STA REPORT

LIVERPOOL REVIEWS AND IMPLEMENTATION GROUP (LRiG)

Palbociclib in combination with an aromatase inhibitor for previously untreated metastatic, hormone receptor-positive, HER2-negative breast cancer [ID915]

Confidential until published

This report was commissioned by the NIHR HTA Programme as project number 15/194/07

Completed 30 November 2016

CONTAINS [redacted] AND [redacted] DATA

UNIVERSITY OF LIVERPOOL

LIVERPOOL REVIEWS AND IMPLEMENTATION GROUP

A MEMBER OF THE RUSSELL GROUP

Title: Palbociclib in combination with an aromatase inhibitor for previously untreated metastatic, hormone receptor-positive, HER2- breast cancer [ID915]

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

Authors: Nigel Fleeman, Research Fellow (Clinical Effectiveness), LR*i*G, University of Liverpool

Angela Stainthorpe, Research Associate (Health Economics Modelling), LR*i*G, University of Liverpool

Adrian Bagust, Professor (Health Economics Modelling), LR*i*G, University of Liverpool

Marty Richardson, Research Fellow (Medical Statistician), LR*i*G, University of Liverpool

Sarah Nolan, Research Assistant (Medical Statistician), LR*i*G, University of Liverpool

Rachel Houten, Research Associate (Health Economics Modelling), LR*i*G, University of Liverpool

Angela Boland, Associate Director, LR*i*G, University of Liverpool

Sophie Beale, Research Associate (Decision Analysis), LR*i*G, University of Liverpool

Eleanor Kotas, Information Specialist, LR*i*G, University of Liverpool

Joanne McEntee, Senior Medicines Information Pharmacist, North West Medicines Information Centre, Pharmacy Practice Unit, Liverpool

Zafar Malik, Consultant Clinical Oncologist, Consultant Clinical Oncologist, The Clatterbridge Cancer Centre NHS Foundation Trust

Correspondence to: Nigel Fleeman, Research Fellow, Liverpool Reviews and Implementation Group, University of Liverpool, Room 2.09, Whelan Building, The Quadrangle, Brownlow Hill, Liverpool L69 3GB

Date completed: 30 November 2016

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 15//194/07.

Declared competing interests of the authors: None.

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, Stainthorpe A, Bagust A, Richardson M, Nolan S, Houten R, Boland A, Beale S, Kotas E, McEntee J, Malik Z. Palbociclib in combination with an aromatase inhibitor for previously untreated metastatic, hormone receptor-positive, HER2- breast cancer [ID915]: A Single Technology Appraisal. LR*i*G, University of Liverpool, 2016

Contributions of authors:

Fleeman N	Project lead, drafted clinical results section and supervised the final report
Stainthorpe A	Critique of the company economic model and proposal of alternative interpretations of the economic evidence
Bagust A	Checking and validation of the economic model and critique
Richardson M	Critical appraisal of the statistical evidence
Nolan S	Critical appraisal of the statistical evidence
Houten R	Summary and critical appraisal of economic evidence
Boland A	Critical appraisal of the clinical and economic evidence
Beale S	Critical appraisal of the clinical and economic evidence
Kotas E	Critical appraisal of the company's database searching
McEntee J	Critical appraisal of the submission
Malik Z	Clinical advice and critical appraisal of the clinical sections of the company's submission

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

Table of contents

List of tables

List of figures

LIST OF ABBREVIATIONS

ABC	Advanced breast cancer
AE	Adverse event
ASCO	American Society of Clinical Oncology
BCS	Breast cancer specific
BICR	Blinded independent central review
BPI	Brief pain inventory
CBR	Clinical benefit rate
CDK	Cyclin-dependent kinase
CEAC	Cost effectiveness acceptability curve
CI	Confidence interval
CNS	Central nervous system
CS	Company submission
CSR	Clinical study report
DFI	Disease-free interval
DSU	Decision support unit
ECOG	Eastern cooperative oncology group
EMA	European medicines agency
EQ-5D	EuroQol-5D
ER	Oestrogen receptor
ER+	Oestrogen receptor-positive
ERG	Evidence review group
ESMO	European Society for Medical Oncology
FACT-B	Functional Assessment of Cancer Therapy-Breast
FACT-G	Functional Assessment of Cancer Therapy-General
HER2-	Human epidermal growth factor receptor 2 negative
H-H	Cumulative hazard
HR	Hazard ratio
HRQoL	Health-related quality of life
ITT	Intention-to-treat
K-M	Kaplan-meier
LET	Letrozole
MID	Minimally important difference
NICE	National Institute for Health and Care Excellence
OS	Overall survival
PALOMA	Palbociclib: Ongoing trials in the Management of breast cancer 1 (trial acronymn for the PALOMA-1 and PALOMA-2 trials)
PAL+LET	Palbociclib in combination with letrozole (palbociclib arm of PALOMA-2 trial)
PFS	Progression-free survival

PH	Proportional hazard
PLACEBO+LET	Placebo plus letrozole (LET arm of PALOMA-2 trial)
PPS	Post-progression survival
PR+	Progesterone receptor positive
PS	Performance status
PSA	Probabilistic sensitivity analysis
QALY	Quality adjusted life year
RCT	Randomised controlled trial
SAE	Serious adverse event
STA	Single technology appraisal
TOI	Trial outcome index
TD	Time to deterioration
TSAP	Trial statistical analysis plan
TTP	Time to progression

1 SUMMARY

The remit of the Evidence Review Group (ERG) is to comment on the clinical and cost effectiveness evidence submitted to the National Institute for Health and Care Excellence (NICE) as part of the single technology appraisal (STA) process. Clinical and economic evidence have been submitted to NICE by Pfizer in support of the use of palbociclib (Ibrance®) for the treatment of postmenopausal people with metastatic, hormone receptor-positive, human epidermal growth factor receptor 2-negative (HER2-) breast cancer previously untreated in the metastatic setting.

Palbociclib received a marketing authorisation, from the European Medicines Agency (EMA), on 9th November 2016, for the treatment of hormone receptor-positive, HER2- locally advanced or metastatic breast cancer (MBC) in combination with an aromatase inhibitor (which is the focus of this appraisal) or in combination with fulvestrant in women who have received prior endocrine therapy (which is expected to be the focus of a separate appraisal).

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

The population specified in the final scope issued by NICE is postmenopausal people with metastatic, hormone receptor-positive, HER2- breast cancer previously untreated in the metastatic setting. The ERG is satisfied that the evidence presented in the company submission (CS) is generalisable to the patient population in England and Wales that is described in the final scope issued by NICE. The intervention of interest in this appraisal is palbociclib (PAL) in combination with an aromatase inhibitor.

Evidence is appropriately presented for palbociclib in combination with letrozole (PAL+LET) versus letrozole (LET). Palbociclib is self-administered orally at a dose of 125mg each day for the first 21 days of a 28-day cycle. It is taken alongside letrozole. The latter is also self-administered orally, but at a dose of 2.5mg per day, every day of a 28-day cycle. LET is a commonly used aromatase inhibitor that is considered to be of equal efficacy to other aromatase inhibitors (anastrozole and exemestane) commonly used in NHS clinical practice in England and Wales.

The outcomes specified in the final scope issued by NICE include overall survival (OS), progression-free survival (PFS), response rates, adverse events (AEs) and health-related quality of life (HRQoL); these are standard outcomes used in oncology clinical trials and the company has presented data for all of these outcomes. The focus of this ERG report, however, is on the outcomes that the ERG considers are most relevant to understanding the clinical and cost effectiveness data submitted by the company for this appraisal, i.e. OS, PFS/time to progression (TTP), AEs and HRQoL. As specified in the final scope issued by NICE, the cost

effectiveness of treatments is expressed in terms of incremental cost per quality adjusted life year (QALY) gained. Outcomes are assessed over a 40-year time horizon (equivalent to a lifetime horizon) and costs are considered from an NHS perspective.

1.2 Summary of clinical effectiveness evidence submitted by the company

Clinical effectiveness data have been derived from two international multi-centre RCTs, the open-label PALOMA-1 trial (N=165; phase I/II) and the double-blind PALOMA-2 trial (N=666; phase III). Patients participating in the PALOMA-1 trial were sequentially enrolled into two cohorts. Recruitment into cohort 1 was based on patients having oestrogen receptor-positive (ER+) and HER2- tumours and recruitment into cohort 2 was based on the combination of ER+/HER2- status and amplification of cyclin D1 and/or loss of p16, or both. Across both cohorts, patients were randomised 1:1 to either the PAL+LET arm or the LET arm of the trial. Patients participating in the PALOMA-2 trial were randomised 2:1 to either the PAL+LET arm or the PLACEBO+LET arm of the trial.

The primary outcome for both trials was investigator assessed PFS; however, in both trials, assessments were also carried out by blinded independent central review (BICR). OS was a secondary outcome for both trials. All of the PALOMA-1 trial data presented by the company correspond to the data available on the cut-off date of 29 November 2013, and all of the PALOMA-2 trial data correspond to the data available on the cut-off date of 26 February 2016.

In both trials, patients randomised to receive PAL+LET spent more time in PFS and, therefore, more time on treatment than patients randomised to receive LET or PLACEBO+LET. In both trials, treatment with PAL+LET was shown to statistically significantly improve investigator assessed PFS compared to treatment with the comparator, by around 10 months (PALOMA-1 median PFS: 20.2 months versus 10.2 months; hazard ratio [HR]=0.488; 95% confidence interval [CI] 0.319 to 0.748, p=0004; PALOMA-2 median PFS: 24.8 months versus 14.5 months; HR=0.576; 95% CI 0.463 to 0.718, one-sided p<0.000001). Unlike results from the PALOMA-2 trial, results from the PALOMA-1 trial BICR assessed PFS analysis did not show a statistically significant median PFS benefit when treatment with PAL+LET was compared with LET (PALOMA-1 trial: HR=0.621; 95% CI 0.378 to 1.019, one-sided p=0.0286; PALOMA-2 trial: HR=0.653; 95% CI 0.505 to 0.844, one-side p=0.000532).

Results from subgroup analyses carried out using data from both trials, generally support the overall results. The analyses undertaken by the company include the subgroup of patients presenting with de novo metastases as well as those previously treated in the (neo)adjuvant setting. In the PALOMA-1 trial the PFS HR for patients with de novo disease was lower than the PFS HR in the ITT population (HR=0.341). In the PALOMA-2 trial, for patients with de

novo metastases, the PFS HR was slightly higher than the PFS HR for patients in the ITT population (HR=0.674). Therefore, for patients with de novo disease, the benefit was more pronounced in the PALOMA-1 trial and less pronounced in the PALOMA-2 trial. The HR estimates for patients previously treated in the (neo)adjuvant setting were similar in both trials (PALOMA-1 trial: HR=0.539; PALOMA-2 trial: HR=0.520).

Analyses of PALOMA-1 trial data suggest that treatment with PAL+LET leads to a large and statistically significant PFS benefit when compared with treatment with LET. However, this benefit is not mirrored in the OS results from this trial (median OS: 37.5 months versus 33.3 months; HR=0.813; 95% CI 0.492 to 1.345, stratified one-sided p=0.2105). The OS data from the PALOMA-1 trial are immature. The company claims that due to the variety of post-progression therapies given to patients, which were not accounted for in the analyses, the OS gain experienced by patients in the PAL+LET arm of the PALOMA-1 trial does not represent the true comparative OS benefit that occurs when the efficacy of treatment with PAL+LET is compared with treatment with LET. OS data were not available from the PALOMA-2 trial because, at the time of the planned analysis, an insufficient number of deaths had occurred to allow the final OS analysis to take place.

All patients in the PAL+LET arm of the PALOMA-1 trial reported an AE, and nearly all patients in the PALOMA-2 trial who were randomised to receive PAL+LET reported an AE (98.9%). AEs were also very common for patients in the trials who were randomised to receive either LET or PLACEBO+LET (84.4% and 95.5% respectively). Severe AEs (Grade 3 to 4 AEs) and serious AEs (SAEs) were more common in the cohort of patients treated with PAL+LET than in the cohort of patients treated with either LET or PLACEBO+LET. The two most common AEs reported for patients treated with PAL+LET in the PALOMA-1 and PALOMA-2 trials were neutropenia (74.7% and 79.5% respectively) and leukopenia (43.3% and 39.0% respectively). Neutropenia was also the most common Grade 3 to 4 AE reported by patients (54.2% and 66.4% of patients in the PAL+LET arm of the PALOMA-1 and PALOMA-2 trials, respectively). However, for the most part, neutropenia was asymptomatic and reversible. None of the cases of neutropenia that occurred in patients in either arm in the PALOMA-1 trial developed into febrile neutropenia. In the PALOMA-2 trial, only 8 of 444 patients (1.8%) in the PAL+LET arm developed febrile neutropenia, compared with no patients in the PLACEBO+LET arm. Febrile neutropenia was the most common SAE for patients treated with PAL+LET in the PALOMA-2 trial. The company argues that despite a high incidence of non-febrile neutropenia reported in the PALOMA-1 and PALOMA-2 trials, dose interruptions and dose reductions enabled patients to remain on PAL+LET.

HRQoL was captured through patient reported outcomes collected as part of both the PALOMA-1 and PALOMA-2 trials. As part of the PALOMA-1 trial, outcomes in relation to pain

(pain severity and pain interference with daily activities) were assessed using the modified Brief Pain Inventory (BPI). As part of the PALOMA-2 trial, HRQoL was captured by the Functional Assessment of Cancer Therapy-Breast (FACT-B) and EuroQol-5D (EQ-5D) questionnaires. Results from the PALOMA-1 trial demonstrate that the addition of PAL to LET does not significantly alter pain severity or pain interference with daily activities. Results from the PALOMA-2 trial show that there [redacted] between trial arms when change in baseline scores for FACT-B score, total Functional Assessment of Cancer Therapy-General (FACT-G) score, FACT-G subscale scores (for each of the four domains), Trial Outcome Index (TOI) score or Breast Cancer Specific (BCS) score are assessed. Results from [redacted]

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

Overall, the ERG is satisfied with the company's clinical effectiveness systematic review process and considers that the company's approach to evidence synthesis is appropriate. The ERG is confident that, despite excluding exemestane from the systematic review of RCT evidence, all relevant studies were included in the company's systematic review.

The ERG notes the possible PALOMA-1 trial biases identified by the EMA, as demonstrated by differences in investigator assessed and BICR assessed PFS estimates in the analysis of patients in cohort 1. As stated in the European Public Assessment Report, the EMA concluded that only findings from cohort 2 should be considered relevant to the efficacy assessment. Notwithstanding these possible biases, the ERG considers that the patient populations included in both the PALOMA-1 and PALOMA-2 trials are broadly similar and are relevant to the decision problem. The ERG considers that the PALOMA-2 trial was generally well designed and well conducted.

The ERG considers that the proportional hazards (PH) assumption is valid for the analyses of PFS data from the PALOMA-1 and PALOMA-2 trials. However, the ERG does not consider that the assumption of PH holds for the analysis of OS data from the PALOMA-1 trial. The OS hazard ratio should, therefore, be interpreted with caution.

The ERG observes that median PFS estimates, calculated using data from both trials, for patients treated with the comparator are within the range of median PFS estimates reported in previous trials of treatment with letrozole for patients with MBC treated in the first-line setting.

The ERG considers the results of the company's subgroup analyses should be treated with caution due to the small numbers of patients included in each analysis. This point is particularly important when results have been generated using data from the PALOMA-1 trial. Furthermore, the EMA's statement that only findings from cohort 2 of the PALOMA-1 trial should be considered relevant to the PFS efficacy assessment should be kept in mind when interpreting these results as the PALOMA-1 trial subgroup analyses include patients from both cohort 1 and cohort 2.

Across both trials, slight imbalances, in terms of the post-progression treatments received by patients in each treatment arm are noted, but given the small numbers of patients this finding is not unexpected. Therefore, although patients received a variety of different post-progression treatments, clinical advice to the ERG is that these treatments are reflective of treatments that are routinely offered to patients with MBC in clinical practice, and any benefit from treatment with PAL+LET in comparison to treatment with LET should, therefore, be reflected in the OS results.

The ERG agrees with the company's view that the main difference in safety profiles between patients treated with PAL+LET versus those treated with LET or PLACEBO+LET is largely a result of increased rates of non-febrile neutropenia in the cohort of patients treated with PAL+LET. The ERG also agrees with the company's view that the majority of cases of neutropenia are reversible and manageable and that the safety profile of PAL+LET is acceptable.

As patients participating in the PALOMA-1 and PALOMA-2 trials were only required to complete HRQoL questionnaires when they were progression-free and, therefore, the numbers of patients completing questionnaires decreased in each cycle. Thus, in later cycles, the numbers of patients responding are very small and the data are only reflective of the experience of relatively healthy patients. Nonetheless, the data from the earlier cycles that were collected during both trials appear to show [redacted] between treatment arms for patients in either the PALOMA-1 or PALOMA-2 trials.

1.4 Summary of submitted cost effectiveness evidence

To generate cost effectiveness results for this appraisal, the company developed a de novo partitioned survival model in Microsoft Excel. The model comprises three health states: pre-progression (stable) disease, progression (which is sub-divided into four different states: first, second and third subsequent lines of treatment and best supportive care [BSC]) and dead. All patients enter the model in the pre-progression health state and are treated with either PAL+LET or LET. Variants of this model structure have been used in previous NICE STAs. The model time horizon is 40 years. As recommended by NICE, a discount rate of 3.5% is

used for both costs and outcomes; outcomes are measured in quality adjusted life years (QALYs) and the model perspective is that of the UK NHS.

Pre-progression survival estimates for both treatment arms are based on Kaplan-Meier (K-M) data from the PALOMA-2 trial. Separate Weibull parametric functions, chosen on the basis of statistical fit and external validation, have been fitted to the PAL+LET and LET arms. Estimates of OS for both treatment arms are based on K-M data from the PALOMA-1 trial (K-M data are not available from the PALOMA-2 trial). The company observed that data from the PALOMA-2 trial show that median PFS for patients treated with PAL+LET is longer than that for patients treated with PLACEBO+LET. The company modelled OS in a way that preserved this survival benefit. A Weibull function was fitted to the K-M data from the PAL+LET arm of the PALOMA-1 trial. Then, to model OS for patients receiving LET, this Weibull function was scaled in such a way as to preserve the PALOMA-1 trial PFS benefit.

The health state utility values used to reflect patient quality of life in the pre-progression state were derived from EQ-5D scores collected, at baseline, from patients participating in the PALOMA-2 trial. This resulted in the pre-progression utility value used in the company model to represent quality of life for patients receiving PAL+LET being slightly higher than that for patients receiving LET. HRQoL in the post-progression state was estimated by adjusting the average baseline utility score for all patients participating in the PALOMA-2 trial using a published disease progression decrement. Resource use and costs were estimated using information from published sources and advice from clinical experts.

The company's base case incremental cost effectiveness ratio (ICER) for the comparison of the cost effectiveness of treatment with PAL+LET versus LET is £150,869 per QALY gained. Treatment with PAL+LET is more expensive (£94,101 versus £31.68) and more effective (+0.78 life years versus +0.63 QALYs) than treatment with LET. The company carried out a range of deterministic sensitivity analyses. The most influential adjustments were those made to the distributions used to model PFS and OS, and limiting the model time horizon to 5-years. The company's probabilistic sensitivity analysis (PSA) ICER (£151,058 per QALY gained) is very similar to the company's deterministic ICER. The PSA results also show that, when any threshold up to £100,000 per QALY gained is used, treatment with PAL+LET has zero probability of being cost effective compared with LET. The company performed scenario analyses using different approaches to modelling survival, health state utility values, resource use and costs.

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

The two fundamental issues relating to the company's cost effectiveness model are: that there are no OS data available from the PALOMA-2 trial; and the issues regarding the reliability of

survival data from the PALOMA-1 trial. Other important issues relate to the estimation of time on treatment, and the calculation of pre- and post-progression utility values.

The company's attempts to overcome the lack of OS data from the PALOMA-2 trial are methodologically flawed, and result in inconsistencies both within the data and between the assumptions underpinning the company's methods and their implementation. The use of PFS and OS data from different trials, due to the lack of OS data in the PALOMA-2 trial, is methodologically flawed, as it assumes independence between the outcomes. PFS and OS are not independent measurements; they are taken from the same individuals at different times. The ERG considers the use of PFS and OS data from the same trial to be more methodologically robust, whilst noting the limitations of the data available from the PALOMA-1 trial.

The ERG considers that the evidence of a shorter post-progression survival (PPS) for treatment with PAL+LET than for LET in the PALOMA-1 trial does not justify the assumption of equal PPS in the base case, which in fact manifests as a small gain in PPS for PAL+LET in the model.

The company has assumed that all patients are treated until progression and has, therefore, used PFS to estimate the proportion of patients receiving treatment in each cycle. The true time to treatment discontinuation (TTD) can be overestimated if patients withdraw from treatment for reasons other than disease progression, or underestimated if patients are permitted to continue treatment after progression. The ERG re-estimated treated duration, and thus the cost of first-line treatment, based on TTD data provided by the company from the PALOMA-1 trial.

Since the difference between the average utility values from patients in the PALOMA-2 trial was not statistically significant, the ERG does not consider it appropriate to use different pre-progression utility values for treatment with PAL+LET and LET. The ERG advocate pooling the baseline EQ-5D values reported in the PALOMA-2 trial. In addition, the method of estimating post-progression utility from published disutility values has been implemented incorrectly and therefore the ERG has provided a new estimate of post-progression utility.

Other issues identified by the ERG include: the lack of half-cycle correction; the incorrect application of AE costs and calculation of AE incidence; the method of discounting; and the number of days per year. The ERG has also provided a scenario analysis to investigate the impact of using data from the PALOMA-2 trial to estimate PFS and TTD. Finally, the ERG has concerns about the approach taken by the company to estimate post-progression treatment costs and undertook a sensitivity analysis to investigate the impact of varying post-progression treatment costs.

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

1.6.1 Strengths

Clinical evidence

- The comparator arm of the PALOMA-1 and PALOMA-2 trials was LET, an aromatase inhibitor used to treat patients with untreated MBC in NHS clinical practice, that is a valid comparator for this appraisal
- The EMA considers that it is reasonable to generalise the clinical effectiveness results associated with LET to other commonly used aromatase inhibitors in NHS clinical practice (i.e. exemestane and anastrozole)
- Results from the PALOMA-2 trial show that the median PFS for patients in the PAL+LET arm of the trial was higher than that for patients in the PLACEBO+LET arm.
- Despite an increased incidence of non-febrile neutropenia in the PAL+LET arms of both the PALOMA-1 and PALOMA-2 trials, there are no statistically significant differences in HRQoL between the arms

Cost effectiveness evidence

- The economic model was generally well constructed and easy to navigate
- The company has undertaken a large number of sensitivity and scenario analyses to explore uncertainty

1.6.2 Weaknesses and areas of uncertainty

Clinical evidence

- The discrepancy between PALOMA-1 trial investigator assessed and BICR assessed PFS may bias the findings from this trial, possibly in favour of treatment with PAL+LET rather than treatment with LET
- When comparing PFS HRs from the ITT populations with subgroup PFS HRs (de novo disease and patients previously treated in the adjuvant setting), the findings from the PALOMA-1 and PALOMA-2 trials are not consistent
- Analysis of data from the PALOMA-1 trial indicates a 10 month improvement in investigator assessed PFS for the cohort of patients treated with PAL+LET compared with those treated with LET; however, there is no corresponding statistically significant improvement in OS
- OS data from the PALOMA-1 trial are immature and are 3 years old (data cut-off date of 29 November 2013)
- PALOMA-2 trial OS data are not currently available due to there being too few events (deaths) to allow the final OS analysis to take place
- The PALOMA-1 trial is a relatively small trial compared to the PALOMA-2 trial and this may explain why there are some apparent imbalances in terms of baseline characteristics and treatments received on disease progression

Cost effectiveness evidence

- Modelling survival using data from two different trials is methodologically unsound

- There is no trial evidence to support the assumption that 100% of PFS gain for treatment with PAL+LET will translate into OS gain
- There is no trial evidence to support the assumption of equal PPS (zero PPS gain) for treatment with PAL+LET and treatment with LET
- The method used to adjust OS data from the PALOMA-1 trial to incorporate the assumptions of (i) PFS gain is equal to OS gain and (ii) zero PPS gain, results in neither of these assumptions holding in the model
- The Weibull model used to project PFS results in implausible hazard profiles in the long-term
- The company's use of PFS data rather than TTD data as the basis for calculating first-line drug acquisition costs leads to inaccurate cost estimates
- There is no valid basis for the company's assumption that, prior to disease progression, the HRQoL of patients prescribed PAL+LET is better than that of patients prescribed LET and, therefore, only one utility value should have been used to represent patient HRQoL in this health state
- An error in the company's calculation of the utility value used to represent the HRQoL of patients in the PPS state renders the company's estimate invalid
- The company model does not include a half-cycle correction
- The method employed by the company to discount costs and benefits was incorrect (per cycle rather than annually)
- The AE costs used in the company model are unreliable as they are based on annual rather than per cycle incidence rates and an average treatment cost (rather than AE-specific treatment costs)
- The algorithm used by the company to generate PSA results did not take into account any correlated uncertainty in the key model parameters (Weibull model scale and shape parameters)
- Within the company model a year comprises 364 rather than 365.25 days.

1.7 Summary of exploratory and sensitivity analyses undertaken by the ERG

The ERG made 12 individual changes to the submitted model, namely: re-modelling OS; re-modelling PFS and TTD based on the PALOMA-1 trial data; re-modelling PFS and TTD based on the PALOMA-2 trial data; re-calculating pre- and post-progression utility values; adding a half-cycle correction; re-calculating AE costs and probabilities; changing discounting to annual rather than per cycle; and changing the number of days per year to 365.25.

The various changes implemented by the ERG for the comparison of treatment with PAL+LET versus treatment with LET yield a mixture of effects. When implemented individually, these revisions both increase and decrease the size of the ICERs per QALY gained. The combined effect of all of the ERG's revisions, when using PALOMA-1 trial data as the basis for modelling

PFS and TTD, is to decrease the company's base case ICER per QALY gained by £17,997 to £132,872. However, the combined effect of all of the ERG's revisions, when using the PALOMA-2 trial data as the basis for modelling PFS and TTD, increases the company's base case ICER per QALY gained by £62,337 to £213,206.

The ERG considers that it is unclear whether the company's base case cost effectiveness results overestimate or underestimate the size of the most probable ICER per QALY gained. The available data from the PALOMA-1 trial are flawed, but allow for the most methodologically robust approach to modelling survival; the available data from the PALOMA-2 trial are more robust, but require the application of methodologically unsound approaches to modelling survival to compensate for the absence of OS data from that trial.

The company's base case cost effectiveness results and the results generated following the application of either of the ERG's combined revision scenarios, are all considerably higher than the range normally considered acceptable by NICE.

2 BACKGROUND

2.1 Critique of company's description of underlying health problem

Key points from the description of the underlying health problem presented in the company submission (CS)[1] are reproduced (as bulleted items) in Box 1.

Box 1 Summary of company's description of underlying health problem

- Breast cancer is a heterogeneous disease composed of a growing number of biological subtypes that vary not only in aetiology and prognosis, but also in their responses to current anti-hormonal [endocrine therapy] and chemotherapy treatments
- Determination of hormone receptor and HER2 status of breast cancer tumours currently serves as the initial basis for most clinical decisions, and it has both prognostic and predictive importance in breast cancer
- Oestrogen and progesterone drive tumour growth, and tumours that express one or both receptors are typically referred to as hormone receptor-positive
- Most hormone receptor-positive tumours are both ER+ and PR+, while approximately 15% to 20% are only ER+
- Hormone receptor-positive breast cancers tend to grow more slowly than hormone receptor-negative tumours and are much more likely to respond to hormonal therapy [i.e. endocrine therapy] that lowers the amount of available oestrogen, or blocks existing oestrogen from binding its receptor
- The most common type of ABC is ER+/HER2-
- A substantial proportion of patients initially diagnosed with early-stage or locally advanced breast cancer go on to suffer recurrence or metastases. In 2009, NICE estimated that up to 40% of those diagnosed with early breast cancer develop advanced disease within 10 years
- National-level data on ABC incidence in the UK are lacking; regional data suggest that 5% of women with breast cancer have metastatic disease at first diagnosis (de novo disease)
- Prognosis of patients with ABC is poor compared with that of patients with early-stage breast cancer, and survival rates fall as the disease advances: 5-year OS is 99% for women in the UK with stage I breast cancer, 90% for stage II, 60% for stage III, and 15% for stage IV (metastatic)
- Studies from European countries and the US consistently report average OS for patients with hormone receptor-positive/HER2- ABC as <5 years. Median OS of women receiving their first post-adjuvant systemic therapy can range from 32 to 48 months.

ABC=advanced breast cancer; ER+=oestrogen receptor positive; HER2=human epidermal growth factor receptor 2; HER2-=human epidermal growth factor receptor 2 negative; NICE=National Institute for Health and Care Excellence; OS=overall survival; PR+= progesterone receptor positive
Source: CS, Section 3.1 and Section 3.4

The Evidence Review Group (ERG) notes that for women with hormone receptor-positive breast cancer, the company presents information describing breast cancer in general, advanced breast cancer (ABC) and metastatic breast cancer (MBC). Hormone receptor-positive breast cancer is synonymous with oestrogen receptor-positive (ER+) breast cancer since most hormone receptor-positive tumours are ER+. MBC is a specific type of ABC; ABC incorporates stage III and stage IV disease, whilst MBC is defined as only stage IV breast cancer. Throughout the CS, the company uses the terms ABC and MBC interchangeably. Clinical opinion received by the ERG is that the health problems associated MBC are reflected by the tumour burden. However, the options available to treat patients with ABC and MBC, are effectively the same but may have differing effects in the metastatic population. The ERG,

therefore, considers that the company's description of the underlying health problem represents a reasonable summary of the issues facing patients with MBC.

2.1.1 Impact of metastatic breast cancer on quality of life

In Section 3.2 of the CS, the company highlights the effects of MBC on patients' health-related quality of life (HRQoL). In particular, the company argues that prolonging progression-free survival (PFS) also maintains a patient's HRQoL. Reasons for this include:

- Symptoms associated with disease progression are avoided while patients remain progression-free[2-4] and disease progression has been found to have a negative impact on HRQoL[5]
- Remaining progression-free delays the onset of chemotherapy which may be associated with many toxicities and reduced HRQoL[5-10]
- There exists among patients a perceived fear of chemotherapy[11,12] which can have a detrimental effect on HRQoL; in particular, high levels of anxiety have been reported[13,14]
- Patients who are progression-free are alive[15] and able to work[16] and maintain 'normality', e.g. fulfilling one's caring duties as partners, friends and mothers.[17]

In addition to the effect of MBC on a patient's HRQoL, diagnosis with MBC and subsequent treatment can negatively affect the caregivers of patients.[18] In particular, carers are at increased risk of depression and reduced quality of life compared to the general population.[19]

2.1.2 Correlation between progression-free survival and overall survival in patients with metastatic breast cancer

The company cites the results of seven studies that suggest that length of PFS correlates strongly with overall survival (OS).[20-25] However, the company acknowledges that it is uncertain whether OS can be directly predicted from PFS, noting biases in the modelling that was carried out in a review of 144 studies of treatment for MBC published in 2014.[20] Indeed, a review undertaken on behalf of NICE by the Decision Support Unit (DSU)[26] found that the level of evidence available to support a relationship between PFS and OS varies considerably by cancer type and is not always consistent even within one specific cancer type. The authors of a 2014 review of the literature on PFS and OS for various types of cancer concluded that only in metastatic colorectal and ovarian cancer treated with cytotoxic agents was there "…acceptable evidence of surrogacy" between PFS and OS.[27]

2.2 Critique of company's overview of current service provision

Key points from the overview of current service provision presented in the CS are reproduced (as bulleted items) in Box 2. Currently, the mainstay of treatment for patients presenting with early breast cancer is endocrine therapy, which, in the NHS, entails treatment with either tamoxifen, a nonsteroidal aromatase inhibitor (letrozole or anastrozole) or a steroidal aromatase inhibitor (exemestane). These endocrine therapies, which are all generic drugs, are also the mainstay of treatment for patients presenting with MBC who do not have imminently life-threatening disease (the focus of this appraisal). Overall, the ERG considers that the company's description of current service provision represents an accurate summary of the treatments available, their efficacy in terms of PFS and the importance of HRQoL as a treatment goal for the MBC patient population.

Box 2 Summary of company's overview of current service provision

Early breast cancer (postmenopausal women)

- The majority of early breast cancers are diagnosed within the UK National Breast Cancer Screening program
- Women diagnosed with early invasive breast cancer, regardless of age, are usually treated with surgery, and may be treated with chemotherapy-based regimens before surgery (neo-adjuvant) to downsize the tumour
- After surgery, most women with early invasive ER+ breast cancer, who are not at low risk of relapse typically receive adjuvant endocrine therapy for at least 5 years
- Several endocrine drugs are in clinical use for adjuvant therapy, including tamoxifen and aromatase inhibitors
- The aromatase inhibitors (anastrozole, exemestane and letrozole) are recommended by NICE for the adjuvant treatment of postmenopausal women with early invasive ER+ breast cancer
- Women at high risk of disease relapse are offered adjuvant chemotherapy before receiving adjuvant endocrine therapy.

Advanced breast cancer (ABC)

- ABC is a life-threatening disease that cannot be cured; the clinical goals are to delay disease progression while maintaining quality of life, alleviating symptoms and improving survival-related outcomes
- For ABC patients whose disease has progressed rapidly and/or has already spread to visceral organs, first-line chemotherapy is recommended
- Patients presenting with ABC who do not have imminently life-threatening disease should be treated with endocrine therapy [NICE guidance CG81 and ESMO and ASCO guidelines]
- Despite being standard of care in ER+ ABC, median PFS associated with treatment with currently approved endocrine therapies generally remains less than 1 year
- The ability to prolong PFS while maintaining HRQoL is, therefore, an important unmet medical need in the ER+/HER2- ABC setting.

ABC=advanced breast cancer; ASCO=American Society of Clinical Oncology; ER+=oestrogen receptor positive; ESMO=European Society for Medical Oncology; HER2-=human epidermal growth factor receptor 2 negative; HRQoL=health-related quality of life; NICE=National Institute for Health and Care Excellence; PFS=progression-free survival
Source: CS, Sections 2.5, 3.1, 3.3

2.3 Aromatase inhibitors for the treatment of metastatic breast cancer

In the ABC setting, letrozole is indicated as first-line treatment or following treatment with an anti-oestrogen, such as tamoxifen.[28] The indication for exemestane in the ABC setting is only following anti-oestrogen therapy (such as tamoxifen),[29] whereas anastrozole is indicated for the treatment of ABC in general,[30] with no restrictions to its place within the treatment pathway. Currently, aromatase inhibitors are only approved by NICE for use in the ABC setting where patients have had no prior treatment with endocrine therapy, or for those patients who have been previously treated with tamoxifen.[31] However, clinical advice to the ERG is that any aromatase inhibitor may be given as a first-line or subsequent line of treatment for post-menopausal patients, irrespective of whether patients have received treatment with tamoxifen or an aromatase inhibitor for early breast cancer.

In a retrospective study of medical record data of patients treated for ER+/HER2- MBC from four countries (United Kingdom [n=209], Belgium and the Netherlands [n=102] and Canada [n=127]) between 2008 and 2014 conducted by Mitra et al,[32] [redacted] were reported to be the two most commonly used treatments. The third most common treatment reported in this review was [redacted]. UK market research data[33] used by the company to estimate the potential number of patients eligible for treatment with palbociclib suggest that between [redacted] and the [redacted] as a first-line treatment for ER+/HER2- MBC [redacted]. By [redacted], the most common aromatase inhibitors (accounting for [redacted] of all treatments) were [redacted] followed by [redacted], and then [redacted]. The two most common chemotherapy regimens were [redacted] and [redacted].

Evidence reported in systematic reviews suggests that all aromatase inhibitors are of superior efficacy to tamoxifen for treating patients with MBC.[34-36] Evidence from an indirect treatment comparison showed that there were no differences in terms of OS, PFS or adverse events (AEs) between aromatase inhibitors that were used for the first-line treatment of patients with MBC.[36] Common AEs associated with treatment with aromatase inhibitors include arthralgia and bone pain.[36] Treatment with aromatase inhibitors may result in a loss of bone density, increasing the risk of osteoporosis and bone fractures.[37]

The ERG is aware of two trials[38,39] that compare exemestane versus anastrozole for the first-line treatment of MBC which were not included in the aforementioned systematic reviews.[34-36] The findings from one trial,[39] a randomised, open-label, phase II trial conducted in Spain, led the authors to conclude that no significant differences in clinical activity were observed in favour of exemestane over anastrozole, despite apparent numerical differences in median time to progression (TTP) between the arms (6.1 versus 12.1 months but not reported to be

statistically significant). The authors of the other trial,[38] a multi-centre, randomised, double-blind, phase III trial conducted in Japan reported almost identical TTP between arms (13.8 versus 13.7 months). The ERG is not aware of any trials that compare the use of letrozole with either anastrozole or exemestane as a first-line therapy for a MBC population.

As reported in Box 2 of this ERG report, the company reports that the median PFS with currently approved endocrine therapies for treating ER+ ABC is generally less than 1 year. Results of the five randomised controlled trials (RCTs) cited to support this statement[40-44] are summarised in a review published in 2013 by Cardosa et al.[45] The findings from the trials suggest a median PFS/TTP of between 5.6 and 8.3 months for patients treated with tamoxifen and between 8.2 months and 12.0 months for patients treated with aromatase inhibitors. As noted by the ERG, results from a recent Japanese trial show median PFS in excess of 12 months for patients treated with exemestane and anastrozole, whereas, more recently published results from trials that included patients receiving letrozole show PFS/TTP results of up to 15.6 months (in combination with placebo in CALGB 40503[46]). However, information presented in the retrospective study of medical record data by Mitra et al,[32] showed that, between 2008 and 2014, patients in UK clinical practice receiving first-line endocrine therapies had a median TTP of 12.17 months.

