RPSFT model to adjust for OS, the ERG made a fourth alteration. This involved amending the model so that survival for both patient groups was equivalent during the PD phase and resulted in an ICER of £170,057. Combining all of the ERG's changes results in an ICER of £207,850 per QALY gained.

It should be noted that any ICER estimate based on the modelled patient population may be optimistic as the modelled population is a subgroup of RIBBON-1 who, at baseline, appear to be younger and healthier than the licensed (ITT) population.

2 BACKGROUND

2.1 Critique of manufacturer's description of underlying health problem

In the context section of the MS (section 2) the manufacturer describes the key issues relating to the underlying health problem and associated challenges. The MS provides an overview of the clinical problem, including epidemiology and the challenge of addressing the heterogeneity of response in patients with mBC. It is noted that patients with a high risk of a poor prognosis include those who have previously received prior taxane and anthracycline treatment. A summary of this section is provided in Box 1and Box 2.

Box 1 Epidemiology (taken directly from the MS)

Breast cancer is the most common cancer in women worldwide with around 1.38 million women diagnosed in 2008; the highest rate of occurrence is in Western Europe and North America. In 2008 39,681 new cases of breast cancer were diagnosed in women in England alone.[1] There were 10,065 deaths from breast cancer in England in 2008 with breast cancer currently accounting for around 16% of female deaths in the UK.

Death from breast cancer is a consequence of metastatic disease, which is estimated to be present in 5-10% of women at the time of first presentation, metastatic disease will also affect 30-40% of patients initially diagnosed with early or localised breast cancer confined to the breast and its draining lymph nodes.[2-4]

Box 2 Heterogeneity of response in patients with mBC (taken directly from the MS)

The median time from diagnosis with metastatic disease to death is reported to be about 2 years,[5] but such median figures hide considerable heterogeneity. For example, post-menopausal women with tumours bearing large numbers of both oestrogen and progesterone hormone receptors (ER, PgR) typically have disease that follows a relatively indolent course and these patients may survive for a prolonged period of time.[6-8]

Other patients with more aggressive forms of the disease have a poorer prognosis with higher risk of early relapse and short overall survival. Human epidermal growth factor receptor 2 (HER2) negative [HER2-ve] breast cancer patients with such a poor prognosis represent a very significant unmet medical need for new therapies. These patients tend to relapse rapidly after their response to first-line therapy and have a short overall survival, due to a lack of durable response to subsequent therapies. Although a number of different patient types may be assigned to the 'high risk of poor prognosis' group, in general they include patients with triple negative disease (they lack hormonal receptors ER and PgR, as well as the HER2/neu gene), with positive lymph nodes at diagnosis, a high grade histology, or who have previously received prior taxane and anthracycline treatment.

The ERG is of the opinion that the manufacturer's description of the underlying health problem is a reasonably accurate account.

2.2 Critique of manufacturer's overview of current service provision

The manufacturer provides a summary of treatment objectives of mBC (Box 3) and choice of cytoxic treatment available (Box 4), which highlight that the "challenge facing oncologists in the UK is how to manage patients presenting with metastatic disease who have been treated in the adjuvant setting with taxanes and/or anthracyclines, as currently, clinicians have a very limited armoury of therapies with which to treat such patients and their outlook may be very poor" (p17 of the MS). The MS notes that there are currently no specific recommendations in the NICE clinical guideline for advanced breast cancer[3] for patients with mBC who have been given both anthracycline and taxane as prior adjuvant therapy. Further, it is argued (p19 of the MS): "It is well recognised by clinicians that for such patients, when they relapse after taxane adjuvant therapy, the outcomes of taxane treatment in the metastatic setting may be very poor. If patients are unable to tolerate a taxane or have relapsed from adjuvant taxane therapy the next available first-line treatment options are capecitabine or vinorelbine monotherapy." Further, it is noted that currently, the majority of patients are given CAPE as opposed to VIN and it is this group of patients for whom BEV+CAPE is intended.

Box 3 Treatment objectives of mBC

The treatment of mBC typically consists of the sequential challenge of a series of treatments with the intention of shrinking the size of the tumour. Unfortunately, the benefits of treatment in this setting tend not to be long term and response rates and the duration of response decline with each successive line of treatment.[2, 9, 10] Cancer survivors whose disease recurs have a lower quality of life in most quality of life indices than those who remain disease-free[11] and the most important distress factor among cancer survivors was found to be the fear of disease progression.[12] Therefore, the major objective of each successive line of therapy is to increase the proportion of patients who respond to first-line therapy and to prolong their disease remission.

Box 4 Choice of cytotoxic treatment

For women who are not candidates for hormonal therapy and whose tumours are HER2 -ve, cytotoxic chemotherapy is still the treatment of choice for locally advanced or metastatic disease. Existing monotherapy treatments such as anthracyclines, taxanes, vinca alkaloids, and anti-metabolites such as capecitabine, are used in the first-line metastatic setting and are capable of prolonging both disease-free and overall survival. Combination of these cytotoxic drugs provides a higher objective response rate (ORR) and a longer progression free survival (PFS) compared to monotherapy treatments; however, the gains of combination treatment generally come at the expense of increased side effects and overlapping toxicities. As a result, the use of sequential single agent cytotoxic chemotherapy remains a frequent approach.

In patients who have received an anthracycline and taxane in the adjuvant setting, treatment choices are potentially more limited, as patients may have acquired resistance to taxane therapy, resulting in poor outcomes when re-treated with a taxane in the metastatic setting. Having used two of the most effective treatment options in the adjuvant setting, these patients are often left with less efficacious monotherapy treatment options such as capecitabine and vinorelbine. Therefore, there is a large unmet clinical need to provide effective doublet treatment to first line metastatic patients who have already received two of the most effective treatments for breast cancer.

The ERG agrees that the manufacturer provides a reasonably accurate overview of current service provision. However, the ERG believes the number of patients estimated to be eligible may be open to some variation.

Based on assumptions listed in Table 1, the manufacturer estimates that around 770 patients will be eligible for treatment with BEV+CAPE. This may be an overestimate since it is assumed that all of those who are potentially eligible for BEV and who relapsed more than 12 months after initial anthracycline and taxane treatment would receive BEV+CAPE. In clinical practice, many of these patients would be considered for a re-challenge with a taxane.

Table 1 Manufacturer's estimated number of patients eligible for treatment with BEV+CAPE

Assumption	N
Annual incidence of mBC in England and Wales [a]	10,913
32% of these mBC patients receive taxanes in an adjuvant setting [b]	3,492
76% of these are HER2-ve [b]	2,654
72% are treated with chemotherapy [b]	1,911
92% of these are not enrolled in a clinical trial [b]	1,758
55% are treated with CAPE (monotherapy or in combination with another agent) [b]	967
96% are not contraindicated for BEV [c]	928
83% of these have relapsed more than 12 months after initial anthracycline and taxane treatment [d]	770

[a] Calculated from the total population for England and Wales (mid-2010 population estimates)[13] and the age-standardised incidence rate of breast cancer [1]
[b] Roche Data on File
[c] Calculated from the proportion of women in England who do not have CHD[14]
[d] Using data on relapse rates for patients with triple negative breast cancer[15] as a proxy for "poor prognostic patients who receive both anthracycline and taxane therapy in the adjuvant setting" (see Box 2)

The figure may be further inflated since this assumes the proportion of patients not contraindicated for BEV is 96% based on the prevalence of women with coronary heart disease (CHD) in England (4%).[14] Given safety concerns relating to cardiac disorders associated with BEV (see Appendix 1), there is logic to using this estimate. However, the ERG notes that this is the prevalence rate for all women of all ages, the prevalence of CHD increasing with age (see Table 37 in Appendix 2) and given the majority of women diagnosed with mBC would be aged 45 and over, this is likely to be an underestimate. Furthermore, given that in addition to cardiac disorders there are also vascular disorders of concern, it could be argued that basing the estimate on the cardiovascular disease (CVD) prevalence estimate (13% of women in England) from the same source[14] may be more appropriate. Thus the proportion of patients not contraindicated for BEV would be at least 87% (as shown in Appendix 2, the prevalence of CVD also rises with age and so this may also be an underestimate).

Finally, estimates are based on the assumption that 72% of HER2-ve patients receive chemotherapy, an estimate derived from market research conducted for the manufacturer. Clinical advisors to the ERG believe that a lower proportion may receive treatment with chemotherapy since the majority of patients would be ER+ve and therefore initially treated with endocrine therapy although it is recognised that the majority of patients with ER+ve mBC eventually become resistant to endocrine therapy.

For illustrative purposes only, assuming the proportion of patients not to be contraindicated to BEV to be 87% and the proportion of HER2-ve patients eligible for chemotherapy to be 60%, the ERG estimates just under 600 patients may be eligible for BEV+CAPE (Table 2).

Table 2 Modified estimate of number of patients eligible for treatment with BEV+CAPE

Assumption	N
Annual incidence of mBC in England and Wales [a]	10,913
32% of these mBC patients receive taxanes in an adjuvant setting [b]	3,492
76% of these are HER2-ve [b]	2,654
60% are treated with chemotherapy [c]	1,592
92% of these are not enrolled in a clinical trial [b]	1,465
55% are treated with CAPE (monotherapy or in combination with another agent) [b]	806
87% are not contraindicated for BEV [d]	701
83% of these have relapsed more than 12 months after initial anthracycline and taxane treatment [e]	582

[a] Calculated from the total population for England and Wales (mid-2010 population estimates)[13] and the age-standardised incidence rate of breast cancer [1]
[b] Roche Data on File
[c] Informed estimate based on clinical advice to the ERG
[d] Calculated from the proportion of women in England who do not have CVD[14]
[e] Using data on relapse rates for patients with triple negative breast cancer[15] as a proxy for "poor prognostic patients who receive both anthracycline and taxane therapy in the adjuvant setting" (see Box 2)

3 CRITIQUE OF MANUFACTURER'S DEFINITION OF DECISION PROBLEM

The manufacturer has based the economic evidence on a more stringent patient population than for whom the majority of the clinical evidence is derived. The differences are summarised in Table 3 and in more detail below.

Table 3 Decision problem specified by NICE and addressed in the MS

Parameter	Final scope issued by NICE	Clinical section of the MS	Economic section of the MS
Population	Adults with HER2-ve mBC previously untreated in the metastatic setting: • for whom treatment with other chemotherapy options, including taxanes or anthracyclines, is not considered appropriate and • who have not received taxane or anthracycline-containing regimens in the adjuvant setting within the last 12 months	As per scope	Subgroup of patients who have previously received a taxane in the adjuvant setting (and have most likely also received an anthracycline in the adjuvant setting) The rationale for concentrating on this subgroup is that these patients are considered to have 'failed' these previous treatments and thus the population closely represents a subgroup of patients who are indicated by the licence for CAPE
Intervention	BEV+CAPE	As per scope	As per scope
Comparator (s)	CAPE VIN	CAPE VIN is excluded on the grounds that it is rarely used	CAPE VIN is included in a scenario analysis in which identical outcomes to CAPE are assumed; drug costs are specific to VIN
Outcomes	The outcome measures to be considered include: • OS • PFS • ORR • AEs • HRQoL	As per scope although PFS and OS are used as proxy outcomes for HRQoL PFS and OS are also presented for selected subgroups of patients	PFS and OS from the clinical section are used to inform the economic model
Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per QALY The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be	-	As per scope

Parameter	Final scope issued by NICE	Clinical section of the MS	Economic section of the MS
	sufficiently long to reflect any differences in costs or outcomes between the technologies being compared Costs will be considered from an NHS and Personal Social Services perspective.		
Subgroups to be considered	Potential subgroups such as by histology and hormone receptor status will be considered if evidence allows.	As per scope	Economic model is focused on one subgroup for whom results are presented in the clinical section It is argued that these patients are considered to have 'failed' an anthracycline and taxane in the adjuvant setting and are thus representative of the population licensed to receive CAPE

3.1 Population

The scope specifies the following population:

> *Adults with HER2-negative metastatic breast cancer previously untreated in the metastatic setting:*
> - *for whom treatment with other chemotherapy options, including taxanes or anthracyclines, is not considered appropriate and*
> - *who have not received taxane or anthracycline-containing regimens in the adjuvant setting within the last 12 months*

Evidence presented in the MS is derived from one clinical trial (RIBBON-1[16]). This was a relatively large RCT in which patients were randomised to receive either BEV+CAPE or CAPE. Around 98% of patients who were enrolled were HER2-ve. The choice of chemotherapy treatment (anthracycline/taxane or CAPE) was determined by the attending clinician prior to randomisation. Thus for all patients randomised to BEV+CAPE or CAPE, it may be assumed taxanes or anthracyclines were not considered appropriate treatment options. Furthermore, an additional exclusion criterion was "Prior adjuvant or neo-adjuvant chemotherapy within 12 months prior to Day 0" (p38 of the MS). Thus the ERG considers that the population of patients randomised to BEV+CAPE or CAPE in RIBBON-1[16] are representative of the patient population identified in the scope.

For the economic model, the manufacturer restricts its population to a subgroup of patients who had received a prior taxane. It is argued that these are likely to have 'failed' on an anthracycline and a

taxane and that this subgroup is therefore broadly representative of patients who meet the licence for treatment with CAPE (see section 3.2).

The ERG agrees that the majority of patients in this subgroup would most likely have previously received an anthracycline in addition to a taxane. However the ERG questions whether they would be considered to have 'failed' on these treatments since the RIBBON-1[16] trial excluded patients who had received an adjuvant taxane or anthracycline within the last 12 months. Given most clinicians would consider a DFI>12 months as long enough to consider a re-challenge with either an anthracycline or a taxane, it is debatable whether such patients should be considered to have 'failed' on these treatments. Furthermore, the ERG notes that in clinical practice, while CAPE is only licensed for patients who have 'failed' an anthracycline or a taxane, in clinical practice it is given to patients who are not considered appropriate for an anthracycline or taxane, regardless of whether they have 'failed' on these treatment regimens in the past.

Most crucially, however, this subgroup is clearly a more stringent population than is licensed for treatment with BEV+CAPE (see section 3.2). Therefore the ERG believes the ITT population from the RIBBON-1[16] trial to be most appropriate. The ERG requested, from the manufacturer, an appropriate economic model based on the ITT/safety populations of the RIBBON-1[16] trial addressed in the clinical section of the MS rather than on the subgroup of people who have had a prior taxane. In their response, the manufacturer stated that:

> *Since the submitted health economic analysis calculated an ICER of approximately £77,000 per QALY for the "failed anthracycline and taxane therapy" subgroup, analysis of the ITT population would result in a larger ICER and therefore clearly not considered to be a cost-effective use of NHS resources.*

Finally, the ERG notes that the number of estimated patients who have received a prior anthracycline and taxane in the adjuvant setting would result in a much lower estimated number of patients eligible for BEV+CAPE than presented in section 2.2. In RIBBON-1,[16] it is noted that 40% of the ITT population were represented in the subgroup of patients who had received a prior taxane.

3.2 Intervention

The intervention is defined in the scope as BEV+CAPE and the intervention addressed in the MS is BEV+CAPE.

BEV (Avastin, Roche) is a humanised anti-vascular endothelial growth factor (VEGF) monoclonal antibody that inhibits VEGF-induced signalling and inhibits VEGF-driven angiogenesis. This reduces vascularisation of tumours, thereby inhibiting tumour growth. BEV is available as a 25mg/ml solution for infusion and is administered 10mg/kg every 2 weeks or 15mg/kg every 3 weeks, in combination with chemotherapy

CAPE (Xeloda, Roche) is an orally-administered chemotherapeutic agent. It is an anti-metabolite (prodrug) which is enzymatically converted to 5-fluorouracil (5-FU) in the tumour, thereby inhibiting tumour growth. CAPE is available as a 150 mg or 500 mg film-coated tablet. The recommended dose of CAPE is 1250 mg/m^2 administered orally twice daily (morning and evening; equivalent to 2500 mg/m^2 total daily dose) for 2 weeks followed by a 1-week rest period given as 3-week cycles.

It should be noted that while the licensed dose for CAPE is 1250 mg/m^2, in the RIBBON-1[16] trial the dose was 1000mg/m^2. The ERG notes that 1000 mg/m^2 may in fact be a more appropriate dose since it reduces toxicity without any apparent reduction in efficacy[17] and that in clinical practice a significant number of patients in the UK cannot tolerate a dose of 1250 mg/m^2. Typically many clinicians start administration of CAPE at 1000 mg/m^2 and then dose escalate to 1250 mg/m^2 if the first cycle is tolerated; it is also common for doses to be reduced after several cycles due to the occurrence of accumulative palmar-plantar erythrodysesthesia syndrome (hand-foot syndrome). Furthermore, the lower dose for CAPE may reflect the fact that a lower dose may be preferred in combination with BEV to reduce toxicities; it is noted that a previous single-arm phase II study (XCALIBr[18]) also combined CAPE with BEV at a dose of 1000mg^2 as does the ongoing TURANDOT study.[19]

According to section 4.1 of the EMA Summary of Product Characteristics (SPC) for BEV, the therapeutic indication specific to this appraisal is as follows:[20]

> *for first-line treatment of patients with metastatic breast cancer in whom treatment with other chemotherapy options including taxanes or anthracyclines is not considered appropriate. Patients who have received taxane and anthracycline-containing regimens in the adjuvant setting within the last 12 months should be excluded from treatment [with BEV+CAPE].*

The MS notes that this marketing authorisation was granted in June 2011. BEV is also indicated for the first-line treatment of mBC in combination with paclitaxel (a taxane). As a result of concerns related to toxicity (see Appendix 1), BEV is no longer licensed in the United States in combination with paclitaxel for mBC and has never been licensed in combination with CAPE. It does however retain its licence for the treatment of other cancers.

For CAPE, the ERG notes the following therapeutic indication for mBC: [21]

> *as monotherapy for the treatment of patients with locally advanced or metastatic breast cancer after failure of taxanes and an anthracycline-containing chemotherapy regimen or for whom further anthracycline therapy is not indicated.*

Although not specified in the MS, the ERG also notes that CAPE is also indicated for the first-line treatment of mBC in combination with docetaxel (a taxane) after failure of cytotoxic chemotherapy. Previous therapy should have included an anthracycline.

3.3 Comparators

The comparators listed in the NICE scope are CAPE and VIN. The marketing indication for CAPE is discussed above. VIN (Navelbine) has the following indication for mBC:[22]

> *Treatment of advanced breast cancer stage 3 and 4 relapsing after or refractory to an anthracycline containing regimen.*

In the MS, it is stated that compared to CAPE, VIN is rarely used for the treatment of patients with HER2-ve mBC previously treated with anthracyclines and taxanes. It is therefore only included as a comparator in the economic evaluation and only as a scenario analysis. To support its statement that CAPE is more commonly used than VIN, the manufacturer cites market research data from interviews with 43 clinical oncologists and 27 medical oncologists conducted in April 2010 which found that 55.4% of patients received first-line CAPE (48% as monotherapy, 7.4% with another agent) and 12.3% of patients received first-line VIN (10% as monotherapy, 2.3% with another agent).

As p23 of the MS clearly states, identical outcomes from the use of CAPE and VIN is only an assumption. However, no research evidence is presented to support this in the clinical section of the MS where it is stated (p61):

> *An indirect comparison of bevacizumab in combination with capecitabine compared with vinorelbine was not necessary in this setting given the findings of the recent clinical guideline, NICE CG81 which assumed no significant difference in survival outcomes for vinorelbine compared to capecitabine based on a single under-powered study in women who had been heavily pre-treated.*[3, 23]

The ERG notes that the cited study was a phase II trial (EORTC 10001[23]) in which all patients had been pre-treated with an anthracycline and taxane but were not necessarily receiving their treatment as first-line for mBC. EORTC 10001[23] was prematurely closed due to low accrual and planned expansion to a phase III trial was not undertaken. Nevertheless, the outcomes were similar between the two arms in terms of OS and PFS (median 9.3 and 2.8 months respectively for CAPE and 11.0 and 2.6 months for VIN). The safety profiles differed however, grade 3/4 AEs being more common in the VIN arm, particularly neutropenia (46% VIN vs 4% CAPE).

