STA REPORT

LIVERPOOL REVIEWS AND IMPLEMENTATION GROUP (LRiG)

Everolimus in combination with an aromatase inhibitor for the treatment of advanced or metastatic HER2 negative, oestrogen receptor positive breast cancer after prior endocrine therapy

This report was commissioned by the NIHR HTA Programme as project number 08/206/01

Completed 1 February 2013

UNIVERSITY OF LIVERPOOL

LIVERPOOL REVIEWS AND IMPLEMENTATION GROUP

A MEMBER OF THE RUSSELL GROUP

Title: Everolimus in combination with an aromatase inhibitor for the treatment of breast cancer after prior endocrine therapy

Produced by: Liverpool Reviews and Implementation Group

Authors: Nigel Fleeman, Research Fellow, LRiG, University of Liverpool

Adrian Bagust, Professor (Health Economic Modelling), LRiG, University of Liverpool

Sophie Beale, Research Associate (Decision Analysis), LRiG, University of Liverpool

Michaela Blundell, Research Associate (Medical Statistics), LRiG, University of Liverpool

Kerry Dwan, Research Associate (Medical Statistics), LRiG, University of Liverpool

Gerlinde Pilkington, Research Assistant, LRiG, University of Liverpool

Chris Proudlove, Director, North West Medicines Information Centre, Pharmacy Practice Unit, Liverpool

Yenal Dundar Research Fellow, LRiG, University of Liverpool

Fabio Vecchio, Research Assistant, LRiG, University of Liverpool

Nicky Thorp, Consultant in Clinical Oncology, Clatterbridge Cancer Centre, Wirral

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: 1 February 2013

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 11/137/01

Declared competing interests of the authors: None

Acknowledgements:
Thanks to James Mahon and Dr Susan O'Reilly who read and commented on a draft version of this report.

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

This report should be referenced as follows: Fleeman N, Bagust A, Beale S, Blundell M, Dwan K, Pilkington G, Proudlove C, Dundar Y, Vecchio F, Thorp N. Everolimus in combination with an aromatase inhibitor for the treatment of breast cancer after prior endocrine therapy: A Single Technology Appraisal. LRiG, The University of Liverpool, 2013

Contributions of authors:

Nigel Fleeman	Project lead, drafted the clinical results section and supervised the production of the final report
Adrian Bagust	Critical appraisal of the economic model
Sophie Beale	Critical appraisal of the economic evidence
Michaela Blundell	Critical appraisal of clinical statistical approach
Kerry Dwan	Critical appraisal of clinical statistical approach
Gerlinde Pilkington	Summary and critical appraisal of the clinical evidence
Chris Proudlove	Critical appraisal of the manufacturer's submission
Yenal Dundar	Cross checking of manufacturer's search strategies
Fabio Vecchio	Cross checking clinical sections of the report
Nicky Thorp	Critical appraisal of the clinical sections of the manufacturer's submission

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

Table of contents

Abbreviations

AE	Adverse event
AI	Aromatase inhibitor
BNF	British National Formulary
BOLERO-2	Breast cancer trials of OraL EveROlimus-2
CBR	Clinical benefit rate
CEAC	Cost-effectiveness acceptability curve
CHMP	Committee for Medicinal Products for Human Use
CI	Confidence interval
CIC	Commercial in confidence (redacted in this ERG report as follows: [redacted])
CONFIRM	COmparisoN of FaslodexTM In Recurrent or Metastatic breast cancer
CR	Complete response
ECOG	Eastern Cooperative Oncology Group
EFECT	Evaluation of Faslodex vs Exemestane Clinical Trial
EMA	European Medicines Agency
EPAR	European Public Assessment Report
ERG	Evidence Review Group
FDA	Food and Drug Administration
HER2(+/-)	Human epidermal growth factor receptor 2 (positive/negative)
HR(+/-)	Hormone Receptor (positive/negative)
HRQoL	Health-related quality of life
ICER	Incremental cost-effectiveness ratio
ITT	Intention to treat
KM	Kaplan-Meier
LYG	Life year gained
MS	Manufacturer's submission
NICE	National Institute for Health and Clinical Excellence
NSAI	Non-steroidal aromatase inhibitor
ORR	Overall response rate
OS	Overall survival
PDT	Post-discontinuation therapy
PgR(+/-)	Progesterone receptor (positive/negative)
PR	Partial response
PFS	Progression-free survival
PPS	Post-progression survival
PS	Performance status
PSA	Probabilistic sensitivity analysis
PSS	Personal Social Services
QALY	Quality adjusted life year
RCT	Randomised controlled trial
SD	Stable disease
SoFEA	Study Of Faslodex vs Exemestane with/without Arimidex
STA	Single Technology Appraisal
TAMRAD	TAMoxifen and RAD001-everolimus
TTP	Time to progression
vs	Versus

1 SUMMARY

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

In the scope issued by the National Institute for Health and Clinical Excellence (NICE), the title for this single technology appraisal (STA) is 'Everolimus in combination with an aromatase inhibitor for the treatment of breast cancer after prior endocrine therapy.' More specifically, the marketing authorisation by the European Commission in July 2012 is for the treatment of hormone receptor positive (HR+), human epidermal growth factor receptor 2 negative (HER2/neu-negative (HER2–)) advanced breast cancer, in combination with exemestane, in postmenopausal women without symptomatic visceral disease after recurrence or progression following a non-steroidal aromatase inhibitor (NSAI). The patient population and intervention addressed by the manufacturer's submission (MS) accurately reflects this license although the MS also presents supporting evidence for the effectiveness of everolimus (Afinitor) in combination with tamoxifen which is outside the scope of the decision problem. Specified comparators in the MS are: endocrine therapy (exemestane (Aromasin), tamoxifen (Nolvadex) and fulvestrant (Faslodex)) and chemotherapy (docetaxel (Taxotere), capecitabine (Xeloda) and doxorubicin (Adriamycin)). All are used in clinical practice and are therefore considered to be appropriate comparators although it is noted that fulvestrant is not recommended by NICE. The outcomes addressed in the decision problem include overall survival (OS), progression-free survival (PFS), response rates, adverse events (AEs) and health-related quality of life (HRQoL). These outcomes are standard in this disease area and are appropriate.

1.2 Summary of clinical effectiveness evidence submitted by the manufacturer

Only one trial was identified for inclusion in the systematic review (BOLERO-2) which compared everolimus in combination with exemestane to placebo in combination with exemestane. This was a multicentre, international, double-blind, phase III trial of 724 patients with HR+, HER2- advanced breast cancer whose disease was refractory to previous NSAIs (letrozole (Femara) or anastrozole (Arimidex)). Supporting evidence was provided from the TAMRAD trial which compared everolimus in combination with tamoxifen to tamoxifen alone. This was an open-label phase II trial of 111 patients in France with HR+, HER2– metastatic breast cancer with prior exposure to aromatase inhibitors (AIs), and experiencing progressive disease. In order to compare everolimus to fulvestrant, a mixed treatment comparison was required and, to compare everolimus to chemotherapy, a 'naïve chained indirect analysis' was conducted.

In BOLERO-2, patients were stratified according to the presence of visceral metastasis and previous sensitivity to endocrine therapy. They were randomised 2:1 to receive everolimus in combination with

exemestane (hereafter referred interchangeably to everolimus, everolimus in combination with exemestane and the everolimus + exemestane arm) or placebo in combination with exemestane (hereafter referred interchangeably to exemestane, exemestane alone, placebo + exemestane and the placebo arm). Analyses were conducted after a median of 7, 12 and 18 months follow-up and, for interim OS, at 16 months median follow-up. A numerical improvement in OS was reported in the everolimus + exemestane arm vs the placebo + exemestane arm but it should be noted that the OS data are not yet mature (hazard ratio 0.77, 95% confidence interval (CI): 0.57 to 1.04, p=0.046). At 18 months, a significant improvement in local investigator assessed PFS of 4.6 months was reported for everolimus in combination with exemestane over exemestane alone (7.8 months vs 3.2 months; hazard ratio 0.45, 95% CI 0.38 to 0.54, p< 0.0001); improvements in PFS according to the independent central radiology committee were even greater (11.0 months vs 4.1 months; hazard ratio 0.38, 95% CI: 0.31 to 0.48, p< 0.0001) largely because patients who had discontinued study treatment as a result of the local assessment went on to receive a new anticancer therapy and so were censored on the date of the last valid radiologic assessment. From evidence submitted as a poster to the American Society of Clinical Oncology 48th Annual Meeting, particularly large differences between treatment groups for locally assessed PFS were apparent for 2/13 pre-specified subgroups: patients with bone-only metastases at baseline (12.9 months vs 5.2 months; hazard ratio 0.33, 95% CI: 0.21 to 0.53) and those with only one organ involved (11.5 months vs 4.4 months; hazard ratio 0.40, [redacted]). Significant improvements in overall response rate (ORR) (12.6% vs 1.7%, p < 0.0001) and clinical benefit rate (CBR) (51.3% vs 26.4%, p < 0.0001) were also reported in the everolimus arm. More AEs were reported by patients who received everolimus than by those who received only exemestane. The most frequently reported Grade 3/4 AE in the everolimus arm was stomatitis (8% vs <1%). Health-related quality of life data included bone analyses and time to definitive deterioration of HRQoL. All HRQoL findings favour everolimus over exemestane alone.

In the mixed treatment comparison, all studies were double-blind, placebo-controlled, multicentre, phase III RCTs which reported on OS, and PFS or Time To Progression (TTP) and included between 693 and 736 postmenopausal patients with advanced, locally advanced or metastatic HR+ breast cancer. In all studies, patients had received previous endocrine therapy but, in one study, this was not necessarily an AI. Not all patients in one study had HER2- tumours and in two other studies, data on HER2 status were not provided. The results suggest poorer outcomes for patients treated with fulvestrant 500mg than everolimus + exemestane. For OS, the treatment difference is not statistically significant [redacted]; in terms of PFS, the difference is statistically significant [redacted].

The results from the 'naïve chained indirect analysis' suggest that for OS, chemotherapy is more efficacious than tamoxifen (hazard ratio 0.94) but not everolimus (hazard ratio 2.09); the hazard ratio

for tamoxifen vs everolimus taken from TAMRAD is 1/0.45 = 2.22). For PFS it is assumed that chemotherapy has the same efficacy as TTP reported for tamoxifen in the TAMRAD trial. In TAMRAD, TTP was 4.5 months with tamoxifen alone and 8.6 months with tamoxifen + everolimus (hazard ratio for everolimus vs tamoxifen 0.54; hence hazard ratios for both chemotherapy and tamoxifen vs everolimus are 1/0.54 = 1.85).

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

The BOLERO-2 trial included in the direct comparison of everolimus + exemestane vs exemestane alone was of good quality with low risk of bias although it is noted that the blinding may have been compromised by AEs (stomatitis and rash) more typical to the everolimus arm. The ERG believes the manufacturer should have included more subgroup analysis at 18-months in the MS, rather than simply providing a reference to a poster presentation. Additional analysis of subgroups was provided following the ERG's clarification letter, however this was deemed to be commercial in confidence (CIC). Since tamoxifen is not an AI, the ERG does not believe that the TAMRAD trial is directly relevant to the decision problem.

The ERG is satisfied that the mixed treatment comparison methodology adopted by the manufacturer is acceptable, but has some concerns about the methodological quality of the studies included in the mixed treatment comparison. There were differences between studies in terms of the HER2 status of the patient population and proportions of patients previously treated with AIs in the adjuvant, metastatic or any setting. It should also be noted that the hazard ratio for the central assessment in BOLERO-2 was used for everolimus in combination with exemestane, whereas other studies included in the mixed treatment comparison used local investigator assessments. Therefore, the findings from the mixed treatment comparison should be treated with caution.

The robustness and reliability of the 'naïve chained indirect analysis' is questionable, in particular, concerning the use of an old, and now outdated, systematic review of chemotherapy vs endocrine therapy (in which the studies include endocrine therapies that are not typical of clinical practice and which do not measure PFS) and the assumption that the efficacy of everolimus + exemestane is the same as everolimus + tamoxifen. Therefore the findings from the 'naïve chained indirect analysis' should be treated with extreme caution.

1.4 Summary of cost-effectiveness evidence submitted by the manufacturer

The manufacturer developed a *de novo* partitioned survival model. It is constructed in Microsoft Excel and structured using three patient health states (progression free survival (stable disease), progressed disease and death). Variants of this model structure have been used in the modelling of metastatic oncology for a number of previous NICE STAs. The model population is based on the participants enrolled in the BOLERO-2 trial. Parametric survival models have been used to allow PFS and OS estimates to be made for the lifetime of the model. In the base case, the economic evaluation adopts a time horizon of 10 years, and the perspective is that of the UK NHS. Resource use, costs and utilities are estimated based on information from trial data and published sources.

For the comparison of everolimus + exemestane with exemestane alone the manufacturer's incremental cost-effectiveness ratio (ICER) per quality-adjusted life year (QALY) gained is £32,417 and the ICER per life year gained is £22,486 (figure taken directly from the manufacturer's model and not the MS). The manufacturer carried out a wide range of deterministic sensitivity analyses for this comparison. These generated ICERs ranging from £20,386 to £98,640 per QALY gained (figures taken directly from the manufacturer's model and not the MS). The manufacturer's probabilistic sensitivity analyses (PSA) suggests that there is a 41.6% chance that the ICER for everolimus + exemestane compared with exemestane alone is less than £30,000 per QALY.

The efficacy of everolimus + exemestane was also compared with tamoxifen, fulvestrant and chemotherapy (capecitabine, doxorubicin and docetaxel). The ICERs/QALY generated by the manufacturer's model are £29,109, £27,147, £24,362, £20,253 and £11,000 respectively. These are figures obtained from Table B56 of the MS. The ERG notes that a number of these values differ from those generated by the manufacturer's submitted model and also from results reported in Tables B54 and B55 of the MS.

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

The ERG is satisfied with the search strategy employed by the manufacturer to identify cost-effectiveness studies, and is reasonably confident that no relevant published articles exist.

The ERG found the manufacturer's submitted model difficult to navigate. The flow of logic between, and within, worksheets is not obvious. Labelling is minimal and often uninformative. Particular difficulty was experienced in deciphering cell formulae, which make no use of range labels. In view of these difficulties the ERG cannot be confident that all logic errors or questionable assumptions have been identified.

The ERG notes that some base case incremental QALY and ICER values provided in the report differ between tables, and also that a number of values are different from those generated by the manufacturer's model. The ERG also notes that deterministic sensitivity analyses results are not presented in the MS (although they are generated by the manufacturer's model). Furthermore, the manufacturer's model is not able to generate a cost-effectiveness acceptability curve that includes all comparators.

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 (BOLERO-2) that compares the intervention of interest (everolimus + exemestane) to one of the comparators of interest (exemestane). The population of patients included in BOLERO-2 is the same group of patients who are specified in the decision problem and for which everolimus has received a marketing licence from the European Commission.

1.6.2 Weaknesses and areas of uncertainty

Following receipt of the MS, the ERG submitted requests for specified Kaplan-Meier analyses of the latest BOLERO-2 trial data, to assess the extent to which the submitted decision model accurately reflects the experience of patients in the trial. The manufacturer claimed that they were unable to provide these data because. The ERG considers the justification for refusing these requests to be unhelpful and, in most respects, ill-founded.

Since tamoxifen is not an AI, the ERG does not believe that the TAMRAD trial is directly relevant to the decision problem. Furthermore, the methods used to derive hazard ratios to allow everolimus to be compared with fulvestrant and, in particular, chemotherapy (capecitabine, doxorubicin and docetaxel) result from analyses which, at best, should be viewed with caution. The ICERs derived using these hazard ratios cannot, therefore, be considered reliable.

In view of difficulties in navigating and understanding the flow of logic between and within worksheets in the submitted economic model, the ERG cannot be confident that all logic errors or questionable assumptions have been identified.

1.7 Summary of alterations/corrections and exploratory analyses undertaken by the ERG

The ERG made a number of alterations/corrections to the model, namely:

- Including all Grade 3/4 AEs reported to in the European Medicines Agency (EMA) Committee for Medicinal Products for Human Use (CHMP) European Public Assessment Report (EPAR)
- Correcting utility values
- Inclusion of assessment of response to treatment / disease progression for patients in PFS
- Removal of a 20% hazard ratio reduction to the modelled OS for everolimus + exemestane
- Applying discount to costs and outcomes annually
- Removal of background deaths
- Replacement of linear regression trends for the proportion of PFS time that patients spend on treatment
- Using PFS as measured by local assessment
- Application of exploratory survival models

Taken together, the ERG amendments increase the estimated ICER to £39,320 per QALY gained without the adjustment for PFS or exploratory survival models, to £52,285 per QALY gained using locally assessed PFS, and to £66,476 per QALY gained when the exploratory models are applied.

2 BACKGROUND

2.1 Critique of manufacturer's description of underlying health problem

The context section of the MS[1] (Section 2), appropriately presents the key issues relating to the underlying health problem, including pathology and prognosis. Summaries of the underlying health problem and the size of the problem taken directly from the MS[1] (p24-28) are presented in Box 1 and Box 2 respectively.

Box 1 Description of underlying health problem

Breast cancer, a malignant cellular growth in the tissues of the breast, is a heterogeneous disease, diverse in its natural history and pathology, and in its responsiveness to treatments.[2] There are several histological and molecular subtypes of breast cancer, with the main molecular subtypes being classified based on gene expression profiles.[2, 3]

HR status is important for choice of therapy with HR+ tumours … being more likely to respond to endocrine therapy. Similarly, HER2 status of the tumour influences likely outcome and responsiveness to treatment, with HER2 + tumours potentially being responsive to anti-HER2 therapies, and HER2− tumours associated with a less favourable prognosis.[4, 5]

Stage of disease is defined according to the size of the primary tumour, the extent of lymph node involvement and the presence of metastases[6] [and also] strongly influences treatment options and prognosis, with treatments for early stage disease being curative in intent, whereas those for metastatic or recurrent disease are generally palliative. In this submission, advanced breast cancer is considered to be represented by stage IV disease. Prognosis and survival among patients with breast cancer is related to stage of disease. Those with advanced disease tend to have a very poor prognosis, with patients typically surviving between 1 and 3 years.[7]

The symptoms of metastatic breast cancer typically result from the spread and growth of tumour cells within distant tissues and all are likely to significantly impact on patient health-related quality of life (HRQoL).[8]

Five-year OS results for patients with advanced breast cancer (stage IV disease) are approximately 13%,[7, 9] and data from the West Midlands Cancer Intelligence Unit indicate that median OS for these patients is approximately 12 months.[10]

Box 2 Size of underlying health problem

It is estimated that approximately 75% to 84% of breast cancer tumours are HR+ .[11, 12] Approximately 55% of breast cancers are HR+ and HER2−, and around 50% occur in postmenopausal women.[13]

There are no national data on the incidence of advanced breast cancer; however regional data from the West Midlands Cancer Intelligence Unit indicates that approximately 5% of patients diagnosed with breast cancer between 1992 and 1994 had metastases at the time of their primary diagnosis (stage IV) and a further 35% of all those with a primary diagnosis of early disease went on to develop advanced breast cancer in the 10 years following diagnosis.[14]

2.2 Critique of manufacturer's overview of current service provision

It is stressed in the MS[1] that: 'the treatment goals are palliative, centred on prolonging PFS and providing symptomatic relief, comfort and the maintenance or improvement of patient quality of life.' (MS,[1] p29). The MS[1] (p28-31) accurately describes options for the treatment of advanced breast cancer (summarised in Box 3). The proposed place for everolimus is following treatment with a previous AI, a point at which many will have developed resistance to an AI (MS[1] p30, Box 4) The MS[1] also includes a schematic of the current treatment pathway as defined by NICE (MS[1] p34) and reproduced in Figure 1.

Box 3 Summary of treatment options for patients with metastatic breast cancer

Advanced breast cancer: diagnosis and treatment' (NICE CG81) was published in 2009 and describes the current therapeutic management of postmenopausal women with HR+ tumours.[15]

According to the most recent UK guidelines,[5] together with guidance published by the European Society of Medical Oncology (ESMO)[16] and the National Comprehensive Cancer Network (NCCN)[17] in the USA and recent international consensus guidelines,[18] anti-oestrogen therapies (endocrine therapies) such as NSAIs or tamoxifen, are the option of choice for most postmenopausal women with HR+ advanced breast cancers who do not have immediate life-threatening visceral disease, as they are generally well tolerated. Such patients have inoperable disease and most have stage IV disease, though some with stage III disease may also be inoperable. As cytotoxic agents are associated with substantial toxicity, chemotherapy is not considered except for patients with immediately life-threatening visceral disease who require a rapid tumour response or on failure of endocrine therapy.[5, 16, 17]

Today, third-generation AIs – including the NSAIs such as letrozole (Femara®) and anastrozole (Arimidex®) and the steroidal AI exemestane − are the accepted standard of care for adjuvant therapy in postmenopausal women and for first-line treatment of metastatic disease in postmenopausal women.[5, 16-18]

Fulvestrant is not recommended by NICE within its licensed indication, as an alternative to aromatase inhibitors for the treatment of oestrogen-receptor positive, locally advanced or metastatic breast cancer in postmenopausal women whose cancer has relapsed on or after adjuvant anti-oestrogen therapy, or who have disease progression on anti-oestrogen therapy. UK clinical consensus is that the use of fulvestrant following AI failure is a valid clinical option for postmenopausal patients.[5]

However, when considering lines of endocrine therapy, there is little evidence to indicate the best sequence in which to use the available endocrine therapies, which include both steroidal and NSAIs, as well as tamoxifen.[5, 15, 18]

Treatment options following failure of AI therapy are unclear; there is no standard of care for postmenopausal women with advanced/metastatic HR+ breast cancer following failure of AI therapy.[5, 15, 18]

Box 4 Place of proposed treatment in the care pathway

As stated in the NICE Guideline CG81 on diagnosis and treatment of advanced breast cancer, there is a need to establish effective endocrine therapy for postmenopausal women with HR+ tumours who progress on treatment with an AI.[15] This is particularly relevant since approximately 50% of women with HR+ breast cancer present with *de novo* resistance to AIs and hence need an effective alternative therapy, while nearly all initial responders to endocrine therapy develop resistance at some point, resulting in disease progression.[19-22] Furthermore, around 50 to 60% of women in whom their first-line endocrine therapy fails will not respond adequately to subsequent lines of endocrine therapy, leaving limited treatment options.[23] In addition, with increasing use of AIs in the adjuvant setting, more patients who progress to metastatic disease have already been exposed to AIs and may therefore have already developed resistance to these agents.[5]

In clinical studies, the addition of everolimus to endocrine therapy (exemestane or tamoxifen) has been shown to induce tumour responses and significantly prolong PFS and OS (as discussed in section 6).[24, 25] The 2012 1st international consensus guidelines for advanced breast cancer and the updated 2012 NCCN guidelines both acknowledge the recent, emerging evidence that the addition of everolimus to treatment with an AI improves outcomes in patients with endocrine resistance.[17, 18]

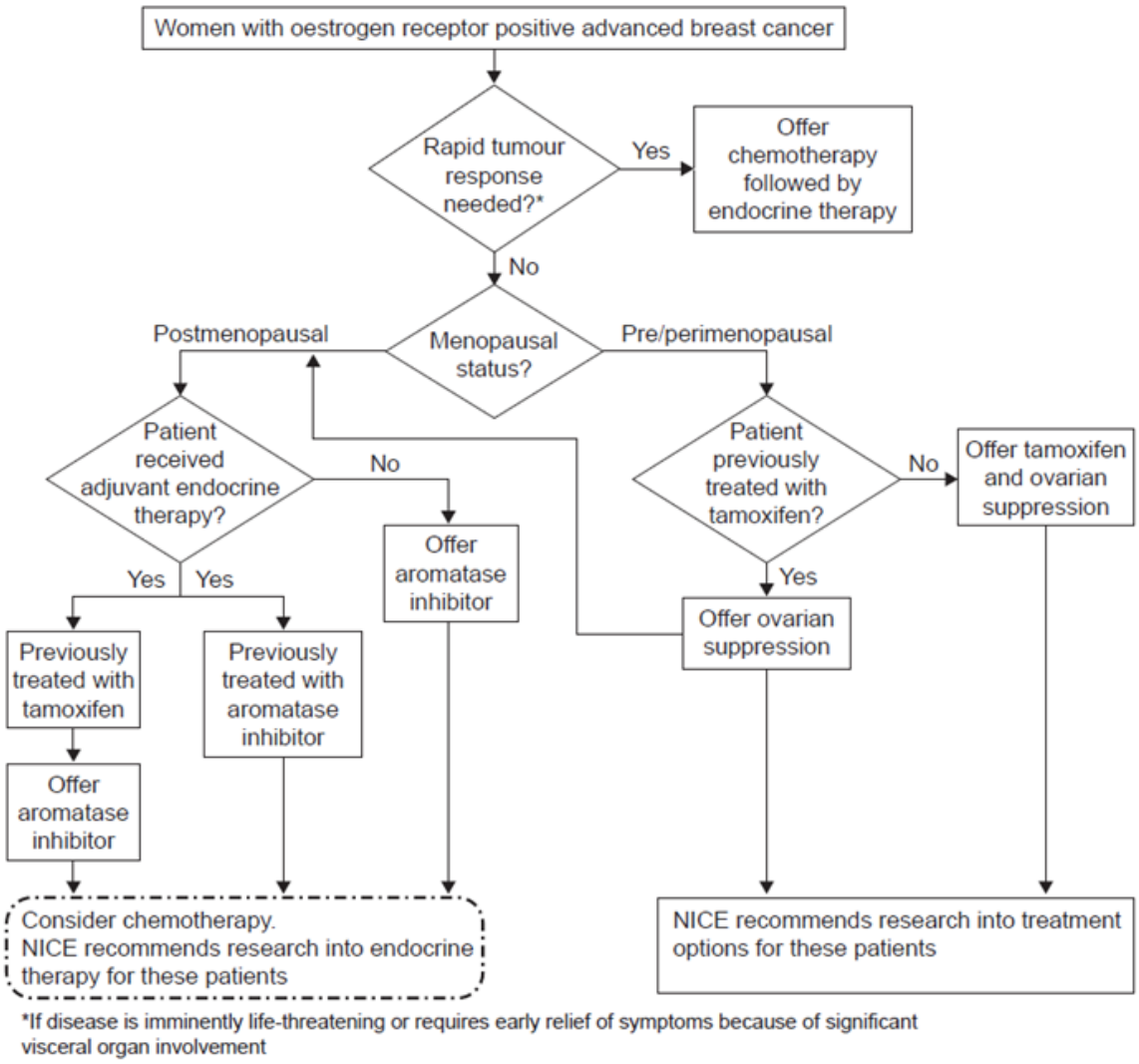


Figure 1 Current treatment pathway as defined by NICE

The manufacturer estimates that 1548 patients in England and Wales would be eligible for treatment with everolimus (see Table 1). Assuming a 1.7% growth rate in the overall population of England and Wales, this would rise to 1574 in 2013, 1601 in 2014, 1628 in 2015 and 1656 in 2016. The assumptions made to derive these figures appear to be reasonable.