Patients previously treated with endocrine therapy may become resistant to treatment with aromatase inhibitors.[47] Primary resistance is typically defined as relapse occurring within 2 years of starting endocrine therapy.[47] Results from the BIG 1-98 trial[48] of adjuvant endocrine therapy, show that primary resistance occurred in 3.1% of patients treated with letrozole and 4.4% of patients treated with tamoxifen. Disease recurrence that takes place within a set period of time after completing treatment may also be considered as resistance; for example, the company considers that patients who had a disease-free interval (DFI) <12 months after completing treatment with an aromatase inhibitor in the adjuvant setting have resistant disease (CS, Section 4.8.1). Patients who have become resistant to a particular aromatase inhibitor in the adjuvant setting are, therefore, likely to be treated with a different aromatase inhibitor if they need treatment in the first-line MBC setting.[47]

Whilst in clinical practice patients may be treated more than once with aromatase inhibitors (i.e. for early breast cancer and for MBC), it is argued that there are no robust RCT data to support this approach.[47] The CALGB 40503 trial[46] is one of the first trials of patients with MBC to be published that includes patients who have been previously treated with aromatase inhibitors in the adjuvant setting. The PALOMA-1 trial[49] which compared palbociclib in combination with letrozole with letrozole alone also permitted patients to have had prior treatment with aromatase inhibitors (providing there was a DFI >12 months in the case of prior treatment with letrozole). The investigators of the BOLERO-2 trial[50] comparing everolimus in

combination with exemestane to exemestane (in combination with placebo) have also reported a subgroup analysis of patients treated in the first-line setting for MBC.[51] Patients in the BOLERO-2 trial had to be refractory to aromatase inhibitors (defined as recurrence during or within 12 months after the end of adjuvant treatment or progression during or within 1 month after the end of treatment for advanced disease).

The company highlights (CS, Section 3.3) that subsequent treatment following recurrence or progression in the first-line setting for MBC includes additional hormonal therapy (endocrine therapy) or chemotherapy. According to the recent study by Mitra et al,[32] the most common treatments for second-line treatment of MBC were [redacted]. The company highlights that other treatment options that are commonly received in the second-line setting include treatments that are not currently recommended by NICE, such as everolimus in combination with exemestane (which has been available via the Cancer Drugs Fund) and fulvestrant.

2.4 Palbociclib

This appraisal considers a new treatment option for patients with previously untreated metastatic, hormone receptor-positive, HER2- MBC: palbociclib in combination with an aromatase inhibitor. As highlighted in the CS, it is now recommended in the American Society of Clinical Oncology (ASCO) 2016 guideline[52] that a nonsteroidal aromatase inhibitor (i.e. letrozole or anastrozole) and palbociclib may be offered to postmenopausal women with treatment-naive hormone receptor-positive MBC.

Palbociclib is an oral anti-neoplastic agent. It is a selective small-molecule inhibitor of cyclin-dependent kinase 4 and 6 (CDKs 4 and 6)[53] as well as the redundant CDK 6/cyclin D1 kinase. Through its mechanism of action, palbociclib enhances the anti-proliferative efficacy of endocrine treatments through inhibition of the oestrogen receptor (ER) in breast cancer cells.[53] The company highlights results from the PALOMA-1,[49] PALOMA-2[15] and PALOMA-3[54] trials to demonstrate that this synergistic enhancement leads to improvements in PFS when treatment is compared with endocrine therapy alone. The company also argues that, by extending PFS, palbociclib should postpone the need for potentially burdensome chemotherapy. So, by prolonging PFS, patients experience a lower pain burden, stable HRQoL, and fewer severe AEs than would be expected when patients progress on endocrine therapy and start treatment with chemotherapy.

The company suggests that the mechanism by which palbociclib causes cell cycle arrest is important when considering palbociclib-induced neutropenia. Unlike chemotherapy-induced neutropenia, which is caused through irreversible human bone marrow cell death, results from the PALOMA-1, PALOMA-2 and PALOMA-3 trials show that, in most cases, cellular

proliferation resumed to near pre-treatment levels when the palbociclib dose was interrupted. Thus, the company considers that palbociclib represents an important change in terms of the treatment available to patients in the first-line ER+/HER2- MBC setting. The company suggests that this is the most important change, in terms of available treatments in this setting, since the introduction of aromatase inhibitors over 10 years ago.

The ERG notes that, alongside palbociclib, other oral CDK4/6 inhibitors are currently being investigated for their efficacy in clinical trials, including phase III trials. For patients with hormone receptor-positive, HER2- MBC previously untreated in the metastatic setting, the authors of the recently published MONALEESA-2 study[55] reported promising results for patients treated with ribociclib in combination with letrozole. Outside of clinical trials, CDK4/6 inhibitors are not, however, currently available to NHS patients treated in England and Wales. Therefore, palbociclib represents the first-in-class CDK4/6 inhibitor to be potentially available to these patients.

2.5 Number of patients potentially eligible for treatment with palbociclib

Company estimates, based on observed frequencies of different breast cancer subtypes and on the incidence of menopause, suggest that 48,867 women in England and Wales have breast cancer, of whom almost 7000 have ER+/HER2- ABC, of whom 5435 would be eligible to receive palbociclib (Table 1).

Table 1 Company's estimate of numbers of patients previously untreated in the metastatic setting

Population		Number	Assumption	Source
#	Decsription			
1	Women with breast cancer in England and Wales	England: 46,085 Wales: 2782	-	ONS data for 2014[56] Welsh Cancer and Surveillance Unit Intelligence data for 2001 to 2014[57]
2	Women with invasive breast cancer	44,061	90% of #1	NICE 2015 [58]
3	Women with early and locally advanced invasive breast cancer	41,858	95% of #2	NICE 2015[58]
4	Women presenting with advanced breast cancer at diagnosis (de novo disease)	2203	5% of #2	NICE 2015[58]
5	Women presenting with early breast cancer that die before disease progression	12,557	30% of #3	
6	Women with early and locally advanced breast cancer progressing into advanced stage	10,255	35% of (#3 - #5)	NICE 2015[58]
7	Total number of women developing advanced breast cancer per year	12,458	#4 + #6	NICE 2015[58]
8	Women with ER+/HER2- advanced breast cancer	6977	56% of #7	De Koven et al 2012[59]*
9	Postmenopausal women with ER+/HER2- advanced breast cancer†	5721	82% of #8	World Health Organization International Agency for Research on Cancer GLOBOCAN project[60]
10	Percentage women treated with first-line therapy (i.e. previously untreated in the metastatic setting)	6628	95% of #8	Pfizer, data on file
11	Percentage women treated with first-line therapy (i.e. previously untreated in the metastatic setting, ER+/HER2- advanced breast cancer)	5435	95% of #9	Pfizer, data on file

ER+=oestrogen receptor positive; HER2-=human epidermal growth factor receptor 2 negative; NICE=National Institute for Health and Care Excellence; ONS=Office for National Statistics
*The proportion in the published paper is of patients with MBC
†Women aged ≥50 years were considered to be postmenopausal
Source: CS, adapted from Table 8 and Table 87

In Section 6 of the CS it is stated that, based on market research data,[33] [redacted] of patients with ER+/HER2- ABC received an aromatase inhibitor in the last quarter of 2015. The company anticipates that [redacted].

The ERG calculates that [redacted] of 5435 equates to [redacted] patients. However, the company suggests (CS, Section 6) that the number of patients treated with palbociclib in 2017 would be [redacted] since a positive NICE recommendation can only have an effect part way through the calendar year i.e., that only [redacted] of all potentially eligible women, not just those receiving an aromatase inhibitor, would receive treatment with palbociclib.

If recommended by NICE, the company estimates that the number of patients treated with palbociclib would rise to [redacted] in 2018 and to [redacted] in 2019. These estimates constitute [redacted] and [redacted] of all potentially eligible women (not just those currently receiving an aromatase inhibitor) respectively, assuming a 0.6% increase in annual breast cancer incidence. The assumption of the rise in incidence is based on statistics obtained from the Cancer Research UK website[61] that indicate that there was a 6% rise in incidence in the UK between 2002-2004 and 2011-2013. However, the ERG observes that the Cancer Research UK website notes that "almost all of this entire rise" occurred "before the mid-2000s".

3 CRITIQUE OF COMPANY'S DEFINITION OF DECISION PROBLEM

Table 2 summarises the decision problem, described by the company in the CS, in relation to the final scope issued by NICE.[62] Each parameter is discussed in more detail in the text following the table (Section 3.1 to Section 3.7).

Table 2 NICE scope and company's decision problem

Parameter	Final scope issued by NICE	Decision problem addressed in the company's submission
Population	Postmenopausal people with metastatic, hormone receptor-positive, HER2- breast cancer previously untreated in the metastatic setting	As per final scope issued by NICE The ERG notes that patients in the trials who had previously received (neo)adjuvant treatment had a disease free interval of >12 months following (neo)adjuvant treatment with letrozole (PALOMA-1 and PALOMA-2 trials) and anastrozole (PALOMA-2 trial) before being treated for MBC
Intervention	Palbociclib in combination with an aromatase inhibitor	Palbociclib in combination with letrozole
Comparator (s)	Aromatase inhibitors (such as letrozole or anastrozole)	Letrozole
Outcomes	The outcome measures to be considered include: • overall survival (OS) • progression free survival (PFS) • response rate (RR) • adverse effects of treatment • health-related quality of life (HRQoL)	The outcome measures to be considered include: • overall survival (OS) • progression free survival (PFS) • response rate (RR) • clinical benefit rate (CBR) • adverse effects of treatment • health-related quality of life (HRQoL)
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	As per the final scope issued by NICE
Subgroups to be considered	None specified	Patients with MBC previously treated in the adjuvant setting compared with those who are presenting for the first time with MBC (de novo)
Other considerations	Guidance will only be issued in accordance with the marketing authorisation Where the wording of the therapeutic indication does not include specific treatment combinations, guidance will be issued only in the context of the evidence that has underpinned the marketing authorisation granted by the regulator	No special considerations, including issues related to equity or equality, were identified Palbociclib in combination with an aromatase inhibitor is not considered by the company to meet NICE End of Life criteria The company has not submitted a Patient Access Scheme proposal

ERG=Evidence Review Group; NICE=National Institute for Health and Care Excellence; HER2=human epidermal growth factor receptor 2; HER2-=human epidermal growth factor receptor 2 negative; MBC=metastatic breast cancer
Source: NICE Final scope and CS, Table 1

3.1 Population

The population specified in the final scope issued by NICE is postmenopausal people with metastatic, hormone receptor-positive, HER2- breast cancer previously untreated in the metastatic setting. The evidence presented by the company is for postmenopausal women with ER+/HER2- MBC (as noted in Section 2.1 of this ERG report, most women with hormone receptor-positive disease have ER+ tumours). However, the anticipated European Medicines Agency (EMA) licence for palbociclib in combination with an aromatase inhibitor will not specify the menopausal status of patients (see Section 3.2 of this ERG report for a description of the anticipated licence). The vast majority (■) of patients referred to in the CS have untreated metastatic, hormone receptor-positive, HER2- breast cancer. The exceptions are:

1. Three patients (2%) in the PALOMA-1 trial had stage III disease (which is categorised as ABC, not MBC). Similarly, in the PALOMA-2 trial, ■ patients (■) had locoregional recurrence, local recurrence or regional recurrence (which is categorised as ABC, not MBC).
2. Patients were not permitted to have relapsed on neo(adjuvant) therapy with LET (PALOMA-1), or LET or anastrozole (PALOMA-2) within 12 months of receiving treatment with these aromatase inhibitors. However, as noted in Section 2.3 of this ERG report, results from the BIG 1-98 trial[48] show the proportion of patients treated with LET who relapsed within 2 years in the adjuvant setting is 3.1%, and that patients who have relapsed whilst being treated with LET are unlikely to be re-treated with LET again. Therefore, patients who have relapsed whilst being treated with LET are outside of the scope of this appraisal.
3. The proportions of patients in the PALOMA-1 and PALOMA-2 trials presenting with de novo disease (49.1% and 37.2% respectively) are higher than seen in clinical practice in England and Wales (5%,[58,63] see also Section 2, Box 1). This discrepancy is, however, is a common feature of trials conducted in the untreated MBC setting (with many trials of LET including approximately 30% to 50% of patients with de novo disease[42,46,64-66] or even more patients with de novo disease[67]). The ERG notes that the company has conducted subgroup analyses, using data from both the PALOMA-1 and PALOMA-2 trials, which allow findings for patients presenting with de novo disease to be compared with those from MBC patients who have previously been treated in the (neo)adjuvant disease setting.

Overall, the ERG is satisfied that the evidence presented in the CS is generalisable to the patient population in England and Wales that is described in the final scope issued by NICE.

3.2 Intervention

The intervention of interest in this appraisal is palbociclib in combination with an aromatase inhibitor. Palbociclib is self-administered orally at a dose of 125mg each day for the first 21 days of a 28-day cycle. It is taken alongside LET which is self-administered orally at a dose of 2.5mg per day, each and every day of the 28-day cycle. Treatment is stopped only on disease progression, or if patients can no longer tolerate the combination.

In accordance with the treatments administered in the PALOMA-1 and PALOMA-2 trials, the company has presented evidence for palbociclib in combination with letrozole (PAL+LET). As described in Section 2.3 of this ERG report, other aromatase inhibitors including anastrozole and exemestane are available to patients treated in the UK NHS, and all aromatase inhibitors are considered to be of equal efficacy and safety.[36] It is, therefore, expected that, in clinical practice, while palbociclib would most likely be given in combination with LET, it may possibly be given with other aromatase inhibitors. Indeed, the ERG observes that the EMA considers that it is reasonable to generalise the clinical effectiveness results associated with LET to other aromatase inhibitors.[68]

3.2.1 Licensing

Palbociclib received a positive opinion from the EMA on 16 September 2016 for the treatment of hormone receptor-positive, human epidermal growth factor receptor 2 (HER2)-negative locally advanced or MBC in combination with an aromatase inhibitor (which is the focus of this appraisal) or in combination with fulvestrant in women who have received prior endocrine therapy (which is expected to be the focus of a separate appraisal). In pre- or peri-menopausal women, the endocrine therapy should be combined with a luteinizing hormone-releasing hormone (LHRH) agonist. EMA marketing authorisation was granted on 9 November 2016.

3.2.2 Implications for practice

The company states that managing the administration of palbociclib is expected to be similar to that of managing the administration of other oral agents currently available in the NHS for patients with MBC (such as aromatase inhibitors). However, additional monitoring of complete blood count on days 1 and 14 of the first two cycles and day 1 of all subsequent cycles is required. Since palbociclib has myelosuppressive properties which may, therefore, predispose patients to infections, patients should also be monitored for signs and symptoms of infection (and treated as medically appropriate). In particular, while 3-monthly visits to see a consultant are typical for patients receiving endocrine therapy, more frequent visits may be required for patients treated with palbociclib if they have palbociclib-induced neutropenia and leukopenia. No concomitant therapies are administered with palbociclib for managing AEs. However, to ensure appropriate management of the AEs, the company states that health care

professionals will need to be taught how to use dose-modification guidelines and be informed about the fundamental differences between palbociclib-induced neutropenia and chemotherapy-induced neutropenia. Palbociclib-induced neutropenia is asymptomatic and reversible, whereas chemotherapy-induced neutropenia is not reversible and, therefore, requires recovery by re-population from the original haemopoietic stem cells. This often means that a patient with chemotherapy-induced neutropenia needs to receive growth factor stimulation (such as the use of granulocyte-colony stimulating factor 7) to support bone marrow recovery.[69]

Dose modification of palbociclib is recommended based on concerns with regard to a patient's safety and tolerability of the drug. For example, management of some AEs may require temporary dose interruptions and/or dose reductions, or permanent discontinuation. In total, two dose reductions are permitted: 125mg to 100mg each day and 100mg to 75mg each day. If further reductions are required then treatment with palbociclib should be discontinued. Tables 1 to 3 of the draft summary of product characteristics provided by the company in Appendix 1 to the CS) provide more detailed information on dose-modification guidelines.

3.3 Comparators

The comparators listed in the final scope issued by NICE are 'aromatase inhibitors (such as LET or anastrozole'; exemestane is not specifically mentioned. The evidence presented by the company focuses on the comparison of PAL+LET with LET. As all aromatase inhibitors are considered to be of equivalent efficacy and safety,[36] the relative efficacy and safety of PAL+LET compared with LET is expected to be the same as that of PAL in combination with any aromatase inhibitor compared with any aromatase inhibitor. Indeed, as noted in Section 3.2 of this ERG report, the ERG observes that the EMA considers that it is reasonable to generalise the clinical effectiveness results associated with LET to other aromatase inhibitors.[68]

The ERG notes that LET has been a treatment option for over 10 years and is now available as a generic drug. Other aromatase inhibitors used in clinical practice in England include anastrozole and exemestane. Both drugs are also available as generic agents. The ERG considers that the comparators specified in the final scope issued by NICE, and addressed by the company, represent the current standard of care for the patient population specified in the final scope issued by NICE.

3.4 Outcomes

The outcomes specified in the final scope issued by NICE are OS, PFS, response rates, AEs and HRQoL; these are standard outcomes used in oncology clinical trials and are the most important outcome measures for this appraisal. In addition to these endpoints, the company

has also reported data for clinical benefit rate (CBR). The company argues that CBR, which captures complete response (CR), partial response (PR) as well as the absence of progression (stable disease) for at least 24 weeks, is regarded as a well-established robust measure of anti-tumour activity that is well suited to measuring the benefit of breast cancer drugs.[70] The focus of this ERG report, however, is on the outcomes that the ERG considers are most relevant to understanding the clinical effectiveness data and also to the cost effectiveness data submitted by the company for this appraisal, i.e. OS, PFS/TTP, AEs and HRQoL. Nonetheless, for completeness, information relating to other outcomes are reported in the appendices to this ERG report.

3.5 Economic analysis

As specified in the final scope issued by NICE, the cost effectiveness of treatments is expressed in terms of incremental cost per quality adjusted life year (QALY) gained. Outcomes are assessed over a 40-year time horizon (equivalent to a lifetime horizon), and costs are considered from an NHS perspective.

3.6 Subgroups

The company has presented PFS findings for a number of pre-specified subgroups, including (but not limited to) comparisons of results for MBC patients in the PALOMA-1 trial with and without de novo disease and for patients in the PALOMA-2 trial with a DFI of ≤12 months, >12 months or patients with de novo disease. During the clarification process, the ERG asked the company to provide PFS findings from both the PALOMA-1 and PALOMA-2 trials, for MBC patients with de novo disease and for MBC patients who had previously undergone (neo)adjuvant therapy for early-stage disease.

3.7 Other considerations

The company has stated that there are no issues relating to equity and equality and no other considerations have been raised. The company does not consider that palbociclib in combination with an aromatase inhibitor meets NICE's End of Life criteria. Nor has the company submitted a Patient Access Scheme application.

4 CLINICAL EFFECTIVENESS

The company conducted two systematic reviews to identify clinical effectiveness evidence: one to find evidence from RCTs, and the other to find evidence from non-randomised and non-controlled studies.

4.1 Methods

Overall, the ERG is satisfied with the clinical effectiveness systematic review process as described in the CS for both reviews (see Sections 4.1.1 to 4.1.4 of this ERG report). The ERG considers that the company's approach to evidence synthesis (see Section 4.1.5 of this ERG report) is appropriate.

4.1.1 Literature search methods

Full details of the strategies used to locate clinical evidence are reported in Section 4.1, Section 4.11, Appendix 4 and Appendix 11 of the CS. The clinical effectiveness searches were originally designed to identify studies published between database inception and January 2015. They were then updated in January 2016 and again in April 2016. The ERG considers updating the searches to be good practice and the date range of the final searches to be appropriate. The company searched the following databases: MEDLINE, MEDLINE in Process, Embase and The Cochrane Library (all databases). Search terms used appeared to be relevant and included medical subject headings and free text terms as well as an RCT filter in the search for RCTs. Searches were limited to finding English language and human studies.

In addition to searches of electronic databases, the company reported results from hand searches of three conference sites: ASCO, European Society for Medical Oncology (ESMO) and American Association for Cancer Research (AACR). The company included details of the search terms used to search these additional resources in the CS (Appendix 4, table 5) and the ERG considers that these search terms were relevant. The company also reports having searched two clinical trial registries: clinicaltrials.gov and International Clinical Trial Registry Platform (ICTRP).

The ERG considers that the company's searches were reported and carried out to an adequate standard. The searches accurately reflect the population and the indication described in the final scope issued by NICE. The ERG is confident that no relevant references were missed.

4.1.2 Eligibility criteria

The company provides a detailed report of the inclusion/exclusion criteria applied to the selection of potentially relevant studies for the two systematic reviews (RCTs and non-

randomised and non-controlled studies) in the CS. These criteria are summarised in Table 3. Two independent reviewers applied eligibility criteria. Disputes relating to eligibility were resolved through discussion between reviewers until consensus, or through consultation with a third reviewer.

Table 3 Summary of eligibility criteria

Parameter	Review of RCT evidence	Review of non-randomised and non-controlled study evidence
Population	Postmenopausal women with ER+, HER2- ABC or MBC Studies had to include ≥50% patients with ER+ or hormone receptor-positive disease, and ≥50% postmenopausal women; or outcomes had to be reported separately for patients in these subgroups	
Intervention	Anastrozole, letrozole or palbociclib (as monotherapy or in combination) in a first-line setting	Palbociclib (as a monotherapy or in combination with any other drug)
Comparator	Anastrozole, letrozole or palbociclib (as a monotherapy or in combination) in a first-line setting	Any or none
Outcomes	A range of pre-specified efficacy (including, but not limited to, OS, PFS, ORR and CBR), AE and HRQoL outcomes*	A range of pre-specified efficacy, AE and HRQoL outcomes*
Study design	Phase II and phase III RCTs only†	Non-randomised, controlled, prospective clinical trials; long-term follow-up studies; prospective observational studies; phase I studies; retrospective studies†
Language	English only	English only
Date	No limit	No limit

*For full details of all efficacy, AE and HRQoL outcomes, see CS, Table 9 and Table 27
† Systematic reviews and meta-analyses were included at the title/abstract screening stage and used for identification of any additional primary studies not identified through the database searches, but were excluded during the full-text assessment
ABC=advanced breast cancer; AE=adverse event; CBR=clinical benefit rate; ER+=oestrogen receptor positive; HER2-=human epidermal growth factor receptor 2 negative; HRQoL=health-related quality of life; MBC=metastatic breast cancer; ORR=objective response rate; OS=overall survival; PFS=progression-free survival; RCT=randomised controlled trial
Source: CS, adapted from Table 9 and Table 27

The ERG notes that studies reporting the safety and efficacy of treatment with exemestane were not considered to be eligible for inclusion into either of the company's reviews. As noted in Section 3.3 of this ERG report, the comparators in the final scope issued by NICE were: 'aromatase inhibitors (such as LET or anastrozole)'. The final scope did not, therefore, explicitly include or exclude exemestane as a comparator. As noted in Section 2.3 of this ERG report, according to its indication [redacted] exemestane is more likely to be used in the second-, rather than first-line setting. However, [redacted], it is also used in the first-line setting [redacted]. The ERG is not aware of any studies that have investigated palbociclib in combination with exemestane, or which have included exemestane in the comparator arm of a relevant trial. Therefore the ERG is confident that, despite excluding exemestane from its systematic review of RCT evidence, all studies relevant to enable a

comparison of palbociclib to an aromatase inhibitor have been identified by, and included in, the company's systematic review.

The ERG notes that the eligibility criteria applied by the company enabled reviewers to exclude studies based on reported trial outcomes. This could, theoretically, introduce outcome selection bias by excluding any study that measured, but did not report, specific outcomes.[71] However, the ERG also notes that as a range of outcomes were specified and as there was no need for included studies to report *all* outcomes but just one of these outcomes, in this instance, including or excluding studies based on outcomes is unlikely to be an important issue with regard to bias.

4.1.3 Data extraction

It is stated in the CS that, for both systematic reviews, data from studies included in the systematic review were extracted into a pre-specified extraction grid developed in Microsoft Excel. It is unclear if data extraction was conducted by one, two, or more reviewers and if this was conducted independently or extracted by one reviewer and cross-checked by another. However, the ERG notes that for studies included in the company's cost effectiveness review, data were extracted by a single reviewer and verified by a second individual.

4.1.4 Quality assessment methods

A risk of bias assessment of the RCTs included in the systematic review of clinical effectiveness was undertaken by the company using the method recommended by NICE[72] (based on the Centre for Reviews and Dissemination's guidance[73]). The company also assessed the methodological quality of the non-randomised and non-controlled studies that they provided as supportive evidence using the Down and Black's checklist for non-randomised studies.[74] This checklist is cited as a checklist to consider using in Appendix H of the manual for developing NICE guidelines.[75,76] It is unclear whether the quality assessment of RCTs and/or non-randomised and non-controlled studies was completed by one reviewer, or independently by two reviewers.

4.1.5 Approach to evidence synthesis

The company's literature search for RCTs led to the identification of two trials that were considered to be directly relevant to the decision problem (the PALOMA-1 and PALOMA-2 trials). The company did not carry out a meta-analysis of efficacy outcomes or pool data for AEs from the two trials (although the company did present pooled data for some AEs occurring in patients treated with PAL+LET); instead the company described and reported findings from the studies narratively. As stated in the company response to the ERG during the clarification

process, its reason for this was that it considered that the PALOMA-2 trial (the larger, confirmatory, later phase trial) was the most robust data source.

Seven citations[77-83] reporting on four studies were considered relevant to the company's systematic review of non-randomised and non-controlled studies. Within the CS, the company has described the studies and reported findings narratively.

The ERG considers that the company's approach to evidence synthesis was appropriate for both systematic reviews. The ERG also considers that, for completeness, a meta-analysis of OS and PFS outcomes from the PALOMA-1 and PALOMA-2 trials, and pooling of the AE data from these two trials, may have been informative. However, the ERG also considers that the reporting of the PALOMA-1 and PALOMA-2 trial data narratively was also appropriate, and sufficient for the purposes of this appraisal.

4.2 Identified studies in the systematic reviews

4.2.1 Randomised controlled trial evidence

Two relevant trials were included in the systematic review of RCT evidence, the phase I/II, multi-centre, randomised, open-label PALOMA-1 trial (N=165) and the larger (N=666) phase III, multi-centre, randomised, double-blind, placebo-controlled PALOMA-2 trial. Both trials included postmenopausal women with ER+/HER2- ABC who had not received previous systemic treatment in the advanced or metastatic setting. The PALOMA-1 trial was designed to compare the efficacy and safety of treatment with PAL+LET with LET, whilst the PALOMA-2 trial was designed to compare the efficacy and safety of PAL+LET with placebo in combination with LET (PLACEBO+LET).

Patients were randomly allocated to treatment in a 1:1 ratio in the PALOMA-1 trial. Randomisation was performed using an interactive web-based randomisation system, stratified by disease site (visceral versus only bone versus other) and by DFI (>12 versus ≤12 months between completion of the last adjuvant treatment and disease recurrence) or de novo.

Patients were randomly assigned 2:1 to the PALOMA-2 trial via an interactive randomisation technology system. Patients were stratified by disease site (visceral versus non-visceral), DFI since completion of prior (neo)adjuvant therapy (de novo metastatic versus ≤12 months versus >12 months), and nature of prior (neo)adjuvant anti-cancer treatment (prior hormonal therapy versus no prior hormonal therapy).

The primary results from the PALOMA-1 trial have been published in a peer reviewed journal.[49] In addition, results relating to pain severity and pain interference,[84] and an expanded analysis

of subgroup data[85] have also been published. The company also cites conference presentations of subgroup analyses by age,[86] previous systemic treatment,[87] bone metastases,[88] long term safety[89] and pain severity and pain interference.[90] In addition to the published data, the company has also presented data from the Clinical Study Report (CSR).[91]

At the time of its systematic reviews, findings from the PALOMA-2 trial have been presented at the ASCO 2016 conference.[15] The company has also included data extracted from the CSR[92] within the CS. Subsequent to the company's submission to NICE, efficacy and safety findings from the PALOMA-2 trial have been published in a peer review journal[93] and HRQoL data presented at the ESMO conference in October 2016.[94]

The ERG considers that both the PALOMA-1 and PALOMA-2 trials are relevant to the NICE decision problem.

4.2.2 Non-randomised and non-controlled evidence

As noted in Section 4.1.5, seven citations[77-83] reporting on four studies were included in the systematic review of non-randomised and non-controlled evidence. The four non-randomised and non-controlled studies were all phase I or phase II studies investigating the use of palbociclib for the treatment of breast cancer, and are described using the following trial identifiers: NCT01320592,[77,78]NCT00141297,[79] NCT00721409 (phase 1)[80] and NCT01037790 (UPCC03909).[82,83] In total, the four studies only included 81 patients with ABC.

The ERG does not consider any of the identified studies to be relevant to the NICE decision problem since none of them included treatment with palbociclib in combination with an aromatase inhibitor. The ERG does, however, note that one of the studies,[80,81] investigated the use of palbociclib monotherapy during the first cycle followed by subsequent cycles of PAL+LET. This is the phase I part (n=12) of the phase I/II RCT, referred to as the PALOMA-1 trial. As noted above, the ERG considers the PALOMA-1 trial to be relevant to the decision problem.

In the remainder of this ERG report, the ERG only critiques the RCT evidence presented by the company.

4.3 Statistical approach used for the conduct and analysis of included studies

In this section, the ERG provides a description and critique of the statistical approaches used to analyse data collected during the PALOMA-1 and PALOMA-2 trials that relate to the outcomes stipulated in the final scope issued by NICE. Information relevant to the statistical approach taken by the company has been extracted from the clinical study reports (CSRs),[91,92] the trial statistical analysis plans (TSAPs),[95,96] the trial protocols [97,98] and the CS.

4.3.1 Analysis populations of the PALOMA-1 and PALOMA-2 trials

Outcome data were collected from different study populations as shown in Table 4.

Table 4 PALOMA-1 and PALOMA-2 trial outcome populations

Analysis	Study population
Efficacy	The ITT population was the primary population for evaluating all efficacy endpoints and patient characteristics. This population included all randomised patients The ITT population, with measurable disease at baseline, was also used for the analysis of ORR in the PALOMA-1 trial, and for the analysis of ORR, CBR and DOR in the PALOMA-2 trial
PROs	PALOMA-1: All analyses were performed on the PRO evaluable population, i.e. all randomised patients who completed the baseline PRO assessment received at least one dose of study treatment and completed at least one post-baseline PRO assessment PALOMA-2: Completion rates are reported for the ITT population, all other analyses were performed on the PRO evaluable population i.e. patients who completed a baseline assessment and at least one post-baseline assessment
Safety	The as-treated population was the primary population for evaluating safety. This population included all patients who received at least one dose of any agent of the combination therapy
Biomarker analyses	The subset of as-treated patients for which baseline assessment of at least one biomarker was available.

CBR=clinical benefit rate; DOR=duration of response; ITT=intention-to-treat; ORR=objective response rate; PRO=patient-reported outcome
Source: CS, adapted from Table 18

4.3.2 Outcomes analysed in the PALOMA-1 and PALOMA-2 trials

The PALOMA-1 trial

The PALOMA-1 trial is a phase I/II trial, meaning that initially, a single-arm phase I study was carried out to assess the safety of PAL+LET and to determine a recommended dose for the PAL+LET combination to be used in the phase II study. The primary outcome for the PALOMA-1 trial was investigator assessed PFS, although assessments were also carried out by blinded independent central review (BICR). TTP and OS were secondary outcomes. The definitions and methods of analysis for PFS, OS and TTP are provided in Table 5.

The following additional endpoints were also measured in this trial: CBR, ORR and duration of response (DOR). For completeness, these are described in appendices to this ERG report (Section 10.1).

Table 5 Description and method of analysis for key efficacy outcomes (PALOMA-1 trial)

Outcome	Description	Statistical analysis
Primary efficacy outcome		
PFS	Time from randomisation to radiological disease progression or death on study. Documentation of progression was by objective disease assessment calculated from the lesion measurements, as defined by RECIST 1.0	Hypothesis: [redacted]
Secondary efficacy outcomes		
OS	Time from the date of randomisation to the date of all-cause death. Patients last known to be alive were censored at date of last contact. Survival was assessed up until approximately 28 days from the last dose of study treatment	[redacted]
TTP	Time from the date of randomisation to the date of first documentation of objective progression	[redacted]

CI=confidence interval; DFI=disease-free interval; H0=null hypothesis; HA=alternative hypothesis; ITT=intention-to-treat; K-M=Kaplan-Meier; OS=overall survival; PFS=progression-free survival; PH=proportional hazards; RECIST=response evaluation criteria in solid tumors; TTP=time to progression
Source: CS, adapted from Table 13, Table 19 and Table 20

The ERG is satisfied that the analysis method for each of these efficacy outcomes was pre-specified in the TSAP, and that all results were reported fully in the CSR. The ERG notes that one-sided hypothesis testing was used to assess PFS and TTP and, as part of the clarification process, asked the company to justify the use of this approach to hypothesis testing. The company states that one-sided hypothesis testing was deemed suitable due to there being "sufficient confidence" that treatment with PAL+LET was more effective than treatment with LET, and that it was more efficient (from a statistical perspective) in light of the expected small sample size and under the null hypothesis to use one-sided testing. The ERG is satisfied that the use of one-sided testing was appropriate, although it considers that more justification could have been provided regarding the basis for the company's confidence that PAL+LET is more effective than LET. Furthermore, the rationale for such an important statistical decision should have been provided in the protocol and/or in the TSAP.

The company states that the assumption of proportional hazards (PH) was verified for PFS, and that the results were satisfactory, referring to Figures 19 and 20 of the CS (CS, Section 4.4.1.2). However, these figures show data from the PALOMA-2 trial. The company does not mention whether any PH testing was conducted for OS. The ERG, therefore, requested clarification from the company on whether any PH testing had been conducted for the PFS or OS data from the PALOMA-1 trial. In the company's response to the ERG clarification letter, the company stated that figures demonstrating the assessment of PH (i.e. a log-cumulative hazard plot and a Schoenfeld residual plot) were not presented in the CS for the PFS data from the PALOMA-1 trial, because PFS data from the PALOMA-1 trial were not used in the economic evaluation. The company did not clarify whether any assessment of PH had been performed for either the PFS or OS data from the PALOMA-1 trial. Consequently, the ERG performed their own assessments of PH using PFS and OS data from the PALOMA-1 trial (see appendices to the ERG report, Section 10.2). The ERG considered that the PH assumption was valid for PFS data, but not for OS data. Therefore, the use of HRs to summarise treatment effect for OS is not appropriate.

The PALOMA-2 trial

The primary outcome of the PALOMA-2 trial was investigator assessed PFS, although assessments were also made by BICR. OS was a secondary outcome; TTP was not pre-specified as an endpoint. The definitions and methods of analysis for PFS and OS are listed in Table 6.

The following additional endpoints were also measured in this trial: ORR, CBR, DOR. For completeness, these are described in appendices to this ERG report (Section 10.1).

The ERG is satisfied that the analysis method for each of the reported efficacy outcomes was pre-specified in the TSAP, and that all results were reported fully in the CSR.

The company demonstrated that the assumption of PH was valid for PALOMA-2 trial PFS data by providing a log-cumulative hazard plot and a Schoenfeld residual plot (Figure 19 and Figure 20 of the CS, respectively). The ERG agrees that proportionality appears to hold for the PFS data and that the use of a HR to demonstrate PFS benefit is appropriate

Table 6 Description of efficacy outcomes reported (PALOMA-2 trial)

Outcome	Description	Statistical analysis
Primary efficacy outcome		
PFS	Time from randomisation to radiological disease progression or death on study. Documentation of progression was by objective disease assessment calculated from the lesion measurements, as defined by RECIST 1.1	[redacted]
Secondary efficacy outcomes		
OS	Time from the date of randomisation to the date of all-cause death. Patients last known to be alive were censored at date of last contact.	[redacted]

CI=confidence interval; H0=null hypothesis; HA=alternative hypothesis; ITT=intention-to-treat; K-M=Kaplan-Meier; OS=overall survival; PFS=progression-free survival; PH=proportional hazards; RECIST=response evaluation criteria in solid tumors
Source: CS, adapted from Table 16, Table 19 and Table 20, and the company's response to the ERG clarification letter

.

4.3.3 Interim analyses of progression-free survival

The PALOMA-1 trial

As stated in Section 4.2 of this ERG report, the PALOMA-1 trial was a phase I/II trial. The phase I element was described as the NCT00721409 (phase 1) study.[80,81] The phase II element was designed to assess the safety and efficacy of treatment with PAL+LET in comparison to treatment with LET alone.

During phase II patients were sequentially enrolled into two cohorts to determine whether selecting patients based on the ABC-associated biomarkers cyclin D1 (CCND1) or p16 might identify subpopulations that would be more likely to benefit from treatment with PAL+LET than the general population of patients eligible for inclusion in the trial. Patients were recruited to the first cohort (cohort 1) based solely on ER+/HER2- status. The second cohort (cohort 2) of patients was recruited based on the combination of ER+/HER2- status and amplification of cyclin D1 and/or loss of p16 or both. Across both cohorts, a total of 84 patients were randomised to receive PAL+LET, and 81 were randomised to receive LET.