More recently, the ERG also notes the publication of a recent systematic review of phase II or phase III studies of palliative chemotherapy by Oostendorp et al 2011.[24] To be eligible, included studies were required to have at least 80% of patients with advanced breast cancer pre-treated with anthracyclines and taxanes. From ten studies of CAPE monotherapy, weighted mean values were reported to be 13.5 months for median OS and 4.2 months for median PFS. From nine studies of VIN monotherapy, weighted mean values were reported to be 12.6 months for median OS and 3.8 months for median PFS. Caution must be taken in attempting to compare the findings across treatment arms

because apart from the study of Pajk et al 2008,[23] all reports provided information from only one study group and, therefore, as the authors stated (p1058): "no differences could be assessed or ratios calculated, and standard meta-analytical techniques were not applicable."[24]

Clinical advisors to the ERG agreed that CAPE is usually preferred to VIN because it is believed to have a more favourable safety profile (e.g. in addition to neutropenia, occurrences of alopecia are much greater for patients on VIN[25]) and requires fewer out-patient visits. Thus in the absence of any studies comparing BEV+CAPE to VIN and in the absence of evidence to suggest that VIN is superior to CAPE, the ERG is satisfied that CAPE is considered the main comparator and that it is appropriate for VIN to be only used as a comparator in a scenario analysis.

3.4 Outcomes

The outcomes listed in the final scope are OS, PFS, response rates, AEs and health-related quality of life (HRQoL). These outcomes are standard in this disease area.

According to p23 of the MS, all of these outcomes are addressed by the MS but no results are presented in the MS that directly measure HRQoL. Indeed, p93 of the MS states that as EQ-5D data was not collected in the RIBBON-1,[16] no HRQoL data consistent with the NICE reference case were available. However, the manufacturer does argue on p16, p69 and p93 of the MS that: "Cancer survivors whose disease recurs have a worse quality of life in most indices than those who remain disease-free[11] and the most important distress factor among cancer survivors was found to be the fear of disease progression.[12]" Thus the implication is that both PFS and OS may be treated as indicators of HRQoL. Indeed, on p88 of the MS, it is stated that: "Both PFS and OS are clinically relevant outcomes that are highly relevant to a patient's length and quality of life." On p77 of the MS, in relation to the economic model, the manufacturer states that the PFS health state is designed to capture an mBC patient's relatively high HRQoL prior to their disease progression. The PD state is designed to capture the relatively poor HRQoL following disease progression/relapse. The ERG accepts the manufacturer's argument with regard to the relationship between PFS, OS and HRQoL. However, the ERG also believes that given the importance of HRQoL outcomes, in particular to patients, the collection of HRQoL data should be encouraged in all breast cancer trials e.g. via FACT-B and the EORTC-QOL C30 questionnaires.

Progression Free Survival and OS (but no other outcomes) are also presented for selected (*a priori* and *post-hoc*) subgroups of patients. These include the subgroup of patients previously treated with a taxane which constitutes the population in the economic model

4 CLINICAL EFFECTIVENESS

4.1 Critique of the methods of review(s)

The clinical effectiveness section of the MS is derived from a single RCT (RIBBON-1[16]) identified from a systematic review of the literature. Table 4 provides an outline of the manufacturer's approach in terms of deriving evidence for the clinical effectiveness of BEV+CAPE and its location within the MS. Its purpose is to signpost the reader to the main areas of clinical information within the MS.

Table 4 Location of key clinical effectiveness information in the MS

Key information	Page number	Tables/figures
Searches	25, 146-154	Figure 1
Eligibility criteria	25-31, 154-162	Tables 3- 4, Figures 2-3
Methods for conducting and analysing relevant RCT	32-49	
Quality assessment of relevant RCT	49-50, 163	
Efficacy results of the relevant RCTs	50-60	Tables 5-6, Figures 4-8
Safety results of the relevant RCTs	61-63	Table 7
Interpretation of clinical evidence	64-72	

4.1.1 Searches

The manufacturer described the literature searches conducted on 14th and 15th November 2011. Searches were conducted to identify relevant RCTs. There were no additional searches for identifying AEs or non-RCT evidence. No searches were conducted to identify studies for indirect and/or mixed treatment comparisons as such analyses were not deemed necessary by the manufacturer.

Major electronic databases were searched including MEDLINE, BIOSIS, EMBASE, and the Cochrane Library. All clinical abstracts for the past two years from relevant American Society of Clinical Oncology (ASCO), San Antonio Breast Cancer Symposium (SABC) and European Cancer Organisation (ECCO) / European Society for Medical Oncology (ESMO) abstracts were also reviewed. The search strategy used index and text words which included bevacizumab, capecitabine and breast cancer as descriptors and was limited to studies published in English and relating to humans and clinical trials. Where possible the search was restricted to mBC or advanced breast cancer. For MEDLINE, BIOSIS, EMBASE, the date span for the searches was from 1993 up until the date of each search. No date limits were specified for the Cochrane Library. For ASCO, the date span was from 2004 until present and for SABC and ESMO, it was from 2007 until present. There has only been one relevant ECCO/ESMO conference, this was in 2011. The ERG considers the range of databases selected and search strategies employed to be appropriate.

From its own searches of Ovid MEDLINE, Scopus (which includes EMBASE), ASCO and SABC on 6th January 2012, the ERG is confident all potentially relevant studies were identified by the manufacturer.

4.1.2 Eligibility criteria

The inclusion/exclusion criteria employed by the manufacturer are clearly described in the MS (reproduced below in Table 5) and appear to be appropriate. Abstracts were obtained for each of the RCT records identified and assessed for relevance. Where it was not possible to determine relevance from the abstract, the full paper or record was obtained and evaluated in more detail. For each excluded RCT, a rationale was recorded. It is not explicit whether the application of inclusion/exclusion criteria was cross checked by a second reviewer.

Table 5 Eligibility criteria for studies to be included in the manufacturer's systematic review

Inclusion criteria	Exclusion criteria
Published papers or abstracts which evaluated the following were included: • BEV had to be the major focus of the study, in order to eliminate references which merely mentioned BEV as part of a discussion of treatments for mBC or other cancers • MBC had to be a major focus of the study, in order to eliminate papers addressing the use of BEV in other types of breast cancers, e.g., inflammatory breast cancer, or in other settings, e.g., neoadjuvant/adjuvant breast cancer, early breast cancer • Studies in which patients received BEV+CAPE, to be consistent with the BEV licence. Data addressing the efficacy of BEV in combination with other agents are not in line with this submission. • Studies in which patients received study therapy for the first-line treatment of mBC, to be consistent with the BEV licence. Data addressing the efficacy of BEV+CAPE in second or later lines of treatment are not in line with the licence. • Patient population had to consist predominantly of HER2-ve patients (≥90%), as this is the patient population of interest for this appraisal • Efficacy endpoints associated with the treatment of mBC were the focus for the data, i.e., PFS, OS, response rates • Clinical trial data – rather than case reports, retrospective reviews, etc. • Controlled studies • Documents relating to humans – since work in animal models is not relevant to this application	Published papers or abstracts which evaluated the following were excluded: • References which were not randomised, controlled phase II/III trials (such as phase I or safety studies or reviews) • Studies where CAPE was not included, or where the difference between treatment arms was the addition of an agent other than BEV (e.g., BEV+CAPE vs BEV+CAPE+agent A) • Studies which were in non-relevant populations, i.e. non first-line setting in metastatic disease, neoadjuvant/adjuvant therapy, early breast cancer, locally advanced breast cancer only or inflammatory breast cancer, HER2+ve disease • Studies where the dose or regimen of BEV or CAPE used was not UK standard practice • References from ongoing studies providing insufficient data e.g. patients demographics/study designed described, but no efficacy data available

Applying the eligibility criteria, the manufacturer identified two potentially relevant clinical trials: RIBBON-1[16] and the ongoing TURANDOT.[19] As TURANDOT[19] is a trial comparing BEV+CAPE

with BEV in combination with a taxane (paclitaxel), the ERG does not believe this second trial meets the inclusion criteria. However, as the trial is ongoing, no data from this trial were presented in the MS.

In addition, a single-arm phase II safety and efficacy study, XCALIBr[18] was also identified by the ERG. While correctly excluded because this was not an RCT, this study could have been used to provide additional data on AEs.

4.1.3 Data extraction

No details of data extraction are provided by the manufacturer.

4.1.4 Quality assessment

The evidence for clinical effectiveness is derived from only one manufacturer supported RCT (RIBBON-1[16]). Its quality was appropriately assessed using the minimum criteria for assessment of risk of bias in RCTs recommended in the Centre for Reviews and Dissemination guidance.[26] This is reproduced below in Table 6.

The ERG generally agrees with the manufacturer in relation to how the study questions were assessed. However, in relation to the third study question, the ERG notes some imbalances (differences of 5% or more) between the two treatment arms (see section 4.2).

Table 6 Quality Assessment of RIBBON-1

Study question	How is the question addressed in the study? (manufacturer response)	ERG comment
1. Was randomisation carried out appropriately?	After written consent was obtained and eligibility established, the study site obtained the patient's identification number and randomisation to treatment arm from the interactive voice response system	Agree
2. Was the concealment of treatment allocation adequate?	The side effect profile of BEV may have given the investigators some insight into which treatment the patients had been allocated. A placebo control was used to minimise bias	Agree
3. Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	The patient demographics and characteristics were generally well balanced in both arms of the CAPE cohort. However there were slightly less triple negative patients and slightly more hormone receptor positive patients in the BEV+CAPE arm.	A few other imbalances were noted. However, none of these were expected to bias the results
4. Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	The study was designed as "double-blind". A placebo control was used to minimise bias in the assessment of disease response and adverse event reporting	Agree
5. Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	There were no unexpected imbalances in drop outs	Agree
6. Is there any evidence to suggest that the authors measured more outcomes than they reported?	There is no evidence to suggest this	Agree
7. Did the analysis include an ITT analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Efficacy analyses were conducted on the ITT population. Safety analyses were conducted on patients who received at least one dose of study medication	Agree

4.1.5 Evidence synthesis

As noted above, the manufacturer identified only one appropriate RCT (RIBBON-1[16]) to measure efficacy and safety. This compared BEV+CAPE to CAPE. No RCTs were identified that compared BEV+CAPE to VIN. Thus no meta-analysis could have been undertaken to compare BEV+CAPE to CAPE or BEV+CAPE to VIN.

In the absence of direct comparisons of BEV+CAPE to VIN, the manufacturer may have attempted to conduct a mixed-treatment comparison. However, as noted in section 0 above, the only known RCT to compare VIN to CAPE was a small RCT that was halted prematurely. In this study,[23] patients were not necessarily receiving their chemotherapy as first-line treatment for mBC and the HER2 status of patients is unknown. Thus a more complex network analysis would have been required if a mixed

treatment comparison was to be attempted. The ERG doubts whether suitable trials would have been available (e.g. HER2 status was unlikely to have been considered important and therefore known in previous trials of VIN and few if any trials would have been limited to the first-line setting for mBC). There also appear to be valid practical reasons why VIN is a less suitable comparator than CAPE, i.e. it is much less likely to be preferred by clinicians in clinical practice. Therefore the ERG does not believe a mixed treatment comparison would have added any value to the evidence base.

4.2 Critique of trials of interest, their analysis and interpretation

The RIBBON-1[16] trial was a relatively large phase III randomised, double-blind, placebo controlled, multi-centre, international study designed to evaluate the efficacy and safety of first-line chemotherapy in combination with BEV vs chemotherapy with placebo in HER2-ve mBC patients. Eligible patients were randomised in a 2:1 ratio to receive BEV or placebo in combination with either an anthracycline- or taxane- based chemotherapy or CAPE. The study characteristics of the included RIBBON-1[16] trial are summarised in Table 7.

It is important to note that the choice of the chemotherapy was at the discretion of the investigator and was specified prior to randomisation for use as a stratification variable. The chemotherapies included several standard cytotoxic chemotherapies such as two taxanes (docetaxel and paclitaxel), four anthracycline-based therapies (doxorubicin, epirubicin, cyclophosphamide, and 5-FU) or CAPE. RIBBON-1[16] was designed to power two separate cohorts for the efficacy analyses of taxane/anthracycline (cohort 1) and CAPE (cohort 2). The primary objective of the study was to determine the clinical benefit of the addition of BEV to standard chemotherapy regimens for previously untreated mBC, as measured by PFS based on investigator tumour assessment. According to the clinical study report (CSR), this was assessed in a parallel manner as follows:

1. Determination of the clinical benefit, as measured by PFS based on investigator tumour assessment, of the addition of BEV to taxane therapy (docetaxel or paclitaxel protein-bound particles administered every 3 weeks) and anthracycline-based therapy, compared with these chemotherapies alone, in subjects who are receiving first-line therapy for locally recurrent or mBC.

2. Determination of the clinical benefit, as measured by PFS based on investigator tumour assessment, of BEV+CAPE vs CAPE, in subjects who are receiving first-line therapy for locally recurrent or mBC.

The MS presents efficacy and safety data from this second cohort, i.e. the CAPE cohort. Given anthracyclines or taxanes were a possible treatment options prior to randomisation, it seems reasonable to assume these patients would be considered unsuitable for an anthracycline or a taxane.

Table 7 Study characteristics of RIBBON-1

Trial design and patients	Intervention	Comparator	Inclusion criteria	Exclusion criteria
• Phase III placebo-controlled RCT • 309 patients were enrolled at 113 sites in the United States, and 306 patients were enrolled at 65 sites outside the United States; 50 (8.1%) patients were recruited from 4 sites in the UK[a]	BEV+CAPE (n=409) • CAPE 1000mg/m^2 twice daily on day 1-14 followed by 7 day break • BEV 15mg/kg every 3 weeks • Both drugs continued until disease progression or a maximum of 48 months • Dose adjustment for BEV not permitted (in cases of serious BEV-related toxicity, BEV would be either temporarily or permanently discontinued) • Dose adjustment for CAPE permitted	CAPE (n=206) • CAPE 1000mg/m^2 twice daily on day 1-14 followed by 7 day break • Placebo every 3 weeks • CAPE continued until disease progression or a maximum of 48 months • Dose adjustment for CAPE permitted	• Histologically or cytologically confirmed adenocarcinoma of the breast, with measurable or non-measurable locally recurrent or metastatic disease. Locally recurrent disease must not have been amenable to resection with curative intent • Patients with HER2-ve mBC (patients with HER2+ mBC were eligible only if they received prior treatment with trastuzumab, unless trastuzumab therapy was contraindicated or unavailable) • Signed Informed Consent Form • Age ≥ 18 years • For women of childbearing potential, use of accepted and effective method of non-hormonal contraception • ECOG performance status of 0 or 1	• Pregnancy or lactation • Inadequate organ function – liver, haematology, coagulation, renal • Prior chemotherapy for locally recurrent or metastatic disease • Prior adjuvant or neo-adjuvant chemotherapy within 12 months prior to Day 0 • Investigational therapy within 28 days of Day 0 • Major surgery within 28 days prior to day 0 or minor surgery within 7 days of day 0. • Prior therapy with BEV, sorafenib, sunitinib, or other VEGF pathway-targeted therapy • Known brain or other central nervous system metastases • Blood pressure ≥150/100 mmHg • Unstable angina; congestive heart failure; history of myocardial infarction, stroke or transient ischemic attack (within 6 months); clinically significant peripheral vascular disease • Evidence of bleeding diathesis or coagulopathy • History of abdominal fistula, gastrointestinal perforation, or intra-abdominal abscess (within 6 months) • History of anaphylactic reaction to monoclonal antibody therapy not controlled with treatment premedication • Serious non-healing wound, ulcer, or bone fracture

All data taken from MS except [a] taken from Table 14.1/7 of the CSR

In the CAPE cohort, a total of 615 patients were enrolled at 178 centres. In order for the results of a trial with so many centres to be meaningfully interpreted, the manner in which the protocol is implemented should be clear and similar across all centres. With so many investigators in different countries, general clinical practice will always be an issue and the results of a trial can only be generalisable if it is executed efficiently. The proportion of patients with protocol deviations was small and approximately similar across the two arms. In total, there were 64 (15.6%) patients in the BEV+ CAPE arm and 17 patients (8.3%) in the CAPE arm that had at least one protocol deviation. These were only minor deviations which suggests the trial was executed efficiently.

According to the CSR, recruitment to RIBBON-1[16] was initiated on 15 December 2005 up until 4 May 2007. The database was locked on 31 July 2008. Although not reported in the MS, during this period, there were a number of amendments to the original protocol following initiation of patient recruitment. These amendments included sample size revision, inclusion of secondary outcomes (1-year survival rate and an IRC assessment of PFS), extension of maximum duration of BEV treatment from 24 months to 48 months, clarity on study therapy and safety reporting. Several exploratory analyses were included as protocol amendments. One such exploratory analysis was the inclusion of the subgroup of patients previously treated with a taxane that was used as the patient population in the economic model.

The RIBBON-1[16] trial inclusion/exclusion criteria were, as would be expected for a clinical trial, relatively stringent to protect patients at greatest risk of AEs and therefore only those with ECOG performance status 0 or 1 were included. In addition, it is noted that patients who might otherwise have been eligible for CAPE were excluded from RIBBON-1.[16] This was because it was unknown at the time of recruitment whether patients would receive BEV or not and therefore a number of exclusion criteria were applied in relation to previous BEV safety concerns, in particular cardiovascular co-morbidity, bleeding diathesis or coagulopathy, history of abdominal fistula, gastrointestinal perforation, or intra-abdominal abscess (within 6 months), history of anaphylactic reaction to monoclonal antibody therapy not controlled with treatment premedication and serious non-healing wound, ulcer, or bone fracture.

Once recruited, patients received treatment every 3 weeks until disease progression, treatment-limiting toxicity, or death. The BEV dose in RIBBON-1[16] was 15mg/kg every 3 weeks as indicated in the licence. As noted in section 3.2, the CAPE dose was not the licensed dose of 1250mg/m^2 in RIBBON-1[16] but rather 1000mg/m^2 which is commonly used in clinical practice and has been used in other BEV+CAPE studies.[18, 19]

Following progression, patients were permitted to move to an open-label post-progression phase consisting of treatment as per investigator discretion (which included chemotherapy and BEV).

Patients who did not opt into this post-progression phase were followed up in a survival follow-up phase, a phase which patients who discontinued from the post-progression phase also entered. Patients were followed up throughout all phases of the trial for survival and subsequent anti-cancer therapies every 4 months until death, withdrawal of consent, loss to follow-up, or study termination, regardless of participation in the optional open-label post-progression phase. However, patients who discontinued from treatment during the blinded treatment phase for reasons other than disease progression were followed up with tumour assessment every 9 weeks, until documented disease progression or death. A maximum of 48 months of treatment with BEV (blinded treatment phase plus optional open-label post-progression phase) was allowed. According to the published paper[16] and CSR, the median follow-up time was 15.6 months (minimum 0 months, maximum 30.6 months).

From the CSR for RIBBON-1,[16] the ERG notes that a minority of patients actually received non-protocol specified antineoplastic therapies (NPTs) prior to disease progression. A greater proportion of these were in the BEV+CAPE arm (9.3% vs 6.3% in the CAPE arm). The two most commonly used NPTs were hormonal therapy (3.7% vs 2.9% in BEV+CAPE and CAPE arms, respectively) and chemotherapy (3.7% vs 1.9% in BEV+CAPE and CAPE arms, respectively).

Randomisation was stratified according to DFI (< 12 months, >12 months since completion of adjuvant chemotherapy or surgery if no adjuvant chemotherapy), prior adjuvant chemotherapy (yes, no), number of metastatic sites (< 3, ≥ 3) and choice of chemotherapy (taxane, anthracycline-based, CAPE). The ERG is of the opinion that these stratification factors are appropriate.

Around half the patients in RIBBON-1[16] were recruited from the United States, with around 8% of the patients recruited from the UK. While clinical practice does differ slightly in the United States and other countries compared to the UK (e.g. there may be differences in treatment options post-progression), the ERG believes the results of the trial are likely to be generaliseable to patients in the UK.