Table 1 Manufacturer's estimate of eligible patients in England and Wales

Description	Proportion	Population	Source
Women in England and Wales aged 15 years or older in mid-2011		23,699,000	Estimated resident population by single year of age and sex; based on the results of the 2011 census[26]
Female population ≥15 with invasive breast cancer	0.18%	42,658	NICE early breast cancer (CG80) and advanced breast cancer (CG81) clinical guidelines[11, 15] and advanced breast cancer costing template[27]
Women with early and locally advanced invasive breast cancer	95.00%	40,525	
Women with advanced invasive breast cancer	5.00%	2133	
Women presenting with early breast cancer that die before disease progresses	30.00%	12,797	
Women with early and locally advanced breast cancer progressing to advanced stage	35.00%	10,451	
Total number with advanced breast cancer mid-2011	120.40%	12,584	Calculation flow in CG80 and CG81[11, 15]
Estimated number of women with advanced breast cancer in 2012	101.70%	12,798	Office for National Statistics Population projections[26]
Postmenopausal Women (aged ≥55yrs)	69.00%	8831	NICE advanced breast cancer costing template[27]
Women with hormone receptor positive breast cancer	83.80%	7418	West Midlands Cancer Intelligence Unit[10]
Women with HER2 negative breast cancer	75.00%	5563	NICE advanced breast cancer costing template[27]
Women with Asymptomatic Visceral metastases (without Visceral Crisis)	74.60%	4172	Sharma *at al* 2011 ASCO abstract[28]
Women with hormone receptor-positive advanced breast cancer for whom (hormonal) therapy is appropriate	70.00%	2921	NICE advanced breast cancer clinical guideline[15]
Women in whom disease progresses or relapses while on, or after receiving a AI	52.8%	1548	Novartis Data on file[29]

3 CRITIQUE OF MANUFACTURER'S DEFINITION OF DECISION PROBLEM

Table 2 displays the decision problem presented in the MS[1] (pp29-30) and the manufacturer's rationale for any deviation from it. Each parameter is discussed in more detail in the text following the table.

Table 2 Decision problem specified by NICE and addressed in the MS

Parameter	Final scope issued by NICE	Decision problem addressed in the MS[1]	Rationale if different from scope
Population	Postmenopausal women with HER2-, oestrogen receptor-positive locally advanced or metastatic breast cancer whose disease has recurred or progressed after prior therapy which has included a non-steroidal aromatase inhibitor (NSAI)	Postmenopausal women with HER2-, hormone receptor-positive, advanced breast cancer without symptomatic visceral disease after recurrence or progression following a NSAI.	To reflect the licence
Intervention	Everolimus in combination with an aromatase inhibitor	Everolimus in combination with exemestane	To reflect the licence
Comparator (s)	Exemestane Tamoxifen Fulvestrant Chemotherapy (in accordance with NICE guidance)	Exemestane Tamoxifen Fulvestrant Chemotherapy (in accordance with NICE guidance, specifically docetaxel, capecitabine and doxorubicin)	n/a
Outcomes	Overall survival Progression free survival Response rate Adverse effects of treatment Health-related quality of life	Overall survival Progression free survival Response rate Adverse effects of treatment Health-related quality of life	n/a
Economic analysis	The reference case stipulates that the cost-effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. The reference case stipulates that the time horizon for estimating clinical and cost-effectiveness should be 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.	Cost-effectiveness presented as incremental cost per quality-adjusted life year (QALY) Time horizon: lifetime (10 years) base case in line with other late stage cancer models Perspective: NHS and Personal Social Services	n/a
Subgroups to be considered	None stated	None	n/a
Special considerations, including issues related to equity or equality	Guidance will only be issued in accordance with the marketing authorisation		

n/a, not applicable

3.1 Population

The patient population addressed by the MS[1] accurately reflects that of the marketing authorisation for everolimus in combination with exemestane for the treatment of metastatic breast cancer. There are two differences from the decision problem specified by NICE in its final scope:

1. The final scope specified prior treatment should be endocrine therapy

2. There was no mention of visceral disease in the final scope

The first difference has very little consequence for considering the evidence presented in the MS[1] since, in clinical practice, most patients who have received prior endocrine therapy would have received an AI. The second difference to specify patients should not have symptomatic visceral disease was made to reflect the marketing authorisation in Europe which was granted based on much of the evidence that is presented in the MS.[1]

It should, however, be noted that evidence submitted in the MS[1] did include patients who had visceral involvement – 56% in the pivotal BOLERO-2[24] trial - but would not necessarily have been symptomatic. The ERG sought further clarification from the manufacturer with regard to this issue, including a request for a breakdown at baseline of patients who had symptomatic and non-symptomatic disease. In response, which is considered to be commercially in confidence, the manufacturer stated that there is no standard clinical definition of symptomatic visceral disease and hence the license wording with regard to exclusion of patients with symptomatic visceral disease was agreed between Novartis and the EMA in order to exclude patients with immediately life-threatening visceral disease, for whom chemotherapy may be the preferred treatment option. The ERG agrees that the patients included in the MS[1] do not appear to have life-threatening visceral disease. Furthermore, this is a population of patients who would be considered for treatment with everolimus in combination with exemestane. In accordance with NICE CG81,[15] as an alternative, patients with symptomatic visceral disease could be treated with chemotherapy, particularly where visceral disease was considered life-threatening, whereas those with no symptomatic visceral disease would most likely be treated with endocrine therapy, probably an AI.

3.2 Intervention

The intervention specified in the final scope was everolimus (Afinitor) in combination with an AI. In the MS[1] it is specified as everolimus in combination with exemestane, which the ERG agrees is a more appropriate intervention since this is the only combination currently licensed and used in clinical practice. Everolimus was granted marketing authorisation by the European Commission in July 2012 for the treatment of hormone receptor positive, human epidermal growth factor receptor 2 negative

(HER2/neu-negative (HER2–)) advanced breast cancer, in combination with exemestane, in postmenopausal women without symptomatic visceral disease after recurrence or progression following a NSAI.

The manufacturer also presents (and uses) supporting evidence' for the effectiveness of everolimus in combination with endocrine therapy, by drawing on data from the phase II TAMRAD[25] study of everolimus + tamoxifen. Tamoxifen is not an AI and this combination is not licensed, nor is it used in clinical practice. Therefore the ERG does not consider the evidence from this trial to be directly relevant to the decision problem. The appropriateness of how the evidence for everolimus in combination with tamoxifen is employed in the MS[1] is addressed further by the ERG in sections 4.2.1 and 0.

3.3 Comparators

Three endocrine therapies (exemestane, tamoxifen and fulvestrant) are specified as comparators in NICE's final scope. All three are commonly used in clinical practice, although the ERG notes that in the TA239 appraisal of fulvestrant,[30] fulvestrant was not recommended by NICE. In addition, chemotherapy is included as a comparator. In the MS,[1] the manufacturer defines chemotherapy as follows: docetaxel, capecitabine and doxorubicin 'in accordance with NICE guidance.' The ERG believes there is some argument for including vinorelbine (Navelbine) as a comparator and so sought clarification from the manufacturer for excluding this. In their response, the manufacturer reported that vinorelbine was not included because feedback from clinicians suggested that the three main chemotherapy treatments used in the UK were capecitabine, doxorubicin and docetaxel. The ERG agrees that it is likely that these three agents are the most commonly used for this group of patients in clinical practice (with vinorelbine probably being the fourth most common).

It should be noted that, in the clinical section of the MS,[1] it was only possible to compare everolimus in combination with exemestane to exemestane and to fulvestrant, whereas in the cost-effectiveness section, the manufacturer also compared everolimus (in combination with exemestane or tamoxifen) to tamoxifen, capecitabine, doxorubicin and docetaxel. The appropriateness of these approaches are addressed further by the ERG in sections 4.3.3.

3.4 Outcomes

The outcomes listed in the final scope are OS, PFS, response rates, AEs and health-related quality of life (HRQoL) and all are included in the MS.[1] These outcomes are standard in this disease area. The ERG notes that although OS is considered to be the most robust outcome in trials of cancer treatments, very few trials of treatments for metastatic breast cancer employ OS as the primary endpoint; indeed BOLERO-2,[24] from which the majority of the evidence in the submission is derived,

specified PFS as its primary outcome. Progression free survival is also presented for selected (a priori) subgroups of patients and exploratory analyses of bone markers are also presented.

3.5 Other relevant factors

No equity issues were identified in the MS,[1] or by the ERG. The ERG is unaware of any on-going patient access scheme application.

3.6 Innovation

Everolimus is a protein kinase inhibitor that inhibits the mammalian target of rapamycin (mTOR), a key component of the phosphatidylinositol 3-kinase/AKT/mTOR (PI3K/Akt/mTOR) signalling pathway that controls cell survival, growth and proliferation. This pathway is believed to be important in allowing tumour cells expressing hormone receptors to escape from hormone dependence and hence develop resistance to endocrine therapy. Thus the ERG agrees with the manufacturer's assertion that everolimus has the potential to usefully impact upon the treatment of HR+ metastatic breast cancer, particularly for patients who develop resistance to endocrine therapy.

4 CLINICAL EFFECTIVENESS

Table 3 provides an outline of the manufacturer's approach in terms of deriving evidence for the clinical effectiveness of everolimus in combination with exemestane and its location within the MS.[1] The purpose of the table is to signpost the reader to the areas of clinical information within the MS.[1]

Table 3 Location of clinical information in the MS

Key information	Page number	Key tables/figures
Description of the technology	18-23	
Context	24-32	
Statement of decision problem	36-39	
Literature search	40,209-216, 221-226	
Study selection	40-42	Table B1
Data extraction	216-217	
Quality assessment	64-65, 89 217-220, 229-231	Table B10, Table B21
Clinical effectiveness evidence key trial	43-77, 99-115	Table B11, Table B31
Other clinical effectiveness evidence, including mixed treatment comparison	78-98, 227-228	Table B20, Table B22, Tables B25-B28

4.1 Critique of the methods of review(s)

From section 6.2.1 and 10.2.6 of the MS,[1] it is apparent that the manufacturer actually conducted a systematic review of everolimus in combination with endocrine therapy (exemestane, fulvestrant or tamoxifen) rather than more specifically in combination with an AI. A mixed treatment comparison was also conducted (section 6.7 of the MS[1]). For the cost-effectiveness section of the MS,[1] the manufacturer also included chemotherapy (docetaxel, capecitabine or doxorubicin) as comparators via 'a naïve chained indirect analysis'. The manner in which these reviews were conducted is explored in section 4.1.1.

4.1.1 Searches

Systematic review

Section 10.2 of the MS[1] describes the search strategies employed for the systematic review. The following databases were searched, 8 to 9 March 2012:

- MEDLINE and MEDLINE In-Process (OvidSP)
- EMBASE (OvidSP)
- Science Citation Index (ISI Web of Science)
- Conference Proceedings Citation Index – Science (ISI Web of Science)
- Cochrane Library (Wiley Interscience):
    - Cochrane Database of Systematic Reviews
    - Cochrane Central Register of Controlled Trials
    - Database of Abstracts of Reviews of Effects (DARE)
    - Health Technology Assessment Database (HTA)
- ClinicalTrials.gov (www.clinicaltrials.gov)
- International Clinical Trials Registry Platform (http://www.who.int/ictrp/en/)
- metaRegister of Controlled Trials (http://www.controlled-trials.com/mrct/)
- US Food and Drug Administration (www.fda.gov/)
- European Medicines Agency (www.ema.europa.eu/)
- National Institute for Health and Clinical Excellence (http://www.nice.org.uk/)
- American Society for Clinical Oncology (ASCO) annual meeting (www.asco.org)
- European Society for Medical Oncology (ESMO) annual meeting (www.esmo.org/)
- International Society for Pharmacoeconomics and Outcomes Research (www.ispor.org)
- European CanCer Organisation (ECCO) and European Breast Cancer Conference (EBCC) annual meeting (www.ecco-org.eu/)
- San Antonio Breast Cancer Symposium (SABC) (www.sabcs.org/)

For all databases, search terms included the term 'everolimus'. For MEDLINE and MEDLINE In-Process, EMBASE and Science Citation Index, searches were also limited to second line or recurrent advanced breast cancer or metastatic breast cancer. No language, study or date restrictions were employed, nor were any search filters used.

The search strategies employed appear to be comprehensive. The ERG also conducted its own searches of MEDLINE and MEDLINE In-Process (Ovid SP), EMBASE (Ovid SP), ASCO and SABCS on 5 December 2012 and did not identify any additional potentially relevant studies.

Mixed treatment comparison

Section 10.4 of the MS[1] describes the search strategies employed to inform the mixed treatment comparison. A series of searches were undertaken by the manufacturer on 22 March 2012 to identify systematic reviews and trials which could be used to provide indirect comparisons. The first searches were undertaken in the following databases:

- MEDLINE and MEDLINE In-Process (OvidSP);
- EMBASE (OvidSP)
- The Cochrane Library (Wiley Interscience);
- ClinicalTrials.gov (http://clinicaltrials.gov/);
- ICTRO (www.who.int/ictrp/).

A second series of searches were undertaken on 26 March 2012 in the Cochrane Library databases via the Wiley Interscience interface, specifically the Cochrane Database of Systematic Reviews (CDSR) and the Database of Abstracts of Reviews of Effects (DARE) and Health Technology Assessment (HTA) database. Searches were also undertaken on the National Horizon Scanning Centre website, and the NICE website.

For all databases, search terms were limited to identify breast cancer studies and, where databases allowed, attempts were made to limit the searches to identify RCTs and systematic reviews/meta-analyses. No language, study or drug restrictions were employed. Date restrictions were only employed for searches of the DARE and HTA databases (to 2010-2012 publications).

The search strategies employed appear to be appropriate. The ERG also conducted its own searches of the Cochrane Library (Wiley Interscience), MEDLINE and MEDLINE In-Process (Ovid SP) on 16 November 2012 and did not identify any additional potentially relevant studies.

Naïve chained indirect analysis

In order to conduct the 'naïve chained indirect analysis', a 'rapid search' of the Cochrane Library (Cochrane Database of Systematic Reviews, DARE and HTA databases) was conducted to identify systematic reviews and health technology assessments of chemotherapy and advanced or metastatic breast cancer. The precise detail of the search strategy is not provided in the MS[1] but it is stated that it 'was designed to be sensitive in order to identify all systematic reviews and health technology assessments about advanced or metastatic breast cancer' (p92) and, for DARE and HTA, limited to reviews published from 2010–2012.

It is not possible to assess the appropriateness of the search strategy employed from the level of detail provided. However the ERG conducted its own searches of the Cochrane Library (Wiley

Interscience), MEDLINE and MEDLINE In-Process (Ovid SP) on 16 November 2012 and did not identify any additional potentially relevant reviews.

4.1.2 Inclusion criteria

Systematic review

The MS[1] presented the inclusion and exclusion criteria for its systematic review in Table B1 in section 6.2.1 (p41, replicated again in Table C1 in section 10.2.6, p215). These are reproduced below in Table 4. The criteria appear to be appropriate to the decision problem.

Table 4 Eligibility criteria used in search strategy for everolimus vs endocrine therapy

	Clinical effectiveness
Inclusion criteria	**Population:** postmenopausal women with HR+, HER2- locally advanced or metastatic breast cancer whose disease had recurred or progressed following endocrine therapy, including treatment with non-steroidal aromatase inhibitors **Intervention:** everolimus in combination with exemestane, fulvestrant or tamoxifen **Comparator:** exemestane, fulvestrant or tamoxifen **Outcomes:** CBR, response rate (complete, partial, stable disease), OS, PFS or TTP, AEs and discontinuations due to AEs, HRQoL, time to treatment discontinuation **Study design:** RCTs of any duration and crossover RCTs if data were presented at crossover; non-randomised comparative and uncontrolled studies reporting AEs were also eligible for inclusion **Language:** there was no language restriction applied to the search; studies with English abstracts, but whose full reports were in languages other than English were not extracted but were listed for information only **Publication status:** published, unpublished and grey literature was eligible; studies published as abstracts or conference presentations were included if an associated published full paper could not be found and adequate data were presented
Exclusion criteria	None specified

Mixed treatment comparison

The MS[1] describes how studies were selected for its mixed treatment comparison in section 10.4.6 (p226). The search results were assessed for relevance to drug interventions for women with HR+ advanced or metastatic breast cancer. To achieve a network, the following eligibility criteria were relaxed for record selection from the results of the second searches:

- HER2- status: trials with mixed populations and where the HER2 status was not reported were considered eligible.
- Treatment lines other than second line were considered eligible.

A critique of the conduct of the mixed treatment comparison, including the manner in which the network was achieved, is presented in section 4.3.3.

Naïve chained indirect analysis

Explicit inclusion and exclusion criteria for the 'naïve chained indirect analysis' were not presented in the MS.[1] However, it is stated that having identified potentially relevant reviews, these 'were then sifted to remove those reviews that were not about drug interventions for advanced or metastatic breast cancer: surgery, radiotherapy, non-drug treatments, screening, prevention, etc.' (MS[1] p92) Without greater detail, it is not possible to comment on the appropriateness of the inclusion and exclusion criteria employed. However, the ERG does not believe that any relevant reviews were excluded.

4.1.3 Data extraction

Systematic review

The manufacturer described the data it planned to extract for its systematic review in section 10.2.7 (p216) of the MS.[1] These data appear to be appropriate. It is not clear whether the data extracted from any study was cross-checked. For the main trial (BOLERO-2[24]) that provided the majority of the evidence in the MS,[1] the ERG has cross-checked much of the data extracted with the published paper[24] and the EMA CHMP EPAR.[31] It is difficult to determine if any relevant data has not been extracted without access to the Clinical Study Report but, based on the information provided in the protocol and statistical analysis plan, it appears that the majority of the analyses that were planned were reported in the MS.[1] However, these were not always reported for the most recent data cut off (18-months). In its clarification letter to the manufacturer, the ERG therefore requested the following data at 18-months:

- Duration of exposure to study treatment
- Time to response
- Duration of response
- Treatment received after discontinuation

All of these data were provided by the manufacturer, although all were deemed to be CIC and therefore, wherever possible, the ERG has attempted to report only data that are not CIC.

Mixed treatment comparison

The manufacturer described the data it planned to extract for its mixed treatment comparison in section 10.4.7 (p226) of the MS.[1] These data appear to be appropriate. However, it is not clear whether the data extracted from any study was cross-checked. The ERG has cross-checked the data extracted with the published papers (and where applicable, previous documentation for NICE STAs and Conference slides) for each study and in some instances, identified some minor errors; where appropriate, these have been corrected in the tables throughout this ERG report.

Naïve chained indirect analysis

The manufacturer does not describe its data extraction strategy for the 'naïve chained indirect analysis' in the MS. However, it would appear that, in addition to the data from the TAMRAD[25] study extracted for the manufacturer's systematic review, only the value of the hazard ratio for the comparison between chemotherapy and endocrine therapy from another systematic review by Wilcken *at al*[32] was extracted. It is not clear whether the data extracted was cross-checked. However, the ERG has cross-checked the extracted data and has not identified any errors.

4.1.4 Quality assessment

Systematic review

The manufacturer conducted an assessment of risk of bias for studies included in its systematic review. It was conducted by using a checklist recommended by the Cochrane Collaboration[33] and presented in the MS in Tables C10 and C11 in the appendices (section 10.3, pp219-220). The ERG conducted its own assessment of risk of bias for these studies using the same checklist and largely reached conclusions that were similar to those of the manufacturer.

Mixed treatment comparison

The manufacturer conducted an assessment of risk of bias for all included studies in the mixed treatment analysis. It was conducted by using a checklist recommended by the Cochrane Collaboration[33] and presented in the MS in Tables C12-C15 in the appendices (section 10.5, p229-231). The ERG conducted its own assessment of risk of bias for these studies using the same checklist and reached conclusions which were similar to those of the manufacturer.

Naïve chained indirect analysis

No assessment of risk of bias was specifically presented for the 'naïve chained indirect analysis' although the primary studies included did not differ from those in the systematic review or mixed treatment comparison and so had already been assessed for risk of bias. However, it is not clear if the quality of the identified systematic review was assessed. Some limitations of this review were, however, raised in the MS[1] and are explored by the ERG in section 0.

4.1.5 Evidence synthesis

Systematic review

Because the studies identified for inclusion into the systematic review had different interventions and comparators, the manufacturer appropriately synthesised the data by reporting on each trial individually (section 6.5 of the MS[1]) and did not attempt a meta-analysis (section 6.6.2, p82 of the MS[1]).

Mixed treatment comparison

The manufacturer performed mixed treatment comparison analyses on two outcomes; PFS (or TTP) and OS. Log hazard ratios were used to inform the analyses and the results were presented as hazard ratios for fulvestrant vs exemestane and fulvestrant vs everolimus in combination with exemestane. The ERG believes this was the most appropriate way to synthesise the data.

Naïve chained indirect analysis

No data synthesis of the 'naïve chained indirect analysis' was undertaken in the clinical section of the MS.[1] The findings from this analysis were used to inform the cost-effectiveness analysis (see section 7.2.15 of the MS,[1] in particular p133).

4.2 Critique of trials of the technology of interest, their analysis and interpretation (and any standard meta-analyses of these)

4.2.1 Identified studies

The systematic review identified one main study (a phase III trial, BOLERO-2[24] and one supporting study (a phase II trial, TAMRAD[25]). BOLERO-2[24] compared everolimus in combination with exemestane to placebo + exemestane in postmenopausal women with HR+, HER2– advanced breast cancer whose disease was refractory to prior treatment with an AI (letrozole or anastrozole). TAMRAD[25] compared everolimus in combination with tamoxifen to tamoxifen alone in postmenopausal women with HR+, HER2– metastatic breast cancer with prior exposure to AIs, and experiencing progressive disease. Since tamoxifen is not an AI, the ERG does not believe that the TAMRAD[25] trial is directly relevant to the decision problem.

As only one RCT investigating the efficacy of everolimus in combination with exemestane is available, the manufacturer did not conduct a meta-analysis. The ERG agrees that it would not have been possible to perform a meta-analysis.

4.2.2 Trial characteristics

BOLERO-2[24] was a multicentre, double blind, international RCT that compared exemestane + everolimus (n=485) with placebo + exemestane (n=239) in postmenopausal women with HR+, HER2- advanced breast cancer whose disease was refractory to previous AIs (letrozole or anastrozole).

According to the MS,[1] patients were randomised (2:1) to receive treatment using a centralised allocation (i.e. interactive web response system/ interactive voice response system). The ERG notes that the method of allocation generation was not reported but is otherwise satisfied that randomisation was carried out appropriately and allocations were adequately concealed in this trial. Randomisation was stratified with the following baseline factors:

- Presence of visceral metastasis;
- Previous sensitivity to endocrine therapy, defined as at least 24 months of endocrine treatment in the adjuvant setting prior to recurrence, or either a response or disease stabilisation lasting for at least 24weeks with endocrine therapy given for advanced disease.

The key trial characteristics are described in Table B3 of the MS[1] (p46) and are summarised in this ERG report in Table 5.

Although designed as a double-blind trial, with a number of necessary checks in place to ensure this (see assessment of risk of bias in Appendix), 59% of patients who received everolimus reported stomatitis and 39% reported rash, compared to 12% and 7% in the exemestane arm (see section 4.4). As a result, the blinding may have been compromised given these AEs are considered to be characteristic of treatment with everolimus.