An unplanned interim analysis of cohort 1 based on 32 PFS events was conducted after it was noted that almost twice as many patients in the control group were discontinuing treatment because of disease progression. The results of the interim analysis showed clinically meaningful activity of the PAL+LET combination compared with LET (hazard ratio [HR]=0.35, 95% confidence interval [CI] 0.17 to 0.72, p=0.006). The company states that these preliminary results from cohort 1 suggested that further patient selection based upon CCND1 amplification or p16 loss was unlikely to further improve patient outcomes in comparison to patient selection based on ER+/HER2- status alone. As a result, further enrolment into cohort 2 (i.e. based upon CCND1 amplification or p16 loss) was stopped, and the TSAP was amended so that all primary and secondary endpoints would be analysed in cohort 1 and 2 combined.

At the time recruitment was stopped, 165 patients had been randomised in total: 66 to cohort 1 and 99 to cohort 2. The sample size had been estimated to provide 80% power to detect a HR for PFS of 0.67 based on 114 PFS events, assuming that PFS would be increased from 9 months for LET patients to 13.5 months for PAL+LET patients. However, after 57 PFS events had occurred across both cohorts, the study protocol was amended to include a second interim analysis. This interim analysis, based on 61 PFS events, reported a HR for PFS of 0.37 (95% CI 0.21 to 0.63, one-sided p<0.0001). The investigators noted that events were being observed at a slower pace than anticipated, and consequently the protocol was amended to state that the final analysis would be performed after 95 PFS events had occurred. This

number of events would give >98% power to detect a HR for PFS of 0.50 at a one-sided α of 0.10, or 75% power to detect a HR for PFS of 0.67.

To take the results of the interim analyses into consideration, the significance level for the final analysis was adjusted using the Lan-DeMets procedure with an O'Brien-Fleming stopping boundary. The adjusted level of significance for PFS was 0.0938.

The PALOMA-2 trial

The PALOMA-2 trial was designed to have one interim analysis, which was to be performed after 226 PFS events had occurred (approximately 65% of total PFS events expected). To account for this interim analysis in the overall significance level for the analysis of PFS, which was to be preserved at 0.025 (one-sided test), hierarchical group sequential testing was performed with an error spending function at a level of 0.025. Specifically, a p-value of 0.000013 was used as the efficacy boundary for the interim analysis. The interim analysis was conducted in October 2015 when 236 PFS events had occurred, corresponding to approximately 68% of the expected events for the study. At this time point, the Data Monitoring Committee (DMC) recommended that the study continue. The company was, and remains to be, blinded to the results of the interim analysis.

4.3.4 ERG critique of statistical approach of the PALOMA trials

A summary of the checks made by the ERG in relation to the statistical approach used by the company to analyse data from the PALOMA-1 and PALOMA-2 trials is provided in Table 7. Having carried out these checks, the ERG is satisfied with the statistical approach employed by the company, with the exception that, despite asking for clarification from the company, it remains unclear to the ERG whether the company performed any testing of the PH assumption for PALOMA-1 PFS and OS data. The ERG's own assessments of the assumption of PH demonstrate that the PH assumptions hold for PFS, but not for OS (see appendices to the ERG report, Section 10.2).

Table 7 ERG assessment of statistical approach used to analyse data from the PALOMA-1 and PALOMA-2 trials

Component	Statistical approach with ERG comments	
	PALOMA-1 trial	**PALOMA-2 trial**
Protocol amendments	Protocol amendments are provided in the CSR (pages 101-104) The protocol was amended several times to include interim analyses and to make changes based on the results of these interim analyses, as outlined in Section 4.3.3. The company states that these interim analyses were not performed with the intention of possibly stopping the trial; rather, they were performed to obtain information and to inform phase III study design (CS, page 63). The ERG believes it is preferable for phase II studies to make amendments to study design in order to inform phase III studies, rather than amendments being made at phase III level, and so is not concerned by the PALOMA-1 protocol amendments. Furthermore, all amendments were made before conduct of the final analysis	Protocol amendments are provided in the CSR (108-112) Protocol amendments are outlined in detail and rationale is provided for these changes. Amendments were made before conduct of the final analysis, and so were unlikely to have been driven by results of the trial
Sample size calculation	Provided in the CSR (page 100) The ERG is satisfied with the company's original sample size calculation. The ERG noted that the company recalculated the power the study would have at the final analysis when the number of events that the final analysis would be based on was amended due to information obtained from the second interim analysis	Provided in the CSR (page 90) The ERG is satisfied with the performed sample size calculation, and noted that the calculation accounted for the one planned interim analysis.
Analysis of AEs	Type, incidence, severity and seriousness of adverse events, their relationship to study medications and any laboratory abnormalities were investigated (CS, Table 13). Many different summaries of AEs are provided in the CSR; a complete list of the different summary tables is provided on pages 230-232 of the CSR The ERG is satisfied that the methodology used to analyse the AEs is appropriate	Type, incidence, severity and seriousness of adverse events, their relationship to study medications and any laboratory abnormalities were investigated (CS, Table 16). Many different summaries of AEs are provided in the CSR; a complete list of the different summaries is provided on pages 101-104 of the CSR The ERG is satisfied that the methodology used to analyse the AEs is appropriate
Sensitivity analyses for PFS	The CSR (page 93) lists 7 sensitivity analyses that were carried out for PFS. All sensitivity analyses were performed using both investigator assessed and BICR outcome data The ERG is satisfied that all sensitivity analyses were pre-specified in the TSAP (page 32) and the results of these analyses were fully reported in the CSR (page 148 and 151)	The CSR (pages 132-133) lists 14 sensitivity analyses that were carried out for PFS. All sensitivity analyses were performed using both investigator assessed and BICR outcome data The ERG is satisfied that sensitivity analyses were pre-specified in the TSAP (pages 36-38) and the results of these analyses were fully reported in the CSR (page 134).

Component	Statistical approach with ERG comments	
	PALOMA-1 trial	**PALOMA-2 trial**
Subgroup analyses for PFS	Subgroup analyses of PFS were performed for the following baseline and prognostic factors (CS, Table 26): Age (<65 years, ≥65 years) Baseline ECOG (0 or 1) Disease site (visceral, bone only, other) Previous chemotherapy (yes, no) Previous endocrine therapy (yes, no) Previous systemic therapy (yes, no) Previous chemotherapy only (yes, no) Previous chemotherapy and endocrine therapy (yes, no) DFI (≤12 months, ≤12 months + de novo, >12 months; ≤5 years, >5 years) Biomarker status (positive, negative, unknown) Region (North America, Europe) Histopathological grade (1/2, 3) Progesterone receptor (positive, negative) Number of disease sites involved (<2, ≥2) De novo advanced disease (yes, no) The ERG notes that a complete list of subgroup analyses was not pre-specified in the TSAP. It is stated that subgroup analyses of PFS may be performed for the baseline stratification factors, baseline patient characteristics, and selected biomarkers (TSAP, p 32)	Subgroup analyses of PFS were performed for the following baseline and prognostic factors (CS, Table 26): Age (<65 years, ≥65 years) Baseline ECOG (0 or 1/2) Disease site (visceral, non-visceral) Region (North America, Europe, Asia/Pacific) Ethnicity (White, Asian) Number of disease sites (1, 2, ≥3) DFI (≤12 months, >12 months, de novo) Previous chemotherapy (yes, no) Previous endocrine therapy (yes, no) Most recent therapy (aromatase inhibitor, anti-estrogen) Biomarker expression (yes/no or low/high) Bone-only disease at baseline (yes, no) Measurable disease (yes, no) The ERG notes that a complete list of subgroup analyses was not pre-specified in the TSAP. It is stated that the potential influences of the stratification factors and baseline patient characteristics such as age, ethnic origin, ECOG performance status, geographical region/country, and selected biomarkers on the primary PFS endpoint would be evaluated (TSAP, page 24)
Analysis of PROs	PROs of pain severity and pain interference with various activities of daily life were assessed in the phase II portion of the study using the mBPI-sf. The mBPI-sf pain severity and interference scales were summarized by cycle using observed values as well as changes from baseline, displaying univariate statistics such as mean, median, SD, and 95% CI of the mean (CSR, page 79 and page 99). No adjustments for multiple testing were performed despite the large number of statistical tests performed, therefore the issue of multiplicity ought to be considered when interpreting p-values from these analyses. The ERG is generally satisfied that the methodology used to analyse PROs data is appropriate	PROs were assessed using the breast cancer specific HRQoL questionnaire (FACT-B) and generic EQ-5D. Comparisons of change from baseline scores between treatment arms were based on a repeated-measures analysis using a mixed-effects model. The variables in the model were treatment, time, and treatment-by-time; baseline was a covariate (CS, section 4.7.3.2 and Table 20). Two-sided hypothesis testing was used for analyses (except for time to deterioration analyses). No adjustments for multiple testing were performed despite the large number of statistical tests performed, therefore the issue of multiplicity ought to be considered when analysis results. The ERG is generally satisfied that the methodology used to analyse PROs data is appropriate

AE=adverse event; BICR=blinded independent central review; CI=confidence interval; CS=company submission; CSR=clinical study report; DFI=disease-free interval; ECOG= Eastern Co-operative Oncology Group; EQ-5D=EuroQol-5 Dimensions; ERG=evidence review group; FACT-B=Functional Assessment of Cancer Therapy-Breast; HRQoL=health related quality of life; mBPI-sf=modified Brief Pain Inventory short form; PFS=progression-free survival; PRO=patient-reported outcome; SD=standard deviation; TSAP=trial statistical analysis plan
Source: CS, PALOMA-1 CSR, PALOMA-2 CSR, the company's response to the ERG clarification letter, and ERG comment

4.4 Quality assessment of included studies

Appendix 8 to the CS includes an assessment of the risk of bias for the PALOMA-1 and PALOMA-2 trials. The ERG has summarised this assessment in Table 8. The ERG's examination of the patient flow in both trials (CS, Figure 7 and Figure 8) shows that none of the patients in either trial were lost to follow-up. In both trials, the reasons for withdrawing treatment were generally similar across both arms, the most common reason being disease progression or relapse.

Table 8 Company's assessment of risk of bias for PALOMA-1 and PALOMA-2 trials

Study question	Company assessment		ERG Comment
	PALOMA-1	PALOMA-2	
Was randomisation carried out appropriately?	Yes	Yes	-
Was the concealment of treatment allocation adequate?	Yes	Yes	-
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	Yes	Yes	PALOMA-1 trial: the company and the ERG noted some slight imbalances. As reported in Section 4.5 of this ERG report, overall, these imbalances are not considered likely to result in bias PALOMA-2 trial: it is unclear if differences in geographic region (See Section 4.5 of this ERG report) would introduce any bias
Were the care providers, participants and outcome assessors blind to treatment allocation?	No	Yes	PALOMA 1 trial: bias may have been introduced due to the open-label design. To mitigate bias, retrospective assessments of tumour response and disease progression were made by independent radiologic review and were blinded to treatment group in 161 of 165 (97.6%) of randomised patients
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	Yes / Yes	No / not applicable	PALOMA-1 trial: the company reports that twice as many patients in the control arm of cohort 1 discontinued the study compared with patients in cohort 2 because of disease progression, so an unplanned interim analysis was performed. The ERG notes that the findings from a final analysis of PFS reported by the EMA shows large differences between investigator assessed PFS and BICR assessed PFS for cohort 1 which the EMA state may indicate bias
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No	No	The company highlights that at the time of PFS analysis, survival events had not reached the pre-specified number of events for a survival analysis to be conducted
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes	Yes	-

EMA=European Medicines Agency; ERG=evidence review group; PFS=progression-free survival
Source: CS, adapted from Appendix 8 (Table 13)

As noted in Table 8, the ERG notes that the findings from a final analysis of cohort 1 shows large differences between investigator assessed PFS and BICR assessed PFS. These findings were reported by the EMA. According to the EMA, these results indicate that findings from cohort 1 may be significantly biased to the extent that the findings from the PALOMA-1 trial are not suitable for licensure. The EMA also conclude that only findings from cohort 2 should be considered relevant to the efficacy assessment.

Overall, the ERG considers that the PALOMA-2 trial was generally well designed and well conducted. In addition, the ERG agrees with the company's conclusion that this trial has a low risk of bias.

4.5 Characteristics of the patients in the included studies

Patients participating in the PALOMA-1 trial were recruited from 50 sites in Canada, France, Germany, Hungary, Ireland, Italy, Russia, South Africa, South Korea, Spain, Ukraine and USA. Patients participating in the PALOMA-2 trial were recruited from 186 sites in Australia, Belgium, Canada, France, Germany, Hungary, Ireland, Italy, Japan, Republic of Korea, Poland, Russian Federation, Spain, Taiwan, Ukraine, UK (seven sites) and USA. The vast majority of patients in both trials had MBC (98% in the PALOMA-1 trial and [redacted] in the PALOMA-2 trial).

In general, the trial eligibility criteria for the PALOMA-1 and PALOMA-2 trials were similar. In both trials, patients had to have ER+/HER2- ABC not amenable to resection or radiation therapy with curative intent. All patients in the PALOMA-1 and PALOMA-2 trials were required to have measurable disease according to response evaluation criteria in solid tumors criteria or bone-only disease. Prior treatment for ABC was not permitted. The presence of brain/central nervous system (CNS) metastases was also an exclusion criterion. Radiation covering <25% of bone marrow at least 2 weeks prior to initiation of study treatment was permitted in the PALOMA-1 trial; however, in the PALOMA-2 trial, patients who received prior radiotherapy to ≥25% of bone marrow were not eligible, regardless of when it had been administered.

While patients in both trials were not permitted to have brain/CNS metastases at trial entry, in the PALOMA-2 trial, patients with a history of CNS metastases or cord compression were eligible if they had been definitively treated with local therapy (e.g. radiotherapy, stereotactic surgery) and had remained clinically stable whilst not taking anticonvulsants and steroids for at least 4 weeks before randomisation. The recruitment criteria for the PALOMA-2 trial explicitly stated that patients with advanced, symptomatic, visceral spread, who were at risk of life-threatening complications in the short-term, including patients with massive uncontrolled effusions (pleural, pericardial, peritoneal), pulmonary lymphangitis, and >50% liver

involvement were to be excluded. It was also explicitly stated that, for patients entering the trial, chemotherapy was not clinically indicated.

Given the eligibility criteria for the PALOMA-1 trial, although not explicitly stated, it is likely that patients at risk of life-threatening complications, and for whom chemotherapy would be clinically indicated, would not have been included in the PALOMA-1 trial.

Other differences in eligibility criteria between the two trials relate to Eastern Cooperative Oncology Group (ECOG) performance status (PS) and prior treatment with an aromatase inhibitor in the adjuvant setting. The PALOMA-1 trial recruitment criteria excluded patients with ECOG PS 2, whereas the PALOMA-2 trial criteria included patents with ECOG PS 0 to 2; however, only 12 (1.8%) patients in the PALOMA-2 trial had ECOG PS 2. Patients included in the PALOMA-1 trial had to have a DFI >12 months following treatment with LET, whilst patients included in the PALOMA-2 trial had to have a DFI of >12 months following treatment with LET or anastrozole. This means that patients in both trials were unlikely to be resistant to LET (and those in the PALOMA-2 trial were also unlikely to be resistant to anastrozole). Advice received by the ERG is that, in clinical practice, most patients who receive aromatase inhibitors as first-line treatment for MBC have ECOG PS 0 to 2. However, in clinical practice, patients with ECOG PS>2 would be considered for treatment.

The company states that in both trials, baseline characteristics of patients were well balanced between the arms although it notes that there were slight imbalances in the proportions of patients with visceral disease, DFI, and previous treatment in the neo(adjuvant) setting in the PALOMA-1 trial. These differences all appear to favour the PAL+LET arm over the LET arm. However, the company states that these differences were not considered to be of clinical significance by the UK clinicians who were part of an advisory board. The ERG also notes additional apparent imbalances also identified by the EMA, namely time since diagnosis of breast cancer which may also favour the PAL+LET arm, proportion of patients with Grade 3 tumours which may favour the LET arm and differences in the proportion of patients with progesterone receptor-positive disease. The ERG notes that since the numbers of patients in the PALOMA-1 trial are relatively small, apparent imbalances in percentage terms may be exaggerated. The EMA also highlights possible differences by age and weight. It is stated that the differences in age may favour the LET arm.

The EMA highlights that apparent imbalances by treatment arm in the PALOMA-1 trial were due to the incorrect stratification factors being used at the time of randomisation which were discovered retrospectively during data review and source data verification. Sensitivity analyses using Case Report Form data were conducted to investigate the impact of the imbalances on the PFS results, using multivariate Cox PH models by investigator and BICR

assessments. These indicated that having additional patients with visceral disease in the LET arm may favour the PAL+LET arm in the comparison (BICR HR 0.4 for non-visceral versus visceral). However, the difference in mean and medians of age may favour the LET arm (BICR HR 0.5 for age ≥ 65 years versus < 65 years). These imbalances appear to add uncertainty to the results.

The ERG also notes imbalances in the PALOMA-2 trial [redacted]. It is unclear if differences by treatment arm according to geographic region would introduce any bias. In terms of PS, given that all patients had ECOG PS 0 to 1, these imbalances are not considered by the ERG to result in bias.

Patient baseline characteristics presented in the CS are summarised by the ERG in Table 9. The ERG notes the following minor differences between the two trials:

- The PALOMA-2 trial included proportionately [redacted] than the PALOMA-1 trial
- The PALOMA-2 trial included proportionately [redacted] patients with de novo ABC and proportionately [redacted] patients with DFI >12 months than the PALOMA-1 trial
- Compared with patients included in the PALOMA-1 trial, proportionately [redacted] patients included in the PALOMA-2 trial had received previous treatment with hormonal therapy (i.e. endocrine therapy)
- Compared with patients included in the PALOMA-1 trial, proportionately [redacted] patients in the PALOMA-2 trial had received hormonal therapy as their last therapy
- In patients whose last treatment was hormonal therapy, compared with patients in the PALOMA-1 trial, proportionally [redacted] patients included in the PALOMA-2 trial received an aromatase inhibitor.

In the CS (Section 4.14), the company argues that despite a high proportion of patients in the PALOMA-1 and PALOMA-2 trials presenting with de novo disease, clinical opinion, in the form of advisory boards, had supported the high external validity of the trial populations in terms of generalisability to clinical practice in England and Wales.[99,100] Despite slight differences in the patient populations of the two trials (as highlighted above), the ERG agrees with the company that the patient populations in both trials are representative of the patients who would be treated in clinical practice in the NHS in England and Wales. However it should be noted that the number of patients presenting with de novo MBC in England and Wales is likely to be considerably less than in the two trials.

Table 9 Baseline characteristics of the PALOMA-1 and PALOMA-2 trials

Characteristics	PALOMA-1		PALOMA-2	
	PAL+LET (n=84)	LET (n=81)	PAL+LET (n=444)	PLACEBO+LET (n=222)
Median age (range), years	63 (54 to 71)	64 (56 to 70)	62 (30 to 89)	61 28 to 88)
Ethnicity				
White	[redacted]	[redacted]	344 (77.5%)	172 (77.5%)
Black	[redacted]	[redacted]	8 (1.8%)	3 (1.4%)
Asian	[redacted]	[redacted]	65 (14.6%)	30 (13.5%)
Other	[redacted]	[redacted]	27 (6.1%)	17 (7.7%)
ECOG performance status				
0	46 (54.7%)	45 (55.6%)	257 (57.9%)	102 (45.9%)
1	38 (45.3%)	36 (44.4%)	178 (40.1%)	117 (52.7%)
2	0	0	9 (2.0%)	3 (1.4%)
Measurable disease at baseline	[redacted]	[redacted]	338 (76.1%)	171 (77.0%)
Disease site*				
Visceral	37 (44.0%)	43 (53.1%)	214 (48.2%)	110 (49.5%)
Non-visceral	47 (56.0%)	38 (46.9%)	230 (51.8%)	112 (50.5%)
Bone only	17 (20.2%)	12 (14.8%)	Not reported	Not reported
Other§	30 (35.7%)	26 (32.1%)	Not reported	Not reported
DFI*				
>12 months	25 (29.8%)	30 (37.0%)	178 (40.1%)	93 (41.9%)
≤12 months or de novo	59 (70.2%)	51 (63.0%)	266 (59.9%)	129 (58.1%)
Previous systemic treatment				
None (de novo)	44 (52.4%)	37 (45.7%)	167 (37.6%)	81 (36.5%)
Chemotherapy	34 (40.5%)	37 (45.7%)	213 (48.0%)	109 (49.1%)
Hormonal	27 (32.1%)	28 (34.6%)	249 (56.1%)	126 (56.8%)
Tamoxifen	24 (28.6%)	24 (29.6%)	Not reported	Not reported
Anastrozole	8 (9.5%)	11 (13.6%)	Not reported	Not reported
Letrozole	2 (2.4%)	1 (1.2%)	Not reported	Not reported
Exemestane	4 (4.8%)	2 (2.5%)	Not reported	Not reported
Most recent therapy				
Chemotherapy	[redacted]	[redacted]	Not reported	Not reported
Hormonal	[redacted]	[redacted]	249 (56.1%)	126 (56.8%)
Anti-oestrogen¥	[redacted]	[redacted]	154 (61.8%)	75 (59.5%)
Aromatase inhibitor	[redacted]	[redacted]	91 (36.5%)	44 (34.9%)
Other	[redacted]	[redacted]	4 (1.6%)	7 (5.6%)

DFI=disease-free interval; ECOG=Eastern Co-operative Oncology Group
*Data reported for disease site and DFI based on Case Report Form in the PALOMA-1 trial
§ [redacted]
¥ Reported as tamoxifen in the PALOMA-1 trial
Source: CS, Table 21 with additional data from CSR for PALOMA-1 trial (Tables 18, 19 and 22)

4.6 Results

All the data from the PALOMA-1 trial presented in this section correspond to the data cut-off date of 29 November 2013, which was the date of the final analysis of the primary outcome (i.e. PFS). All the data from the PALOMA-2 trial correspond to the data cut-off date of 26 February 2016, which was the date of the primary analysis of the primary outcome (i.e. PFS).

4.6.1 Time on treatment

In both trials, patients spent more time on treatment with PAL+LET than with LET or PLACEBO+LET (Table 10). The ERG notes that while median relative dose intensity (RDI) was similar between trials, time on treatment was longer in both arms of the PALOMA-2 trial than in the equivalent arms of the PALOMA-1 trial. There also appear to be differences in rates of cycle delay and dose interruptions in the PAL+LET arms of the two trials; rates of cycle delay and dose interruptions were also notably fewer in the PLACEBO+LET arm. However, rates of RDI for palbociclib/placebo and LET were similar in all arms of both trials.

Table 10 Time on treatment for patients in the PALOMA-1 and PALOMA-2 trials who received at least one dose of study treatment

Duration, delay and relative dose intensity	**PALOMA-1**			**PALOMA-2**			
	PAL+LET (n=83)		**LET (n=77)**	**PAL+LET (n=444)**		**PLACEBO+LET (n=222)**	
	PAL	**LET**	**LET**	**PAL**	**LET**	**PLACEBO**	**LET**
Median duration of treatment, days	420	428	231	603	617	413	420
Number (%) of patients with at least one							
Cycle delay	70 (84.3)	--	--	[redacted]	[redacted]	[redacted]	[redacted]
Dose reduction	33 (39.8)	--	--	[redacted]	[redacted]	[redacted]	[redacted]
Dose interruption	47 (56.6)	32 (38.6)	23 (29.9)	[redacted]	[redacted]	[redacted]	[redacted]
Relative dose intensity %*							
Mean (Standard deviation)	94.1 (26.2)	99.5 (1.1)	99.5 (2.2)	--	--	--	--
Median (Range)	95.4	100.0	100.0	93.0 (40.3 to 109.5)	99.9 (73.4 to 100.2)	99.6 (56.1 to 104.5)	100.0 (79.0 to 100.0)

* Defined as (actual dose / intended dose) x 100%
Source: CS, adapted from Tables 40 and 42

4.6.2 Progression-free survival / time to treatment progression

While the primary outcome of both trials was investigator assessed PFS, the company also provided BICR results for PFS in the intention-to-treat (ITT) population for both trials. Subgroup analyses for PFS were also conducted in both trials. As highlighted by the ERG in Section 4.3.2 of this report, TTP was a secondary outcome in the PALOMA-1 trial but not in the PALOMA-2 trial. The results of the analyses of PFS and TTP in the ITT populations of both trials are summarised in Table 11.

Table 11 Progression-free survival and time to treatment progression results in the PALOMA-1 and PALOMA-2 trials

Outcome	PALOMA-1		PALOMA-2	
	PAL+LET (n=84)	LET (n=81)	PAL+LET (n=444)	PLACEBO+LET (n=222)
PFS				
Median PFS, months (95% CI) – investigator assessment	20.2 (13.8 to 27.5)	10.2 (5.7 to 12.6)	24.8 (22.1 to NE)	14.5 (12.9 to 17.1)
Hazard ratio (95% CI) for progression or death – investigator assessment	0.488 (0.319 to 0.748, one-sided p=0.0004[a])		0.576 (0.463 to 0.718, one-sided p<0.000001[b])	
Median PFS, months (95% CI) – BICR[c]	25.7 (17.7 to NE)	14.8 (9.3 to 20.4)	30.5 (27.4 to NE)	19.3 (16.4 to 30.6)
Hazard ratio (95% CI) for progression or death – BICR[c]	0.621 (0.378 to 1.019, one-sided p=0.0286[a])		0.653 (0.505 to 0.844) one-sided p=0.000532[b])	
TTP				
Median TTP, months – investigator assessment	20.2	10.2	-	-
Hazard ratio (95% CI) for progression – investigator assessment	0.399 (0.265 to 0.601, p<0.0001)		-	-
Median TTP, months – BICR[c]	25.7	14.8	-	-
Hazard ratio (95% CI) for progression – BICR[c]	0.621 (0.378 to 1.019, stratified log rank p=0.0286)		-	

[a]P<0.0938 indicated a statistically significant result
[b]P<0.025 indicated a statistically significant result
[c]BICR was conducted on 97% of the ITT population for PALOMA-1, and the entire ITT population for PALOMA-2
'-'=not reported, BICR=blinded independent review; CI=confidence interval; ITT=intention-to-treat; NE=not evaluable; PFS=progression-free survival; TTP=time to progression
Source: CS, adapted from Tables 22 to 24, CSR, Table 36 and EMA[68] Table 29 and Table 34

Progression-free survival and time to progression results (ITT populations)

Compared to treatment with LET (PALOMA-1 trial) and PLACEBO+LET (PALOMA-2 trial), treatment with PAL+LET was shown to statistically significantly improve median PFS by around 10 months. The company also provided the Kaplan-Meier (K-M) data, from both trials, for the analysis of investigator assessed PFS (CS, Figure 9 and Figure 12). In both instances, the K-M data for the two treatment arms diverge early (from approximately 2 months in the PALOMA-1 trial, and approximately 3 months in the PALOMA-2 trial), and the treatment benefit for patients treated with PAL+LET is sustained over time.

The ERG notes that investigator assessed median PFS for patients in the PALOMA-2 trial treated with PLACEBO+LET is numerically higher than the investigator assessed median PFS reported for patients in the LET arm of the PALOMA-1 trial. The ERG notes that median PFS is considerably longer in both arms of the PALOMA-1 trial when assessed by BICR rather than by the investigator; the difference between arms is not however statistically different for BICR assessed PFS. Median PFS is also considerably longer in both arms of the PALOMA-2 trial when assessed by BICR rather than by the investigator. The investigator assessed median PFS is within the range of median PFS reported for LET or PLACEBO+LET in previous trials of first-line endocrine therapy for treating MBC[42,46,64-67,101,102] but only the BICR assessed PFS in the PALOMA-1 trial falls within this range. It should be noted that not all trials necessarily include patients with similar characteristics, however. For example, four trials[42,46,65,102] have permitted the use of chemotherapy for treating MBC prior to first-line hormonal treatment for MBC (although in two trials,[42,65] was received by <10% of patients).

Consistent with this PFS benefit, the median TTP calculated from PALOMA-1 trial data is 20.2 months in the PAL+LET arm and 10.2 months in the LET arm (HR=0.399; 95% CI 0.265 to 0.601, p<0.0001). The BICR results are broadly consistent with these results.

The univariate and multivariate analyses of PFS in the PALOMA-2 trial were in accordance with the results from the primary analysis, demonstrating a statistically significant improvement in PFS for PAL+LET in comparison to LET, for both investigator assessed and BICR data. These results are provided in appendices to this ERG report (Section 10.3, Table 37). Pre-specified progression-free survival subgroup analyses

PFS subgroup analyses were performed for various pre-specified demographic and prognostic factors (see Section 4.3.4 of this report [Table 7] and the company provided the results from these analyses in Figure 1 of the company response to the ERG clarification letter (PALOMA-1 trial) and in Figure 14 of the CS (PALOMA-2 trial).

The results of all subgroup analyses demonstrate a statistically significant treatment benefit for patients treated with PAL+LET in comparison to patients treated with LET, with the following exceptions:

- PALOMA-1 trial: DFI ≤12 months subgroup (excluding patients with de novo disease) - a trend was demonstrated favouring PAL+LET, although statistical significance was not achieved
- [redacted].

The company postulates that the treatment effect estimate for patients in the PALOMA-1 trial with a DFI ≤12 months may not have reached statistical significance due to the small number of patients in this subgroup (n=15 in the PAL+LET arm, n=14 in the LET arm). The ERG agrees that the small sample size may be the reason for the non-significant effect estimate, and notes that the p-value for the test for subgroup differences between this subgroup (patients with a DFI ≤12 months) and the subgroup of patients with a DFI >12 months is non-significant. Therefore, there is no evidence to suggest that there is a statistically significant difference between these groups (patients with DFI ≤12 months and patients with a DFI<12 months).

Regarding the subgroup of [redacted], the treatment effect estimate favoured treatment with PAL+LET over treatment with PLACEBO+LET ([redacted]) but the study was not powered to detect significant differences in this subgroup. The ERG, therefore, considers that the fact that the treatment effect estimate for this subgroup did not achieve statistical significance should not be a cause for concern.

The company highlights that results from the PALOMA-1 and PALOMA-2 trials indicate similar PFS benefit for the intervention arm compared with the comparator arm for the subgroups of women older than 65 and those younger than 65. The company states that these results are of particular importance as treatment advances for breast cancer have traditionally benefited younger women more than older women. The ERG notes that in the PALOMA-1 trial, the subgroup analysis results do suggest a greater treatment benefit for younger (age<65) women than older (age≥65) women, but that treatment with PAL+LET statistically significantly improves PFS in comparison to treatment with LET for both groups of women, and the p-value for the test for subgroup differences was non-significant. Data from the PALOMA-2 trial show the treatment effect estimates for these two subgroups are extremely similar, suggesting that older women gain as much benefit as younger women from treatment with PAL+LET in comparison to treatment with PLACEBO+LET.

The ERG notes the EMA's conclusion that only findings from cohort 2 should be considered relevant to the efficacy assessment in the PALOMA-1 trial (see Section 4.4 of this ERG report). Therefore, the results of all subgroup analyses should be treated with caution.

Progression-free survival subgroup analyses requested by the ERG

The company argues that because regional data suggest that only 5% of women in the UK with breast cancer have de novo metastases, the PFS HR for the PALOMA-2 ITT population may conservatively reflect the efficacy of PAL+LET in the context of the UK population. This is because in the PALOMA-2 trial, for patients with de novo metastases, the PFS HR was slightly higher than the PFS HR for patients in the ITT population, i.e. in patients with de novo disease, the benefit was less pronounced. As evident from data requested by the ERG (Table 12), this is in contrast to the results of the PALOMA-1 trial as the PFS HR for patients with de novo disease was lower than the PFS HR in the ITT population.

The findings must however be treated with caution due to the small numbers of patients included in the analyses, particularly in the PALOMA-1 trial. Furthermore, the ERG again notes the EMA's conclusion that only findings from cohort 2 should be considered relevant to the efficacy assessment in the PALOMA-1 trial (see Section 4.4). These subgroup analyses include patients from both cohort 1 and cohort 2 of the PALOMA-1 trial.

Table 12 Progression-free survival in the subgroup analyses requested by the ERG for the PALOMA-1 and PALOMA-2 trials

Outcome	PALOMA-1		PALOMA-2	
	PAL+LET (n=84)	LET (n=81)	PAL+LET (n=444)	PLACEBO+LET (n=222)
ITT population				
Median PFS, months (95% CI)	20.2 (13.8 to 27.5)	10.2 (5.7 to 12.6)	24.8 (22.1 to NE)	14.5 (12.9 to 17.1)
Hazard ratio (95% CI)	0.488 (0.319 to 0.748)		0.576 (0.463 to 0.718)	
Patients with de novo disease				
Median PFS, months (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]
Hazard ratio (95% CI)	0.341 (0.194 to 0.599)		0.674 (0.457 to 0.993)	
Patients who have received prior neo(adjuvant) therapy				
Median PFS, months (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]
Hazard ratio (95% CI)	0.539 (0.302 to 0.962)		0.520 (0.399 to 0.680)	

CI=confidence interval; ITT=intention-to-treat population; NE=not estimable; PFS=progression-free survival
Source: Company response to ERG clarification letter, A3

Biomarker analyses

The company also provided the results of the biomarker analyses for the PALOMA-1 and PALOMA-2 trials in appendix 10 of the CS. Analyses were performed on the subset of as-treated patients for which baseline assessment of at least one biomarker was available These exploratory analyses did not indicate that there were any particular biomarkers that should guide the use of PAL+LET in clinical practice.

Other analyses of progression-free survival in the PALOMA-1 trial

As noted in Section 4.4 of this ERG report, the EMA have stated that only findings from cohort 2 should be considered relevant to the efficacy assessment of the PALOMA-1 trial.[68] The investigator assessed and BICR assessed PFS findings for the two cohorts are summarised in Table 13. It can be clearly seen form the results that there is a large discordance between investigator assessed PFS and BICR assessed in cohort 1, which is less pronounced in cohort 2. In part, the large difference may again be attributable to small sample size in cohort 1 (n=66).

Table 13 Progression-free survival by cohort in the PALOMA-1 trial

Outcome	Cohort 1		Cohort 2	
	PAL+LET (n=34)	LET (n=32)	PAL+LET (n=50)	LET (n=49)
Investigator assessed PFS				
Median PFS, months (95% CI)	26.1 (11.2 to NE)	5.7 (2.6 to 10.5)	18.1 (13.1 to 27.5)	11.1 (7.1 to 16.4)
Hazard ratio (95% CI)	0.299 (0.156 to 0.572)		0.508 (0.303 to 0.853)	
One-sided p-value	p=0.0001		p=0.0046	
BICR assessed PFS				
Median PFS, months (95% CI)	31.6 (11.2 to NE)	38.6 (7.5 to 38.6)	20.3 (12.2 to NE)	14.6 (8.1 to 20.0)
Hazard ratio (95% CI)	0.731 (0.300 to 1.779)		0.576 (0.316 to 1.050)	
One-sided p-value	p=0.2442		p=0.0342	

BICR= blinded independent central review; CI=confidence interval; ITT=intention-to-treat population; NE=not estimable; PFS=progression-free survival
Source: EMA European Public Assessment Report, adapted from Figure 17

The ERG notes that if only the findings from cohort 2 are considered from the PALOMA-1 trial, then the gain in investigator assessed median PFS is reduced from approximately 10 months to 7 months. The difference based on BICR assessed median PFS is reduced from nearly 11 months to 5.7 months.

The company also refers to an analysis of treatments given to patients in the PALOMA-1 trial after their disease progressed[103] to demonstrate how the use of PAL+LET may delay the onset of subsequent therapies in comparison to LET. The company states that delaying chemotherapy is psychologically beneficial to patients in many ways (see Section 2.1 of this ERG report). This analysis showed that the median time from randomisation to first subsequent treatment was longer in the PAL+LET arm than in the LET arm when the subsequent treatment was endocrine therapy (428 days versus 369 days) and when it was chemotherapy (280 days versus 119 days). Additionally, the first subsequent chemotherapy was administered earlier to patients who had received PAL+LET (57 days) than to patients who received LET (136 days).

ERG comment on progression-free survival findings

The ERG considers the PFS data from the PALOMA-1 trial (whether from the ITT population or from subgroup analyses) to be less robust than the PFS findings from the PALOMA-2 trial. This is because the PALOMA-1 trial appears to be at greater risk of bias for reasons highlighted in Section 4.4 of this ERG report and because of the large differences reported by the EMA[68] in terms of investigator assessed PFS and BICR assessed PFS in cohort 1 of the PALOMA-1 trial. In the PALOMA-2 trial, median PFS in both arms of the trial appears to be substantially higher according to BICR when compared with investigator assessed PFS. However, the HR for BICR assessed PFS is not too dissimilar to the HR observed with investigator assessed PFS. Furthermore, differences between arms are statistically significant for both investigator assessed PFS and BICR assessed PFS in the PALOMA-2 trial. In the PALOMA-1 trial, statistically significant differences were only observed with investigator assessed PFS.

4.6.3 Overall survival

PALOMA-1 trial

The median follow-up was 29.6 months in the PAL+LET arm and 27.9 months in the LET arm. The median OS in the PAL+LET arm was 37.5 months (95% CI 28.4 to not reached [NR]) and in the LET arm was 33.3 months (95% CI 26.4 to NR). The probability of survival was higher for patients receiving PAL+LET than for those receiving LET at 1 year (89.0% versus 87.0%), at 2 years (77.1% versus 70.2%), and at 3 years (53.0% versus 44.0%). The company also provided the K-M curves for the analysis of OS (CS, Figure 11). The observed HR for the comparison of PAL+LET and LET for OS was 0.813 (95% CI 0.492 to 1.345, p=0.2105). However, the ERG notes that the K-M curves cross, and therefore the assumption of PH, which is used to generate the HR, does not hold. The OS hazard ratio should, therefore, be interpreted with caution.