Table 8 Baseline characteristics of RIBBON-1

Demographic or disease-related variable		CAPE (n = 206) [a]	BEV+CAPE (n = 409) [a]	Total (n = 615) [a]
Age (years)	Median (range)	57 (23−88)	56 (28−91)	56 (23−91)
Sex	Female	204 (99.0%)	408 (99.8%)	612 (99.5%)
Region	United States	95 (46.1%)	214 (52.3%)	309 (50.2%)
	Other	111 (53.9%)	195 (47.7%)	306 (49.8%)
ECOG performance status	0	110 (53.4%)	214 (52.7%)	324 (52.9%)
	1	96 (46.6%)	192 (47.3%)	288 (47.1%)
Sites of involvement [a]	Bone [b]	130 (63.4%)	281 (68.7%)	411 (66.9%)
	Lung [b]	107 (52.2%)	177 (43.3%)	284 (46.3%)
	Local-regional [b]	87 (42.4%)	163 (39.9%)	250 (40.7%)
	Liver [b]	76 (37.1%)	168 (41.1%)	244 (39.7%)
	Distant nodes [b]	83 (40.5%)	156 (38.1%)	239 (38.9%)
	Effusion/ascites [b]	39 (19.0%)	78 (19.1%)	117 (19.1%)
	Ipsilateral supraclavicular nodes [b]	26 (12.7%)	38 (9.3%)	64 (10.4%)
	Distant skin/subcutaneous [b]	13 (6.3%)	26 (6.4%)	39 (6.4%)
	Opposite breast [b]	6 (2.9%)	18 (4.4%)	24 (3.9%)
	Adrenal [b]	7 (3.4%)	13 (3.2%)	20 (3.3%)
	Bone marrow [b]	3 (1.5%)	5 (1.2%)	8 (1.3%)
	Other [b]	21 (10.2%)	36 (8.8%)	57 (9.3%)
No. metastatic sites	< 3	113 (54.9%)	232 (56.7%)	345 (56.1%)
	≥ 3	93 (45.1%)	177 (43.3%)	270 (43.9%)
Bone lesion only	Yes	21 (10.2%)	36 (8.8%)	57 (9.3%)
	No	185 (89.8%)	373 (91.2%)	558 (90.7%)
Hormone receptor) and HER2 status	ER+ve and/or PgR+ve [a]	146 (73.7%)	312 (77.4%)	458 (76.2%)
	HER2-ve [a]	196 (97.0%)	392 (98.0%)	588 (97.7%)
	Triple negative (ER-ve/PgR-ve/HER2-ve) [a]	50 (25.3%)	87 (21.7%)	137 (22.9%)
DFI	≤ 24 months [c]	-	-	205 (33.3%)
	≤12 months	45 (21.8%)	109 (26.7%)	154 (25.0%)
	>12 months	161 (78.2%)	300 (73.3%)	461 (75.0%)
	>24 months [c]	-	-	410 (66.7%)
Prior treatment for primary breast cancer	Any	190 (92.2%)	374 (91.4%)	564 (91.7%)
	Surgery	188 (91.3%)	365 (89.2%)	553 (89.9%)
	Chemotherapy	156 (75.7%)	288 (70.4%)	444 (72.2%)
	Taxane	84 (40.8%)	161 (39.4%)	245 (39.8%)
	Anthracycline-based agent	143 (69.4%)	247 (60.4%)	390 (63.4%)
	Radiotherapy	140 (68.0%)	254 (62.1%)	394 (64.1%)
	Hormonal therapy	109 (52.9%)	203 (49.6%)	312 (50.7%)
Prior treatment for mBC	Any	98 (47.6%)	207 (50.6%)	305 (49.6%)
	Hormonal therapy	89 (43.2%)	188 (46.0%)	277 (45.0%)
	Radiotherapy	49 (23.8%)	113 (27.6%)	162 (26.3%)

[a] For data on ECOG performance status, n=615 (n=206 CAPE, n=406 BEV+CAPE), sites of involvement, n=614 (n=205 CAPE, n=419 BEV+CAPE), hormone receptor status, n=601 (n=198 CAPE, n=403 BEV+CAPE) HER2 status, n=602 (n=202 CAPE, n=400 BEV+CAPE) and triple negative, n=599 (n=198 CAPE, n=401 BEV+CAPE)
All data taken from Table 6 of the MS except [b] Table 11 of the CSR and [c] conference abstract[27]

Generally, baseline characteristics within RIBBON-1[16] appeared to be balanced across the treatment groups (Table 8). A few imbalances (differences of ≥5% between treatment arms) were identified by the ERG. A greater proportion of patients were from the United States in the BEV+CAPE arm than the CAPE arm and there were also differences in some sites of involvement. A greater proportion of BEV+CAPE patients had bone metastases whereas a greater proportion of CAPE patients had lung metastases. None of these differences were considered to be clinically relevant by clinical advisors to the ERG. However, the manufacturer also highlighted there were proportionately fewer triple negative and proportionately more hormone receptor positive patients in the BEV+CAPE arm (although these differences were not≥5%) while the ERG notes slightly greater proportion of patients who have received adjuvant chemotherapy and/or radiotherapy in the CAPE arm. Taken together, these factors could suggest a better prognosis for the BEV+CAPE patients. On the other hand, comparatively more patients had DFI ≤12 months in the BEV+CAPE arm which could suggest that these patients may have a slightly worse prognosis.

RIBBON-1[16] was a superiority trial in which the primary endpoint was investigator assessed PFS according to RECIST criteria.[28] A number of secondary outcomes common to breast cancer trials were also utilised, including OS. In addition, investigator assessed PFS, not censored for non-protocol specified NPTs was also conducted as a sensitivity analysis. All outcome measures and their definitions from the MS are provided in Table 38 in Appendix 3. These endpoints are commonly employed for the disease area specified in the decision problem issued by NICE and which the ERG considers to be appropriate.

The primary outcome of investigator assessed PFS demonstrated a statistically significant benefit for patients in the BEV+CAPE arm of 2.9 months. The IRC assessed PFS showed a slightly greater benefit of 3.5 months and a sensitivity analysis of investigator assessed PFS that was not censored for NPTs resulted in a PFS benefit that was between the two estimates (Table 9). As can be seen from Table 9, far fewer patients were classified as having an event by the IRC than by the investigator. The ERG notes that since the HRs were almost identical for investigator and IRC assessed PFS, the evidence suggesting a benefit in terms of PFS appears to be robust.

Table 9 Primary, secondary and sensitivity analyses for PFS in RIBBON-1

Outcome	CAPE (n=206)	BEV+CAPE (n=409)
Primary outcome – investigator assessed, ITT population		
PFS events	162 (78.6%)	291 (71.1%)
PFS, median (range) months	5.7 (4.3 to 6.2)	8.6 (8.1 to 9.5)
	Stratified HR=0.69 (95% CI: 0.56 to 0.84); p=0.0002 Unstratified HR=0.67 (95% CI: 0.55 to 0.82); p<0.0001	
Secondary outcomes - **IRC assessed**		
PFS events	119 (57.8%)	219 (53.5%)
PFS, median (range) months	6.2 (4.7 to 7.8)	9.8 (8.4 to10.4)
	Stratified HR=0.68 (95% CI: 0.54 to 0.86); p=0.0011	
Key sensitivity analysis - Investigator assessed, not censored for NPTs		
PFS events	168 (81.6%)	309 (75.6%)
PFS, median months	5.5	8.8
	Stratified HR=0.66 (95% CI: 0.55 to 0.81); p<0.0001	

Significant improvements were also reported for patients in the BEV+CAPE arm compared to those in the CAPE arm for ORR, with just over a third in the BEV+CAPE arm responding compared to around a quarter in the CAPE arm. In both arms, the majority of patients achieved a partial, rather than complete, response. The duration of overall response was two months greater in the BEV+CAPE arm (Table 10).

Table 10 Analyses for objective response in RIBBON-1

Outcome	CAPE (n=206)	BEV+CAPE (n=409)
Number of patients with measurable disease	161	325
ORR	38 (23.6%)	115 (35.4%)
	Between arm difference: 11.8% (95% CI: 3.4% to 20.2%) p=0.0097	
Complete response	1 (0.6%)	7 (2.2%)
Partial response	37 (23.0%)	108 (33.2%)
Duration of overall response, median (95% CI) months	7.2 (5.1 to 9.3)	9.2 (8.5 to 10.4)

Significant improvements in PFS and ORR did not translate into significant improvements in OS. However, there appeared to be a trend towards improved OS, the survival benefit being around 3 months and 81% surviving in the first year in the BEV+CAPE arm compared to around 75% in the CAPE arm (Table 11).

Table 11 Analyses for OS in RIBBON-1[a]

Outcome	CAPE (n=206)	BEV+CAPE (n=409)
Number of patients who died	99 (48.1%)	186 (45.4%)
OS, median (range) months [a]	22.8 (20.5 to 28.4)	25.7 (22.0 to 28.4)
	Unstratified HR=0.88 (95% CI: 0.69 to 1.13); p=0.33	
One–year survival rate[a]	74.8%	81.0%
	Between arm difference: 6.2% (95% CI: -1.0% to 13.4%) p=0.092	

[a] Data cut off points were 31st July 2008 for main analysis and 23rd February 2009 for updated analysis (OS and one-year survival rate only). The estimated stratified HR for OS, based on the earlier cut-off, was 0.85 (95% CI: 0.63 to 1.14; p = 0.27) according to the published paper[16]

To assess the consistency of treatment benefit with respect to PFS and OS across a number of subgroups, forest plots (including estimated HRs using unstratified Cox proportional hazards regression model) were presented in the MS. From the CSR, the ERG discovered that there were a number of pre-specified subgroups and *post-hoc* exploratory subgroups, as well as subgroups specified once the study was begun but before the analysis was completed (see Table 12). All subgroup analyses suggested improvements in terms of PFS for the BEV+CAPE arm compared to the CAPE arm. A similar pattern was reported for OS although exceptions presented in the forest plot (Figure 8 of the MS) were in the subgroups of patients who were aged ≥65 years, had DFI ≤12 months, time from diagnosis of local recurrent disease / mBC to diagnosis of primary cancer ≤ 12 months. While the majority of differences in PFS were statistically significant by subgroup, the only statistically significant subgroups reported for OS (all in favour of BEV+CAPE) were those aged <50 years and in subgroups of patients previously treated with a taxane, anthracycline or neoadjuvant/adjuvant chemotherapy. The ERG urges caution in interpreting all subgroup results because no statistical adjustments were performed to control for multiple testing in all subgroups and of all outcomes presented in Table 12, thus increasing the likelihood of significant results emerging by chance.

Table 12 Subgroups for which PFS and/or OS data were presented in the MS

Pre-specified subgroups	Subgroups specified once the study was begun but before the analysis was completed	*Post-hoc* exploratory subgroups
• age < 50 vs ≥ 50 years), • race (white vs non-white), • region (United States vs ex−United States) • baseline ECOG performance status (0 vs 1) • prior adjuvant chemotherapy (yes vs no), • prior adjuvant hormone therapy (yes vs no), • prior hormonal therapy for locally recurrent or metastatic disease (yes vs no),	• age (< 40, 40–64, ≥ 65 years) • region (North America, Latin America, Eastern Europe, Western Europe, Asia) • menopausal status (premenopausal vs perimenopasual vs postmenopausal) • number of metastatic sites (< 3 vs ≥ 3) • sites of involvement (non-visceral only vs any visceral; liver involvement vs others; bone only vs others) • disease measurability (yes vs no), • sum of longest diameters of target lesions (≤ median vs > median) • HR status (ER+ve and/or PgR+ve vs ER-ve and/or PgR-ve) • triple-negative status (ER-ve, PgR -ve and HER2-ve vs ER+ve or PgR +ve or HER2+ve) • DFI (≤ 12 months vs > 12 months) • prior neoadjuvant/adjuvant chemotherapy (yes vs no), • prior taxane therapy (yes vs no), • prior anthracycline therapy (yes vs no)	• Prior adjuvant chemotherapy or hormonal therapy (yes vs no) • Time from diagnosis of local recurrent disease / mBC to diagnosis of primary cancer (≤24 months vs >24 months) • OS calculated using the RPSFT model for the prior taxane therapy (yes vs no) subgroup

For the ITT analyses of OS, a formal testing of OS was first performed at the time of the final PFS analyses (31st July 2008) and later updated on 23rd February 2009. A stratified log-rank test was used to compare OS between treatment arms. The Kaplan−Meier (KM) method was used to estimate median OS for each treatment arm. The HRs were estimated using the stratified and unstratified Cox proportional hazards regression model. However, the trial was designed so that if patients in any treatment arm had disease progression, they could cross-over/continue to receive BEV or other anti-cancer therapies.

As current treatment pathways for mBC incorporate many active agents delivered in a sequential fashion, the manufacturer argues it is therefore very difficult to demonstrate a statistically significant OS advantage for a therapy used in the first-line setting. The ERG agrees that differences in subsequent treatments do create problems in understanding the true impact on OS of an intervention, particularly where subsequent treatments differ by treatment arm. This is a particular problem where patients are able to cross-over from one treatment arm to another as bias may be introduced.

In RIBBON-1[16] patients in the CAPE arm and those in the BEV+CAPE arm were also able to receive subsequent BEV and other anti-cancer therapies on progression and, in a minority of instances, prior to progression. Thus PFS may be a suitable proxy measure of benefit of the first-line treatment. However, ideally data on subsequent lines of anti-cancer therapies should be presented and attempts should also be made to adjust OS estimates to allow for subsequent lines of treatment.

Detailed data on treatment received following progression were not presented in the MS; simply data on those who crossed over to receive additional BEV were presented. Data discussed in the peer reviewed publication of RIBBON-1[16] were also provided by the manufacturer following the ERG's clarification request. While no adjustment for cross-over and other post-progression treatment with anti-cancer drugs were pre-specified in the protocol of the RIBBON-1[16] trial, a *post-hoc* adjusted analysis based on the RPSFT model[29] was presented in the economics section of the MS.

As expected for patients relapsing after first-line treatment of mBC, the majority of patients received additional lines of systemic treatment, either with hormonal agents or with chemotherapy and for some patients this second-line chemotherapy was combined with BEV (Table 13). A greater proportion of patients in the CAPE arm received subsequent BEV and/or chemotherapy than in the BEV+CAPE arm. Confusingly, the proportion of patients who received subsequent BEV cited in the MS exceeded the proportion cited in the published paper[16] although in both instances the difference between treatment arms was around 15% with a greater proportion of CAPE patientsthan BEV+CAPE patients receiving subsequent BEV.

The OS estimate using the RPSFT model was only conducted for the subgroup of patients who had received a prior taxane. The OS findings using the RPSFT model are reported in section 4.3 alongside a critique of this method.

Table 13 Subsequent anti-cancer therapy (RIBBON-1 ITT population)

Subsequent therapy received	CAPE (n=206)	BEV+CAPE (n=409)
Patients who 'crossed over' to receive additional BEV [a]	120 (59.7%)	184 (45.5%)
Patients who received subsequent therapy [b]	142 (68.9%)	251 (61.4%)
Type of therapy:		
• BEV [b]	112 (54.4%)	160 (39.1%)
• Chemotherapy [b]	135 (65.5%)	226 (55.3%)
• Hormonal therapy [b]	28 (13.6%)	51 (12.5%)
• Radiotherapy [b]	12 (5.8%)	35 (8.6%)
• Surgery [b]	4 (1.9%)	3 (0.7%)
• Other [b]	8 (3.9%)	12 (2.9%)

[a] Data taken from Table 14 of the MS
[b] Data taken from Table 3 of the published paper[16]

The primary safety analyses were based on all patients who received any study treatment, defined as at least one full or partial dose of either study treatment. This population was referred to as the safety population and differed to the ITT population in that patients were analysed based on their initial treatment. In this population, a greater proportion of patients reported any AEs, serious adverse events (SAEs) and NCI-CTC grade 3-5 AEs (Table 14). The CSR (Table 14.3/23) suggests that the difference between the two arms may be largely attributed to differences in grade 3 AEs, (27% vs 14%) with the proportion of patients experiencing grade 4 AEs (6% vs 5%) and grade 5 AEs (3% vs 4%) being similar. The proportion of AEs leading to the discontinuation of BEV or placebo was also similar in both arms (Table 14).

Table 14 Patients experiencing at least one AE in RIBBON-1 (blinded treatment phase, safety population) [a]

Type of AE	CAPE (n = 201)	BEV+CAPE (n = 404)
Any AE [a]	54 (26.9%)	162 (40.1%)
Grade 3–5 AE	46 (22.9%)	148 (36.6%)
SAE	41 (20.4%)	102 (25.2%)
AE leading to discontinuation of BEV or placebo	24 (11.9%)	51 (12.6%)
All deaths (including disease progression)	97 (48.3%)	185 (45.8%)
Deaths unrelated to disease progression [b]	5 (2.5%)	6 (1.5%)

[a] AEs collected as per study protocol based on the later clinical cut off date of 23rd February 2009
[b] Deaths occurring within 30 days of the last dose of study drug due to a reason other than disease progression

Patients in the BEV+CAPE arm were more likely to report AEs 'of special interest' than in the CAPE arm (Table 15). Although not explicitly defined in the MS, AEs 'of special interest' were identified by the manufacturer via clinical review and appear to relate to AEs highlighted in the SPC;[20] in the CSR they are referred to as 'selected' AEs (see also Table 40 of the CSR). Overall, 92 (22.8%) patients experienced one of these AEs in the BEV+CAPE arm compared to 18 (9.0%) in the CAPE arm. In addition to these AEs 'of special interest' identified by the manufacturer, the ERG has included AEs identified from previous studies of CAPE in Table 15. These were not presented in the MS but were taken from Table 14.3/23 of the CSR. It was noticeable that there were a greater proportion of patients experiencing hypertension, proteinura and sensory neuropathy; however, only hypertension was reported in ≥5% patients. All other AEs 'of special interest', including AEs identified by the ERG, were relatively rare.

Table 15 Patients experiencing at least one AE 'of special interest' [a, b] in RIBBON-1 (blinded treatment phase, safety population) [a]

AE 'of special interest' [a, b]	CAPE (n = 201)	BEV+CAPE (n = 404)
Arterial thromboembolic event [a]	3 (1.5%)	8 (2.0%)
Bleeding [a]	1 (0.5%)	1 (0.2%)
Febrile neutropenia [a]	0 (0.0%)	0 (0.0%)
Fistula [a]	1 (0.5%)	1 (0.2%)
Gastrointestinal perforations [a]	0 (0.0%)	0 (0.0%)
Hypertension [a]	2 (1.0%)	43 (10.6%)
Left ventricular systolic dysfunction [a]	1 (0.5%)	6 (1.5%)
Neutropenia [a]	2 (1.0%)	5 (1.2%)
Proteinuria [a]	0 (0.0%)	9 (2.2%)
Reversible posterior leukoencephalopathy syndrome [a]	0 (0.0%)	0 (0.0%)
Sensory neuropathy [a]	1 (0.5%)	12 (3.0%)
Venous thromboembolic event [a]	7 (3.5%)	20 (5.0%)
Wound dehiscence [a]	0 (0.0%)	3 (0.7%)
Diarrhoea [b]	4 (2.0%)	6 (1.5%)
Fatigue [b]	0 (0.0%)	2 (0.5%)
Nausea [b]	0 (0.0%)	2 (0.5%)
Palmar-plantar erythrodysesthesia [b]	4 (2.0%)	4 (1.0%)
Vomiting [b]	3 (1.5%)	3 (0.7%)

[a,] AEs identified through clinical review in the MS and collected as per study protocol (AEs 'of special interest', AEs resulting in treatment discontinuation, SAEs) based on the later clinical cut off date of 23rd February 2009 taken from Table 7 of the MS
[b] AEs identified as commonly experienced by CAPE patients in other studies, highlighted by clinical advisors to the ERG and taken from Table 14.3/23 of the CSR

In addition to AEs identified as 'of special interest', the ERG notes from Table 14.3/23 of the CSR there were 15 (3.7%) cardiac disorders in the BEV+CAPE arm and 4 (2.0%) in the CAPE arm. These AEs include left ventricular systolic dysfunction which was identified as an AE 'of special interest' (Table 15). Other cardiac disorders reported in the BEV+CAPE arm included myocardial infarction (4 [1.0%]), cardiac arrest (2 [0.5%]), cardio−respiratory arrest, cardiac failure, cardiogenic shock, angina pectoris, pericardial effusion, sinus arrest and restrictive cardiomyopathy (all 1 [0.2%]). Pericardial effusion (2 [1.0%]) and atrial fibrillation (1 [0.5%]) were reported in the CAPE arm.