Table 5 BOLERO-2 trial characteristics

Characteristic	Description
Location	International (189 centres in 24 countries) including the UK (13 patients in 6 centres)
Design	Randomised, double-blind, placebo-controlled, phase III trial
Duration	Treatment continued until disease progression, the development of unacceptable toxicity or withdrawal of consent
Method of randomisation	Randomisation at a 2:1 ratio in favour of the everolimus in combination with exemestane arm. Patients were assigned to treatment arm by centralised allocation (i.e., interactive web response system/interactive voice response system)
Method of blinding (care provider, patient and outcome assessor)	Patients, investigator staff, persons performing the assessments, all Novartis personnel and individuals at central laboratories (including central imaging) were to remain blinded to the identity of the treatment from the time of randomisation until database lock
Intervention(s) and comparator	Everolimus (10 mg/day) in combination with exemestane (25 mg/day) (n = 485) Exemestane (25 mg/day) in combination with matched everolimus placebo (n = 239)
Primary outcomes	PFS based on local and central assessment Pre-planned analyses of PFS were to be undertaken after 317 and 528 local PFS events Tumour assessments based on the RECIST v1.0 criteria[34] were carried out locally every 6 weeks until progression
Secondary outcomes	OS ORR CBR Time to response Duration of response Safety (AEs, biomarker analysis, vital signs, time to deterioration of ECOG performance status) Quality of life, evaluated using the EORTC QLQ-C30 and breast cancer module BR23
Duration of follow-up	PFS analyses: 1st interim analysis – median follow-up of 7 months 2nd interim analysis – median follow-up of 12 months Final PFS analysis – median follow-up of 18 months OS analysis – median follow-up of 16 months

4.2.3 Participant characteristics

The key patient characteristics of the BOLERO-2[24] trial are presented in Table B5 of the MS[1] (p.51) and are summarised in Table 6 of this report. The patient population appears to reflect that specified in NICE's final scope. The ERG largely agrees with the manufacturer's statement that participant's baseline characteristics were well balanced between the two treatment arms for all major baseline characteristics. However, the ERG notes that a slightly higher proportion in the placebo arm had been most recently treated for metastatic disease (84% vs 79%) or received a NSAI for metastatic disease (76% vs 71%). In addition, from the EMA CHMP EPAR[31] it is noted that a slightly greater proportion of patients were aged ≥65 in the everolimus + exemestane arm (40% vs 34%).

Table 6 BOLERO-2 participant characteristics

Baseline characteristic	Everolimus in combination with exemestane (n=485)	Placebo in combination with exemestane (n=239)
Age, median (range)	62 years (34–93)	61 years (28–90)
Ethnic background, n (%)		
White	361 (74)	186 (78)
Black	13 (3)	3 (1)
Asian	98 (20)	45 (19)
Other	13 (3)	5 (2)
Performance status, n (%)		
ECOG 0	293 (60)	142 (59)
ECOG 1	174 (36)	84 (35)
ECOG 2	9 (2)	7 (3)
Missing	9 (2)	6 (3)
Disease-free interval, mediana (range)	58 months (1–340)	57 months (5–316)
Hormone receptor status, n (%)		
ER +, PgR +	351 (72)	173 (72)
ER +, PgR−	134 (28)	66 (28)
ER−, PgR +	0	0
Number of metastatic sites, n (%)		
1	155 (32)	69 (29)
2	152 (31)	81 (34)
≥ 3	175 (36)	89 (37)
Type of metastatic sites, n (%)		
Visceral	281 (58)	143 (60)
Lung	140 (29)	79 (33)
Liver	160 (33)	72 (30)
Bone	369 (76)	184 (77)
Bone only	105 (22)	50 (21)
Measurable disease, n (%)	338 (70)	162 (68)
Previous sensitivity to endocrine therapy, n (%)	409 (84)	201 (84)
Prior treatment with aromatase inhibitor, n (%)		
Adjuvant	142 (29)[b]	57 (24)[c]
Metastatic	343 (71)[d]	182 (76)[e]
Prior chemotherapy, n (%)		
Adjuvant	211 (44)	95 (40)
Metastatic	125 (26)[f]	61 (26)[g]
Prior anti-oestrogen treatment, n (%)		
Any	276 (57)	140 (59)
Tamoxifen	230 (47)	118 (49)
Fulvestrant	80 (17)	39 (16)
Number of prior therapies, n (%)[h]		
1	76 (16)	42 (18)
2	146 (30)	71 (30)
≥ 3	263 (54)	126 (53)

Baseline characteristic	Everolimus in combination with exemestane (n=485)	Placebo in combination with exemestane (n=239)
Purpose of most recent treatment n (%)		
Adjuvant therapy	(21)	(16)
Treatment of advanced/ metastatic disease	(79)	(84)

a Disease-free interval is defined as the time from diagnosis of breast cancer to first relapse in patients who received adjuvant therapy (308 patients in the everolimus + exemestane arm and 153 patients in the exemestane alone arm); b Five patients were miscoded as having received AI therapy for prevention only; c Two patients were miscoded as having received AI therapy for prevention only; d Twenty patients received AI therapy in both adjuvant and metastatic settings; e Twelve patients received AI therapy in both adjuvant and metastatic settings; f 58 patients received chemotherapy in both adjuvant and metastatic settings; g 38 patients received chemotherapy in both adjuvant and metastatic settings; h Previous therapies include those used in the adjuvant setting or to treat advanced disease.

4.2.4 Description and critique of the statistical approach

Outcomes

The pre-specified primary endpoint of the study was PFS derived from investigator assessment of radiology data, defined as the time from randomisation to the date of the first documented disease progression or death due to any cause. The ERG notes that the original plan had been to use central radiological review and that the change to investigator assessment was made after patient recruitment had started. While it is not generally advised to alter the method of assessment of the primary outcome once the trial is underway, the ERG is satisfied that this change was justified, and local investigator assessments are likely to be more in line with clinical practice than central radiological reviews.

The following pre-specified secondary efficacy outcomes are presented in the statistical analysis plan:

- **OS**: defined as the time from the date of randomisation to the date of death due to any cause
- **ORR:** defined as the proportion of patients with best overall response of complete (CR) or partial (PR) response according to RECIST and determined by the local investigator's tumour assessment
- **ECOG performance status (PS)**: assessed and recorded at screening, on treatment day 1 (prior to administration of the study drug), at week 6 and every 6 weeks thereafter as well as at discontinuation from study treatment
- **HRQoL**: the EORTC QLQ-C30 questionnaire, along with the breast module (BR23), were used to collect patients' HRQoL data and were administered on treatment day 1 and every 6 weeks thereafter until progression
- **Duration of response (CR or PR)**: defined as the time from the date of first documented response (CR or PR) to the date of first documented progression or death due to underlying cancer; only applies to patients whose best overall response was CR or PR
- **Time to response**: defined as the time from the date of randomisation to the date of the first documented response (CR or PR) as determined by the local investigator's tumour assessment; all patients were included. Patients who did not achieve CR or PR were censored
- **CBR**: defined as the proportion of patients with either a best overall response of CR, PR or stable disease (SD) lasting for 24 weeks or longer. A patient was considered to have SD for 24 weeks or longer if a SD response was recorded at 24 weeks or later from randomisation

A number of safety analyses, pharmacokinetic analyses and biomarker analyses are also detailed in the statistical analysis plan. Data is presented for all pre-specified outcomes in the MS[1] other than ECOG PS which is only presented at baseline.

Sample size calculation

The sample size calculation is based on the primary outcome, PFS. Based on the EFECT[35] study, the median PFS in the control arm was assumed to be 3.7 months. It was hypothesized that everolimus + exemestane would provide a clinically meaningful 26% reduction in the hazard ratio (corresponding to a 35% increase in the median PFS to 5 months). To detect a hazard ratio of 0.74 with 90% power, using a log rank test and a 2-look Lan-Demets group sequential design with an O'Brien-Fleming type boundary at one-sided cumulative 2.5% level of significance, 528 PFS events are required for the final analysis of PFS. In order to observe these events it was estimated that 633 patients were needed. Assuming approximately 10% of patients would be lost to follow-up or withdraw consent, 705 patients were required to be randomised.

Although the study was not specifically designed to detect a difference in OS, calculations were performed to ensure that an adequate number of OS events occurred to allow sufficient power to detect any differences, provided that there was a statistically significant difference in PFS. To detect a hazard ratio of 0.74 with 80% cumulative power, using a log-rank test and a 3-look Lan-Demets group sequential design with an O'Brien-Fleming type boundary at one-sided cumulative 2.5% level of significance, 392 events were required for the final analysis of OS. In March 2012 a protocol amendment was made to add an additional interim analysis after [redacted], resulting in a change to the number of deaths required for the final analysis from 392 to 398.

Statistical analyses

According to the statistical analysis plan, the analysis population for primary and secondary efficacy analyses is the full analysis set, consisting of all randomised patients. Analyses followed the intention-to-treat (ITT) principle, analysing patients according to the treatment and stratum they were allocated to at randomisation.

The safety population consists of all patients who received at least one dose of the study treatment and who have at least one valid post-baseline safety assessment. Patients were analysed according to the treatment they actually received.

The statistical methods used to analyse the efficacy outcomes in the trial are presented in Table 7. The ERG is generally satisfied that these methods of analysis are appropriate but is unclear as to why a one-sided 2.5% significance level has been used for the log rank tests when 95% confidence intervals have been reported (which would require a two-sided test to be performed).

According to the statistical analysis plan, OS is hierarchically tested in the following way: if the test of PFS was significant, OS would be tested for significance. If the test of OS did not yield a significant result, the OS endpoint would be tested again at subsequent analyses driven by the number

of deaths until it was found significant or at the final OS analysis. If the test of PFS was not found to be significant at the final analysis, OS would not be statistically evaluated. No mention of any other adjustments for multiple testing was made.

Table 7 Efficacy analyses

Outcome	Method of Analysis
PFS	The distribution of PFS was compared between two treatment arms using a stratified log-rank test at one-sided 2.5% significance level. The distribution function was estimated using the Kaplan-Meier method and displayed for each treatment arm. The median PFS along with the survival probabilities at 2, 4, 6 and 9 months and the associated 95% confidence intervals were presented for each treatment arm. The stratified Cox regression model was used to estimate to estimate the hazard ratio of OS along with 95% confidence interval where the baseline hazard function is allowed to vary across strata. PFS was censored at the last adequate tumour assessment if one of the following occurred: absence of event; event occurred after new anticancer therapy was given; event occurred after two or more missing tumour assessments. Discontinuation of study treatment is not considered a reason for censoring.
OS	The distribution of OS was compared between two treatment arms using a stratified log-rank test at one-sided 2.5% significance level. The distribution function was estimated using the Kaplan-Meier method and displayed for each treatment arm. The median OS along with the proportion of patients alive at 12, 18, 24 and 30 months and the associated 95% confidence intervals were presented. The stratified Cox regression model was used to estimate to estimate the hazard ratio of OS along with 95% confidence interval where the baseline hazard function is allowed to vary across strata. If a death was not observed by the date of analysis cut-off, then OS was censored at the date of last contact.
ORR	Proportions of subjects with ORR were presented by treatment arm along with 95% confidence intervals. The Cochran-Mantel Haenszel chi-square test (strata based on the randomisation stratification factors) was used to compare the two treatment arms at one-sided 2.5% level of significance.
ECOG PS	Descriptive statistics were used to summarise ECOG PS data at each scheduled assessment time point. Additionally, change from baseline at the time of each assessment was summarised. An analysis of time to definitive deterioration of the ECOG PS by one category of the score from baseline was also performed.
HRQoL	The number of patients filling in HRQoL data and the number of patients missing/expected to have HRQoL assessments were summarised by each treatment arm for scheduled assessment time points. The amount and the pattern of missing data were explored by treatment arm and over time. Descriptive statistics were used to summarise the individual item and scored sub-scale scores of HRQoL data at each scheduled assessment time point. Time to definitive 5% deterioration in global health status/ quality of life scale and in each of the three secondary scales were compared between the two treatment arms using the stratified log-rank test. The survival distributions were presented descriptively using Kaplan-Meier curves. Median time to definitive 5% deterioration and the proportions of patients without deterioration at 3 and 6 months were presented along with 95% confidence intervals.
Duration of response	DoR was summarised by treatment arm. Distribution of DoR was estimated using the Kaplan-Meier method and the median response duration was presented along with 95% confidence interval. If a patient had not had an event, duration is censored at the date of last adequate tumour assessment. No inferential analysis that compares duration of response between the treatment arms was performed.
Time to response	Time to response data was listed and summarised by treatment arm. Distribution of time to response was estimated using the Kaplan-Meier method and the median time to response was presented along with 95% confidence interval. Patients who did not achieve a confirmed PR or CR were censored • at maximum follow up for patients who had a PFS event • at last adequate tumour assessment date otherwise No inferential analysis that compares time to response between the treatment arms was performed.
CBR	CBR was summarised for the two treatment arms using descriptive statistics. The Cochran-Mantel-Haenszel test was used to compare the treatment arms.

Sensitivity analyses

Sensitivity analyses of PFS were performed to address the impact of missing/unknown tumour assessments and to assess the impact of censoring due to another anti-cancer therapy. A sensitivity analysis to assess the impact of stratification on PFS was also performed where treatment arms were compared using an unstratified log-rank test and a hazard ratio (with associated 95% confidence

interval) was obtained using an unstratified Cox regression model. A number of further supportive analyses were also implemented.

A sensitivity analysis of OS was performed where a stratified Cox proportional hazard model was fitted, adjusting the treatment difference for key potential prognostic factors.

Subgroup analyses

The following 13 pre-specified subgroup analyses for PFS were described in the statistical analysis plan:

- sensitivity to prior hormonal therapy (yes vs no)
- presence of visceral metastasis (yes vs no) - this was also a stratification factor for randomisation
- baseline ECOG PS (0 vs 1, 2)
- bone only lesions at baseline (yes vs no).
- age (< 65 years and ≥ 65 years)
- Japanese patients (refer to patients randomized in Japan sites)
- race (Caucasian, Asian and other)
- region (Europe, North America, Asia and other)
- prior chemotherapy (yes vs no)
- number of prior therapies (1 vs 2 vs ≥3)
- number of organs involved (1 vs 2 vs ≥3)
- prior use of hormonal therapy other than NSAI (yes vs no)
- PgR status (positive vs negative)

A further four exploratory subgroup analyses were described in the MS[1] :

- Measurable disease (yes vs no)
- Most recent therapy (AI vs anti-oestrogen vs other)
- Purpose of most recent therapy (adjuvant therapy vs treatment for advanced or metastatic disease)
- Previous treatment with fulvestrant (yes vs no)

The ERG notes that there is a large number of subgroups. Having a large number of subgroups is not generally recommended as it increases the risk of a statistically significant result, which is not clinically significant, being identified by chance.

.

4.2.5 Results

Progression-free survival

The planned final primary analysis for PFS was performed based on a data cut from December 2011, when there had been a total of 510 PFS events (median follow-up of 18-months). The final analysis for OS has not yet been carried out. A number of interim analyses have been conducted. These are presented in the MS[1] (Table B7, p58) and in Table 8.

Table 8 Timing of analyses

Analysis	Follow-up, months	Cut-off date	No. PFS events	No. OS events
First PFS interim analysis	7	11-Feb-2011	359 (68%)	83 (21%)
PFS Update[a]	12	8-Jul-2011	457 (87%)	137 (35%)
Second OS interim analysis	16	31-Oct-2011	Not performed	182 (46%)[b]
Final PFS analysis[a]	18	15-Dec-2011	510	200
Final OS analysis	—	—	—	392

a not pre-planned, requested by FDA; b data available in the EMA CHMP EPAR[31]

Progression free survival data are presented in the MS[1] (Table B12, p68) and in Table 9. The primary analysis was based on the local assessments; the central assessments were analysed to provide support. The ERG notes a substantial difference between the median PFS measured by each assessment. In the clarification response the manufacturer explained that this was because some of the disease progressions recorded by the local review were not considered as such by the independent central radiology committee but because these patients had discontinued study treatment and most went on to receive a new anticancer therapy, subsequent scans were not available and so patients had to be censored on the date of the last valid radiologic assessment. This meant that the estimate for PFS in both groups was longer based on the central assessment than the local assessment. While both local and central analyses at each time point achieve statistically significant results, strongly in favour of everolimus in combination with exemestane, the ERG believes that the local review is more likely to reflect clinical practice. Survival probabilities at 2, 4, 6 and 9 months and their respective 95% confidence intervals are not reported in the MS.[1]

Table 9 Progression-free survival based on local and central assessments at 7, 12 and 18-month analyses

Follow-up (months)	Assessor	PFS, median (months)		Hazard ratio (95% CI)	p-value
		Everolimus in combination with exemestane	Placebo in combination with exemestane		
7	Local	6.9	2.8	0.43 (0.35 to 0.54)	< 0.001
	Central	10.6	4.1	0.36 (0.27 to 0.47)	< 0.001
12	Local	7.4	3.2	0.44 (0.36 to 0.53)	< 0.0001
	Central	11.0	4.1	0.36 (0.28 to 0.45)	< 0.0001
18	Local	7.8	3.2	0.45 (0.38 to 0.54)	< 0.0001
	Central	11.0	4.1	0.38 (0.31 to 0.48)	< 0.0001

Sensitivity and subgroup analyses for progression-free survival

The ERG notes that none of the sensitivity analyses that are pre-specified in the statistical analysis plan (see above) are reported in the MS.[1]

The MS[1] provides a figure showing a number of the subgroup analyses pre-specified in the statistical analysis plan (Figure B5, p71) but does not provide any values for the associated hazard ratios. Furthermore, these analyses are from the 7-month cut-off. However, 18-months follow-up data for 12 of the 13 pre-specified subgroups were presented at the ASCO 48th Annual Meeting in June 2012,[36] sensitivity to prior hormonal therapy (yes vs no) being the subgroup for which data were not presented.[36] In addition, a forest plot containing all pre-specified subgroup analyses, along with the number of events, hazard ratios and confidence intervals for the 18-months analysis was provided by the manufacturer in their response to the clarification letter (and marked as CIC). These data are reproduced in Table 10 where particularly large differences between treatment groups for locally assessed PFS were apparent for the following two subgroups: bone-only lesions at baseline (around 7.5 months) and one organ involved (around 7 months). Interestingly, given the emphasis on patients without symptomatic visceral disease in the marketing authorisation for the EU, the ERG notes that those with no visceral metastasis had more favourable findings in both treatment arms.[1] The difference was around 4 months for those with, and around 5.5 months for those without, visceral metastasis.

[redacted] However, it should be noted that the trial was not powered to detect significant differences in subgroups of patients.

Table 10 Subgroup analyses for progression-free survival in BOLERO-2 at 18-months

Subgroup	N	Hazard Ratio (95% CI)	Median PFS, months	
			Everolimus + exemestane	Placebo + exemestane
All	724	0.45 (0.37 to 0.54)	7.82	3.19
Age				
< 65	449	0.38 [redacted]	8.31	2.92
≥ 65	275	0.59 [redacted]	6.83	4.01
Region				
Asia	137	0.60 [redacted]	8.48	4.14
Europe	275	0.45 [redacted]	7.16	2.83
North America	274	0.38 [redacted]	8.41	2.96
Other	38	0.40 [redacted]	4.53	1.48
Japanese patients				
Japan	106	0.58 [redacted]	8.54	4.17
Non-Japan	618	0.42 [redacted]	7.16	2.83
Race				
Asian	143	0.62 [redacted]	8.48	4.14
Caucasian	547	0.42 [redacted]	7.36	2.96
Other	34	0.25 [redacted]	6.93	1.41
Baseline ECOG PS				
0	435	0.48 [redacted]	8.25	4.11
1,2	274	0.39 [redacted]	6.93	2.76
PgR status				
Positive	184	0.51 [redacted]	6.93	2.83
Negative	523	0.41 [redacted]	8.08	3.32
Number of organs involved				
1	219	0.40 [redacted]	11.50	4.37
2	232	0.52 [redacted]	6.70	3.45
≥3	271	0.41 [redacted]	6.93	2.56
Presence of visceral metastasis				
No	318	0.41(0.31 to 0.55)	9.86	4.21
Yes	405	0.47 (0.37 to 0.60)	6.83	2.76
Bone-only lesions at baseline				
No	573	0.48 (0.39 to 0.58)	6.90	2.83
Yes	151	0.33 (0.21 to 0.53)	12.88	5.29
Number of prior therapies				
1	118	0.60 [redacted]	8.05	4.37
2	217	0.45 [redacted]	6.93	2.96
≥3	389	0.41 [redacted]	8.18	2.96
Prior chemotherapy				
No	231	0.53 [redacted]	6.97	3.45
Yes	493	0.41 [redacted]	8.18	3.19
Prior use of hormonal therapy other than NSAI				
No	326	0.52 [redacted]	7.00	4.11
Yes	398	0.39 [redacted]	8.11	2.76

Overall survival

Overall survival data are presented in the MS[1] (Table B13, p72) and Table 11. At 16 months, the hazard ratio of everolimus + exemestane compared with placebo + exemestane was 0.77 (95% CI: 0.57 to 1.04, p=0.046). The ERG is unsure why the p-value indicates a significant result but the confidence interval does not. Neither median OS, nor proportions of patients alive at 12, 18, 24 and 30 months and the corresponding 95% confidence intervals, were reported in the MS.[1]

Table 11 Overall survival at 7, 12 and 18-month analyses*

Follow-up, months	Number of deaths			
	Total	Everolimus in combination with exemestane (n = 485), n (%)	Placebo in combination with exemestane (n = 239), n (%)	Difference, %
7	83	52 (10.7)	31 (13.0)	2.3
12	137	83 (17.2)	54 (22.7)	5.5
16	182	112 (23)	70 (29)	6.0; p=0.046

* Although the most recent data (last row of the table) is marked as CIC in Tables B11 and B13 of the MS,[1] the same data is also reported in the text on p13 of the MS where the data is not marked as CIC

It is important to note that the OS data are not mature (median OS has yet to be reached) and this analysis is only an exploratory interim analysis. Furthermore, the possible impact of treatment received following treatment progression on either everolimus + exemestane or placebo + exemestane should be considered. Although not presented in the MS,[1] the ERG notes that data on subsequent treatment was provided in the EMA CHMP EPAR.[31] The ERG notes that chemotherapy was the most common therapy, particularly in the placebo + exemestane arm (see Table 12 adapted from Table 32 (p38) of the EMA CHMP EPAR.[31])

Table 12 Subsequent anti-cancer therapy received following progression at 7-month cut off[a]

Post-treatment therapy	Everolimus + exemestane N=279		Placebo + exemestane N=189	
	n	%	n	%
Chemotherapy	162	58.1	130	68.7
Hormonal therapy	137	49.1	83	43.9
Radiotherapy	24	8.6	13	6.9
Targeted therapy[b]	13	4.7	18	9.5
Immunotherapy	2	0.7	0	-
Surgery	2	0.7	0	-
Other	8	2.9	2	1.1

a The proportion of patients who received each therapy does not add up to 100% because patients could receive more than one therapy e.g. chemotherapy + hormone therapy; b bevacizumab, denosumab, lapatinib, monoclonal antibodies, sorafenib, trastuzumab and everolimus (commercial use)

In the clarification letter the ERG requested similar data for the 18-month cut-off. This was provided by the MS[1] but marked as CIC.

[redacted]

[redacted]

Clinical benefit and overall response

Clinical benefit rates and overall response rates (ORRs) are presented in the MS[1] (Table B14, p73) and in Table 13. The ERG notes that the CBR and ORRs are considerably higher in the everolimus + exemestane arm than in the placebo + exemestane arm.

Table 13 Clinical benefit rate and overall response rate at 7, 12 and 18-month analyses

Follow-up, months	ORR, n (%)[a]		CBR, n (%)[b]	
	Everolimus in combination with exemestane (n = 485)	Placebo in combination with exemestane (n = 239)	Everolimus in combination with exemestane (n = 485)	Placebo in combination with exemestane (n = 239)
7	46 (9.5)*	1 (0.4)	162 (33.4)**	43 (18.0)
12	58 (12.0)**	3 (1.3)	245 (50.5)**	61 (25.5)
18	61 (12.6)**	4 (1.7)	249 (51.3)**	63 (26.4)

a ORR defined as patients with either CR or PR; b CBR defined as CR, PR or SD at ≥ 24 weeks; *p < 0.001; **p < 0.0001.

At the 7-month interim analysis, time to response ranged from 5.1 to 37.1 weeks for the everolimus + exemestane arm compared with 7.4 weeks for the single patient who had a response in the placebo + exemestane arm. Duration of overall response data, presented in the MS[1] and at 7- months, ranged from 6.0 to 66.1 weeks for the everolimus + exemestane arm compared to 12.1 weeks for the single patient who had a response in the placebo + exemestane arm. Distributions of time to response and duration of response estimated using Kaplan Meier methods are not reported in the MS,[1] but were addressed in the clarification response (marked as CIC). [redacted]

Exploratory analyses

The MS[1] includes an analysis of bone markers (risk of disease progression in the bone, increase in bone resorption associated with exemestane, incidence of fractures) and time to deterioration of global health status. It was reported that by adding everolimus to exemestane, bone turnover was suppressed and the increase in bone absorption associated with exemestane was reversed. The change in bone turnover markers at 6 and 12 weeks is presented in the MS[1] (Figure B7, p74) and is reproduced in Figure 2. The figure shows that there are positive changes in bone turnover for patients treated with everolimus + exemestane whereas for placebo + exemestane, the changes are negative.