It is important to note that the OS data reported in the PALOMA-1 trial are immature; the analysis was performed on OS data taken from a cut-off date of 29 November 2013, based on only 61 deaths among 165 patients and so, at the time, the trial was not powered to detect significant differences between the two treatments. The company states that a further OS analysis will become available on an event-driven basis, however the company did not report whether any analyses have been conducted in the subsequent three years to the OS analysis presented in the CS.

PALOMA-2 trial

OS data were not available from the PALOMA-2 trial. In accordance with the TSAP, OS was to be tested for significance when interim and final PFS analyses were performed, provided

PFS was statistically significant at this time. The interim PFS analysis was conducted on data available on 01 May 2015; however, PFS had not reached statistical significance at this time and, therefore, an OS analysis was not conducted.

At the time of the final PFS analysis (26 February 2016), data showed that an insufficient number of deaths had occurred and so the final OS analysis could not be carried out (only [redacted] deaths from 666 patients had occurred, which equates to only [redacted] of the required 390 total deaths pre-specified for the final OS analysis). The External Data Monitoring Committee reviewed the results and did not propose early closure of the trial for efficacy or express any safety concerns. Since the company remains blinded to the results of the interim OS analysis, the K-M OS curves and censoring information, part of the interim OS analysis, are unavailable at this time.

Treatment received on disease progression in the PALOMA-1 trial

The company claims that due to the variety and frequency of post-progression therapies received by patients, which were not accounted for in the analyses, OS data from the PALOMA-1 trial do not represent the true comparative survival gain by patients treated with PAL+LET when compared to patients treated with LET. While the ERG agrees with the company that the health of individual patients deteriorates at different rates post-progression, and so all patients may not be best suited to the same post-progression therapies, the ERG does not agree that the PALOMA-1 trial was unable to capture true OS benefit. By definition, an RCT such as the PALOMA-1 trial consists of balanced treatment groups, with a variety of patients with different baseline characteristics and prognostic factors in each treatment arm. Furthermore, the ERG considers that the population included in the PALOMA-1 trial is reflective of the population seen in clinical practice (see Section 4.5). Therefore, although patients receive a variety of different treatments post-progression, these post-progression treatments will be reflective of clinical practice, and any benefit from treatment with PAL+LET in comparison to treatment with LET alone should be, therefore, reflected in the OS results.

The ERG notes that data reported in a poster presented at the 38th San Antonio Breast Cancer Symposium in December 2015[103] (summarised in appendices to this ERG report, Section 10.4) appear to show some imbalances by treatment arm in terms of treatments received post-progression. A greater proportion of patients in the PAL+LET arm received subsequent chemotherapy than in the LET arm (51.5% versus 39.6% respectively) whereas a smaller proportion received subsequent endocrine therapy (45.4% versus 60.4% respectively) or other therapy (18.2% versus 24.5% respectively). These results may reflect slight differences in ECOG PS by treatment arm recorded at the time of progression. Data presented by the company during the clarification process show that at the time of disease progression, [redacted] of

patients had an ECOG PS ≥2 in the PAL+LET arm compared with [redacted] of patients in the LET arm. However, the numbers of patients in both arms who received subsequent treatment were very small (n=33 and n=53 respectively) as was the number of patients for whom ECOG PS was available for ([redacted] and [redacted] respectively). The ERG notes that small differences in actual numbers can result in large differences in proportions and therefore suggests that the data from the PALOMA-1 trial must be treated with caution.

Treatment received on disease progression in the PALOMA-2 trial

During the clarification process the company provided data showing that [redacted] in both arms of the PALOMA-2 trial. In this trial a large number of patients received subsequent treatments ([redacted] in the PAL+LET arm and [redacted] in the PLACEBO+LET arm). The most common post-progression hormonal treatments received by patients in the PAL+LET and PLACEBO+LET arms respectively were [redacted] and the most common chemotherapies were [redacted]. ECOG PS at time of progression by arm was [redacted] in this trial than in the PALOMA-1 trial: [redacted].

ERG comment on overall survival findings

The ERG considers that the post-progression treatments received by patients in both trials are treatments that are routinely offered to patients with MBC in clinical practice. However, clinical opinion received by the ERG is that patients in England and Wales are more likely to receive anthracycline based treatments on disease progression, especially when patients do not receive an anthracycline treatment as a component of adjuvant treatment. Baseline characteristics reported for the PALOMA-1 and PALOMA-1 trials include details of prior chemotherapy, not prior anthracycline based chemotherapy.

4.6.4 Other secondary efficacy outcome results

The company reported a number of other secondary outcomes, including ORR, CBR and DOR. These are described and critiqued in appendices to this ERG report.

4.6.5 Safety

Safety data for patients in the PALOMA-1 and PALOMA-2 trials treated with PAL+LET are reported in the CS.

Overview of treatment emergent adverse events (including death)

The company's overview of treatment emergent AEs reported in the CS are summarised by the ERG in Table 14. All patients in the PAL+LET arm of the PALOMA-1 trial reported an AE and in the PALOMA-2 trial, nearly all patients reported an AE. AEs were also common in the LET and PLACEBO+LET arms of the trials. The company reported the proportion of serious AEs (SAEs) and Grade 3 to 4 AEs in each arm for the PALOMA-1 and PALOMA-2 trials. Compared with LET and PLACEBO+LET arms, SAEs and Grade 3 to 4 AEs were more common with PAL+LET. Deaths from AEs were relatively uncommon in both trials. [redacted]

Table 14 Treatment emergent adverse events in the PALOMA-1 and PALOMA-2 trials

Adverse events	PALOMA-1		PALOMA-2	
	PAL+LET (n=83)	LET (n=77)	PAL+LET (n=444)	PLACEBO+LET (n=222)
	%	%	%	%
Patients with any AE	100.0	84.4	98.9	95.5
Patients with SAEs	21.7	6.3	19.6	12.6
Patients with Grade 3 or 4 AEs	75.9†	20.8	77.5	25.2
Patients with Grade 5 AEs (deaths)	1.2	0.0	2.3	1.8

AE=adverse event; SAE=serious adverse event
Source: CS, Sections 4.12.1 and 4.12.2 and EMA,[68] adapted from Table 49

Types of treatment-emergent adverse events and serious events

Treatment-emergent AEs that occurred in the PALOMA-1 and PALOMA-2 trials are presented in the CS (Table 39 and Table 41 respectively) and summarised in the appendices to this ERG report (Section 10.6, Table 41). The most commonly experienced AEs with PAL+LET were haematological toxicities, particularly neutropenia (74.7%) and leukopenia (43.4%). In the PALOMA-2 trial, the proportions were 79.5% and 6.3%. In the PAL+LET arm of the PALOMA-1 trial, neutropenia was the most common Grade 3 to 4 AE (54.2%). In the PALOMA-2 trial, the most common Grade 3 to 4 AE with PAL+LET was also neutropenia (66.4%).

In the PALOMA-1 trial, [redacted] were the only SAEs reported [redacted]. In the LET arm, [redacted]. In the PALOMA-2 trial, the most commonly reported all-causality SAE in the PAL+LET arm was [redacted] and in the PLACEBO+LET

arm it was [redacted]. All other all-causality SAEs were reported [redacted] of the patients in either arm of the PALOMA-2 trial.

Overall, therefore, the main difference between the treatment arms in terms of types of AEs reported appears to relate to incidence of neutropenia (to a large extent) and leukopenia (to a lesser extent).

Managing neutropenia

The company highlights that none of the cases of neutropenia in either arm in the PALOMA-1 trial developed into febrile neutropenia and that all cases of neutropenia in this trial were asymptomatic. In the PALOMA-2 trial, it is reported in the CS that only seven of 444 patients (1.6%) in the PAL+LET arm developed febrile neutropenia compared with none of the patients in the PLACEBO+LET arm; the recently published paper[93] reports that eight of 444 patients (1.8%) in the PAL+LET arm developed febrile neutropenia compared with none of the patients in the PLACEBO+LET arm. Additionally, it is stated by the company that the results of a subgroup analysis from the PALOMA-1 trial (data not presented or referenced in the CS) indicate that neutropenia, especially of more severe grades, tended to occur less frequently with increasing treatment cycles. Overall, the company considers that palbociclib-associated neutropenia is relatively uncomplicated. The ERG concurs that the data appear to support this assertion.

Treatment discontinuation due to adverse events

As shown in Table 15, [redacted]

The company highlights that treatment duration was longer with PAL+LET than with LET/PLACEBO+LET (see Section 4.6.1 of this ERG report). Therefore the company argues that despite a high incidence of neutropenia reported in the PALOMA-1 and PALOMA-2 trials, dose interruptions and dose reductions enabled patients to remain on PAL+LET, helping to prolong PFS as a result.

Table 15 Treatment discontinuation associated with adverse events the PALOMA-1 and PALOMA-2 trials

Discontinuation type due to adverse events	PALOMA-1		PALOMA-2	
	PAL+LET (n=83)	Letrozole (n=77)	PAL+LET (n=444)	PLACEBO+ LET (n=222)
	%	%	%	%
Permanent discontinuation from trial	[redacted]	[redacted]	2.5	1.8
Permanent discontinuation of palbociclib/placebo	[redacted]	[redacted]	9.2	5.4
Permanent discontinuation of letrozole	[redacted]	[redacted]	6.1	5.0
Temporary discontinuation of palbociclib/placebo	[redacted]	[redacted]¥	74.8	15.8
Temporary discontinuation of letrozole	[redacted]	[redacted]	17.3	9.9
Dose reduction of palbociclib/placebo	[redacted]	[redacted]	36.0	1.4

¥ [redacted]

Source: CSR for the PALOMA-1 trial, adapted from Table 68 and EMA,[68] adapted from Table 49

Subgroup analysis of adverse events

The company states that the results of subgroup analysis by age (younger or older than 65 years) in the PALOMA-1 trial suggest similar rates of Grade 3 to 4 AEs and rates of dose reductions and discontinuations regardless of age. The company argues that these results (which are not presented in the CS but have been presented in a journal publication[85]) further support the ability of palbociclib to benefit both younger and older patients. The ERG concurs with the company.

ERG comment on adverse events

The ERG concurs with the company that the main difference in the safety profiles of the treatments (PAL+LET compared with LET or PLACEBO+LET) is largely the result of increased rates of neutropenia in the palbociclib treated patients. The ERG also concurs with the company that the majority of cases of neutropenia experienced in the two trials are reversible and manageable, resulting in relatively few permanent treatment discontinuations and that the safety profile of PAL+LET is therefore acceptable.

4.6.6 Health-related quality of life

As part of the PALOMA-1 trial, outcomes in relation to pain (pain severity and pain interference with daily activities) were assessed using the modified Brief Pain Inventory (BPI). As has been recognised in a publication reporting results from the PALOMA-1 trial: "The BPI is not an instrument that can measure quality of life broadly; as such, this study was not designed to provide an analysis of patients' general well-being, emotional and physical functioning, global quality of life, or utility associated with study treatment."[84] However, a broader HRQoL analysis was conducted in the PALOMA-2 trial using the using the Functional Assessment of Cancer Therapy-Breast (FACT-B)[104] and EuroQol-5D (EQ-5D)[105] questionnaires.

The PALOMA-1 trial

As noted in Section 4.3.1 (Table 4) of this ERG report, all analyses were performed on the PRO evaluable population i.e. all randomised patients who completed the baseline PRO assessment, received at least one dose of study treatment and completed at least one post-baseline PRO assessment: 76 patients in the PAL+LET arm and 74 patients in the LET arm. Assessments were carried out on day 1 of each treatment cycle and at withdrawal or at the end of treatment. An examination of findings presented at the 2014 San Antonio Breast Cancer Symposium[90] and published this year in a peer reviewed journal[84] show that:

- Baseline observed mean pain severity and pain interference scores were similar between the two treatment arms
- Patients in the PAL+LET arm generally showed a consistently greater numeric reduction from baseline in pain severity and pain interference until "later" cycles; the ERG observes that the data appear to be less consistent after cycle 23, when 27.6% and 14.9% of all PRO patients in the PAL+LET and LET arms respectively completed the BPI
- The difference between treatment arms in the mean change of pain severity score from baseline was statistically significant at some of the earlier cycles (cycles 5, 6, 7, 8, 10, 12; $p<0.05$; no adjustments were made for multiplicity) representing a numerically greater decrease in the pain experienced by patients in the PAL+LET arm compared with those in the LET arm
- There were no statistically significant differences between treatment arms for mean change of pain severity score from baseline in the later cycles
- There were no statistically significant differences in change from baseline for mean change of pain interference score from baseline
- There were no statistically significant differences between treatment arms in pain severity score or pain interference score
- Whilst the change-from-baseline analyses were pre-specified, the between arm (mixed model) comparisons in the PALOMA-1 trial were post-hoc analyses
- A limitation of the study is that results were not adjusted for the concomitant use of opioids or other medications used to control pain.

The PALOMA-2 trial

All analyses were performed on the PRO evaluable population: [redacted] patients in the PAL+LET arm and [redacted] patents in the PLACEBO+LET arm. All possible outcomes that can be derived from the FACT-B and EQ-5D questionnaires were pre-specified outcomes in the PALOMA-2 trial. A large number of analyses were conducted. The key findings are as follows:

[redacted]

[redacted]

Health-related quality of life subgroup analyses

Results from a post-hoc subgroup analysis of patients, with and without bone disease baseline, participating in the PALOMA-1 trial are included in the CS. In addition, results from a post-hoc subgroup analysis of patients who were de novo or had disease recurrence >12 months from the end of adjuvant treatment in the PALOMA-1 trial have also been presented.[90] As with the HRQoL analyses for all trial patients, findings between arms in the PALOMA-1 trial were reported to be similar for all measures of pain reported.

The company also assessed the impact of neutropenia on HRQoL for patients in the PALOMA-2 trial in which patients in the PAL+LET arm were classified by neutropenia status. [redacted].

ERG comment on health-related quality of life

Common to trials that report HRQoL outcomes, patients in the PALOMA-1 and PALOMA-2 trials were only asked to complete questionnaires up until the time of disease progression. The number of patients eligible to complete questionnaires decreases with each cycle and the high HRQoL response rates reported by the company in the CS only apply to the number of eligible patients in any given cycle. For example, in the PALOMA-1 trial, it can be observed from the published data[84] that by cycle 16 and cycle 9 of the PAL+LET and LET arms respectively, only 50% of all originally eligible patients completed a questionnaire. The number of eligible patients had fallen to 25% by cycle 25 and cycle 18 in the PAL+LET and LET arms respectively. Thus, in later cycles, the numbers of patients responding are very small and the data are only reflective of the experience of relatively healthy patients. Nonetheless, the data from the earlier cycles in both trials do appear to show there is no difference in HRQoL between treatment arms for patients in either the PALOMA-1 trial or the PALOMA-2 trial.

4.7 Conclusions of the clinical effectiveness section

The primary sources of clinical evidence for this appraisal are the phase I/II PALOMA-1 trial and phase III PALOMA-2 trial. Evidence is presented for PAL+LET versus LET and PLACEBO+LET respectively. The EMA considers that it is reasonable to generalise the clinical effectiveness results associated with LET to other aromatase inhibitors; the ERG concurs with this viewpoint.

All patients in both trials had ABC (and [redacted] had MBC) that had been previously untreated in the metastatic setting. Patients in the trials did not have immediately life-threatening disease and so, if these patients were to be treated currently in clinical practice, they would most likely be given an aromatase inhibitor, as per the treatment of patients in the control arm of both trials. Despite a higher proportion of patients in the PALOMA-1 and PALOMA-2 trials presenting with de novo disease than would be seen in clinical practice (49.1% and 37.2% respectively compared with 5% seen in clinical practice in England and Wales[58,63]), the ERG is generally satisfied that the evidence derived from both trials is generalisable to the patient population in England and Wales described in the scope issued by NICE.

Both trials were international multi-centre RCTs. The PALOMA-2 trial was considered by the ERG to be of superior quality and lower risk of bias than the PALOMA-1 trial as this trial was designed as a double-blind trial (whereas the PALOMA-1 trial was designed as an open-label trial). The PALOMA-2 trial was also much larger than the PALOMA-1 trial and the findings from the PALOMA-2 trial therefore appear to be more robust than those from the PALOMA-1 trial.

Compared with LET or PLACEBO+LET, both trials demonstrated a large improvement in median PFS. The improvement in PFS was generally consistent across subgroups analysed by the company for the PALOMA-1 trial and was generally consistent across subgroups analysed for the PALOMA-2 trial. This included patients presenting with de novo disease and those who had received prior neo(adjuvant) therapy, although the magnitude of the effects differed by subgroups (albeit based on very few numbers of patients, particularly in the de novo subgroup of the PALOMA-1 trial). However, the improvements in PFS did not translate into a statistically significant improvement in median OS for patients in the PALOMA-1 trial and an estimate of median OS is not yet available for patients in the PALOMA-2 trial. It is not clear why there was no gain in OS in the PALOMA-1 trial given there was such a large gain in PFS although it should be noted, the OS data were are immature (37% of deaths) and are from a data cut-off date of 29 November 2013. A possible reason may be attributed to the quality of the PFS data in the PALOMA-1 trial. Investigator assessed PFS findings reported for cohort 1 of the PALOMA-1 trial differed markedly to BICR assessed PFS. This has led the EMA to conclude that only findings from cohort 2 should be considered relevant to the efficacy assessment.

Across the two trials, differences between the treatment arms in terms of safety were mostly attributable to a much higher rate of haematological toxicities, particularly neutropenia in patients treated with PAL+LET. While this included high rates of Grade 3 to 4 neutropenia, for the most part, neutropenia was asymptomatic and reversible with febrile neutropenia being reported by <2% of patients (all incidence occurring in the PALOMA-2 trial). The data suggest

neutropenia rarely results in permanent discontinuation of treatment with PAL+LET. Therefore the safety profile of PAL+LET is considered by the company and ERG to be acceptable. Importantly, compared with LET or PLACEBO+LET, patients remained progression-free for longer and were therefore treated with PAL+LET for longer; despite patients having an increased risk of neutropenia, there were no differences in patients' HRQoL estimates between the trial arms.

5 COST EFFECTIVENESS

5.1 Introduction

This section provides a structured critique of the economic evidence submitted by the company in support of using PAL+LET to treat postmenopausal patients with locally advanced or metastatic, ER+/HER2- breast cancer that has been previously untreated in a metastatic setting. The two key components of the economic evidence presented in the CS are (i) a systematic review of the relevant literature and (ii) a report of the company's de novo economic evaluation. The company has also provided an electronic version of their economic model, which was developed in Microsoft Excel.

5.2 ERG comment on the company's review of cost effectiveness evidence

5.2.1 Objective of cost effectiveness review

The objective of the company's literature search was to identify published estimates of the cost effectiveness of palbociclib to treat postmenopausal women with ER+/HER2- locally advanced or MBC who had received no prior systemic anti-cancer treatment for advanced disease.

Company searches

The company searched MEDLINE, MEDLINE In-Process, Embase, The Cochrane Library (The Health Technology Assessment [HTA] Database and the NHS Economic Evaluation Database only) and EconLit in January 2016. These searches were supplemented, in March 2016, by searches of conference proceedings from the 2014 and 2015 European Breast Cancer Conference, ESMO congress, International Health Economics Association (iHEA) conference and International Society for Pharmacoeconomics and Outcomes Research annual European and International meetings. In addition, in March 2016, the NICE and Scottish Medicines Consortium websites were searched for any relevant HTA submissions. The search strategies employed by the company are provided in Appendix 14 of the CS.

5.2.2 Eligibility criteria used in study selection

The inclusion/exclusion criteria used by the company to facilitate study selection are provided in the CS and reproduced in Table 16.

Table 16 Eligibility criteria for the cost effectiveness systematic review

Domain	Inclusion	Exclusion	Rationale
Population	First-line population: postmenopausal women with ER+, HER2- locally advanced or metastatic breast cancer who have not received any prior systemic anticancer treatment for advanced disease	Population not relevant, or outcomes not reported separately for the population of interest	This is the patient population relevant to the NICE decision problem for this submission
Intervention	Palbociclib	Studies not evaluating palbociclib	This is the intervention specified in the NICE decision problem for this submission
Comparator	Any pharmacological intervention	Non-pharmacological comparators	This encompasses all relevant comparators specified in the NICE decision problem for this submission
Outcomes	The outcomes of relevant study designs, including: costs life years QALYs incremental costs and QALYs ICERs	Studies presenting irrelevant outcomes only	These outcomes encompass the economic outcomes specified as relevant in the NICE decision problem for this submission
Study design	Economic evaluations, specifically one of the following analysis types: cost effectiveness cost utility cost benefit cost minimisation cost consequence	Any other study design	The study designs and publication types specified as eligible for inclusion were those considered most likely to report relevant data for this systematic review
Publication type	Economic evaluations and HTAs	Any other publication type, including non-systematic reviews, editorials and case reports	
	Systematic reviews of economic evaluations were included at the title/abstract screening stage and used for identification of any additional primary studies not identified through the database searches, but were excluded during the full-text review stage.		
Language	English	Any other language	The review team did not have the linguistic capability to review non-English language articles; however, studies were not limited to those conducted in specific geographical locations

ER=oestrogen receptor; HER2=human epidermal growth factor receptor 2; NICE=National Institute for Health and Care Excellence; QALYs=quality adjusted life years; ICER=incremental cost effectiveness ratio; HTA=health technology assessment
Source: CS, Table 47

5.2.3 Included and excluded studies

Ten papers were identified from the company's literature searches; however, none of these met the review inclusion criteria. Nine of the studies were excluded at title and abstract stage; eight were ineligible due to the publication type or study design, and the remaining study was not conducted in the relevant population. The only paper[106] that was screened at full text level was excluded from the review, as the authors did not report economic outcomes.

5.2.4 Findings from cost effectiveness review

No cost effectiveness studies designed to support the use of palbociclib to treat postmenopausal women with ER+/HER2- locally advanced or MBC who had not received any prior systemic anti-cancer treatment for advanced disease were identified during the review process.

5.3 ERG critique of the company's literature review

Full details of the strategies used to locate cost effectiveness evidence were reported in Section 5.1 and Appendix 14 of the CS. The economic searches were conducted in January 2016. This search included population terms but did not include any indication terms; the ERG considers this approach to be appropriate. The search also included an economics filter. The ERG considers that the detail provided by the company, in relation to the literature reviews that were carried out to identify and assess published cost effectiveness evidence (including information on HRQoL, costs and resource use), was very useful.

5.4 Summary and critique of company's submitted economic evaluation by the ERG

The economic evaluation undertaken by the company is designed to compare the costs and benefits (in terms of QALYs) of treatment with PAL+LET versus LET in postmenopausal women with ER+/HER2- locally advanced or MBC. Data from the PALOMA-1 trial have been to estimate survival for patients receiving first-line treatment whilst data from the PALOMA-2 trial have been used to model post-progression survival. Data from the PALOMA-2 trial have also been used to estimate the incidence of AEs and, in conjunction with published figures, HRQoL. Published sources and expert advice have been used to estimate the value of model resource use and cost parameters.

In addition to base case results, the company has also presented results from one-way deterministic, probabilistic and scenario analyses.

5.4.1 Model structure

The company de novo model is a partitioned survival model that comprises three health states; pre-progressed (stable) disease, progression (which is sub-divided into four different states: first, second, and third subsequent lines of treatment and best supportive care [BSC]) and dead. All patients enter the model in the pre-progressed health state and are treated with either PAL+LET or LET. In each cycle patients can either remain in their current health state or, if their disease progresses, move to a worse health state (i.e. a further line of treatment or BSC) or to the death state (see Figure 1). Within the model it is assumed that each post-progression treatment sequence/line lasts for up to six cycles. After completing up to four lines of treatment, it is assumed that patients receive BSC up to the point of death.

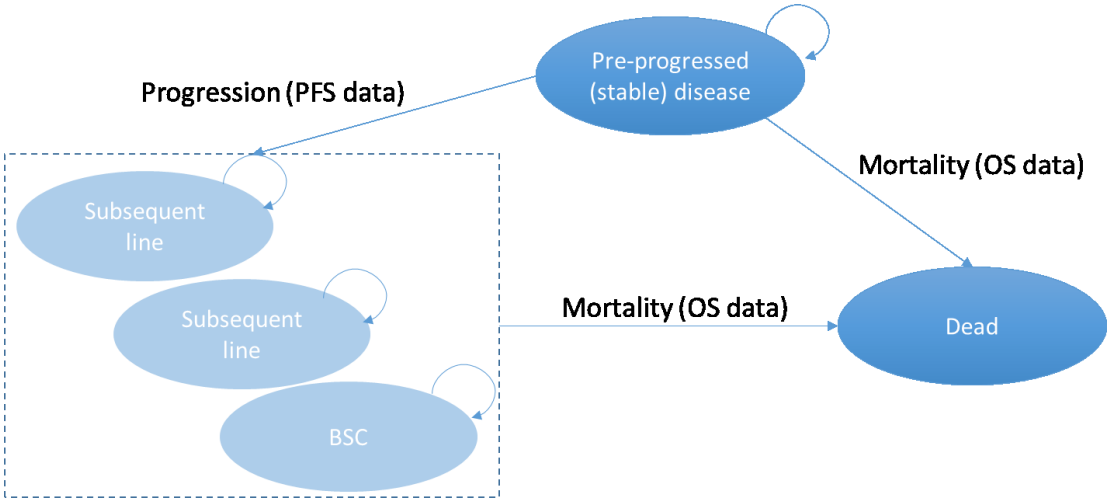

Figure 1 Model schematic

Source: CS, Figure 18

The model cycle length is 28 days (13 cycles per year, 364 days) and, due to the short length of the treatment cycle, a half-cycle correction was not implemented.

The company model structure is similar to that of other models that have been submitted to NICE as part of an STA process that have considered new treatments for advanced or metastatic cancers.[107]

5.4.2 Population

The population reflected in the company model is postmenopausal women with ER+/HER2- ABC who have never received systemic therapy in the LABC/MBC setting (i.e. those receiving first-line treatment).

5.4.3 Interventions and comparators

Intervention

PAL is supplied as a tablet and is used to treat patients in the model in line with its expected EMA marketing authorisation (i.e. 125 mg daily for 21 consecutive days with the subsequent 7 days off treatment until disease progression).

Comparators

It is stated within the final scope issued by NICE that the comparators for this appraisal are aromatase inhibitors; however, LET is the only aromatase inhibitor included as a comparator in the cost effectiveness analysis. The company suggests that, as LET is the most commonly used aromatase inhibitor in the NHS, and as the effectiveness of the other aromatase inhibitors are not significantly different from that of LET, modelling only one of the comparator options detailed in the final scope issued by NICE is justified.

LET is supplied as a tablet and is used to treat patients in the model in line with its EMA marketing authorisation, which reflects the dosage used in UK clinical practice (i.e. 2.5 mg daily, without a break until progression).

Subsequent lines of treatment

Doses of subsequent lines of treatment are not included in the company model. Only the monitoring costs of subsequent lines of therapy are included in the model.

5.4.4 Perspective, time horizon and discounting

The company states that the economic evaluation is undertaken from the perspective of the NHS and PSS (Personal Social Services) and the model time horizon is 40 years. The company states both costs and benefits are discounted at a rate of 3.5% per annum.

5.4.5 Treatment effectiveness and extrapolation

Extrapolation method

To model effectiveness over a lifetime horizon, the company extrapolated survival data from the PALOMA-1 and PALOMA-2 trials. Regression modelling was used to fit parametric curves to K-M data. Six different models were considered: exponential, Weibull, log-normal, log-logistic, Gompertz and generalised gamma. Model selection was based on standard statistical criteria (Akaike and Bayesian information criteria [AIC and BIC respectively]) and clinical plausibility (assessed through consultation with clinical experts and comparison with previously published curves).

Progression-free survival

Data from the PALOMA-2 trial were used as the basis for identifying a statistical model to represent pre-progression survival. In the base case, separate Weibull models were fitted to the PAL+LET and LET arms. Alternative models were explored in sensitivity analyses.

Overall survival

Overall survival data from the PALOMA-2 trial were unavailable and therefore the company based their survival estimates on data from a mix of data from the PALOMA-1 and PALOMA-2 trials. To estimate OS for patients treated with PAL+LET, the company fitted a Weibull distribution to the K-M OS data from the intervention arm of the PALOMA-1 trial.

Results from the PALOMA-2 trial demonstrate a median PFS difference of 10.3 months between the two arms of the trial. However, examination of the Weibull distributions used to model PFS (which were based on data from the PALOMA-2 trial) indicated that the difference in median PFS between PAL+LET and PLACEBO+LET was 9.2 months.

The company model representation of OS for patients receiving LET was then derived by scaling the Weibull distribution used to represent the OS of patients receiving PAL+LET (based on data from PALOMA-1) in such a way as to preserve the 9.2 month median PFS survival gain observed in the PALOMA-2 trial (Figure 2) from PAL+LET.

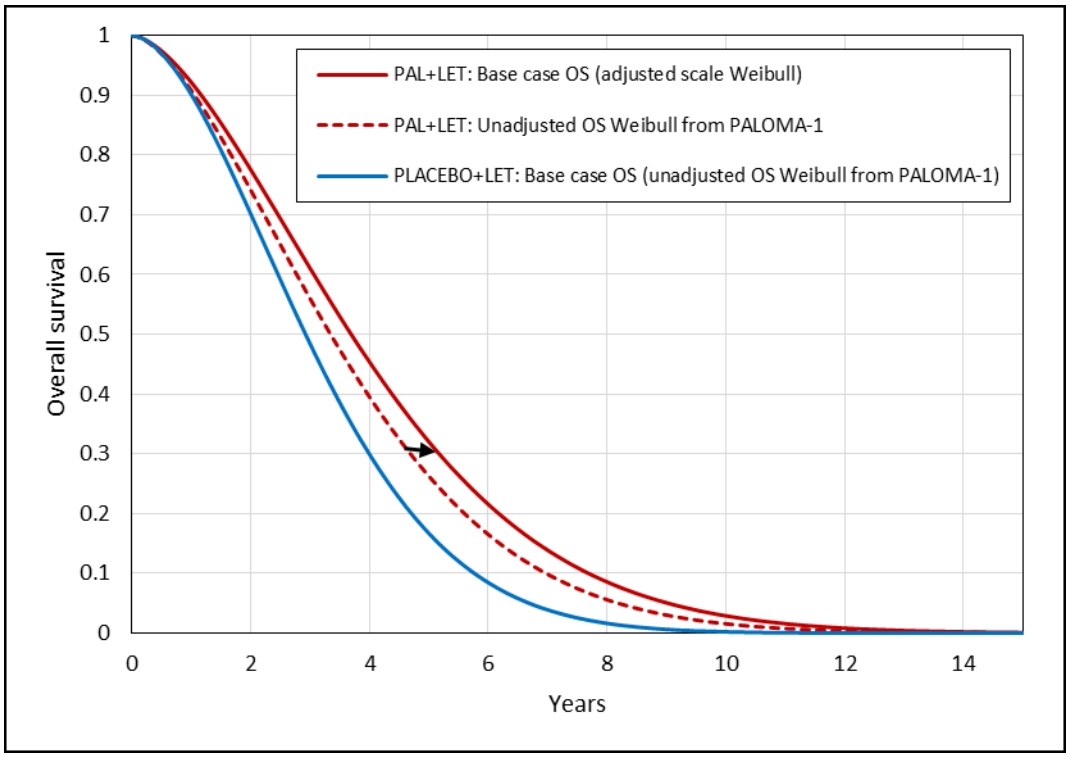

Figure 2 Company overall survival models using PALOMA-1 trial data: base case adjusted Weibull model (PAL+LET) and unadjusted Weibull model (PAL+LET and LET)

LET=letrozole; PAL+LET=palbociclib+letrozole; OS=overall survival
Source: Company model

Efficacy of subsequent treatments

The OS distributions implemented in the company model are based on K-M data from the PALOMA-2 trial. These data incorporate the influence of subsequent treatments and, therefore, no additional modelling was required to represent the effect of subsequent treatments on survival.

5.4.6 Adverse events

The company states that all Grade 3 and Grade 4 AEs observed during the course of the PALOMA-2 trial that have a measurable impact on costs and QALYs are included within their model. The probability of an AE occurring was calculated based on incidence and median exposure to first-line treatments. No account was taken of any AEs experienced as a result of receiving any subsequent therapy, as the inclusion of such AEs would have had a comparable impact on both treatment arms (as the length of time exposed to these treatments was the same for patients in both PALOMA-2 trial arms). Figures relating to the probability of a Grade 3 or a Grade 4 AE occurring in the model are presented in Table 17.

Table 17 Adverse event probabilities used in the company model

	PAL+LET	LET
Probabilities used in the model		
Any Grade 3 AE	44.38%	19.44%
Any Grade 4 AE	8.39%	1.95%

AE=adverse event; LET=letrozole; PAL=palbociclib
Source: CS, Table 60

5.4.7 Health-related quality of life

During the PALOMA-2 trial, HRQoL data were collected using the EuroQol five-dimensions (EQ-5D), three-levels questionnaire. A summary of the utility values used in the company model is presented in Table 18.

Pre-progression utility values

No statistically significant differences in baseline or on treatment EQ-5D index scores were estimated when the company compared results from the PAL+LET and the PLACEBO+LET arms of the PALOMA-2 trial. However, the company used the individual treatment baseline utility values to represent HRQoL for the duration of the pre-progression state ([redacted] for patients receiving PAL+LET and [redacted] for patients receiving LET). The company considers that treatment with palbociclib delivers benefits to HRQoL that may not be captured by the EQ-5D questionnaire (see CS Section 3.2.1 and Appendix 11.8 of this report).

The utility values derived from the data collected during the PALOMA-2 trial include decrements to HRQoL that may be caused by AEs; therefore, in the company base case, no

disutilty adjustments have been applied (as to do so would be considered double counting). However, disutility adjustments (based on data reported in the Lloyd et al (2006)[5] paper) are applied in a scenario analysis.

The company undertook a systematic literature review to identify alternative estimates of utility values that might be used to represent the HRQoL of patients in the pre-progression and post-progression health states. No appropriate alternative utility values were identified.

Post-progression utility value

In the base case, the company assumed that utility values for all subsequent post-progression states (three lines of treatment and BSC) are assumed to be equal. The company considers this assumption to be conservative as, in the PALOMA-2 trial, patients treated with PAL+LET had a utility at baseline that was higher than that of patients treated with LET. The utility value applied throughout all post-progression health states has been calculated using the Lloyd (2006)[5] disease progression decrement. This decrement has been applied to the average baseline utility value which was calculated from data that were collected from patients in both arms of the PALOMA-2 trial.

Table 18 Summary of utility values for cost effectiveness analysis

Health state	PAL+LET		LET		Source
	Mean	95% CI	Mean	95% CI	
Pre-progression	[redacted]	[redacted]	[redacted]	[redacted]	PALOMA-2 EQ-5D data on file
Post-progression (all lines)	0.4492	-	0.4492	-	PALOMA-2 EQ-5D data on file adjusted using Lloyd 2006[5] disease progression multiplier

CI=confidence interval, EQ-5D=EuroQol-five dimensions questionnaire; LET=letrozole; PAL=palbociclib
Source: CS, Table 62

5.4.8 Resources and costs

The company carried out literature searches to identify published papers that reported UK NHS costs, PSS costs and resource use of relevance to a model designed to explore the cost effectiveness of PAL+LET. Only one relevant study[108] was identified. This study[108] was carried out at a single centre in Wales.[108] Details relating to this study are provided in the CS (Table 64).

Drug acquisition costs

The drug acquisition costs (for first-line treatments) used in the company model are detailed in Table 19. Costs associated with subsequent lines of therapy were not included in the model.

Table 19 Drug acquisition costs

Technology	Licensed dose	Package information	Cost per package	Source
PAL	125 mg daily used in model (100 mg and 75mg also available)	125 mg tablets, 21 tablets in pack	Proposed list price: £2,950	Unpublished. Note, the same price for all mg
LET	2.5 mg daily	2.5 mg tablets, 28 tablets in pack	£1.52 (SD: £1.47)	eMIT 2016[109]

LET=letrozole; mg=milligram; PAL=palbociclib; SD=standard deviation
Source = CS Table 65

Drug wastage

Both PAL and LET are available in cycle packs (21 days and 28 days respectively). Once a pack has been opened, another patient cannot use the same pack. Drugs are costed on the basis that each patient in the pre-progressed health state is issued with a pack of PAL and/or LET on the first day of each cycle and, therefore, if the patient ceases treatment at any point before the end of that cycle any unused treatment is wasted.

Monitoring and administration costs

As both PAL and LET are provided in tablet form, the company assumed that there are no costs associated with drug administration.

The company assumed that patients who are treated with PAL require a monthly blood test; the company assumes that monthly monitoring of patients treated with LET is not required. The resource use and monitoring cost associated with monthly blood tests are detailed in Table 20.

Table 20 Resource use and costs for patients receiving LET

Resource use		Source
Assumption	1 full blood count every month	Draft SPC (CS, Appendix 1)
Cost	£3.01	DAPS05 (Haematology outpatient appointment) NHS Reference Costs 2014/15[110]

SPC=summary of product characteristics
Source: CS, Table 66 and Table 67

Health state resource use and unit costs

In the model, the company has assumed that the level of resource depends on the patient's health state and their treatment. The estimates of resource use are based on levels reported in the NICE Clinical Guideline for Advanced Breast Cancer (2009),[31] with adjustments made on the advice of Clinical Nurse Specialists (CNSs) to reflect current NHS practice, and any differences to resource use associated with receipt of different lines of treatment.