The MS states that the incidence of AEs in the open-label phase of RIBBON-1[16] was similar to that in the blinded treatment phase. Although not presented in the MS, an inspection of the CSR (Table 14.3/47) by the ERG suggests this to be so.

Because the safety of BEV has been identified as a potential concern in the past, the severity of AEs according to the NCI-CTC, in both the blinded and open-label phases of the trial are described in Appendix 4. Furthermore, because the proportion of AEs thought to be typically associated with CAPE appeared to be low in RIBBON-1,[16] the ERG made a crude comparison of these AEs across

two other phase II studies (one RCT[23] and one single arm study[18]) and a systematic review.[24] The data are summarised in Table 39 in Appendix 5 where it is evident that AEs of grade 3 or higher are also relatively uncommon in other studies but nevertheless were lower still in RIBBON-1.[16] The ERG proposes two reasons for this:

1. With the exception of RIBBON-1[16] and the single-arm XCALIBr[18] study, patients in the other studies (included in the systematic review[24]) had to have previously been treated with an anthracycline and/or a taxane and with the exception of these two studies patients were not necessarily receiving first-line treatment for mBC. Therefore it is probable that many of the patients in the other trials were receiving second or third-line treatment and therefore more susceptible to AEs.

2. With the exception of RIBBON-1[16] and the single-arm XCALIBr[18] study, all patients in other studies received CAPE at the licensed dose of 1250mg/m^2 or higher. These patients would therefore be expected to be more likely to experience AEs than at the 1000mg/m^2 dose.

Nevertheless, in the single-arm XCALIBr[18] study, 13% of BEV+CAPE patients reported hand-foot syndrome (i.e. palmar-plantar erythrodysesthesia syndrome). In RIBBON-1,[16] the proportion of BEV+CAPE patients who reported palmar-plantar erythrodysesthesia syndrome was 1% (and when CAPE patients are included, 1.3%). This discrepancy is harder to explain given the patient population appears to be broadly similar. It has been speculated in the report by the EMA[30] that regarding palmar-plantar erythrodysesthesia syndrome, these AEs were most likely classified as sensory neuropathy. However incidences of sensory neuropathy were still relatively rare (3% in the BEV+CAPE arm).

Considering all the data on AEs above, the manufacturer concludes that no new safety signals were noted and that the addition of BEV to CAPE does not lead to a clinically relevant increase in AEs that were typically associated with other chemotherapy regimens, such as febrile neutropenia, neutropenia, and sensory neuropathy. From the data presented in the MS and CSR, the ERG is in agreement with this view.

4.3 Clinical data for the manufacturer's subgroup included in the economic model

The only subgroup findings that were highlighted in the text of the MS were the PFS and OS findings for patients treated with a prior taxane. As highlighted in section 3.1 above and section 5.2.3 below, this population is that upon which the economic evaluation was modelled and which is considered a population of patients for whom the CAPE indication is most appropriate, i.e. is believed to be a proxy population for those who 'failed' a prior taxane and anthracycline.

Because the only data presented in the MS for this subgroup in the clinical section were the findings for PFS and OS, the ERG requested additional data on baseline characteristics, post-progression treatment and for all other outcomes, including AEs. The manufacturer did not supply further data, indicating in their response that PFS and OS findings were presented in the MS and that data on baseline characteristics, post-progression treatment and AEs were presented in Microsoft Excel worksheets submitted as part of the economic model. These data were extracted, interpreted and critiqued by the ERG.

As can be seen from Table 16, the population of patients who received a prior taxane appears to be quite different to the ITT population. In particular the ERG notes that it appears to be a younger and healthier population than the ITT population.

Table 16 Comparison of selected baseline characteristics presented in the model for the prior taxane subgroup and the ITT population of RIBBON-1

Demographic variable		Prior taxane subgroup		ITT population	
		CAPE (n=84)	BEV+CAPE (n=161)	CAPE (n=206)	BEV+CAPE (n=409)
Age (years)	Mean (SD)	53.4 (11.5)	53.4 (10.2)	57.1 (12.1)	56.6 (11.5)
	Median (range)	52 (23 to 78)	52 (30 to 84)	57 (23 to 88)	56 (28 to 91)
Age category	<40 years	9 (10.7%)	12 (7.4%)	15 (7.3%)	21 (5.1%)
	40-64 years	61 (72.6%)	126 (78.3%)	137 (66.5%)	289 (70.7%)
	>=65 years	14 (16.7%)	23 (14.3%)	54 (26.2%)	99 (24.2%)
Age group	<50 years	33 (39.3%)	59 (36.6%)	54 (26.2%)	119 (29.1%)
	>=50 years	51 (60.7%)	102 (63.4%)	152 (73.8%)	290 (70.9%)
Menopausal Status	Premenopausal	35 (41.6%)	60 (37.3%)	60 (29.1%)	120 (29.3%)
	Perimenopausal	4 (4.8%)	10 (6.2%)	11 (5.3%)	26 (6.4%)
	Postmenopausal	40 (47.6%)	85 (52.8%)	125 (60.7%)	245 (59.9%)
	Not Applicable	1 (0.1%)	0 (0.0%)	1 (0.5%)	1 (0.2%)
	Unknown	5 (5.9%)	6 (3.7%)	9 (4.4%)	17 (4.2%)
Sex	Female	83 (98.8%)	161 (100%)	204 (99.0%)	408 (99.8%)
Race/ ethnicity	White	58 (69.0%)	115 (71.4%)	157 (76.2%)	308 (75.3%)
	Black	7 (8.3%)	14 (8.7%)	10 (4.9%)	21 (5.1%)
	Other	19 (22.6%)	28 (17.3%)	39 (19.0%)	80 (19.5%)
Geographical region	North America	52 (61.9%)	118 (73.3%)	104 (50.5%)	226 (55.3%)
	Latin America	9 (10.7%)	10 (6.2%)	24 (11.7%)	42 (10.3%)
	Eastern Europe	2 (2.4%)	6 (3.7%)	32 (15.5%)	53 (13.0%)
	Western Europe	7 (8.3%)	11 (6.8%)	28 (13.6%)	57 (13.9%)
	Asia	14 (16.7%)	16 (9.9%)	18 (8.7%)	31 (7.6%)
ECOG performance status	0	48 (57.2%)	94 (58.8%)	110 (53.4%)	214 (52.7%)
	1	36 (42.8%)	66 (41.2%)	96 (46.6%)	192 (47.3%)

[a] All subgroup data is taken from the economic model (Microsoft Excel worksheet)
[b] All ITT data is taken CSR (Table 9 and Table14.1/34)
[c] For data on ECOG performance status, for the prior taxane group, n=84 in the CAPE arm and n=160 in the BEV+CAPE arm, for the ITT population, n=206 in the CAPE arm and n=406 n=615 (n=206 CAPE, n=406 BEV+CAPE)

In terms of PFS and OS, the differences between the BEV+CAPE and CAPE arms appeared to be greater in the subgroup than in the ITT population (Table 17). Indeed, for OS, the difference between the two arms was statistically significant unlike the difference between the two arms for the ITT population. However, the ERG reiterates that no statistical adjustments were performed to control for multiple testing in all subgroups and of all outcomes and so these findings may have therefore occurred by chance and must be interpreted with caution.

As noted in section 4.2, while no adjustment for cross-over and other post-progression treatment with anti-cancer drugs were pre-specified in the protocol of the RIBBON-1[16] trial, a *post-hoc* adjusted analysis based on the RPSFT model[29] was conducted. These findings are summarised in Table 17 where it can be seen no relative measures for the analysis using the RPSFT model were presented.

Table 17 Comparison of PFS and OS reported for the prior taxane subgroup and for the ITT population in RIBBON-1

Endpoint	Prior taxane subgroup		ITT population	
	CAPE (n=84)	BEV+CAPE (n=161)	CAPE (n=206)	BEV+CAPE (n=409)
PFS events	63 (75.0%)	115 (71.4%)	162 (78.6%)	291 (71.1%)
PFS (median, months)	4.2	8.7	5.7	8.6
	HR=0.62 (95% CI: 0.45 to 0.84)		HR=0.67 (95% CI: 0.55 to 0.82)	
Number of patients who died	44 (52.4%)	70 (43.5%)	99 (48.1%)	186 (45.4%)
OS (median, months)	20.5	28.4	22.8	25.7
	HR=0.67 (0.48 to 0.98)		HR=0.88 (95% CI: 0.69 to 1.12)	
OS (median, months) using RPSFT model [a]	15.0	24.0	-	-

[a] Estimate using RPSFT model taken from Table 32 of the MS,

The RPSFT model presented by the manufacturer for adjusting for cross-over is one of the approaches that have increasingly been used in the technology appraisals submitted to NICE (e.g. Pazopanib for the first-line treatment of advanced renal cell carcinoma[31]). This is an approach where an adjustment is made to the survival time of each patient who crossed over with the purpose of correcting their actual survival to reflect what their survival would have been had they not crossed over to the active treatment. As recently highlighted at an International Society for Pharmacoeconomics and Outcomes Research workshop,[32] this approach is unsuitable when a large proportion of patients cross-over from the control arm and when those in the intervention arm also 'cross-over' (in the BEV+CAPE arm, patients were able to receive additional BEV) and when other therapies are permitted as in RIBBON-1[16]. Considering these limitations, the ERG requested justification for the use of the RSPFT over other approaches such as the Inverse Probability of Censoring Weights (IPCW) method[33] from the

manufacturer. The ERG also requested estimates using both methods in order to compare the estimates from both approaches to adjusting for cross-over.

In response, the manufacturer acknowledged that the most appropriate method for accounting for cross-over in clinical trials is the subject of an ongoing academic debate. However the manufacturer perceived the IPCW method to involve more subjective choices than the RPSFT model, in that the calculation of the stabilised weights used in the weighted Cox proportional hazard regression model in IPCW may depend on the choice of the baseline covariates and the time-dependent covariates. In addition, the IPCW method requires that patients not crossing over are weighted more strongly to compensate for censoring of those who receive treatment following progression. In situations where such a large proportion of patients cross-over, the number of patients not crossing over is reduced and therefore their weighting is increased, potentially magnifying consequences of small errors. Taking all of these factors into consideration, the manufacturer determined that the RPSFT model was most appropriate for this submission and no estimates were provided using the IPCW method. Furthermore, given their justification for preferring the RPSFT model over the IPCW method, the manufacturer added they: "have doubts that such an analysis would result in a significant enough change in incremental cost effectiveness to affect the final decision."

The ERG does recognise the limitations of the IPCW alongside those of the RPSFT model. However, given the limitations of the RPSFT model highlighted above and in the absence of any other estimate to adjust for cross-over, the ERG is unable to confirm the likely effect of the cross-over and post-progression therapies on OS in this subgroup. In particular, the ERG is unable to ascertain whether the RPSFT model results in bias towards any particular treatment arm or not. Thus, additional caution should be exercised when interpreting the OS results presented in Table 17.

Around half the patients who initially received CAPE in the subgroup subsequently received BEV, slightly fewer than in ITT population but nevertheless, still a considerable proportion (Table 18). In the BEV+CAPE arm, around 45% subsequently received BEV in both the subgroup and ITT populations, again a large proportion of patients. Because the only data on specific therapies was provided in the model and because treatments were not mutually exclusive, it is not possible to determine the number of patients who received specific types of therapies as listed for the ITT population (e.g. for chemotherapy, adding up the numbers for each type of chemotherapy regimen together results in a total greater than the number of patients), other than for BEV (Table 18). However for each treatment, there did not appear to be any noticeable differences between the treatment arms. Confusingly, the proportion of patients in the subgroup who received subsequent BEV cited in the published paper[16] differed to the proportion cited in the economic model. However both sources report a greater proportion of CAPE patients received subsequent BEV.

Table 18 Subsequent anti-cancer therapy for the prior taxane subgroup and for the ITT population in RIBBON-1

Subsequent therapy received	Prior taxane subgroup		ITT population	
	CAPE (n=84)	BEV+CAPE (n=161)	CAPE (n=206)	BEV+CAPE (n=409)
Patients who 'crossed over' to receive additional BEV [a]	44 (52.4%)	72 (44.7%)	120 (59.7%)	184 (45.5%)
Patients who received subsequent therapy [b]	-	-	142 (68.9%)	251 (61.4%)
Type of therapy:				
• BEV [b]	43 (51.2%)	67 (41.6%)	112 (54.4%)	160 (39.1%)
• Chemotherapy [b]	-	-	135 (65.5%)	226 (55.3%)
• Hormonal therapy [b]	-	-	28 (13.6%)	51 (12.5%)
• Radiotherapy [b]	-	-	12 (5.8%)	35 (8.6%)
• Surgery [b]	-	-	4 (1.9%)	3 (0.7%)
• Other [b]	-	-	8 (3.9%)	12 (2.9%)

[a] Data taken from Table 14 of the MS
[b] Data taken from Table 3 of the published paper[16] for the ITT population – data not provided for subsequent therapy for the prior taxane subgroup; although data on subsequent therapy is provided in the economic model (Microsoft Office worksheet), the treatments are not mutually exclusive and the number of patients cannot therefore be determined other than for BEV

Regarding the safety of BEV+CAPE compared to CAPE in the prior taxane subgroup, it was not possible to compare the proportion of patients who experienced any AE, any grade 3–5 AE, any SAE or any AE leading to discontinuation of BEV or placebo because the manufacturer did not present these data. Nor did the manufacturer present data for all deaths (including disease progression) and deaths unrelated to disease progression. However, from the data extracted from the economic model, it is known there was one (0.62%) sudden death in the BEV+CAPE arm.

From the same data source, it was also possible to extract data on AEs 'of special interest' which on the whole appeared to be similar in frequency as in the safety population (Table 19). In addition to AEs 'of special interest', according to the economic model the proportion of patients reporting a cardiac disorder of grade 3 or higher was greater in the BEV+CAPE arm (4.4%) than the CAPE arm (no events reported). Cardiac disorders reported were: cardiac arrest (two [1.2%] patients), cardiac failure, cardio-respiratory arrest, cardiogenic shock, myocardial infarction and pericardial effusion (all one patient [0.6%]). All were grade 4 or grade 5 AEs. For the BEV+CAPE arm, the overall proportion of cardiac disorders is a slightly greater proportion than reported in the safety population for cardiac disorders (2.1%). However, the ERG urges caution in interpreting the findings because of the small numbers of patients (and therefore smaller number of AEs) in this subgroup.

Table 19 Patients experiencing at least one AE 'of special interest' [a] (NCI-CTC grade ≥3) in blinded treatment phase for the prior taxane group [b] and safety population [c]

AE 'of special interest'	Prior taxane subgroup		Safety population	
	CAPE (n=84)	BEV+CAPE (n=161)	CAPE (n = 201)	BEV+CAPE (n = 404)
Arterial thromboembolic event [b, c]	0 (0.0%)	0 (0.0%)	3 (1.5%)	6 (1.5%)
Bleeding [b, c]	0 (0.0%)	0 (0.0%)	1 (0.5%)	1 (0.2%)
Febrile neutropenia [b, c]	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Fistula [b, c]	0 (0.0%)	0 (0.0%)	1 (0.5%)	1 (0.2%)
Gastrointestinal perforations [b, c]	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Hypertension [b, c]	0 (0.0%)	14 (8.7%)	2 (1.0%)	38 (9.4%)
Left ventricular systolic dysfunction [b, c]	0 (0.0%)	0 (0.0%)	1 (0.5%)	4 (1.0%)
Neutropenia [b, c]	1 (1.2%).	2 (1.2%)	2 (1.0%)	5 (1.2%)
Proteinuria [b, c]	0 (0.0%)	6 (3.7%)	0 (0.0%)	9 (2.2%)
Reversible posterior leukoencephalopathy syndrome [b, c]	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Sensory neuropathy	0 (0.0%)	4 (4.3%)	1 (0.5%)	12 (3.0%)
Venous thromboembolic event [b, c]	0 (0.0%)	0 (0.0%)	7 (3.5%)	19 (4.8%)
Wound dehiscence [b, c]	0 (0.0%)	1 (0.6%)	0 (0.0%)	3 (0.7%)
Diarrhoea [b,d]	2 (2.5%)	2 (1.2%)	4 (2.0%)	6 (1.5%)
Fatigue [b,d]	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Nausea [b,d]	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (0.2%)
Palmar-plantar erythrodysesthesia [b,d]	0 (0.0%)	4 (2.5%)	3 (1.5%)	4 (1.0%)
Vomiting [b,d]	1 (1.2%)	0 (0.0%)	2 (1.0%)	2 (0.5%)

[a] AEs identified through clinical review in the MS and collected as per study protocol (AEs 'of special interest', AEs resulting in treatment discontinuation, SAEs) based on the later clinical cut off date of 23 February 2009 (taken from Table 7 of the MS) or AEs identified as commonly experienced by CAPE patients in other studies, highlighted by clinical advisors to the ERG and taken from Table 14.3/23 of the CSR

[b] AEs taken from economic model (Microsoft Excel worksheet) and Table 28 of MS for prior taxane subgroup

[c] AEs taken from Table 60 of the MS and [d] AEs taken from Table 14.3/23 of the CSR for ITT population

4.4 Conclusions of the clinical effectiveness section

The clinical effectiveness evidence is derived from a single, relatively large, well conducted, manufacturer supported RCT (RIBBON-1[16]) which compares BEV+CAPE to CAPE. The trial reported an additional PFS and OS benefit of around 3 months for patients in the BEV+CAPE arm over the CAPE arm (investigator assessed median PFS: 8.6 vs 5.7 months; IRC assessed median PFS: 9.8 vs 6.2 months; median OS: 25.7 vs 22.8 months). Since the HRs for investigator and IRC assessed PFS were almost identical (HR=0.69 [95% CI: 0.56 to 0.84] and HR=0.68 [95% CI: 0.54 to 0.86] respectively), the evidence suggesting a benefit in terms of PFS does appear to be robust. However, despite significant improvements also in ORR for the BEV+CAPE arm (35.4% compared to 23.6%), only the PFS and not the OS findings were statistically significant. The lack of a statistically significant difference in OS between the groups may be explained by differences in the nature and frequency of subsequent treatments received in both arms of the trial following diseaseprogression.

Given anthracyclines or taxanes were possible treatment options prior to randomisation to receive BEV+CAPE or CAPE, it seems reasonable to assume patients in this cohort would be those considered unsuitable for an anthracycline or a taxane. Therefore the ERG believes the study population of RIBBON-1[16] was that which was specified in the decision problem in the scope issued by NICE. The intervention (BEV+CAPE) and comparator (CAPE) were also as specified in the decision problem although the CAPE dose was 1000mg/m^2 rather than the licensed dose of 1250mg/m^2. Given this lower dose is commonly used in clinical practice, the ERG has no concerns about this. With the exception of HRQoL, the outcomes measured in RIBBON-1[16] were also in accordance with the decision problem in the scope issued by NICE. The manufacturer addressed HRQoL indirectly by arguing that improving PFS (and OS) results in improved HRQoL for patients.

Compared to CAPE patients, a greater proportion of patients in the BEV+CAPE arm in RIBBON-1[16] reported any AE (40% vs 27% in the CAPE arm), SAEs (25% vs 20%) and grade 3-5 AEs (37% vs 23%). The ERG believes the difference between the two arms can largely be attributed to differences in grade 3 AEs (27% vs 14%). While a greater number of AEs were reported in the BEV+CAPE arm than the CAPE arm, including AEs 'of special interest' (23% vs 9%), no new safety concerns were identified.

As highlighted in section 3.1, the manufacturer does not use the population of all patients treated with BEV+CAPE or CAPE in RIBBON-1[16] to derive its evidence for cost effectiveness (see section 5), rather it uses a subgroup of patients who had received a prior taxane. As emphasised in section 3.1 above, the ERG does not believe the subgroup of patients who received a prior taxane is a more appropriate group of patients than the ITT population since it is this larger ITT population that meets the criteria specified in the NICE scope and the marketing indication for BEV+CAPE. In addition, the ERG has identified that there appear to be baseline differences between this subgroup and the entire ITT population. In particular, the population of patients who received a prior taxane appear to be younger and healthier.