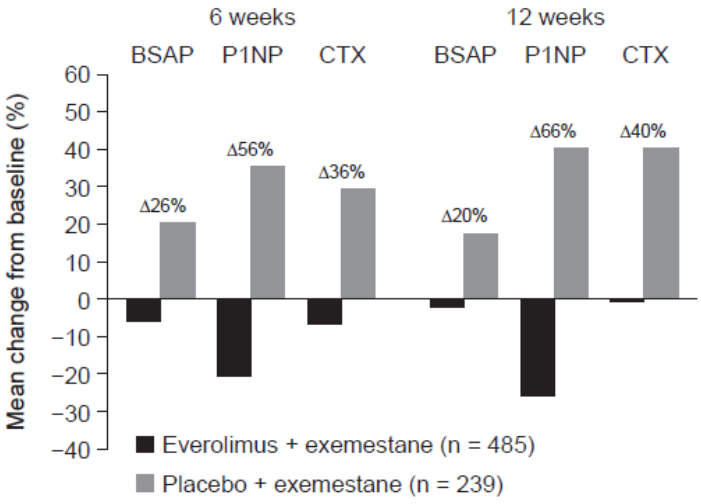


Figure 2 Change in bone turnover at 6 and 12 weeks

4.3 Critique of indirect analyses

4.3.1 Identified studies and their characteristics

Mixed treatment comparison

The following four trials were included in the mixed treatment comparison:

- BOLERO-2[24]
- CONFIRM[37]
- EFECT[35]
- SoFEA[38]

The manufacturer summarised similarities and differences of patient and trial characteristics in Table B20 of the MS[1] (p87). All data were cross-checked by the ERG and Table 14 reports these data, amended where appropriate.

All studies were double-blind, placebo-controlled, multicentre, phase III RCTs which reported on PFS or TTP as the primary outcome and included OS, CBR, ORR and safety as secondary outcomes. However, the OS findings from the EFECT[35] study were reported only as a poster presentation at the San Antonio Breast Cancer Symposium 2007[39] and not used by the manufacturer in its mixed treatment comparison.

The studies included between 693 and 736 postmenopausal patients with advanced, locally advanced or metastatic breast cancer. In all studies, patients had received previous endocrine therapy; in three studies it was a requirement that this was an AI. All patients had HR+ tumours but the two older studies (EFECT[35] and CONFIRM[37]) did not provide data on the HER2 status of patients.

Regarding the quality of the trials, the manufacturer concluded: 'In most studies, baseline characteristics were well balanced between the treatment groups, although many studies did not report sufficient data to adequately assess randomisation, concealment of treatment allocation and blinding. All studies used an intent-to-treat analysis, and there was little evidence of outcome reporting bias.' (MS,[1] p89) The ERG agrees with this assessment of the risk of bias.

Naïve chained indirect analysis

The 'naïve chained indirect analysis' included a systematic review of chemotherapy vs endocrine therapy by Wilcken *at al*[32] TAMRAD[25] which compared everolimus + tamoxifen with tamoxifen monotherapy was also included. Unlike the trials included in the mixed treatment comparison, this

was an open-label phase II trial in which TTP was only a secondary outcome, CBR being the primary outcome. This trial was conducted in France in a patient population of postmenopausal women with HR+, HER2– metastatic breast cancer who had received prior AI therapy in the adjuvant or metastatic setting and who had developed progressive disease. The characteristics of this trial are summarised in Table B3 (p46) and Table B5 (p51) of the MS[1] and are summarised here in Table 14.

The manufacturer notes that the TAMRAD[25] trial did not report sufficient details of the trial methodology to fully address the quality criteria evaluated; in particular, it was unclear how randomisation was carried out. Where it was possible to assess the quality, it is noted that ECOG PS of zero was more common in the everolimus + tamoxifen arm (59% versus 40%) and there was also a higher percentage of drop-outs due to AEs in this arm (22% vs 7%). However the trial analyses were correctly based on the principle of ITT and adequate detail was included in the final publication to suggest there was no selective reporting. The ERG agrees with this assessment of the risk of bias.

Table 14 Characteristics of trials included in indirect analyses

Parameters and characteristics	**Mixed treatment comparison**				**'naïve chained indirect analysis'**
	BOLERO-2[24]	**CONFIRM[37]**	**EFECT[35]**	**SoFEA[38a]**	**TAMRAD[25]**
Patients (N)	724	736	693	723	111
Key baseline characteristics (% except where stated):					
Median age (years)	61-62	61	63	63-66	63-66
HER2−	100	NR	NR	56	95
HR+	100	100	100	100	100
Prior adjuvant AI	27	15[b]	60[c]	18	41[c]
Prior AI in advanced setting	73	28[b]	87[c]	82	67[c]
Prior adjuvant chemotherapy	42	52[b]	45	NR	25[c]
Prior chemotherapy in advanced setting	26	20[b]	23	NR	51[c]
Intervention and comparator(s)	Everolimus (10mg/day) in combination with exemestane (25mg/day) (n = 485)	Fulvestrant (500mg on days 0, 14, 28 and every 28 days thereafter) (n=362) Fulvestrant (250mg on days 0, 28 and every 28 days thereafter) in combination with placebo (n=374)	Fulvestrant (500mg on day 0, 250mg on days 14 and 28 and 250mg every 28 days thereafter) in combination with matching placebo (n=351)	Fulvestrant (500mg on day 0, 250mg on days 14 and every 28 days thereafter) in combination with anastrozole (1mg/day) (n = 243) Fulvestrant (500mg on day 0, 250mg on days 14 and every 28 days thereafter) in combination with placebo (once daily) (n=231)	Everolimus (10 mg/day) in combination with tamoxifen (20 mg/day) (n=54);
	Exemestane (25mg/day) in combination with matched placebo (10mg/day) (n = 239)		Exemestane (25mg/day) in combination with matching placebo (n = 342)	Exemestane 25mg/day (n=249)	Tamoxifen (20 mg/day) (n =57)
Primary endpoint	PFS	PFS	TTP	PFS	CBR

a Data extracted from slides provided by Novartis for presentation at the European Breast Cancer Conference, 2012; b Data extracted from AstraZeneca 2010;[40] c Patients may have received prior therapy in both the adjuvant and advanced setting

4.3.2 Individual study findings

The results of the RCTs selected for the mixed treatment comparison are summarised in Table B22 of the MS[1] (p91), whereas the results for TAMRAD[25] which informs the 'naïve chained indirect analysis' are reported in the text (pp78-80) and in Tables B16-B18 and Figures B11 and B12 of the MS;[1] these are summarised in Table 15. Not all relevant results are included in Table B22 and so a modified version is presented in Table 15. As reported in the MS,[1] results from the BOLERO-2[24] trial show that everolimus in combination with exemestane led to a significant improvement in PFS compared with exemestane alone. By contrast, the EFECT[35] and SoFEA[38] trials showed no significant difference in TTP and PFS, respectively, between exemestane and fulvestrant. No study reported a significant difference in OS and the only study to report a significant difference in both ORR and CBR was BOLERO-2[24] . However, not all studies applied statistical tests for all secondary endpoints.

Table 15 Summary of results of trials included indirect analyses

Trials in mixed treatment comparison or 'naïve chained indirect analysis'	Follow-up, months	Median PFS or TTP, months	Median OS, months [deaths, n (%)]	ORR, n (%)[a]	CBR, n (%)[b]
BOLERO-2[24]	18				
Everolimus + exemestane		Central:11.0 Local: 7.8	n/a[c] [112 (23)]	61 (12.6)	249 (51.3)
Placebo + exemestane		Central: 4.1 Local: 3.2	n/a[c] [70 (29)]	4 (1.7)	63 (26.4)
CONFIRM[37]	9[d]				
Fulvestrant 500 mg		6.5	25.1 [NR (NR)][e]	33 (9.1)	165 (45.6)
Fulvestrant 250 mg + placebo		5.5	22.8 [NR (NR)][e]	38 (10.2)	148 (39.6)
EFECT[35]	13				
Fulvestrant 250mg + placebo		3.7[f]	24.4[g] [209 (59.5)]	20 (7.4)	113 (32.2)
Exemestane + placebo		3.7[f]	22.6[g] [197 (57.9)]	18 (6.7)	108 (31.5)
SoFEA[38h]	NR				
Fulvestrant 250mg + anastrozole		4.4	20.2 [168 (69)]	18 (7.4)	82 (33.7)
Fulvestrant 250mg + placebo		4.8	19.4 [167 (72)]	16 (6.9)	73 (31.6)
Exemestane		3.4	21.6 [173 (69)]	9 (3.6)	67 (26.9)
TAMRAD[25]	24				
Everolimus + tamoxifen		8.6	n/a [16 (30)]	5 (9.3)	33 (61.1)
tamoxifen		4.5	32.9 [31 (54)]	5 (8.7)	24 (42.1)

n/a, not applicable (median OS not reached); a ORR defined as patients with either CR or PR; b CBR defined as CR, PR or SD at ≥ 24 weeks; c OS results presented are based on the 16-month analysis (CIC); d data not presented in published paper but extracted from Fleeman *at al* 2011;[30] e updated OS results were presented at the 2012 San Antonio Breast Cancer Symposium, these data (26.4 months vs 22.3 months, p=0.016) were not available at the time the manufacturer undertook its search of the literature; f Time to progression rather than PFS; g data presented in poster presented to San Antonio Breast Cancer Symposium by Chia *at al* 2007[39] and not used in mixed treatment comparison presented by manufacturer, median follow-up of 20.9 months, median OS in months for ER + patients; h Data extracted from slides of the Conference presentation provided by Novartis alongside the MS[1]

4.3.3 Description and critique of the statistical approach

Mixed treatment comparison

The mixed treatment comparison used a Bayesian approach and was performed using the Markov chain Monte Carlo software package WinBUGs. This approach combines a prior probability distribution that reflects a prior belief of the possible values of the pooled relative effects with a likelihood distribution of the pooled effect based on the observed data in the different studies to obtain a posterior distribution of the pooled relative treatment effect.

The treatments evaluated in the mixed treatment comparison are connected as shown in Figure 3. It can be seen that the treatments fail to form a closed loop as they do not all have a common comparator. As a result, some treatments are connected by a longer path and this reduces the reliability of their comparison.

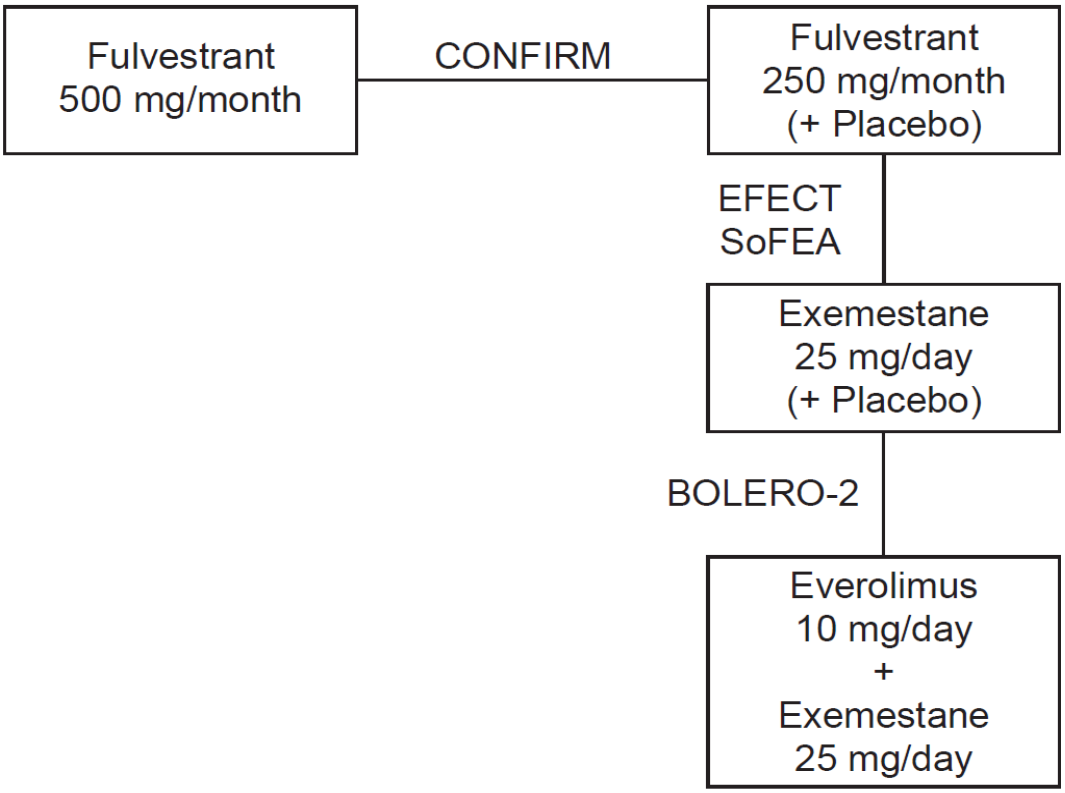


Figure 3 Evidence network used to inform the mixed treatment comparison

The data used in the mixed treatment comparison were the log hazard ratios and their precision (the reciprocal of the variance) was calculated from the hazard ratios and confidence intervals of the included studies. A Bayesian fixed effects model was used for the analysis. Exemestane was adopted as the baseline treatment in the model because it is used in clinical practice and it was the treatment

with most information in the evidence network. The basic parameters used in the model are the log hazard ratios compared with exemestane. These parameters were given vague priors so that the data from the studies would have greater influence than any prior belief on the results of the analysis.

For each outcome measure, the analysis was performed with three chains, run for a series of 20000 burn-in simulations to allow for convergence and then a further 20000 iterations were run for each chain and the estimates were obtained from the updated iterations.

The ERG is satisfied that the mixed treatment comparison methodology adopted by the manufacturer is acceptable.

With regard to the suitability of the inclusion of the trials, there is the possibility of clinical heterogeneity. The ERG notes that, contrary to the statement in the MS,[1] only a minority (43%) of patients in the CONFIRM[37] trial were reported to have been previously treated with an AI (Table B20, p87 of the MS[1] erroneously states that all patients had received an AI in the adjuvant setting). In the other trials, between 73% and 82% of patients had received an AI in the advanced setting compared with 28% in CONFIRM.[37] This may be an important difference because exploratory subgroup analyses from this trial have suggested that patients whose last treatment was an AI do not have as favourable outcomes as those whose last treatment was an anti-oestrogen (such as tamoxifen).[30] Another possible source of clinical heterogeneity relates to the HER2 status of patients, which is not known in two of the trials and is markedly different between the BOLERO-2[24] and SoFEA[38] trials. The ERG believes that the inclusion of patients with HER+ tumours is likely to result in reduced PFS and OS. In many other respects, however, the trials are relatively similar (including in terms of the line of treatment).

Naïve chained indirect analysis

To compare everolimus with chemotherapy, the manufacturer conducted a 'naïve chained indirect analysis' as shown in Figure 4 (a partial reproduction of Figure B23, MS[1] p133). No recognised statistical approach appears to have been performed for the 'naïve chained indirect analysis'. Data on the efficacy of chemotherapy vs tamoxifen was derived from a systematic review [32] and data on the efficacy of tamoxifen vs everolimus was derived from the TAMRAD[25] trial. To allow chemotherapy to be compared with everolimus, the hazard ratio reported for chemotherapy vs endocrine therapy was multiplied by that for everolimus vs tamoxifen.

Figure 4 Naïve chained indirect comparison of chemotherapy with everolimus

While if possible, a mixed treatment comparison would have been preferable to this approach, in response to a query raised in the ERG's clarification letter, the manufacturer highlighted a number of assumptions that would be required in order to perform this:

- Relaxing the inclusion criteria to allow any menopausal status
- Allowing studies without the line of therapy stated, or studies using first-line treatments only
- Including data that are almost thirty years old for chemotherapy studies

The Manufacturer therefore argued that the 'naïve chained indirect analysis' was no less robust than attempting to conduct a mixed treatment comparison given the additional assumptions that would be required. The ERG agrees that, given the above, a mixed treatment comparison would have questionable robustness and reliability. However, the ERG also notes that a number of the limitations still exist with the assumptions required for the 'naïve chained indirect analysis', namely

- In the systematic review,[32] no restrictions were placed on menopausal status, hormone receptor status or line of treatment
- A 'class effect' was assumed for the three main chemotherapy treatments in the analysis (docetaxel, capecitabine and doxorubicin); data on chemotherapy vs endocrine therapy was derived from a meta-analysis of studies published between 1978 and 1992, none of these studies included docetaxel or capecitabine
- The outcomes for tamoxifen were assumed to be equivalent to those of all endocrine therapies; endocrine therapies were considered by the authors of the systematic review to be atypical[32]
- For PFS, the efficacy of chemotherapy was assumed to be that of tamoxifen as reported in TAMRAD[25]
- The clinical effectiveness of everolimus in combination with exemestane was the same as everolimus in combination with tamoxifen – this was necessary to complete the chained network to derive a hazard ratio for chemotherapy

These results should, therefore, be viewed with extreme caution.

4.3.4 Results

Mixed treatment comparison

The manufacturer performed mixed treatment comparison analyses on two outcomes; PFS (or TTP) and OS. The manufacturer recognises that PFS and TTP are different outcomes but justifies combining them based on the reasoning that TTP is assumed to be similar to PFS in diseases with short survival times. As the median survival time in metastatic breast cancer is only approximately 2 years, most deaths would be disease-related and therefore count towards progression. The ERG supports the assumption made by the manufacturer.

The mixed treatment comparison for PFS was based on the data presented in the MS[1] (Table B25, p94) and reproduced in Table 16. The ERG notes that the data extracted from the BOLERO-2[24] study are based on the central assessment rather than the local assessment. The ERG also notes that the hazard ratio obtained from the central assessment was more favourable to the everolimus in combination with exemestane arm than that obtained from the local assessment.

Table 16 Hazard ratios for PFS in studies included in mixed treatment comparison

Trial	Treatment	Comparator	Hazard ratio	95% CI	
BOLERO-2[24]	Everolimus in combination with exemestane	Exemestane	0.38	0.31	0.48
CONFIRM[37]	Fulvestrant (500 mg)	Fulvestrant (250 mg)	0.80	0.68	0.94
EFECT[35]	Fulvestrant	Exemestane	0.96	0.819	1.133
SoFEA[38]	Fulvestrant (250 mg)	Exemestane	0.95	0.79	1.14

The results of the mixed treatment comparison for PFS can be found in the MS[1] (Table B27, p96) and are reproduced in Table 17. Hazard ratios greater than 1 indicate that the comparator treatment is less effective than the intervention. Therefore, the results suggest that fulvestrant (at either 250mg or 500mg) is more efficacious than exemestane but less efficacious than everolimus in combination with exemestane. It should be noted that the 500mg dose is the dose that is now licensed and used in clinical practice, although neither dose is recommended by NICE.[41]

In the MS[1] it is stated that 'For PFS and TTP, everolimus in combination with exemestane was found to perform better than all other comparators, and this difference was statistically significant compared with exemestane, fulvestrant (both 250 mg and 500 mg) and tamoxifen' (section 6.7.7, p96). The ERG is confused by the claim that everolimus was found to perform better than tamoxifen as no studies containing tamoxifen are included in the manufacturer's mixed treatment comparison. It is also noted that the manufacturer stated in their clarification response that tamoxifen could not be included in the mixed treatment comparison and so it is possible that this conclusion is drawn from the TAMRAD[25] trial rather than the mixed treatment comparison.

Table 17 Mixed treatment comparison results for PFS

Comparator treatment	Hazard ratio vs exemestane (95% CI)	Hazard ratio vs everolimus in combination with exemestane (95% CI)
Fulvestrant 250 mg	[redacted]	[redacted]
Fulvestrant 500 mg	[redacted]	[redacted]

The mixed treatment comparison for OS was based on the data in the MS[1] (Table B26, p94) and reproduced in Table 18. The ERG notes that OS data from the EFECT[35] trial were available[39] when the review was conducted and is unclear as to why these data were not included in the mixed treatment comparison.

Table 18 Hazard ratios for OS in studies included in mixed treatment comparison

Trial	Treatment	Comparison	Hazard ratio	95% CI	
BOLERO-2[24]	Everolimus in combination with exemestane	Exemestane	0.77	0.57	1.04
CONFIRM[37]	Fulvestrant (500 mg)	Fulvestrant (250 mg)	0.84	0.69	1.03
SoFEA[38]	Fulvestrant (250 mg)	Exemestane	1.05	0.84	1.29

The results of the mixed treatment comparison for OS can be found in the MS[1] (Table B28, p96) and are reproduced in Table 19. Hazard ratios >1 indicate that the comparator treatment is less effective than the intervention. Everolimus in combination with exemestane is found to perform better than both doses of fulvestrant in terms of OS but the treatment differences are not statistically significant.

Table 19 Mixed treatment comparison results for OS

Comparator treatment	Hazard ratio vs exemestane (95% CI)	Hazard ratio vs everolimus in combination with exemestane (95% CI)
Fulvestrant 250 mg	[redacted]	[redacted]
Fulvestrant 500 mg	[redacted]	[redacted]

The manufacturer was unable to perform a complete statistical assessment of heterogeneity. For the PFS analysis there was only one pair-wise comparison that was supported by evidence from more than one trial. Two trials (EFECT[35] and SOFEA[38]) compared fulvestrant 250mg with exemestane. The hazard ratios from these trials were extremely similar, 0.96 and 0.95. Combining these trials in a meta-analysis resulted in an I^2 value of 0% indicating that very little heterogeneity was present. For OS, because the manufacturer did not include the data from the EFECT[35] trial, there were no pair-wise comparisons supported by evidence from more than one trial. The ERG is generally satisfied with the manufacturer's approach to the assessment of heterogeneity; however, had they included the OS data

from the EFECT[35] trial they could have assessed whether there was heterogeneity present between this trial and the SOFEA[38] trial.

Naïve chained indirect analysis

The results of the 'naïve chained indirect analysis' are not presented in the clinical section of the MS[1] but are presented in the cost-effectiveness section (Section 7.2.15, p133). For chemotherapy, PFS is assumed to be the same as tamoxifen (MS,[1] p92). The hazard ratio for TTP (which is assumed to equal PFS) for everolimus vs tamoxifen in TAMRAD[25] is reported to be 0.54; 95% CI, 0.36 to 0.81; p = 0.0021 (MS,[1] p79). The inverse of the hazard ratio for tamoxifen vs everolimus is reported as 1/0.54 = 1.85 (MS,[1] Table B39, p134). Results for OS are presented in Figure B23 of the MS[1] (p133) and report that the hazard ratio for chemotherapy vs tamoxifen is 0.94, for tamoxifen vs everolimus the hazard ratio is 1/0.45 = 2.22 (as in TAMRAD[25] the hazard ratio is reported as 0.45 (05% CI: 0.24 to 081)) and for chemotherapy vs everolimus the hazard ratio is 2.09. The ERG notes that these hazard ratios were not presented with confidence intervals as is the norm and sought clarification from the manufacturer. The manufacturer stated that the 'naïve chained indirect analysis' was used to calculate a simple hazard ratio of chemotherapy vs everolimus in combination with exemestane and as such, confidence intervals were not calculated. Given the assumptions required to generate this hazard ratio, this stance was probably reasonable as a confidence interval may have provided false confidence in these results. For reasons highlighted in section 4.3.3 above, these results should be seen as extremely exploratory.

A further word of caution is necessary in interpreting the 'naïve chained indirect comparison' results. Firstly, the limitations of the Wilcken *at al*[32] systematic review must be considered. This systematic review was published in 2003 and treatment has changed considerably since then. Furthermore, even at that time, the authors noted: 'The trials were generally old (published between 1963 and 1995) and small (median 70 participants, range 50 to 226 women). The chemotherapy regimens used were reasonably conventional, although taxanes were not included. Endocrine therapies were less conventional.' (Wilcken *at al*[32] p5), indeed, none of the endocrine therapies comprised AIs. Four studies included tamoxifen as part of a treatment (three studies tamoxifen alone) and two studies compared a regimen including doxorubicin with tamoxifen alone, but none of the studies included docetaxel or capecitabine. Only OS (not PFS) was assessed in the review which is why the PFS for all chemotherapy regimens was considered to be the same as that for tamoxifen as reported in the TAMRAD trial.[25] Secondly, the limitations of the TAMRAD[25] trial and the assumptions surrounding its use in the 'naïve chained indirect comparison' should be recognised. Namely, this was an open-label phase II trial only conducted in France and the assumption that the efficacy of everolimus + exemestane is the same as everolimus + tamoxifen (which is being assumed to complete this chain) is totally untested.