In the base case 75% of patients are assumed to receive subsequent treatment on disease progression and that, after each line of subsequent treatment, 75% of patients go on to receive another line of subsequent treatment. The remaining patients move directly to BSC, where they remain until death. To estimate resource use for patients receiving subsequent lines of

treatment, the duration of time spent in each subsequent line of treatment was assumed to be six cycles. This assumption is based on clinical expert opinion that, either by choice or for health reasons, not all surviving patients continue to receive active treatment. Background health state costs are provided in Table 21.

Table 21 Background health state unit costs

Resource use	Unit cost	Source
Community nurse visit	£55.50	PSSRU 2015[111]
Community nurse travel time	£27.75	Assumption
Consultant visit (oncologist) – first visit	£177.83	NHS Reference Costs 2014/15[110]
Consultant visit (oncologist) – follow-up visit	£131.97	NHS Reference Costs 2014/15[110]
GP contact (surgery visit)	£38.50	PSSRU 2015[111]
GP contact (home visit)	£198.00	PSSRU 2015[111]
Clinical nurse specialist	£86.00	PSSRU 2015[111]
Social worker visit	£67.00	PSSRU 2015[111]
Social worker travel time	£33.50	Assumption
Palliative care	£55.50	Assumption
CT scan	£121.68	NHS Reference Costs 2014/15[110]
Therapist (community occupational therapist and hospital occupational therapist)	£39.00	PSSRU 2015[111]
Physiotherapist (hospital occupational therapist)	£36.00	PSSRU 2015[111]
Lymphoedema nurse	£55.50	Assumption

CT= computerised tomography scan; GP=general practitioner; PSSRU=personal social services research unit
Source: CS, Table 69

The company assumed that resource use during the final 2 weeks of life (terminal care) is the same for all patients but differs depending on whether this period is spent in hospital, in a hospice or at home. The proportion of patients assumed to reside in hospital, hospice and at home, along with the unit costs associated with spending 2 weeks in any of these settings, are shown in Table 22.

Table 22 Terminal care resource use and unit costs (last 2 weeks of life)

Setting	Percentage cohort in each setting	Source for clinical setting	Unit cost	Source unit cost
Hospital	40%	NICE CG 81 Package 3[31]	£5,521.73	NICE CG 81 Package 3[31] unit costs, inflated from 2006/07 to 2014/15 values
Hospice	10%		£6,883.98	
Home	50%		£2,848.87	

CG=clinical guideline; NICE=National Institute for Health and Care Excellence
Source: CS, Table 70

Adverse events

Within the company model, patients who have multiple AEs occurring simultaneously within a single cycle only incur one cost (and one disutility value.

Neutropenia was the most common Grade 3 and Grade 4 event experienced by patients in the PALOMA-2 trial and the estimated resource use required to treat this AE is used within the company model to represent the resource use required to treat all Grade 3 and Grade 4 AEs. The cost is implemented at the start of each cycle and is assumed to last no more than than one cycle. The resource use assumptions and unit costs used in the company model are detailed in Table 23.

Table 23 Resource use assumptions and unit costs for grade 3 or 4 adverse events

Neutropenia	Resource use assumption	Unit cost	Note about unit cost	Source
Grade 3	1 oncologist visit per event (20 min visit) for patient management and dose modification	£43.99	WF01A service code 800 Clinical Oncology (Previously Radiotherapy) Non-Admitted Face to Face Attendance, Follow-up	NHS Reference Costs 2014/15[110]
Grade 4	1 oncologist visit per event (30 min visit) for patient management and dose modification	£65.99		

Source: CS, Table 71

5.4.9 Cost effectiveness results

Estimates, generated by the company model, for total costs, life years gained (LYG), QALYs and incremental cost effectiveness ratios (ICERs) per QALY gained for the comparison of the cost effectiveness of treatment with PAL+LET versus LET are shown in Table 24. In the base case, treatment with PAL+LET generates more benefits than treatment with LET (+0.78) and +0.63 QALYs) but at an increased cost of £94,853. The company base case ICER for the comparison of treatment with PAL+LET versus LET is £150,869 per QALY gained.

Table 24 Base case deterministic results for PAL+LET vs LET

Technologies	Total costs	Total LYG	Total QALYs	Incremental			ICER per QALY gained
				Costs	LYG	QALYs	
LET	£21,843	3.02	1.77				
PAL+LET	£116,696	3.79	2.40	£94,853	0.78	0.63	£150,869

LET=letrozole; LYG=life years gained; PAL=palbociclib; QALYs=quality adjusted life years
Source: CS, Table 74

A summary of the predicted resource use for each of the cost categories is presented in Table 25. Over 97% of the difference in costs between the intervention and comparator technologies is due to the difference in the costs of the first-line therapies.

Table 25 Summary of predicted resource use by category of cost

Item	Cost		Increment	Absolute increment	% absolute increment
	PAL+LET	PAL			
Drug acquisition costs	£92,101.27	£31.68	£92,069.59	£92,069.59	97.07%
Within cycle wastage costs	£0.00	£0.00	£0.00	£0.00	0.00%
Drug administration costs	£0.00	£0.00	£0.00	£0.00	0.00%
Drug monitoring costs	£93.79	£0.00	£93.79	£93.79	0.10%
AE costs	£782.02	£205.10	£576.92	£576.92	0.61%
Pre-progression health state costs	£5,290.91	£3,533.90	£1,757.01	£1,757.01	1.85%
Second-line treatment background health state costs	£495.84	£626.26	-£130.42	£130.42	0.14%
Third-line treatment background health state costs	£791.83	£982.17	-£190.35	£190.35	0.20%
Fourth-line treatment background health state costs	£1,016.85	£1,223.99	-£207.14	£207.14	0.22%
BSC	£12,365.25	£11,366.38	£998.86	£998.86	1.05%
Terminal care	£3,758.38	£3,873.67	-£115.29	£115.29	0.12%
Total	£116,696.13	£21,843.16	£94,852.97	£94,852.97	**100.00%**

AE=adverse events; BSC=best supportive care; LET=letrozole; PAL=palbociclib
Source: CS, Table 79

5.4.10 Sensitivity analyses

Deterministic sensitivity analysis

The company carried out one-way sensitivity analyses to explore the sensitivity of model results to variations in the magnitude of 12 model inputs. Results are presented in the CS as a tornado diagram, which is reproduced in Figure 3. The results show that varying the OS and PFS parametric model coefficients has the biggest effect on the company's cost effectiveness results.

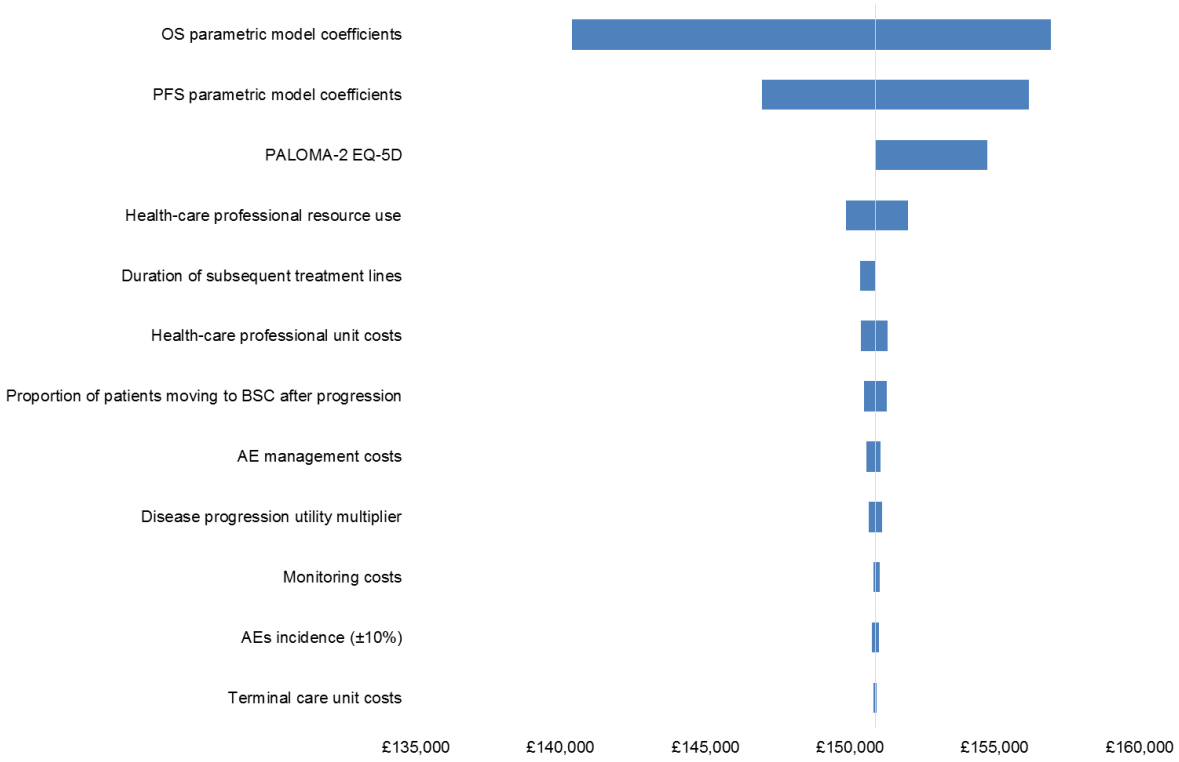

Figure 3 Tornado diagram of one-way sensitivity analyses (PAL at list price)
AE=adverse event; BSC=best supportive care; OS=overall survival; PAL=palbociclib; PFS=progression-free survival
Source: CS, Figure 28

In addition, the company carried out a further 10 one-way sensitivity analyses to explore the effect on model results of varying model assumptions. Results displayed in Table 26 show that, apart from the scenario in which a 5-year time horizon was implemented (which the company states is too short to fully capture all of the relevant costs and benefits in this patient population) amendments to OS and PFS assumptions have the largest influence on the resultant ICERs per QALY gained.

Table 26 List of sensitivity analyses varying model assumptions (PAL at list price)

Scenario	Parameter varied	Incremental costs	Incremental QALYs	Deterministic ICER
13	Use the Beauchemin linear regression method	£100,711	0.86	£116,806
14	Use unadjusted OS from PALOMA-1- Weibull for both arms	£91,384	0.49	£187,881
15	Use unadjusted OS from PALOMA-1 - Log-logistic for both arms	£95,112	0.63	£150,273
16	PFS parametric models - Gompertz for both arms	£84,696	0.44	£193,312
17	AEs: include AE disutility values	£94,853	0.57	£166,954
18	Model horizon: 5 years	£84,718	0.42	£199,943
19	Model horizon: 10 years	£94,201	0.61	£153,485
20	Model horizon: 15 years	£94,834	0.63	£150,934
21	Exclude discounting costs and benefits	£102,608	0.73	£140,954
22	Baseline utility (pre-progressed state): assume same value	£94,853	0.57	£166,802
23	Disease progression multiplier: use Nafees[112] value	£94,853	0.63	£150,334
24	Assume gradual utility decrease with every line of progression	£94,853	0.62	£152,781
25	Assume no post-progression sequential modelling: direct move to BSC	£94,121	0.63	£149,704
26	Use the health state costs from the NICE TA295 submission[113]	£94,522	0.63	£150,342

AE=adverse event; BSC=best supportive care; CI=confidence interval; ICER=incremental cost effectiveness ratio; OS=overall survival; PAL=palbociclib; PFS=progression-free survival; QALY=quality adjusted life year
Source: CS, Table 84

Probabilistic sensitivity analysis

The company undertook a probabilistic sensitivity analysis (PSA) to derive the mean ICER per QALY gained for the comparison of PAL+LET versus LET. The PSA was run for 1000 iterations. Results from the deterministic analysis and the PSA are shown in Table 27. The probabilistic ICER per QALY gained for PAL+LET versus LET is £151,058, which is very similar to the deterministic ICER per QALY gained (£150,869).

Table 27 PSA results for PAL+LET versus LET (PAL at list price)

	Incremental costs	Incremental QALYs	ICER per QALY gained
Deterministic result	£94,853	0.63	£150,869
Average value from PSA	£94,951	0.63	£151,058

ICER=incremental cost effectiveness ratio; LET=letrozole; PAL=palbociclib; QALY=quality adjusted life year; PSA=probabilistic sensitivity analysis
Source: CS, Table 80

The results from the PSA are presented as a scatter plot (cost effectiveness plane) in Figure 4. An examination of this figure shows that, at a cost effectiveness threshold of £30,000 per QALY gained, PAL+LET has a 0% probability of being cost effective compared with LET. The

cost effectiveness acceptability curve (CEAC) is shown in Figure 5. It is not until beyond a threshold of £100k per QALY that PAL+LET has any probability of being cost effective compared to LET.

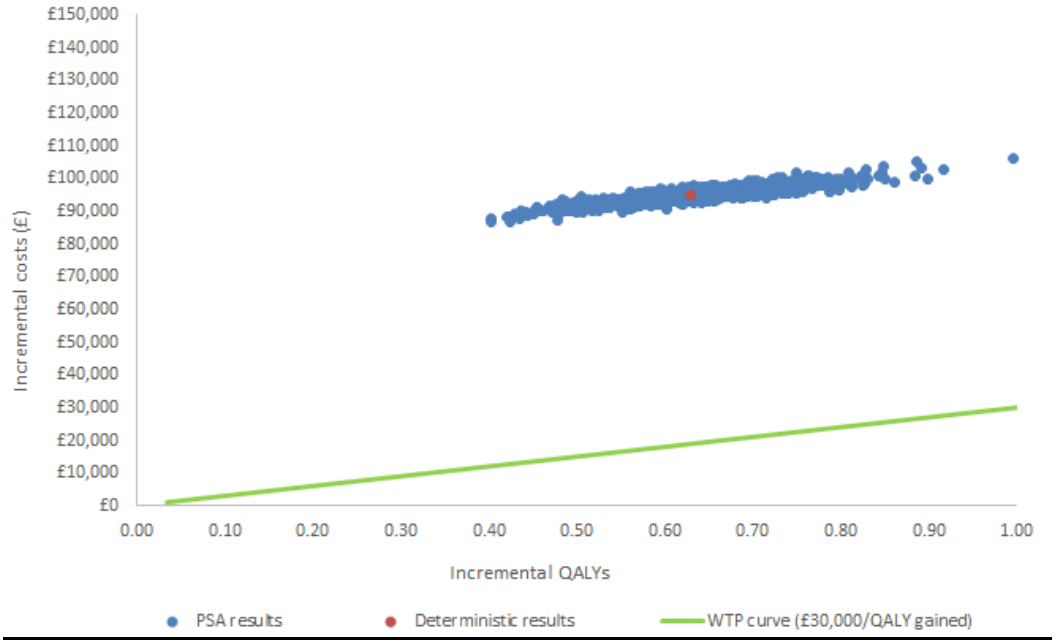

Figure 4 Cost effectiveness plane for the comparison of PAL+LET vs LET (PAL at list price)

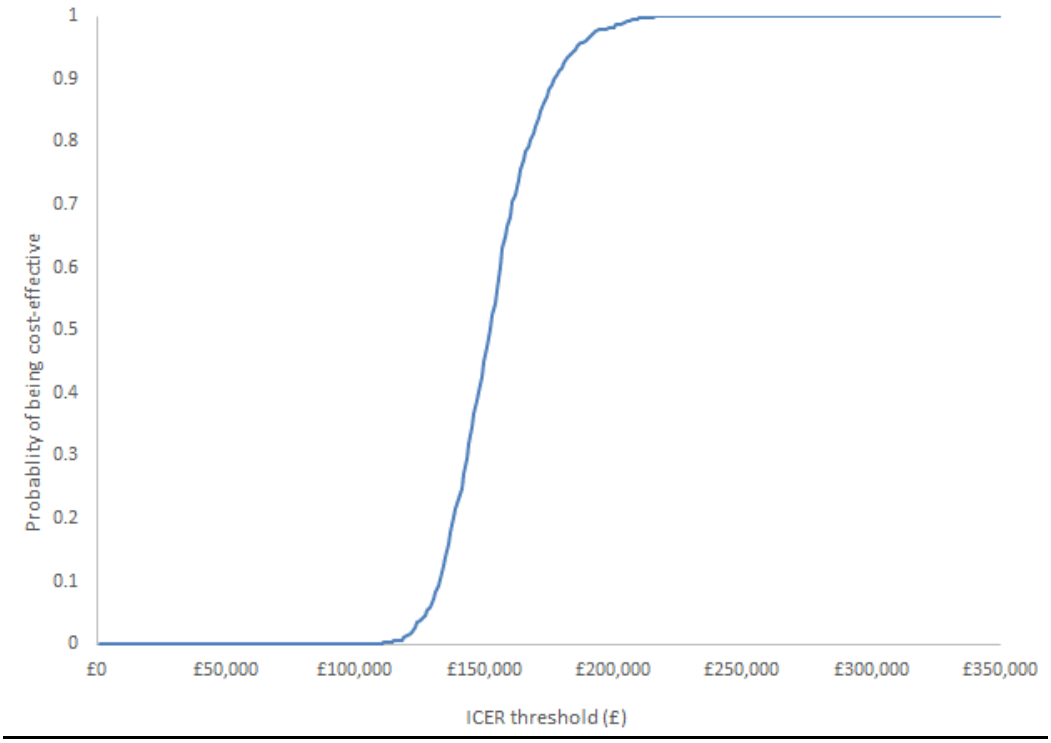

Figure 5 Cost effectiveness acceptability curve for PAL+LET vs LET (PAL at list price)

5.4.11 Scenario analyses

The company presented scenario analyses in two parts; the first five scenarios explored assumptions that drive the base case ICER beyond a £30,000 per QALY threshold and the second five scenarios demonstrate the impact to model results of combinations of amendments to parameter values or assumptions.

Results in Table 28 show the changes in ICERs per QALY that result from varying assumptions. The removal of the OS gain for PAL+LET increases the ICER per QALY gained by approximately £162,000.

Table 28 Exploratory scenario analyses varying model assumptions (PAL at list price)

#	Assumptions varied	Change in ICER from base case
Base case deterministic ICER		£150,869 per QALY
27	Only PFS gain for PAL (10.3 months) No OS gain for PAL (0 months)	+ £161,766
28a	Increased OS improvements with PAL: a 5-year incremental gain	- £89,047
28b	Increased OS improvements with PAL: a 5-year incremental gain, *but removing post-progression costs*	- £108,075
29	Increase in utility of +0.1 for patients in the PFS state	- £16,735
30	A comparator with the same monthly acquisition costs *(i.e. fixed cost of £2,951.52 per month, but only for respective treatment durations)*	- £97,795
31	Reduced treatment duration by 12 months in each arm *(PFS reduced from 15.7 to 3.7 months for LET, and from 24.9 to 12.9 months for PALb)*	- £64,450

ICER=incremental cost effectiveness ratio; LET=letrozole; OS=overall survival; PAL=palbociclib; PFS=progression-free survival; QALY=quality adjusted life year
Source: CS, Table 85

The ICERs per QALY gained displayed in Table 29 result from implementing combinations of changes to baseline assumptions. Scenario 36 is the only scenario that generates an ICER below a threshold of £30,000 per QALY. In this scenario, the cost of LET is assumed to be the same as that for PAL, there are no costs associated with post-progression, there is an OS gain of 24 months for patients receiving PAL+LET compared with LET and the utility value associated with being in the pre-progression state is increased by 0.1.

Table 29 Combining scenarios to evaluate exploratory ICERs per QALY gained (PAL at list price)

#	Assumptions changed	Incremental costs	Incremental QALYs	ICER per QALY gained
32	Comparative monthly acquisition costs Value of PFS utility increase (+0.1) *No change to base case OS assumption*	£33,013	0.82	£47,187
33	Comparative monthly acquisition costs (#30) Value of PFS utility increase (+0.1) Incremental OS gain of 12 months	£35,734	0.82	£43,819
34	Comparative monthly acquisition costs Value of PFS utility increase (+0.1) Incremental OS gain of 12 months Removal of post-progression costs	£33,013	0.82	£40,482
35	Comparative monthly acquisition costs (#30) Value of PFS utility increase (+0.1) Incremental OS gain of 24 months	£45,963	1.27	£36,194
36	Comparative monthly acquisition costs Value of PFS utility increase (+0.1) Incremental OS gain of 24 months Removal of post-progression costs	£33,013	1.27	£26,996

PFS=progression-free survival; OS=overall survival; PAL=palbociclib; ICER=incremental cost effectiveness ratio; QALY=quality adjusted life year
Source: CS, Table 86

5.4.12 Model validation and face validity check

The company took a number of steps to try and ensure the validity of the extrapolations and parameter values employed in their model:

- Utility values were sourced directly from the phase III trial (PALOMA-2) and from a source[5] established in previous STA submissions for people with ABC[113,114]
- Clinical opinion was sought to validate the estimates of resource use, and national databases (NHS Reference Costs,[110] PSSRU[111] and eMIT[109]) were used to source costs
- Detailed modelling of subsequent treatment lines allowed the complexity of subsequent therapies to be explored
- Validation of the model and its findings were undertaken internally by the model developers on behalf of the company and by an external independent health economist.

5.5 Detailed critique of the company's economic model

5.5.1 NICE Reference Case checklist

Table 30 NICE Reference Case checklist completed by ERG

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	Letrozole is the only aromatase inhibitor compared to palbociclib although there are others available for the indication described in the scope
Perspective costs	NHS and PSS	PSS costs were not fully considered in the CS
Perspective benefits	All health effects on individuals	Yes
Form of economic evaluation	Cost effectiveness analysis	Yes
Time horizon	Sufficient to capture differences in costs and outcomes	Yes
Synthesis of evidence on outcomes	Systematic review	The company uses data from PALOMA-1 and PALMOA-2 trials to estimate survival and HRQoL estimates for initial therapy. A systematic review was conducted to estimate the outcomes of subsequent therapy
Outcome measure	Quality adjusted life years	Yes
Health states for QALY	Described using a standard and validated instrument	Yes
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	The company used an annual discount rate of 3.5% per annum for costs and benefits. Discounting is implemented per cycle, rather than on an annual basis, within the model
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

HRQoL=health-related quality of life; PSS=Personal Social Services

5.5.2 Drummond checklist

Table 31 Drummond critical appraisal checklist completed by the ERG

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?	Partly	In the model, PFS and OS were estimated using survival data from different trials. Projecting OS from PFS data from a different trial adds uncertainty to the effectiveness evidence used in the model (and therefore adds uncertainty to the size of the ICER per QALY gained)
Were all the important and relevant costs and consequences for each alternative identified?	Partly	Costs of subsequent therapy and AEs whilst on subsequent lines of treatment are not included in the model
Were costs and consequences measured accurately in appropriate physical units?	Partly	The days of the year modelled equated to 364 rather than the ERGs preferred 365.25. The HRQoL multiplier for progressed disease was implemented incorrectly. The annual incidence rate of AEs was implemented each cycle in the model.
Were the cost and consequences valued credibly?	Partly	An oncologists consultation was used as the cost to treat neutropenia taken from NHS reference costs and was assumed to last 60 minutes. This cost was weighted according to the Grade of neutropenia with Grade 3 incurring a 20 minute appointment and Grade 4 a 30 minute appointment thus cutting the reference cost by two-thirds and half respectively.
Were costs and consequences adjusted for differential timing?	Yes	Costs and benefits were not discounted on an annual basis
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	

PFS=progression free survival, OS=overall survival, AE=adverse events

5.6 Detailed critique and exploratory analyses undertaken by the ERG

The company's Microsoft Excel spreadsheet model is constructed according to conventional practice and is generally implemented correctly.

5.6.1 Key issues in the company model

The two fundamental issues relating to the company's cost effectiveness model are: the absence of OS data from the PALOMA-2 trial; and issues regarding the reliability of survival data from the PALOMA-1 trial (Section 4.4). The company's attempts to overcome the lack of OS data from the PALOMA-2 trial are methodologically flawed, and result in inconsistencies (i) within the survival data used in the company model and (ii) between the assumptions underpinning the company's survival projection methods and their implementation.

Specific issues in the model connected to the lack of reliable survival data are:

- use of data from two different trials (PFS from PALOMA-2 and OS from PALOMA-1) introduces inconsistencies in the model estimates of survival
- no evidence to support the assumption that 100% of PFS gain for treatment with PAL+LET versus LET translates into OS gain
- assumption that there is no difference in PPS between treatment with PAL+LET and treatment with LET when evidence suggests that PPS is shorter for patients treated with PAL+LET than for those treated with LET
- method used to adjust OS data from the PALOMA-1 trial to incorporate the assumptions of (i) PFS gain is equal to OS gain and (ii) zero PPS gain, results in neither of these assumptions holding in the model.

Other issues identified by the ERG include:

- using PFS as a proxy for time on treatment, when TTD data provide a more accurate basis for estimating treatment acquisition costs
- calculating pre-progression health state utility values using data from the PALOMA-2 trial ITT population when using values collected from just the European population would have been more relevant to the NHS
- using different pre-progression health state utility values to reflect the quality of life of patients in the intervention and comparator arms when evidence from the PALOMA-2 trial indicated that there was no statistically significant difference between the two values
- absence of half-cycle correction
- incorrect use of a published method for calculating a post-progression health state utility value
- unjustified proportionate use of a NHS Reference Cost for costing the treatment of AEs
- incorrect calculation of the incidence of AEs
- discounting on a per cycle rather than on an annual basis.

The ERG has also included a sensitivity analysis which allows investigation of the impact of including the drug acquisition and administration costs associated with subsequent lines of treatment within the model. These costs are not included in the company's base case model.

There are no OS data available from the phase III PALOMA-2 trial. The company has modelled patient survival using PFS data from the PALOMA-2 trial and (adjusted) OS data from the smaller, phase I/II PALOMA-1 trial. However, using PFS from one trial and OS from another is methodologically flawed as it assumes independence between the outcomes. PFS and OS are not independent measurements; they are taken from the same individuals at different times. There is a relationship between PFS and OS are within trials because the data points come from the same set of individuals, however, the nature of their relationship is not necessarily generalisable between trials or across indications.[26]

The company's implementation of the assumptions that PFS gain for treatment with PAL+LET translates into equal OS gain and that PPS is equal for patients treated with either PAL+LET or LET is flawed. Neither of these assumptions hold in the model, as the company's method of adjusting OS to ensure median OS gain equals median PFS gain results in a mean PPS gain for treatment with PAL+LET (and thus a greater OS gain than PFS gain for treatment with PAL+LET).

The assumptions that OS gain for treatment with PAL+LET is equal to PFS gain and that there is no difference in PPS between treatment with PAL+LET and treatment with LET also ignore a pertinent feature of the data from the PALOMA-1 trial: that patients treated with PAL+LET seem to have a shorter life expectancy after progression than those treated with LET.

The ERG has investigated alternative methods of modelling of time-to-event data using PFS and OS from the PALOMA-1 trial only, in order to maintain consistency between PFS and OS. This method is also subject to uncertainties, as the data from the PALOMA-1 trial used for modelling has limitations and the results based on data from the PALOMA-1 trial should be treated with caution (Section 4.4); despite these limitations, the ERG considers using PFS data and OS data from the same trial to be a more methodologically sound approach than the one taken by the company. The ERG notes that the EMA has identified discrepancies between the investigator assessed and BICR-assessed PFS data from the PALOMA-1 trial (Table 8) and has declared only part of the data from that trial to be relevant for efficacy assessment. In light of the EMA's view, the ERG has also provided a scenario analysis in Section 5.6.13 to investigate the use of PFS data from the PALOMA-2 trial.

The company also includes with arguments alongside its base case cost effectiveness analysis to suggest that the current NICE methodology[115] for estimating cost effectiveness underestimates the benefit of the intervention. The ERG does not agree that the NICE

methodology is especially punitive to the intervention in this submission nor that the scenarios provided by the company to address its concerns are meaningful. The assumptions and scenarios put forward by the company are examined in detail in the appendices to this ERG report, Section 10.9.

5.6.2 Re-censoring Kaplan-Meier data

During the clarification process, the ERG requested that the company provide K-M data re-censored using the following rules:

- Patients without a documented event (TTD, PFS, OS) at the point of data cut-off should be re-censored at data cut-off
- Patients who have withdrawn from the trial for any reason and are no longer considered to be part of the trial should be re-censored at the time of withdrawal.

The conventional censoring rule applied to survival data is to censor on the date of last known contact any patients who have not experienced a given event (treatment discontinuation, disease progression, death) at the time of data cut-off. However, this rule can distort results when the data are immature. The ERG requested during the clarification process that K-M data be re-censored to limit potential bias from the application of the conventional censoring rule.

When trials are stopped early or are subject to early analysis, the conventional censoring rule (censor when last contacted/reviewed) always understates the time patients are exposed to risk but is much less likely to understate events, especially deaths. That is, at the time of an interim or early data cut-off, there are many patients still at risk in the trial who are still being followed up beyond data cut-off and will feature in later analyses, but who are censored weeks or months before data cut-off in an interim analysis because that is the last time that they were contacted. But, if a patient dies between the time of their last contact and the time of data cut-off, that death would likely still be recorded as an event. Thus, in the period between last tumour assessment and data cut-off, there may be fewer people recorded at risk than there are in reality, whereas the number of events such as deaths will still likely reflect the true frequency.

The result is that the inter-event period hazard rates calculated by the K-M algorithm are exaggerated when multiple patients are censored in any period. The resulting K-M estimated time-to-event trends may therefore be distorted by ‘informative censoring’ (patients are more likely to be censored early if they are still alive at data cut-off) and poorly reflect the true profile of time-to-event hazards.

All of the ERG's analyses of PFS, PPS, OS and TTD are based on re-censored K-M data. The company's analyses of PFS, PPS, OS and TTD are based on K-M data censored according to the conventional rule.

5.6.3 Time-to-event evidence: overall survival and post-progression survival

The company's modelling of OS in the base case is informed by the assumption that 100% of PFS gain translates into OS gain and that there is no difference in PPS between treatment with PAL+LET and treatment with LET. This is an important assumption because patients continue to accrue QALYs and costs beyond progression that can have a substantial effect on the overall ICER per QALY gained. If there is no difference in PPS between the two treatments, the costs and benefits of the drug are limited to those that accrue in PFS. The ERG does not agree that the company's assumption is justified.

The company provides no evidence for the assumption of zero PPS gain. The assumption of zero PPS gain is not even a conservative one, as evidence from the PALOMA-1 trial indicates that PPS is shorter for treatment with PAL+LET than for treatment with LET (a PPS loss). Re-censored K-M data provided by the company during the clarification process indicate that mean PFS gain in the PALOMA-1 trial, until the data cut on 29 November 2013, was [redacted] months and mean OS gain was [redacted] months. Mean PPS *loss* for treatment with PAL+LET was [redacted] months. Although data are sparse (18 deaths in the post-progression state in the PAL+LET arm and 26 in the LET arm),

Figure 6 shows that patients treated with LET in the PALOMA-1 trial tend to live longer after progression than patients treated with PAL+LET.

Figure 6 PPS K-M data for PAL+LET and LET (PALOMA-1)

Source: Clarification response B4

To implement the assumption of zero PPS gain, the company has attempted to reconcile OS data from the PALOMA-1 trial and PFS data from the PALOMA-2 trial. This approach is methodologically flawed, as PFS and OS data are measurements from the same set of individuals in a trial and so are not independent of one another. The company fitted separate Weibull models to data from both arms of the PALOMA-1 trial, but adjusted only the curve for the treatment with PAL+LET in order to increase the modelled median OS gain so that it matched median PFS gain from the PALOMA-2 trial. The company justifies leaving the OS curve unadjusted for treatment with LET by comparing it to the results of other trials in the published literature.[40,42,116] However, the company does not compare the relationship of PFS to OS for treatment with LET in these trials.

Figure 7 shows that there is a pronounced difference between PFS for the LET arm of the PALOMA-1 and PALOMA-2 trials, but that the PAL+LET arms in the two trials are similar. Given that the difference between investigator assessed PFS in the LET arms of the PALOMA-1 and PALOMA-2 trials is substantial, the ERG does not consider that the company is justified in leaving OS for treatment with LET unadjusted in order to create an OS curve to fit alongside PFS modelled from the PALOMA-2 trial.

Figure 7 Comparison of PFS K-M data from the PALOMA-1 and PALOMA-2 trials

Source: Clarification response B4

The company's assumption of zero PPS gain is flawed when implemented in the model. The company has adjusted the OS curve fitted to data from the PALOMA-1 trial for treatment with PAL+LET so that median OS gain in the model equals median (modelled) PFS gain from the

PALOMA-2 trial. This method does not, however, result in equality between *mean* OS gain and *mean* PFS gain. Because of the way the shape and scale parameters interact in the Weibull model, increasing the median of a curve to a predefined level has a proportionately larger effect on the mean value of that same curve. This means that, by adjusting projected OS for treatment with PAL+LET, the company model actually includes a small (0.49 months) gain in PPS for treatment with PAL+LET. The appendices to this ERG report, Section 10.8, include a more detailed discussion of the effect of adjusting a Weibull model.

The company has attempted to justify its extension of OS for treatment with PAL+LET beyond what is seen in the PALOMA-1 trial with reference to, first, issues of potential confounding in the PALOMA-1 trial and, second, literature identifying a correlation between PFS and OS in advanced breast cancer. The company notes that OS was a secondary outcome measure in the PALOMA-1 trial and that data are immature, and states that the study was substantially underpowered to detect statistically significant differences in OS. The ERG understands by this that the company is arguing that OS data from the PALOMA-1 trial are too flawed to be used for modelling purposes.

The ERG agrees with the company that the PALOMA-1 data have limitations for modelling. However, the company's approach, first, does little to mitigate the problems inherent in the OS data from the PALOMA-1 trial and, second, adds further uncertainties by adjusting the model for treatment with PAL+LET. The company still uses data from the LET arm to model OS for treatment with LET without adjustment and uses the shape of the OS data from the PAL+LET arm to model OS for treatment with PAL+LET. The only amendment the company makes to the OS data from the PALOMA-1 trial is an adjustment of the scale parameter in the Weibull model for treatment with PAL+LET.

ERG exploratory analyses

The ERG considers it unnecessary to introduce further uncertainties into the model by adjusting the OS data from the PALOMA-1 trial, especially when there are already concerns about the robustness of OS K-M data (few recorded events, old data cut) from the PALOMA-1 trial. The ERG's preferred approach to projecting time-to-event data is based on using the re-censored K-M data directly from the PALOMA-1 trial and appending a parametric projection beyond the limits of the trial data to project OS across the model time horizon.

The ERG analysed the re-censored OS K-M data provided by the company during the clarification process (

Figure 8) and did not observe a statistically significant difference between the two arms of the trial (log rank test p=0.488, Mann-Whitney U p=0.734). The ERG notes that the PALOMA-1 trial had not been powered to detect differences in OS, and so considered it appropriate to

produce separate projections for the intervention and the comparator. However, the difference in the ERG's revised estimates of mean OS for the two treatments should be treated with caution, as they are based on data that are not statistically significantly different.

Since OS hazards are proportional after the curves cross at approximately 8 months (Section 10.2), the ERG concluded it was justified to pool the data to produce a more robust estimate of the overall OS trend than could have been found by modelling the arms separately, before applying HRs from a Cox PH regression analysis (of data after the crossing of the curves at 8 months) to the pooled trend to fit separate projections.

Figure 8 OS K-M data for patients treated with PAL+LET and LET (PALOMA-1 trial)

LET=letrozole; PAL+LET=palbociclib+letrozole; OS=overall survival
Source: Clarification response B4

The pooled OS data from the PALOMA-1 trial exhibit apparently increasing hazards over time, which can in fact be modelled as two sections of constant, but different, hazards that change at around 20 months. These constant hazards are represented by straight lines in the cumulative hazard plot in Figure 9 and translate into piecewise exponential OS estimates.

Figure 9 Cumulative hazard plot of pooled OS with two-part exponential trend

OS=overall survival
Source: Clarification response B4, ERG calculations

The ERG used HRs from the Cox PH regression analysis to adjust the exponential model from the second half of the pooled analysis to forecast OS for treatment with PAL+LET and treatment with LET. The ERG then fitted these adjusted exponential tails to the relevant OS K-M data for the intervention and comparator (Figure 10).

Figure 10 ERG OS projections and company model base case OS

LET=letrozole; PAL+LET=palbociclib+letrozole; OS=overall survival
Source: Company model; Clarification response B4; ERG calculations

The ERG's revised OS model for treatment with PAL+LET yields lower estimates than the company's model until around 8 years, after which the ERG's model estimates higher OS than the company base case for patients treated with PAL+LET. The ERG's revised model also yields lower estimates of OS than the company base case in the early part of the model for patients treated with LET, but yields higher estimates than the company base case after approximately 4 years.

Mean OS in the ERG's revised model is 47.7 months for PAL+LET and 41.2 months for LET, which gives a projected mean OS gain of 6.6 months for treatment with PAL+LET. This is in comparison to a mean gain of 11.2 months in the company base case. The ERG notes that this projected OS gain is based on data whose means are not statistically significantly different, therefore there is considerable uncertainty in the estimate. Applying the ERG's revised OS estimates in the model increases the ICER per QALY gained by £38,441 to £189,310.

5.6.4 Time-to-event evidence: progression-free survival

The two key problems with the company's estimates of PFS are: first, that it uses data (from the PALOMA-2 trial) to inform its modelling of PFS that are inconsistent with the data (from the PALOMA-1 trial) used to model OS; and second, that the Weibull model used in the base case produces implausible results.