In terms of clinical effectiveness for this subgroup, the differences in PFS (4.5 months) and OS (around 8 months) between the BEV+CAPE and CAPE arms appeared to be greater in the subgroup (median PFS: 8.7 vs 4.2 months; median OS: 28.4 vs 20.5 months) than in the ITT population. Both differences in the subgroups were statistically significant. However no statistical adjustments were performed to control for multiple testing in all subgroups and of all outcomes, thus increasing the likelihood of significant results emerging by chance. On the whole, AEs 'of special interest' appeared to be similar in frequency in the subgroup as safety population. A slightly greater proportion of patients in the subgroup reported cardiac disorders (4.4%) than in the overall trial population (2.1%). However, the ERG urges caution in interpreting the findings because of the small numbers of patients (and therefore smaller number of AEs) in this subgroup.

5 COST EFFECTIVENESS

This section provides a structured critique of the economic evidence submitted by Roche in support of BEV+CAPE for the treatment of mBC. The two key components of the economic evidence presented in the MS are (i) a systematic review of the relevant literature and (ii) a report of the manufacturer's *de novo* economic evaluation. Table 20 contains details of the location of key information within the MS. The manufacturer also provided an electronic version of their economic model which was developed in Microsoft Excel.

Table 20 Location of key cost-effectiveness information in the MS

Key information	Page number	Tables/figures
Details of the systematic review of the economic literature	73-75, 164-169	
De novo analysis	76-80	Tables 11-12, Figure 10
Clinical evidence used in economic evaluation	81-92	Tables 13-17, Figures 11-15
Measurement and valuation of health effects	93-102	Tables 18-21, Figures 16
Resource identification, measurement and valuation	103-116	Tables 22-28, Figures 17-20
Methods of sensitivity analysis	117-120	Tables 29-31
Results - base-case analysis	121-132	Tables 32-39
Results - sensitivity analysis	133-136	Tables 40-42, Figures 21-22
Validation	137	Figure 23
Interpretation of economic evidence	138-139	
Assessment of factor relevant to the NHS and other parties	140-141	

5.1 ERG comment on manufacturer's review of cost-effectiveness evidence

5.1.1 Objective of the manufacturer's cost-effectiveness literature review

The MS states that the search was designed to evaluate whether *de novo* modelling was necessary in order to answer the decision problem set out in the scope issued by NICE. Outline details of the manufacturer's search strategy are presented in the MS with full details in an appendix (Appendix 10, pg 164-169). Dialogue Data-Star was used to search Embase, Medline, Medline (R) In-Process and EconLit, whilst the NHS EED database was searched using the University of York's Centre for Reviews and Dissemination website. The Data-Star searches were carried out on 16th November 2011 and the EconLit and NHS EED searches on 2nd December 2011. The date span for the searches was from 1993 up until the date of each search.

The manufacturer appears not to have undertaken any searches of the unpublished literature; however, the ERG considers that finding any relevant unpublished studies is unlikely and concludes that the search strategy used by the manufacturer was appropriate.

5.1.2 Inclusion and exclusion criteria used in study selection

The inclusion/exclusion criteria used in the study selection are presented in Table 21.

Table 21 Economic evaluation search inclusion and exclusion criteria

Parameter	Inclusion criteria	Exclusion criteria
Population	Previously untreated advanced breast cancer patients	Non-breast cancer patients, previously treated patients
Intervention	BEV+CAPE	-
Comparator	CAPE	-
Outcome	Cost per QALY gained Cost per LY gained	-
Study design [a]	Economic evaluation (cost effectiveness analyses, cost utility analyses, cost minimisation analyses)	RCTs, observational data, budget impact assessments

[a] During the record sifting process records were excluded if they were not a cost-utility analysis

5.1.3 Included and excluded studies

No relevant studies were identified.

5.1.4 Conclusions of the cost-effectiveness literature review

The manufacturer's review of the published cost-effectiveness literature describing BEV+CAPE vs CAPE for previously untreated advanced breast cancer patients did not identify any relevant cost-effectiveness studies. The ERG is satisfied with the manufacturer's search strategy and is reasonably confident that the manufacturer did not miss any relevant published articles.

5.2 Summary and critique of manufacturer's submitted economic evaluation by the ERG

5.2.1 NICE reference case checklist

Table 22 shows how closely the manufacturer's submitted economic evaluation accords with the requirements for a base-case analysis as set out in the NICE reference case checklist. In general the manufacturer's analysis matches the requirements set by NICE.[34]

Table 22 NICE reference case checklist

Attribute	Reference case	Does the *de novo* economic evaluation match the reference case?
Decision problem	The scope developed by the Institute	Yes
Comparator(s)	Alternative therapies routinely used in the NHS	Partially - VIN is only considered in a scenario analysis
Perspective costs	NHS and PSS	Partial - PSS costs are not considered
Perspective benefits	All health effects on individuals	Yes
Form of economic evaluation	Cost-effectiveness analysis	Yes
Time horizon	Sufficient to capture differences in costs and outcomes	Partial -a time horizon of 15 years is used but subsequent lines of therapy are not modelled
Synthesis of evidence on outcomes	Systematic review	N/A – the manufacturer only uses data from the RIBBON-1[16] trial
Outcome measure	Quality adjusted life years	Yes
Health states for QALY	Described using a standardised and validated instrument	Yes. The manufacturer uses values from published literature that have been used in previous STAs
Benefit valuation	Time-trade off or standard gamble	Yes
Source of preference data for valuation of changes in HRQoL	Representative sample of the public	Yes
Discount rate	An annual rate of 3.5% on both costs and health effects	Yes
Equity	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	Yes
Sensitivity analysis	Probabilistic sensitivity analysis	Yes

5.2.2 Model structure

Three health states are used to model disease progression. All patients enter the model in the PFS health state and in each month can either progress to a 'worse' health state (i.e. from PFS to Progressed or Death, or from Progressed to Death) or remain in the same health state. Second-line therapy is not considered in the model. The model has been developed in Microsoft Excel and has a one month cycle length, includes a half-cycle correction as recommended by NICE and the time horizon is set at 15 years. The model structure is shown in Figure 1.

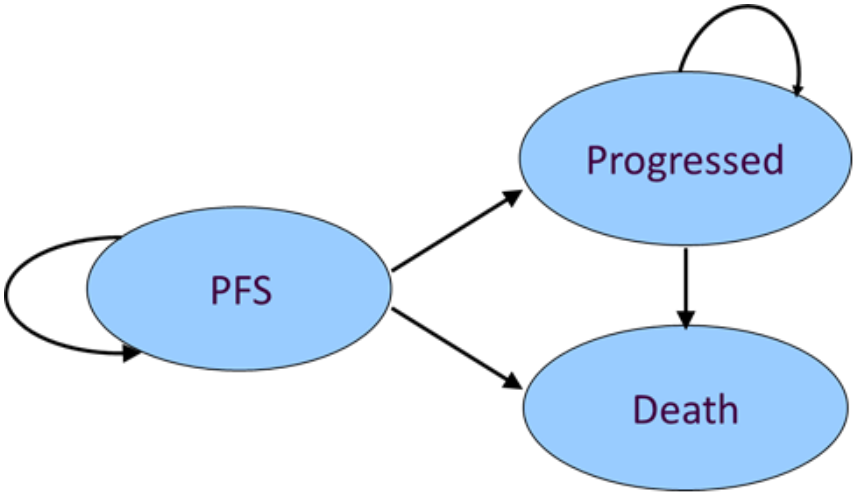

Figure 1 Schema of manufacturer's model

Parameters and values

Key population parameters used in the cost-effectiveness analysis are presented in Table 23.

Table 23 Key parameters in the cost-effectiveness analysis

Patient variables	Value	Source
Patient age	53.0years	RIBBON-1[16] study prior taxane subgroup
Patient weight	72.1kg	RIBBON-1[16] study prior taxane subgroup
Patient height	160.89cm	RIBBON-1[16] study prior taxane subgroup
Body surface area	1.7609m^2	RIBBON-1[16] study prior taxane subgroup

Note: Where there are discrepancies, the values in the table are those used in the model rather than those reported in the MS

5.2.3 Population

The modelled population is a subgroup of the RIBBON-1[16] trial population. The ERG has concerns that this subgroup may not be representative of the population licensed to receive BEV+CAPE. A full discussion of the issues may be found in section 4.3 of this report; the key points are that:

- Patients in the modelled subgroup are younger and healthier than the ITT population;
- Detailed trial data are not available on the treatments received post progression for this subgroup.

5.2.4 Interventions and comparators

Modelling of BEV+CAPE

BEV+CAPE is modelled as administered in the RIBBON-1[16] study (i.e. BEV 15 mg/kg every 21 days and 1000 mg/m^2 CAPE administered twice daily for 14 days of each 21 day cycle followed by a 7 day 'rest' period. For both drugs, treatment is continued until disease progression or unacceptable toxicity and treatment duration is a maximum of 48 months).

Modelling of CAPE

CAPE is modelled as administered in the RIBBON-1[16] study (i.e.1000 mg/m^2 CAPE administered twice daily every day until progression, unacceptable toxicity or a maximum of 48 months of treatment). The manufacturer points out that this dose is different from the SPC specified dose for CAPE[21] in which it is recommended that CAPE be given at a dose of 1250 mg/m^2 (25% higher than the dose used in RIBBON-1[16]). The ERG has explored the impact of use of this higher dose in a sensitivity analysis.

Modelling of VIN

Market research data from interviews of 43 clinical oncologists and 27 medical oncologists conducted in April 2010 indicate that CAPE holds a market share approximately five times that of VIN. Additionally, VIN has previously been assumed to have equivalent efficacy when compared with CAPE (NICE clinical guideline for advanced breast cancer[3]). These factors led to the manufacturer concluding that VIN was not an appropriate comparator. However, the manufacturer does explore the cost effectiveness of BEV+CAPE against different formulations of VIN in scenario analyses.

ERG expert advice supports the manufacturer's decision to omit VIN from the main analysis.

5.2.5 Perspective, time horizon and discounting

The economic evaluation is undertaken from the perspective of the NHS and Personal Social Services. Outcomes are expressed in terms of gains in life-years and quality adjusted life-years (QALYs). The time horizon is set at 15 years and, in line with the NICE Methods Guide to Technology Appraisal,[34] both costs and benefits are discounted at 3.5%.

5.2.6 Treatment effectiveness and extrapolation

The most recent data-cut of the RIBBON-1[16] RCT (23rd February 2012) was used in the model.

PFS

The base case model uses the probability of remaining in PFS observed in the RIBBON-1[16] trial for each arm directly until the twelfth month of treatment, after which the survival is extrapolated according to an exponential function. The number of patients in each treatment arm dying from any cause while in PFS is used to derive a constant rate and probability of mortality (see Table 24) which is assumed to be at least as great as the underlying sex-and age-related mortality in the general population.

Table 24 Transition probabilities in the cost-effectiveness analysis

Comparator	Estimation method	Value	Source
Probability of dying in PFS			
BEV+CAPE	Derived from RIBBON-1[16]	0.00527	MS Section 6.3.1
CAPE	Derived from RIBBON-1[16]	0.00640	MS Section 6.3.1

PD

During the PD phase patients in the RIBBON-1[16] trial receive a variety of different therapies; however, the model does not differentiate between these different drugs. The manufacturer models survival in PD based on 'uncrossed' RIBBON-1[16] trial data up to month 12, with curves extrapolated according to an exponential function thereafter. The data have been 'uncrossed' to try to take account of the effect resulting from patients in the BEV+CAPE continuing on BEV and patients in the CAPE arm crossing over to receive BEV. A critique of the method used to uncross the data can be found in section 4.3 of this report. No account is taken of the possible impact of receipt of other drugs.

RIBBON-1[16] trial KM data show that patients who enter PD experience an increasing probability of dying each month they spend in this state. To take account of this, new patients with PD enter an array of monthly temporary states, called tunnel states, and the probability of death increases with increasing time in PD. The tunnel states are arranged so that each state has a transition only to death or the next temporary state.

OS

Monthly OS is derived by adding together the estimated proportions of patients in PFS and PD.

5.2.7 Health related quality of life

EQ-5D data were not collected in the RIBBON-1[16] study and the manufacturer therefore undertook a review of literature to identify relevant HRQoL data to use in the economic evaluation. The search found three studies.[35-37] The manufacturer concluded that it would be most appropriate to calculate utilities from the results of the mixed model analysis presented by Lloyd et al[37] as values from this study had been used in previous health technology appraisals for mBC.[38] The utility values used in the model are displayed in Table 25.

Table 25 Utility values in the cost-effectiveness analysis

State	Utility value	Source
Reported utility values		
PFS BEV+CAPE	0.784	Calculated from results reported by Lloyd et al 2006[37]
PFS CAPE	0.774	Calculated from results reported by Lloyd et al 2006[37]
PD	0.496	Calculated from results reported by Lloyd et al 2006[37]

5.2.8 Resources and costs

Intervention costs

Intervention costs are made up of the cost of BEV and the cost of CAPE and the associated administration and pharmacy costs. For the purposes of the base case the manufacturer assumes that no vial sharing takes place.

Comparator costs

Comparator costs are made up of the cost of CAPE and the associated administration and pharmacy costs. Intervention and comparator costs are summarised in Table 26.

Table 26 Intervention and comparator drug costs in the cost-effectiveness analysis

Costs	Value	Source
BEV+CAPE drug costs	£4001.53 per month	BNF 62[39]
Month 1: BEV+CAPE administration and pharmacy cost	£348.82 per month	Millar 2008[30] NHS Reference Costs 2009/10 (SB13Z: Deliver more complex parenteral chemotherapy at first attendance (Day Case))[40] PSSRU 2010[41]
Subsequent months: BEV+CAPE administration and pharmacy cost	£205.99 per month	Millar 2008[30] NHS Reference Costs 2009/10 (SB97Z: Same day chemotherapy admission/attendance (Day case and Regular Day/Night))[40] PSSRU 2010[41]
CAPE drug cost	£312.41 per month	BNF 62[39]
CAPE administration and pharmacy cost	£255.32 per month	Millar 2008[30] NHS Reference Costs 2009/10 (SB11Z: Deliver exclusively oral chemotherapy)[40] PSSRU 2010[41]

Health care costs

Resource use in the economic evaluation is not derived from data collected as part of the RIBBON-1[16] trial. Monthly supportive care costs associated with PFS are assumed to be the same as described in NICE clinical guideline for advanced breast cancer[3] 'Package 1',which have been interpreted as:

- Two visits from a community nurse (each of duration 20 minutes)
- One consultation with a GP (in surgery)
- One visit from a community nurse specialist (duration 1 hour).

Additionally, to assess response to treatment and/or progression of disease, it is assumed that each patient will have an outpatient consultation with an oncologist and a CT scan every three months. Monthly supportive care costs associated with PD are assumed to be the same as described in NICE clinical guideline for advanced breast cancer[3] 'Package 2', which have been interpreted as:

- Four visits from a community nurse (each of duration 20 minutes)
- One home visit from a GP
- Four visits from a community nurse specialist (each of duration 1 hour)
- Two sessions (each of duration 1 hour) with an NHS community occupational therapist.

The ERG notes that the cost associated with the terminal phase of mBC is not included in the model. Furthermore, the impact on social services of supporting people with mBC and their families is not considered.

Costs and sources for health state costs are displayed in Table 27.

Table 27 Health care costs in the cost-effectiveness analysis

Health state	Value	Source
PFS	£263.55 per month	NICE CG81[3] NHS Reference Costs 2009/10[40] PSSRU 2010[41]
PD	£804.00 per month	NICE CG81[3] NHS Reference Costs 2009/10[40] PSSRU 2010[41]

Adverse event costs

Only those AEs occurring in >2% of patients at grade 3/4 severity are incorporated into the analysis. Where clinical advice indicated that the usual treatment pathway was discontinuation of treatment it was assumed that this had been accounted for elsewhere in the model and no additional costs were accrued. All AEs were assumed to occur in month 1 and so were not discounted. Costs and sources for AEs are displayed in Table 28.

Table 28 Key model parameters: AEs

Adverse event	Cost per episode	Source
Deep vein thrombosis	£388.84	NHS Reference Costs 2009/10 (QZ20Z - Non-elective short stay deep vein thrombosis)[40]
Hypertension	£455.40	NHS Reference Costs 2009/10 (EB041 - Non-elective short stay hypertension w/out cc)[40]
Peripheral sensory neuropathy	N/A [a]	Expert opinion
Diarrhoea	N/A	Assumed to be treated with generic rehydration therapies and/or anti-motility agents which have a negligible contribution to costs.
Palmar-plantar erythrodysesthesia syndrome	N/A [a]	Expert opinion
Proteinuria	N/A [a]	Expert opinion

[a] Treatment is discontinuation of chemotherapy
cc = complications and comorbidities

5.2.9 Cost-effectiveness results

The base case incremental results generated by the manufacturer's model are presented in Table 29. The ICER for the target population is £77,318 per QALY gained and £45,073 per life year gained. A summary of predicted resource use by category of cost is presented in Table 30.

Table 29 Base-case results

Technologies	Total costs	Total LYG	Total QALYs	Inc. costs	Inc. LYG	Inc. QALYs	ICER (cost/LY)	ICER (cost/ QALY)
CAPE	£12,721	1.3648	0.8346					
BEV+CAPE	£51,645	2.2283	1.3381	£38,924	0.8636	0.5034	£45,073	£77,318

Table 30 Summary of predicted resource use by category of cost for the base case

Unit Cost	Cost BEV+CAPE	Cost CAPE	Increment	Absolute increment	% absolute increment
Mean total treatment cost (BEV+CAPE)	£30,840	0	£30,840	£30,840	79.2%
Administration cost (BEV+CAPE)	£1,779	0	£1,779	£1,779	4.6%
Mean total treatment cost (CAPE)	£2,612	£1,714	£898	£898	2.3%
Administration cost (CAPE)	£83	£1,401	-£1,318	£1,318	3.4%
Mean supportive care cost of PFS	£2,553	£1,737	£816	£816	2.1%
Mean supportive care cost of PD	£13,711	£7,869	£5,841	£5,841	15.0%
Cost of AEs	£68	0	£68	£68	0.2%
Total	**£51,645**	**£12,721**	**£38,924**	**£38,924**	**100%**

5.2.10 Sensitivity analyses

The manufacturer varied costs (± 40%), utilities (± 25%), discount rates (0-6%) and the time horizon (± 5 years); and also fitted alternative survival curves (using Gompertz curves for PFS and PD and a Weibull curve for time to off treatment (TTOT)). The results presented in Table 31 demonstrate that the ICER for BEV+CAPE in these patients is most sensitive to assumptions concerning all utilities, particularly those relating to PD and the PD parametric curve.

Table 31 Deterministic univariate sensitivity analysis results

Costs and outcomes	Base case Value	High Value	Low Value	High Result (ICER)	Low Result (ICER)
Costs				**£77,318**	
BEV administration (1st month)	£338.94	£474.52	£203.36	£77,581	£77,055
BEV administration (other months)	£196.09	£274.52	£117.65	£78,468	£76,168
CAPE administration in BEV arm	£9.88	£13.83	£5.93	£77,384	£77,252
CAPE administration in CAPE arm	£255.32	£357.45	£153.19	£76,205	£78,431
PFS BSC	£263.55	£368.97	£158.13	£77,967	£76,669
PD BSC	£804.00	£1,125.60	£482.40	£81,959	£72,677
All costs				£82,974	£71,662
Outcomes					
PFS utility (BEV)	0.78419	0.98419	0.58419	£68,665	£88,466
PFS utility (CAPE)	0.77364	0.97364	0.57364	£79,175	£75,546
PD utility	0.49612	0.69612	0.29612	£62,327	£101,804
All utilities above				£57,568	£117,698
PFS parametric fit	KM	Gompertz			£74,879
PD parametric fit	KM	Gompertz			£110,092
TTOT parametric fit	KM	Weibull			£83,030
Other					
Cost discount rate	0.035	0.06	0	£76,209	£79,064
Health outcomes discount rate	0.035	0.06	0	£81,020	£72,068
Time horizon	15	10	5	£77,457	£82,351

TTOT, time to off treatment

Scenario analyses

The manufacturer undertook two scenario analyses, the first using utility values published by Peasgood et al[42] and the second comparing BEV+CAPE with three different VIN combinations.

Figures in Table 32 show that, using the utility values reported by Peasgood et al[42] has little effect on the incremental cost-effectiveness ration per QALY gained (ICER range:£77,815 to £79,991, compared with the base case of £77,318).