4.4 Critique of the adverse events data

In BOLERO-2,[24] there was a greater proportion of serious AEs, Grade 3/4 AEs and withdrawal due to AEs in the everolimus + exemestane arm compared with the placebo + exemestane arm (see Table 20). The manufacturer suggests that the higher rate of withdrawals 'is likely to reflect the increased time on study drug and lower rates of withdrawal due to disease progression observed in the everolimus groups … at the 12-month analysis for BOLERO-2,[24] the duration of exposure to everolimus was approximately double that for placebo (24 weeks vs 13 weeks), possibly reflecting the lower rate of disease progression in the everolimus group.' (MS[1] p100 and p103) At the ERG's request, exposure to study treatment was also provided for the 18-months analysis (marked as CIC): [redacted]. However since the majority of AE data reported in the MS[1] is from the 7-month analysis at which point patients had received everolimus for a median of 14.6 weeks compared with 12 weeks of placebo (Table B32, section 6.9.3 of the MS[1] (p103)), the ERG believes that increased time on study drug is unlikely to be a major factor.

Table 20 Summary of adverse events in BOLERO-2

class/AEs	7-month follow-up, n (%)		18-month follow-up, n (%)	
	Everolimus in combination with exemestane (n=485)	Placebo in combination with exemestane (n=239)	Everolimus in combination with exemestane (n=485)	Placebo in combination with exemestane (n=239)
Serious AEs, n (%)	110 (22.8)	29 (12.2)	NR	NR
Grade 3 or 4 AEs, n (%)	211 (43.8)	61 (25.6)	NR	NR
Withdrawal due to AEs, n (%)	32 (6.7)	7 (2.9)	44 (9.1)	8 (3.3)

Additional relevant AE data are provided in a published paper[24] and on p42-49 of the EMA CHMP EPAR.[31] Seven (1%) deaths are reported in the everolimus arm which are attributed to AEs occurred during treatment, or within 28 days of stopping treatment: two deaths from sepsis and one each from pneumonia, tumour haemorrhage, cerebrovascular incident, renal failure, and suicide. One death was suspected by the investigator to be related to study treatment. The remaining six were attributed to the underlying malignancy and are not suspected to be related to study treatment; however, four were due to events that reflect known risks of everolimus therapy. In the exemestane alone arm, one (<1%) death during treatment was reported, the cause being attributed to pneumonia.

Compared with the exemestane alone arm, a much higher rate of suspected drug-related AEs occurred in the everolimus + exemestane arm for all Grade AEs (96% vs 60%) and for Grade 3/4 AEs (8% vs <1%). There were 52 patients (10.8%) in the everolimus + exemestane arm compared with 3 patients

(1.3%) in the placebo + exemestane group who experienced SAEs that were suspected to be adverse drug reactions. The most common serious adverse drug reactions are reported to be pneumonitis (12 (3%) vs 0), renal failure (5 (1%) vs 0) and hyperglycaemia (4 (<1%) vs 0).

In terms of types of AEs, the MS[1] only presents data on Grade 3/4 AEs. Stomatitis and anaemia are the only Grade 3/4 AEs observed in more than 5% of patients receiving everolimus in combination with exemestane (8% vs <1% and 6% vs <1% respectively). The only Grade 3/4 AE that is reported by more than 5% of patients in the placebo arm was an increase in gamma-glutamyltransferase level [redacted]. The published paper[24] and the EMA CHMP EPAR[31] provide data on types of All Grade AEs from the BOLERO-2[24] trial (figures not reported in the MS[1]). These sources report the following AEs occurred in at least 5% more patients in the everolimus arm compared with the exemestane arm: stomatitis (56% vs 10%), rash (36% vs 6%), fatigue (33% vs 26%), diarrhoea (30% vs 16%), decreased appetite (29% vs 10%), anaemia (16% vs 4%), aspartate aminotransferase level increase (13% vs 6%), hyperglycaemia (13% vs 2%), pneumonitis (12% vs 0), thrombocytopenia (13% vs <1%) and alanine aminotransferase level increased (11% vs 3%). A relatively large difference is also reported for dyspnoea (4% vs <1%). The proportion of patients reporting nausea was similar in both arms (27% vs 27%).

Data for selected AEs after 18-months follow-up were also presented at the ASCO 48th Annual Meeting in June 2012[36] and provided by the manufacturer in response to the ERG's clarification letter. These data are reproduced in Table 21 and suggest that there are no new differences emerging from this latter data cut, only modest increases in the number of reported AEs. Additional 18-month AE data were supplied by the manufacturer in their response to the ERG's clarification letter. These data (also presented in Table 21) are considered to be CIC.

Overall, the manufacturer argues that the data show that everolimus is generally well tolerated in postmenopausal women with HR+, HER2– advanced breast cancer. The ERG agrees with this statement.

Table 21 Proportion of adverse events in BOLERO-2 at 18-months

AE (preferred term)	Everolimus in combination with exemestane (n=485)					Placebo in combination with exemestane (n=239)				
	Grade					Grade				
	All	1	2	3	4	All	1	2	3	4
Any	100	7	40	44	9	91	26	36	23	5
Stomatitis	59	29	22	8	0	12	9	2	<1	0
Rash	39	29	9	1	0	7	5	2	0	0
Fatigue	37	18	14	4	<1	27	16	10	1	0
[redacted]										
[redacted]										
Diarrhoea	34	26	6	2	<1	19	14	4	<1	0
Nausea	31	21	9	<1	<1	29	21	7	1	0
Decreased appetite	31	19	10	1	0	13	8	4	1	0
Weight decreased	28	10	16	2	0	7	3	5	0	0
Cough	26	21	4	1	0	12	8	3	0	0
Pneumonitis	16	7	6	3	0	0	0	0	0	0
Hyperglycaemia	14	4	5	5	<1	2	1	1	<1	0
[redacted]										
[redacted]										
[redacted]										
[redacted]										
[redacted]										
[redacted]										
[redacted]										

The manufacturer argues that the safety profile of everolimus compares favourably with that of chemotherapy (see Table B33, p112 of MS[1]). The ERG agrees that this is likely to be the case. Adverse events for tamoxifen from the TAMRAD[25] trial are presented in the MS (Table B31, p102). Here it is evident that, after 24 months, AEs are more common in patients treated with everolimus than those treated with tamoxifen, although serious AE rates are similar (32% in each arm). There are no direct trials comparing everolimus with fulvestrant, and AEs associated with fulvestrant are not reported in the MS. However, AEs for fulvestrant have been reported in the CONFIRM[37] and EFECT[35] trials and presented for SoFEA[38] at the European Breast Cancer Conference in 2012. The ERG believes that, aside from AEs related to injection administration, the safety profile for everolimus is less favourable than that for fulvestrant.

4.5 Critique of the health related quality of life data

Health-related quality of life data reported by manufacturer from the BOLERO-2[24] trial included an assessment of patients' experience, as measured by the EORTC QLQ-C30 questionnaire, and are presented in the MS[1] (Table B15, p76) and in Table 22. The difference between the two groups at 18-months is in favour of everolimus + exemestane and is shown to be statistically significant. The median time to definitive deterioration is 8.3 months in the everolimus + exemestane arm compared

with 5.8 in the placebo + exemestane arm. However, neither the number of patients completing the HRQoL nor the number expected to have completed the HRQoL are reported in the MS.[1] No exploration of missing HRQoL data is reported in the MS,[1] neither are descriptive statistics of the individual item or scored sub-scale scores of HRQoL data. The ERG also notes that pre-planned results at 3- and 6-months have not been presented.

Table 22 Health related quality of life measured by EORTC QLQ-C30 at 7, 12 and 18-month analyses

Follow-up, months	EORTC QLQ-C30 median time to deterioration of global health status/HRQoL domain score (≥ 5%), months			
	Everolimus in combination with exemestane	Placebo in combination with exemestane	Hazard ratio (95% CI)	p-value
7	4.5	4.4	0.91 (0.68 to 1.20)	0.217
12	7.0	5.6	0.81 (0.62 to 1.06)	0.040
18	8.3	5.8	0.74 (0.58 to 0.95)	0.0084

4.6 Discussion of the clinical effectiveness section

Findings from the direct evidence from BOLERO-2[24] suggest that there is a significant improvement in PFS for patients in the everolimus arm. The ERG believes that the local investigator assessment is more reliable than the independent central assessment, although noting that the manufacturer uses the central assessment to inform its economic model. Differences in OS were not significant but the data for OS is not yet mature. Alongside these improvements in PFS and OS were improvements in other secondary outcomes including CBR, ORR and HRQoL. These improvements need to be considered alongside the less favourable safety profile for patients undergoing treatment with everolimus compared with those receiving exemestane alone.

Although not reported directly in the MS,[1] the ERG notes from evidence presented to ASCO 48th Annual Meeting[36] that in BOLERO-2,[24] patients with bone lesions may benefit the most from treatment with everolimus. However, the BOLERO-2[24] trial was not powered to detect significant differences and therefore these findings should only be considered exploratory.

Findings from the mixed treatment comparison also suggest improved PFS, but not OS, when everolimus is compared with fulvestrant; however, this is based on the more favourable findings from the central assessment of PFS reported in the BOLERO-2[24] trial. The ERG notes that there do appear to be some clinical differences between the studies in terms of patient population (HER2 status) and the proportions of patients previously treated with AIs (in the adjuvant, metastatic, or any setting). Furthermore, the hazard ratio for the central assessment reported in the BOLERO-2[24] trial was used for everolimus in combination with exemestane, whereas other studies included in the mixed

treatment comparison used local investigator assessments. Therefore, these findings should be treated with caution.

Evidence from the 'naïve chained indirect analysis' comparing everolimus to tamoxifen and chemotherapy also suggest an improvement in PFS and OS for patients treated with everolimus. However, the evidence is derived both from an outdated systematic review using atypical treatment regimens[32] and the TAMRAD[25] trial that is of lesser methodological quality than BOLERO-2[24] according to the assessment of risk of bias. It is also noted that a number of important, yet untested assumptions are made, including that the efficacy of everolimus + tamoxifen can be assumed to be that of everolimus + exemestane and that the PFS of chemotherapy regimens can be assumed to be equivalent to that of the TTP of tamoxifen as reported in TAMRAD.[25]

As well as its inclusion in the 'naïve chained indirect analysis', evidence from TAMRAD[25] is also presented as supporting evidence to the direct evidence presented in BOLERO-2.[24] The ERG does not consider the TAMRAD[25] trial to be directly relevant to the decision problem as tamoxifen is not an AI.

4.7 Conclusions of the clinical effectiveness section

The great majority of the evidence presented in the MS[1] is relevant to the decision problem. However, the ERG considers the quality of the evidence from the direct comparison of everolimus + exemestane to placebo + exemestane from BOLERO-2[24] to be of the highest quality and therefore the most robust and relevant.

BOLERO-2[24] reports improvement in PFS but not OS for patients who received everolimus in combination with exemestane over exemestane alone in postmenopausal women with HER2-, HR+ advanced breast cancer after recurrence or progression following a NSAI. The safety profile is less favourable for the everolimus arm; however, everolimus + exemestane is generally well tolerated and the safety profile compares favourably with that of chemotherapy.

Evidence from the mixed treatment comparison suggests an improvement in PFS for everolimus compared with fulvestrant. However, given possible differences in the trial populations and the use of the more favourable central assessment PFS from BOLERO-2,[24], these findings should be treated with caution.

A 'naïve chained indirect analysis' is used by the manufacturer to compare everolimus with tamoxifen. In order to do this, a number of untested assumptions based on largely inappropriate data are required. Therefore the ERG does not consider these findings to be robust or reliable and they should be treated with extreme caution.

5 COST-EFFECTIVENESS

This section provides a structured critique of the economic evidence submitted by Novartis in support of everolimus + exemestane for the treatment of post-menopausal women with HER2-negative oestrogen receptor positive locally advanced or metastatic breast cancer whose disease has recurred or progressed after prior therapy which has included a non-steroidal aromatase inhibitor. The two key components of the economic evidence presented in the MS[1] are (i) a systematic review of the relevant literature and (ii) a report of the manufacturer's *de novo* economic evaluation. Table 23 contains details of the location of key information within the MS.[1] The manufacturer has also provided an electronic version of their economic model which was developed in Microsoft Excel.

Table 23 Location of key cost-effectiveness information in the MS

Key information	Page number	Key tables/figures
Details of the systematic review of the economic literature	116-117, 234-239	
De novo analysis	117-122	Table B34, Figure B16,
Clinical evidence used in economic evaluation	123-135	Tables B35-B40, Figures B17-B23
Measurement and valuation of health effects	136-154	Tables B41-B46
Resource identification, measurement and valuation	155-164	Tables B47-B49
Methods of sensitivity analysis	165-167	Tables B50
Results - base-case analysis	168-185	Tables B51-B56, Figures B24-B37
Results - sensitivity analysis	185-189	Tables B57-B58, Figures B38-B39
Validation	189, 250-251	
Subgroup analysis	190-191	
Interpretation of economic evidence	192-193	
Assessment of factor relevant to the NHS and other parties	194-199	Tables C1-C6

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 manufacturer carried out a search to identified studies reporting the cost-effectiveness of everolimus in postmenopausal women with HR+, HER2-, advanced (locally advanced or metastatic) breast cancer who had already received endocrine therapy.

The databases searched included: MEDLINE, MEDLINE In-Process, EMBASE, EconLit and the NHS Economic Evaluation Database (NHS EED). All searches were carried out on 8 and 9 March 2012. The search strategy used did not include an economic search filter because scoping searches had indicated that the amount of literature for everolimus was very small. The search strategies comprised the drug name in combination with search terms for advanced or metastatic breast cancer. No date or language limits were applied. Full details of the search strategies, as well as the databases and resources searched, are provided in the MS[1] (Appendix 10, p234-239).

No economic evaluations of everolimus were identified.

5.1.2 Conclusions of the cost-effectiveness literature review

The manufacturer's search to identify studies reporting the cost-effectiveness of everolimus in postmenopausal women with HR+, HER2-, advanced (locally advanced or metastatic) breast cancer who had already received endocrine therapy 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 Checklists

Table 24 tests 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[42] and Table 25 summarises the ERG's appraisal of the economic evaluation conducted by the manufacturer using the Drummond checklist.[43]

Table 24 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	The use of vinorelbine has not been included in the model. The ERG recognises that it is likely that capecitabine, docetaxel and doxorubicin are the most commonly used chemotherapy for this group of patients, with vinorelbine probably being the fourth most common chemotherapy option.
Perspective costs	NHS and Personal Social Services	The perspective of the model is that of the NHS.
Perspective benefits	All health effects on individuals	Partially - the costs of AEs are not included in the base case analysis.
Form of economic evaluation	Cost-effectiveness analysis	Yes
Time horizon	Sufficient to capture differences in costs and outcomes	Yes
Synthesis of evidence on outcomes	Systematic review	A number of different approaches were used: trial data (everolimus, exemestane (BOLERO-2[24]) and tamoxifen (TAMRAD[25]); indirect comparison (fulvestrant); and literature review (chemotherapy). It is noted that no studies reporting PFS were identified from the literature review and that the manufacturer has assumed that the PFS hazard ratio for chemotherapy compared with everolimus + exemestane is the same as that for tamoxifen compared with everolimus + exemestane. The use of the information from the one study identified by the review that reported OS information in a 'naïve chained indirect analysis' to generate an OS hazard ratio for chemotherapy compared with everolimus + exemestane generates a value that cannot be considered robust.
Outcome measure	Quality adjusted life years	Yes
Health states for QALY	Described using a standardised and validated instrument	Partially – utility values were obtained from a published source, but disutilities due to AEs were excluded from the base case analysis.
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, although this is applied monthly, rather than annually.
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

Table 25 Critical appraisal checklist for the economic 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	-
Was the effectiveness of the programme or services established?	No	Although PFS gain was established, the immaturity of OS trial data calls into question the magnitude and significance of any OS gains
Were all the important and relevant costs and consequences for each alternative identified?	No	No patient monitoring costs prior to disease progression were included in the model. Adverse event treatment costs and disutilities were excluded from the base case analysis
Were costs and consequences measured accurately in appropriate physical units?	No	An arbitrary hazard ratio adjustment to modelled OS was applied which generates unjustified additional survival gain. A simplistic adjustment has been applied to drug cost estimates, which artificially reduces costs in the intervention arm.
Were the cost and consequences valued credibly?	No	Most of the unit costs for AEs cannot be verified from the quoted sources. The estimated utility value for patients in PFS is not correctly calculated from the source model.
Were costs and consequences adjusted for differential timing?	Yes	-
Was an incremental analysis of costs and consequences of alternatives performed?	Yes	-
Was allowance made for uncertainty in the estimates of costs and consequences?	Yes	-
Did the presentation and discussion of study results include all issues of concern to users?	Yes	-

5.2.2 Model structure

A schematic of the model structure is shown in Figure 5. Variants of this structure have been used in the modelling of metastatic oncology for numerous NICE STAs (for example, eribulin (TA 250)[44] and fulvestrant (TA 239)[30]).

Three health states are used to model disease progression. All patients enter the model in the Stable (PFS) health state and in each month can either progress to a 'worse' health state (i.e. from Stable to Progressed or Dead, or from Progressed to Dead) or remain in the same health state. Subsequent lines of therapy are not considered in the model.

The model has been developed in MS[1] Excel and has a one month cycle length. It includes a half-cycle correction and the base case time horizon is 10 years. A discount rate of 3.5% has been used for both costs and outcomes. The perspective is that of the NHS.

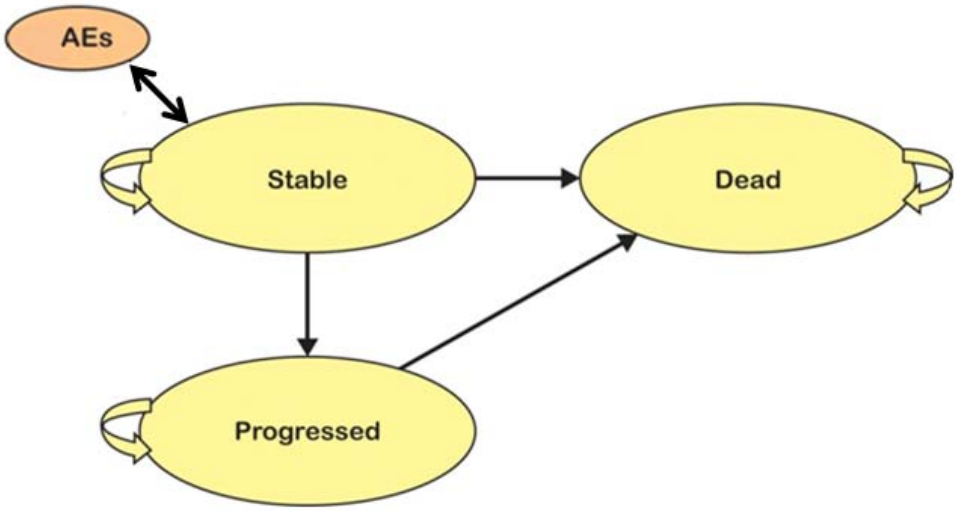


Figure 5 Schema of manufacturer's model

5.2.3 Population

The patient group considered in the base case is postmenopausal women with HR+ HER2- metastatic breast cancer who have progressed on therapy with a NSAI. The base-case uses data from the BOLERO-2[24] trial. The manufacturer considers that this population is representative of the patients who will receive everolimus in the UK. The median age of patients in the BOLERO-2[24] trial is 62 years (standard deviation=10.14 years).

5.2.4 Interventions and comparators

The technology considered in this analysis is everolimus in combination with exemestane. The main comparator is exemestane alone; however, a number of other comparators are also modelled, namely tamoxifen, fulvestrant, and chemotherapy (docetaxel, doxorubicin and capecitabine). In their response to the ERG's clarification questions the manufacturer reported that vinorelbine was not included as a chemotherapy option because feedback from clinicians suggested that the three main chemotherapy treatments used in the UK were capecitabine, doxorubicin and docetaxel. The manufacturer states that all treatments are implemented as per their marketing authorisations. However, this is only achieved implicitly as none of the treatments are modelled directly.

5.2.5 Perspective, time horizon and discounting

The economic appraisal is undertaken from the perspective of the NHS. Outcomes are expressed in terms of gains in life years and quality-adjusted life years (QALYs). The time horizon is set at 10 years and, in line with the NICE Methods Guide to Technology Appraisal,[42] both costs and benefits are discounted at 3.5%.

5.2.6 Treatment effectiveness and extrapolation

Everolimus + exemestane and exemestane alone

The same approach was used to model both the effectiveness of treatment with everolimus + exemestane and the effectiveness of treatment with exemestane alone.

Overall survival

Observed data from the BOLERO-2[24] trial were not used in the model. Instead, the manufacturer fitted a series of parametric curves to the Kaplan Meier analysis of OS data from the BOLERO-2[24] trial (median follow-up of 16 months). A number of curves were generated and statistical tests suggested that the curve generated by the log-logistic function provided the best fit, followed by that generated by the Weibull function. Following consultation with clinicians, the manufacturer chose to use the curve generated by the Weibull function as their base case. They reasoned that, compared with the Weibull function, the log-logistic function potentially over-estimates survival.

However, the manufacturer found that using the fitted Weibull function resulted in the post-progression survival (PPS) for everolimus in combination with exemestane being less than that predicted for some of the comparator drugs. To address this problem the manufacturer introduced a multiplication factor of 80% for everolimus + exemestane OS. This means that the mortality hazard in all time periods in the parametric function that models the efficacy of everolimus + exemestane is reduced by 20% so that the OS estimate is increased for the treatment arm compared with the parametric curve originally fitted to trial data. The manufacturer states that the magnitude of the multiplication factor is in line with the Beauchemin *at al*[45] review of metastatic breast cancer which drew conclusions about the extent to which differences in median PFS would translate into differences in median OS.

After 48 months, age related mortality calculated from Office for National Statistics data[26] is applied to all those patients alive in each cycle, but only to the everolimus + exemestane arm.

Progression Free Survival

The approach used to estimate PFS was similar to that used to estimate OS. Independent central assessment data from the BOLERO-2[24] trial were available up to a median follow-up of 18 months. Parametric curves were fitted to Kaplan Meier analysis data using the exponential, Weibull, log-logistic and Gompertz functions. The curve generated from the log-logistic function was found, statistically, to be the best fit; however, the Weibull function was used in the base case. The manufacturer explained that this choice was made based on guidance received from clinical experts and visual inspection of the fit to the trial data.

PFS and OS for the other comparators

Tamoxifen

The manufacturer reports that PFS and OS data were extracted from the TAMRAD[25] trial.

Fulvestrant

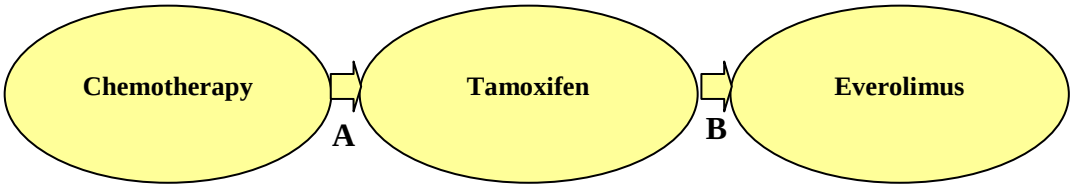
Since there were no head-to-head trials comparing everolimus + exemestane with fulvestrant the manufacturer derived hazard ratios from the mixed treatment comparison analysis (see section 0).

Chemotherapy

The manufacturer assumed a 'class effect' for the three chemotherapy treatments in the analysis, i.e. all drugs are equally efficacious for this group of patients in both the PFS and OS settings.

In the model, the manufacturer has assumed that PFS for patients receiving chemotherapy is the same as that for patients receiving tamoxifen. No justification is provided to support this assumption, although, in their response to a query raised in the clarification letter the manufacturer argued that this approach was preferable to conducting a mixed treatment comparison which would have required a number of additional assumptions (see section 4.3.3).

A search of the literature failed to identify any studies comparing endocrine therapy alone with chemotherapy alone for patients with metastatic breast cancer. It did however identify one systematic review (Wilcken *at al* 2003[32]) which considers patients with metastatic breast cancer and compares OS for patients receiving endocrine therapy with that of patients receiving chemotherapy. The trials included in this review are old (published between 1963 and 1995) and small (median 70 participants, range 50 to 226 women). The hazard ratio of 0.94 reported in the study for endocrine therapy vs chemotherapy is derived from a meta-analysis of six trials, only two of which compare a regimen including doxorubicin with tamoxifen alone. Furthermore, the 95% CI for this hazard ratio includes 1 (0.79 – 1.12) and it is, therefore, not possible to conclude from the review that the efficacy of endocrine and chemotherapies are different. Despite these limitations this figure is used by the manufacturer in a 'naïve chained indirect analysis' to represent the effectiveness of tamoxifen compared with chemotherapy (see Figure 6).



(A) Wilcken *at al*:[32] hazard ratio for chemotherapy versus tamoxifen = 0.94
(B) TAMRAD:[25] hazard ratio for tamoxifen versus everolimus = 1/0.45 = 2.22 (0.45 = hazard ratio for everolimus vs tamoxifen)
Naïve chained comparison of chemotherapy versus everolimus (A x B) = 0.94 x 2.22 = 2.09.

Figure 6 Naïve chained comparison of chemotherapy with everolimus

All the hazard ratios used in the economic model are presented in Table 26.