The ERG considers it methodologically sound to use data from the same trial to estimate PFS and OS, as this approach maintains consistency between PFS, PPS and OS within the model.

The Weibull models used by the company to model PFS for treatment with PAL+LET and treatment with LET each have monotonically increasing hazards. This means that, the longer a patient remains progression free, the more likely they are to progress or die than they were previously (

Figure 11). The logic here is that patients who have done well following treatment, either because of the treatment itself or because of some underlying characteristic, and who have lived for many years after beginning treatment are actually at greater risk of progression (or death) than patients who were sicker or less responsive and died earlier – that is, the further a patient is from randomisation, the more likely they are to progress or die. The impact of increasing general mortality due to age only accounts for a small proportion of these increasing hazards, so the model effectively forecasts that patients will be at greater risk from the disease several years after randomisation than they were when first diagnosed with advanced or MBC. The ERG considers the phenomenon of monotonically increasing hazards, continued over the 40 years of the model time horizon, to be implausible.

Figure 11 Hazard profiles for company base case PFS

LET=letrozole; PLACEBO+LET=placebo+letrozole; PFS=progression free survival
Source: Company model; ERG calculations

ERG exploratory analyses

The ERG considers it preferable to use data from the PALOMA-1 trial as the basis for modelling PFS to maintain consistency with the OS data from the PALOMA-1 trial used for modelling survival. The ERG acknowledges that the data from the PALOMA-1 trial have some limitations (Section 4.4). The ERG urges caution in the interpretation of its revised PFS estimates due to the unreliability of the PFS data from the PALOMA-1 trial.

The ERG prefers to use direct trial K-M data, when available, to model early events and only use later data to model a projection once a long-term trend has been established. This means that early features of the data that can be awkward to model parametrically, such as deaths due to AEs or administrative issues such as time to first assessment, are captured by the trial data. It also means that the most accurate data available are used and no assumptions are required that add to the uncertainty in the model.

The company provided the ERG with re-censored investigator assessed PFS data from the PALOMA-1 trial during the clarification process. Mean PFS gain for patients treated with PAL+LET versus LET in the PALOMA-1 trial was [redacted] months.

Examination of the re-censored K-M data reveals clear exponential trends in both the PAL+LET and LET arms of the PALOMA-1 trial (Figure 12 and

Figure 13). The steep drop in PFS at around 3 months (Figure 12) indicates that treatment with PAL+LET appears to offer protection against early progression in around 20% of patients versus treatment with LET.

Figure 13 shows that patients treated with PAL+LET have a lower hazard of progression in the long-term versus those treated with LET (the gradient of the exponential trend applied to the cumulative hazard is steeper for treatment with LET than for PAL+LET).

Figure 12 PFS K-M data and exponential trend (PALOMA-1 trial)

LET=letrozole; PAL+LET=palbociclib+letrozole; PFS=progression free survival
Source: Clarification question B4; ERG calculations

Figure 13 PFS cumulative hazard plot of K-M data and exponential trend (PALOMA-1 trial)

LET=letrozole; PAL+LET=palbociclib+letrozole; PFS=progression free survival
Source: Clarification question B4; ERG calculations

The well-established exponential trend in the PAL+LET treatment arm of the PALOMA-1 trial allows projection of PFS beyond the limits of the available K-M data for treatment with PAL+LET. The ERG extrapolated PFS for treatment with PAL+LET by appending the exponential trend established early in the K-M data to a data point close to the end of the K-M data. The data point chosen as the first point of extrapolation (16.6 months) was identified using the smallest of the weighted squared residuals calculated from the K-M data and fitted exponential curve. Extrapolation was not necessary for treatment with LET, as the final patient at risk died at [redacted] (Figure 12).

The ERG's projected PFS yielded estimates below those in the company model throughout the model time horizon for both treatments, except for a brief period in the first year for treatment with PAL+LET (

Figure 14).

Figure 14 ERG PFS projections using PALOMA-1 trial data vs company model PFS

ERG=Evidence Review Group; K-M=Kaplan-Meier; LET=letrozole; PAL+LET=palbociclib+letrozole; PFS=progression free survival
Source: Company model; Clarification response B4; ERG calculations

Mean PFS gain increased for PAL+LET in the ERG's model versus mean PFS gain in the company base case (13.3 months in the ERG's revised model versus 10.7 months in the base case). Applying the ERG's PFS projections based on the re-censored PALOMA-1 K-M data decreases the ICER per QALY gained by £29,461 to £121,408.

5.6.5 Time-to-event evidence: time to treatment discontinuation

The company has assumed that all patients in the model are treated to progression and has, accordingly, used PFS to estimate the proportion of patients receiving treatment in each cycle.

Figure 15, however, shows that some patients in the PALOMA-1 trial stopped treatment for reasons other than progression or death, which indicates that the time spent on treatment in this trial was less than the time spent in the progression-free state. It is unclear whether the TTD data for the PAL+LET arm of the PALOMA-1 trial represent PAL alone (that is, patients may have continued treatment with LET monotherapy) or whether it represents the discontinuation of all first-line treatments.

It is important to model time on treatment using trial TTD data where possible, as using PFS as a proxy can lead to an overestimation of the costs of treatment acquisition and administration (or an underestimation, if patients are permitted to continue treatment after progression).

Figure 15 shows how, at around 3 months, some patients treated with LET actually received treatment for a brief period after their progression was confirmed. Treatment beyond progression was not specified in the trial protocol.[97]

The company provided the ERG with TTD data from the PALOMA-1 trial during the clarification process. The difference between PFS and TTD was greater for patients treated with PAL+LET than for patients treated with LET (

Figure 15). The difference between PFS and TTD can be explained in the most part by the proportion of patients discontinuing treatment due to AEs: [redacted] of patients who discontinued treatment with PAL+LET in the PALOMA-1 trial did so due to AEs,[91] in comparison to [redacted] of patients who discontinued treatment with LET due to AEs.

Figure 15 PFS and TTD K-M data (PALOMA-1 trial data re-censored)
LET=letrozole; PAL+LET=palbociclib+letrozole; PFS=progression free survival; TTD=time to treatment discontinuation
Source: Clarification response B4

ERG exploratory analyses

To amend the model to calculate treatment costs using TTD rather than PFS, the ERG investigated methods of projecting the TTD K-M data provided by the company during the clarification process. The ERG found exponential trends established in the TTD data from the PALOMA-1 trial from around 9 months in the PAL+LET arm and around 5 months in the LET arm (

Figure 17).

The ERG used the trend in the PAL+LET arm to append exponential extrapolations to points near the end of the K-M data for treatment with PAL+LET. The extrapolation point was identified by choosing the K-M data point with the smallest weighted squared residual of the difference between the K-M data and the fitted exponential curve. The final K-M data point in the LET arm of the re-censored PALOMA-1 data set was censored, but, rather than extrapolating an estimate for this point, the ERG used the final PFS K-M point from the PALOMA-1 trial as a proxy in order that patients in the model did not receive treatment with

LET beyond progression when the ERG's PFS revisions were also applied.

Figure 16 Cumulative hazard plot of TTD K-M data and exponential trends (PALOMA-1 trial)

LET=letrozole; PAL+LET=palbociclib+letrozole; TTD=time to treatment discontinuation
Source: Clarification response B4

Figure 17 TTD K-M data and exponential trends PALOMA-1 trial data (re-censored)

LET=letrozole; PAL+LET=palbociclib+letrozole; TTD=time to treatment discontinuation

Source: Clarification response B4

Figure 18 shows the ERG's TTD projections using PALOMA-1 trial data alongside the ERG's PFS projections. The ERG's revisions using TTD K-M data from the PALOMA-1 trial reduce mean time on treatment by 10.2 months to 20.7 months for treatment with PAL+LET and by 7.3 months to 12.9 months for treatment with LET.

Figure 18 ERG TTD and PFS projections (PALOMA-1 trial data)

ERG=Evidence Review Group; K-M=Kaplan_Meier; LET=letrozole; PAL+LET=palbociclib+letrozole; PFS=progression free survival; TTD=time to treatment discontinuation
Source: Company model; ERG calculations

Applying the ERG's TTD projections based on the re-censored PALOMA-1 K-M data in the model alongside the company's base case PFS projections decreases the ICER per QALY gained by £47,941 to £102,928.

5.6.6 Health state utility values: pre-progression

The ERG does not consider the company to be justified in using a [redacted] for treatment with PAL+LET versus LET ([redacted]), as [redacted] found between the utility values calculated from the responses to the EQ-5D questionnaire in the two arms PALOMA-2 trial.[92]

The EQ-5D questionnaire was completed by patients on [redacted], and at the end of randomised treatment. [redacted] in each cycle in the ITT population completed the EQ-5D from baseline to cycle 21 in the PAL+LET arm ([redacted]) and from [redacted] in the PLACEBO+LET arm, after which [redacted] (

Figure 19). A [redacted] proportion of patients in the PLACEBO+LET arm [redacted] the EQ-5D questionnaire at each time point than did patients in the PAL+LET arm.

Figure 19 EQ-5D utility values and completion rates over time (PALOMA-2 trial)

Source: Company clarification responses B4 and B6; ERG calculations

Since [redacted] from patients in the PALOMA-2 trial was [redacted], the ERG considers that utility values should have been pooled and an overall average should have been used for both treatments. The company investigates in Scenario 22 the impact on the ICER per QALY gained of using an average of the two pre-progression utilities by applying a utility value of [redacted]. Using a pre-progression utility value of [redacted] for both treatments, the company's Scenario 22 increases the company's base case ICER per QALY gained by £14,991 to £165,860.

ERG exploratory analyses

The ERG has attempted to replicate the calculation of the pre-progression utility values used in the model using the data provided by the company during the clarification process, but was not able to identify the method used to yield the values [redacted]. The ERG has instead calculated alternative pre-progression utility values using the mean utility values from European patients in the PALOMA-2 trial. The ERG considers that using responses from European patients alone is likely to be a better approximation of responses of UK patients than using responses from the full ITT population, whilst still retaining a large enough data set to give a reliable average.

The ERG is also satisfied that it is valid to use utility values calculated from EQ-5D responses from the PALOMA-2 trial alongside time-to-event data from the PALOMA-1 trial in the absence

of EQ-5D data from the PALOMA-1 trial. This is because utility data are less prone to serious differences than time-to-event data provided the disease area and stage of disease are broadly similar.

The ERG calculated a weighted average utility value using the mean values per cycle and the the number of respondents per cycle from both arms of the PALOMA-2 trial for the first 21 cycles of treatment (since [redacted] of each arm in [

Figure 19], so can be considered reliable).

The average pooled cycle utility for European patients in the first 21 cycles in the PALOMA-2 trial was [redacted]. Applying the recalculated pre-progression utility values for PAL+LET and LET in the model increases the ICER per QALY gained by £16,858 to £167,727.

5.6.7 Health state utility values: post-progression

The company has made an error in the calculation of post-progression utility values using the published results of a study by Lloyd et al.[5] The company used the utility decrement associated with disease progression in the Lloyd[5] paper to derive a multiplier, which it then applied to the (average) pre-progression utility value from the PALOMA-2 trial. The company's resulting post-progression utility value used for both treatments in the base case is 0.4492.

This method assumes that the utility decrement associated with progressed disease can be applied linearly. However, a logistic transformation was applied to the data used in the Lloyd[5] study before analysis in order that it approximated the normal distribution necessary to allow use of a standard regression analysis. This means that the resulting utility gains and decrements reported in the paper cannot be directly applied or linearly adjusted and must be re-calculated to take into account the logistic transformation.

The ERG has recalculated the post-progression utilities using the results of the mixed model analysis given in the Lloyd[5] paper, including the logistic transformation of the data, and calibrated the result to the UK average age (48.52 years[117]) in the UK value set. The ERG's recalculated post-progression utility value is 0.5052. Applying this recalculated post-progression utility value in the model increase the ICER per QALY gained by £277 to £151,146.

5.6.8 Half-cycle correction

The company did not include a half-cycle correction to improve the accuracy of the cost and outcomes estimates. All patients progression-free and/or alive at the beginning of a cycle are assumed by the company to accrue costs and benefits throughout the entire cycle. However, some patients progress or die during a cycle and do not accrue the full costs and benefits for

that cycle. It is more accurate to assume costs and benefits apply to the average number of patients progression-free and/or alive in a cycle, which can be achieved by averaging the number of patients at the beginning and end of a cycle (mid-cycle correction).

Applying a mid-cycle correction to PFS and OS in the model reduces both incremental costs and incremental QALYs, and reduces the base case ICER per QALY gained by £2,182 to £148,687.

5.6.9 AE costs

The company is not justified in using a proportion of the relevant NHS Reference Cost[110] to represent a meeting of 20 minutes (Grade 3) or 30 minutes (Grade 4) with a consultant oncologist. This is because NHS Reference Costs[110] provide a currency for payment for the average patient[118] and do not represent an hourly cost (unless that is how much of the resource the average patient uses).

The ERG has amended the model to apply the full NHS Reference Cost[110] of £132 (Healthcare resource group currency code WF01A service code 800) to both Grade 3 and Grade 4 AEs. This increases the ICER per QALY gained by £1,603 to £152,472.

5.6.10 AE incidence calculation

The company has made two errors when calculating the incidence of AEs: first, the company used the median rather than mean time on treatment to calculate the probability of an AE; second, the company has applied annual rather than cycle AE probabilities to each cycle in the model. The ERG has amended these errors, which increases the time on treatment used in the probability calculations for both treatments, and substantially reduces the probability of AEs in each cycle.

Applying both of the ERGs corrections to the AE incidence calculation decreases the ICER per QALY gained by £854 to £150,015.

5.6.11 Discounting

In the company model, discounting of costs and outcomes is applied on a per cycle basis, rather than annually in line with NHS budgeting and accounting years. This has the effect of increasing the incremental QALYs more than the incremental costs.

The ERG has amended discounting to be applied on an annual basis. Application of this amendment decreases both incremental costs and incremental QALYs, and decreases the ICER per QALY gained by £159 to £150,710.

5.6.12 Days per year

The company has assumed 364 days per year in the model as a basis for several calculations, as there are 364 days in 13 28-day cycles. The ERG does not agree with using 364 days to approximate the number of days per year and has amended the value to 365.25 days.

Amending the number of days per year to 365.25 increases the ICER by £2 to £150,871.

5.6.13 ERG scenario analysis: using PALOMA-2 trial data

The ERG notes that the findings from a final analysis of cohort 1 from the PALOMA-1 trial shows large differences between investigator assessed PFS and BICR assessed PFS (Table 8). These findings were reported by the EMA. According to the EMA, these results indicate that findings from cohort 1 may be significantly biased to the extent that the findings from the PALOMA-1 trial are not suitable for licensure. The EMA also conclude that only findings from cohort 2 should be considered relevant to the efficacy assessment.

The ERG did not request K-M data from the PALOMA-1 trial to be split by cohort, so was unable to model PFS for cohort 2 from the PALOMA-1 trial. The ERG has instead provided a scenario analysis using re-censored, investigator assessed PFS data from the PALOMA-2 trial, along with TTD data from the same trial, as an alternative to using investigator assessed PFS data from the PALOMA-1 trial. This scenario analysis is subject to some of the same methodological flaws present in the company's base case, as it introduces inconsistencies into the relationship between PFS and OS.

The ERG has used re-censored K-M data from the PALOMA-2 trial directly for the first 19.2 months for treatment with PAL+LET and 18.1 months for treatment with LET, after which it appended exponential projections that had been calibrated using respective K-M data from 5 months onwards. The ERG's revised PFS projections based on data from the PALOMA-2 trial yield higher estimates of PFS for treatment with both PAL+LET and LET versus the company base case (

Figure 20). Mean PFS gain for treatment with PAL+LET in the ERG scenario analysis is 11.5 months versus 13.3 months in the ERG's revised model using PFS data from the PALOMA-1 trial and versus 10.7 months in the company base case.

Figure 20 ERG revised PFS model (scenario analysis: PALOMA-2 trial data) and company base case PFS projections

Source: Company model; Clarification response B4; ERG calculations

The ERG has also remodelled TTD using data from the PALOMA-2 trial to maintain consistency (

Figure 21). The ERG used the same approach to modelling TTD from the PALOMA-2 trial as it used to model PFS (K-M data plus exponential extrapolation). The ERG's remodelling of TTD from the PALOMA-2 trial reduces time on treatment versus PFS by 3.4 months for treatment with PAL+LET and 2.7 months for treatment with LET.

Figure 21 ERG revised PFS and TTD models (scenario analysis: PALOMA-2 trial data)

Source: Clarification response B4; ERG calculations

In the ERG scenario, applying the ERG's revised PFS using data from the PALOMA-2 trial increases the ICER per QALY gained versus the company base case by £6,155 to £156,984. Applying the ERG's revised TTD using data from the PALOMA-2 trial decreases the ICER per QALY gained versus the base case by £4,631 to £146,238.

5.6.14 ERG sensitivity analysis: subsequent treatments costs

The ERG does not agree with the company that it is reasonable to omit drug acquisition costs for subsequent treatments post-progression and considers that the company should have carried out a more thorough costing of post-progression treatments in this appraisal. The ERG's revised PFS and OS estimates increase time spent in PPS, and thus the cost of PPS, substantially more for patients treated with LET than they do for patients treated with PAL+LET (Table 32), which indicates that the model is sensitive to the cost of subsequent treatments when PPS is not assumed to be equal for the intervention and comparator. The ERG was not able to perform a full costing of post-progression treatments, so carried out a simple sensitivity analysis to investigate the magnitude of the impact of adding drug costs to the subsequent therapy calculations.

Mean PPS gain for treatment with PAL+LET in the company's base case model is 0.49 months, which decreases to a 6.7 month mean PPS loss when the ERG's revised PFS and OS estimates are applied. The incremental cost of subsequent treatment more than doubles versus the base case when the ERG's PFS and OS estimates are applied; subsequent treatment costs are £528 lower for treatment with PAL+LET than for LET in the base case, but are £1,487 lower for treatment with PAL+LET when the ERG's PFS and OS estimates are applied.

Table 32 Cost of subsequent treatment for PAL+LET and LET (excluding BSC)

	Total subsequent treatment costs (excluding BSC)		
	PAL+LET	LET	Difference
Company base case	[redacted]	[redacted]	[redacted]
Using ERG preferred PFS and OS estimates	[redacted]	[redacted]	[redacted]

BSC=best supportive care; ERG=Evidence Review Group; LET=letrozole; PAL+LET=palbocicliob+letrozole; OS=overall survival; PFS=progression free survival
Source: Company model, ERG calculations

The ERG used its revised PFS and OS estimates based on PALOMA-1 trial data in order to introduce a reduced time spent in PPS for patients treated with PAL+LET versus those treated with LET (-6.7 months). The difference in subsequent treatment costs for patients treated with PAL+LET versus LET ranged from -£1,841 if drugs cost £100 per cycle to -£36,840 if drugs cost £10,000 per cycle (

Table 33). The impact on the ICER per QALY gained ranged from -£3,606 for drugs costing £100 per cycle to -£70,047 for drugs costing £10,000 per cycle.

The ICER per QALY gained decreases with an increase in subsequent treatment costs because the analysis uses ERG estimates of PFS and OS in order that the model includes a mean PPS loss for treatment with PAL+LET. This reduces the time spent both on first-line and subsequent treatment for patients receiving PAL+LET in particular, which substantially reduces the total cost of treatment for these patients. However, the key conclusion of the sensitivity analysis is that the ICER per QALY gained changes substantially depending on the cost of subsequent treatment. The ERG thus considers that the company should have included a more through costing of post-progression treatments in its model.

Table 33 Subsequent treatment cost sensitivity analysis

Drug acquisition and administration cost per cycle	Total subsequent treatment costs (excluding BSC)			ICER per QALY gained	ICER difference from base case
	PAL+LET	LET	Difference		
£100	[redacted]	[redacted]	[redacted]	£147,262	-£3,606
£1,000	[redacted]	[redacted]	[redacted]	£141,222	-£9,646
£10,000	[redacted]	[redacted]	[redacted]	£80,822	-£70,047

BSC=best supportive care; ERG=Evidence Review Group; ICER=incremental cost effectiveness ratio; LET=letrozole; PAL+LET=palbocicliob+letrozole; QALY=quality adjusted life year
Source: Company model, ERG calculations

5.6.15 Company probability sensitivity analysis

Figure 4 shows the CEAC for the company's base case. The scatterplot is essentially one-dimensional along the QALY axis, with very little variability in the cost axis. Tthis result is due to the way in which the company has formulated the PSA. The PSA macro is set up to exclude any correlated uncertainty in the key model parameters (Weibull model scale and shape parameters). This leads to apparently minimal uncertainty in the estimate of the probabilistic ICER and therefore virtually no spread in the CEAC. The ERG therefore places no confidence in the PSA results which are inconsistent with the use of multiple Weibull models in projecting future costs and outcomes.

5.7 Conclusions of the cost effectiveness section

The various changes implemented by the ERG for the comparison of treatment with PAL+LET versus treatment with LET yield a mixture of effects. When implemented individually, these revisions both increase and decrease the size of the ICERs per QALY gained. The combined effect of all of the ERG's revisions, when using PALOMA-1 data as the basis for modelling PFS and TTD, decreases the ICER per QALY gained by £17,997 to £132,872. However, the

combined effect of all of the ERG's revisions, when using PALOMA-2 data as the basis for modelling PFS and TTD, increases the ICER per QALY gained by £62,337 to £213,206.

The ERG considers that there is considerable uncertainty as to whether the company's base case results overestimate or underestimate the size of the most probable ICER per QALY gained. The available data from the PALOMA-1 trial is flawed, but allows for the most methodologically robust approach to modelling survival; the available data from the PALOMA-2 trial is more robust, but requires the application of methodologically unsound approaches to modelling survival to compensate for the absence of OS data from that trial.

The cost effectiveness results that are generated in the company's base case and following the application of either of the ERG's combined revision scenarios are all considerably higher than the range normally considered acceptable by NICE.

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

The ERG has made ten changes to the submitted model to address the points raised in Section 5.6. The combined impact on the ICER per QALY gained as a result of the following changes are given in Scenario B in Table 34:

R1) OS estimates based on data from the PALOMA-1 trial

R2) PFS estimates based on data from the PALOMA-1 trial

R3) TTD estimates based on data from the PALOMA-1 trial

R4) re-calculate pre-progression utility values from the PALOMA-2 trial data

R5) re-calculate post-progression utility values using information in the Lloyd study[5]

R6) use mid-cycle correction

R7) re-calculate cost of treating AEs using full NHS Reference Costs

R8) correct AE incidence calculation

R9) change discounting to annual

R10) use 365.25 days per year instead of 364

The ERG has made a further two changes to the submitted model to provide alternatives to using PALOMA-1 trial data to model PFS and TTD. The combined impact on the ICER per QALY gained as a result of the substituting the following changes for R2) and R3) in Scenario B are given in Scenario C in Table 34:

R11) PFS estimates based on data from the PALOMA-2 trial

R12) TTD estimates based on data from the PALOMA-2 trial

Details of all Microsoft Excel revisions made by the ERG to the company's model are presented in the appendices to this report (Section 10.10).

6.1 Summary of ERG revisions to company model

The cost effectiveness results obtained by applying each of the ERG's model revisions are shown in Table 34.

The ERG's revised base case scenario encompassing all of the ERG's revisions to the company's model, using the ERG's revised PFS and TTD estimates based on data from the PALOMA-1 trial (Scenario B in Table 34) yields an ICER per QALY gained of £132,872, which is £17,997 lower than in the company's base case. The ERG's revised base case for the comparison of treatment with PAL+LET versus treatment with LET using PALOMA-1 trial data

generates both incremental costs (£59,934) and benefits (0.451 QALYs) that are lower than those generated by the company. The ERG's revised base case using PALOMA-1 trial data to model PFS and TTD also reduces incremental life years gained (0.454 years) compared to the company's base case.

The reduction in the ICER per QALY gained in Scenario B, when all the ERG's revisions are applied simultaneously and using PALOMA-1 trial data, is principally a result of the reduction in treatment costs due to using TTD rather than PFS to estimate the proportion of patients receiving treatment in each cycle. The reduction in treatment costs is proportionately much larger for patients receiving PAL+LET than patients receiving LET, which decreases the ICER per QALY gained. The substantial decrease in the ICER per QALY gained due to lower treatment costs is mitigated, however, by decreases in the incremental QALYs accrued for treatment with PAL+LET due to less time spent in PPS for these patients and to equal pre-progression utility values assumed to apply to both the intervention and comparator.

The ERG's revised base case scenario encompassing all of the ERG's revisions to the company's model, using the ERG's revised PFS and TTD estimates based on data from the PALOMA-2 trial (Scenario C in Table 34), yields an ICER per QALY gained of £213,206, which is £62,337 higher than in the company's base case. The ERG's revised base case for the comparison of treatment with PAL+LET versus treatment with LET using PALOMA-2 trial data generates both incremental costs (£88,452) and benefits (0.415 QALYs) that are lower than those generated by the company. The ERG's revised base case using PALOMA-2 trial data to model PFS and TTD reduces incremental life years gained (0.454 years) compared to the company's base case.

Table 34 Cost effectiveness results: ERG revisions to company base case

Model scenario ERG revision	PAL+LET			LET			Incremental			ICER	ICER
	Cost £	QALYs	Life years	Cost £	QALYs	Life years	Cost £	QALYs	Life years	£/QALY[+]	Change
A. Company original base case	***£116,696***	***2.402***	***3.793***	***£21,843***	***1.773***	***3.016***	***£94,853***	***0.629***	***0.777***	***£150,869***	
R1) ERG OS estimates based on data from PALOMA-1	£114,359	2.314	3.598	£23,381	1.834	3.152	£90,977	0.481	0.447	£189,310	+£38,441
R2) ERG PFS estimates based on data from PALOMA-1	£107,386	2.314	3.793	£25,458	1.639	3.016	£81,928	0.675	0.777	£121,408	-£29,461
R3) ERG TTD estimates based on data from PALOMA-1	£86,544	2.402	3.793	£21,831	1.773	3.016	£64,712	0.629	0.777	£102,928	-£47,941
R4) ERG recalculated pre-progression utility values from PALOMA-2 trial	£116,696	2.353	3.793	£21,843	1.787	3.016	£94,853	0.566	0.777	£167,727	+£16,858
R5) ERG recalculated post-progression utility values using Lloyd 2006[5]	£116,696	2.480	3.793	£21,843	1.852	3.016	£94,853	0.628	0.777	£151,146	+£277
R6) Use mid-cycle correction	£115,308	2.376	3.759	£21,875	1.748	2.982	£93,433	0.628	0.778	£148,687	-£2,182
R7) Use full reference costs for AEs	£118,088	2.402	3.793	£22,227	1.773	3.016	£95,861	0.629	0.777	£152,472	+£1,603
R8) Correct AE incidence calculation	£115,962	2.402	3.793	£21,646	1.773	3.016	£94,317	0.629	0.777	£150,015	-£854
R9) Change discounting to annual	£118,449	2.438	3.851	£22,187	1.800	3.062	£96,262	0.639	0.789	£150,710	-£159
R10) Use 365.25 days per year	£116,698	2.402	3.793	£21,844	1.773	3.016	£94,854	0.629	0.777	£150,871	+£2
B. ERG revised base case using PALOMA-1 PFS, OS and TTD (R1:R9)	***£87,478***	***2.280***	***3.619***	***£27,544***	***1.829***	***3.164***	***£59,934***	***0.451***	***0.454***	***£132,872***	***-£17,997***
R11) ERG PFS estimates based on data from PALOMA-2	£121,946	2.452	3.793	£20,708	1.808	3.016	£101,238	0.645	0.777	£156,984	+£6,115
R12) ERG TTD estimates based on data from PALOMA-2	£113,783	2.402	3.793	£21,842	1.773	3.016	£91,942	0.629	0.777	£146,238	-£4,631
C. ERG revised base case using PALOMA-2 PFS and TTD (R1 & R4:R9)	***£110,970***	***2.386***	***3.619***	***£22,518***	***1.971***	***3.164***	***£88,452***	***0.415***	***0.454***	***£213,206***	***+£62,337***

Costs, QALYs and life years discounted

N.B. incremental undiscounted life years are 0.931 in the company base case and 0.549 in the ERG's revised base case estimates

ERG=Evidence Review Group; ICER=incremental cost effectiveness ratio; LET=letrozole; OS=overall survival; PAL+LET=palbociclib+letrozole; PFS=progression-free survival; QALYs=quality adjusted life years; TTD=time to treatment discontinuation

[+] Rounding errors account for difference between ICERs calculated using the incremental cost and QALY values given in the table and ICERs in this column

7 END OF LIFE

The company has not made a case for PAL+LET to be considered under NICE's End of Life criteria.

8 OVERALL CONCLUSIONS

8.1 NHS clinical practice

Despite a higher proportion of patients in the PALOMA-1 and PALOMA-2 trials presenting with de novo disease than would be seen in NHS clinical practice, the ERG is generally satisfied that the evidence derived from both trials is generalisable to the patient population in England and Wales described in the final scope issued by NICE. The EMA considers that it is reasonable to generalise the clinical effectiveness results associated with LET to other aromatase inhibitors; the ERG concurs with this viewpoint.

8.2 Clinical effectiveness

Efficacy evidence is derived from two trials. The phase III PALOMA-2 trial was considered by the ERG to be of superior quality and lower risk of bias than the phase I/II PALOMA-1 trial as the former trial was larger and designed as a double-blind trial, whereas the PALOMA-1 trial was designed as an open-label trial. Furthermore, investigator assessed PFS findings reported for cohort 1 of the PALOMA-1 trial differed markedly to BICR assessed PFS. This has led the EMA to conclude that only findings from cohort 2 should be considered relevant to the efficacy assessment. OS data from the PALOMA-1 trial are also immature and are from a data cut-off date of 29 November 2013. There are no OS data currently available from the PALOMA-2 trial. Despite a large gain in investigator assessed median PFS (of approximately 10 months) for patients treated with PAL+LET versus LET or PLACEBO+LET in both the PALOMA-1 and PALOMA-2 trials, no statistically significant improvement in median OS for patients in the PALOMA-1 trial or the PALOMA-2 trial was observed.

Differences between the treatment arms in terms of safety were mostly attributable to a much higher rate of haematological toxicities, particularly neutropenia in patients treated with PAL+LET. While this included high rates of Grade 3 to 4 neutropenia, for the most part, neutropenia was asymptomatic and reversible, with febrile neutropenia being reported by <2% of patients (all incidence occurring in the PALOMA-2 trial). These data suggest neutropenia rarely results in permanent discontinuation of treatment with PAL+LET. Therefore, the safety profile of PAL+LET is considered by the company and the ERG to be acceptable.

There were no statistically significant differences between trial arms in terms of HRQoL measures reported in either of the PALOMA-1 trial or the PALOMA-2 trial. Thus, while the trials did not demonstrate that prolonging PFS improved HRQoL over time, the trials did suggest an increase in incidence of AEs for patients treated with PAL+LET compared with LET or PLACEBO+LET; however, this increase in incindence of AEs did not appear to affect HRQoL.

8.3 Cost effectiveness

There is considerable uncertainty as to whether the company's base case cost effectiveness results overestimate or underestimate the size of the most probable ICER per QALY gained. When implemented individually, the ERG's revisions both decrease and increase the estimated ICER per QALY gained versus the company base case. However, the company's base case cost effectiveness results, as well as those generated following the application of all the ERG's revisions, are considerably higher than the range normally considered acceptable by NICE.

The available data from the PALOMA-1 trial are flawed, but allow for the most methodologically robust approach to modelling survival and yields an ICER estimate of £132,872 per QALY gained (£17,997 lower than in the company's base case); the available data from the PALOMA-2 trial are more robust, but require the application of methodologically unsound approaches to modelling survival to compensate for the absence of OS data from that trial, and yield an ICER estimate of £213,306 per QALY gained (£62,337 higher than in the company's base case).

8.4 Implications for research

While LET, anastrozole and exemestane, the aromatase inhibitors currently used in NHS clinical practice can be considered to be of equal efficacy, studies comparing palbociclib in combination with, and versus, other aromatase inhibitors would add to the evidence base. The ERG notes that the EMA highlight that ongoing clinical studies examining palbociclib in combination with anastrozole and exemestane are underway.

More evidence for the impact of palbociclib in combination with an aromatase inhibitor on OS is required. While OS data from the PALOMA-2 trial will add to the evidence base when the data become available, more mature data from the PALOMA-1 trial would also be informative.

9 REFERENCES

1. Pfizer. National Institute for Health and Care Excellence. Single Technolgoy Appraisal. Palbociclib for Treating Metastatic, Hormone Receptor-positive, HER2-negative Breast Cancer [ID915]. Company Evidence Submission. 2016.
2. Irvin W, Muss HB, Mayer DK. Symptom Management in Metastatic Breast Cancer. Oncologist. 2011; 16:1203-14.
3. Wyatt G, Sikorskii A, Tamkus D, You M. Quality of life among advanced breast cancer patients with and without distant metastasis. Eur J Cancer Care. 2013; 22:272-80.
4. Reed E, Simmonds P, Haviland J, Corner J. Quality of life and experience of care in women with metastatic breast cancer: a cross-sectional survey. J Pain Symptom Manage. 2012; 43:747-58.
5. Lloyd A, Nafees B, Narewska J, Dewilde S, Watkins J. Health state utilities for metastatic breast cancer. Br J Cancer. 2006; 95:683-90.
6. Al-Batran SE, Hozaeel W, Tauchert FK, Hofheinz RD, Hinke A, Windemuth-Kieselbach C*, et al.* The impact of docetaxel-related toxicities on health-related quality of life in patients with metastatic cancer (QoliTax). Ann Oncol. 2015; 26:1244-8.
7. Jerusalem G, Gupta S, Zhang J. The association of chemotherapy versus hormonal therapy and health outcomes by geographic region among patients with hormone receptor-positive, HER2-negative metastatic breast cance. 9th European Breast Cancer Conference; Glasgow, Scotland, 19-21 March, 2014.
8. Vardy J, Tannock I. Cognitive function after chemotherapy in adults with solid tumours. Crit Rev Oncol Hematol. 2007; 63:183-202.
9. Milanti A, Metsala E, Hannula L. Reducing psychological distress in patients undergoing chemotherapy. Br J Nurs. 2016; 25:S25-30.
10. Dodd MJ, Cho MH, Cooper BA, Miaskowski C. The effect of symptom clusters on functional status and quality of life in women with breast cancer. Eur J Oncol Nurs. 2010; 14:101-10.
11. Chen YC, Huang HM, Kao CC, Sun CK, Chiang CY, Sun FK. The Psychological Process of Breast Cancer Patients Receiving Initial Chemotherapy: Rising From the Ashes. Cancer Nurs. 2016; 39:E36-e44.
12. Singer S, Blettner M, Kreienberg R, Janni W, Wockel A, Kuhn T*, et al.* Breast Cancer Patients' Fear of Treatment: Results from the Multicenter Longitudinal Study BRENDA II. Breast Care. 2015; 10:95-100.
13. Lim CC, Devi MK, Ang E. Anxiety in women with breast cancer undergoing treatment: a systematic review. Int J Evid Based Healthc. 2011; 9:215-35.
14. Bidstrup PE, Christensen J, Mertz BG, Rottmann N, Dalton SO, Johansen C. Trajectories of distress, anxiety, and depression among women with breast cancer: Looking beyond the mean. Acta Oncol. 2015; 54:789-96.
15. Finn RS, Martin M, Rugo HS, Jones SE, Im S-A, Gelmon KA*, et al.* PALOMA-2: Primary results from a phase III trial of palbociclib (P) with letrozole (L) compared with letrozole alone in postmenopausal women with ER+/HER2- advanced breast cancer (ABC). J Clin Oncol. 2016; 34:507.
16. Andre F, Marinsek N, Ricci J-F, Etchberger J, Degun R, Benelli G*, et al.* Patterns of Clinical Management and Resource Utilisation for Postmenopausal Hormone-

Receptor-Positive HER2-Negative (HR+ HER2-) Advanced Breast Cancer (BC) in Europe. Value Health. 2015; 15:A419.

17. Beaver K, Williamson S, Briggs J. Exploring patient experiences of neo-adjuvant chemotherapy for breast cancer. Eur J Oncol Nurs. 20:77-86.

18. Breast Cancer Working Group (GOV.UK). Improving Outcomes: A Strategy for Cancer. December 2010. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/213772/dh_123414.pdf (Accessed 27 October 2016).

19. Heidari Gorji MA, Bouzar Z, Haghshenas M, Kasaeeyan AA, Sadeghi MR, Ardebil MD. Quality of life and depression in caregivers of patients with breast cancer. BMC Res Notes. 2012; 5:310.

20. Beauchemin C, Cooper D, Lapierre ME, Yelle L, Lachaine J. Progression-free survival as a potential surrogate for overall survival in metastatic breast cancer. Onco Targets Ther. 2014; 7:1101-10.

21. Burzykowski T, Buyse M, Piccart-Gebhart MJ, Sledge G, Carmichael J, Luck HJ*, et al.* Evaluation of tumor response, disease control, progression-free survival, and time to progression as potential surrogate end points in metastatic breast cancer. J Clin Oncol. 2008; 26:1987-92.

22. Hackshaw A, Knight A, Barrett-Lee P, Leonard R. Surrogate markers and survival in women receiving first-line combination anthracycline chemotherapy for advanced breast cancer. Br J Cancer. 2005; 93:1215-21.

23. Miksad RA, Zietemann V, Gothe R, Schwarzer R, Conrads-Frank A, Schnell-Inderst P*, et al.* Progression-free survival as a surrogate endpoint in advanced breast cancer. Int J Technol Assess Health Care. 2008; 24:371-83.