Table 32 Incremental cost-effectiveness per QALY gained using a range of utility values

Comparator	Lloyd et al [37]	Peasgood et al [42]		
		Model 1	Model 2	Model 3
PFS BEV+CAPE	0.7842	0.7435	0.9132	0.8817
PFS CAPE	0.7736	0.7569	0.8984	0.8679
PD	0.4961	0.4880	0.4350	0.4350
ICER	**£77,318**	**£79,991**	**£77,851**	**£79,147**

The incremental cost-effectiveness ration per QALY gained for BEV+CAPE against different VIN formulations are shown in Table 33. In these analyses VIN was assumed to have an equivalent efficacy and safety profile to CAPE, with different costs of acquisition and administration.[3]

Table 33 Incremental cost-effectiveness per QALY gained of BEV+CAPE vs VIN formulations

Comparator	ICER	Comment
Intravenous branded VIN regimen	£76,199	Very similar to the base case (£77,318)
Generic VIN	£80,260	Less cost-effective than the base case
Oral formulation of VIN	£58,198	More cost effective than the base case

Main findings from the univariate sensitivity and scenario analyses

The sensitivity analyses undertaken by the manufacturer show that the key drivers of the cost-effectiveness results are:

- Utilities, especially those associated with PD;
- The parametric functions describing PFS and PD;
- Drug consumption as portrayed by the TTOT parametric fit.

Additionally, the manufacturer states that the key drivers of cost-effectiveness results include the cost of BEV (MS pg 136).

Probabilistic sensitivity analysis

The manufacturer also undertook probabilistic sensitivity analysis (PSA) to derive the mean ICER of BEV+CAPE vs CAPE. The distributions used in the PSA are summarised in Table 34.

Table 34 Distributions used in the cost-effectiveness PSA

Costs/outcomes	Distribution	Logic
Costs		
PFS BSC and monitoring	Gamma (7.11, 37.061719)	Gamma is positively constrained (as costs are) and allows the possibility of high 'outlier' values.
PD BSC	Gamma (7.11, 113.06250)	
Health outcomes		
PFS utility	Beta (0.78, 0.000169)	Utility value far enough from 0 to warrant a transformed (1-x) normal function unnecessary. Constrained at the upper end by 1.
PD utility	Beta (0.50, 0.000250)	

The manufacturer states that the PSA results suggest that there is less than a 0.1% chance that the ICER for BEV+CAPE is less than £50,000 per QALY A scatter plot (incremental cost vs QALY) and a cost-effectiveness acceptability curve are included in the MS and reproduced in Figure 2 and Figure 3 respectively.

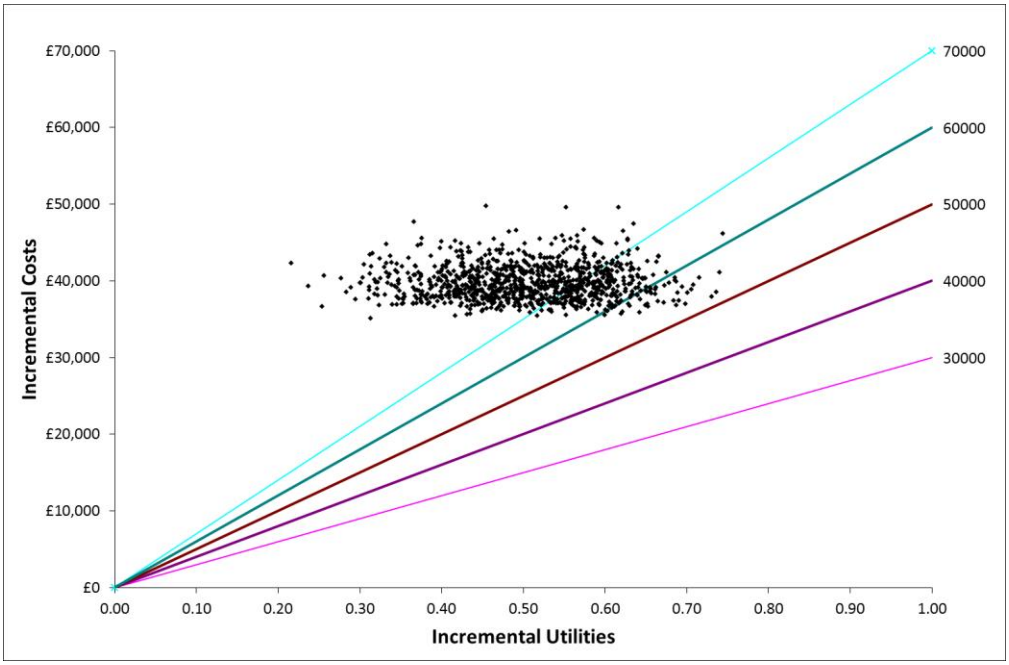

Figure 2 Cost-effectiveness plane for the addition of BEV to CAPE in mBC

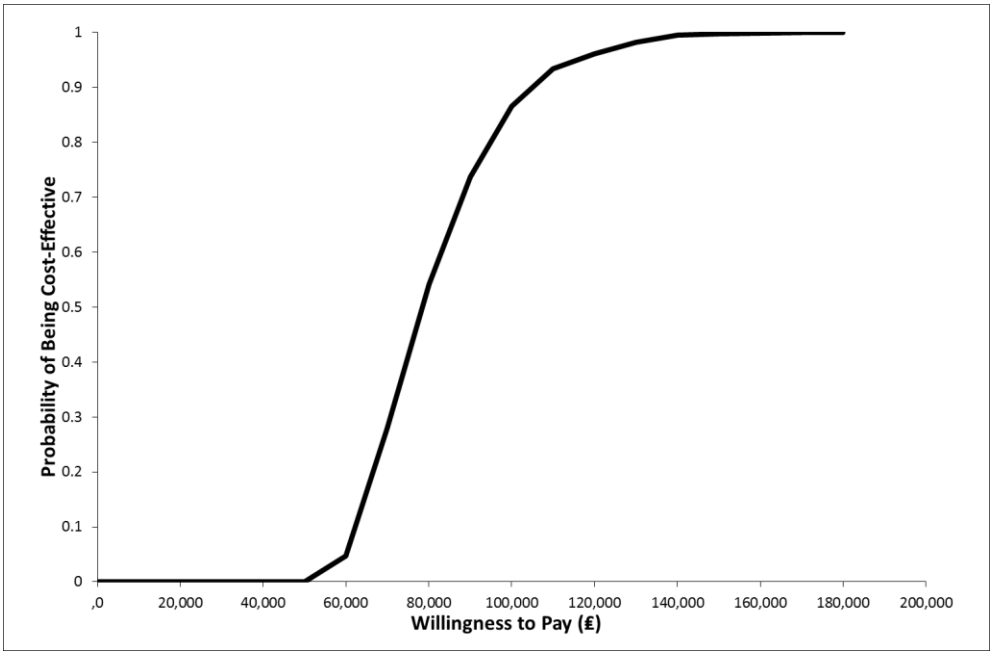

Figure 3 Cost-effectiveness acceptability curves for the addition of BEV to CAPE in mBC

5.2.11 Model validation and face validity check

It is reported that no clinical experts were consulted in the development of this economic model. The manufacturer felt that having recently held two advisory boards to obtain validation of the assumptions and inputs utilised in other mBC economic models (BEV in combination with a taxane and trastuzumab in combination with an aromatase inhibitor) rendered further validation of resource use inputs unwarranted.

5.3 Detailed critique of manufacturer's economic model

Table 35 summarises the ERG's appraisal of the economic evaluation conducted by the manufacturer using the Drummond 10-point checklist.[43]

Table 35 Critical appraisal checklist for the cost-effectiveness analysis

Question	Critical appraisal	ERG comment
Was a well-defined question posed in answerable form?	Yes	
Was a comprehensive description of the competing alternatives given?	Yes	ERG agrees that with the manufacturer that VIN is not a valid comparator
Was the effectiveness of the programme or services established?	No	The modelled population is a subgroup of the licensed population. RIBBON-1[16] trial data indicate that, at baseline, this subgroup is younger and healthier than the licensed population. Therefore, model results may be optimistic.
Were all the important and relevant costs and consequences for each alternative identified?	Mostly	The ERG notes that the economic model does not include terminal care costs. Additionally, social care costs are not considered.
Were costs and consequences measured accurately in appropriate physical units?	Not always	Sources of resource use and cost data were appropriate (e.g. NHS Reference Costs 09-10[40] PSSRU 2010[41] and NICE guidelines CG81[3]) ERG prefers to incorporate distribution of body surface area/weight and UK patient characteristics into cost calculations where appropriate
Were the cost and consequences valued credibly?	Not always	The ERG identified an error in the calculation of utility values. The approach used by the manufacturer to model survival in PD may not accurately reflect the effectiveness of BEV.
Were costs and consequences adjusted for differential timing?	Yes	
Was an incremental analysis of costs and consequences of alternatives performed?	Yes	Plenty of detail is presented by the manufacturer as per the NICE template
Was allowance made for uncertainty in the estimates of costs and consequences?	Yes	SA and PSA were undertaken
Did the presentation and discussion of study results include all issues of concern to users?	Mostly	It would have been informative if the costs and benefits of subsequent lines of treatment had been explicitly included in the economic model

5.3.1 Model structure and design

The manufacturer has adopted a simple three-state model design, adapting a model structure previously used in several submissions to NICE appraisals of cancer drugs. The model is driven by survival models governing PFS and PD, calibrated against data from the RIBBON-1[16] clinical trial.

An important limitation of the approach taken is that although the model covers a period of 15 years, no further chemotherapy is considered within the model following disease progression after treatment with either BEV+CAPE or CAPE. This could lead to substantial bias as, if there is better PFS in one arm than the other, the discounted costs and benefits of subsequent treatments will differ. Furthermore, if the proportion of patients able to receive subsequent lines of therapy differs between the arms then the costs and outcomes will also differ. Additionally, the omission of subsequent lines of treatment from the model is contrary to the expectations in the NICE Methods Guide to Technology Appraisal[34] that models will encompass all likely consequences of an innovative treatment over a whole lifetime.

5.3.2 Model implementation

The manufacturer's model is implemented as a series of Microsoft Excel worksheets. The layout of the model is generally clear and tables are adequately labelled; however, the inclusion of superfluous sheets and formulas relating to parametric models which were considered during the model development process but not actually implemented in the final version of the model can make navigation confusing.

5.3.3 Estimation of patient outcomes

PFS

The modelling approach used by the manufacturer to estimate PFS involves the direct use of KM data from the RIBBON-1[16] trial for the first 12 months and a fitted exponential curve thereafter. This approach appears credible.

PD

Although a similar approach, in terms of using trial data for the first 12 months and a parametric curve thereafter, was used to model PD, the ERG is concerned that the design of the RIBBON-1[16] trial allowed patients to receive BEV (subject to the consulting physician's discretion) post progression. The manufacturer felt that this may have introduced bias in estimation of treatment effects as patients randomised to the control arm may have had their survival prolonged due to receiving the study drug after disease progression. The manufacturer has therefore used the RPSFT model to 'uncross' data prior to modelling survival in PD. However, as highlighted in section 4.3, this approach is unsuitable when a large proportion of patients cross-over from the control arm and when those in the

intervention arm also 'cross-over'. For the modelled population 44.7% of the BEV+CAPE arm and 52.4% of the CAPE arm received BEV after progression. Furthermore, although exact proportions are unclear, patients in the modelled subgroup also received other therapies after progression. The ERG is also concerned by the fact that the 'manufacturer's base case estimates a total incremental life-year gain of 0.863 for the BEV+CAPE arm and most of this (60.87%) accrues during PD. In view of the limitations of the RFSPT method, such gains may be overly optimistic.

The ERG carried out analysis of the original PD trial data to explore survival during this phase. The analyses (see Figure 5) show that there was little discernible difference in the cumulative hazard trends for the following three groups during PD (log-rank test: chi-square = 0.419, p=0.811):

- BEV+CAPE patients who did not cross (n=72,46.8% of the BEV+CAPE population)
- BEV+CAPE patients who crossed (n=82, 53.2% of the BEV+CAPE population)
- CAPE patients who crossed (n=44, 56.4% of the CAPE population)

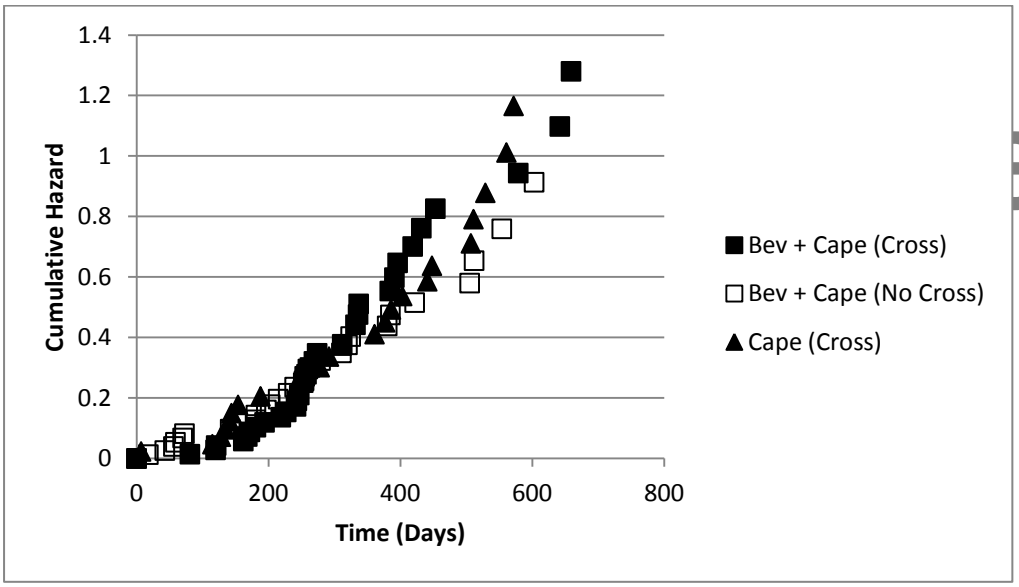

Figure 4 Kaplan-Meier analysis of PD from RIBBON-1 clinical trial for BEV+CAPE patients and CAPE patients who crossed

The log-rank test did suggest that there is a difference between CAPE patients who crossed and those who did not (chi-square= 4.458, p=0.035) but the cumulative HR plot (see Figure 5) shows that for patients surviving the progression event both trajectories are similar up until approximately 1 year, at which point there are only about six patients left in the CAPE no cross group.

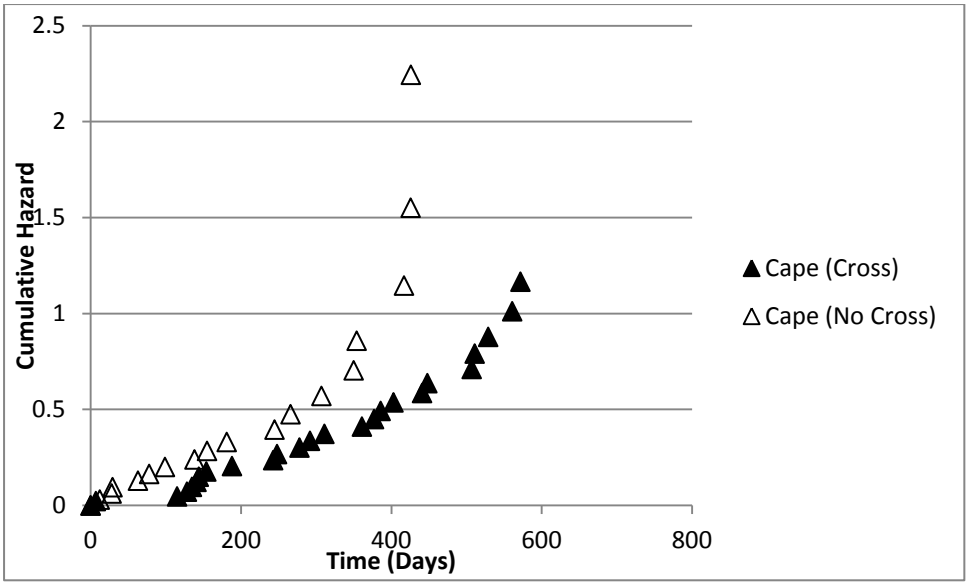

Figure 5 Kaplan-Meier analysis of PD from RIBBON-1 clinical trial by CAPE arm (crossed and uncrossed)

As a result of these analyses the ERG concluded that it would be appropriate to apply the same model to PD irrespective of trial arm or whether patients had, or had not, crossed. Two different models were developed, one based on all data and the second excluding data from patients in the CAPE arm who had not crossed. The KM plot (see Figure 6) using data from the whole population suggests that trial data can be used directly in the model, and, as the survival curve covers the entire range until all patients have died, there is no need for a parametric curve to be fitted.

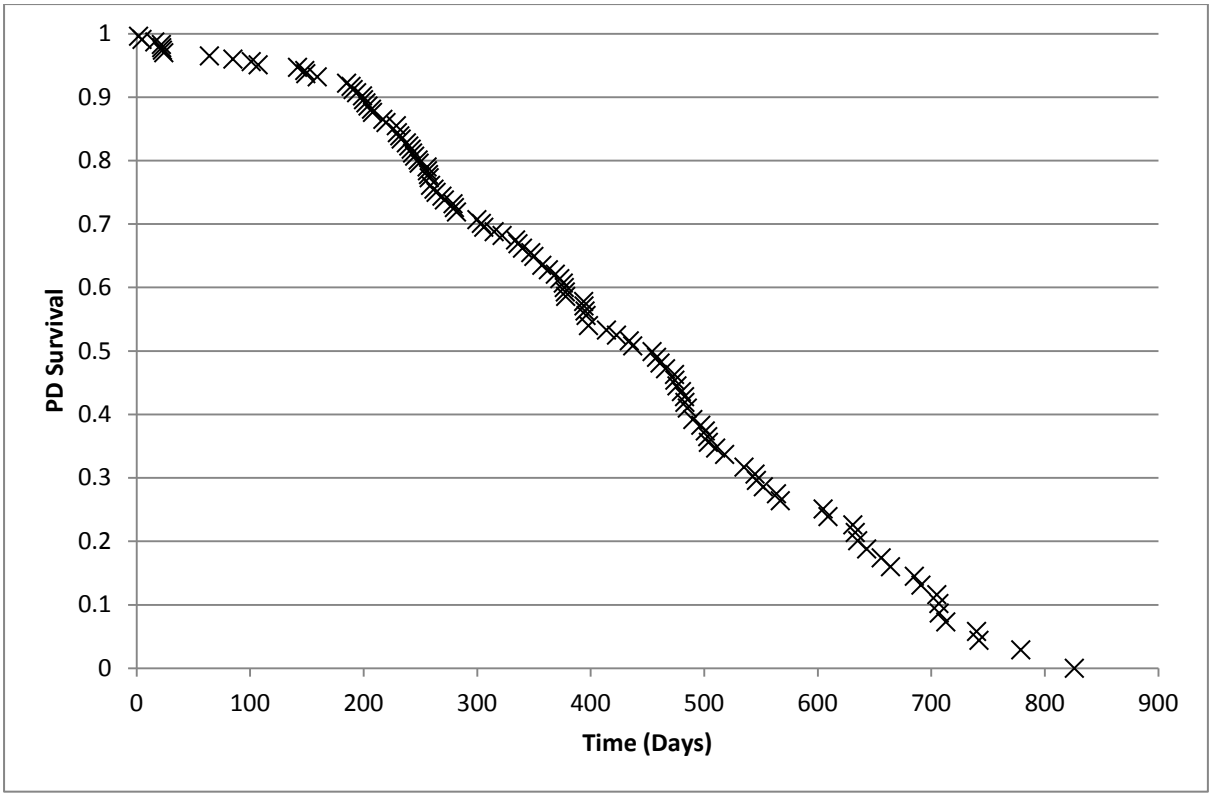

Figure 6 Kaplan-Meier analysis of PD from RIBBON-1 clinical trial all modelled patients

PD survival data for all BEV+CAPE patients and those CAPE patients who crossed over was modelled directly from trial data up to 15 months, with a linear extrapolation ($y = -0.0012x+1.0856$, $R^2=0.9919$) thereafter (see Figure 7).