Table 26 Hazard ratios used in the model

Treatment	OS	PFS	Source
Everolimus in combination with exemestane	1.00	1.00	
Tamoxifen	N/A	N/A	
Fulvestrant	[redacted]	[redacted]	Indirect comparison
Doxorubicin	1.85	2.09	Naïve comparison using information from Wilcken *at al*[32] and the TAMRAD[25] trial
Capecitabine	1.85	2.09	
Docetaxel	1.85	2.09	

Note: A hazard ratio >1 indicates worse effectiveness of the comparator treatment compared with everolimus in combination with exemestane

5.2.7 Health related quality of life

The manufacturer carried out a search of the literature to identify studies reporting utility values for postmenopausal women with HR+ advanced (including locally advanced) or metastatic breast cancer receiving second, and subsequent, lines of therapy. The manufacturer concluded that all of the studies identified had limitations and chose to base the figures used on values reported in previous STAs (fulvestrant[30] and eribulin[44]). The figures used in the model are 0.798 for PFS and 0.496 for progressed disease. Unfortunately, the calculations to arrive at values for PFS are incorrect (see section 5.3.9).

In the base case analysis the AEs are not included in the model. The manufacturer states that very few AEs were reported in the BOLERO-2[24] trial and that excluding AEs is consistent with the ERG's report that informed the appraisal of fulvestrant,[30] where the disutilities and associated costs of AEs were excluded from the model on the basis that very few AEs were reported, and no significant differences were found between therapies. Given AEs were relatively common in BOLERO-2[24] and differed by treatment arm (see section 4.4) the ERG considers this argument to be unsustainable and analyses relating to the disutility associated with AEs may be found in section 5.3.4.

5.2.8 Resources and costs

Drug costs

Drug costs are calculated using cost information extracted from British National Formulary 63 (BNF63).[46] In the base case, drug costs are multiplied by the proportion of patients in the relevant arm of the BOLERO-2[24] trial who receive the drug each month. In the scenario analyses, the proportion of patients who receive the comparator drug each month is assumed to be equal to the proportion of patients in the BOLERO-2[24] trial who received exemestane each month.

Dosing information is taken from published BOLERO-2[24] trial data and BNF63.[46] Administration costs are only applied to those treatments delivered by injection or infusion. Pharmacy costs are not included.

Intervention and comparator costs (taken directly from the manufacturer's model) are summarised in Table 27.

Table 27 Intervention and comparator drug costs in the cost-effectiveness analysis

Treatment	Dose (mg)	Dose intensity	Unit size (mg per unit)	Pack size (no. units)	Cost per pack	Cost per unit	Units required per dose	Cost per dose	Doses per month	Cost per month	Cost for initial administration and dose	Sub-sequent admin-istrations per month	Cost per subse-quent admin-istration	Total cost of admin-istration per month
Everolimus	10.0	76.40%	10	30	£2,970.00	£99.00	1.0	£99.00	30.44	£2,302.17				
Exemestane	25.0	98.36%	25	90	£266.40	£2.96	1.0	£2.96	30.44	£88.62				
Everolimus + exemestane										£2,390.79				£0.00
Exemestane	25.0	99.52%	25	90	£266.40	£2.96	1.0	£2.96	30.44	£89.66				£0.00
Tamoxifen	20.0	76.40%	20	30	£2.95	£0.10	1.0	£0.10	30.44	£2.29				£0.00
Doxorubicin	129.1	76.40%	50	1	£96.86	£96.86	3.0	£290.58	1.45	£739.36	£251.60	1.45	£288.11	£417.59
Fulvestrant	500.0	76.40%	500	2	£522.41	£261.21	2.0	£522.41	1.00	£687.23	£899.41	1.00	£288.11	£288.11
Capecitabine	60231.8	76.40%	500	120	£265.55	£2.21	120.5	£266.58	1.45	£295.19				£0.00
Docetaxel	172.1	76.40%	80	1	£508.01	£508.01	3.0	£1,524.03	1.45	£2,105.22	£251.60	1.45	£288.11	£417.59

Health care costs

Monthly supportive care costs during the PFS (stable disease) and PPS (progressed disease) phases are stated to be as described in NICE clinical guideline for advanced breast cancer[15] 'Package 1' and 'Package 2' respectively, and terminal care costs are reported to have been extracted from the same source. Details of these costs are displayed in Table 28.

Table 28 Costs associated with different health states

Health state	Items	Value (per month)	Reference in submission
Stable disease	Community nurse: home visit 20 minutes	£46.38	NICE CG81,[11] PSSRU Unit Costs
	GP contact: 1 surgery visit	£36.00	NICE CG81[11]
	Clinical nurse specialist: 1 hour contact time	£82.00	PSSRU Unit Costs
	Social worker: 1 hour	£38.00	NICE CG81[11]
	Total	**£202.38**	
Progressed disease	Community nurse home visits	£92.76	NICE CG81[11]
	Clinical nurse specialist	£356.55	PSSRU Unit Costs
	GP contact: 1 home visit	£263.07	NICE CG81[11]
	Therapist: 1 hour	£89.90	PSSRU Unit Costs
	Total	**£802.28**	
Death	Terminal care costs	**£3785.00**	NICE CG81,[11] PSSRU Unit Costs

Adverse event costs

Adverse event costs (and disutilities) are only incorporated into the model as sensitivity analyses.

5.2.9 Cost-effectiveness results

The base case incremental results generated by the manufacturer's model for everolimus + exemestane compared with exemestane alone are presented in Table 29. The ICER for this comparison is £32,417 per QALY gained and £22,486 per life year gained. Figures associated with life-years are not included in the MS[1] but have been taken directly from the manufacturer's model. A summary of the corresponding predicted resource use by category of cost is presented in Table 30.

Table 31 includes summary results for the base case comparison of everolimus + exemestane with the other comparator drugs identified by the manufacturer. The results in this table have been taken from the MS[1] (Tables B53, B55 (p178-181) and Table B56 (p185)). Results displayed in Table 32 were generated, by the ERG, from the manufacturer's model. It is evident that results for only three of the comparisons (everolimus + exemstane with exemestane alone, with tamoxifen (based on TAMRAD data) and with fulvestrant are the same in both Table 31 and Table 32).

Table 29 Base-case results

Technologies	Total costs	Total LY	Total QALYs	Inc. costs	Inc. LY	Inc. QALYs	ICER (cost/LY)	ICER (cost/ QALY)
Everolimus + Exemestane	£48,821	3.82	2.142					
Exemestane	£21,736	2.41	1.306	£27,086	1.41	0.836	£22,486	£32,417

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

Unit Cost	Cost Everolimus + exemestane	Cost exemestane	Increment	Absolute increment	% absolute increment
Technology cost	£22,074	£628	£21,446	£21,446	77.5%
PFS	£3,349	£1,436	£1,913	£1,913	6.9%
PSS	£20,199	£16,186	£4,013	£4,013	14.5%
Terminal care	£3,200	£3,486	-£286	£286	1.0%
AEs	£0	£0	£0	£0	0.0%
Total	**£48,821**	**£21,736**	**£27,086**	**£27,658**	**100%**

Table 31 Summary results for all comparator therapies (extracted from the MS)

Technologies	Total costs	Total QALYs	Inc. costs	Inc. QALYs	ICER (cost/ QALY)
Exemestane vs					
Everolimus	£48,821	2.142			
Exemestane	£21,736	1.306	£27,086	0.836	£32,417
Tamoxifen vs					
Everolimus	£58,231	2.658			
Tamoxifen	£24,065	1.481	£34,256	1.18	£29,109
Fulvestrant vs					
Everolimus	£48,821	2.142			
Fulvestrant	£27,885	1.371	£20,937	0.77	£27,147
Chemotherapy vs					
Everolimus	£48,913	2.119			
Capecitabine	£19,317	0.904	£29,597	1.22	£24,362
Doxorubicin	£23,687	0.874	£25,227	1.25	£20,253
Docetaxel	£35,549	0.904	£13,364	1.22	£11,000

Table 32 Summary results for all comparator therapies (generated by the ERG from the manufacturer's model)

Technologies	Total costs	Total QALYs	Inc. costs	Inc. QALYs	ICER (cost/ QALY)
Using everolimus + exemestane treatment effectiveness based on BOLERO-2[24] trial data, vs					
Everolimus	£48,821	2.142			
Exemestane	£21,736	1.306	£27,086	0.84	£32,417
Tamoxifen	£24,065	1.481	£24,756	0.66	£37,446
Fulvestrant	£27,885	1.371	£20,937	0.77	£27,147
Capecitabine	£17,735	1.104	£31,086	1.04	£29,955
Doxorubicin	£21,599	1.104	£27,222	1.04	£26,232
Docetaxel	£33,968	1.104	£14,854	1.04	£14,313
Using everolimus + exemestane treatment effectiveness based on TAMRAD[25] trial data, vs					
Everolimus	£58,321	2.658			
Tamoxifen	£24,065	1.481	£34,256	1.18	£29,109

5.2.10 Sensitivity analyses

Sensitivity analyses

The manufacturer undertook a wide range of sensitivity analyses. Results of their deterministic sensitivity analyses are not included in the MS.[1] The figures in Table 33 have, therefore, been generated from the model by the ERG. They show that the ICER/QALY values range from £20,368 to £98,640. It is noted that in three cases (Fixed PPS (4 to 48 months, AE: unknown cost assumption (£25; £200) and AE: unknown disutility assumption (-0.01;-0.10)) the parameter variation has no effect on the value of the baseline ICER.

Table 33 Deterministic univariate sensitivity analysis results (everolimus + exemestane compared with exemestane alone)

Parameter	Base case value (£32,417)	Low value		High Value	
		Value	ICER/ QALY	Value	ICER/ QALY
PFS: Everolimus + exemestane (-50% ; +50%)		50%	£20,367.77	150%	£37,318.00
PFS: exemestane (-50% ; +50%)		50%	£50,002.62	150%	£28,503.75
OS: Everolimus + exemestane (-50% ; +50%)		50%	£27,060.41	150%	£68,061.74
OS: exemestane (-50% ; +50%)		50%	£71,733.77	150%	£29,175.64
Fixed PPS (6 - 48 months)	12	6.00	£32,417.09	48.00	£32,417.09
Utility: stable (0.36 ; 0.90)	0.80	0.36	£55,220.35	0.90	£29,573.15
Background costs: PFS (£50 ; £500)	£202.38	£50	£29,399.25	£500	£38,311.32
Background costs: PPS (£500 ; £3000)	£802.28	£500	£23,308.64	£3,000	£98,640.34
AE costs: Everolimus + exemestane (£38 ; £133)	£0.00	£38	£32,462.43	£133	£32,575.79
AE costs: exemestane (£10 ; £34)	£0.00	£10	£32,405.32	£34	£32,375.89
AE disutilities: Everolimus + exemestane (-0.01 ; -0.04)	0.000	-0.01	£32,861.72	-0.04	£34,028.54
AE disutilities: exemestane (0.00 ; -0.01)	0.000	0.00	£32,340.65	-0.01	£32,151.10
AE: unknown cost assumption (£25 ; £200)	£100.00	£25	£32,417.09	£200	£32,417.09
AE: unknown disutility assumption (-0.01 ; -0.10)	-0.050	-0.013	£32,417.09	-0.100	£32,417.09
Fixed PPS applied	Without	-	-	With	£33,731.23

Scenario analyse

Results from the scenario analyses carried out by the manufacturer are presented in Table 34.

The presented scenarios that have the greatest effect on the ICER/QALY are using OS as per BOLERO-2[24] trial (increase of £5,302/QALY gained), and the inclusion of lost productivity costs (decrease of £12,124/QALY gained).

Table 34 Results from scenario analyses (everolimus + exemestane compared with exemestane alone)

Parameter	Incremental costs	Incremental QALYs	ICER/QALY
Base case	£27,086	0.84	£32,417
OS as per BOLERO-2[24] trial	£22,670	0.60	£37,719
Same PPS for treatment and comparator	£25,912	0.77	£33,731
Age adjusted Lloyd utilities	£27,068	0.81	£33,303
6% discount for costs & 0% discount for benefits	£25,796	0.95	£27,050
Include costs of lost productivity	£16,955	0.84	£20,293
PFS measured by local assessment	£25,371	0.73	£34,684
OS as per TAMRAD[25] trial	£38,346	1.44	£26,697
Log-logistic function for PFS and OS from BOLERO-2[24]	£23,685	0.90	£26,329

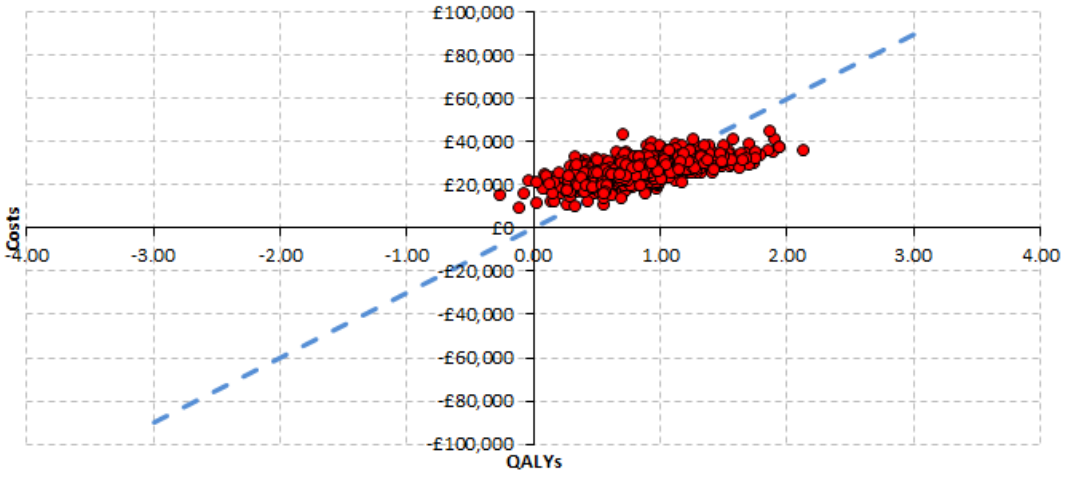
Probabilistic sensitivity analysis

The manufacturer also undertook probabilistic sensitivity analysis (PSA) to derive the mean ICER per QALY of everolimus + exemestane compared with exemestane alone. The distributions used in the PSA are summarised in Table 35.

Table 35 Distributions used in the cost-effectiveness PSA

Parameter	Mean	Standard error	Alpha	Beta	Distribution type
Utilities					
Utility: stable	0.80	0.2	2.418	0.612	Beta
Utility: progressed	0.50	0.2	2.606	2.644	Beta
Cost (treatment)					
Cost: stable	£202	£40	£25	8.095	Gamma
Cost: progressed	£802	£160	25.000	32.091	Gamma
Cost (comparator)					
Cost: stable	£202	£40	£25	8.095	Gamma
Cost: progressed	£802	£160	25.000	32.091	Gamma
Effectiveness					
RR - Tx PFS	100%	10%	n/a	n/a	Lognormal
RR - Tx OS	80%	10%	n/a	n/a	Lognormal
RR - Cx PFS	100%	10%	n/a	n/a	Lognormal
RR - Cx OS	100%	10%	n/a	n/a	Lognormal

The manufacturer's PSA results suggest that there is a 41.6% chance that the ICER for everolimus + exemestane compared with exemestane alone is less than £30,000 per QALY. A scatter plot (incremental cost vs QALY) and a cost-effectiveness acceptability curve are included in the MS[1] and reproduced in Figure 7 and Figure 8 respectively.



1000/1000

Treatment	Everolimus + exemestane	Incremental costs	£26,806
Comparator	Exemestane	Incremental QALYs	0.848
Time horizon	10 years	ICER	£31,614
Sub-group	All patients		
		% cost-effective	41.60%

Figure 7 Cost-effectiveness plane for the comparison of everolimus + exemestane with exemestane alone: Willingness to pay £30,000

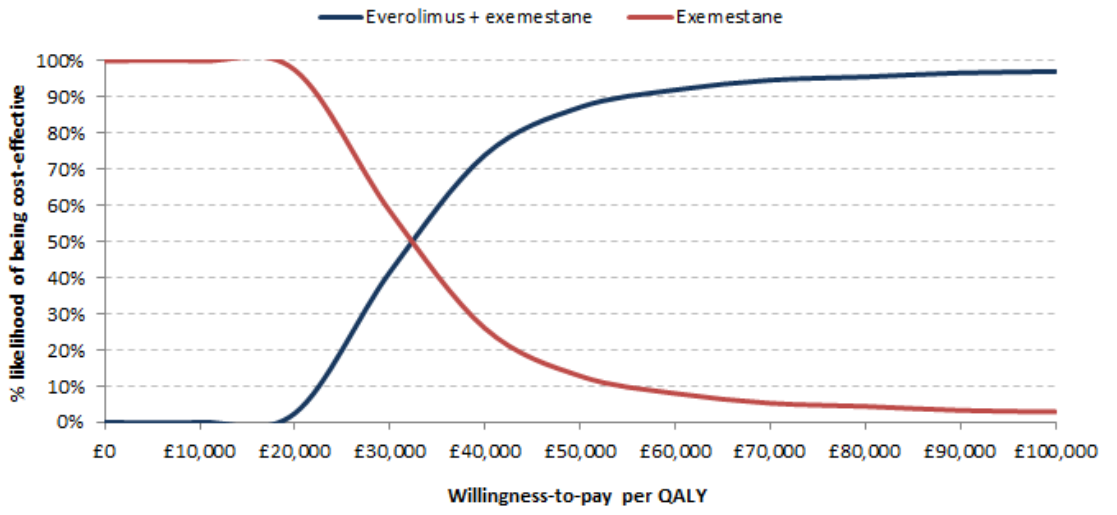


Figure 8 Cost-effectiveness acceptability curve (everolimus + exemestane compared with exemestane alone

The ERG notes that the cost-effectiveness acceptability curve provided in the MS[1] (reproduced in Figure 2) includes only a binary comparison of everolimus + exemestane and exemestane monotherapy. However, when there are two or more comparators, a true representation of relative

cost-effectiveness can only be achieved by generating a cost-effectiveness acceptability plot that includes all comparators. A request for such a plot was included in the clarification letter. In their response, the manufacturer explained that as the model was not designed in a way that allows all comparators to be assessed in one run of the model, a multi-comparator cost-effectiveness acceptability curve would have significant limitations and hence one was not generated.

5.2.11 Model validation and face validity check

The manufacturer reports that the model was subjected to a rigorous 'pressure test' to identify potential errors. Internal validation was undertaken by varying an extensive list of inputs and comparing the impact against expected results. In addition, detailed testing of the model's formulae and functionality was undertaken. A summary of the tests conducted is provided in the MS[1] (Appendix 14, pp250-251).

5.3 Exploratory and sensitivity analyses undertaken by the ERG

The exploratory and sensitivity analyses undertaken by the ERG focus on the comparison of everolimus + exemestane with exemestane alone. The reasons for this are that since tamoxifen is not an AI, the ERG does not believe that the TAMRAD[25] trial is directly relevant to the decision problem. Furthermore, the methods used to derive hazard ratios to allow everolimus to be compared with fulvestrant, capecitabine, doxorubicin and docetaxel result from analyses which, at best, should be viewed with caution (see sections 4.3.3 and 4.7). The ICERs derived using these hazard ratios cannot, therefore, be considered reliable.

The ERG found the manufacturer's submitted model difficult to navigate. The user instructions provided are brief and enigmatic. The flow of logic between and within worksheets is not obvious. Furthermore, the layout of worksheets frequently appears poorly structured with additional columns introduced without clear explanation of content or purpose. Key control parameters appear in different locations on each sheet, sometimes collected into input sheets and sometimes appearing in unexpected locations. There is little labelling, with much of the little provided being uninformative. In some places it is clear that potentially important functions have been designed and then abandoned, or partially deleted (e.g. subgroup analysis). Particular difficulty was experienced in deciphering cell formulae, which make no use of range labels.

The ERG has identified a number of logical errors and questionable assumptions relating to the model and these areas are addressed in sections 5.3.1 to 5.3.13 . However, in view of the difficulties described in the previous paragraph, the ERG cannot be confident that all issues have been identified.

5.3.1 Access to analyses of the BOLERO-2 clinical trial

Following receipt of the MS,[1] the ERG submitted requests for specified Kaplan-Meier analyses of the latest BOLERO-2[24] trial data, to assess the extent to which the submitted decision model accurately reflects the experience of patients in the trial. This involved results of PFS, PPS and OS relating to the whole trial population, and then for three mutually exclusive subgroups: patients with bone-only metastases, patients with visceral metastases, and patients with non-visceral metastases excluding bone-only metastases (see also section 5.3.12, Table 37).

The ERG considered the first request to be necessary in view of the substantial amount of survival gain arising after patients have suffered disease progression, and the need to assess whether such a claim is consistent with the trial evidence. The ERG consider that having access to consistent evidence for OS and its components is essential to enable a definitive conclusion to be reached on this question, which is one of the most influential factors governing the estimated ICER.

The second request concerns the need to assess relative cost-effectiveness in three important subgroups (described in detail in section 5.3.11 below), where there is strong published evidence suggesting large differences which are likely to be influential in addressing the decision problem.

The manufacturer's local UK representatives were unable to fulfil these requests directly and referred the requests to the company's international centre. After considerable delay, a negative response was received. The ERG considers the justification offered by the manufacturer for refusing these requests to be unhelpful and in most respects ill-founded. The consequences of the manufacturer's refusal to accommodate these requests are detailed in the following sections of the report.

5.3.2 Wastage of oral medication

In the manufacturer's model the cost of everolimus + exemestane is calculated on the basis of the average number of patients in the stable health state in each period (month). This does not take account of wastage of oral drugs which occurs when a patient suffers a progression event during the course of the period. Since a monthly supply of tablets will be dispensed at the beginning of each month, and any unused tablets must be disposed of, the correct cost of treatment requires that the

calculation be based on the number of patients in PFS at the start of each month. Although some patients may require dose reductions (e.g. from 10mg to 5mg everolimus) or may miss doses, this is already incorporated in an average dose intensity figure applied to reduce the drug total cost in the model. The more accurate method of calculating medication costs, which includes wastage, results in an increase in incremental cost per patient of £581, and an increase in the manufacturer's base case ICER for everolimus + exemestane compared with exemestane alone of £696 per QALY from £32,417/QALY to £33,113/QALY.

5.3.3 Adjustment for Time on Treatment

A multiplication factor is applied to the cost of systemic treatment within the manufacturer's model. The purpose of this is to account for patients who discontinue treatment prematurely, due to intolerance to AEs or other personal factors, but remain in a non-progressive state. The modellers have drawn on BOLERO-2[24] data for time on treatment for the first 7 months of the trial, and fitted simple linear regression equations to the available data. The resulting equations are then applied indefinitely until all patients have progressed. This leads to an anomalous result in that all everolimus + exemestane patients are off treatment by month 32, but some exemestane only patients are continuing on treatment for more than 10 years, despite the clear advantage the patients receiving combination therapy gain in terms of additional PFS.

Examination of the 7-month 'time on treatment' data (Figure 9 reproduced from the manufacturer's clarification response) suggests that by month 5 a stable situation has been reached whereby almost all PFS patients randomised to exemestane are continuing on treatment, and about 80% of everolimus + exemestane PFS patients remain on treatment. This is in line with the expectation that the greatest effect on patients choosing to withdraw early from treatment is related to AEs which are likely to occur early in the trial, whereas those who find AEs tolerable are likely to persist with treatment. If these alternative approximate figures are applied to the model, the cost of treatment in the everolimus + exemestane arm increases substantially as does the ICER. However, this serves only to indicate the substantial uncertainty associated with estimating treatment costs.

The ERG requested detailed BOLERO-2[24] data from the manufacturer for time on treatment, in order to obtain direct evidence of the time patients are considered to be on medication. This is a more reliable approach to costing than that employed in the model, which first estimates those patients still in the stable (non-progressed) state and then applies an estimated time-varying proportion of these remaining actively on treatment. Figure 10 shows Kaplan-Meier analyses of the time on treatment data, together with exponential projective models fitted by the ERG to allow long-term estimates of patients continuing to receive treatment. When these figures are applied to the model, drug costs in both arms are reduced and the base case ICER for everolimus + exemestane compared with exemestane alone also reduces by £1,616 per QALY from £32,417/QALY to £30,801/QALY. This

amendment automatically supercedes the wastage calculations described above (5.3.2) which are therefore not considered separately hereafter.