24. Petrelli F, Barni S. Surrogate endpoints in metastatic breast cancer treated with targeted therapies: an analysis of the first-line phase III trials. Med Oncol. 2014; 31:776.

25. Sherrill B, Amonkar M, Wu Y, Hirst C, Stein S, Walker M*, et al.* Relationship between effects on time-to-disease progression and overall survival in studies of metastatic breast cancer. Br J Cancer. 2008; 99:1572-8.

26. Davis S, Tappenden P, Cantrell A. A review of studies examining the relationship between progression-free survival and overall survival in advanced or metastatic cancer. Report by the Decision Support Unit. Sheffield: School of Health and Related Research, University of Sheffield 2012.

27. Ciani O, Davis S, Tappenden P, Garside R, Stein K, Cantrell A*, et al.* Validation of surrogate endpoints in advanced solid tumors: systematic review of statistical methods, results, and implications for policy makers. Int J Technol Assess Health Care. 2014; 30:312-24.

28. eMC. Femera: Summary of product characteristics. Available from: https://www.medicines.org.uk/emc/medicine/1285 (Accessed 14 November 2016).

29. eMC. Aromasin: Summary of Product Chracteristics. Available from: https://www.medicines.org.uk/emc/medicine/2484 (Accessed 14 November 2016).

30. eMC. Arimidex: Summary of product characteristics. Available from: https://www.medicines.org.uk/emc/medicine/3845 (Accessed 14 November 2016).

31. National Institute for Health and Care Excellence. Advanced breast cancer: diagnosis and treatment. Clinical guideline [CG81]. Published date: February 2009. Last updated: July 2014. Available from: www.nice.org.uk/guidance/cg81 (Accessed 11 November 2016).

32. Mitra D, Kurosky S, Zanotti G, Kaye J. Real-world treatment patterns and clinical outcomes in ER+/HER2- metastatic breast cancer: results from a multicountry retrospective medical record review. ISPOR 21st Annual International Meeting (poster). 2016.

33. IPSOS Healthcare. Ipsos EU5 Oncology Monitor – Breast Cancer Treatment Overview - data on file. 2015.

34. Gibson L, Lawrence D, Dawson C, Bliss J. Aromatase inhibitors for treatment of advanced breast cancer in postmenopausal women. Cochrane Database Syst Rev. 2009; (4):CD003370. doi: 10.1002/14651858.CD003370.pub3.

35. Kleijnen J, Riemsma R, Amonkar MM, Lykopoulos K, Diaz JR, Forbes CA*, et al.* An indirect comparison of aromatase inhibitors (AIs) in the first line treatment of post menopausal women with hormone receptor positive (HR+) metastatic breast cancer (MBC). EJC Suppl. 2010; 8 (3):200.

36. Riemsma R, Forbes CA, Kessels A, Lykopoulos K, Amonkar MM, Rea DW*, et al.* Systematic review of aromatase inhibitors in the first-line treatment for hormone sensitive advanced or metastatic breast cancer. Breast Cancer Res Treat. 2010; 123:9-24.

37. Shapiro CL. Aromatase Inhibitors and Bone Loss: Risks in Perspective. J Clin Oncol. 2005; 23:4847-9.

38. Iwata H, Masuda N, Ohno S, Rai Y, Sato Y, Ohsumi S*, et al.* A randomized, double-blind, controlled study of exemestane versus anastrozole for the first-line treatment of postmenopausal Japanese women with hormone-receptor-positive advanced breast cancer. Breast Cancer Res Treat. 2013; 139:441-51.

39. Llombart-Cussac A, Ruiz A, Anton A, Barnadas A, Antolin S, Ales-Martinez JE*, et al.* Exemestane versus anastrozole as front-line endocrine therapy in postmenopausal patients with hormone receptor-positive, advanced breast cancer: Final results from the Spanish Breast Cancer Group 2001-03 phase 2 randomized trial. Cancer. 2012; 118:241-7.

40. Paridaens RJ, Dirix LY, Beex LV, Nooij M, Cameron DA, Cufer T*, et al.* Phase III study comparing exemestane with tamoxifen as first-line hormonal treatment of metastatic breast cancer in postmenopausal women: The European Organisation for Research and Treatment of Cancer Breast Cancer Cooperative Group. J Clin Oncol. 2008; 26:4883-90.

41. Chernozemsky I, Kalinov K, Tzekov H, Racheva M, Hristova S, Tomova A*, et al.* Randomized phase III trial of exemestane or tamoxifen in first-line hormonal treatment of postmenopausal women with metastatic breast cancer. 30th Annual San Antonio Breast Cancer Symposium; 2007 13-16 December; San Antonio, TX, USA.

42. Mouridsen H, Gershanovich M, Sun Y, Perez-Carrion R, Boni C, Monnier A*, et al.* Phase III study of letrozole versus tamoxifen as first-line therapy of advanced breast cancer in postmenopausal women: analysis of survival and update of efficacy from the International Letrozole Breast Cancer Group. J Clin Oncol. 2003; 21:2101-9.

43. Bonneterre J, Thürlimann B, Robertson JF, Krzakowski M, Mauriac L, Koralewski P*, et al.* Anastrozole versus tamoxifen as first-line therapy for advanced breast cancer in 668 postmenopausal women: results of the Tamoxifen or Arimidex Randomized Group Efficacy and Tolerability study. J Clin Oncol. 2000; 18:3748-57.

44. Nabholtz JM, Buzdar A, Pollak M, Harwin W, Burton G, Mangalik A*, et al.* Anastrozole is superior to tamoxifen as first-line therapy for advanced breast cancer in postmenopausal women: results of a North American multicenter randomized trial. J Clin Oncol. 2000; 18:3758-67.

45. Cardoso F, Bischoff J, Brain E, Zotano ÁG, Lück HJ, Tjan-Heijnen VC*, et al.* A review of the treatment of endocrine responsive metastatic breast cancer in postmenopausal women. Cancer Treat Rev. 2013; 39:457-65.

46. Dickler MN, Barry WT, Cirrincione CT, Ellis MJ, Moynahan ME, Innocenti F*, et al.* Phase III trial evaluating letrozole as first-line endocrine therapy with or without bevacizumab for the treatment of postmenopausal women with hormone receptor-positive advanced-stage breast cancer: CALGB 40503 (Alliance). 2016; 34:2602-9.

47. Cameron D. Clinical Use of Aromatase Inhibitors in Breast Cancer: History and Present. In: Larionov A, editor. Resistance to Aromatase Inhibitors in Breast Cancer. Cham: Springer International Publishing; 2015. p. 1-11.

48. Mauriac L, Keshaviah A, Debled M, Mouridsen H, Forbes JF, Thurlimann B*, et al.* Predictors of early relapse in postmenopausal women with hormone receptor-positive breast cancer in the BIG 1-98 trial. Ann Oncol. 2007; 18:859-67.

49. Finn RS, Crown JP, Lang I, Boer K, Bondarenko IM, Kulyk SO*, et al.* The cyclin-dependent kinase 4/6 inhibitor palbociclib in combination with letrozole versus letrozole alone as first-line treatment of oestrogen receptor-positive, HER2-negative, advanced breast cancer (PALOMA-1/TRIO-18): a randomised phase 2 study. Lancet Oncol. 2015; 16:25-35.

50. Baselga J, Campone M, Piccart M, Burris HAI, Rugo HS, Sahmoud T*, et al.* Everolimus in Postmenopausal Hormone-Receptor–Positive Advanced Breast Cancer. N Engl J Med. 2012; 366:520-9.

51. Beck JT, Hortobagyi GN, Campone M, Lebrun F, Deleu I, Rugo HS*, et al.* Everolimus plus exemestane as first-line therapy in HR+/HER2- advanced breast cancer in BOLERO-2. Breast Cancer Res Treat. 2014; 143:459-67.

52. Rugo HS, Rumble RB, Burstein HJ. Endocrine Therapy for Hormone Receptor Positive Metastatic Breast Cancer: American Society of Clinical Oncology Guideline Summary. J Oncol Pract. 2016; 12:583-7.

53. Finn RS, Dering J, Conklin D, Kalous O, Cohen DJ, Desai AJ*, et al.* PD 0332991, a selective cyclin D kinase 4/6 inhibitor, preferentially inhibits proliferation of luminal estrogen receptor-positive human breast cancer cell lines in vitro. Breat Cancer Res. 2009; 11:R77-R.

54. Cristofanilli M, Turner NC, Bondarenko I, Ro J, Im SA, Masuda N*, et al.* Fulvestrant plus palbociclib versus fulvestrant plus placebo for treatment of hormone-receptor-positive, HER2-negative metastatic breast cancer that progressed on previous endocrine therapy (PALOMA-3): final analysis of the multicentre, double-blind, phase 3 randomised controlled trial. Lancet Oncol. 2016; 17:425-39.

55. Hortobagyi GN, Stemmer SM, Burris HA, Yap Y-S, Sonke GS, Paluch-Shimon S*, et al.* Ribociclib as First-Line Therapy for HR-Positive, Advanced Breast Cancer. New Eng J Med. Published online 8 October 2016. Available from: http://www.nejm.org/doi/full/10.1056/NEJMoa1609709 (Accessed 25 October 2016).

56. Office for National Statistics. Cancer Registration Statistics, England. Available from: http://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/datasets/cancerregistrationstatisticscancerregistrationstatisticsengland (Accessed 27 October 2016).

57. Welsh Cancer Intelligence and Surveillance Unit. Available from: http://www.wcisu.wales.nhs.uk/officical-statistics-exel-files-of-trend (Accessed 27 October 2016).

58. National Institute for Health and Care Excellence. Advanced breast cancer: costing template. Available from: https://www.nice.org.uk/guidance/cg81/resources (Accessed 27 October 2016).

59. DeKoven M, Bonthapally V, Jiao X, Ganguli A, Pathak P, Lee WC, *et al.* Treatment pattern by hormone receptors and HER2 status in patients with metastatic breast cancer in the UK, Germany, France, Spain and Italy (EU-5): results from a physician survey. J Comp Eff Res. 2012; 1:453-63.

60. GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11 [database on the Internet]. International Agency for Research on Cancer. 2013. Available from: http://globocan.iarc.fr (Accessed 25 October 2016).

61. Cancer Research UK. Breast cancer incidence trends over time (females). Available from: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer/incidence-invasive#heading-Two (Accessed 8 November 2016).

62. National Institute for Health and Care Excellence (NICE). Final scope for the appraisal of Palbociclib in combination with an aromatase inhibitor for previously untreated metastatic, hormone receptor-positive, HER2-negative breast cancer . Issue Date: August 2016. London: NICE; Available from: https://www.nice.org.uk/guidance/indevelopment/gid-ta10068/documents (Accessed 15 November 2016).

63. Lyratzopoulos G, Abel GA, Barbiere JM, Brown CH, Rous BA, Greenberg DC. Variation in advanced stage at diagnosis of lung and female breast cancer in an English region 2006-2009. Br J Cancer. 2012; 106:1068-75.

64. Huober J, Fasching PA, Barsoum M, Petruzelka L, Wallwiener D, Thomssen C, *et al.* Higher efficacy of letrozole in combination with trastuzumab compared to letrozole monotherapy as first-line treatment in patients with HER2-positive, hormone-receptor-positive metastatic breast cancer - results of the eLEcTRA trial. Breast. 2012; 21:27-33.

65. Paul D, Vukelja SJ, Holmes FA, Blum J, McIntyre KJ, Kumar AR, *et al.* Letrozole plus dasatinib improves progression-free survival (PFS) in hormone receptor-positive, HER2-negative postmenopausal metastatic breast cancer (MBC) patients receiving first-line aromatase inhibitor (AI) therapy. Cancer Res. 2013; 73(24 suppl. 1): Available from: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/044/CN-01064044/frame.html.

66. Wolff AC, Lazar AA, Bondarenko I, Garin AM, Brincat S, Chow L, *et al.* Randomized phase III placebo-controlled trial of letrozole plus oral temsirolimus as first-line endocrine therapy in postmenopausal women with locally advanced or metastatic breast cancer. J Clin Oncol. 2013; 31:196-202.

67. Goss P, Bondarenko IN, Manikhas GN, Pendergrass KB, Miller WH, Jr., Langecker P, *et al.* Phase III, double-blind, controlled trial of atamestane plus toremifene compared with letrozole in postmenopausal women with advanced receptor-positive breast cancer. J Clin Oncol. 2007; 25:4961-6.

68. European medicines Agency. EMA/652627/2016. Committee for Medicinal Products for Human Use (CHMP). Assessment report. IBRANCE. International non-proprietary name: palbociclib. Procedure No. EMEA/H/C/003853/0000. 15 September 2016. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Public_assessment_report/human/003853/WC500217198.pdf (Accessed 30 November 2016).

69. Hu W, Sung T, Jessen BA, Thibault S, Finkelstein MB, Khan NK, *et al.* Mechanistic Investigation of Bone Marrow Suppression Associated with Palbociclib and its Differentiation from Cytotoxic Chemotherapies. Clin Cancer Res. 2016; 22:2000-8.

70. European Medicines Agency. Guideline on evaluation of anticancer medicinal products in man (draft, MA/CHMP/205/95 Rev.5) 2015. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2016/03/WC500203320.pdf (Accessed 28 October 2016).

71. Williamson PR, Gamble C, Altman DG, Hutton JL. Outcome selection bias in meta-analysis. Stat Methods Med Res. 2005; 14:515-24.

72. National Institute for Health and Care Excellence (NICE). Single technology appraisal: User guide for company evidence submission template. NICE Process [PMG24]. Published date: January 2015 Available from: https://www.nice.org.uk/process/pmg24/chapter/1-instructions-for-companies (Accessed 11 August 2016).

73. Centre for Reviews and Dissemination (CRD). CRD's guidance for undertaking reviews in healthcare: Systematic Reviews (3rd Edition). York: CRD, University of York 2008.

74. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. J Epidemiol Community Health. 1998; 52:377-84.

75. National Institute for Health and Care Excellence (NICE). Developing NICE guidelines: the manual. Process and methods [PMG20]. October 2014. Available from: https://www.nice.org.uk/media/default/about/what-we-do/our-programmes/developing-nice-guidelines-the-manual.pdf (Accessed 31 October 2016).

76. National Institute for Health and Care Excellence (NICE). Developing NICE guidelines: the manual appendices A–I. July 2015. Available from: ttps://www.nice.org.uk/media/default/About/what-we-do/NICE-guidance/NICE-guidelines/developing-NICE-guidelines-the-manual.pdf (Accessed 31 October 2016).

77. Clark A, O'Dwyer P, Heitjan D, Lal P, Feldman M, Gallagher M, *et al.* A phase I trial of palbociclib and paclitaxel in metastatic breast cancer. J Clin Oncol. 2014; 32:5s (suppl; abstr 527).

78. Clark A, O'Dwyer P, Troxel A, Lal P, Feldman M, Gallagher M, *et al.* Palbociclib and paclitaxel on an alternating schedule for advanced breast cancer: Results of a phase Ib trial. San Antonio Breast Cancer Symposium, 12 December 2015.

79. Flaherty KT, Lorusso PM, Demichele A, Abramson VG, Courtney R, Randolph SS, *et al.* Phase I, dose-escalation trial of the oral cyclin-dependent kinase 4/6 inhibitor PD 0332991, administered using a 21-day schedule in patients with advanced cancer. Clin Cancer Res. 2012; 18:568-76.

80. Slamon D, Hurvitz S, Applebaum S, Glaspy J, Allison M, DiCarlo B, *et al.* Phase I study of PD 0332991, cyclin-D kinase (CDK) 4/6 inhibitor in combination with letrozole for first-line treatment of patients with ER-positive, HER2-negative breast cancer. J Clin Oncol. 2010; 28:15s (suppl; abstr 3060).

81. Pfizer. NCT00721409: Study Of Letrozole With Or Without Palbociclib (PD-0332991) For The First-Line Treatment Of Hormone-Receptor Positive Advanced Breast Cancer Available from: https://clinicaltrials.gov/ct2/show/NCT00721409 (Accessed 31 October 2016).

82. DeMichele A, Clark AS, Tan KS, Heitjan DF, Gramlich K, Gallagher M, *et al.* CDK 4/6 inhibitor palbociclib (PD0332991) in Rb+ advanced breast cancer: phase II activity, safety, and predictive biomarker assessment. Clin Cancer Res. 2015; 21:995-1001.

83. ODwyer P, Abramson Cancer Center of the University of Pennsylvania. NCT01037790: Phase II Trial of the Cyclin-Dependent Kinase Inhibitor PD 0332991 in Patients with Cancer Available from: https://clinicaltrials.gov/ct2/show/NCT01037790 (Accessed 31 October 2016).

84. Bell T, Crown JP, Lang I, Bhattacharyya H, Zanotti G, Randolph S, *et al.* Impact of palbociclib plus letrozole on pain severity and pain interference with daily activities in patients with estrogen receptor-positive/human epidermal growth factor receptor 2-negative advanced breast cancer as first-line treatment. Curr Med Res Opin. 2016; 32:959-65.

85. Finn RS, Crown JP, Ettl J, Schmidt M, Bondarenko IM, Lang I, *et al.* Efficacy and safety of palbociclib in combination with letrozole as first-line treatment of ER-positive, HER2-negative, advanced breast cancer: expanded analyses of subgroups from the randomized pivotal trial PALOMA-1/TRIO-18. Breat Cancer Res. 2016; 18:67.

86. Crown J, Finn RS, Ettl J, Boer K, Patel R, Thummala A, *et al.* Efficacy and safety of first-line palbociclib plus letrozole compared with letrozole alone in patients aged > 65 years with estrogen receptor-positive, HER2-negative advanced breast cancer: A subgroup analysis by age of the PALOMA-1/TRIO-18 trial. J Clin Oncol. 2015; 33 (suppl; abstr 571).

87. Finn RS, Crown J, Ettl J, Pinter T, Thummala A, Shparyk YV, *et al.* Clinical efficacy and safety profile of palbociclib (P) in combination with letrozole (L) as first-line treatment in patients (pts) with ER+ and HER2-advanced breast cancer (ABC) who have not received any systemic treatment (ST): A subgroup analysis of PALOMA-1/TRIO-18. J Clin Oncol. 2015; 33 (suppl; abstr 575).

88. Finn RS, Crown J, Lang I, Kulyk SO, Schmidt M, Patel R, *et al.* The effect of palbociclib (P) in combination with letrozole (L) on bone metastases in women with ER+/ HER2-metastatic breast cancer (MBC): Subanalysis from a randomized phase II study. J Clin Oncol. 2015; 33 (suppl; abstr 572).

89. Slamon D, Crown J, Lang I, Kulyk SO, Schmidt M, Patel R, *et al.* Longterm safety profile of palbociclib (P) in combination with letrozole (L) as firstline treatment for postmenopausal patients with ER+ and HER2 advanced breast cancer (ABC) (PALOMA1/TRIO18). J Clin Oncol. 2015; 33.

90. Bell T, Crown JP, Lang I, Bhattacharyya H, Zanotti G, Randolph S, *et al.* Abstract P5-19-19: Impact of adding palbociclib to letrozole on pain severity and pain interference with various activities of daily life in patients with ER+, HER2- metastatic breast cancer as first line treatment. Cancer Res. 2015; 75 (9 Suppl P5-19-19).

91. Pfizer. Clinical study report on trial A5481003: Phase 1/2, open-label, randomized study of the safety, efficacy and pharmacokinetics of letrozole plus PD 0332991 (oral CDK 4/6 inhibitor) and letrozole single agent for the first-line treatment of ER-positive, HER2-negative advanced breast cancer in postmenopausal women - data on file. . 2015.

92. Pfizer. Clinical study report on trial A5481008: A Randomized, Multicenter, Double-Blind Phase 3 Study of PD-0332991 (Oral CDK 4/6 Inhibitor) Plus Letrozole Versus Placebo Plus Letrozole for the Treatment of Postmenopausal Women With ER (+), HER2 (-) Breast Cancer Who Have Not Received Any Prior Systemic Anti-Cancer Treatment for Advanced Disease - data on file2016 11 August.

93. Finn RS, Martin M, Rugo HS, Jones S, Im S-A, Gelmon K*, et al.* Palbociclib and Letrozole in Advanced Breast Cancer. N Engl J Med. 2016; 375:1925-36.

94. Practice Update News 12 October 2016. Palbociclib + Letrozole Maintains Health-Related Quality of Life vs Letrozole Alone in Treatment-Naïve Postmenopausal ER-Positive, HER2-Negative Metastatic Breast Cancer Available from: https://www.practiceupdate.com/content/palbociclib-letrozole-maintains-health-related-quality-of-life-vs-letrozole-alone-in-treatment-naive-postmenopausal-er-positive-her2-negative-metastatic-breast-cancernbsp/45107/8/1/2 (Accessed 18 October 2016).

95. Pfizer. Protocol A5481003. Phase 1/2, Open-label, Randomized Study of The Safety, Efficacy, And Pharmacokinetics of Letrozole Plus PD 0332991 (Oral CDK 4/6 Inhibitor) And Letrozole Single Agent for The First-line Treatment of ER Positive, Her2 Negative Advanced Breast Cancer in Postmenopausal Women. Statistical Analysis Plan (SAP). Version 4.0. July 31, 2013.

96. Pfizer. Protocol A5481008. A Randomized, Multicenter, Double-Blind Phase 3 Study Of PD-0332991 (Oral CDK 4/6 Inhibitor) Plus Letrozole Versus Placebo Plus Letrozole For The Treatment Of Postmenopausal Women With ER (+), HER2 (-) Breast Cancer Who Have Not Received Any Prior Systemic Anti-Cancer Treatment For Advanced Disease. Statistical Analysis Plan (SAP). Version 3.0. March 15, 2016.

97. Pfizer Inc. Paloma-1. Protocol. A Phase 1 Clinical, Pharmacokinetic, and Pharmacodynamic Evaluation of 2 Schedules of Oral PD 0332991, a Cyclin-Dependent Kinase Inhibitor, in Patients With Advanced Cancer2004.

98. Pfizer Inc. Paloma-2. Protocol. A Randomized, multicenter, double-blind phase 3 study of PD-0332991 (Oral CDK 4/6 inhibior) plus letrozole versus placebo plus letrozole for the treatment of postmenopausal women with ER (+), HER2 (-) breast cancer who have not received any prior systemic anti-cancer treatment for advanced disease 2015.

99. Pfizer. Meeting report: UK Metastatic Breast Cancer Service Advisory Board. DoubleTree by Hilton Hotel, London. 21 April 2015.

100. Pfizer. Meeting report prepared for Pfizer UK by Sudler Medical Communications: UK Metastatic Breast Oncology Advisory Board Meeting. Thistle Euston Hotel, London, Tuesday 5th July, 2016.

101. Johnston S, Arteaga C. Lapatinib (Tykerb) plus letrozole (Femara) or letrozole alone for postmenopausal hormone receptor-positive metastatic breast cancer. P and T. 2009; 34:102.

102. Gershanovich M, Chaudri HA, Campos D, Lurie H, Bonaventura A, Jeffrey M*, et al.* Letrozole, a new oral aromatase inhibitor: Randomised trial comparing 2.5 mg daily, 0.5 mg daily and aminoglutethimide in postmenopausal women with advanced breast cancer. Ann Oncol. 1998; 9:639-45.

103. Finn RS, Crown JP, Ettl J, Pinter T, Thummala A, Shparyk Y*, et al.* Treatment patterns of postdisease progression in the PALOMA-1/TRIO-18 trial. 38th Annual San Antonio Breast Cancer Symposium December 8-12; 2015; San Antonio.

104. Brady MJ, Cella DF, Mo F, Bonomi AE, Tulsky DS, Lloyd SR*, et al.* Reliability and validity of the Functional Assessment of Cancer Therapy-Breast quality-of-life instrument. J Clin Oncol. 1997; 15:974-86.

105. The EuroQol Group EuroQol--a new facility for the measurement of health-related quality of life. Health Policy. 1990; 16:199-208.

106. National Institute for Health Research Horizon Scanning Centre. Palbociclib for advanced ER-positive/HER2-negative breast cancer in post-menopausal women -

first line in combination with letrozole (Structured abstract),. Health Technology Assessment Database: NIHR Horizon Scanning Centre (NIHR HSC),. 2014,.

107. National Institute for Health and Care Excellence (NICE). Dabrafenib for treating unresectable, advanced or metastatic BRAFv600 mutation-positive melanoma (TA 321). 2014.

108. Holt S, Bertelli G, Humphreys I, Valentine W, Durrani S, Pudney D*, et al.* A decision impact, decision conflict and economic assessment of routine Oncotype DX testing of 146 women with node-negative or pNImi, ER-positive breast cancer in the U.K (Provisional abstract). Br J Cancer. 2013; 108(11): Available from: http://onlinelibrary.wiley.com/o/cochrane/cleed/articles/NHSEED-22013043900/frame.html.

109. Department of Health. Drugs and pharmaceutical market information (eMIT). [updated Updated 4 May 2016]; Available from: www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit.

110. National Health Service. Reference costs 2014-20152015 November.

111. Curtis L, Burns A. Unit Costs of Health and Social Care. 2015.

112. Nafees B, Patel C, Ray D, Gray L, Lau H, Lloyd A. An Assessment of Health-State Utilities in Metastatic Breast Cancer in the United Kingdom. ISPOR 21st Annual International Meeting, Washington, DC, USA, May 21 - May 25, 2016. 2016.

113. National Institute for Health and Care Excellence. Everolimus in combination with exemestane for treating advanced HER2-negative hormone-receptor-positive breast cancer after endocrine therapy (Technology appraisal guidance 295, 28 August 2013)2013.

114. National Institute for Health and Care Excellence (NICE). Fulvestrant for the treatment of locally advanced or metastatic breast cancer (TA 236). 2011.

115. National Institute for Health and Care Excellence (NICE). Guide to the methods of technology appraisal 2013. Available from: https://www.nice.org.uk/process/pmg9/chapter/foreword (Accessed 24 November 2016).

116. Bergh J, Jonsson P-E, Lidbrink EK, Trudeau M, Eiermann W, Brattstrom D*, et al.* FACT: An open-label randomized phase III study of fulvestrant and anastrozole in combination compared with anastrozole alone as first-line therapy for patients with receptor-positive postmenopausal breast cancer. J Clin Oncol. 2012; 30:1919-25.

117. Kind PH, Geoffrey; Macran, Susan. UK population norms for EQ-5D: Centre for Health Economics, University of York 1999.

118. Department of Health. A simple guide to payment by results. 2012; Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/213150/PbR-Simple-Guide-FINAL.pdf (Accessed 24 November 2016).

119. Davis S. Assessing technologies that are not cost-effective at a zero price: report by the Decision Support Unit. 2014; Available from: https://www.ncbi.nlm.nih.gov/pubmedhealth/PMH0088909/pdf/PubMedHealth_PMH0088909.pdf (Accessed 30 November 2016).

10 APPENDICES

10.1 Additional secondary efficacy endpoints reported in the PALOMA-1 and PALOMA-2 trials

The PALOMA-1 trial

The company focuses on the investigator assessed results for the secondary outcomes, ORR, CBR, DOR and TTP, although BICR results were also provided in the CS for comparison. The definitions and methods of analysis for these secondary efficacy outcomes are provided in Table 5.

Table 35 Description and method of analysis for secondary efficacy outcomes (other than time to progression and overall survival) reported in the PALOMA-1 trial

Outcome	Description	Statistical analysis
ORR	Defined according to RECIST 1.0 from the lesion measurements	[redacted]
CBR	Defined as per RECIST 1.0 as complete response, partial response or stable disease lasting at least 24 weeks	[redacted]
DOR	Time from first documentation of complete or partial response to date of first documentation of objective progression or death	[redacted]

CBR=clinical benefit rate; CI=confidence interval; DOR=duration of response; ITT=intention-to-treat; K-M=Kaplan-Meier; OR=odds ratio; ORR=objective response rate; RECIST=response evaluation criteria in solid tumors
Source: CS, adapted from Table 13, Table 19 and Table 20

The ERG is satisfied that the analysis method for each of these efficacy outcomes was pre-specified in the TSAP, and that all results were reported fully in the CSR. The ERG notes that one sided hypothesis testing was used for the outcomes of ORR and CBR, and asked the company to provide to justify the use of this approach to hypothesis testing. As part of their response to the ERG clarification letter, the company confirmed that one-sided hypothesis testing was deemed suitable due to there being sufficient confidence that PAL+LET was more effective than LET alone, and additionally, that it was more efficient statistically, considering an expected small sample size, under the null hypothesis to use one-sided testing. The ERG is satisfied with the company's justification, although the ERG considers that rationale for such an important statistical decision ought to have been provided in the protocol and/or TSAP.

The PALOMA-2 trial

Although the company focuses on the investigator assessed results for the secondary outcomes, ORR, CBR, and DOR, BICR results were also provided in the CS for comparison. The definitions and methods of analysis for these efficacy outcomes are listed in Table 6.

Table 36 Description of efficacy outcomes reported in the PALOMA-2 trial

Outcome	Description	Statistical analysis
ORR	Defined according to RECIST 1.1 from the lesion measurements	[redacted]
CBR	Defined as per RECIST 1.1 as complete response, partial response or stable disease lasting at least 24 weeks	[redacted]
DOR	Time from first documentation of complete or partial response to date of first documentation of objective progression or death	[redacted]

CBR=clinical benefit rate; CI=confidence interval; CR=complete response; DOR=duration of response; ITT=intention-to-treat; K-M=Kaplan-Meier; OR=odds ratio; ORR=objective response rate; PR=partial response; RECIST=response evaluation criteria in solid tumors; SD=stable disease
Source: CS, adapted from Table 16, Table 19 and Table 20, and the company's response to the ERG clarification letter

The ERG is satisfied that the analysis method for each of these efficacy outcomes was pre-specified in the TSAP, and that all results were reported fully in the CSR.

10.2 ERG assessment of proportional hazards in the PALOMA-1 trial

The ERG requested clarification from the company on whether any PH testing had been conducted for the PFS or OS data from the PALOMA-1 trial. In the company's response to the ERG clarification letter, it was not clear whether the company had performed an assessment of PH for either the PFS or OS data. Consequently, the ERG performed their own assessments of PH using PFS and OS data from the PALOMA-1 trial. The ERG produced cumulative hazard versus cumulative hazard (H-H) plots and log-log plots for PFS (Figure 22 and Figure 23) and OS data (Figure 24 and Figure 25). To demonstrate proportionality of hazards, the H-H plot should demonstrate a straight line trend, with individual data points distributed close to and on either side of the trend line, which should pass through the graph origin (zero value on both axes). The log-log plots should show that the curves for both treatments are approximately parallel if the PH assumption is valid.

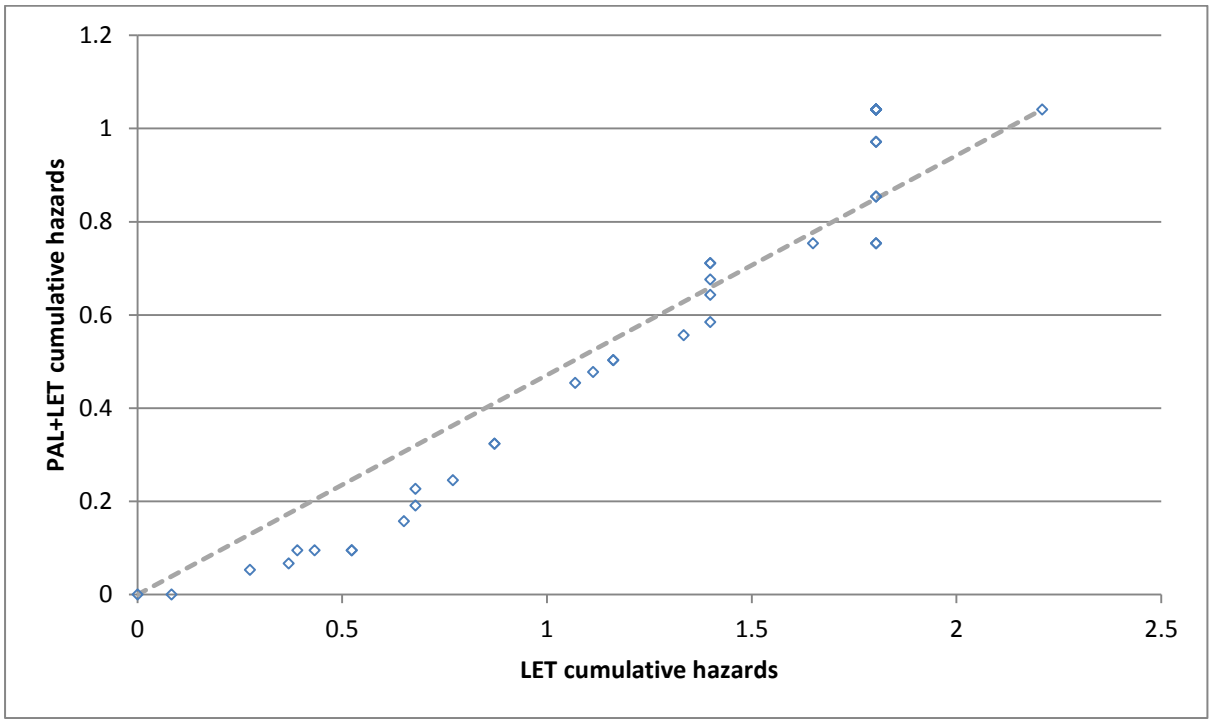

Figure 22 PALOMA-1 PFS H-H plot

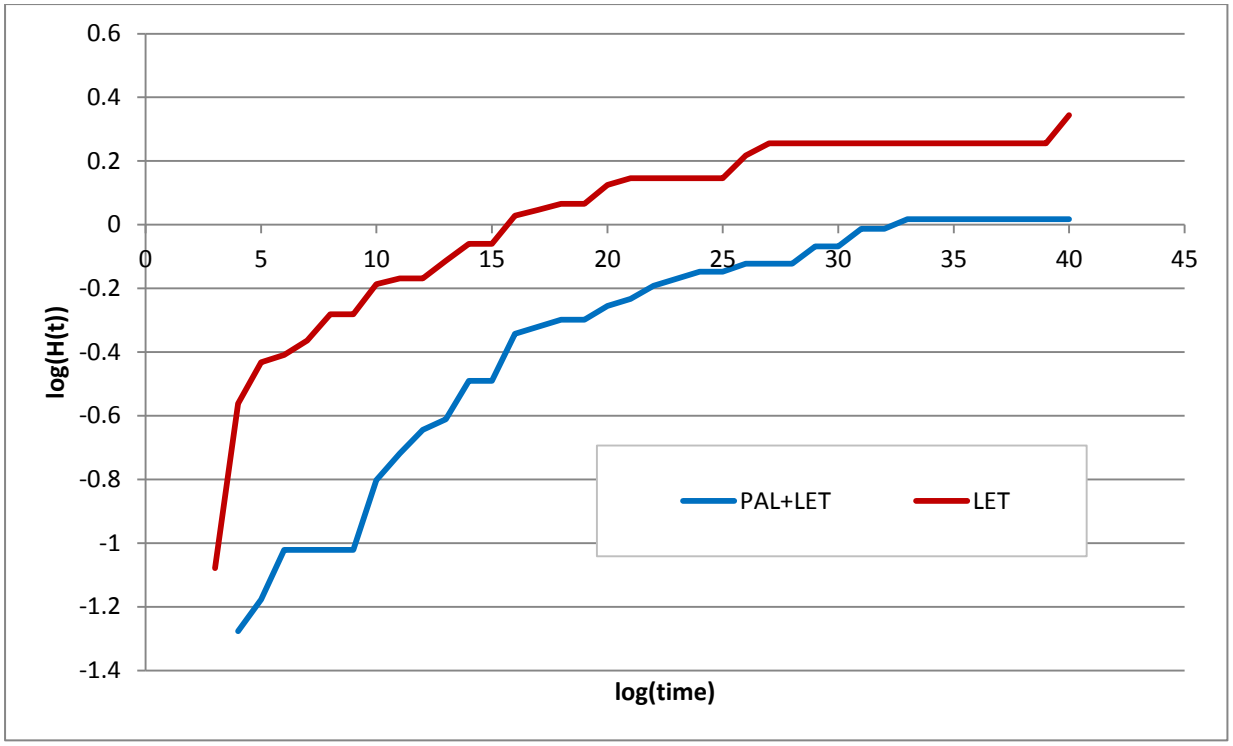

Figure 23 PALOMA-1 PFS log-log plot

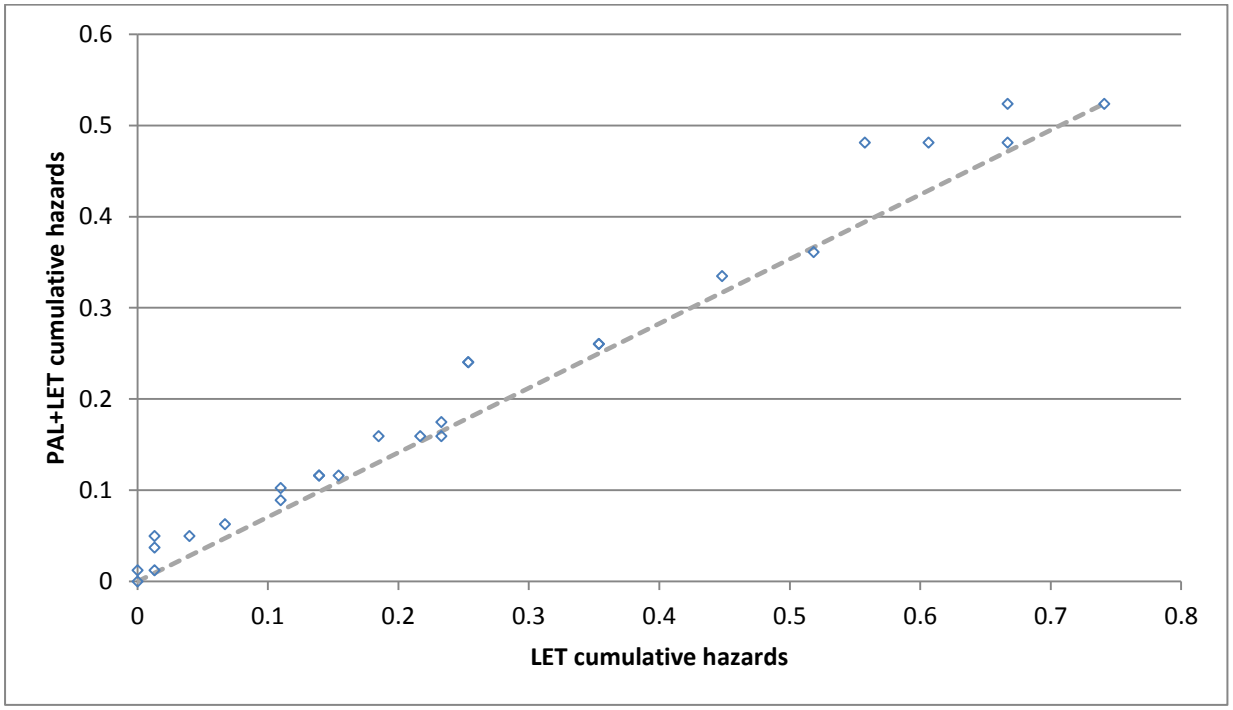

Figure 24 PALOMA-1 OS H-H plot

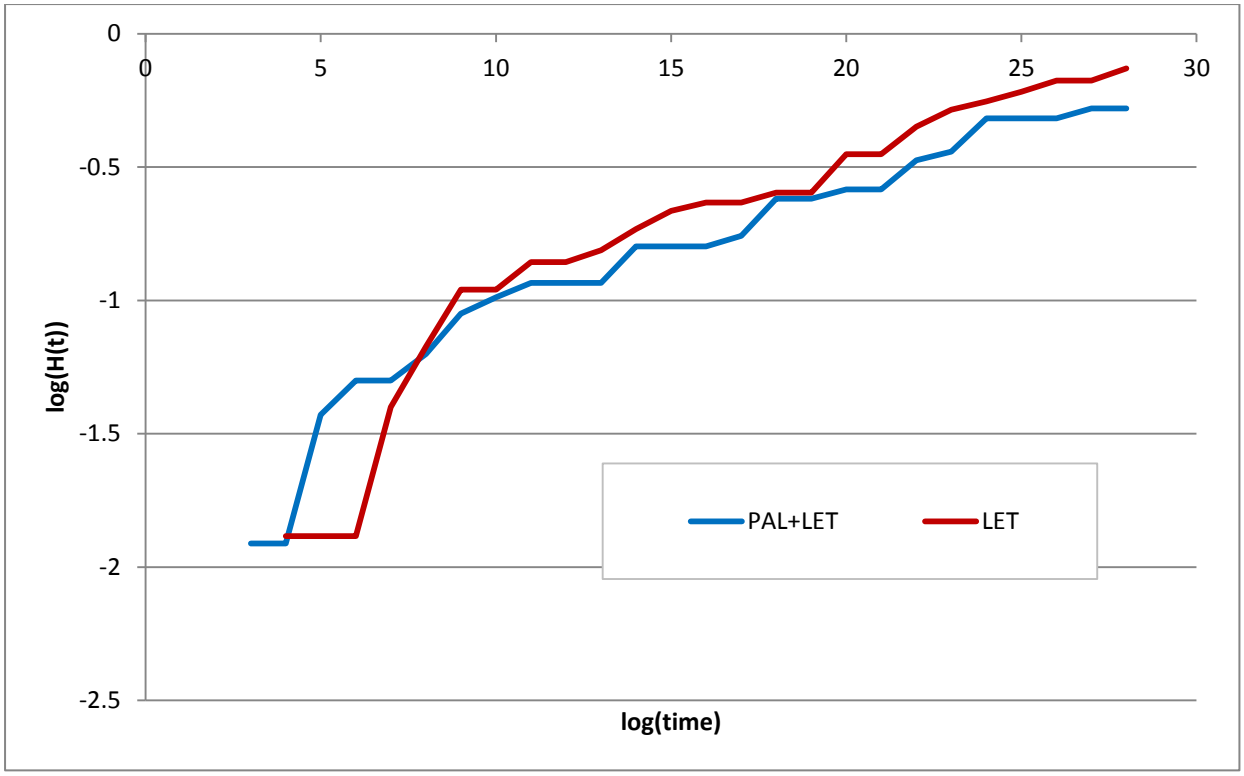

Figure 25 PALOMA-1 OS log-log plot

The ERG considered that it is reasonable to assume that the PH assumption is valid for PFS data, as the log-log plot (Figure 23) demonstrates that the curves are approximately parallel for PAL+LET and LET. Although, individual data points aren't quite randomly scattered about the trend line in the H-H plot (Figure 22), further investigation revealed that when considering data from 100 days onwards, the H-H plot is satisfactory (data not shown). In the first 100 days, the PH assumption does not hold due to the drop off in PFS in the LET arm at the time of the first tumour assessment. The ERG considers that PH is valid for the remainder of the trial period.