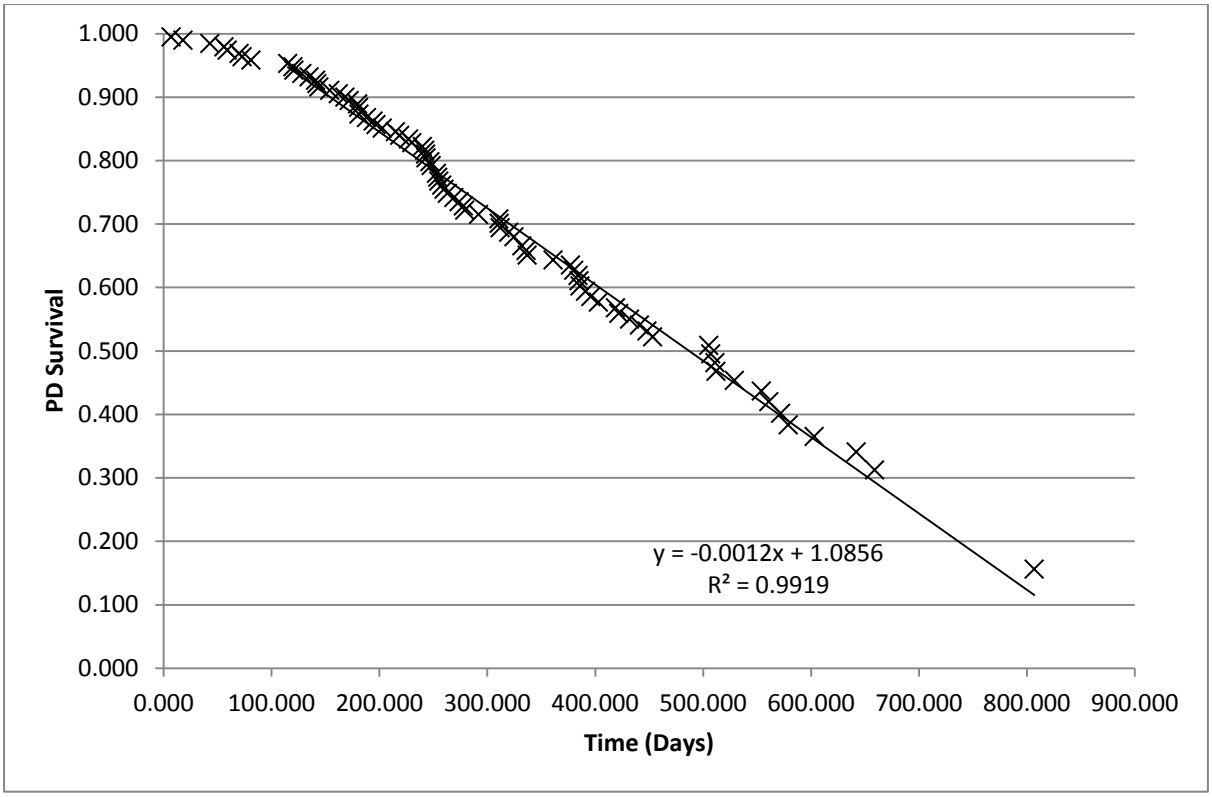

Figure 7 Kaplan-Meier analysis of PD from RIBBON-1 clinical trial all modelled patients except CAPE patients who did not cross

Using the PD survival model based on data from the whole population (excluding CAPE patients who did not cross) for both trial arms decreases the incremental life year gain by 0.625 (0.627) and decreases the QALY gain by 0.310 (0.311) compared with the base case, resulting in an increased ICER of £92,739 (£93,515) per QALY gained.

5.3.4 Cost estimation and parameter values

Active treatment costs

The ERG has re-estimated the costs of therapy based on the distribution of patient body weight and body surface area of a UK specific cohort of patients, rather than the use of simple average based on trial data. Overall these changes increase the drug costs in the BEV+CAPE arm by £2,966 per patient and the drug costs in the CAPE arm by £50 per patient. These adjustments result in a revised ICER that is £5,793 higher per QALY gained than the manufacturer's base case ICER.

Adverse event costs

In the submitted model, the costing of AEs seems to have been implemented without justification of assumptions. Some alternative choices could have been made by the authors of the manufacturer's

model, but the ERG is of the opinion that such changes would have only a very minor impact on the incremental cost and the estimated ICER, and so can be ignored.

Terminal care costs

The manufacturer's model does not include the costs of terminal care during the last 2 weeks of life, as specified in NICE guidelines[34]. This cost was estimated by uplifting costs reported by Remak et al[44] to 2009/10 prices using the Hospital & Community Health Services (HCHS) pay and prices index published by PSSRU. To test its impact, this cost was added into the submitted model and, as a result of the modest improvement in OS attributable to BEV, this modification produced a small discounted cost difference which reduces the incremental cost per patient by £53, and reduces the ICER by £105 per QALY gained.

5.3.5 Utility estimation and parameter values

The utility values used in the submitted model have been estimated using the statistical model detailed in a study by Lloyd et al.[37] This model features several factors including the rate of response to chemotherapy, and the exposure to a set of important AEs.

When using the Lloyd et al[37] model to estimate utility values the manufacturer has not included AE rates for the modelled subgroup. However, examination of the reported frequency of AEs indicates very low rates for the key events and the ERG is, therefore, satisfied that no adjustments for AE disutility are necessary.

The Lloyd et al model[37] reflects the views of a population with a mean age of 40.16 years and the age parameter in their model relates to the views of this surveyed population. However, the manufacturer misinterpreted this parameter as the age of the individual with mBC. The ERG has amended the age parameter in the utility model resulting in revised estimates of 0.7548 and 0.7432 for PFS in the BEV+CAPE and CAPE arms respectively, and a value of 0.4544 for PD. Using these revised figures (and correcting a typing mistake in the formula used for some months in the manufacturer's CAPE arm) results in a decrease in incremental utility gain from use of BEV by 9.4% and increases the base case ICER by £8,012 per QALY gained.

5.3.6 Sensitivity analysis - licensed dose of CAPE

The manufacturer reports that in the RIBBON-1[16] trial CAPE was administered at a dose of 1,000mg/m^2. The manufacturer points out that this differs slightly from the SPC[21] specified dose in which it is recommended that CAPE be given at a dose of 1,250mg/m^2. The ERG expert advisors have suggested that in practice CAPE tends to be administered at the lower dose due to the higher incidence of AEs observed at the higher dose. Comparing AE rates at the lower and higher doses is not straightforward and although a request was made for this information in the first clarification letter the manufacturer was not able to provide it.

The ERG notes that:

- At a dose of 1000mg/m^2 the impact of AEs on overall costs is negligible;
- Assuming that increasing the dose to 1250mg/m^2 has equal impact, in terms of increased incidence of AEs, on both treatment arms, then the modelled effect (in terms of cost rather than patient well being) should cancel itself out.

Bearing in mind these two factors the ERG found that changing the dose of CAPE to 1,250mg/m^2 results in a monthly cost of £398.55, an overall incremental increase in drug costs of £3,782 and an accompanying increase of £7,512 to the ICER estimate.

5.4 Conclusions of the cost-effectiveness section

The manufacturer's review of the published cost-effectiveness literature describing BEV+CAPE vs CAPE for previously untreated advanced breast cancer patients did not identify any relevant cost-effectiveness studies. The ERG is satisfied with the manufacturer's search strategy and is reasonably confident that the manufacturer did not miss any relevant published articles.

The manufacturer's reported base case ICER is £77,318 per QALY gained. It should be noted, however, that the modelled population is a subgroup of the population licensed to receive BEV+CAPE, namely those who have previously received a taxane in the adjuvant setting. The baseline characteristics of this subgroup indicate that they are younger and healthier than the whole licensed population; suggesting that the ICER per QALY gained for the whole licensed population may be somewhat higher than the value generated by the manufacturer's model.

The ERG made three relatively minor amendments/corrections to the manufacturer's model, the impact of which was to increase the manufacturer's ICER by between £105 and £8,011 per QALY gained. However, the ERG's re-estimation of PD survival, and as a consequence OS, suggests that the manufacturer's ICER may be optimistic.

6 IMPACT ON THE ICER OF ADDITIONAL ANALYSES UNDERTAKEN BY THE ERG

The alterations to the submitted economic model described above were implemented by the ERG to assess their influence on the incremental cost-effectiveness ratio for BEV+CAPE vs CAPE. In most cases the required amendments were relatively minor, but the introduction of the ERG's preferred approach to the estimation of survival in PD proved to be more substantial. This involved the assumption that survival during PD was equivalent irrespective of previous treatment and developing a common projection for both regimens, whilst maintaining the differing timings of entry into, and deaths whilst in, PD.

Table 36 shows the results of applying the sensitivity analysis related to using the licensed rather than the trial dose of CAPE and each of the ERG model amendments. Three of these lead to relatively minor alterations to the estimated ICER; with revised ICERs ranging between £77,213 and £85,329 per QALY gained. The new method of estimating survival during PD is much more important since it reduces the base case life-year (utility) gains by 27.66% (38.37%) with only a 15.56% reduction in incremental costs and leads to an ICER estimate of £170,057 per QALY gained. When all the relatively minor amendments are applied simultaneously, the final ERG ICER estimate increases to £91,625 per QALY gained; however, when the revised approach to modelling survival in PD is also included, the ERG's ICER increases to £207,850 per QALY gained.

The manufacturer's reported base case ICER is £77,318. Furthermore, the manufacturer's choice to base their model on a subgroup of patients licensed to receive BEV+CAPE suggests that, as acknowledged by the manufacturer (see section 3.1), the ICER for the whole licensed population is likely to be somewhat higher than the reported base case.

Table 36 Cost-effectiveness results following application of ERG model amendments

Scenario / model change	BEV +CAPE					CAPE					Incremental			
	Life-years	QALYs	Drug costs	Support-ive care costs	All costs	Life-years	QALYs	Drug costs	Support-ive care costs	All costs	Life-years	QALYs	Costs	ICER
SA: Use of CAPE at licensed dose	2.228	1.338	£39,568	£16,264	£55,900	1.365	0.835	£3,588	£9,606	£13,194	0.864	0.503	£42,706	£84,830
Base case	**2.228**	**1.338**	**£35,313**	**£16,264**	**£51,645**	**1.365**	**0.835**	**£3,115**	**£9,606**	**£12,721**	**0.864**	**0.503**	**£38,924**	**£77,318**
ERG drug costs	2.228	1.338	£38,280	£16,264	£54,612	1.365	0.835	£3,165	£9,606	£12,771	0.864	0.503	£41,841	£83,111
Add terminal care costs	2.228	1.338	£35,313	£16,264	£53,351	1.365	0.835	£3,115	£9,606	£14,479	0.864	0.503	£38,871	£77,213
ERG revised utility values	2.228	1.255	£35,313	£16,264	£51,645	1.365	0.799	£3,115	£9,606	£12,721	0.864	0.456	£38,924	£85,329
ERG revised PD survival estimates (whole modelled population)	1.868	1.160	£35,313	£12,792	£48,173	1.630	0.966	£3,115	£12,163	£15,278	0.239	0.193	£32,896	£170,057
ERG Changes to Drug cost, Terminal care & Utility values	2.228	1.255	£38,280	£16,264	£56,317	1.365	0.799	£3,165	£9,606	£14,529	0.864	0.456	£41,788	£91,607
All ERG changes	1.868	1.092	£38,280	£12,792	£52,867	1.630	0.919	£3,165	£12,163	£17,070	0.239	0.172	£35,797	£207,850

Base-case: manufacturer

- The manufacturer reports an ICER of £77,318 per QALY gained for the comparison of BEV+CAPE vs CAPE as a first-line therapy for patients with untreated HER2-ve mBC who have previously received a taxane in the adjuvant setting (and have most likely also received an anthracycline in the adjuvant setting).

- Results of the PSA conducted by the manufacturer suggest that, based on the assumptions made and the evidence available, BEV+CAPE is not a cost-effective treatment compared with CAPE at a willingness to pay of £30,000 or £50,000 per QALY gained in any circumstances (0% probability).

Base case: ERG

- The ERG made three comparatively minor amendments/corrections to the manufacturer's model and these included modifications to drug costs, the addition of terminal care costs, and use of ERG revised utility values. Individually, these resulted in only relatively small changes to the manufacturer's base-case ICER (range: £77,213-£85,329). The ERG made one major amendment to the model (revised PD survival estimates) which resulted in an ICER of £170,057 per QALY gained.

- When all of the ERG's changes are incorporated together into the submitted model, the ICER is estimated to be £207,850 per QALY gained,

7 OVERALL CONCLUSIONS

The clinical effectiveness evidence is derived from a single, manufacturer supported RCT (RIBBON-1[17]). This trial, which was well conducted, compared BEV+CAPE to CAPE. Given anthracyclines or taxanes were a possible treatment options prior to randomisation, it seems reasonable to assume these patients would therefore be considered unsuitable for an anthracycline or a taxane. For all these patients (the ITT population), it can be assumed that a taxane or anthracycline were not considered appropriate. Thus this population can be considered to be a group of patients for whom BEV+CAPE is licensed. Compared with the CAPE arm, a statistically significant increase in PFS (2.9 months) but not OS was reported for BEV+CAPE. Despite there being a greater proportion of AEs reported for patients in the BEV+CAPE arm, no new safety concerns were identified.

The cost-effectiveness analysis is based on a subgroup of patients from RIBBON-1,[16] namely a group of patient who had previously received a taxane (and by implication, it is assumed an anthracycline) for adjuvant treatment. In this subgroup, significant improvements in PFS of 4.5 months and OS of 7.9 months were reported for patients who received BEV+CAPE compared with CAPE. However, this was just one of a number of subgroups for which analyses of PFS and OS were conducted. No statistical adjustments were performed to control for multiple significance testing and so these findings must be treated with caution. In addition, baseline characteristics suggest that this is a younger and healthier group of patients than the ITT population.

The ERG made three relatively minor alterations/corrections to the economic model, namely: a recalculation of drug costs based on the distribution of patient body weight and body surface area of a UK specific cohort of patients, rather than a simple average based on trial data; addition of the cost of terminal care; correction to the calculation of utility values. In addition, because of concerns surrounding the *post-hoc* use of the RPSFT model to adjust for OS, the ERG also adjusted the economic model so that survival for both patient groups was equivalent during the PD phase.

The manufacturer's reported base case ICER is £77,318 per QALY gained. Implementing all three relatively minor changes increased the ICER to £91,607 per QALY gained. Amending the model so that survival for both patient groups was equivalent during the PD phase resulted in an ICER of £170,057 per QALY gained. Combining all four of the ERG changes results in an ICER of £207,850 per QALY gained.

The manufacturer's base case ICER per QALY gained cannot be considered to be generalisable to the whole licensed population. This is because analyses of the RIBBON-1[16] trial data show that the subgroup on which this ICER is based is a more stringent population than the licensed population ho, at baseline, appear younger and healthier than the ITT population. Further, there are caveats around

the subgroup PFS and OS findings. Both the manufacturer and the ERG believe that the ICER per QALY gained for the licensed population would be higher than that for the modelled subgroup.

8 REFERENCES

1. Cancer Research UK. Breast cancer - UK incidence statistics. 2011 [updated 18/07/2011]; Available from: http://info.cancerresearchuk.org/cancerstats/types/breast/incidence/ (Accessed: 03/02/2012).
2. Burstein H, Harris J, Morrow M. Malignant tumors of the breast. In: DeVita V, Lawrence T, Rosenberg S, editors. Cancer: Principles and Practice of Oncology. 8th ed. Philadelphia, PA, USA: Lippincott Williams & Wilkins; 2008. p. 1606-54.
3. National Institute for Health and Clinical Excellence (NICE). Advanced breast cancer: diagnosis and treatment. NICE clinical guideline 81. 2009; Available from: http://www.nice.org.uk/Guidance/CG81 (Accessed: 10/02/2010).
4. O'Shaughnessy J. Extending Survival with Chemotherapy in Metastatic Breast Cancer. Oncologist. 2005; 10(suppl 3):20-9.
5. Stewart B, Kleihues P. World Cancer Report. Lyon: IARC Press; 2003; Available from: http://www.iarc.fr/en/publications/pdfs-online/wcr/2003/index.php (Accessed: 03/02/2012).
6. Anderson W, Jatoi I, Devesa S. Distinct breast cancer incidence and prognostic patterns in the NCI's SEER program: suggesting a possible link between etiology and outcome. Breast Cancer Res Treat. 2005; 90(2):127-37.
7. Muss HB, Smith LR, Cooper MR. Tamoxifen rechallenge: response to tamoxifen following relapse after adjuvant chemohormonal therapy for breast cancer. J Clin Oncol. 1987; 5(10):1556-8.
8. Ravdin PM, Green S, Dorr TM, McGuire WL, Fabian C, Pugh RP, *et al.* Prognostic significance of progesterone receptor levels in estrogen receptor-positive patients with metastatic breast cancer treated with tamoxifen: results of a prospective Southwest Oncology Group study. J Clin Oncol. 1992; 10(8):1284-91.
9. Jones SE. Metastatic breast cancer: The treatment challenge. Clin Breast Cancer. 2008; 8(3):224-33.
10. Wood W, Muss H, Solin L, Olopade O. Malignant tumors of the breast. In: DeVita V, Lawrence T, Rosenberg S, editors. Cancer: Principles and Practice of Oncology. 7th ed. Philadelphia, PA, USA: Lippincott Williams & Wilkins; 2005.
11. Helgeson VS, Tomich PL. Surviving cancer: A comparison of 5-year disease-free breast cancer survivors with healthy women. Psychooncology. 2005; 14(4):307-17.
12. Herschbach P, Keller M, Knight L, Brandl T, Huber B, Henrich G, *et al.* Psychological problems of cancer patients: a cancer distress screening with a cancer-specific questionnaire. Br J Cancer. 2004; 91(3):504-11.
13. Office for National Statistics. Population Estimates for UK, England and Wales, Scotland and Northern Ireland, Population Estimates Timeseries 1971 to Current Year. Newport: Office for National Statistics; 2011; Available from: http://www.ons.gov.uk/ons/rel/pop-estimate/population-estimates-for-uk--england-and-wales--scotland-and-northern-ireland/population-estimates-timeseries-1971-to-current-year/index.html (Accessed: 03/02/2012).
14. Scarborough P, Bhatnagar P, Wickramasinghe K, Smolina K, Mitchell C, Rayner M. Coronary heart disease statistics: 2010 edition. London: British Heart Foundation; 2011; Available from: http://www.bhf.org.uk/publications/view-publication.aspx?ps=1001546 (Accessed: 03/02/2012).
15. Dent R, Trudeau M, Pritchard KI, Hanna WM, Kahn HK, Sawka CA, *et al.* Triple-Negative Breast Cancer: Clinical Features and Patterns of Recurrence. Clin Cancer Res. 2007; 13(15):4429-34.
16. Robert NJ, Diéras V, Glaspy J, Brufsky AM, Bondarenko I, Lipatov ON, *et al.* RIBBON-1: Randomized, double-blind, placebo-controlled, phase III trial of chemotherapy with or without bevacizumab for first-line treatment of human epidermal growth factor receptor 2-negative, locally recurrent or metastatic breast cancer. J Clin Oncol. 2011; 29(10):1252-60.