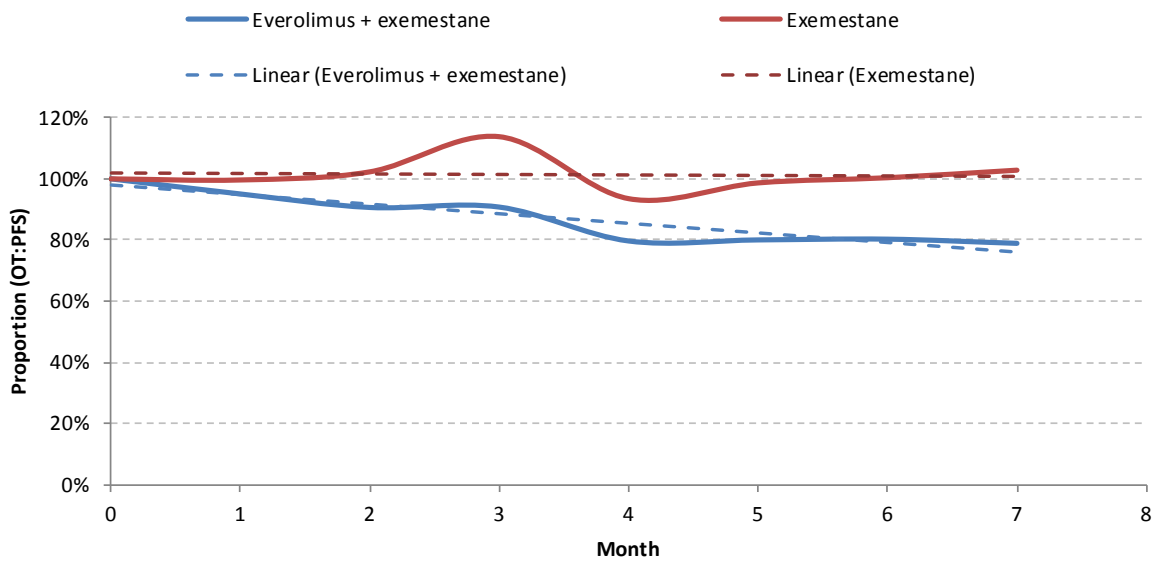


Figure 9 Proportion of progression-free survival patients on treatment each month in BOLERO-2 trial

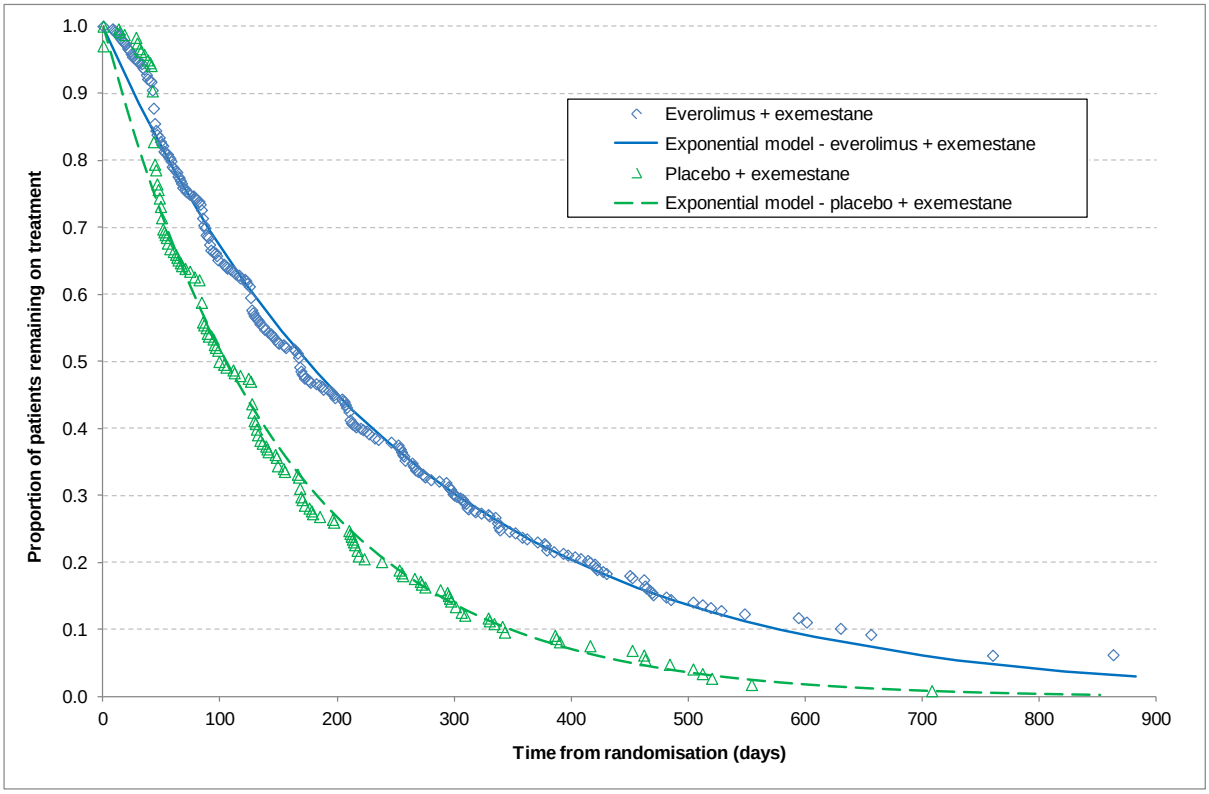


Figure 10 Proportion of patients remaining on treatment in BOLERO-2 trial: Kaplan-Meier analysis and ERG fitted exponential models

5.3.4 Adverse events

The manufacturer's submission describes the approach taken to estimate the disutilities associated with a range of treatment-related AEs. However, in section 7.4.7 of the manufacturer's submission it is argued that since the ERG report relating to the appraisal of fulvestrant (TA239[30]) considered that AEs were sufficiently unimportant to merit inclusion, then AE costs should also be omitted from this appraisal.

The ERG considers this argument to be unsustainable, since the relevance of a particular factor to the decision problem is context specific and depends upon the specific AEs involved, the trial data values for those events, the relative costs and disutilities associated with those events, and the likely influence of the AE data on the estimated ICER in the context of other issues to be considered in the appraisal.

In the manufacturer's base case both the costs and disutilities of AEs are excluded. The ERG considers that, for this appraisal, this is not justifiable. Furthermore, the ERG has compared the frequencies of AEs included in the model with drug-related AEs made available to the EMA and shown as Table 41 in the EMA CHMP EPAR[31] for everolimus + exemestane in treating metastatic breast cancer.[31] As there appear to be several discrepancies, and potentially important omissions from the limited list (10 in total) of AEs considered in the base case scenario, the ERG has carried out analyses using the EMA CHMP EPAR[31] values for all Grade 3/4 AE incidence rates including an additional five types of event. On this basis the incremental cost per patient increases by £142, and QALYs per patient reduce by 0.029 so that the estimated base case ICER for everolimus + exemestane compared with exemestane alone increases by £1,324 per QALY from £32,417/QALY to £33,742/QALY.

There is additional uncertainty in the calculation of unit hospital treatment costs associated with AE episodes. The calculations could only be verified from the descriptions provided using the original sources (NHS Reference Costs for 2009-10[47] or 2010-11[48]) for two of the nine values included in the model.

5.3.5 Hospital monitoring/assessment

The manufacturer's submitted model does not include any costs for regular assessment of response to treatment / disease progression whilst patients remain in the stable health state. For consistency with earlier appraisal the ERG has employed the same approach used in the appraisal of bevacizumab + capecitabine in metastatic breast cancer (TA263).[49] This assumes a response assessment every 3 months involving a face to face meeting with an Oncologist (NHS Reference Cost[48] code 800 - Consultant led follow-up attendance, non-admitted, face to face (clinical oncology)) and a CT scan (NHS Reference Cost[48] RA12Z - outpatient CT Scan (2 areas with contrast)) at a total cost of £254.52

(£121.53+£132.99) per response assessment. When this additional cost is introduced the incremental cost per patient in the everolimus + exemestane arm increases by £798 per patient, and the base case ICER for everolimus + exemestane compared with exemestane alone increases by £955 per QALY from £32,417/QALY to £33,372/QALY.

However, this alteration assumes that the frequency of follow-up does not differ by treatment. Comments submitted by the Royal College of Physicians state:

> 'The comparator arm of exemestane is a well-tolerated treatment, and when given as monotherapy patients may only be seen in an outpatient clinic once every 3 months. Some patients may also be receiving an IV [intravenous] bisphosphonates and therefore seen monthly. The implementation of this treatment (with the side-effect profile discussed) will therefore take up additional clinic capacity, with patients requiring more regular visits, blood tests and assessments. Because the treatment is more intensive than endocrine therapy alone, it is likely that more radiological assessments will be required to document objective responses/stabilisation to justify continued treatment, and to investigate for complications such as pneumonitis.
>
> However it is also important to recognise that this treatment might delay the need for chemotherapy (which would normally be given at progression on exemestane alone) and might even replace chemotherapy for some patients. This could result in resource saving in terms of chemotherapy nursing, outpatient clinic and day unit time.'

In view of these competing possibilities, the ERG has not felt it appropriate to apply any differential follow-up costs to the model, though it should be noted that this is a cautious decision as some net increase in NHS resources may well occur from use of everolimus in combination with exemestane.

5.3.6 Hazard ratio adjustment to everolimus + exemestane projective model

The manufacturer's base case analysis includes an amendment based on a recent conference poster considering correlations between median PFS and median OS in published trials of systemic treatments for metastatic breast cancer. This study indicates an average linear relationship for predicting OS from estimated PFS. To replicate this relationship the OS projective models for everolimus + exemestane based on the BOLERO-2[24] data have been altered by reducing the fitted risk variable by 20%. This results in an increased OS estimate in the everolimus + exemestane arm, exhibited by an enhanced apparent gain in PPS for patients receiving everolimus.

There are several problems with this alteration:

- An average relationship obtained across a heterogeneous selection of 144 trials does not provide a better estimate of OS than that obtained directly from a single well conducted trial; at best it provides a very crude indicator of a range of possible values in cases where no data on OS are available;
- The quoted study is based on meta-analysing median values, which is inherently questionable (a weighted average of means can be meaningfully calculated, but averaging medians does not produce an overall median and is uninterpretable), particularly when the result of interest in modelling is the long-term mean OS;
- The altered projective model of OS in the everolimus + exemestane arm no longer bears any relationship to the BOLERO-2[24] trial data against which the projective models were calibrated;
- Applying such an alteration to the intervention arm whilst leaving the comparator arm unchanged ensures that the model results are necessarily biased in favour of everolimus + exemestane.

The ERG is, therefore, of the view that this post-hoc alteration is without any merit and should be disregarded. Reverting to the original unadjusted base case scenario generates a loss of all outcome gains during the post-progression period with a corresponding reduction in incremental post-progression costs, so that the base case ICER for everolimus + exemestane compared with exemestane alone increases by £5,302 per QALY from £32,417/QALY to £37,719/QALY.

5.3.7 Discounting costs and outcomes

Costs and outcome are discounted in the submitted model on a continuous monthly basis from the time of randomisation. It is conventional in the UK to discount annually (i.e. no discounting in the first year, followed by use of a single discount factor for each successive twelve month period) to match the annual publication of price base information (e.g. NHS Reference Costs[48]) and the annual setting of budgets. Amending the method of discounting in this way leads to minor alterations which increase both incremental costs and QALYs, so that the base case ICER for everolimus + exemestane compared with exemestane alone is reduced by less than £100 per QALY in the base case scenario from £32,417/QALY to £32,326/QALY.

5.3.8 Use of background mortality rate

When estimating future survival the fitted parametric functions are employed in the manufacturer's model throughout the time horizon of the analysis (10 years in the base case scenario). However, after 4 years an additional multiplier is introduced based on the average monthly mortality rate in the overall female population of the same age. Thus for the last 6 years of the analysis the modelled overall mortality observed in the BOLERO-2[24] trial (including causes of death other than breast cancer) is applied as well as the overall population mortality rate, so that deaths from causes other than breast cancer are double-counted. If the chosen parametric function were found to result in long-term mortality rates lower than those of the general female population, then it may be appropriate to replace the modelled estimate by the general population rate on the grounds that it is unrealistic to

expect patients with metastatic disease to achieve better long-term outcomes than others without metastatic disease. However, this is not the case in this model, so the chosen parametric function should be used without modification. At the time of introduction of this modification (4 years) virtually all patients in the comparator arm are projected to have died, so the inflated mortality rate is applied only to the everolimus + exemestane arm leading to incorrectly reduced estimated outcome gains and corresponding under-estimated long-term care costs. When this adjustment is removed the base case scenario ICER for everolimus + exemestane compared with exemestane alone reduces by a small amount (£168 per QALY) from £32,417/QALY to £32.248/QALY.

5.3.9 Calculation of utility values

The manufacturer's model employs utility values drawn from the implementation of the mixed model analysis results reported by Lloyd *at al*,[50] and previously used by the ERG in both the eribulin[44] and fulvestrant[30] STAs. The calculations used to arrive at values for the PFS state are incorrect, and yield a single utility value to be used in both arms of the decision model. A separate value should have been obtained for each arm reflecting the different levels of objective response to treatment reported in the BOLERO-2[24] trial (12.6% vs 1.7%). The corresponding utilities are then 0.7644 for everolimus + exemestane and 0.7571 for placebo + exemestane, which when applied in the model result in reductions in estimated QALYs, and an increase in the estimated base case ICER for everolimus + exemestane compared with exemestane alone of £881 per QALY from £32,417/QALY to £33,299/QALY.

5.3.10 Progression-free survival and overall survival models and trial data

The submitted model relies on parametric projective survival models fitted to the BOLERO-2[24] trial data. Few details are provided for these models, and the ERG's examination of the model spreadsheets suggests that these may not have been calibrated against the latest data cut (December 2011). Few model diagnostics have been provided and without access to more detail on the underlying data it is not possible for the ERG to validate the projective models used in the base case scenario.

When the original survival models are compared in Table 36 (without use of the questionable hazard ratio adjustment critiqued above) it becomes apparent that the bulk of any survival gain estimated by the model occurs in the pre-progression state, regardless of which projective function is employed. Although, in principle, OS is normally considered the more objective and reliable outcome measure, the maturity of the PFS data and the immaturity of the OS data suggest that most attention should be focused on the analyses of PFS data, provided the lack of survival benefit post-progression can be verified.

Table 36 Incremental health gain estimated by parametric functions in the manufacturer's model (excluding hazard ratio adjustment to overall survival for everolimus + exemestane)

Parametric projective function	Mean incremental life years			Estimated OS for comparator
	PFS	PPS	OS	Life years
Weibull (base case)	+ 0.84	+ 0.02	+ 0.85	2.41
Exponential	+ 0.71	+ 0.07	+ 0.78	3.10
Gompertz	+ 1.34	-0.23	+ 1.11	3.45
Log-logistic	+ 1.31	-0.68	+ 0.62	3.33

For these reasons the ERG requested that the manufacturer provide a consistent set of Kaplan-Meier survival analyses for PFS, PPS and OS data from the BOLERO-2[24] trial, but these have not been made available. Without this information it is not possible for the ERG to determine the reliability of the data included in the submitted decision model, nor is it possible for the ERG to determine with confidence the most appropriate base case projective models.

5.3.11 Local or central assessment

In the manufacturer's base case scenario, PFS parametric models are fitted to data based on PFS data using central assessment of the time of progression. As discussed in Section 4, the ERG considers that local assessments of progression are more appropriate as they more closely align with normal clinical conditions. If a model option to switch to PFS estimates based on local assessment data, both incremental costs and QALY reduce as on average less patient time is assigned to the pre-progression phase. This results in an increase in the estimated ICER of £2,266/QALY from £32,417/QALY to £34,684/QALY.

5.3.12 Subgroups

The manufacturer declined to carry out any subgroup analyses on the grounds that from their subgroup analyses of PFS in BOLERO-2,[24] no specific subgroup can be identified for which the PFS benefit is statistically significantly superior to that of the overall BOLERO-2[24] population. This refers to a comparison of PFS hazard ratios shown in Figure B5 of the manufacturer's submission (page 71), implying that the proportionate effect of treatment with everolimus + exemestane is similar across subgroups. However, this is not a valid argument against considering subgroups in cost-utility analysis, which is driven not by the relative magnitude of health gain (hazard ratios) but the absolute magnitude of gain (extra days of survival or QALYs per patient). So a similar proportionate effect can lead to widely differing treatment gains for patients.

The results of subgroup analysis of PFS in the final BOLERO-2[24] results are reported in the ASCO conference poster by Piccart *et al*[36] and the ESMO conference poster by Campone *at al.*[51] These reveal very different risk profiles when comparing those patients suffering visceral metastases with

those suffering non-visceral metastases. Even more extreme differences are evident when comparing those patients suffering visceral metastases with those suffering only bone metastases. It is clear that the latter subgroup of patients have a substantially better prognosis than those with visceral metastases in line with clinical experience, leading to much greater absolute survival gain when the same relative risk is applied to both groups. As a consequence, it can be expected that, in a cost-utility analysis based on subgroups of the BOLERO-2[24] trial, patients with only bone metastases will exhibit a more advantageous ICER than the overall average population, whereas those with visceral metastases will have a correspondingly poorer ICER for the use of everolimus + exemestane compared with exemestane alone.

Believing that this is an important effect that should be quantified and made available to the Appraisal Committee, the ERG requested that additional survival analysis data (PFS, OS and PPS) be provided for three mutually exclusive subgroups (bone-only metastases, visceral metastases, and non-visceral metastases excluding bone-only). The manufacturer has reported difficulty in obtaining the analyses requested, but has provided printouts of similar analyses, but only relating to PFS. The ERG has carried out parametric modelling on these data to estimate the differences in mean PFS which are attributable to combination treatment with everolimus as shown in Table 37. It is clear that patients with only bone metastases fared much better, and those with visceral metastases much worse, than the overall average experience, and these should be reflected in widely different estimated ICERs. Unfortunately, it is not possible to quantify the ICER differences without access to similar OS and PPS survival data for these subgroups.

Table 37 ERG mean estimated progression-free survival (months) for three subgroups of BOLERO-2 patients

Subgroup	Everolimus + exemestane	Placebo + exemestane	PFS gain
Visceral metastases	8.37	4.87	3.50
Non-visceral metastases (excluding bone-only metastases)*	11.06*	5.70*	5.36*
Bone-only metastases	14.93	6.58	8.35

* 10% of bone-only met patients could not be identified and excluded from this subgroup so some contamination remains

5.3.13 Uncertainty in base case survival estimates

To consider the extent of uncertainty in the estimated base case ICER in the absence of the full survival analyses requested by the ERG, graphical information on PFS has been extracted from Piccart *at al*'s ASCO 2012 poster,[36] and from Figure B6 of the MS[1] for OS. These data have been subjected to exploratory projection modelling. The PFS data is sufficiently mature that the fitted exponential models shown in Figure 11 can be used within the model with confidence. However, the

OS data provided in Figure B6 of the MS[1] does not even allow median OS to be estimated in either arm of the BOLERO-2[24] trial. In Figure 12 parametric models have been fitted by the ERG to the available OS data up to 18 months, and are seen to reflect the observation that from 10 months onwards the two arms appear to be subject to similar mortality risks.

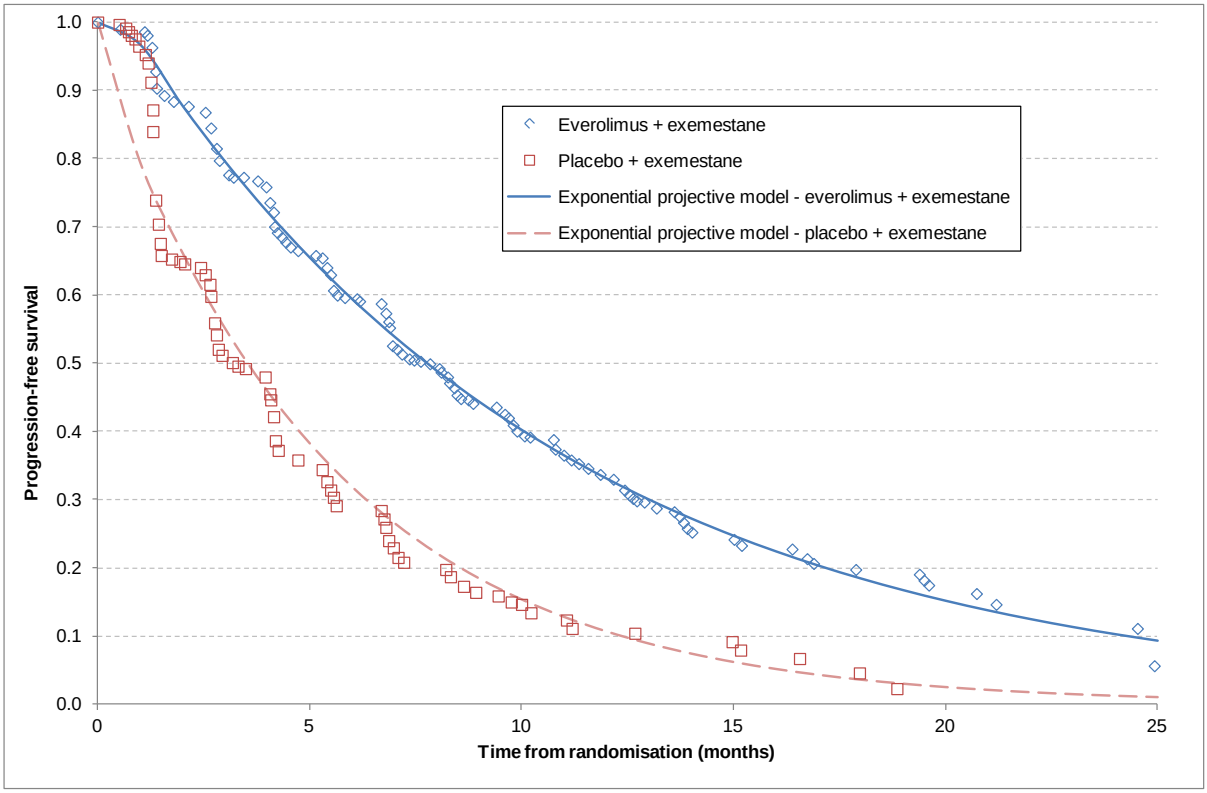


Figure 11 BOLERO-2 progression-free survival Kaplan-Meier results reported in Piccart *at al* poster, with ERG fitted exponential models

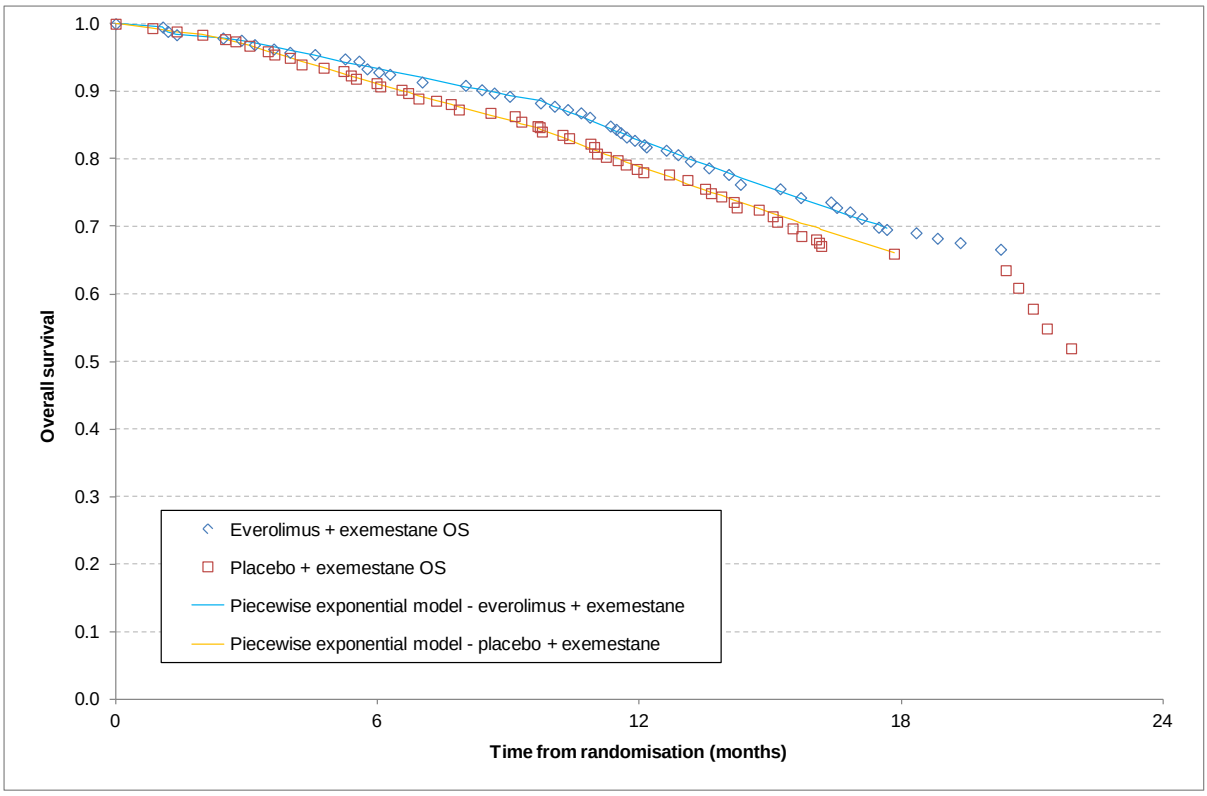


Figure 12 BOLERO-2 progression-free survival Kaplan-Meier results reported in Fig.B6 of MS, with ERG fitted piecewise exponential models

When these survival profiles are applied to the manufacturer's model in place of the original Weibull functions, estimates of both mean PFS and OS are reduced in the combination therapy arm but not in the exemestane arm, resulting in an increase in the base case ICER for everolimus + exemestane compared with exemestane alone from £32,417/QALY to £39,978/QALY.