For OS, the log-log plot clearly demonstrates that the curves for PAL+LET and LET cross at approximately 8 months, indicating that the assumption of PH is not valid. Therefore, the use of HRs to summarise treatment effect for OS is not appropriate.

10.3 Results from univariate and multivariate analyses of progression-free survival in the PALOMA-2 trial

The results of the univariate and multivariate analyses of PFS are provided in Table 37.

Table 37 Results of the univariate and multivariate analyses of PFS

PFS analysis		PAL+LET versus PLACEBO+LET, hazard ratio (95% CI)
Univariate	Investigator assessed	[redacted]
	BICR	[redacted]
Multivariate	Investigator assessed	[redacted]
	BICR	[redacted]

BICR=blinded independent central review; CI=confidence interval; PFS=progression-free survival

10.4 Subsequent treatment received on disease progression in the PALOMA-1 trial

Data on second-line treatment received following disease progression in the PALOMA-1 trial presented at the 38th San Antonio Breast Cancer Symposium in December 2015[103] are summarised in Table 38.

Table 38 First subsequent treatment after progression on study treatment in the PALOMA-1 trial*

Type of treatment received	PAL+LET n=33	LET n=53
Endocrine therapy, n (%)†	15 (45.4)	32 (60.4)
• Exemestane	1 (3.0)	7 (13.2)
• Fulvestrant	9 (27.3)	12 (22.6)
• Letrozole	1 (3.0)	5 (9.4)
• Medroxyprogesterone	4 (12.1)	1 (1.9)
• Tamoxifen	0 (0.0)	7 (13.2)
Chemotherapy, n (%)†	17 (51.5)	21 (39.6)
• Capecitabine	1 (3.0)	4 (7.5)
• Cyclophosphamide	1 (3.0)	3 (5.7)
• Cyclophosphamide/epirubicin/fluorouracil	2 (6.1)	1 (1.9)
• Docetaxel	1 (3.0)	2 (3.8)
• Doxorubicin	1 (3.0)	4 (7.5)
• Epirubicin	2 (6.1)	1 (1.9)
• Fluorouracil	1 (3.0)	2 (3.8)
• Gemcitabine	3 (9.1)	1 (1.9)
• Mitoxantrone	1 (3.0)	1 (1.9)
• Paclitaxel	10 (30.3)	8 (15.1)
• Vinorelbine	1 (3.0)	0 (0.0)
Other therapy, n (%)†	6 (18.2)	13 (24.5)
• Bevacizumab	3 (9.1)	4 (7.5)
• Blinded therapy	3 (9.1)	3 (5.7)
• Everolimus	0 (0.0)	3 (5.7)
• Other	0 (0.0)	5 (9.4)

*These are patients for whom post-progression treatment data were available at data cut-off; note: disease progression on study treatment had occurred in 40 of the 84 patients (47.6%) in the PAL+LET arm and 59 of the 81 patients (72.8%) in the LET-alone arm

† Patients with >1 therapy as the first subsequent therapy after disease progression starting on the same day are reported under each therapy

Source: Finn et al 2015,[103] Table 3

10.5 Other secondary efficacy outcome results from the PALOMA-1 and PALOMA-2 trials

The PALOMA-1 trial

The results of the analyses for the secondary outcomes of PALOMA-1 not reported in the main body of this ERG report are provided in Table 39. ORR was analysed for both the ITT population, and in the subpopulation of patients with measurable disease. All other outcomes were analysed using the ITT population. The company presented both investigator assessed and BICR results where applicable.

Table 39 Additional secondary efficacy outcome results from the PALOMA-1 trial[a]

Outcome	PAL+LET (n=84)	LET (n=81)	p-value between arms[b]
ITT population (n)	84	81	-
Patients with measurable disease (n)	65	66	-
ORR, % (95% CI)			
Investigator assessed	43 (32 to 54)	33 (23 to 45)	p=0.13
BICR[c]	30 (20 to 41)	21 (13 to 32)	p=0.1314
ORR in patients with measurable disease, % (95% CI)			
Investigator assessed	55 (43 to 68)	39 (28 to 52)	p=0.047
BICR[c]	49 (35 to 63)	32.7 (20 to 47)	p=0.0728
CBR, % (95% CI)			
Investigator assessed	81 (71 to 89)	58 (47 to 69)	p=0.0009
BICR[c]	71 (61 to 81)	51 (39 to 62)	p=0.0046
Stable disease lasting at least 24 weeks, %			
Investigator assessed	38.1	24.7	-
BICR[c]	41.7	29.6	-

[a]Results are presented for the ITT population unless otherwise noted
[b]All p-values are one-sided p-values, although no formal testing was performed for secondary endpoints; nominal p-values were reported but no multiplicity adjustments were made for the secondary analyses
[c]BICR was conducted on 97% of the ITT population
BICR=blinded independent central review; CI=confidence interval; CBR=clinical benefit rate; ITT=intention-to-treat; NE=not estimable; ORR=objective response rate; OS=overall survival; TTP=time to progression
Source: CS, adapted from Table 22 and Table 23, and CSR, Table 36

In the ITT population, ORR was higher among patients who received PAL+LET than among those who received LET alone, although this difference was not found to be statistically significant (investigator assessed ORR: 43% versus 33%, p=0.13). The ITT population included patients with both measurable and non-measurable disease. The company states that non-measurable disease was comprised principally by bone-only disease, and that it was important to include these patients in the trial owing to their significant representation of the advanced breast cancer (ABC) population. However, the company states that there are inherent inaccuracies associated with assessing ORR for non-measurable/bone-only disease and that the inclusion of these patients in the ITT population for the analysis of ORR may have contributed to the failure of the ITT population to report significant ORR differences between

the two trial arms. In the measurable disease population, a statistically significant difference was identified for ORR between PAL+LET and LET alone (55% versus 39%, p=0.047).

Results for BICR also suggested a trend in favour of PAL+LET in terms of ORR for both the ITT and measurable disease populations, although the ERG notes that ORR was considerably lower for both treatment groups when assessed by BICR, in comparison to ORR obtained by investigator-assessment.

CBR was found to be statistically significantly higher for PAL+LET patients than LET patients (81% versus 58%, p=0.0009). The company argues that CBR may be a better measure of treatment benefit than ORR for a treatment which has a disease stabilisation component, as CBR incorporates both stable disease for at least 24 weeks, and ORR. Within CBR, the proportion of patients showing stable disease for at least 24 weeks was higher for PAL+LET patients than for those receiving LET alone (38.1% versus 24.7%). BICR results for both CBR and stable disease were broadly comparable to those obtained by investigator-assessment. Clinical advice to the ERG was that CBR is indeed a better tool for assessing efficacy than ORR, as bone only disease is incredibly difficult to assess response rates with existing imaging modalities. The ERG therefore agrees with the company that is appropriate to consider ORR in patients with measurable disease as well as in the ITT population, and also to consider the results of the analyses of CBR.

The PALOMA-2 trial

The results of the analyses for the secondary outcomes of PALOMA-2 not reported in the main body of this ERG report are provided Table 40. ORR and DOR were analysed for both the ITT population, and in the subpopulation of patients with measurable disease. All other outcomes were analysed using the ITT population. The company presented both investigator assessed and BICR results where applicable.

In the ITT population, ORR was higher among patients who received PAL+LET than those who received PLACEBO+LET (42.1% versus 34.7%), although this difference was not found to be statistically significant (odds ratio [OR]=1.40; 95% CI 0.98 to 2.01). The BICR result for this population achieved statistical significance. In the population of patients with measurable disease, a statistically significant difference was identified for ORR between PAL+LET and PLACEBO+LET (55.3% versus 44.4%), corresponding to an OR of 1.55 (95% CI 1.05 to 2.28). For the measurable disease population, the BICR result was in accordance with investigator assessed ORR.

CBR was found to be statistically significantly higher for PAL+LET patients than PLACEBO+LET patients (84.9% versus 70.3%), corresponding to an OR of 2.39 (95% CI 1.58 to 3.59). BICR results for CBR were broadly comparable to those obtained by investigator-

assessment. Within CBR, the proportion of patients showing stable disease for at least 24 weeks was [redacted] for PAL+LET patients than for those receiving PLACEBO+LET [redacted].

Table 40 Additional secondary efficacy outcome results from the PALOMA-2 trial[a]

	PAL+LET (n=84)	LET (n=81)	p-value between arms[b]
ITT population (n)	444	222	-
Patients with measurable disease (n)	338	171	-
ORR, % (95% CI)			
Investigator assessed	42.1 (37.5 to 46.9)	34.7 (28.4 to 41.3)	0.0310
BICR[c]	[redacted]	[redacted]	[redacted]
ORR in patients with measurable disease, % (95% CI)			
Investigator assessed	55.3 (49.9 to 60.7)	44.4 (36.9 to 52.2)	0.0132
BICR[c]	[redacted]	[redacted]	[redacted]
CBR, % (95% CI)			
Investigator assessed	84.9 (81.2 to 88.1)	70.3 (63.8 to 76.2)	<0.0001
BICR[c]	[redacted]	[redacted]	[redacted]
DOR, median (months), (95% CI)			
Investigator assessed	22.5 (19.8-28.0)	16.8 (14.2-28.5)	NA
BICR[c]	[redacted]	[redacted]	[redacted]
DOR in patients with measurable disease, median (months) (95% CI)			
Investigator assessed	22.5 (19.8-28.0)	16.8 (15.4-28.5)	NA
BICR[c]	[redacted]	[redacted]	[redacted]
Stable disease ≥24 weeks in confirmed cases of the ITT population, %			
Investigator assessed	[redacted]	[redacted]	-

[a]Results refer to the ITT population unless otherwise noted
[b]All p-values are one-sided p-values
[c]BICR was conducted on the entire ITT population
BICR=blinded independent central review; CBR=clinical benefit response; CI=confidence interval; DOR=duration of response; ITT=intention-to-treat; NA=not applicable; NE=not estimable; ORR=objective response rate
Source: CS, adapted from Table 24 and 25, and the CSR, Table 27

10.6 Most common adverse events in the PALOMA trials

Table 41 Most common (>20% in any treatment arm) treatment emergent adverse events in the PALOMA-1 and PALOMA-2 trials

Adverse events, n (%)	PALOMA-1						PALOMA-2					
	PAL+LET (n=83)			LET (n=77)			PAL+LET (n=444)			LET (n=222)		
	All-Grade	Grade 3	Grade 4	All-Grade	Grade 3	Grade 4	All-Grade	Grade 3	Grade 4	All-Grade	Grade 3	Grade 4
Neutropenia*	62 (74.7)	40 (48.2)	5 (6.0)	4 (5.2)	1 (1.3)	0	353 (79.5)	249 (56.1)	46 (10.4)	14 (6.3)	2 (0.9)	1 (0.5)
[redacted]												
Leukopenia*	36 (43.4)	16 (19.3)	0	2 (2.6)	0	0	173 (39.0)	107 (24.1)	3 (0.7)	5 (2.3)	0	0
Fatigue	34 (41.0)	2 (2.4)	2 (2.4)	18 (23.4)	1 (1.3)	0	166 (37.4)	8 (1.8)	0	61 (27.5)	1 (0.5)	0
Nausea	21 (25.3)	2 (2.4)	0	10 (13.0)	1 (1.3)	0	156 (35.1)	1 (0.2)	0	58 (26.1)	4 (1.8)	0
Arthralgia	19 (22.9)	1 (1.2)	0	12 (15.6)	2 (2.6)	0	148 (33.3)	3 (0.7)	0	75 (33.8)	1 (0.5)	0
Alopecia	18 (21.7)	0	0	2 (2.6)	0	0	146 (32.9)	0	0	35 (15.8)	0	0
Diarrhoea	17 (20.5)	3 (3.6)	0	8 (10.4)	0	0	116 (26.1)	6 (1.4)	0	43 (19.4)	3 (1.4)	0
Cough	10 (12.0)	0	0	8 (10.4)	0	0	111 (25.0)	0	0	42 (18.9)	0	0
Anaemia	29 (34.9)	4 (4.8)	1 (1.2)	5 (6.5)	1 (1.3)	0	103 (23.2)	23 (5.2)	1 (0.2)	20 (9.0)	4 (1.8)	0
Back pain	12 (14.5)	0	1 (1.2)	12 (15.6)	1 (1.3)	0	96 (21.6)	6 (1.4)	0	48 (21.6)	0	0
Headache	12 (14.5)	0	0	8 (10.4)	0	0	95 (21.4)	1 (0.2)	0	58 (26.1)	4 (1.8)	0
Hot flush	17 (20.5)	0	0	9 (11.7)	0	0	93 (20.9)	0	0	68 (30.6)	0	0

Source: CS, Tables 39 and 41 and published paper for the PALOMA-2 trial[93]

*In the PALOMA-2 trial, neutropenia was categorised according to the Medical Dictionary for Regulatory Activities (MedDRA) preferred terms neutropenia and neutrophil count decreased and leukopenia was categorised according to the MEDRA preferred terms leukopenia and white blood cell count decreased

10.7 Calculation of post-progression utility values

Utility values were transformed using the formula from Lloyd et al:[5]

$$Transformed\ utility = ln\left(\frac{1 - utility}{utility}\right)$$

And back transformed using the following formula:

$$Utility = \frac{1}{1 + \exp(transformed\ utility)}$$

10.8 Adjusting a parametric curve using medians

The primary assumption underlying the company's modelling of OS in the base case is that post-progression survival is equal for patients treated with PAL+LET and PLACEBO+LET; that is, all survival gain is accrued in the progression-free state. However, this assumption is not borne out in the model, as using *medians* to recalibrate the OS curve for PAL+LET has resulted in a *mean* PPS gain for patients treated with PAL+LET. By subtracting PFS from OS on a cycle-by-cycle basis, the ERG has calculated a mean PPS gain for PAL+LET of 0.49 months in the base case.

The reason that mean OS gain increases when a Weibull model is adjusted based on its median is because the ratio of median to mean is based on the interaction of the shape and scale parameters used to specify the curves. The Weibull distribution fitted to the OS K-M data from the PAL+LET arm of the PALOMA-1 trial and the adjusted version of this model used in the base case are both right skew, which means that the mean is greater than the median in both cases. The ratio of median to mean is also different in both of these Weibull models. The combination of the right skew and the dynamic ratio of median to mean means that adjusting the scale parameter, as the company has, in order to achieve a larger median OS gain has a proportionately greater effect on mean OS for PAL+LET and, thus, on mean OS gain.

Table 42 shows how the ratio of median to mean OS gain when using the adjusted base case model for PAL+LET is proportionately greater than when using the unadjusted Weibull model (0.830 versus 0.773).

Table 42 Comparison of median and mean OS between the base case and the unadjusted PALOMA-1 model

	Median in model (months)		Mean in model (months)		Median:Mean	
	OS	OS Gain	OS	OS Gain	OS	OS Gain
PAL+LET (base case adjusted Weibull)	44.3	9.3	49.9	11.2	0.888	0.830
PAL+LET (PALOMA-1 IPD Weibull)	40.1	5.1	45.3	6.6	0.885	0.773
PLACEBO+LET (PALOMA-1 IPD Weibull)	35.0	-	38.7	-	0.904	-

Source: Company model, ERG calculations
*Note: Some values given in the CS differ from those in the model. Model values have been used where discrepancies exist.

10.9 Company scenario analyses 27 to 36

The company argues that limitations it has identified in the ICER per QALY gained calculation, might be mitigated by the adoption of certain assumptions regarding the cost of the comparator, the pre-progression utility value, the modelling of OS, and the cost of care in the post-progression state. The company has put together several assumptions in different combinations that yield ICERs per QALY gained of under £50,000 (company Scenarios 32 to 36). Since each of these assumptions in isolation has flaws and/or breaches standard NICE methods, the ERG considers the uncertainty inherent in the combined scenarios to render them uninformative.

The company presents a variety of exploratory scenarios in which it investigates the effects on the ICER per QALY gained of varying the assumptions in the model beyond the parameters of the standard sensitivity analyses. The ICERs per QALY gained in the company's exploratory scenarios range from £26,996 to £312,635 (Table 43).

Table 43 Company exploratory scenario analyses varying model assumptions (palbociclib at list price)

Scenario # in CS	Assumptions varied	ICER/QALY gained	ICER change
Base case deterministic ICER		**£150,869**	**-**
27	Only PFS gain for PAL+LET (10.3 months) No OS gain for PAL+LET (0 months)	£312,635	+ £161,766
28a	Increase median OS gain for PAL+LET to 5 years	£61,822	- £89,047
28b	Increase median OS gain for PAL+LET to 5 years, *but removing post-progression costs*	£42,794	- £108,075
29	Increase in utility of +0.1 for patients in the PFS state	£134,134	- £16,735
30	A comparator with the same monthly acquisition costs *(i.e. fixed cost of £2,951.52 per month, but only for respective treatment durations)*	£53,074	- £97,795
31	Reduced treatment duration by 12 months in each arm *(PFS reduced from 15.7 to 3.7 months for LET, and from 24.9 to 12.9 months for PAL+LET)*	£86,419	- £64,450
32	• Comparative monthly acquisition costs • Value of PFS utility increase (+0.1) *No change to base case OS assumption*	£47,187	-£103,682
33	• Comparative monthly acquisition costs (#30) • Value of PFS utility increase (+0.1) • Incremental median OS gain of 12 months	£43,819	-£107,050
34	• Comparative monthly acquisition costs • Value of PFS utility increase (+0.1) • Incremental median OS gain of 12 months • Removal of post-progression costs	£40,482	-£110,387
35	• Comparative monthly acquisition costs (#30) • Value of PFS utility increase (+0.1) • Incremental median OS gain of 24 months	£36,194	-£114,675

Scenario # in CS	Assumptions varied	ICER/QALY gained	ICER change
36	• Comparative monthly acquisition costs • Value of PFS utility increase (+0.1) • Incremental median OS gain of 24 months • Removal of post-progression costs	£26,996	-£123,873
From scenarios 33 and 34	Incremental OS gain of 12 months	£134,294	-£16,575
From scenarios 35 and 36	Incremental OS gain of 24 months	£95,656	-£55,213
From scenarios 28b, 34, 35 & 36	Remove all post-progression costs	£150,303	-£566

ICER=incremental cost effectiveness ratio; PFS=progression free survival; QALY=quality adjusted life year; OS=overall survival
Source: CS Table 85; CS Table 86; ERG calculations

The company's exploratory scenarios fall into to one (or a combination) of four categories: OS gain for PAL+LET; acquisition costs of letrozole; PFS utility values; and post-progression costs.

Company exploratory scenarios: OS gain for PAL+LET

The ERG considers it justifiable to explore alternative OS scenarios given the problems inherent in the PALOMA-1 data, however the ERG considers the magnitude of the gains modelled to be implausible given the preliminary data available from the PALOMA-1 and PALOMA-2 trials, and is not aware of any other data that would support such gains.

The company presents these scenarios to demonstrate the importance of OS on the ICER per QALY gained. The company states that that treatment with PAL+LET would need to extend life by approximately 9 years to yield an ICER per QALY gained of around £50,000 (with palbociclib at list price and all other base case assumptions remaining the same), which it notes is not clinically plausible. However, the price of the drug also influences the impact of extended time spent in PFS. If the cost of palbociclib were to increase or decrease, and all other elements of the model were to stay the same, the size of the OS gain required to bring the ICER down towards the NICE threshold would also increase or decrease

The company supports its modelling of improved OS gains for treatment with PAL+LET versus treatment with LET by suggesting that people with stable disease are less likely to die (and thus time in PFS will be reflected in time in OS). Supported by additional evidence, amendments to the model structure could be made to apply differential death rates to the pre-progression and post-progression health states within the model and produce further scenario analyses. However, the lack of maturity of the OS data means any such estimates at present, if calculated, would carry substantial uncertainty.

Company exploratory scenarios: acquisition costs of letrozole

The company argues that the introduction of a new treatment, such as palobociclib, as an add-on therapy or into a therapy area with no new treatment or breakthrough, inherently values that new treatment less than if the therapy area had already benefitted from recent innovation. The company attempts to show that this is the case by running a scenario where the price of LET monotherapy is equal to the price of PAL+LET. Whilst the ERG agrees that a comparison with a generic drug makes it relatively more difficult to demonstrate cost-effectiveness in the mathematical sense, NICE methods do not allow for deviation on this basis as the true opportunity cost for the NHS must be considered in potentially reallocating resources from a generic to a proprietary drug.

The ERG also considers the implementation of this scenario to be methodologically flawed as, rather than changing the price of letrozole to equal that of palbociclib and thus double the cost of the combined therapy, only the price of letrozole when used as monotherapy is amended. The ERG does not therefore consider the comparative acquisition costs scenario as plausible in practice as if letrozole had a higher list price, this would also be the price for use in combination with palbociclib.

Company exploratory scenarios: PFS utility values

The company argues that PFS is undervalued for a number of reasons in Section 3.2.1 of the CS and presents a sensitivity analysis in which the utility of the PFS health state is increased by 0.1 which results in an ICER of £134,134. The ERG considers that many of the arguments put forward by the company are in fact adequately reflected in the utility values used to represent the health states within the model. The benefit of having stable disease (being in the pre-progressed health state) in the model is an improvement in health-related quality of life of more than 0.2 (on the 0-1 utility scale) over the progressed health state, in both the company estimated and the ERG re-calculated utility values. This incremental benefit exists for the duration of any PFS extension offered by PAL + LET treatment in comparison to LET alone. The value used to estimate progressed utility is taken from a study of patients receiving chemotherapy and therefore any difference in AE profiles or psychological impacts between the treatments received pre- and post-progression is represented within the difference between the health-related quality of life values.

The ability to continue to work is captured within the activities of daily living question which forms part of the EQ-5D questionnaire in which patients would indicate a lower score if their normal working pattern was disrupted. The costs to the patients of being unable to undertake paid employment cannot be considered as part of the NICE appraisal process without discriminating in favour of individuals of working-age.

The company argues that the burden on carers of patients with this disease is so substantial that its exclusion contributes to undervaluing the benefit of PFS. The company does not however present any evidence to quantify the health-related quality of life impact of caring for a patient with progressed disease may have, nor explore this as an individual hypothetical scenario within the modelling.

The data used to value PFS in this model are the best available and consistent with the NICE reference case, which is used to benchmark all appraisals. Any departure from EQ-5D values directly obtained from patients would only be supported given significant evidence of the insufficiency of the EQ-5D to capture all elements relevant to patients in this disease area. Given that the arguments put forward by the company do not appear specific to postmenopausal women with ER+/HER2- ABC who have never received systemic therapy in the LABC/MBC setting but could in fact be relevant to all patients with ABC, or the population of people with breast cancer as a whole, any methodological change to the valuation of utility would have implications for all appraisals of breast cancer interventions.

Company exploratory scenarios: post-progression costs

The company includes the removal of post-progression costs as part of their scenarios with combinations of amendments (Scenarios 28b, 34 & 36). As the only post-progression costs that are included within the company model are the costs of monitoring patients undergoing further therapy, the impact of removing these costs is minimal. As shown in Table 43, the ICER decreases by £566.

In addition, the DSU discussion paper regarding cost-effectiveness at zero price[119] considers scenarios in which non-treatment related costs could be excluded however concludes that a narrow perspective does not enable full consideration of the opportunity cost to the NHS of the introduction of a new technology and therefore the ERG does not consider this element of the scenario analyses plausible.

10.10 ERG Revisions to company's model

All revisions are activated by a logic switch with:

0 = unchanged

1 = apply ERG modification

Logic switches are indicated by named range variables Mod_*letter* where letter = A - L.

A menu of revisions and Mod names appears below and on the 'Results_Deterministic' worksheet together with summary results as used to transfer to the ERG report.

Revision #	Modification name	Description
R1)	Mod_A	ERG OS estimates based on data from PALOMA-1
R2)	Mod_B	ERG PFS estimates based on data from PALOMA-1
R3)	Mod_C	ERG TTD estimates based on data from PALOMA-1
R4)	Mod_D	ERG recalculated pre-progression utility values from PALOMA-2 trial
R5)	Mod_E	ERG recalculated post-progression utility values using Lloyd 2006[5]
R6)	Mod_F	Use mid-cycle correction
R7)	Mod_G	Use full reference costs for AEs
R8)	Mod_H	Correct AE incidence calculation
R9)	Mod_I	Change discounting to annual
R10)	Mod_J	Use 365.25 days per year
R11)	Mod_K	ERG PFS estimates based on data from PALOMA-2
R12)	Mod_L	ERG TTD estimates based on data from PALOMA-2

Instructions for modifying the company model

1. Move all sheets from *palbo 915_ERG additional model data.xlsx* into company model

2. For each sheet given in the 'Sheet' column below:
    - copy formulae from the 'Modified formulae' column in the table below
    - paste formulae into the cells referred to in the 'Cells' column in the table below

ERG revision number and description	Modification name	Sheet	Cells	Modified formulae
R1) ERG OS	Mod_A	OS_L1	X58 copy down to X578	Amend PAL+LET OS =IF(Mod_A=0,CHOOSE(OS_model_scenario,W58,U58,V58),'ERG time to event_P1'!N11) N.B. amend formatting to multiple decimal places after pasting
R1) ERG OS	Mod_A	OS_L1	M58 copy down to M578	Amend LET OS =IF(Mod_A=0,K58, 'ERG time to event_P1'!O11)
R2) ERG PFS estimates based on data from PALOMA-1 AND R11) ERG PFS estimates based on data from PALOMA-2	Mod_B Mod_K	PFS_L1	W57 copy down to W577	Amend PAL+LET PFS =IF(AND(Mod_B=0,Mod_K=0),U57,IF(AND(Mod_B=1, Mod_K=0), 'ERG time to event_P1'!F11, IF(AND(Mod_B=0, Mod_K=1), 'ERG time to event_P2'!F11)))

ERG revision number and description	Modification name	Sheet	Cells	Modified formulae
R2) ERG PFS estimates based on data from PALOMA-1 AND R11) ERG PFS estimates based on data from PALOMA-2	Mod_B Mod_K	PFS_L1	M57 copy down to M577	Amend LET PFS =IF(AND(Mod_B=0,Mod_K=0),K57, IF(AND(Mod_B=1, Mod_K=0), 'ERG time to event_P1'!G11, IF(AND(Mod_B=0, Mod_K=1), 'ERG time to event_P2'!G11))) N.B. amend formatting to multiple decimal places after pasting
R3) ERG TTD estimates based on data from PALOMA-1 AND R12) ERG TTD estimates based on data from PALOMA-2	Mod_C Mod_L	EnginePAL_LET	AP11 copy down to AP531	Amend PAL+LET TTD =IF(AND(Mod_C=0,Mod_L=0),$F11*AP$9, IF(AND(Mod_C=1,Mod_L=0),'ERG time to event_P1'!V11*$AP$9, IF(AND(Mod_C=0,Mod_L=1), 'ERG time to event_P2'!V11*$AP$9)))
R3) ERG TTD estimates based on data from PALOMA-1 AND R12) ERG TTD estimates based on data from PALOMA-2	Mod_C Mod_L	EngineLET_PBO	AP11 copy down to AP531	Amend LET TTD =IF(AND(Mod_C=0,Mod_L=0),$F11*AP$9, IF(AND(Mod_C=1,Mod_L=0),'ERG time to event_P1'!W11*$AP$9, IF(AND(Mod_C=0,Mod_L=1), 'ERG time to event_P2'!W11*$AP$9)))
R4) ERG recalculated pre-progression utility values from PALOMA-2 trial	Mod_D	Utility	C18	Amend PAL+LET pre-progression utility =IF(mod_D=0,IF(D18="",CHOOSE(I18,E18,F18,G18,H18),D18), [redacted])
R4) ERG recalculated pre-progression utility values from PALOMA-2 trial	Mod_D	Utility	C12	Amend LET pre-progression utility =IF(Mod_D=0,IF(D12="",CHOOSE(I12,E12,F12,G12,H12),D12), [redacted])

ERG revision number and description	Modification name	Sheet	Cells	Modified formulae
R5) ERG recalculated post-progression utility values using Lloyd 2006	Mod_E	Utility	C19 copy down to C21	Amend PAL+LET post-progression utility =IF(D19="",IF(Mod_E=0,E19,0.5052),D19)
R5) ERG recalculated post-progression utility values using Lloyd 2006	Mod_E	Utility	C13 copy down to C15	Amend LET post-progression utility =IF(D13="",IF(Mod_E=0,E13,0.5052),D13)
R6) Use mid-cycle correction	Mod_F	ERG_mid cycle correction	B11 copy down to B532	Create mid-cycle PFS for PAL+LET =PFS_L1!W57
R6) Use mid-cycle correction	Mod_F	ERG_mid cycle correction	E11 copy down to E532	Create mid-cycle PFS for LET =PFS_L1!M57
R6) Use mid-cycle correction	Mod_F	ERG_mid cycle correction	J11 copy down to J532	Create mid-cycle OS for PAL+LET =OS_L1!X58
R6) Use mid-cycle correction	Mod_F	ERG_mid cycle correction	M11 copy down to M532	Create mid-cycle OS for LET =OS_L1!M58
R6) Use mid-cycle correction	Mod_F	EnginePAL_LET	D11 copy down to D531	Amend PAL+LET OS for mid-cycle correction =IF(Mod_F=0,MAX(1E-50,OS_L1!X58), MAX(1E-50,'ERG_mid cycle correction'!L11))
R6) Use mid-cycle correction	Mod_F	EnginePAL_LET	E11 copy down to E531	Amend PAL+LET PFS for mid-cycle correction =IF(Mod_F=0,PFS_L1!W57,'ERG_mid cycle correction'!D11)
R6) Use mid-cycle correction	Mod_F	EngineLET_PBO	D11 copy down to D531	Amend LET OS for mid-cycle correction =IF(Mod_F=0,MAX(1E-50,OS_L1!M58), MAX(1E-50,'ERG_mid cycle correction'!O11))
R6) Use mid-cycle correction	Mod_F	EngineLET_PBO	E11 copy down to E531	Amend LET PFS for mid-cycle correction =IF(Mod_F=0,PFS_L1!M57,'ERG_mid cycle correction'!G11)
R7) Use full reference costs for AEs	Mod_G	Cost_AE	C46 copy down to C47	Amend AE costs for Grade 3 and Grade 4 neutropenia =IF(Mod_G=0,IF(D46="",CHOOSE(I46,E46,F46,G46,H46),D46), 132)

ERG revision number and description	Modification name	Sheet	Cells	Modified formulae
R8) Correct AE incidence calculation	N/A	AE_P_2	F96 copy across and down to G97	Calculate AE rates =(-LN(1-C80)/$C$73)
R8) Correct AE incidence calculation	N/A	AE_P_2	H96 copy across and down to I97	Calculate AE cycle probabilities =1-EXP(-F96*28)
R8) Correct AE incidence calculation	Mod_H	AE_P_2	C96 copy across and down to D97	Change annualised AE probability to cycle probability =IF(Mod_H=0,1-((1-CHOOSE($C$91,C80,B89*C80,B90*C80,C85))^(GenSettings!$C$62/C$73)),H96)
R8) Correct AE incidence calculation	Mod_H	AE_P_2	C73	Change duration on treatment for PAL+LET from median to mean and make dynamic =IF(Mod_H=0,603, IF(AND(Mod_H=1,Mod_C=0, Mod_L=0), Results_Deterministic!F41*DaysInMonth, IF(AND(Mod_H=1, Mod_C=1,Mod_L=0), 'ERG time to event_P1'!V8, IF(AND(Mod_H=1, Mod_C=0,Mod_L=1), 'ERG time to event_P2'!V8))))
R8) Correct AE incidence calculation	Mod_H	AE_P_2	D73	Change duration on treatment for LET from median to mean and make dynamic =IF(Mod_H=0,420, IF(AND(Mod_H=1,Mod_C=0,Mod_L=0), Results_Deterministic!E41*DaysInMonth, IF(AND(Mod_H=1, Mod_C=1,Mod_L=0), 'ERG time to event_P1'!W8, IF(AND(Mod_H=1, Mod_C=0,Mod_L=1), 'ERG time to event_P2'!W8))))
R9) Change discounting to annual	Mod_I	Discounting	B6 copy down to B526	Change discounting to annual =IF(Mod_I=0,A6/13,ROUNDDOWN(A6/13,0))
R10) Use 365.25 days per year	Mod_J	GenSettings	C62	Change to 365.25 days per year =IF(Mod_J=0,364,365.25)