17. Saif MW, Katirtzoglou NA, Syrigos KN. Capecitabine: An overview of the side effects and their management. Anti-Cancer Drugs. 2008; 19(5):447-64.
18. Sledge G, Miller K, Moisa C, Gradishar W. Safety and efficacy of capecitabine (C) plus bevacizumab (B) as first-line in metastatic breast cancer. J Clin Oncol (Meeting Abstracts). 2007; 25(18 suppl):1013.
19. Lang I, Inbar MJ, Greil R, Steger GG, Beslija S, Kahan Z, *et al.* Bevacizumab (Bev) combined with either capecitabine (X) or weekly paclitaxel (Pac) as first-line chemotherapy (CT) for HER2-negative, locally recurrent or metastatic breast cancer (LR/MBC): Preliminary safety data from the CECOG phase III TURANDOT trial. J Clin Oncol (Meeting Abstracts). 2010; 28(15 suppl):1126.
20. European Medicines Agency. Avastin: Summary of Product Characteristics. Available from: http://www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/human/medicines/000582/human_med_000663.jsp&mid=WC0b01ac058001d124 (Accessed: 03/02/2012).
21. European Medicnes Agency. Xeloda: Summary of Product Characteristics. Available from: http://www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/human/medicines/000316/human_med_001157.jsp&mid=WC0b01ac058001d124 (Accessed: 03/02/2012).
22. Electronic Medicine Compendium. Vinorelbine 10 mg/ml Concentrate for solution for infusion: Summary of Product Characteristics. Available from: http://www.medicines.org.uk/emc/medicine/19453/SPC/vinorelbine%2010%20mg~ml%20concentrate%20for%20solution%20for%20infusion/#PRODUCTINFO (Accessed 03/02/2012).
23. Pajk B, Cufer T, Canney P, Ellis P, Cameron D, Blot E, *et al.* Anti-tumor activity of capecitabine and vinorelbine in patients with anthracycline- and taxane-pretreated metastatic breast cancer: Findings from the EORTC 10001 randomized phase II trial. Breast. 2008; 17(2):180-5.
24. Oostendorp LJM, Stalmeier PFM, Donders ART, van der Graaf WTA, Ottevanger PB. Efficacy and safety of palliative chemotherapy for patients with advanced breast cancer pretreated with anthracyclines and taxanes: a systematic review. The Lancet Oncology. 2011; 12(11):1053-61.
25. Chan A, Verrill M. Capecitabine and vinorelbine in metastatic breast cancer. Eur J Cancer. 2009; 45(13):2253-65.
26. Centre for Reviews and Dissemination (CRD). CRD's guidance for undertaking reviews in healthcare: Systematic Reviews (3rd Edition). York: CRD, University of York 2008.
27. Brufsky A, Ponomarova O, Tjulandin S. Influence of disease free interval on the efficacy of capecitabine− bevacizumab for HER2−negative metastatic breast cancer (MBC) in the RIBBON−1 trial. Eur J Cancer Suppl. 2010; 8(3):201 (Abstract 485).
28. Therasse P, Arbuck SG, Eisenhauer EA, Wanders J, Kaplan RS, Rubinstein L, *et al.* New Guidelines to Evaluate the Response to Treatment in Solid Tumors. J Natl Cancer Inst. 2000; 92(3):205-16.
29. Robins JM, Tsiatis AA. Correcting for non-compliance in randomized trials using rank preserving structural failure time models. Comm Statist Theory. 1991; 20(8):2609-31.
30. European Medicines Agency. Assessment Report For Avastin (bevacizumab). Procedure No.: EMEA/H/C/000582/II/0033. London: European Medicines Agency; 2011; Available from: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Assessment_Report_-_Variation/human/000582/WC500108311.pdf (Accessed 03/02/2012).
31. National Institute for Health and Clinical Excellence. Pazopanib for the first-line treatment of advanced renal cell carcinoma. NICE technology appraisal guidance 215. London: National Institute for Health and Clinical Excellence; 2011; Available from: http://guidance.nice.org.uk/TA215 (Accessed: 09/02/2012).
32. Delea TE, Duh NS, Wei LJ, Robins JM. Methods to Adjust for Selective Crossover in Survival Analysis In Assessments of Cost-Effectiveness of Cancer Therapies. ISPOR workshops: session II. Monday, May 23, 2011. Available from: http://www.analysisgroup.com/uploadedFiles/News_and_Events/News/ISPOR2011_Selected_Crossover_Cost-Effectiveness_Cancer_Therapies.pdf (Accessed: 09/02/2012).

33. Robins J, Finkelstein D. Correcting for noncompliance and dependent censoring in an AIDS clinical trial with Inverse Probability of Censoring Weighted (IPCW) log-rank tests. Biometrics. 2000; 56 (3):779-88.
34. National Institute for Health and Clinical Excellence. Guide to the methods of technology appraisal. 2008; Available from: http://www.nice.org.uk/media/B52/A7/TAMethodsGuideUpdatedJune2008.pdf (Accessed: 10/02/2012).
35. Dranitsaris G, Leung P, Mather J, A O. Cost-utility analysis of second-line hormonal therapy in advanced breast cancer: A comnparison of two aromatase inhibitors to megestrol acetate. Anti-Cancer Drugs. 2000; 11(7):141-54.
36. Leung P, Tannock I, Oza A, Puodziunas A, Dranitsaris G. Cost-utility analysis of chemotherapy using paclitaxel, docetaxel, or vinorelbine for patients with anthracycline-resistant breast cancer. J Clin Oncol. 1999; 17(10):3082-90.
37. Lloyd A, Nafees B, Narewska J, Dewilde S, Watkins J. Health state utilities for metastatic breast cancer. Br J Cancer. 2006; 95(6):683-90.
38. Fleeman N, Bagust A, Boland A, Dickson R, Dundar Y, Moonan M, *et al*. Lapatanib and trastuzumab in combination with an aromatase inhibitor for the first-line treatment of metastatic hormone receptor positive breast cancer which over-expresses HER2. Health Technol Assess. 2011; 15(2).
39. British National Formulary. BNF 62. 2011; Available from: http://www.bnf.org/bnf/ (Accessed:27/01/2012).
40. Department of Health. NHS Reference Costs (2009-2010). 2011: Available from: http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_111591 (Accessed: 21/02/2012).
41. Curtis L. Unit Costs of Health and Social Care 2010. 2011; Available from: http://www.pssru.ac.uk/uc/uc2010contents.htm (Accessed: 21/02/2012).
42. Peasgood T, Ward S, Brazier J. Health-state utility values in breast cancer. Expert Rev Pharmacoecon Outcomes Res. 2010; 10(5):553-66.
43. Drummond M, Jefferson T. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. BMJ. 1996; 313:275-83.
44. Remak E, Brazil L. Cost of managing women presenting with stage IV breast cancer in the United Kingdom. Br J Cancer. 2004; 91(1):77-83.
45. Miller KD. E2100: A phase III trial of paclitaxel versus paclitaxel/bevacizumab for metastatic breast cancer. Clinical Breast Cancer. 2003; 3(6):421-2.
46. Miles DW, Chan A, Dirix LY, Cortés J, Pivot X, Tomczak P, *et al*. Phase III Study of Bevacizumab Plus Docetaxel Compared With Placebo Plus Docetaxel for the First-Line Treatment of Human Epidermal Growth Factor Receptor 2–Negative Metastatic Breast Cancer. J Clin Oncol. 2010; 28(20):3239-47.

APPENDIX 1: BEVACIZUMAB SAFETY CONCERNS

A number of safety concerns have been raised with regard to BEV from systematic reviews (and meta-analyses) and single studies. It is partly for this reason that in the United States, the Food and Drug Administration (FDA) has revoked the licence for any indication for breast cancer. The other reason why the FDA revoked the licence was because it did not believe that the clinical benefit for the treatment of mBC from three RCTs outweighed these perceived harms.

BEV was originally given accelerated approval in the United States for mBC after the E2100 trial[45] reported a statistically significant median improvement in PFS of 5.5 months for patients who received BEV and paclitaxel over patients who received paclitaxel alone (in this trial, there was no improvement in OS). A condition of accelerated approval was that the manufacturer would submit confirmatory studies that confirmed that BEV improved patient outcomes over standard chemotherapy as seen in E2100. In the FDA's opinion, two subsequent trials (RIBBON-1[16] and AVADO[46]) which reported PFS gains of no more than 3 months (and no OS or HRQoL benefits) failed to do this while an increase in SAEs for patients who received BEV over standard chemotherapy were observed.

It should however be noted that BEV is still indicated in the United States as a treatment for colon, lung, kidney, and brain cancers. It should also be emphasised that the view that the harms outweigh the benefits is clearly not shared by the European Union who have approved BEV in combination with a taxane and in combination with CAPE for the treatment of mBC.

Many of the safety concerns relate to cardiac and vascular disorders. The ERG notes that the RIBBON-1[16] trial, from which the evidence presented in the MS is derived, included the following exclusion criteria:

- Blood pressure >150/100 mmHg
- Unstable angina
- New York Heart Association (NYHA) Grade II or greater congestive heart failure
- History of myocardial infarction (within last 6 months)
- History of stroke or transient ischemic attack (within last 6 months)
- Clinically significant peripheral vascular disease

None of these conditions are listed as contraindications in the SPC.[20] The only contraindications listed are:

- Hypersensitivity to the active substance or to any of the excipients.
- Hypersensitivity to Chinese hamster ovary cell products or other recombinant human or humanised antibodies.
- Pregnancy.

Furthermore, while the SPC[20] notes that caution should be exercised when treating patients with clinically significant CVD such as congestive heart failure (CHF), in studies to date, most of the patients who experienced CHF had received previous treatment with anthracyclines, prior radiotherapy to the left chest wall or other risk factors for CHF were present.

Finally, the ERG notes that the majority of studies and systematic reviews published to date that have highlighted safety concerns have included patients who received various chemotherapy regimens (e.g. taxanes) in combination with BEV and for different lines of treatment. In addition, not all patients included in these reviews had mBC, many included patients with different types of cancer. As this current appraisal is concerned with BEV+CAPE for the first-line treatment of mBC, many of these previous safety concerns may therefore not be relevant. For this reason, the ERG has not conducted a separate search to identify studies and systematic reviews examining AEs in patients treated with BEV, it being assumed that relevant studies examining the safety of BEV+CAPE would be identified from the search strategy employed for BEV+CAPE (see section 4.1.1). The ERG notes from Appendix 8 of the MS, the manufacturer also states: "Safety was a secondary endpoint and was covered in results from the RIBBON-1 trial [which compared BEV+CAPE to CAPE]."

APPENDIX 2: PREVALENCE OF CORONARY HEART DISEASE AND CARDIOVASCULAR DISEASE

The manufacturer assumes that 4% of patients would not be eligible for BEV+CAPE because of a history of CHD. This is based on the prevalence rate for CHD for women of all ages in England. As can be seen from Table 38, the rates vary by age.

The ERG notes that in the RIBBON-1[16] trial on which the clinical evidence was based, 94.1% of patients were aged ≥40 years and 71.9% were aged ≥50 years (mean age =56.8 years).

In the prior taxane subgroup of RIBBON-1[16] on which the economic evidence was based, 92.6% were aged ≥40 years and 63.4% were aged ≥50 years (mean age = 53.4 years).

The youngest patient in RIBBON-1[16] was aged 23 and the oldest was 91 (the range was 23 to 84 in the subgroup).

In clinical practice, the ERG notes that an older group of patients are likely to be treated than in RIBBON-1.[16]

Because the prevalence rate was weighted for non-response and only the unweighted base were provided, it was impossible to calculate the prevalence rates for specific age groups of all women aged ≥45, ≥50 or ≥55 years which may have been more informative.

Table 37 Prevalence of CHD and CVD in women in England, 2006

Age (years)	CHD (%)	CVD (%)	Unweighted base
16-24	0.1	4.5	794
25-34	0.1	5.7	1,148
35-44	0.3	7.8	1,494
45-54	1.3	10.3	1,279
55-64	3.5	15.2	1,269
65-74	10.0	21.2	470
75+	19.3	36.9	471
All ages	4.0	13.0	6,925

Note: Prevalence rates are weighted for non-response. Respondents were prompted to recall whether they had ever been diagnosed with each of the conditions by a doctor.
All data taken from Scarborough et al 2011[14]

As noted in Appendix 1, the SPC[20] notes that caution should be exercised when treating patients with clinically significant CVD. For this reason, the prevalence rate for CVD may be more appropriate than CHD.

APPENDIX 3: ENDPOINTS AND DEFINITIONS EMPLOYED IN RIBBON-1

The endpoints employed in the RIBBON-1[16] trial and their definition are summarised in Table 38.

Table 38 Endpoints and definitions employed in RIBBON-1

Endpoint	Definition
PFS	PFS based on investigator assessment was considered the primary efficacy endpoint. PFS was defined as the time from randomisation to first disease progression or death due to any cause, whichever occurs first. Data for subjects without disease progression or death was censored at the time of the last tumour assessment (or, if no tumour assessments were performed after the baseline visit, at the time of randomisation plus 1 day). The primary efficacy analysis population was the ITT population, defined as all subjects who were randomised, regardless of whether they received any study drug or complete the full course of treatment. A subject was considered to be randomly assigned to the treatment when the study site was notified of the subject's treatment arm assignment by the interactive voice response system. Subjects were grouped according to the treatment assignment at randomisation. PFS based on IRC-reviewed data was considered a secondary efficacy endpoint and served as a sensitivity analysis to support the investigator-determined assessment. Radiographic data were sent from investigative sites to the IRC. Disease progression was assessed by IRC according to RECIST. PFS is defined as the time from randomisation to first disease progression, as determined by IRF, or death due to any cause, whichever occurs first. Data for subjects without IRC-determined disease progression or death were censored at the time of the last tumour assessment (or, if no tumour assessments were performed after the baseline visit, at the time of randomisation plus 1 day). Data for subjects who received excluded therapy prior to IRC-determined disease progression were also censored at the time of the last tumour assessment prior to the initiation of the excluded therapy.
ORR	Objective response is defined as a complete or partial response determined on two consecutive assessments ≥4 weeks apart. Objective response rate is the proportion of subjects who have objective response. The primary analysis for objective response rate was performed using only subjects with measurable disease at baseline. The supportive analysis included the ITT population. Subjects without a post-baseline tumour assessment were considered non-responders.
Duration of objective response	For the subset of subjects who achieved objective responses , the time from the first tumour assessment that supports the subject's objective response to the time of disease progression, or death due to any cause, whichever occurs first.
OS	The time from randomisation until death from any cause. For subjects who had not died at the time of analysis, duration of survival was censored as of the date the subject was last known to be alive.
One-year survival rate	The percentage of subjects who are still alive at one year after the randomisation.
Safety	Protocol-defined selected AEs included a number associated with BEV in previous studies. The primary safety analyses were based on all randomised subjects who received any study treatment, defined as at least one full or partial dose of either study treatment or chemotherapy. This population was referred to as the safety population. Subjects were analysed as randomised. Subjects who received chemotherapy other than the cohort they were initially enrolled to were analysed based on their initial chemotherapy assignments for the safety analyses. Verbatim descriptions of adverse events were mapped to MedDRA thesaurus terms and graded according to the NCI-CTC for AEs, Version 3.0. All AEs, SAEs, and AEs leading to death and study treatment discontinuation were summarised by treatment arm and NCI-CTC grade. AEs were also summarised by age, race, and geographic region. For events of varying severity, the highest grade was used in summaries. Additionally, AEs that occurred after disease progression but prior to the optional post-progression phase were summarised by the treatment arm for the subset of subjects who continued on the blinded treatment phase after disease progression.

APPENDIX 4: NCI-CTC GRADES FOR AES REPORTED IN RIBBON-1

Blinded phase

For hypertension, according to Table 60 of the CSR, both (1.0%) AEs in the CAPE arm were grade 3 and 38 (9.4%) in BEV+CAPE were grade 3 and 3 (0.7%) were grade 2. Regarding proteinura, all nine (2.2%) AEs in the BEV+CAPE arm were grade 3 and for sensory neuropathy, all 12 (3.0%) AEs in the BEV+CAPE arm were grade 3, the AE in the CAPE arm (0.5%) was grade 4.

With regard to all cardiac disorders, these were mostly grade 5 (five [1.2%]), grade 4 (five [1.2%]) or grade 3 (three [0.7%]) in the BEV+CAPE arm; two (1.0%) were grade 4 and two (1.0%) were grade 3 in the CAPE arm. It should be noted that within these cardiac disorders are included AEs labelled as left ventricular systolic dysfunction, which were reported by 1.5% in the BEV+CAPE arm and 0.5% in the CAPE arm.

The ERG also notes there were 19 (4.7%) patients who experienced infections and infestations in the BEV+CAPE arm and 5 (2.5%) in the CAPE arm. Most of these AEs were grade 3 or lower although two (0.5%) were grade 5 and two (0.5%) were grade 4 in the BEV+CAPE arm.

Open-label phase

Of the 249 patients, 32.9% reported an AE, most (17.7%) being grade 3 (4.4% were grade 2 or lower, 7.6% were grade 4 and 3.2% were grade 5). The only AEs occurring at a ≥5% incidence were neutropenia (17 [6.8%]; nine [3.6%] were grade 4 and seven [2.8%] were grade 3). In addition, the ERG notes there were 12 (4.8%) patients who experienced hypertension (one [0.4%] grade 4, 11 [4.4%] grade 3), seven (2.8%) patients who experienced febrile neutropenia (of which two [0.8%] were grade 4 and four [1.6%] were grade 3), six patients who experienced peripheral sensory neuropathy (all grade 3), three (1.2%) patients who experienced proteinura (all grade 3) and two (0.8%) patients who experienced vomiting (of which one [0.4%] was grade 3).

In addition to the AEs 'of special interest', from Table 14.3/47 of the CSR, the ERG notes there were three (1.2%) patients who experienced cardiac AEs (one was grade 5 [left ventricular dysfunction], one was grade 3 and one was grade 2). There were also 16 (6.4%) vascular disorders (of which one [0.4%] was grade 4 and 14 [5.6%] were grade 3). Regarding infections and infestations, 14 [5.6%] patients experienced these AEs in the open-label phase, most again being grade 3 (14 [5.6%] with one [0.4%] being grade 4).

APPENDIX 5: COMPARISON OF AES ACROSS STUDIES OF CAPECITABINE

Because AEs commonly associated with CAPE that were reported in RIBBON-1[16] appeared to be less frequent than expected, the ERG conducted a quick comparison with a recently identified systematic review,[24] the only known head-to-head RCT of CAPE with VIN (EORTC 10001[23]) and a single-arm phase II study of BEV+CAPE (XCALiBr, presented only as an abstract[18]). The findings are summarised in Table 39 where it appears that AEs previously associated with CAPE were much lower in both the BEV+CAPE and CAPE arm in RIBBON-1.[16] However, the majority of the evidence is derived from studies of different lines of treatment and at higher doses of CAPE than in RIBBON-1[16] and XCALiBr[18] which were studies of first-line treatment at a CAPE dose of 1000mg/m^2

Table 39 AEs (%) commonly associated with CAPE and VIN compared across selected studies and a systematic review

AE/ Grade	Systematic review[24]		EORTC 10001[23]		RIBBON-1[16]			XCALiB[18]
	CAPE	VIN	CAPE (n=23)	VIN (n=24)	CAPE (n=201)	BEV+ CAPE (n=404)	Open-label (n=249)	BEV+ CAPE (n=103)
Nausea	-	-	-	-	0	<1	0	-
3/4	3	2	4	0	0	<1	0	-
1/2	-	-	39	67	0	<1	0	-
SAE	-	-	-	-	0	<1	-	-
Discontinued	-	-	-	-	0	0	-	-
Vomiting	-	-	-	-	<2	<1	<1	-
3/4	3	2	4	0	1	<1	<1	-
1/2	-	-	22	33	<1	<1	<1	-
SAE	-	-	-	-	<2	<1	-	-
Discontinued	-	-	-	-	0	<1	-	-
Diarrhoea	-	-	-	-	2	<2	0	-
3/4	10	1	0	0	2	<2	0	-
1/2	-	-	48	21	0	0	0	-
SAE	-	-	-	-	1	<2	-	-
Discontinued	-	-	-	-		<2	-	-
Fatigue	-	-	-	-	0	<1	<1	-
3/4	5	13	4	13	0	0	0	-
1/2	-	-	61	50	0	<1	<1	-
SAE	-	-	-	-	0	0	-	-
Discontinued	-	-	-	-	0	0	-	-
hand-foot syndrome [a]	-	-	-	-	2	1	<1	-
3/4	16	0	4	0	<2	1	0	13
1/2	-	-	52	0	<1	0	<1	-
SAE	-	-	-	-	0	<1	-	-
Discontinued	-	-	-	-	<1	<1	-	-
Sensory neuropathy [b]	-	-	-	-	<1	3	<3	-
3/4	-	-	-	-	<1	3	<3	-
1/2	-	-	-	-	0	0	0	-
SAE	-	-	-	-	0	0	-	-
Discontinued	-	-	-	-	0	0	-	-

[a] Also referred to as palmar-plantar erythrodysesthesia syndrome

[b] Data on sensory neuropathy was included because it was speculated in the report by the EMA[30] that palmar-plantar erythrodysesthesia AEs may have been classified as sensory neuropathy in RIBBON-1[16]