When these exploratory profiles are applied together with the other amendments described above (but excluding the switch to local assessment which is no longer needed), the combined effect is to increase the base case ICER from £32,417/QALY to £66,476/QALY. Clearly the decision problem is very sensitive to the approach taken to projecting OS for the lifetime of patients, and the extent of this uncertainty can only be reduced by access to more detailed information on survival, especially for OS and PPS.

5.4 Overview and conclusions of the cost-effectiveness section

The manufacturer's base case analysis results in an estimated ICER for everolimus + exemestane compared with exemestane alone which exceeds the NICE reference range for cost-effectiveness (£20,000 - £30,000 per QALY). When the methodology adjustments described in sections 5.3.2 - 5.3.13 are applied, the base case adjusted ICER is increased to £52,285/QALY (Table 37). However, even after these adjustments, there is substantial unresolved uncertainty concerning long-term survival in the model, especially for the post-progression phase. This is wholly attributable to the immaturity

of the trial data, which have not yet continued to the point at which the difference of OS medians can be estimated. Since the model structure depends on establishing reliable methods of estimating *both* PFS and OS beyond the trial data to the end of life, it can reasonably be argued that any modelling undertaken using such immature data should be considered purely exploratory. The ERG has sought to maximise the trial information made available by the manufacturer, and has demonstrated that an ICER for everolimus + exemestane compared with exemestane alone that exceeds £66,000 per QALY can be considered consistent with the current survival evidence.

Additionally, the ERG has examined the evidence of subgroup differences in the PFS data from BOLERO-2[24] and concludes that there are very large differences in the mean PFS between patients with visceral metastases, patients with only bone metastases, and other patients. These differences will inevitably result in a much better ICER for some patients, and a worse ICER for other patients. It must, therefore, be considered very likely that a cost-effectiveness analysis fully informed with subgroup survival data will show that for patients with visceral metastases the estimated ICER is considerably greater than the NICE reference range. The likely outcome for other subgroups is less clear.

The extent of uncertainty in model results may only be reduced by either:

- the provision of OS and/or PPS survival data from the BOLERO-2[24] clinical trial (overall and at subgroup level), or
- results from an independent confirmatory trial of the same intervention in a similar patient population.

6 IMPACT ON THE ICER OF ADDITIONAL CLINICAL AND ECONOMIC ANALYSES UNDERTAKEN BY THE ERG

The individual effects of the model changes recommended by the ERG compared with the manufacturer's base case scenario are shown in Table 38. The two amendments which have the largest influence on the ICER are the removal of a 20% hazard ratio reduction to the modelled OS for everolimus + exemestane, and the replacement of the manufacturer's OS and PFS survival trends with the exploratory models calibrated by the ERG against published graphs. Taken together, the ERG amendments increase the estimated ICER to £52,285 per QALY gained without the exploratory survival models and to £66,476 per QALY gained with the ERG exploratory models.

The large uncertainty associated with the outcomes of the BOLERO-2[24] trial can only be reduced by access to the additional information requested by the ERG, and not provided by the manufacturer. This affects two key issues: the cost-effectiveness of everolimus + exemestane in the whole population, and its relative cost-effectiveness in the three subgroups where substantially different prognosis and benefit have already been demonstrated.

Table 38 Cost and outcome effects of ERG model amendments relative to the manufacturer's base case analysis

Adjustment	Exemestane				Everolimus + Exemestane				Incremental				
	Therapy cost	Other costs	Survival (months)*	QALYs	Therapy cost	Other costs	Survival (months)*	QALYs	Survival (months)*	Cost	QALYs	ICER (£/QALY)	ICER change
Base case	**£628**	**£21,108**	**28.9**	**1.306**	**£22,074**	**£26,748**	**45.9**	**2.142**	**16.9**	**£27,086**	**0.836**	**£32,417**	**-**
a) Include AEs + EPAR data	£628	£21,149	28.9	1.299	£22,074	£26,931	45.9	2.106	16.9	£27,227	0.807	£33,742	+£1,324
b) Correct utility values	£628	£21,108	28.9	1.282	£22,074	£26,748	45.9	2.096	16.9	£27,086	0.813	£33,299	+£881
c) Include monitoring	£628	£21,585	28.9	1.306	£22,074	£28,023	45.9	2.142	16.9	£27,883	0.836	£33,372	+£955
d) No hazard ratio adjustment	£628	£21,108	28.9	1.306	£22,074	£22,332	39.2	1.907	10.3	£22,670	0.601	£37,719	+£5,302
e) Correct discounting	£635	£21,435	28.9	1.325	£22,343	£27,167	45.9	2.174	16.9	£27,440	0.849	£32,326	-£91
f) No background deaths	£628	£21,108	28.9	1.306	£22,074	£26,926	46.2	2.152	17.2	£27,264	0.845	£32,248	-£169
g) Time on treatment	£504	£21,108	28.9	1.306	£20,600	£26,748	45.9	2.142	16.9	£25,736	0.836	£30,801	-£1,616
ERG amendments a - g	**£509**	**£21,839**	**28.9**	**1.293**	**£20,481**	**£23,809**	**39.4**	**1.860**	**10.5**	**£22,302**	**0.567**	**£39,320**	**+£6,903**
h) Local assessment	£458	£22,246	28.9	1.259	£17,706	£30.369	45.9	1.990	16.9	£25.371	0.732	£34,684	+£2,255
i) Exploratory OS & PFS	£628	£21,108	28.9	1.306	£17,636	£29,089	44.2	1.931	15.3	£24,990	0.625	£39,978	+£7,561
ERG amendments a - h	**£509**	**£22,839**	**28.9**	**1.251**	**£20,841**	**£27,244**	**39.4**	**1.723**	**10.5**	**£24,683**	**0.472**	**£52,285**	**+£19,868**
ERG amendments a – g, & i	**£509**	**£21,839**	**28.9**	**1.293**	**£20,841**	**£25,886**	**37.7**	**1.660**	**8.7**	**£24,378**	**0.367**	**£66,476**	**+£34,059**

* survival is undiscounted, all other figures are discounted

7 OVERALL CONCLUSIONS

Clinical evidence is derived from a well conducted RCT (BOLERO-2[24]) that compares the intervention of interest (everolimus + exemestane) to one of the comparators of interest (exemestane). The population of patients included in BOLERO-2[24] is the same group of patients as are specified in the decision problem and for which everolimus has received a marketing licence from the European Union. However:

- The model submitted by the manufacturer is poorly structured and the projective modelling is only loosely associated with data from the BOLERO-2[24] trial
- There is substantial unresolved uncertainty concerning modelled long-term survival (especially OS). This is wholly attributable to the immaturity of the BOLERO-2[24] trial data, which have not yet continued to the point at which the difference of OS medians can be estimated
- The ERG has not been able to verify the projective survival models used in the manufacturer's base case because the manufacturer has not provided the relevant survival data
- Evidence from Piccart *et al*[36] and Campone *at al*[51] suggests that there may be some additional benefit for the subgroup of patients suffering only bone metastases but the manufacturer has not provided data to allow the ERG to explore the full extent of this effect
- Taken together, the ERG amendments to the manufacturer's economic model increase the estimated ICER from £32,417 to either £52,285 or £66,476 per QALY gained.

7.1 Implications for research

Despite published findings from the BOLERO-2[24] trial there is substantial unresolved uncertainty concerning long-term survival for patients receiving everolimus plus exemestane compared with those receiving exemestane alone, especially during the post-progression phase. Additionally, PFS data from BOLERO-2[24] suggest that there are large differences in mean PFS between patients with visceral metastases, patients with only bone metastases, and other patients. The extent of this uncertainty may only be reduced by either:

- the provision of OS and/or PPS survival data from the BOLERO-2[24] clinical trial (overall and at subgroup level), or
- results from an independent confirmatory trial of the same intervention in a similar patient population.

8 REFERENCES

1. Novartis. *National Institute for Health and Clinical Excellence Single Technology Appraisal (STA): Everolimus (Afinitor) in combination with exemestane for the treatment of advanced or metastatic HER2 negative, hormone receptor positive breast cancer after prior endocrine therapy: Specification for manufacturer/sponsor submission of evidence, 22 November 2012*.
2. Perou CM, Sorlie T, Eisen MB. Molecular portraits of human breast tumours. *Nature* 2000;406:747-52.
3. National Cancer Institute. *General information about breast cancer* http://www.cancer.gov/cancertopics/pdq/treatment/breast/healthprofessional/page1 (Accessed: 10 January 2013)
4. Cooke T. What is HER2? *Eur J Oncol Nurs* 2000;4(2-9).
5. Coleman RE, Bertelli G, Beaumont T, Kunkler I, Miles D, Simmonds PD, et al. UK Guidance Document: treatment of metastatic breast cancer. *Clin Oncol* 2011;24:169-76.
6. Singletary SE, Allred C, Ashley P, W. BL, Berry D, Bland KI, et al. Revision of the American Joint Committee on Cancer Staging System for Breast Cancer. *J. Clin. Oncol.* 2002;20:3628-36.
7. Cancer Research UK. *Statistics and outlook for breast cancer* http://www.cancerresearchuk.org/cancer-help/type/breast-cancer/treatment/statistics-and-outlook-for-breast-cancer (Accessed: 10 January 2013)
8. Cancer Research UK. *Symptoms of secondary breast cancer* http://www.cancerresearchuk.org/cancer-help/type/breast-cancer/secondary/about/symptoms-of-secondary-breast-cancer (Accessed: 10 January 2013)
9. Sant M, Allemani C, Capocaccia R, Hakulinen T, Aareleid T, Coebergh JW, et al. Stage at diagnosis is a key explanation of differences in breast cancer survival across Europe. *Int J Cancer* 2003;106(3):416-22.
10. West Midlands Cancer Intelligence Unit. *The second all breast cancer report*. 2011. http://www.wmciu.nhs.uk/publicationdetails.aspx?id=474 (Accessed: 10 January 2013)
11. National Institute for Health and Clinical Excellence (NICE). *Early and locally advanced breast cancer: diagnosis and treatment*. 2009. http://www.nice.org.uk/Guidance/CG80 (Accessed: 11 January 2013)
12. Murray J, Young OE, Renshaw L, White S, Williams L, Evans DB. A randomised study of the effects of letrozole and anastrozole on oestrogen receptor positive breast cancers in postmenopausal women. *Breast Cancer Res Treat* 2009;114:495-501.
13. Datamonitor. *Epidemiology: Breast Cancer* 2010 http://www.datamonitor.com/store/Product/epidemiology_breast_cancer_the_most_common_cancer_in_women_worldwide_and_both_incidence_and_prevalence_are_increasing?productid=HC00041-001 (Accessed: 24 January 2013)
14. Breast Cancer Care. *Secondary Breast Cancer Taskforce Report* 2007.
15. National Institute for Health and Clinical Excellence (NICE). *Advanced breast cancer: diagnosis and treatment*. 2009. http://www.nice.org.uk/Guidance/CG81 (Accessed: 11 January 2013)
16. Cardoso F, Fallowfield L, Costa A, Castiglione M, Senkus E. Locally recurrent or metastatic breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Annals of Oncology* 2011;22(6):vi25-30.
17. National Comprehensive Cancer Network. *NCCN Clinical Practice Guidelines in Oncology: Breast cancer v.1* 2012. http://www.nccn.org/professionals/physician_gls/f_guidelines.asp (Accessed: 24 January 2013)
18. Cardoso F, Costa A, Norton L, Cameron D, Cufer T, Fallowfield L, et al. 1st International consensus guidelines for advanced breast cancer (ABC 1). *Breast* 2012;21:242-52.
19. Moy B, Goss PE. Estrogen receptor pathway: resistance to endocrine therapy and new therapeutic approaches. *Clin Cancer Res* 2006;12:4790-3.

20. Normanno N, Di Maio M, De Maio E, De Luca A, De Matteis A, Giordano A, et al. Mechanisms of endocrine resistance and novel therapeutic strategies in breast cancer. *Endocrine-Related Cancer* 2005;12:721-47.
21. Pink JJ, Bilimoria MM, Assikis J, Jordan VC. Irreversible loss of the oestrogen receptor in T47D breast cancer cells following prolonged oestrogen deprivation. *Br J Cancer* 1996;74:1227-36.
22. Roche H, Vahdat LT. Treatment of metastatic breast cancer: second line and beyond. *Ann Oncol* 2011;22:1000-10.
23. Hurvitz SA, Pietras RJ. Rational management of endocrine resistance in breast cancer. *Cancer* 2008;113:2385-97.
24. Baselga J, Campone M, Piccart M, Burris HA, Rugo HS, Sahmoud T, et al. Everolimus in Postmenopausal Hormone-Receptor–Positive Advanced Breast Cancer. *N Engl J Med* 2012;366(6):520-9.
25. Bachelot T, Bourgier C, Cropet C, Ray-Coquard I, Ferrero J-M, Freyer G, et al. Randomized phase II trial of everolimus in combination with tamoxifen in patients with hormone receptor-positive, human epidermal growth factor receptor 2-negative metastatic breast cancer with prior exposure to aromatase inhibitors: a GINECO study. *J. Clin. Oncol.* 2012;30(22):2718-24.
26. Office of National Statistics. *UK National Statistics Publication Hub*. 2012. http://www.statistics.gov.uk/hub/population/index.html (Accessed: 11 January 2013)
27. National Institute for Health and Clinical Excellence (NICE). *Advanced breast cancer: costing template*. 2009. http://guidance.nice.org.uk/CG81/CostingTemplate/xls/English (Accessed: 11 January 2013)
28. Sharma PP, Sail K, Delea TE, Forsyth MT, Halm M, Y. M. Nagarwala, et al. Patterns of care and patient characteristics in postmenopausal women with HER2+ and hormone receptor-positive (HR+) metastatic breast cancer (MBC). ASCO annual meeting, Abstract e11107. *Journal of Clinical Ocology* 2011;29(suppl; abstr e11107).
29. Novartis Data on file. Afinitor launch planning: treatment algorithm, IMS, 2012.
30. Fleeman N, Bagust A, Boland A, Blundell M, Dickson R, Dundar Y, et al. *Fulvestrant for the treatment of locally advanced or metastatic breast cancer: A Single Technology Appraisal*. 2011. http://www.hta.ac.uk/erg/reports/2528.pdf (Accessed: 13 January 2013)
31. European Medicines Agency Committee for Medicinal Products for Human Use. *Assessment Report For Afinitor (everolimus). Procedure No.: EMEA/H/C/001038/II/0020*. 2012. http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Assessment_Report_-_Variation/human/001038/WC500130788.pdf (Accessed: 11 January 2013)
32. Wilcken N, Hornbuckle J, Ghersi D. Chemotherapy alone versus endocrine therapy alone for metastatic breast cancer. *Cochrane Database Syst Rev* 2003(2):CD002747.
33. The Cochrane Collaboration. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0* http://www.cochrane.org/training/cochrane-handbook (Accessed: 11 January 2013)
34. 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. *Journal of the National Cancer Institute* 2000;92(3):205-16.
35. Chia S, Gradishar W, Mauriac L, Bines J, Amant F, Federico M, et al. Double-blind, randomized placebo controlled trial of fulvestrant compared with exemestane after prior nonsteroidal aromatase inhibitor therapy in postmenopausal women with hormone receptor-positive, advanced breast cancer: results from EFECT. *J. Clin. Oncol.* 2008;26(10):1664-70.
36. Piccart M, Noguchi S, Pritchard K, Burris III H, Rugo H, Gnant M, et al. Everolimus for postmenopausal women with advanced breast cancer: Updated results of the BOLERO-2 phase III trial. *Journal of Clinical Ocology* 2012;30(suppl; abstr 559).
37. Di Leo A, Jerusalem G, Petruzelka L, Torres R, Bondarenko IN, Khasanov R, et al. Results of the CONFIRM phase III trial comparing fulvestrant 250 mg with fulvestrant 500 mg in postmenopausal women with estrogen receptor-positive advanced breast cancer.(Erratum appears in J Clin Oncol. 2011 Jun 1;29(16):2293). *J. Clin. Oncol.* 2010;28(30):4594-600.

38. Johnston S, Kilburn LS, Ellis P, Cameron D, Dodwell D, Howell A, et al. Fulvestrant alone or with concomitant anastrozole vs exemestane following progression on non-steroidal aromatase inhibitor-first results of the sofea trial (CRUKE/03/021 & CRUK/09/007) (ISRCTN44195747). *Eur J Cancer* 2012;48 (SUPPL. 3 (S2)).
39. Chia S, Piccart M, Gradishar W, on behalf of the EFECT writing committee. Fulvestrant vs exemestane following non-steroidal aromatase inhibitor failure: first overall survival data from the EFECT trial. In: Poster presented at the San Antonio Breast Cancer Symposium, Texas, USA, 13-16 December 2007.
40. AstraZeneca UK Ltd. *National Institute for Health and Clinical Excellence Single technology appraisal (STA): Fulvestrant 500 mg for the treatment of postmenopausal women with oestrogen receptor positive, locally advanced or metastatic breast cancer: Specification for manufacturer/sponsor submission of evidence*. 2010. http://www.nice.org.uk/nicemedia/live/12967/55985/55985.pdf (Accessed: 11 January 2013)
41. National Institute for Health and Clinical Excellence. *Breast cancer (metastatic) - fulvestrant: final appraisal determination document* 2011. http://guidance.nice.org.uk/TA/Wave23/15/FAD/FinalAppraisalDetermination/pdf/English (Accessed: 21 January 2013)
42. National Institute for Health and Clinical Excellence. *Guide to the methods of technology appraisal*. 2008. http://www.nice.org.uk/media/B52/A7/TAMethodsGuideUpdatedJune2008.pdf (Accessed: 16 January 2013)
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. Greenhalgh J, Bagust A, Boland A, Oyee J, Trevor N, Dundar Y, et al. *Eribulin for the treatment of advanced or metastatic breast cancer: A Single Technology Appraisal*. 2011. http://www.hta.ac.uk/erg/reports/2504.pdf (Accessed: 16 January 2013)
45. Beauchemin C, Cooper D, Lapierre ME, Yelle L, Lachaine J. PCN33 Progression-free survival as a surrogate for overall survival in metastatic breast cancer. *Value In Health* 2012;15(7):A414.
46. British National Formulary. *BNF 63*. 2012. http://www.bnf.org/bnf/ (Accessed: 16 January 2013)
47. Department of Health. *National schedule of reference costs year 2009–2010* http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_123459 (Accessed: 23 January 2013)
48. Department of Health. *National schedule of reference costs year 2010–2011* http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_131140 (Accessed: 23 January 2013)
49. Fleeman N, Beale S, Davis H, 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*. 2011. http://www.nice.org.uk/nicemedia/live/12100/58821/58821.pdf (Accessed: 13 January 2013)
50. Lloyd A, Nafees B, Narewska J, Dewilde S, Watkins J. Health state utilities for metastatic breast cancer. *Br J Cancer* 2006;95:683-90.
51. Campone M, Noguchi S, Pritchard KI, Rugo HS, Hortobagyi GN, Baselga J, et al. *Efficacy and Safety of Everolimus in Postmenopausal Women With Advanced Breast Cancer (BOLERO-2): Effect of Visceral Metastases (Abstract 324PD). European Society for Medical Oncology Congress; 28 September-2 October 2012, Vienna, Austria.* http://abstracts.webges.com/viewing/view.php?congress=esmo2012&congress_id=370&publication_id=1895 (Accessed: 16 January 2013)

9 APPENDICES

Assessment of risk of bias for studies included in the MS

Assessment of risk of bias: BOLERO-2

Study question	Manufacturer response		ERG comment
	How is the question addressed in the study?	Grade (yes/no/not clear/N/A)	
Was randomisation carried out appropriately?	Yes; patients were assigned a randomisation number by logging onto an interactive web and voice response system. This randomisation number linked the patient to a treatment arm and specified a unique medication number for the packages of study drug to be dispensed to the patient. Randomisation was stratified by documented prior sensitivity to hormonal therapy and the presence of visceral disease.	Yes	Agree
Was the concealment of treatment allocation adequate?	Yes; allocation achieved by interactive web and voice response system.	Yes	Agree
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	Yes: well balanced for all major baseline characteristics.	Yes	Mostly, although there was a slightly greater proportion of patients aged ≥65 in the everolimus arm (40% vs 34%) and a slightly higher proportion in the placebo arm had received their last treatment for metastatic disease (84% vs 79%) or a NSAI for metastatic disease (76% vs 71%)
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)?	Yes: all patients, investigators, assessors, Novartis personnel and individuals at central laboratories were blinded to randomisation data. Treatment identity was concealed by using study drugs that were identical in packaging, labelling, dosing schedule and appearance.	Yes	Agree. However, it is noted that due to the differences in AEs experienced by patients receiving everolimus compared to exemestane, some unblinding may have occurred

Study question	Manufacturer response		ERG comment
	How is the question addressed in the study?	Grade (yes/no/not clear/N/A)	
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	The percentage of patients discontinued from the study was greater in the placebo in combination with exemestane arm. Disease progression, which was more frequent in the placebo in combination with exemestane arm, was the primary reason for discontinuation from the study. Discontinuations due to AEs and consent withdrawal were higher in the everolimus in combination with exemestane arm. Discontinuation imbalance was not explained.	No	Agree
Is there any evidence to suggest that the authors measured more outcomes than they reported?	The BOLERO-2 CSR refers back to the trial protocol and the published interim analysis does not suggest reporting bias.	No	Agree
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	The patients were analysed according to the treatment and stratum they were assigned to at randomisation (i.e. intention to treat). For the primary outcome (PFS), if a patient did not progress or was known to have died at the date of the analysis cut-off or start of another antineoplastic therapy, the PFS date was censored to the date of last adequate tumour assessment prior to cut-off date or start of antineoplastic therapy.	Yes for outcomes measured using Kaplan Meier	Agree

Assessment of risk of bias: CONFIRM

Study question	Manufacturer response		ERG comment
	How is the question addressed in the study?	Grade	
Was randomisation carried out appropriately?	Not reported. The authors stated that patients were randomly assigned in a 1:1 ratio, with patients stratified by institution site	Unclear	Agree
Was the concealment of treatment allocation adequate?	Not reported	Unclear	Agree
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	Generally well-balanced	Yes	Agree
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 is described as double-blind; data monitoring was performed by an independent data monitoring committee; no other details were reported.	Unclear	Placebo injection added for lower dose fulvestrant so that both treatment groups received 2 injections.
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	No	Yes	Agree
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No	Yes	Agree
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	All patients were included in the analyses.	Yes	Agree

Assessment of risk of bias: EFECT

Study question	Manufacturer response		ERG comment
	How is the question addressed in the study?	Grade	
Was randomisation carried out appropriately?	Not reported	Unclear	Agree
Was the concealment of treatment allocation adequate?	Not reported	Unclear	Agree
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	Overall, the groups were well balanced, except that the fulvestrant cohort had a slightly greater number of women with ER+, PgR+ tumours (67.5%) vs the exemestane cohort (56.4%)	Yes	Agree
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 is described as double-blind; no other details are reported	Unclear	Agree
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	Not reported	Unclear	Agree although it is noted withdrawals due to AEs were similar in each treatment arm
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No	Yes	Agree
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	The authors stated that data for the efficacy parameters were analysed and summarised on an intention-to-treat basis	Yes	Agree

Assessment of risk of bias: SoFEA

Study question	Manufacturer response		ERG comment
	How is the question addressed in the study?	Grade	
Was randomisation carried out appropriately?	Not reported	Unclear	Agree.
Was the concealment of treatment allocation adequate?	Not reported	Unclear	Agree.
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	Overall, the groups were well balanced, except that the exemestane group had a slightly lower percentage of women with ER+, PgR- tumours (9.2%) vs the fulvestrant and fulvestrant + anastrazole groups (14.3 and 15.6%)	Yes	Agree.
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)?	Two treatment arms were blinded to the NSAI.	Unclear	Agree.
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	Not reported	Unclear	Agree.
Is there any evidence to suggest that the authors measured more outcomes than they reported?	Not reported	Unclear	Agree.
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	It appears that all data were included in the analyses	Yes	Agree.

Assessment of risk of bias: TAMRAD

Study question	Manufacturer response		ERG comment
	How is the question addressed in the study?	Grade	
Was randomisation carried out appropriately?	Not reported. The authors stated that patients were randomly assigned in a 1:1 ratio, with patients stratified by primary and secondary hormone resistance.	Unclear	Agree
Was the concealment of treatment allocation adequate?	Not reported.	Unclear	Agree
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	Generally well-balanced except for performance status: ECOG PS of 0 was more common in the tamoxifen in combination with everolimus group (59% vs, 40%), while ECOG status of 1 was more common in the tamoxifen group (49% vs 33%).	No	Agree
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)?	No; open-label study.	No	Agree
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	Discontinuations due to adverse effects were greater in the tamoxifen in combination with everolimus group (22% vs 7%).	No	Agree
Is there any evidence to suggest that the authors measured more outcomes than they reported?	There is no obvious evidence of reporting bias.	No	Agree
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	The patients were analysed according to the principle of intention to treat: per protocol results were also reported.	Yes	Agree